Scene Level Normalization and Harmonization of Planet Dove Imagery

Authors

Joe Kington

Alan Collison

TABLE OF CONTENTS

INTRODUCTION

The PlanetScope (PS) constellation includes satellites with different relative spectral responses (RSRs). Using PS data therefore requires an additional harmonization step to compare between different generations of satellites. Furthermore, PS surface reflectance (SR) data is often variable for a variety of other reasons, resulting in the need for a normalization step in addition to harmonization. Directional reflectance effects are not corrected for in PS SR data, stray light can be present in some satellites, and the radiative transfer models used for SR corrections are based on approximations that may have errors at low sun angles and other adverse conditions. In addition, SR corrections for PS data rely on MODIS estimates of atmospheric optical thickness collected at a different time of day and at a coarse spatial resolution. This leads to a significant degree of variability in PS SR data even within data with the same RSR, as shown in Figure 1.1. In addition, there is variability between generations of satellites even after applying a global harmonization correction, as seen in Figure 1.2. As a result, using Planetscope data often requires additional processing to correct for both factors before analysis. The scene-level normalization tool provides a method to both approximately harmonize between the varying RSRs and to normalize the data to correct for variability due to other parameters. This provides increased consistency, though absolute radiometric accuracy can suffer in some cases.

Figure 1.1: Planetscope SR scene data from a single generation of sensor over a 14-day interval in Dec, 2019 (top) compared with data from the same scenes processed with scene level normalization (bottom).

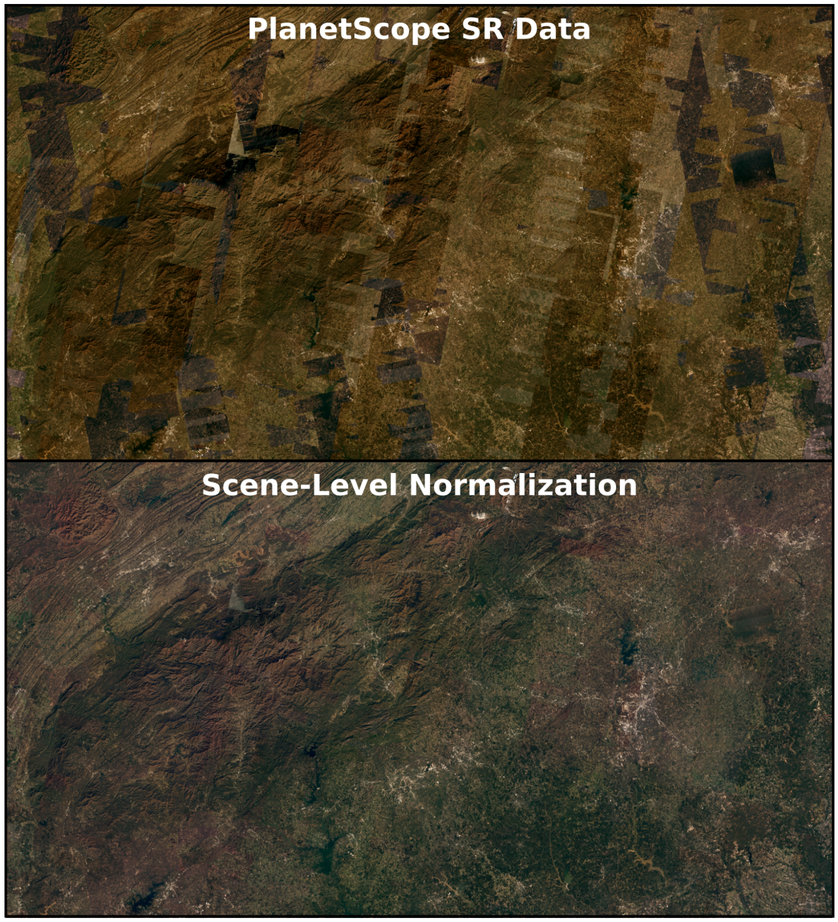
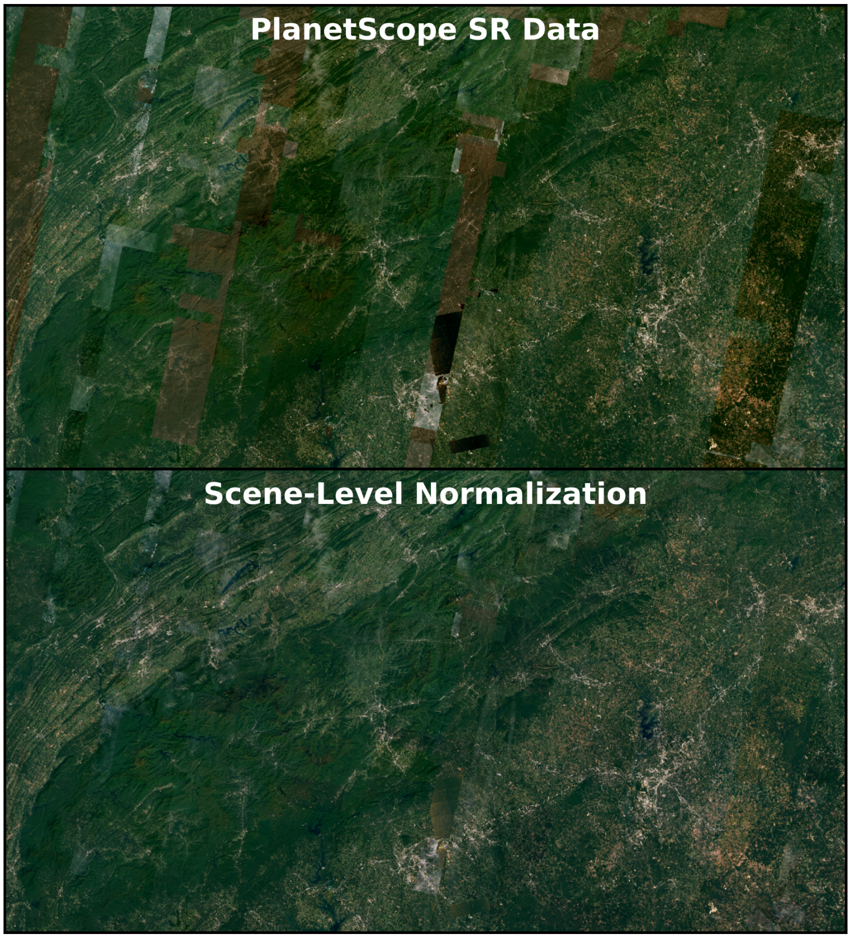


Figure 1.2: Planetscope SR scene data from multiple generations of sensors over a 14-day interval in Sept, 2021 with global harmonization applied (top) compared with data from the same scenes processed with scene level normalization (bottom).



This document describes the rationale and methods behind the “harmonize” operation available in the Planet Orders and Subscriptions APIs. The scene level methods described here apply only to the harmonize operation when “Sentinel-2” is selected as the “target_sensor” operation. By default, the tool applies a top-of-atmosphere-only global harmonization to make newer PS data comparable with older PS data. Scene level normalization makes all PS data consistent and approximately comparable to Sentinel2. Scene level normalization is applied to the red, green, blue and near infrared bands for all PS data.

METHODS

The scene level normalization method minimizes scene-to-scene and sensor-to-sensor variability by loosely matching each scene to a Sentinel2-based seasonal reference dataset. To do so, we fit a linear model for each band in each scene based on co-located, non-cloudy PlanetScope and seasonal model reference pixels. The linear model is tightly constrained to only produce realistic transformations even under adverse conditions such as undetected clouds, sunglint, snow, or significant changes such as burn scars. It is also constrained to minimize large changes in band ratios during normalization. We carefully select both the allowed parameter range and the misfit metrics so that the model is most sensitive to changes in dark features, which we hope are more likely to remain invariant between the multi-year seasonal model and the daily scene data. The final model is unique to each scene and transforms the input Planetscope reflectance coefficients to approximately match Sentinel2's spectral response and to minimize large deviations between scenes from different days and ambient conditions. The same transformation is applied uniformly across the entire scene. As a result, we preserve small and medium scale changes such as a stand of trees dying or a field being harvested. However, we may reduce any changes that affect the entire scene uniformly.

COMPARISON TO BASEMAP METHODS

The methods underlying scene level normalization were originally developed as a part of the Planet Basemap product's processing and was first published in Csillik, et al (2019) and Kington, et al (2019). These publications use Landsat-based seasonal models and slightly different settings than the methods described here, but are conceptually similar. Newer-generation Planetscope Basemaps use the same approach described here, with two key differences: 1) Basemaps apply additional spatial processing such as seamline removal in addition to scene level normalization. 2) Basemap scene-level normalization fits separate models for water and land to avoid issues near coastlines and large water bodies. The scene level normalization methods described here use one model for the entire scene, even if the scene is split between water and land.

MODEL DEVELOPMENT

Normalization is intended to improve spatial and temporal consistency of data from different scenes and different sensors and improve overall radiometric accuracy of a diverse satellite constellation. However, it must also avoid unrealistic values and artifacts in the presence of snow, sunglint, undetected clouds, changes in scene content, and similar issues. As a result, we need to avoid overfitting to our reference data and need a very stable and predictable transformation. Because a model is developed independently for each scene, this is a "train once, use once" problem. Overfitting is difficult to avoid with complex methods and a fundamentally limited, single-scene training data set. Therefore, we chose a parametric approach with low dimensionality over non-parametric methods such as random forest that are prone to overfitting with limited training data.

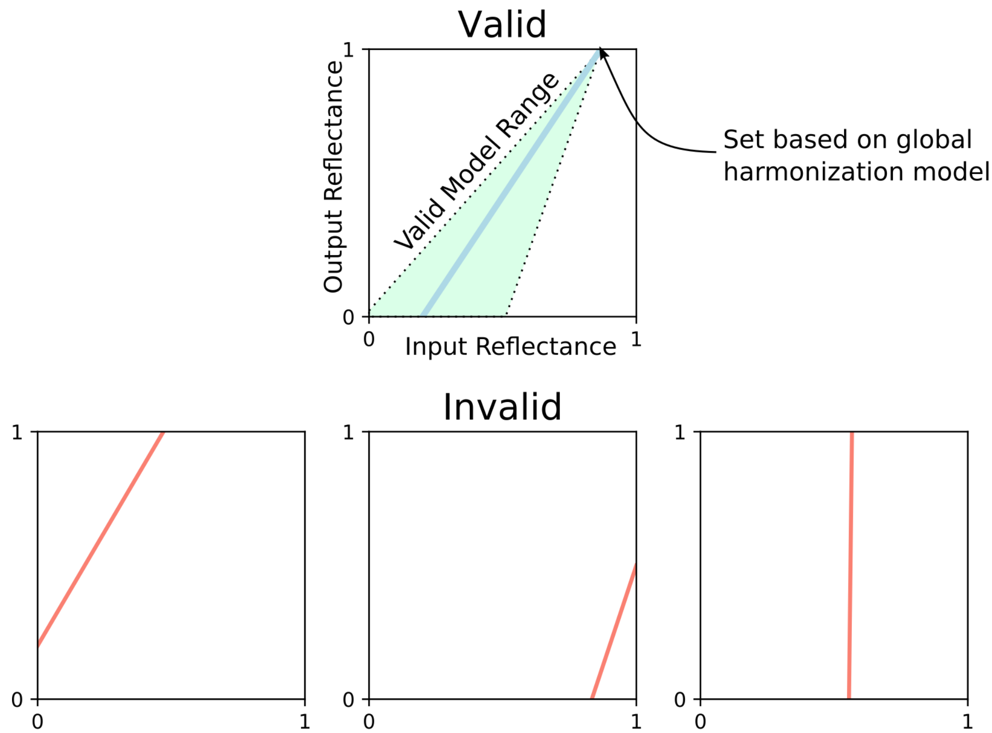
Linear models are frequently used in spectral band adjustment factors to correct for differences between spectral responses of satellites (e.g. Teillet, et al, 2007). However, a single global linear harmonization model is only approximately valid for small differences in spectral response. Ideally, harmonization approaches should use different models for different land cover types. Using a unique model for each scene is less accurate than using different models for each land cover type within the scene, but still improves on a single, global model. A per-scene approach also avoids introducing unrealistic variability by splitting a scene into different land cover types and applying a different model to each one. The boundaries between the different land cover classes, and therefore different models, often produce visible artifacts in the final result. Therefore, we've chosen to fit a unique harmonization model for each scene, but avoid artificially splitting a scene into separate regions with separate models. This ensures we maintain spatial consistency within a scene.

Furthermore, the scene level normalization method is meant to correct for other factors beyond only harmonizing the spectral response differences. Uncertainties in SR corrections and stray light affect all land cover classes similarly, and necessitate a different correction than pure harmonization. Differences due directional reflectance effects – as typically modeled by a bi-directional reflectance distribution function (BRDF) – vary by land cover class and view angle, but are also not easily incorporated into a pure harmonization model. Because our satellites have a very narrow field of view, BRDF effects are mostly constant across a scene, and can also be particularly corrected for with a single linear model. For these reasons, a per-scene linear model is capable of approximately correcting for many of the sources of variability between sensors and scenes within the PlanetScope constellations.

It is critical to avoid using a linear, least squares fit for this use case even for a linear model. A linear fit will frequently produce unrealistic values in adverse situations (e.g. snow or clouds) and least squares misfit measurements are heavily biased by outliers. Furthermore, Most natural ground reflectance values are relatively low (less than 0.5), and very high reflectances in most non-snowy data are limited to small areas such as rooftops. An unconstrained linear model will often produce results that fit the bulk of the data well, but give unrealistic predictions for small bright features such as rooftops. Similarly, we want to avoid overfitting snow or clouds in the scene data that are not present in the seasonal model.

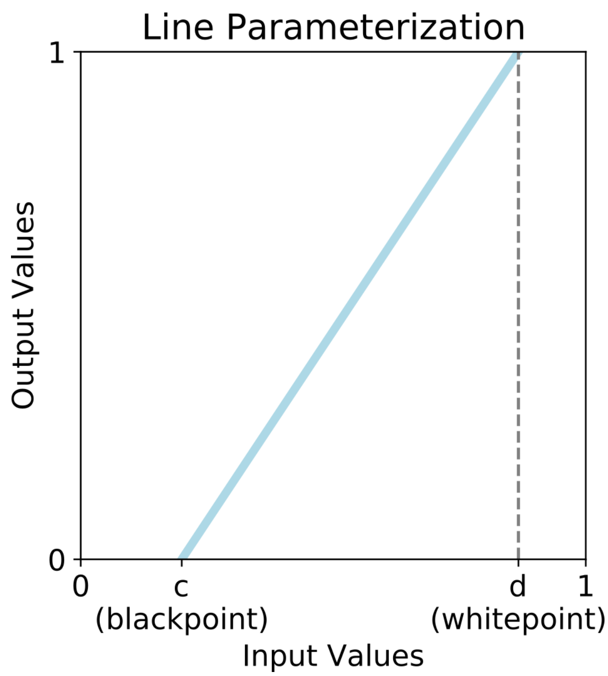
We employ several techniques to ensure reasonable transformations to the scene data for both low-reflectance land cover classes such as vegetation and bright features such as snow or rooftops. First, we do not consider all linear models valid and exclude any that represent clearly incorrect transformations of the scene data to match the seasonal target (Figure 2.1). For example, we need to reject transformations that would brighten the data excessively (Figure 2.1 - bottom left), darken the data excessively (Figure 2.1 - bottom center), or produce a “thresholded” result with little dynamic range (Figure 2.1 - bottom right). Next, we do not want to produce a result that dramatically changes bright objects. This is because bright objects are likely to be transient effects that are not present in the reference seasonal model (e.g. clouds, snow, sunglint, etc) and because bright objects are usually underrepresented in the scene values and the darker portions of the range dominate the fit. To constrain these issues, we “pin” the model such that a specific high input reflectance does not change in the output. The value we choose to pin is based on the global harmonization (a global linear per-band model) chosen for the starting guess. For Superdove, this is a no-change model, so the pinned reflectance in each band is 1.0. For Dove Classic, it is different for each band. Only variation within a defined range from the starting guess is allowed (green region on Figure 2.1 - top). Finally, we expect that there will be significant real differences between the input scene data and our long-term seasonal models. As a result, we want to be most sensitive to differences in features that are likely to be invariant. Broadly speaking, dark features are more likely to be invariant than bright features, so we want to bias our model and results towards removing major differences in dark, hopefully-invariant features on the ground.

Figure 2.1: Valid vs invalid linear normalization models.



It is difficult to enforce the necessary constraints with the typical $Y = aX + b$ line parameterization, where x represents our input scene reflectance values and y represents our normalized reflectance values. It is easiest to enforce constraints in convex optimization problems when the desired behavior can be expressed as simple ranges for the model parameters. It's not possible to define the green region in Figure 2.1 with a bounded range of slopes and offsets in a linear equation, as the desired slope constraints depend on the offset. For that reason, we'll define our transformation in terms of two parameters: c and d (Figure 2.2). c is defined as the input reflectance value that produces an output reflectance of 0.0 ("blackpoint" in common color correction terminology) and d is defined as the input reflectance value that produces an output value of 1.0 ("whitepoint" in color correction terminology). We can then define our linear model as a non-linear equation in terms of c and d (Equation 2.1). This allows easily defining the range of valid model parameters, but requires the use of non-linear optimization methods, as the problem is now non-linear with regards to the c and d values we're solving for. We'll then allow $x0$ to vary within a defined range of c_{min} to c_{max} which allows us to constrain results to the green region shown in Figure x.1. d can either be allowed to vary by a more tightly constrained range of d_{min} to d_{max} ("unpinned whitepoints") or fixed to the value defined by the initial guess ("pinned whitepoints"). In the pinned whitepoints case, we only solve for a single parameter (c) per-band, which reduces the dimensionality of the problem and often produces a more robust result at the expense of reducing the ability of the model to correct for differences in very bright regions. This still produces a very different result than solving for a static shift (e.g. dark object subtraction), as the change from the original value depends on the pixel value. With the pinned case, bright values have less potential to change from their input SR value than darker values.

Figure 2.1: Graphical explanation of linear model parameterization.



Equation 2.1: Linear model parameterization.

$$y = aX + b = \frac{X}{d - c} + \frac{c}{c - d}$$

Using the parameterization in Equation 2.1 requires using a non-linear solver, as the problem is no longer linear in terms of the model parameters we need to solve for. However, a non-linear optimization routine also allows for the use of more robust objective functions than the typical least squares misfit. Minimizing only least squares misfit to the seasonal reference data is often dangerous, as outliers due to fundamental differences between the datasets heavily bias the result. Furthermore, a linear least squares fit would not allow us to enforce constraints about preserving similar correlations between bands as the input data. Because the model parameters are independent for each band, we need to avoid extreme changes in band ratios when fitting the model. Finally, a simple distance metric of misfit is biased towards bright values. Values near zero are unlikely to have as large of a misfit in least square distance terms as bright values. A change of 0.01 in reflectance in a pixel with a value of 0.01 is usually much more significant than a change of 0.01 in a pixel with a reflectance of 0.5. This also serves to ensure we focus on fitting dark features more closely than bright features, as we expect bright features are more likely to be transient and due to undetected clouds, haze, sunglint, and similar issues.

For these reasons, we need to minimize a more complex objective function than least squares misfit. First, we'll use a different distance metric to compute misfit relative to the seasonal reference data.

Let's define misfit in terms of co-located input scene reflectance values ($\mathbf{X}$) and seasonal model reflectance values ($\mathbf{Y}$), each with m sampled points and n spectral bands.

$$\mathbf{X} = \begin{bmatrix} X_{(1,1)} & \dots & X_{(1,n)} \\ \vdots & \ddots & \\ X_{(m,1)} & & X_{(m,n)} \end{bmatrix} \mathbf{Y} = \begin{bmatrix} Y_{(1,1)} & \dots & Y_{(1,n)} \\ \vdots & \ddots & \\ Y_{(m,1)} & & Y_{(m,n)} \end{bmatrix}$$

And following Equation 2.1, the model that we're trying to solve for parameters for based on these co-located points is:

$$model(\mathbf{X}) = \begin{bmatrix} a_1 X_{(1,1)} + b_1 & \dots & a_n X_{(1,n)} + b_n \\ \vdots & \ddots & \\ a_1 X_{(m,1)} + b_1 & & a_n X_{(m,n)} + b_n \end{bmatrix} = \begin{bmatrix} \frac{X_{(1,1)}}{d_1 - c_1} + \frac{c_1}{c_1 - d_1} & \dots & \frac{X_{(1,n)}}{d_n - c_n} + \frac{c_n}{c_n - d_n} \\ \vdots & \ddots & \\ \frac{X_{(m,1)}}{d_1 - c_1} + \frac{c_1}{c_1 - d_1} & \dots & \frac{X_{(m,n)}}{d_n - c_n} + \frac{c_n}{c_n - d_n} \end{bmatrix}$$

Rather than minimizing the squared misfit (L2 norm), we'll minimize the mean absolute value (L1 norm) of the ratio of misfit between values to the sum of the values. This ensures we treat misfit proportionally to the brightness of the pixel. Let's therefore define the misfit metric as:

$$misfit(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{i=m} \sum_{j=1}^{j=n} \frac{1}{m} \left| \frac{X_{i,j} - Y_{i,j}}{X_{i,j} + Y_{i,j}} \right|$$

Next, we need a way to penalize solutions where one band would change dramatically more than others during normalization. While the differing spectral responses between satellites means that we require band ratios to change slightly, cases where one band changes much more than others are unrealistic, even if they reduce misfit. We'll use a regularization metric that measures how much "gray" pixels – pixels with equal values in each band – change during normalization. We'll use a series of varying shades of gray (e.g. 0.1, 0.2, 0.3 etc in each band, defined as $\mathbf{Z}$ in Equation 2.2) as input data to our model, then measure how much the model changes these uniform values with the same metric used to measure misfit from the seasonal model. This results in a regularization metric (referred to as "balance" below) that is sensitive to cases where band ratios would change unrealistically during normalization and cases where normalization would change dramatically from the initial state of the model.

Equation 2.2: Regularization metric.

$$\mathbf{Z} = \begin{bmatrix} 0/n & \dots & 0/n \\ 1/n & \dots & 1/n \\ \vdots & \ddots & \vdots \\ n/n & \dots & n/n \end{bmatrix}, balance(\mathbf{Z}, model) = misfit(\mathbf{Z}, model(\mathbf{Z}))$$

We then combine our misfit metric from the seasonal model with the regularization metric defined above to produce an objective function that will be minimized during fitting. This is a weighted sum of the two factors, with w being the weight applied to the regularization metric. The weight placed on the regularization metric provides an additional hyperparameter for how closely we should fit our seasonal reference models. Higher values lead to smaller changes from the input scene data, and smaller values allow more change to match the seasonal model.

$$objective(\mathbf{X}, \mathbf{Y}, model) = misfit(\mathbf{X}, \mathbf{Y}) + w \cdot balance(model)$$

We then solve for our model parameters independently for each scene using values of co-located pixels in both the scene and the seasonal reference model. We use a bounded gradient descent method (Byrd, et al, 1995) to minimize the objective function subject to the constraints c_{min}, c_{max} and in the "unpinned" case d_{min}, d_{max} as well. This method requires a starting guess for model parameters, which also defines the value of d in the "pinned" case. We use a global harmonization model (see Table x.1) for the starting guess. c_{min}, c_{max}, d_{min}, d_{max} and w are hyperparameters which should be optimized for the use case at hand.

IMPLEMENTATION

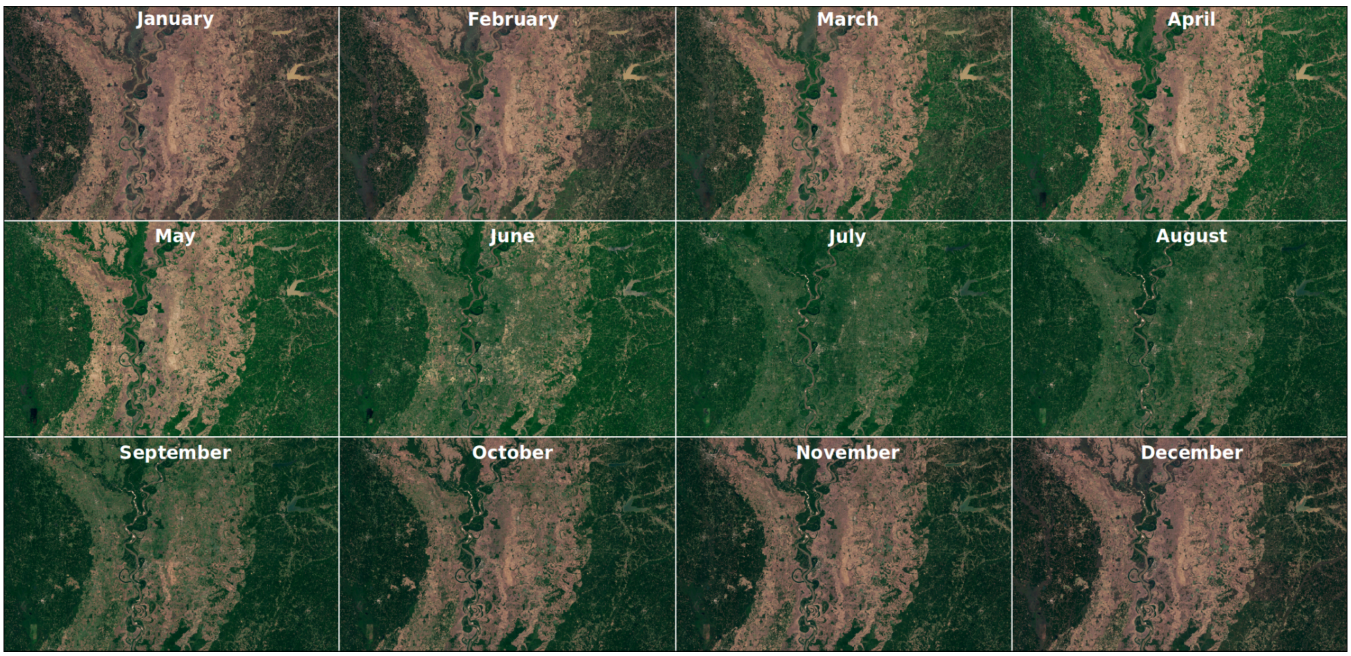
SEASONAL REFERENCE MODELS

The pinned linear model used for scene level normalization is designed to be sensitive to differences in the reflectance of dark objects as compared to the reference data. As a result, haze or cloud contamination of dark objects in the seasonal model can produce unrealistic results. Using only reference data acquired within a relatively short duration around the time of acquisition of the Planetscope scene tends to produce inaccuracies due to haze contamination of dark objects. Higher cadence reference data such as MODIS can produce haze and cloud free reference data in a short interval, but Sentinel2 compatibility is a requirement for the scene level normalized product. As a result, we use static, multi-year seasonal models rather than near-coincident Sentinel2 imagery. This ensures a very cloud, haze, and snow free reference dataset even in very cloudy tropical regions. Extensive testing within the Planet Basemap product has shown this approach to give more consistent and accurate results than using near-coincident reference imagery from Sentinel2. Because the scene level method is designed to not overfit to the reference imagery and be most sensitive to dark objects in both datasets, we are able to accurately match coincident Sentinel2 observations despite using a static model, as shown in the Crossover Analysis section.

To create seasonal reference data used as a baseline for this method we processed the entire Sentinel2 archive to surface reflectance using Frantz's (2019) FORCE method, producing 30m pixel size output. We then made monthly-cadence multi-year composites of the resulting data (Figure 3.1). These composites are based on selecting the observation nearest the 30th percentile of total brightness (mean of all bands) for all cloud free data at each pixel location in each month between

2017 and 2021. The 30th percentile of observations was chosen based on trial and error to avoid haze, clouds, and snow while not inadvertently selecting for cloud shadows and other transient dark effects. A mean value is heavily biased by haze/etc, and while the median is less biased towards those effects, frequently hazy regions necessitated a below-median value to produce a high quality composite. We base blue band normalization on B2 from Sentinel2, green band on B3, red band on B4, and NIR on B8a (narrow NIR).

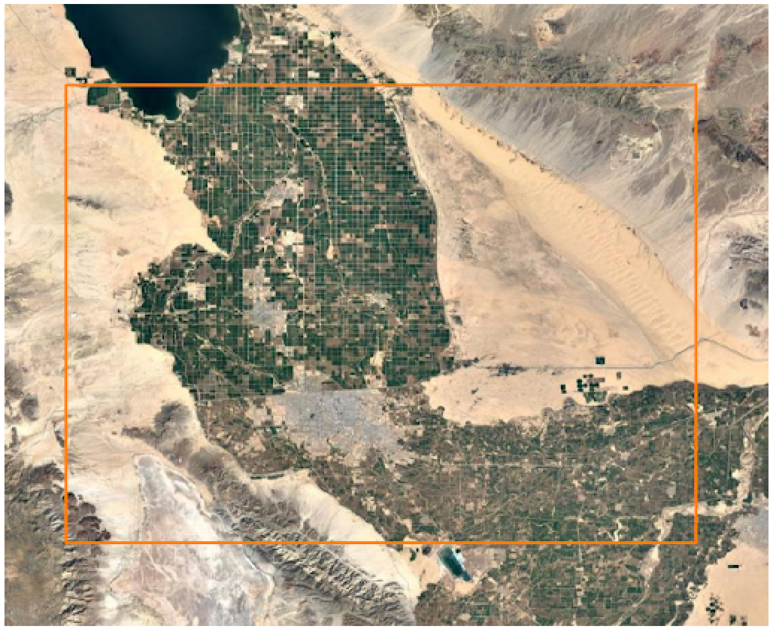
Figure 3.1: Seasonal reference model data for 12 months over a portion of the Mississippi River Valley, central US.



OPTIMIZING HYPERPARAMETERS

The normalization model provides a robust and effective method for transforming Dove imagery into Sentinel-2-like products. To optimize the input hyperparameters for the model and the starting normalization factors, a study was performed for an area covering the Imperial Valley region of southern California/northern Mexico, shown in Figure 4.1.

Figure 4.1: The area of the Imperial Valley region used for initial normalization parameter optimization.

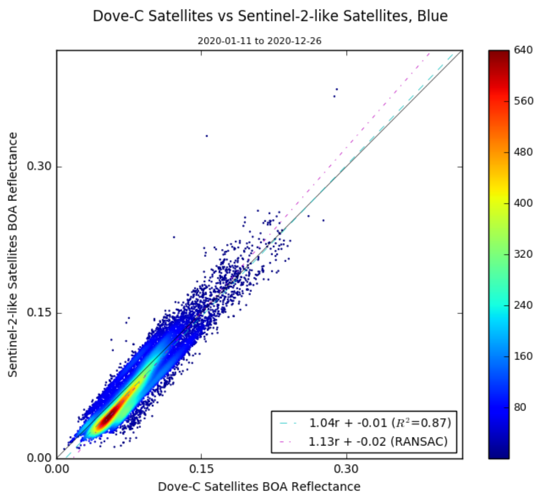


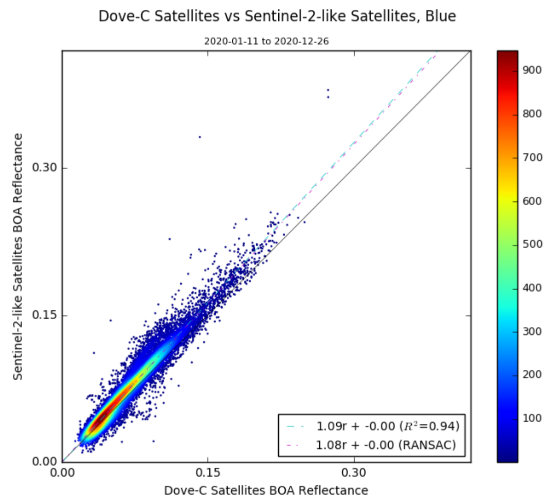
This area was chosen because it provides many cloud-free days over a full year. It also contains cropland with multiple growth/harvest periods during the year, the fields of which provide a convenient sampling grid for time series comparisons. This allowed collection of clear images over most of the year and covered a full cycle of vegetation growth and decay over any subset of the year. In addition to the fixed parameters of the normalization model, we also used these crossovers to determine the starting guess for the normalization gain and offset.

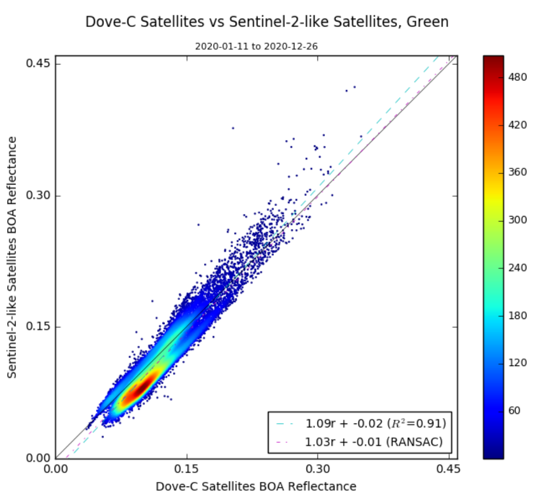
Given the very close match between the spectral bands of Planet's SuperDove satellite sensors and those of Sentinel-2, simultaneous crossovers between those satellites and Dove Classic satellites were used for determining the initial normalization inputs. In all, over 14,000 Dove Classic crossover images were collected for the time period between January 1 and December 31 of 2020 intersecting with 2500 SuperDove images. Scatter plot linear fits for intersecting image pairs covering 36 combinations of the normalization model parameters were generated and compared. Figure 4.2 shows an example of the scatter plots containing the crossover data for one such combination of model parameters for each Dove Classic sensor spectral band.

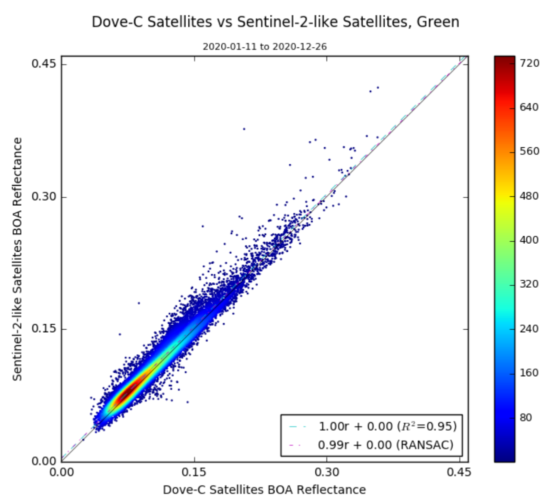
Figure 4.2: The area of the Imperial Valley region used for initial normalization parameter optimization.

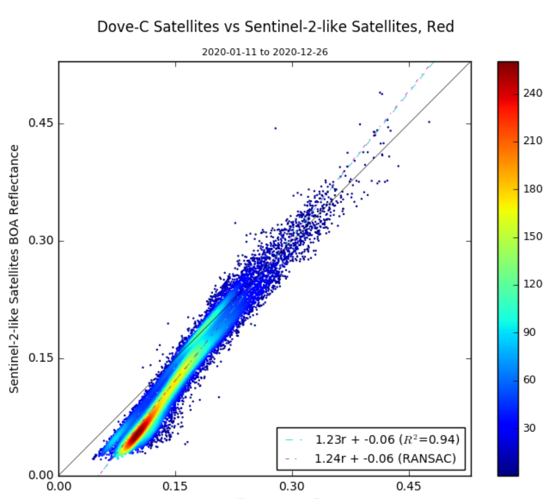
Without Normalization **With Normalization**

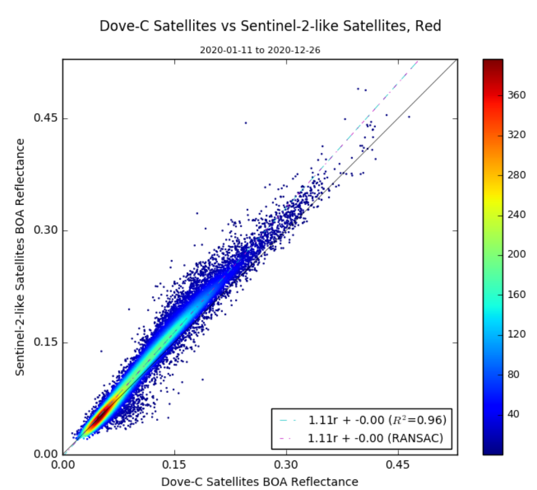

Dove-C Satellites vs Sentinel-2-like Satellites, Blue
2020-01-11 to 2020-12-26
Sentinel-2-like Satellites BOA Reflectance
Dove-C Satellites BOA Reflectance
1.04r + -0.01 (R^2=0.87)
1.13r + -0.02 (RANSAC)

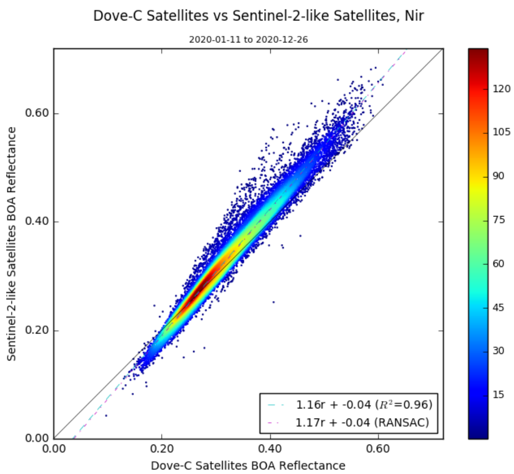

Dove-C Satellites vs Sentinel-2-like Satellites, Blue
2020-01-11 to 2020-12-26
Sentinel-2-like Satellites BOA Reflectance
Dove-C Satellites BOA Reflectance
1.09r + -0.00 (R^2=0.94)
1.08r + -0.00 (RANSAC)

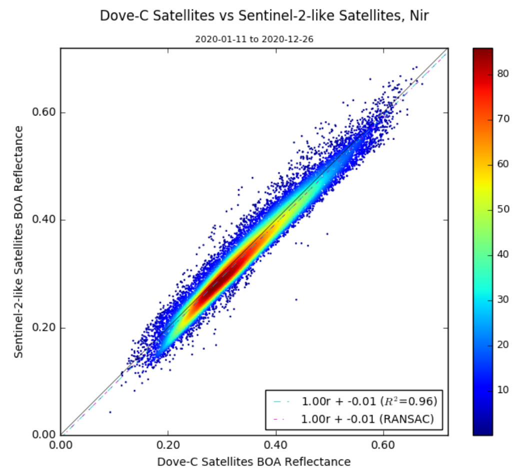

Dove-C Satellites vs Sentinel-2-like Satellites, Green
2020-01-11 to 2020-12-26
Sentinel-2-like Satellites BOA Reflectance
Dove-C Satellites BOA Reflectance
1.09r + -0.02 (R^2=0.91)
1.03r + -0.01 (RANSAC)


Dove-C Satellites vs Sentinel-2-like Satellites, Green
2020-01-11 to 2020-12-26
Sentinel-2-like Satellites BOA Reflectance
Dove-C Satellites BOA Reflectance
1.00r + 0.00 (R^2=0.95)
0.99r + 0.00 (RANSAC)


Dove-C Satellites vs Sentinel-2-like Satellites, Red
2020-01-11 to 2020-12-26
Sentinel-2-like Satellites BOA Reflectance
Dove-C Satellites BOA Reflectance
1.23r + -0.06 (R^2=0.94)
1.24r + -0.06 (RANSAC)


Dove-C Satellites vs Sentinel-2-like Satellites, Red
2020-01-11 to 2020-12-26
Sentinel-2-like Satellites BOA Reflectance
Dove-C Satellites BOA Reflectance
1.11r + -0.00 (R^2=0.96)
1.11r + -0.00 (RANSAC)





Application of the generated normalization factors improves the overall agreement between Dove Classic and SuperDove in every corresponding band, aligning the distribution of data points closer to the 1:1 line and generally reducing the overall scatter.

Table 4.1 shows the combination of input hyperparameters which produced the best comparison results. While these were the optimal parameters, it should be noted that there was very little difference between all the combinations, so individual image normalizations are not expected to be very sensitive to the chosen set of model parameters.

Table 4.1: Optimal normalization model hyperparameters determined from Dove Classic/SuperDove crossovers and the initial guess for the normalization factors to be apple to Dove Classic images.

MODEL HYPERPARAMETERS

Hyperparameter	Optimal Value
w	0.5
c_{min}	-0.1
c_{max}	0.23
d behavior	“Pinned whitepoints”

INITIAL MODEL GAIN/OFFSET (Dove Classic)

Band	Gain	Offset
blue	0.860	0.014
green	0.946	0.021
red	0.961	0.021
nir	1.001	0.007

Given the very close similarity between the sensor band responses of Dove-R, SuperDove and Sentinel 2 satellites, under ideal viewing conditions we expect very little difference in their measured reflectances. We also expect that normalization applied to images from Dove-R and SuperDove will be a small adjustment. Therefore for these satellites, while the same model parameters used for Dove Classic satellites are used for Dove-R and SuperDove, the initial normalization model gain and offset are set to 1.0 and 0.0, respectively.

PERFORMANCE

CROSSOVER ANALYSIS

The normalization model parameter values and the starting per-band gain and offset values have been determined based on comparisons to monthly FORCE generated Sentinel 2 surface reflectance mosaics generated for a region in the Imperial Valley region of North America. This region is a mixture of barren sandy and rocky areas surrounding a large cropland area with a mixture of fields growing alfalfa, which is harvested multiple times a year, and other annual crops. To validate the performance of these model parameters for other areas with somewhat different makeups, we examined two other regions: one in the U.S. midwest (Iowa) and the other in Egypt covering a desert area with crop circles intermingled. Figure 5.1 shows these two regions.

Figure 5.1: The US-Iowa (left) and Egypt (right) regions used to validate the normalization model parameters.

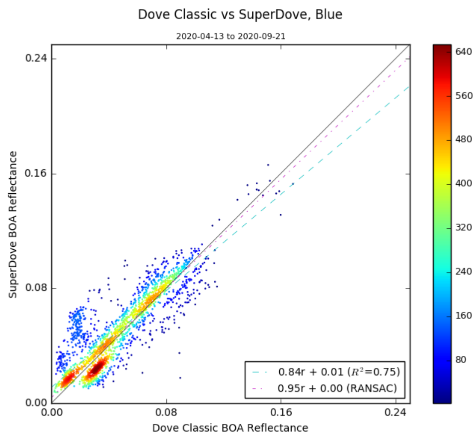
For each region, direct comparisons between Dove Classic and SuperDove satellites were made for the selected sample areas. In the case of the Iowa region, these samples corresponded to common land units (CLUs), which identify individual fields within the US. For Egypt, the samples were from a grid made up of 2500x2500 meter squares. Direct comparisons between the Dove satellites are then made between the average per band reflectance values within each sample. Time series plots are also compared for all generations of Doves and Sentinel 2, where for convenience Planet's atmospheric correction is applied to Sentinel TOA reflectance to produce it's surface reflectance products..

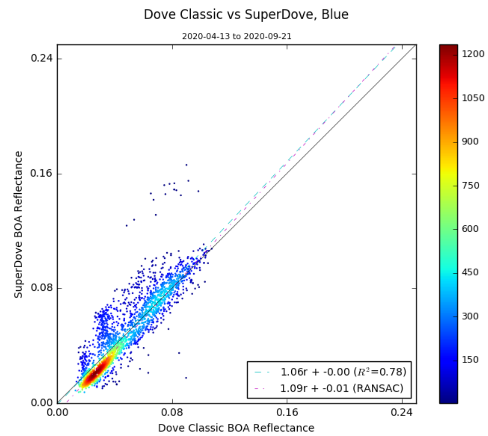
US (Iowa)

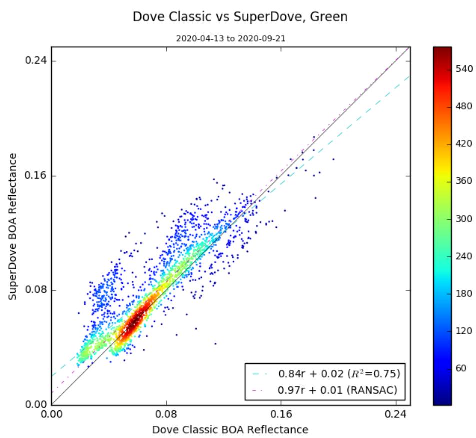
The Iowa region that was examined is dominated by agricultural fields, primarily corn and soybean crops harvested once a year. The time period examined was from April 1 to October 1 for 2020, covering a full growing season. Each sample area (CLU in this case) was analyzed with and without normalization applied and then both scatter plot and time series comparisons were made. Figure 5.2 shows a comparison of reflectances for all Dove Classic satellites compared to SuperDove.   For each spectral band, normalization improves the overall consistency between Dove CLassic and SuperDove.

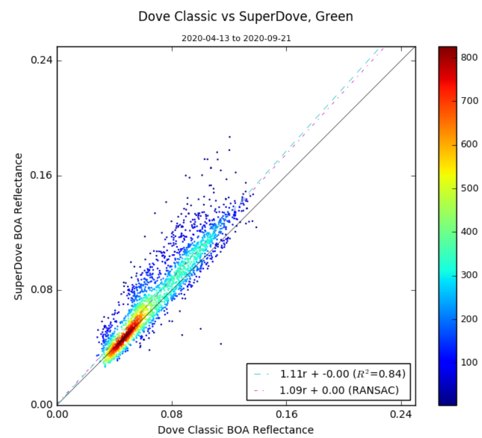
Figure 5.2: Comparison of surface reflectance between Dove Classic and SuperDove with and without normalization applied.

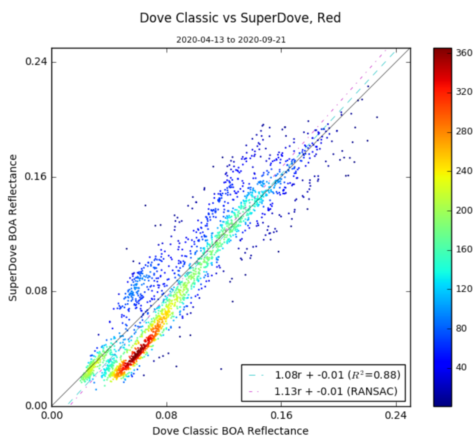
No Normalization **Normalized**

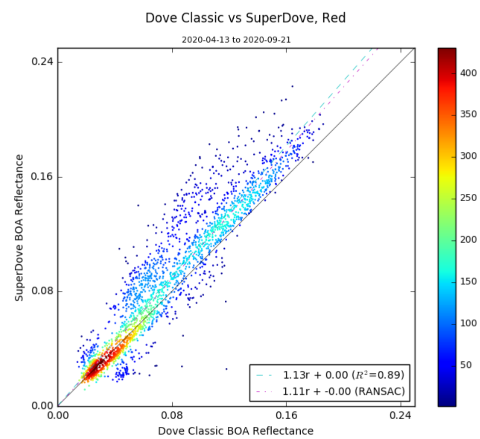

Dove Classic vs SuperDove, Blue
2020-04-13 to 2020-09-21
0.84r + 0.01 (R^2=0.75)
0.95r + 0.00 (RANSAC)
SuperDove BOA Reflectance
Dove Classic BOA Reflectance

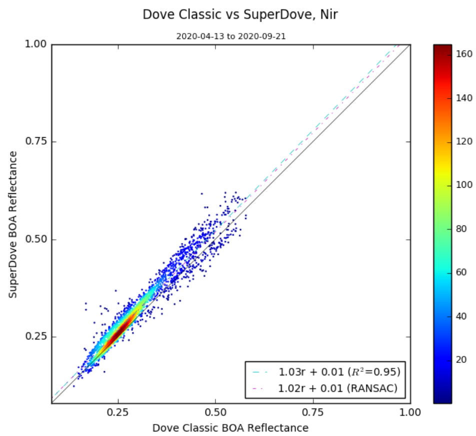

Dove Classic vs SuperDove, Blue
2020-04-13 to 2020-09-21
1.06r + -0.00 (R^2=0.78)
1.09r + -0.01 (RANSAC)
SuperDove BOA Reflectance
Dove Classic BOA Reflectance


Dove Classic vs SuperDove, Green
2020-04-13 to 2020-09-21
0.84r + 0.02 (R^2=0.75)
0.97r + 0.01 (RANSAC)
SuperDove BOA Reflectance
Dove Classic BOA Reflectance


Dove Classic vs SuperDove, Green
2020-04-13 to 2020-09-21
1.11r + -0.00 (R^2=0.84)
1.09r + 0.00 (RANSAC)
SuperDove BOA Reflectance
Dove Classic BOA Reflectance


Dove Classic vs SuperDove, Red
2020-04-13 to 2020-09-21
1.08r + -0.01 (R^2=0.88)
1.13r + -0.01 (RANSAC)
SuperDove BOA Reflectance
Dove Classic BOA Reflectance


Dove Classic vs SuperDove, Red
2020-04-13 to 2020-09-21
1.13r + 0.00 (R^2=0.89)
1.11r + -0.00 (RANSAC)
SuperDove BOA Reflectance
Dove Classic BOA Reflectance



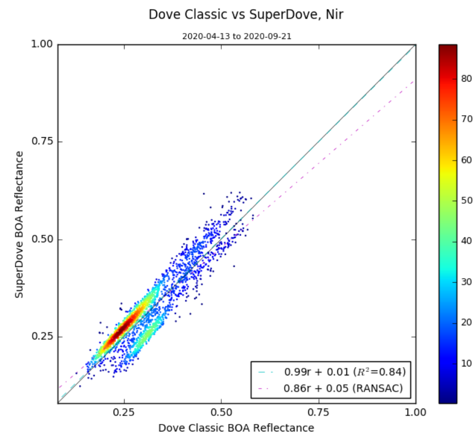
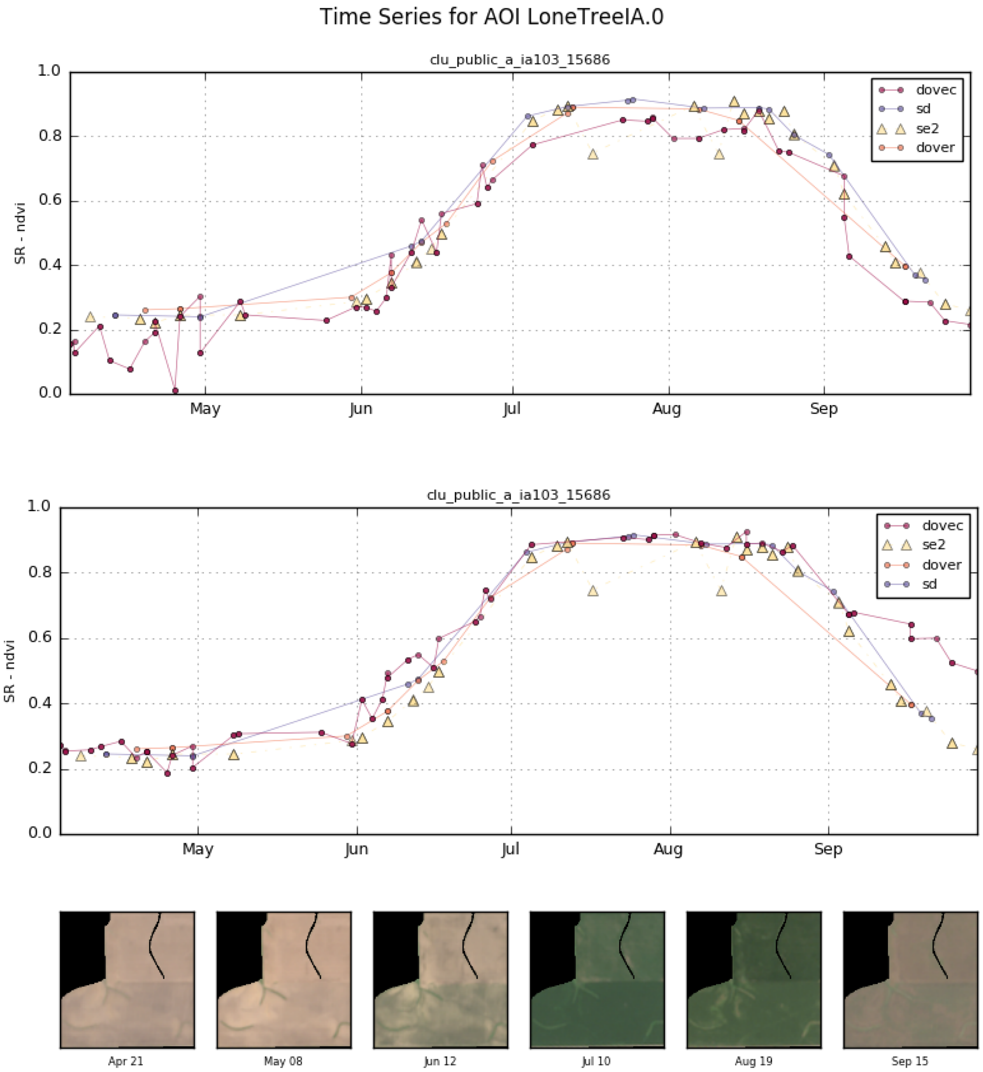


Figure 5.3 shows an example of a time series for a single field over the entire growing season, alternating between standard and normalized reflectances. In this case data from Dove Classic, Dove-R and SuperDove satellites are included and only those crossovers which intersect 100% of the field are used. The Sentinel 2 surface reflectance products here have been generated using Planet's atmospheric correction process and are included to illustrate the overall agreement between all types of satellites. As expected, normalization has only a small effect on the Dove-R and SuperDove satellites which already show good consistency with Sentinel, but has a significant effect on Dove Classic satellites, generally improving the overall agreement with Sentinel, particularly during the peak of the growing season. However, given that the normalization reference favors the surface conditions during the middle of the month, deviations can occur when conditions are changing rapidly, which is the case during early September of the example shown.

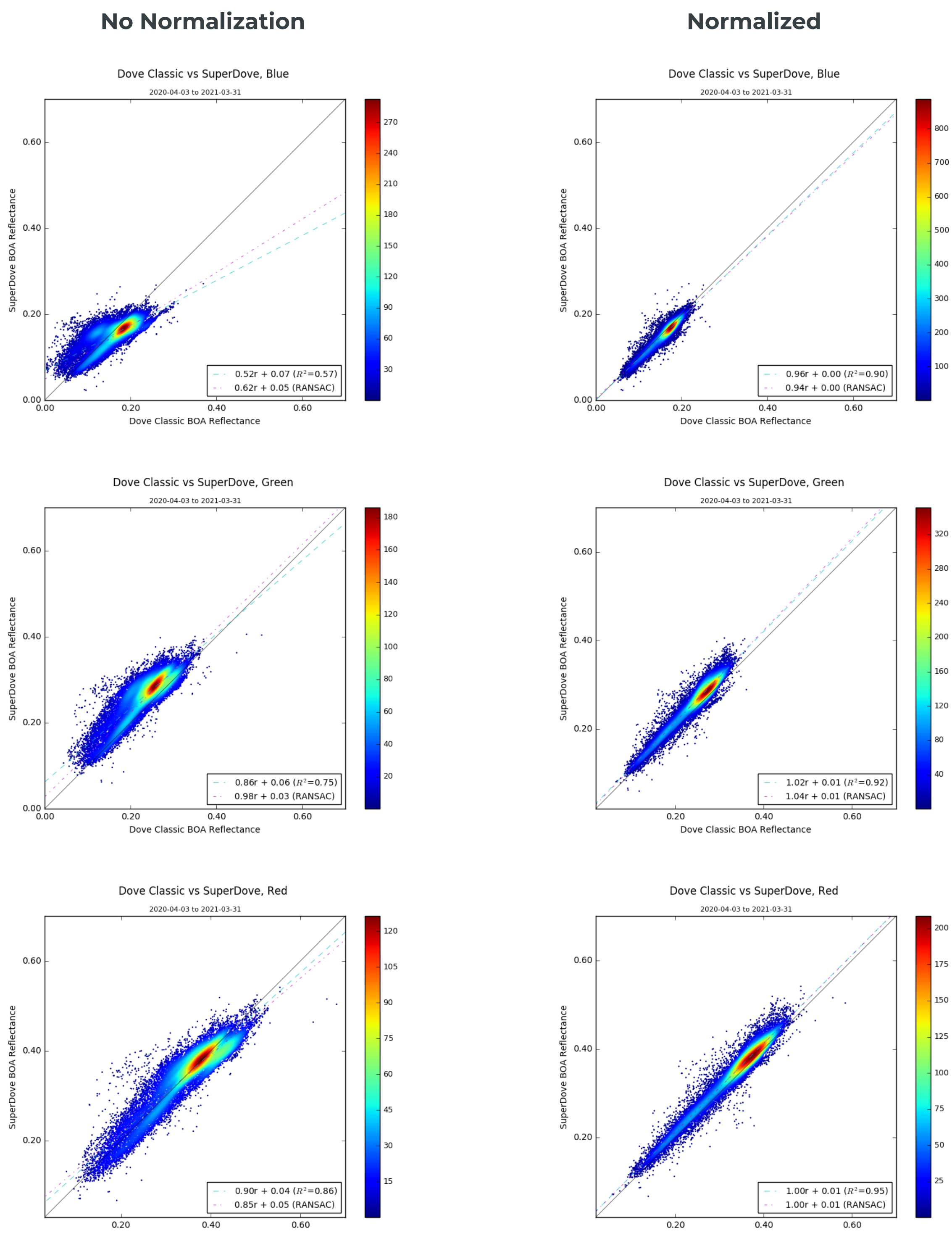
Figure 5.3: Comparison showing the time series of NDVI values for a field in Iowa over the growing season without (top) and with (bottom) normalization applied to the Dove satellites.

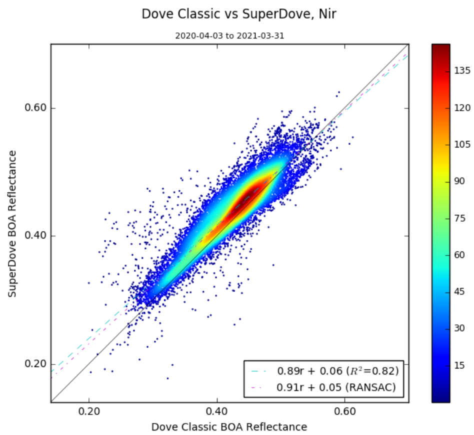


Egypt

The second region selected is in Egypt and is primarily desert with many crop circles arranged within the area. Unlike Iowa, the crops here are grown and harvested continuously throughout the year and overall the region is much brighter. To analyze this area, we divided it into squares each 2500 meters on a side and averaged the band reflectances within each one. Otherwise the processing is the same as for the Iowa example. In general, where there were crop circles, multiple fields were included in each square with crops grown and harvested at different times. Figures 5.4 and 5.5 show the results of one of the sample squares. In every band, normalization of the Dove reflectances shows significant improvement in consistency between Dove Classic and the Sentinel-like satellites.

Figure 5.4: Comparison of surface reflectance between Dove Classic and Sentinel 2 with and without normalization applied.




Dove Classic vs SuperDove, Nir
2020-04-03 to 2021-03-31
SuperDove BOA Reflectance
Dove Classic BOA Reflectance
0.89r + 0.06 (R^2=0.82)
0.91r + 0.05 (RANSAC)

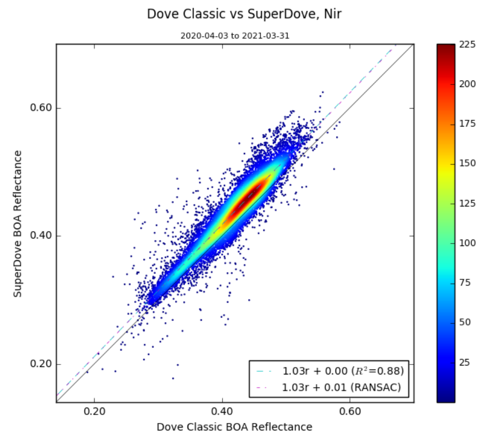
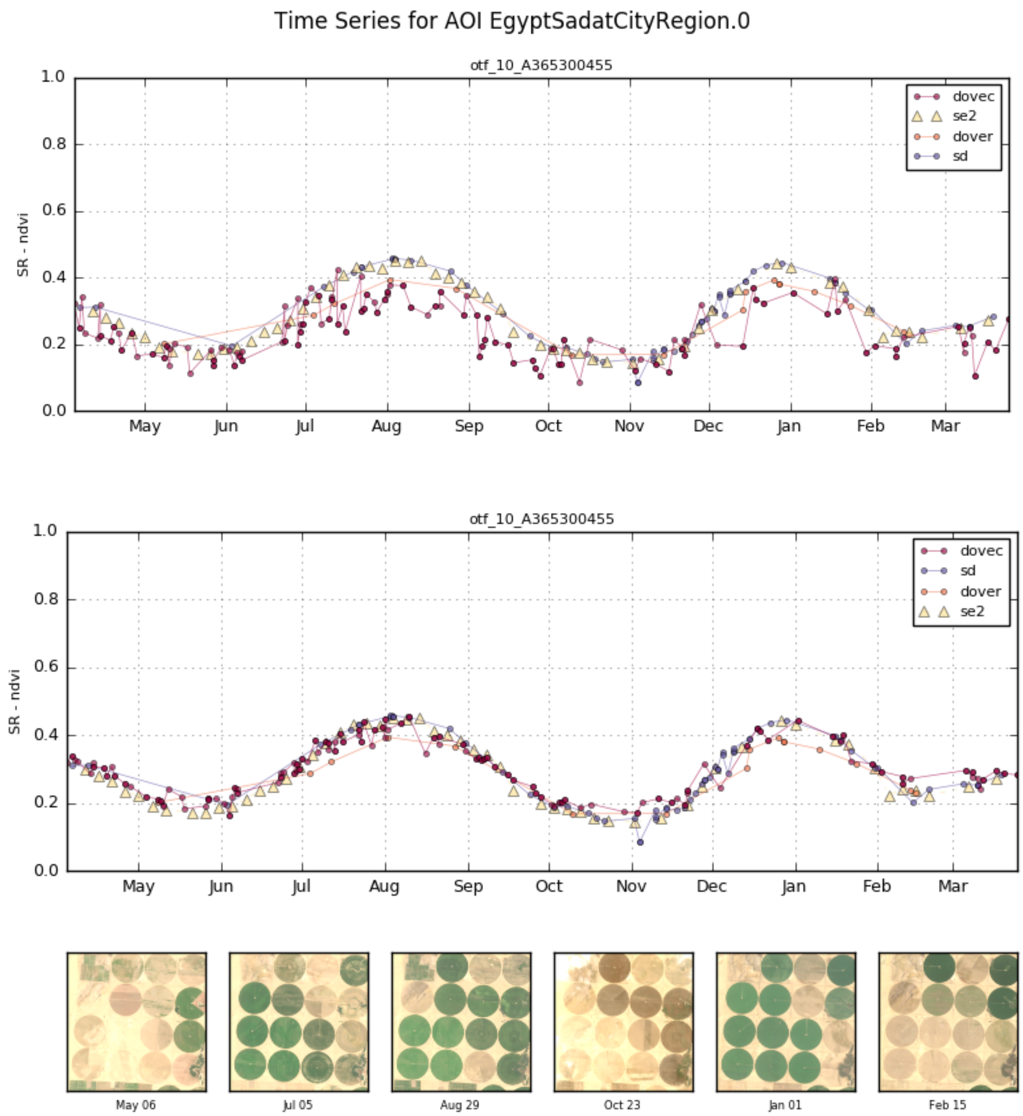

Dove Classic vs SuperDove, Nir
2020-04-03 to 2021-03-31
SuperDove BOA Reflectance
Dove Classic BOA Reflectance
1.03r + 0.00 (R^2=0.88)
1.03r + 0.01 (RANSAC)

Figure 5.5: Comparison showing the time series of NDVI values for a collection of crop circles in Egypt over the full year without (top) and with (bottom) normalization applied to the Dove satellites.



CAVEATS

When a Dove PSScene surface reflectance product is requested for download, by default a normalized product is generated. While this will in general provide reflectance values more consistent with what one would expect from a native Sentinel 2 image, the consistency with which this is the case and the degree to which deviations occur depends on a number of factors, including:

- How different the Dove satellite sensor band responses are from Sentinel 2.
- How quickly the surface conditions change during a month for a particular region.
- Whether an image is dominated by a large body of water, for example one with a large lake or along the coast.
- Whether the atmospheric correction has overcorrected an image.

In addition, given that normalization operates on a per-scene level, it can introduce visible differences between adjacent scenes within a single strip.

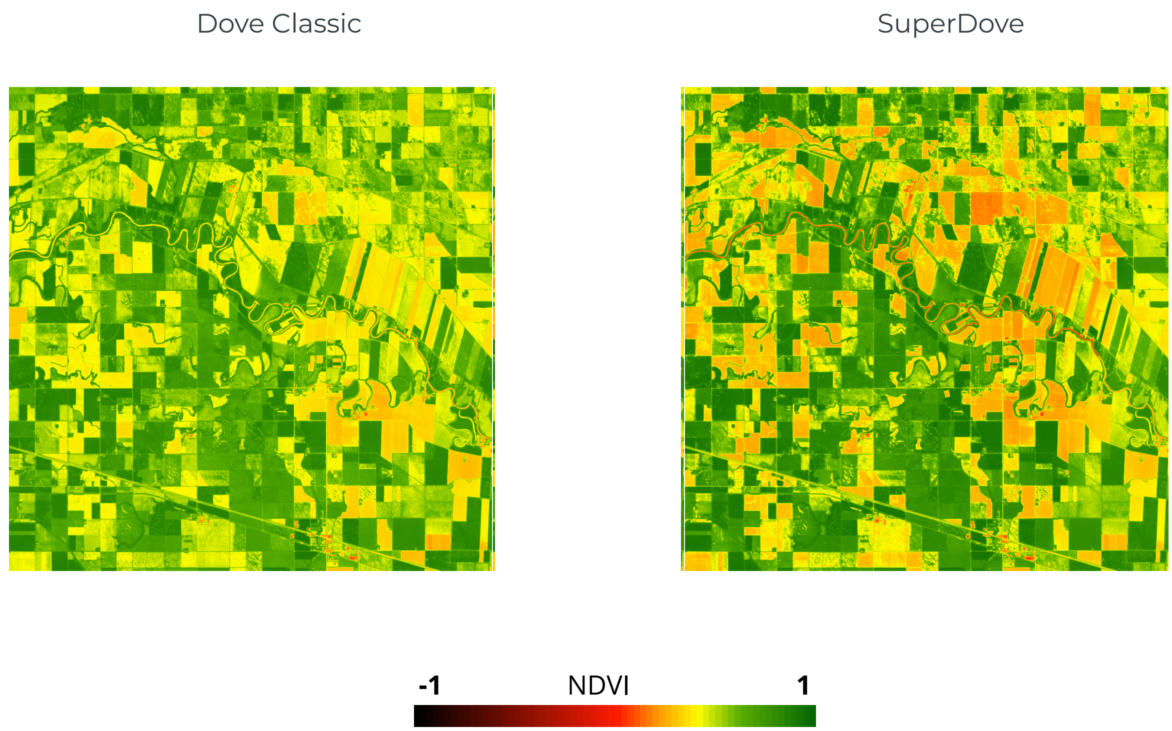
In the following sections we will discuss the potential issues that can occur in Dove imagery where normalization has been applied and show examples of each.

LARGE SPECTRAL RESPONSE DIFFERENCES

All Dove surface reflectance imagery is normalized to the same Sentinel 2 sensor reference spectral response. Because Dove-R and SuperDove relative spectral responses (RSRs) are very similar to Sentinel 2, the normalization for those satellites only requires very small adjustments. However, Dove Classic RSRs are very different from Sentinel 2's and the normalization requires substantial corrections that are constrained by the linear model being used. The result is that Dove Classic data tends to underestimate peak NDVI values and overestimate minimum NDVI values. This leads to small but systematic differences between Dove Classic and other Dove satellites in the normalized products (Figure 6.1).

Figure 6.1: Illustration of systematic differences between Dove Classic and SuperDove normalized NDVI due to differences in their sensor spectral response differences.

Effect of Large RSR Differences



RAPID SEASONAL OR UNEXPECTED CHANGES

The reference for normalization is based on a monthly, seasonal basemap generated from FORCE generated Sentinel 2 surface reflectance. Since the model is not fit exactly to avoid overfitting, the relatively coarse temporal resolution of the model generally causes no issues. However, if there is an extreme mismatch between the model and the image being normalized, normalization can produce inaccurate results. In cases of sudden change that affect the entire scene, such as when there is an unexpected drought or widespread harvest of almost all fields within the image, normalization can overcorrect back to a more "seasonally normal" result. Generally speaking, changes like deforestation, burn scars, snow, clouds, etc are handled well, but sudden regionally-extensive changes can cause issues. Figure 6.2a shows an example of a region where all fields are harvested in a short period of time and the effect that has on the normalized image. Figure 6.2b shows an NDVI time series for a single field in this area, with samples from early in the year through the end of September, illustrating the negative effects on the normalized images both during the rapid growth and the rapid harvesting of the crops.

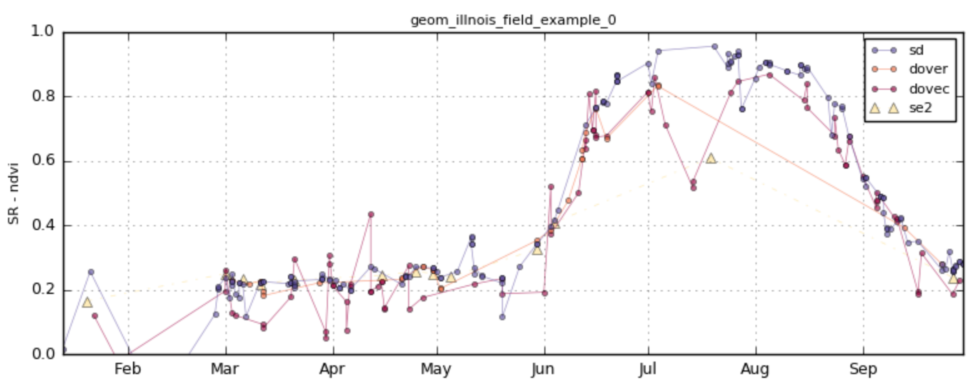
Figure 6.2a: Example of the effect on normalization when there is a widespread, rapid change over a large region.

Effect of Rapid, Widespread Change

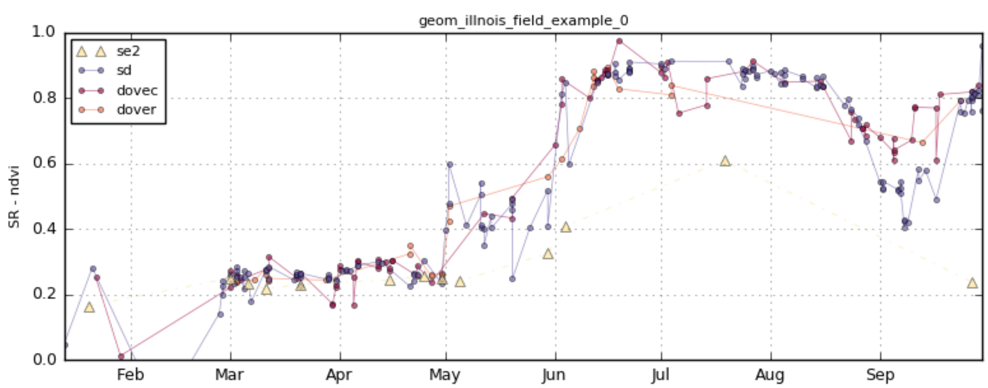
Without Normalization | With Normalization

Figure 6.2b: Example of the effect on a time series analysis when rapid change is present over the same area shown in Figure 6.2a.

No Normalization



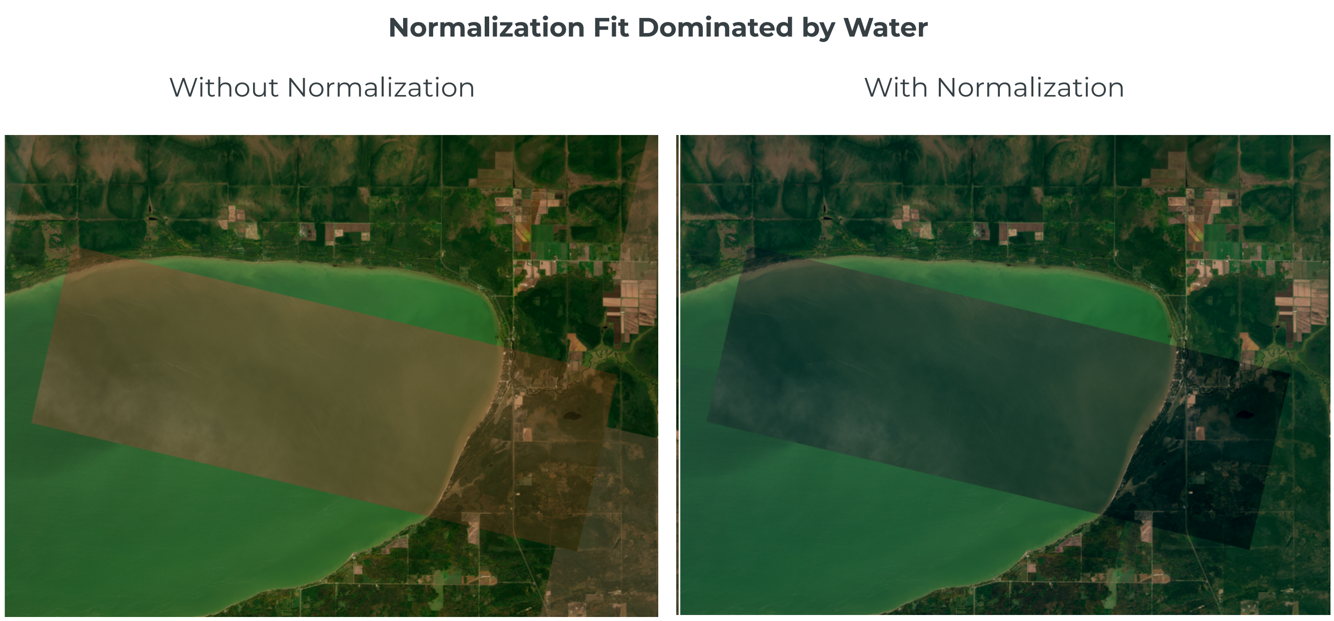
With Normalization



NORMALIZATION FIT DOMINATED BY WATER

When an image has a large body of water, such as a large lake or a coastal area, the water pixels tend to dominate the normalization fit. This will tend to make other scene pixels “more like water” and result in reflectance values with significant errors. Figure 6.3 shows an example in an area dominated by a large lake.

Figure 6.3: Example of the effect on normalization when there is a large body of water covering most of the image.



OVERCORRECTED SURFACE REFLECTANCE

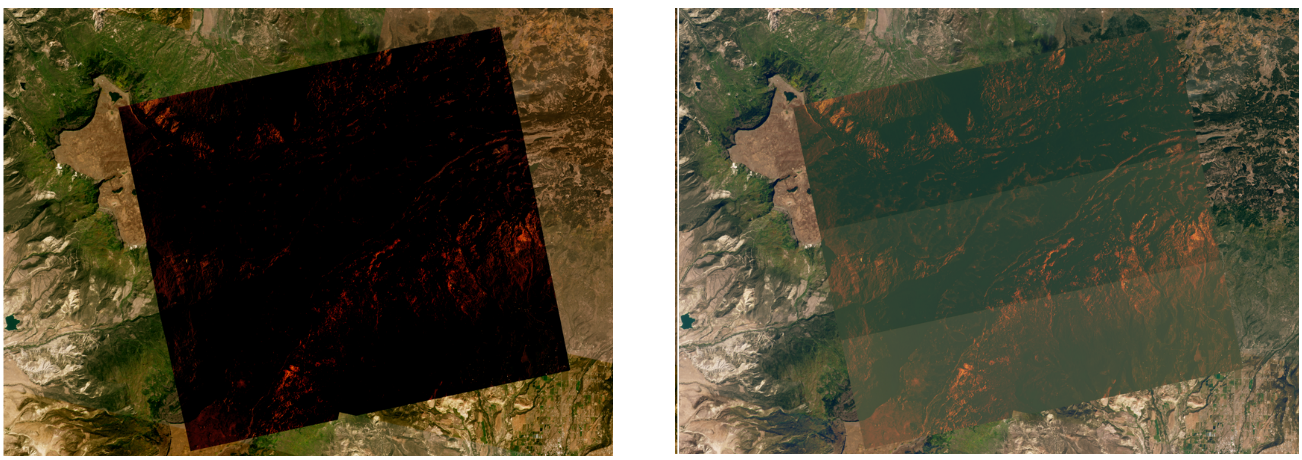
Planet’s surface reflectance products are generated by correcting for the effects on sunlight passing through the atmosphere. These corrections rely on a detailed knowledge of the various components, such as water vapor and aerosol composition, at the time of the image was collected and at its specific location. Because Planet Dove satellites do not have the spectral bands to accurately determine these components, the generation of surface reflectance uses MODIS derived aerosol optical depth (AOD) along with water vapor and ozone concentrations. Any time these inputs to the correction do not reflect the true conditions at the time of the image collect, the image could be overcorrected, sometimes resulting in non-physical negative reflectance values that get set to 1, effectively losing information. It is also known that FORCE-generated surface reflectance is systematically overcorrected at high latitudes and low sun angles, resulting in similar loss of information in the normalization model.

Overcorrection in SR data removes information from the input imagery by clipping negative reflectance values. Normalization cannot fix overcorrection and cannot add back information that was lost due to clipping. Figure 6.4 shows an example where this overcorrection occurs due to a very low sun elevation from an early morning collect. Almost all values in the blue band are clipped to a constant DN of 1 (minimum possible reflectance) by SR processing before input to normalization.

Figure 6.4: Example of normalizing an image where the source reflectance values are the result of overcorrection. In this case normalization will not produce results close to the reference.

Normalization for Overcorrected Image

Without Normalization | With Normalization

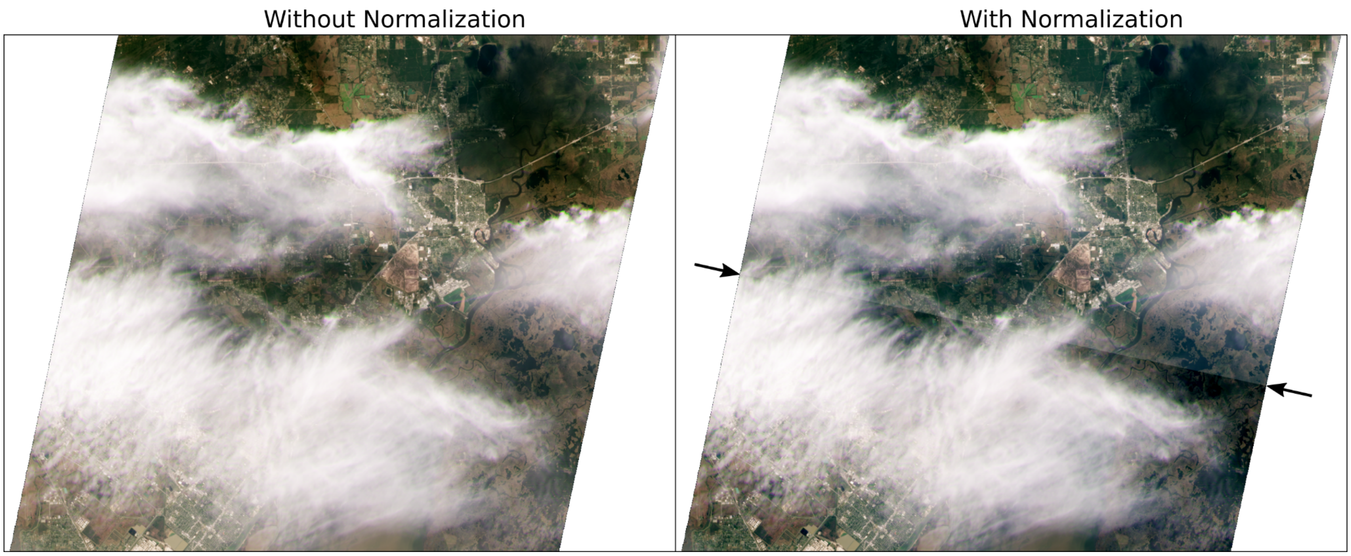


SEAM LINES SCENES WITHIN A STRIP

Normalization is a process applied to individual Dove images and only uses reference data within the image boundary. As a consequence, different normalization corrections will be made between two adjacent scenes from a collect strip. In some cases these corrections are different enough to result in noticeable seam lines between these scenes. Figure 6.5 shows an example of this.

Figure 6.5: Example of normalization producing differences between scenes within the same strip that otherwise would be seamless.

Normalization Produces Differences Within Strips



CONCLUSIONS

The scene level normalization tool significantly reduces variability between scenes and between generations of PlanetScope satellites while maintaining reasonable radiometric accuracy and compatibility with Sentinel2 data.

REFERENCES

Csillik, O., Kumar, P., Mascaro, J., O'Shea, T., & Asner, G. P. (2019). Monitoring tropical forest carbon stocks and emissions using Planet satellite data. *Scientific reports*, *9*(1), 1-12. https://doi.org/10.1038/s41598-019-54386-6

Frantz, D. (2019). FORCE—Landsat+ Sentinel-2 analysis ready data and beyond. *Remote Sensing*, *11*(9), 1124. https://doi.org/10.3390/rs11091124

Kington IV, J. D., Jordahl, K. A., Kanwar, A. N., Kapadia, A., Schönert, M., & Wurster, K. (2019, December). Spatially and Temporally Consistent Smallsat-Derived Basemaps for Analytic Applications. In *AGU Fall Meeting Abstracts* (Vol. 2019, pp. IN13B-0716). https://doi.org/10.6084/m9.figshare.17270117.v1

Teillet, P. M., Fedosejevs, G., Thome, K. J., & Barker, J. L. (2007). Impacts of spectral band difference effects on radiometric cross-calibration between satellite sensors in the solar-reflective spectral domain. *Remote Sensing of Environment*, *110*(3), 393-409. https://doi.org/10.1016/j.rse.2007.03.003