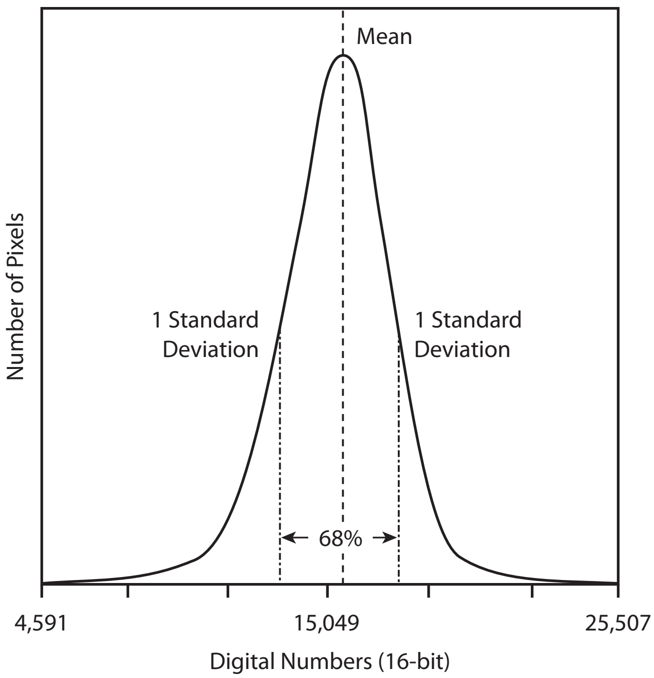

FIGURE 9-3 Landsat OLI band 6 (SWIR1) input histogram with a mean DN of 15,049.

anomalies and patterns of interest. Jensen (2016) provides an in-depth discussion of image statistics.

Univariate Statistics

The histogram for Landsat OLI band 6 (SWIR1) is shown in Figure 9-3. The horizontal axis represents the DNs recorded by the OLI instrument as it orbited across the Thermopolis area. Pixel brightness increases from left to right on the horizontal axis. The minimum brightness is 4,591 and the maximum brightness is 25,507 (the OLI instrument collects 16-bit data or 0 to 65,536 levels of gray). The vertical axis is the number of pixels with a DN value between 4,591 and 25,507. Table 9-1 compiles the minimum, maximum, and mean values for the six Landsat 8 bands in Figure 3-11. The SWIR1 data shown as a histogram in Figure 9-3 is highlighted in gray in Table 9-1. OLI band 2 (blue) has the lowest mean brightness while OLI band 6 (SWIR1) has the highest mean brightness. Contrast stretching will be different for bands with different mean brightness and brightness ranges.

The shape of the histogram is important. Normally distributed data has a bell shaped curve (Figure 9-3). The shape of this curve is expected by many image processing algorithms, however, data that has many dark or bright pixel outliers (far from the mean) is skewed and can degrade processing output. Standard deviation and variance are statistical measures of the dispersion of values about the mean (Jensen, 2016). Variance involves subtracting the DN value for each pixel (or a sample of pixels) from the band's mean DN value, and then squaring the difference. Variance is the sum of the squared differences divided by the number of samples. The SWIR1 band has 6.5 times the variance compared with the blue band (Table 9-1), indicating more spectral richness for the SWIR1 band. Standard deviation (σ) is the common statistical measure of dispersion around the mean and is derived from the positive square root of variance.

$$\text{standard deviation} = \sqrt{\text{variance}} \qquad \textbf{(9-1)}$$

One standard deviation on either side of the mean contains 68% of the population or sample. For our SWIR1 band, 68% (544,000 pixels) of the 800,000 pixels in the image are within one standard deviation of the mean and have DNs between 12,891 and 17,208.

Multivariate Statistics

Multispectral data can have a high band-to-band correlation as the images are collected at the same time over the same area, from the same altitude, and with the same illumination and atmospheric conditions. Data from bands that are highly correlated have less potential when used with image processing algorithms that are designed to bring out subtle spectral patterns, discover anomalies, or highlight spectral targets that cover a small portion of the imaged area.

One band can be compared to another band on a 2-D plot where the horizontal axis represents the range of DNs from one band and the vertical axis represents the range of DNs from the other band (Figure 9-4). This 2-D plot is termed a two-band *feature space plot* (Jensen, 2016). The feature space plot extracts the brightness value from two bands for every pixel in the scene and plots the occurrence in the Landsat OLI 16-bit feature space. The plot can display the frequency of occurrence of unique pairs of values with varying brightness

TABLE 9-1 Band DN statistics associated with Landsat OLI (16-bit) data.

	Band	Minimum	Maximum	Mean	Standard Deviation	Variance
Blue	2	7,442	22,807	9,619.6	845.7	716,018.5
Green	3	6,474	23,233	9,774.9	1,189.9	1,417,120.7
Red	4	5,778	25,165	10,470.3	1,540.4	2,374,469.4
NIR	5	4,420	28,517	13,397.6	1,912.0	3,657,625.1
SWIR1	6	4,591	25,507	15,049.1	2,158.5	4,661,414.4
SWIR2	7	4,777	23,767	12,991.3	2,077.2	4,316,661.9

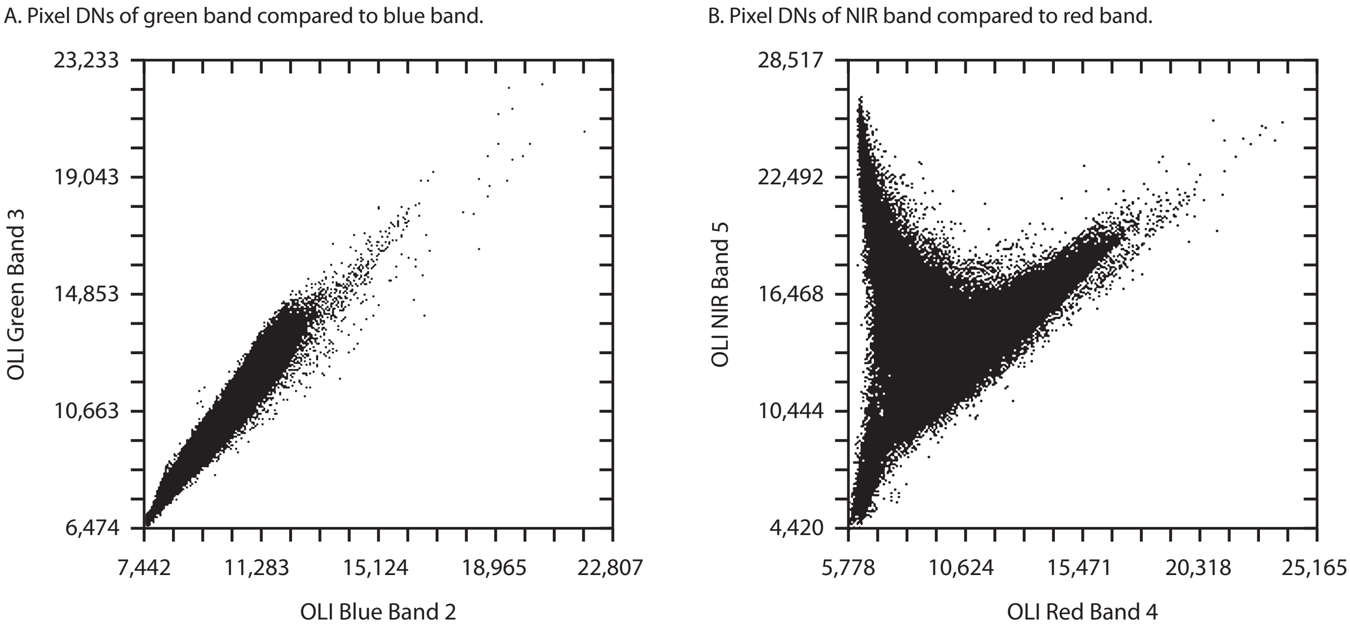

FIGURE 9-4 Feature space plots of Landsat OLI bands.

or color. However, Figure 9-4 is simplified and displays all pairs (whether one deep or hundreds deep) as only one black dot.

Landsat OLI blue band 2 is compared to the green band 3 in the feature space plot of Figure 9-4A. The DNs from the blue band are on the horizontal axis and the DNs from the green band are on the vertical axis. These two visible light bands are highly correlated—where one is bright the other is bright, and vice versa. The bands are highly redundant and that minimizes their potential for new spectral information. In Figure 9-4B, Landsat OLI NIR band 5 (vertical axis) is compared to the red band 4 (horizontal axis). Many pixels in the two bands are not well correlated. Many bright NIR pixels range from dark to bright in the red band. The low correlation indicates there is excellent potential for new spectral information in these two bands. For more information, Figure 9-4A can be evaluated alongside Figures 3-11A and 3-11B, while Figure 9-4B can be evaluated alongside Figures 3-11C and 3-11D.

A correlation matrix based on covariance and standard deviation measurements provides a unitless correlation ratio between bands that varies from –1 to +1. A correlation ratio of 1 indicates a positive, systematic relationship between the brightness values of two bands—as one band increases in brightness the other band's values also increase (Jensen, 2016). A correlation matrix provides a repeatable and defensible number to what is visually interpreted from feature space plots. Table 9-2 is a correlation matrix of the six Landsat bands of the Thermopolis area. The bands used in the feature space plots of Figure 9-4 are highlighted in gray in Table 9-2.

The high correlation seen in Figure 9-4A between the blue and green bands is confirmed by a correlation value of 0.978 in Table 9-2. In contrast, the low correlation seen in Figure 9-4B between the NIR and red bands has a correlation value of 0.423 in Table 9-2. The NIR band contains the most unique spectral information compared to the visible and SWIR bands in the Landsat scene of Thermopolis. The SWIR bands have only moderate correlation with the visible bands (ranging from 0.79 to 0.92 in Table 9-2). The correlation matrix and feature space plots confirm the NIR and SWIR bands will provide the most spectral information for image enhancement techniques.

TABLE 9-2 Correlation matrix of Landsat OLI data.

	Band	Band 2 (Blue)	Band 3 (Green)	Band 4 (Red)	Band 5 (NIR)	Band 6 (SWIR1)	Band 7 (SWIR2)
Blue	2	1					
Green	3	0.978681	1				
Red	4	0.927854	0.962711	1			
NIR	5	0.371338	0.469063	0.423231	1		
SWIR1	6	0.793064	0.837701	0.893752	0.485006	1	
SWIR2	7	0.832312	0.862618	0.926953	0.328594	0.951521	1

Correcting for Atmospheric Scattering

In the preprocessing stage it is important to confirm that your multispectral dataset is in radiance or reflectance. Imagery acquired by aerial and satellite sensors is affected by atmospheric scattering and absorption of light, especially in the shorter wavelengths (Figure 2-26A). Light reflected from the target interacts with scattered and absorbed light from the atmosphere and both are collected by the sensor (Figure 1-18). The combined target reflectance and atmospheric scattering that reaches the detector is called *total radiance*. If the user is processing single-date imagery, comparing maps generated from multi-date imagery, and spectral libraries are not being used, the imagery can remain in total radiance or original DN values provided by the sensor operator (Jensen, 2016). If the user is generating band ratios, comparing spectral measurements from imagery acquired on different dates, building spectral training sites from imagery acquired on different dates, and using spectral libraries, then the imagery needs to be atmospherically corrected and converted into reflectance (Jensen, 2016).

Visual inspection of spectra profiles from multispectral data can help decipher if the data has been atmospherically corrected. In Figure 9-5, the pixels are vegetated and were collected by the Landsat 8 OLI sensor over Haiti. The horizontal axis illustrates the six VNIR-SWIR Landsat bands with blue farthest to the left, while the vertical axis is the DN/brightness value of the pixels. Atmospheric scattering of blue light causes the brightness to be greater in radiance data compared with reflectance data. Green light is reflected by chlorophyll in healthy vegetation while blue and red light are absorbed (Figures 2-26B and 2-31). Reflectance imagery will confirm the absorption of blue and red light with darker DN values compared to the brighter DN values of reflected green light (Figure 9-5).

Absolute Atmospheric Correction

Many complex algorithms and atmospheric models have been developed to convert the digital brightness values recorded by an aerial or satellite multispectral and hyperspectral sensor into scaled surface reflectance values that can be compared to scaled surface reflectance values of any image collected elsewhere on the globe at a different time (Jensen, 2016). A major goal of absolute atmospheric correction is accurate modeling of the atmosphere's properties at the time of image acquisition. NASA's MODIS system collects data daily that can be used for this correction. Jensen (2016) provides details on the processes and software involved in atmospheric correction. Landsat Level-2 science products have pixel brightness atmospherically corrected to reflectance values.

Relative Atmospheric Correction

Chapter 2 described how the atmosphere selectively scatters the shorter wavelengths of light, which causes haze and reduces the contrast ratio of images. For Landsat TM images, band 1 (blue) has the highest component of scattered light and band 7 (reflected SWIR2) has the least. Figure 9-6 shows two techniques for determining the relative correction factor for different TM bands. Both techniques are based on the fact that band 7 is essentially free of atmospheric scattering, which can be verified by examining the DNs for shadows; these typically have very low DN values on band 7. The first technique (Figure 9-6A) employs data from an area within the image that has shadows caused by irregular topography. For each pixel the DN in band 7 is plotted against the DN in band 1, and a straight line is fitted through the plot using a least squares technique. If there was no haze in band 1, the line would pass through the origin. Because there is haze, the intercept is offset along the band 1

A. Landsat OLI radiance values for one vegetated pixel.

B. Atmospherically corrected Landsat OLI reflectance values for one vegetated pixel.

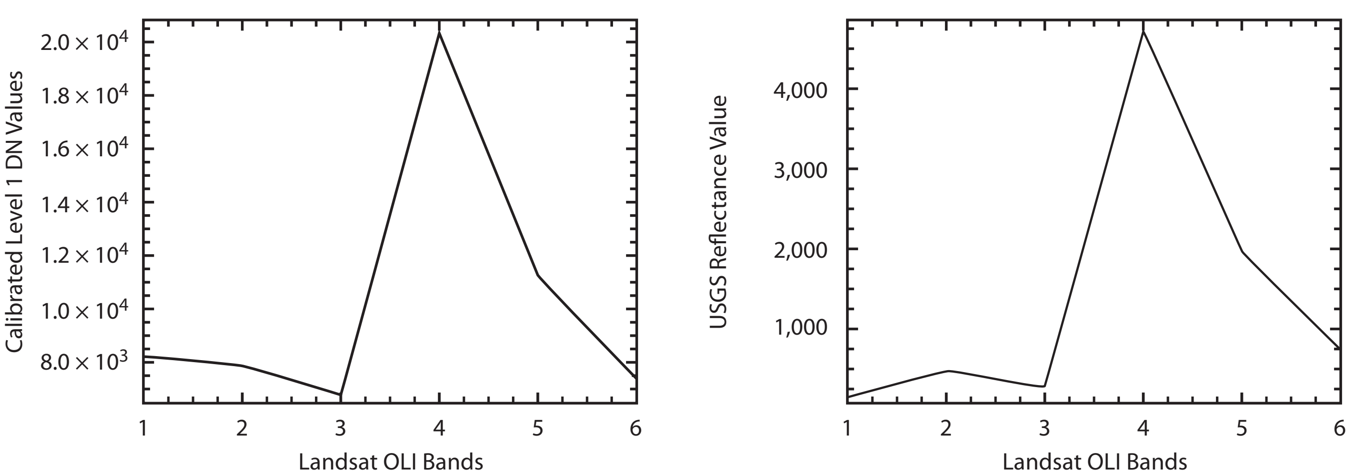

Figure 9-5 Calibrated image DN values compared to atmospherically corrected image reflectance values.