# Data pre-processing in R, part 2/3

## Conversion of raw DN to top-of-atmosphere reflectances

We want to compare two Landsat images to answer our research question. Unfortunately, the two images were taken by different Landsat sensors. They have different gain and offset values as well as a differing radiometric resolution. The raw DN pixel values are thus not comparable. A comparison of data taken by different sensors is only possible after the raw DNs have been converted to physically measurable radiances or – preferably – to reflectance. Radiances quantify the radiation reflected form the Earth's surface and detected by the sensor for a specific band as radiant flux per angle and area (W sr-1 m-2). Reflectance is the ratio between incoming and reflected radiation and hence independent from external factors such as variable sun irradiation. Top-of-atmosphere (TOA) reflectance is the reflectance measured at the sensor.

To convert raw DNs to TOA reflectance, several constants and variables that are provided either in the sensor's documentation or the image meta data are needed. For Landsat 8 OLI data, a conversion of raw DN values to TOA reflectance using eq. 2.1 is quite comfortable because most variable factors have already been considered (see page 54 of Landsat8 Data User Handbook: https://www.usgs.gov/media/files/landsat8datausershandbook).

$$Ref_{TOA} = \frac{raw\ DN * REFLECTANCE\ MULT\ BAND\ x + REFLECTANCE\ ADD\ BAND\ x}{\sin(SEA)} \quad (2.1)$$

Each band has to be multiplied by the band-specific coefficient 'REFLECTANCE MULT BAND X' taken from the meta data file. A second coefficient 'REFLECTANCE ADD BAND X' is subsequently added to the product. To correct for the variable sun elevation angle (SEA), the result is divided by the sinus of the SEA, which is also provided in the meta data.

**Q2.1: Open the meta data file of the Landsat 8 image and identify the coefficients and the SEA.**

For our image, 'REFLECTANCE MULT BAND X' is 2.0000E-05 for all bands and 'REFLECTANCE ADD BAND X' is -0.100000 for all bands. The SEA is 53.37968040°.

Note that in R with a few exceptions all angles have to be converted from degree to radians. Basic mathematical operations can be easily applied to raster objects by treating them as matrix-like objects. The operation is applied individually to each pixel value. Using the clipped raster stack, the conversion can be accomplished in a single line of code:

```
ls82014toa <- (ls82014.subs * 2.0000E-05 - 0.1) / sin (53.37968040 / 180 * pi)
```

For the Landsat 5 TM data, the conversion is a more complex process. First, the raw DNs have to be converted to radiances (eq. 2.2), then the radiances are converted to TOA reflectance.

$$Rad = raw\ DN * RADIANCE\ MULT\ BAND\ x + RADIANCE\ ADD\ BAND\ x \quad (2.2)$$

This time, the rescaling coefficients (listed as RADIANCE_MULT_BAND_X and RADIANCE_ADD_BAND_X in the Landsat 5 image meta data) are really band-specific and we have to apply the conversion for each band separately. A single band in a raster stack object can be addressed by using the [[]] brackets. Using the clipped Landsat 5 raster stack as input data, the conversion to radiances is thus:

```
blue2011rad <- ls52011.subs[[1]] * 0.766 - 2.28583
green2011rad <- ls52011.subs[[2]] * 1.448 - 4.28819
red2011rad <- ls52011.subs[[3]] * 1.044 - 2.21398
nir2011rad <- ls52011.subs[[4]] * 0.876 - 2.38602
swira2011rad <- ls52011.subs[[5]] * 0.120 - 0.49035
swirb2011rad <- ls52011.subs[[6]] * 0.066 - 0.21555
```

The conversion of these radiances to TOA reflectance follows eq. 2.3.

$$Ref_{TOA} = \frac{Rad * \pi * d^2}{ESUN * \cos(SZA)} \tag{2.3}$$

Variable *d* is the distance between Earth and sun in astronomical units for the day of the image acquisition. It can be taken from a table that can be downloaded on the Landsat page (via Landsat7 Data User Handbook on page 80; https://landsat.usgs.gov/sites/default/files/documents/EarthSun_distance.xls).

**Q2.2: Download the table and determine *d* for the acquisition day of the Landsat 5 image.**

*d* = 1.0072409

ESUN is the band-specific mean exoatmospheric irradiance. These variables apply to all Landsat 5 TM images and are not provided in the meta data. Instead, they are published in the literature. Table 2.1 lists the ESUN coefficients taken from Chander & Markham (2003). The sun zenith angle (SZA) can be calculated from the SEA provided in the meta data as SZA = 90-SEA.

```
sza <- (90 - 49.65690900) / 180 * pi
```

*Table 2.1: ESUN coefficients for Landsat 5 TM taken from Chander & Markham (2003).*

| Band | Blue | Green | Red | NIR | SWIR A | SWIR B |
|---|---|---|---|---|---|---|
| ESUN / W/(m² µm) | 1957 | 1826 | 1554 | 1036 | 215 | 80.67 |

The conversion to TOA-reflectance is thus:

```
blue2011ref <- (blue2011rad * pi * 1.0072409^2) / (1957 * cos (sza))
green2011ref <- (green2011rad * pi * 1.0072409^2) / (1826 * cos (sza))
red2011ref <- (red2011rad * pi * 1.0072409^2) / (1554 * cos (sza))
nir2011ref <- (nir2011rad * pi * 1.0072409^2) / (1036 * cos (sza))
swira2011ref <- (swira2011rad * pi * 1.0072409^2) / (215 * cos (sza))
swirb2011ref <- (swirb2011rad * pi * 1.0072409^2) / (80.76 * cos (sza))
```

Finally, we stack the reflectance band to a new raster stack.

```
ls52011toa <- stack (blue2011ref, green2011ref, red2011ref, nir2011ref,
                     swira2011ref, swirb2011ref)
```

The thermal bands of both sensors can only be converted to radiances. The required coefficients are also part of the meta data.

```
the2011rad <- ls5the.subs * 0.055 + 1.18243
thea2014rad <- ls8the1.subs * 3.3420E-04 + 0.10000
theb2014rad <- ls8the2.subs * 3.3420E-04 + 0.10000
```

### Dark pixel subtraction

TOA-reflectance may be affected by atmospheric scattering. Removal of these effects can be achieved with an elaborate atmospheric correction that requires (i) detailed knowledge on the atmosphere at the time of image acquisition and (ii) access to complex radiative transfer models. Since we have neither available, we need to employ a rather pragmatic and less elegant approach called dark pixel subtraction. The assumption behind dark pixel subtraction is that a REALLY dark surface within the image exists. This surface is assumed to reflect none of the incoming light. In consequence, any reflectance measured by the sensor for this surface has to be a result of atmospheric scattering. To remove the atmospheric effects, the reflectance signal measured for these dark pixels is subtracted from each pixel. Of course, the underlying assumption is wrong for multiple reasons. First, REALLY dark pixels hardly exist. Frequently, deep water bodies are used for this purpose but even very deep water bodies reflect some of the incoming radiation and are not really dark. Second, atmospheric scattering has a more complex effect on the spectral signal than resulting in a simple offset that can be subtracted. Still, dark pixel subtraction is a pragmatic approach that is sometimes able to remove effects due to atmospheric scattering.

To apply dark pixel subtraction, we identify for each band the minimum reflectance value. This is achieved by using the minValue() function that returns a vector with the minimum value of each band. If a vector of length *n* is subtracted from a raster stack with *n* bands, the *i*th element of the vector is subtracted from the *i*th band.

```
mn2014 <- minValue (ls82014toa)
ls82014dps <- ls82014toa - mn2014
mn2011 <- minValue (ls52011toa)
ls52011dps <- ls52011toa - mn2011
```

The resulting images are assumed to be free of effects due to the atmosphere.

**Q2.3: Which band had the highest minimum reflectance? Is it the same band in both images? Can you guess why?**

For both images the highest minimum reflectance is observed for the blue band. This reflectance is probably the result of Rayleigh scattering (i.e., scattering of visible light due to gas molecules with a particle size smaller than the wavelength of the scattered light).

### Correction of topographic effects

The Shasta and Trinity lake area is in rugged terrain. Reflectance values in the images are thus affected by shadows due to the topography (Fig. 2.1a). Since these shadows are a function of surface slope, aspect, and the illumination angles (Fig. 2.1b), they can be corrected for with a topographic normalization (Fig. 2.1c). Various approaches for topographic normalization exist; we use the cosine correction, a simple and straight-forward approach for topographic normalization (eq. 2.4).

$$Ref_{topo} = Ref_{dps}\frac{\cos(SZA)}{\cos(SIA)} = Ref_{dps}\frac{\cos(SZA)}{hill\,shade} \tag{2.4}$$

Based on a digital elevation model (DEM) with the same spatial resolution as the imagery and illumination angles (SEA and sun azimuth angle, SAA), the cosine of the sun incidence angle (SIA) is calculated for each pixel. The resulting raster layer corresponds to the well-known hill shade that is often used for 3D illustration purposes. We use the SRTM DEM (SRTM_ffB01_p045r032.tif) with 30 m x 30 m resolution for this purpose. It is required that the pixels of the image and the DEM are exactly matching. Otherwise R is unable to process both data sets together. We therefore have not only to clip the DEM to the same extent but also to resample the pixels. The resample() function is able to do these operations simultaneously. It may, however, take some time until the process is accomplished.

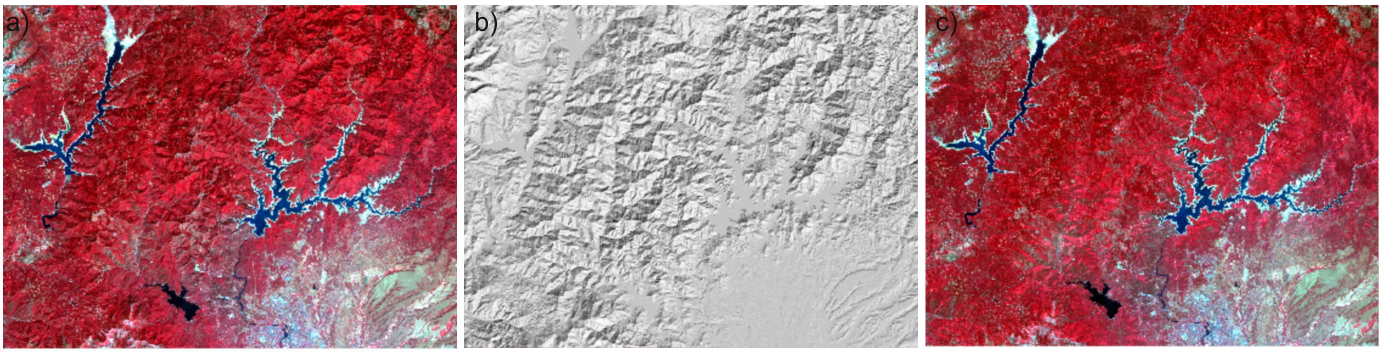


*Fig. 2.1: Topographic normalization. a) shows the TOA-reflectance image before topographic normalization. The cos(SIA) layer is displayed in b). c) shows the normalized image. Shadows induced by the rugged terrain are removed.*

```
dem <- raster ("SRTM_ffB01_p045r032.tif")
dem <- resample (dem, ls82014toa) ## resample and clip the DEM to the image
```

To calculate the hill shade that is identical to the  cosine of the sun incidence angle, we need the slope and aspect of the surface. Both can be calculated with the terrain() function.

```
slp <- terrain (dem, opt="slope")
asp <- terrain (dem, opt="aspect")
```

Both raster layers are then used to calculate the sun incidence angle for the time of image acquisition. Close examination of eq. 2.4 reveals that cos (SIA) = 0 returns Inf as result. We therefore reclassify the hill shade and set values smaller than 0.1 to 0.1.

```
hs2014 <- hillShade (slp, asp, angle=53.37968040, direction=145.67767882)
hs2014 <- reclassify (hs2014, matrix (c (-Inf, 0.1, 0.1), 1, 3))
```

Using the reclassified hill shade, the topographic normalization can be applied.

```
ls82014dps.topo <- ls82014dps * cos ((90 - 53.37968040) / 180 * pi) / hs2014
plotRGB (ls82014dps, 4, 3, 2, stretch=”lin”)
plotRGB (ls82014dps.topo, 4, 3, 2, stretch=”lin”)
```

This operation is repeated for the Landsat 5 TM image.

```
hs2011 <- hillShade (slp, asp, angle=49.65690900, direction=144.56665673)
hs2011 <- reclassify (hs2011, matrix (c (-Inf, 0.1, 0.1), 1, 3))
ls52011dps.topo <- ls52011dps * cos ((90 - 49.65690900) / 180 * pi) / hs2011
```

**Q2.4: Compare the result of the topographic normalization with the original (clipped) images. Has the quality of the images been improved? Are clouds present in the images that have to be masked?**

There is a small cloud in the Landsat 5 image in the northern parts of lake Shasta.

### References

Chander G, Markham B (2003). Revised Landsat-5 TM radiometric calibration procedures and postcalibration dynamic ranges. IEEE Transactions on Geoscience and Remote Sensing 41, 2674-2677.