Satellite Sensor Data Normalization Issues
A User Perspective

Prasad S. Thenkabail
Research Geographer, U.S. Geological Survey (USGS), Flagstaff, Arizona
Lecture @ Boston University, Boston, USA. October 27-29, 2009
Landsat Science Team Data Normalization Workshop
Overview of Today’s Lecture
Two Main Questions that a Sensor Data User Often Asks

1. How do we get “normalized data” for one sensor over time?
2. How do we get “normalized data” for multiple sensors over time?

.........these are 2 questions that always a sensor data user asks (often, without real answers)
Satellite Sensor Data Normalization Issues

What we Mean by “normalized data”?

A. at-sensor reflectance
Corrections for: (a) sensor degradation changes, (b) solar elevation, (c) band-width (spectrum at which irradiance is received), (c) Earth-sun distance.

B. Surface reflectance
Corrections for atmospheric effects: (a) cloud removal composite, (b) haze removal.

C. Inter-sensor Calibrations
Corrections for: (a) pixel resolution (e.g., 30m vs. 80m), (b) band width (e.g., broad-band vs. narrow-band), (c) radiometer (e.g., 8-bit vs. 11-bit).
Data Normalization Issues

1. at-sensor reflectance

Well understood….quite Straightforward…..yet data providers still do not provide this as a product….making Users life difficult
### Satellite Sensor Data Normalization Issues

#### What to Normalize for?

1. **Satellites**
   - Height of acquisition (e.g., 500 km, 700 km, 36,000 km above earth)
   - Orbital parameters

2. **Sensors**
   - Radiometry
   - Band-width
   - Optics/design
   - Degradation over time
   - Nadir, off-nadir viewing

3. **Solar flux or irradiance**
   - Function of wavelength

4. **Sun**
   - Sun elevation @ time of acquisition

5. **Sun-Earth**
   - Distance between earth and sun

6. **Stratosphere or Atmosphere**
   - Ozone, water vapor, haze, aerosol
   - Path radiance

7. **Surface of Earth**
   - Topography

8. **Seasons**
   - Earth-sun distance

---

**Atmospheric corrections**

- Haze (atmospheric)
- Haze (dust)
- Haze (harmattan)
Satellite Sensor Data Normalization Issues

What to Normalize for?: e.g., Data in Digital Numbers vs. Surface Reflectance

Radiometric differences across sensors clearly imply the need for normalizations.
Example: To Convert he ETM+ 8 bit DNs to radiances:

\[
\text{Radiance (W m}^{-2} \text{ sr}^{-1} \mu\text{m}^{-1}) = \text{gain} \times \text{DN} + \text{offset}
\]

Note: see data header files for gains and offsets

Spectral radiance

Spectral radiance (Price, 1987) is computed using the following equation:

\[ R_i = \alpha_i \cdot DNi + \beta_i \rightarrow (1) \]

- \( R_i \) = spectral radiance in \( \text{W m}^{-2} \, \text{\mu m}^{-1} \)
- \( \alpha_i \) = gain or slope in \( \text{W m}^{-2} \, \text{\mu m}^{-1} \)
- \( \beta_i \) = bias or intercept in \( \text{W m}^{-2} \, \text{\mu m}^{-1} \)
- \( DNi \) = digital number of each pixel in TM bands
- \( i = 1 \) to 5 and 7 (except the thermal band 6)

Some References:
Satellite Sensor Data Normalization Issues

Radiance (W m$^{-2}$ sr$^{-1}$ µm$^{-1}$) to at-sensor Reflectance (%)

\[
\text{Reflectance (\%)} = \frac{\pi L_{\lambda} d^2}{ESUN_{\lambda} \cos \theta_S} \times 100
\]

Where, TOA reflectance (at-sensor or at-satellite exo-atmospheric reflectance)

\(L_{\lambda}\) is the radiance (W m$^{-2}$ sr$^{-1}$ µm$^{-1}$),
\(d\) is the earth to sun distance in astronomic units at the acquisition date (see Markham and Barker, 1987),

\(ESUN_{\lambda}\) is irradiance (W m$^{-2}$ sr$^{-1}$ µm$^{-1}$) or solar flux (Neckel and Labs, 1984), and
\(\theta_s\) = solar zenith angle

Note: \(\theta_s\) is solar zenith angle in degrees (i.e., 90 degrees minus the sun elevation or sun angle when the scene was recorded as given in the image header file).
Satellite Sensor Data Normalization Issues

Solar Irradiance or Solar Flux ($W m^{-2} sr^{-1} \mu m^{-1}$) (e.g., across electromagnetic spectrum)
### Table 2. Solar flux or exo-atmospheric irradiances (W m\(^{-2}\) \(\mu m^{-1}\)) for Landsat-5 TM wavebands (Markham and Barker, 1985).

<table>
<thead>
<tr>
<th>Band</th>
<th>Solar Flux or exo-atmospheric irradiances (W m(^{-2}) (\mu m^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1946.48</td>
</tr>
<tr>
<td>2</td>
<td>1812.63</td>
</tr>
<tr>
<td>3</td>
<td>1545.95</td>
</tr>
<tr>
<td>4</td>
<td>1046.70</td>
</tr>
<tr>
<td>5</td>
<td>211.12</td>
</tr>
<tr>
<td>6</td>
<td>10.000</td>
</tr>
<tr>
<td>7</td>
<td>76.91</td>
</tr>
</tbody>
</table>
### Satellite Sensor Data Normalization Issues

**Astronomical Units (dimensionless) for Earth-Sun Distance**

<table>
<thead>
<tr>
<th>Julian Day</th>
<th>Distance</th>
<th>Julian Day</th>
<th>Distance</th>
<th>Julian Day</th>
<th>Distance</th>
<th>Julian Day</th>
<th>Distance</th>
<th>Julian Day</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.9832</td>
<td>74</td>
<td>.9945</td>
<td>152</td>
<td>1.0140</td>
<td>227</td>
<td>1.0128</td>
<td>305</td>
<td>.9925</td>
</tr>
<tr>
<td>15</td>
<td>.9836</td>
<td>91</td>
<td>.9993</td>
<td>166</td>
<td>1.0158</td>
<td>242</td>
<td>1.0092</td>
<td>319</td>
<td>.9892</td>
</tr>
<tr>
<td>32</td>
<td>.9853</td>
<td>106</td>
<td>1.0033</td>
<td>182</td>
<td>1.0167</td>
<td>258</td>
<td>1.0057</td>
<td>335</td>
<td>.9860</td>
</tr>
<tr>
<td>46</td>
<td>.9878</td>
<td>121</td>
<td>1.0076</td>
<td>196</td>
<td>1.0165</td>
<td>274</td>
<td>1.0011</td>
<td>349</td>
<td>.9843</td>
</tr>
<tr>
<td>60</td>
<td>.9909</td>
<td>135</td>
<td>1.0109</td>
<td>213</td>
<td>1.0149</td>
<td>288</td>
<td>.9972</td>
<td>365</td>
<td>.9833</td>
</tr>
</tbody>
</table>
Satellite Sensor Data Normalization Issues

at-sensor Reflectance (%)

Allows us to compare across Sensors
Satellite Sensor Data Normalization Issues

at-sensor Reflectance (%) Model for Landsat ETM+ written in ERDAS Imagine

Dis-advantages of NOT providing data in Reflectance

1. Not all users want to do this;
2. Not all users have expertise to do this;
3. It is time-consuming;
4. Often users may end up using just digital numbers—leading to serious issues with data interpretation;
5. Providing data in reflectance is a big step forward.
Satellite Sensor Data Normalization Issues
at-sensor Reflectance (%) Model for IKONOS written in ERDAS Imagine

Dis-advantages of NOT providing data in Reflectance
1. Not all users want to do this;
2. Not all users have expertise to do this;
3. It is time-consuming;
4. Often users may end up using just digital numbers - leading to serious issues with data interpretation;
5. Providing data in reflectance is a big step forward.
Satellite Sensor Data Normalization Issues
at-sensor Reflectance (%) Model for Hyperion (band 1-70) written in ERDAS Imagine

Dis-advantages of NOT providing data in Reflectance
1. Not all users want to do this;
2. Not all users have expertise to do this;
3. It is time-consuming;
4. Often users may end up using just digital numbers- leading to serious issues with data interpretation;
5. Providing data in reflectance is a big step forward.
### Satellite Sensor Data Normalization Issues

#### At-sensor Reflectance

1. Quite reliable;
2. A must;
3. Most will agree;
4. Good that the satellite data provider provides this instead of making a user convert this.
Data Normalization Issues

2. Surface Reflectance

Clouds......Haze......Confusion......Uncertainty......need clear decisions
Data Normalization Issues

2A. Cloud Removal algorithms

Cloud removal..............data loss........but provides cloud free
data..............only time-compositing over time (e.g., 8-day, monthly)
provides some useful data
Satellite Sensor Data Normalization Issues
Cloud Removal Algorithms

1. Maximum Value NDVI compositing;
2. Blue band reflectivity threshold;
3. Visible band reflectivity threshold; and
4. MODIS First 5 Band reflectivity threshold;
Satellite Sensor Data Normalization Issues

8-day time composites of MODIS 250m Surface Reflectance Product

Observe Clouds in Each 8-day Composite

FCC (RGB): 2,1,6 (NIR,red,SWIR1)
Satellite Sensor Data Normalization Issues

Monthly Maximum Value (NDVI) composite from 8-day time composites of MODIS 250m Surface Reflectance Product to reduce cloud cover

Clouds are about zero!.

Monthly Maximum Value composite (MVC) image: derived from four 8-day composite images
Significant clouds scenario July 27 image of Krishna basin. When reflectance (percent) in band 1 and band 2 and band 3 and band 4 and band 5 is all > 20 percent cloud is present (red areas in right image) else no cloud is present (blue areas in left image). Based on this definition, the image had a high percent of clouds on July 27. The left image is a FCC (RGB) of MODIS bands 2, 1, 6 (858 nm, 648 nm, and 1640 nm) and shows significant clouds. Each of the first 5 bands should have > 20 percent reflectance for cloud to be present. Thereby the formulae in ERMapper is:

\[
\text{If } (i_1 > 20 \text{ and } i_2 > 20 \text{ and } i_3 > 20 \text{ and } i_4 > 20 \text{ and } i_5 > 20) \text{ then 255 else null}
\]
Satellite Sensor Data Normalization Issues
First 5 Band (of MODIS 7 band Reflectance product) composite to reduce cloud cover

If \((i_1 > 20 \text{ and } i_2 > 20 \text{ and } i_3 > 20 \text{ and } i_4 > 20 \text{ and } i_5 > 20)\) then 255 else null

**No cloud scenario**: April 30 image of Krishna basin. When reflectance (percent) in band 1 and band 2 and band 3 and band 4 and band 5 is all > 20 percent cloud is present (red areas in left image) else no cloud is present (blue areas in left image). Based on this definition, left image had zero cloud on April 30. The right image is a FCC (RGB) of MODIS bands 2, 1, 6 (858 nm, 648 nm, and 1640 nm) and shows little or no clouds. Each of the first 5 bands should have > 20 percent reflectance for cloud to be present. Thereby the formulae in ERMapper is:

If \((i_1 > 20 \text{ and } i_2 > 20 \text{ and } i_3 > 20 \text{ and } i_4 > 20 \text{ and } i_5 > 20)\) then 255 else null
Satellite Sensor Data Normalization Issues
Blue Band Minimum Reflectivity Threshold for Cloud Removal

2. **Blue band minimum reflectivity threshold**
   
   If (blue band > 21% reflectance) then null else I

3. **Visible band minimum reflectivity threshold**
   
   If (blue band > 22% reflectance and green band > 21% reflectance and red band > 23% reflectance) then null else I

Results of the first Algorithm
Satellite Sensor Data Normalization Issues

Surface Reflectance: (a) cloud removal

1. Cleans up cloud areas and provides clean data......but data loss;
2. Time compositing (e.g., 8-day, monthly) useful;
3. Cloud removal algorithms does not address haze;
Data Normalization Issues

2B. Atmospheric correction ("eliminate or reduce path radiance" resulting from haze (thin clouds, dust, harmattan, aerosols, ozone, water vapor))
Satellite Sensor Data Normalization Issues

What to Normalize for?

Energy off Target (%)
Reflectance (%) = ............................................
Energy from the Source (%)

Radiance (Wm^{-2}sr^{-1}µm) @ TOA
= Radiance leaving the Ground * Transmission factor + path radiance.

One pass on days: D+10  D+5  D  D-5

Swath observed
60 km

Radiance (Wm^{-2}sr^{-1}µm) @ TOA
= Radiance leaving the Ground * Transmission factor + path radiance.

U.S. Geological Survey
U.S. Department of Interior

Note: Transmission factor assumed 1 except in 6S model. Also in arid and semi-arid regions, it is anyway nearly 1.
Atmospheric correction (“eliminate or reduce path radiance” resulting from haze (thin clouds, dust, harmattan, aerosols, ozone, water vapor))

1. Dark object subtraction technique (Chavez et al.);
2. Improved dark object subtraction technique (Chavez-Milton);
3. Radiometric normalization technique: Bright and dark object regression or (Elvidge et al.); and
4. 6S model (Vermote et al.).
The starting Haze value in NIR band of right image is 9 compared with 1 for the left image in NIR. This is indicative of haze in right image.

**Correction:**
1. simply deduct SHV in right image from each band,
2. Radiometrically correct the right image (haze affected) image to the left image (clear image).
Satellite Sensor Data Normalization Issues

Atmospheric Corrections: Simple dark-object subtraction Technique based on blue band

The starting Haze value in blue band of right image is 73 compared with 62 for the left image in NIR. This is indicative of haze in right image.

**Correction:**
1. simply deduct SHV in right image from each band,
2. Radiometrically correct the right image (haze affected) image to the left image (clear image).
Satellite Sensor Data Normalization Issues

Atmospheric Corrections: Improved dark-object subtraction Technique based on Starting Haze Value in Blue Band

The Chavez procedure uses a number of relative scattering models for different atmospheric conditions. The characteristic of the model:

1. Scattering is wavelength dependant (e.g., Rayleigh scattering); Shorter the wavelength greater the scattering theory;
2. Choose a starting haze value (SHV). Blue band preferred, but green band maybe practical as blue band may over correct;
3. Chavez techniques allows the use of digital numbers as SHV;
4. Model can be worked on a spreadsheet. All you need to do is to provide SHV;
5. The end result is a SHV for all bands from the model that will be used to correct each band of each image (unless it is a clear image)
6. For your study area select all images and categorize them as below.

<table>
<thead>
<tr>
<th>Atmospheric conditions</th>
<th>Exponent of Relative scattering model</th>
<th>TM digital number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very clear</td>
<td>$\lambda^{-4}$</td>
<td>&lt;55</td>
</tr>
<tr>
<td>Clear</td>
<td>$\lambda^{-2}$</td>
<td>56-75</td>
</tr>
<tr>
<td>Moderate</td>
<td>$\lambda^{-1}$</td>
<td>76-95</td>
</tr>
<tr>
<td>Hazy</td>
<td>$\lambda^{-0.7}$</td>
<td>96-115</td>
</tr>
<tr>
<td>Very hazy</td>
<td>$\lambda^{-0.5}$</td>
<td>&gt;115</td>
</tr>
</tbody>
</table>

SHV This stands for the ‘starting haze value’. This is the DN value at which the histogram in a short-wavelength band (usually TM band 1) begins to leave the baseline (see figure below).

Band This is the band from which the SHV is chosen.

Satellite Sensor Data Normalization Issues
Atmospheric Corrections: Radiometric Normalization using the Brightest and Darkest Objects

Regressions: Reference a very clear image (say 1998) to all other images (e.g., 1986 illustrated here) that are relatively hazy.

Note: Second Simulation of the Satellite Signal in the Solar spectrum (6S)

Data needed for the model

From image header files:
- Geometry
- Spectral conditions

Atmospheric information from NVAP and TOMS (course)
- Ozone
- Water vapor concentrations
- Haze
- Aerosols

Are these input model data measured at time of acquisition of the image?

Are these input model data measured at appropriate pixel resolutions?

Satellite Sensor Data Normalization Issues

Surface Reflectance: (a) haze removal

1. Useful data removed?;
2. Over-correction in some places and under-correction in others?;
3. Validation (globally) is key to making this work;
4. Probably, using more than 1 method and cross comparison (apart from point 3) will bring reliability.
Data Normalization Issues

3. Overarching correction using time-invariant sites
Satellite Sensor Data Normalization Issues

Normalize based on time-Invariant Site (e.g., Sahara Desert)

### Calibration factor (method 1) for NDVI

<table>
<thead>
<tr>
<th>NDVI-Calibration factor</th>
<th>Time (Start:July, 1981; End:September, 2001: month by month)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.2</td>
<td>Note: getting a perfect black body within a Landsat image is not easy. This method ideal for large area studies.</td>
</tr>
<tr>
<td>-0.15</td>
<td>Time-invariant sites are ideal to correct for atmospheric as well as sensor related and time related issues. In this way, it is quite an holistic correction technique quite ideal. However, getting a time-invariant site (e.g., site like Sahara desert where reflectance is expected to be constant) is not easy within a Landsat scene area. This approach is ideal for large areas.</td>
</tr>
<tr>
<td>-0.1</td>
<td></td>
</tr>
<tr>
<td>-0.05</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td></td>
</tr>
</tbody>
</table>

1. Establish reflectance factor (fc) for time-invariant site (plot below):
   Calculate calibration factor (fc) for every band and for every date by dividing the reflectance of “this date of given band” with “long term reflectance (e.g., 20 years) of same band”;

2. Use reflectance factor (fc) of time-invariant site (plot below) to multiply with entire image of corresponding dates.
Satellite Sensor Data Normalization Issues

Surface Reflectance: (c) time-invariant sites

1. Very difficult to get time-invariant sites within landsat scene;
2. How “time invariant” are “time invariant sites”?
3. Validation (with ground based measurements) is required for reliability of results.
Data Normalization Issues

4. Overarching correction using Spectral matching Techniques
Spectral Matching Technique: Ground measured vs. Satellite measured

Spectral Measurements made at ground (no atmospheric effects) using a spectroradiometer... exactly at same time as Satellite Overpass (with atmospheric effects).........then “match” ground spectra (no atmospheric effect) with satellite sensor spectra (atmospheric effect........have several 100 or 1000 global ground stations (attached to climate stations?)

2. Spectral matching and rectification
   A. best technique
   B. needs resources
Satellite Sensor Data Normalization Issues
Surface Reflectance: (d) spectral matching technique

1. This will be ideal to correct for “everything”;
2. Costly;
3. But doable if we can tie with global meteorological stations.
Data Normalization Issues
5. Derived products for Correction
Satellite Sensor Data Normalization Issues

Normalize based on Derived Products (e.g., NDVI)

Note: The idea here is that derived products like NDVI ought to be same for same biomass (example) over clear and hazy areas (or other differences like topography) through corrections.

### 1. Atmosphere

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>NIR</th>
<th>NDVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paddy</td>
<td>28</td>
<td>132</td>
<td>0.65</td>
</tr>
<tr>
<td>Paddy</td>
<td>32</td>
<td>149</td>
<td>0.65</td>
</tr>
</tbody>
</table>

### 2. Topography

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>NIR</th>
<th>NDVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paddy</td>
<td>19</td>
<td>164</td>
<td>0.79</td>
</tr>
<tr>
<td>Paddy</td>
<td>17</td>
<td>145</td>
<td>0.79</td>
</tr>
</tbody>
</table>
Data Normalization Issues

6. Inter-sensor Calibrations
Satellite Sensor Data Normalization Issues
What Happens when Sensors Migrate (e.g., AVHRR to MODIS)

Develop inter-sensor relationships for obtaining continuous time-series data when we migrate from one sensor to another.

\[ y = 0.7633x - 0.0483 \]

\[ R^2 = 0.7793 \]
Satellite Sensor Data Normalization Issues

What Happens when Sensors Migrate (e.g., AVHRR to MODIS)

Apply inter-sensor relationships for obtaining continuous time-series data when we migrate from one sensor to another.
Data Normalization Issues
7. Inter-sensor Calibrations
Satellite Sensor Data Normalization Issues
Multiple Sensors: How do we Address Sensor of various pixel-resolutions?

**Note 1**: all datasets geolinked to 4 m IKONOS (which is not in full resolution)
Satellite Sensor Data Normalization Issues
Multiple Sensors: How do we Address Sensor of various band-widths?

Broad-band (e.g., ETM+) vs. Narrow-band (e.g., MODIS)
Lead to differences in radiance measured off the same target.
<table>
<thead>
<tr>
<th>Sensor</th>
<th>NDVI Range</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>IKONOS</td>
<td>0 to 0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>ALI</td>
<td>-0.1 to 0.67</td>
<td>0.68</td>
</tr>
<tr>
<td>ETM+</td>
<td>-0.17 to 0.45</td>
<td>0.62</td>
</tr>
<tr>
<td>Hyperion</td>
<td>-0.2 to 0.62</td>
<td>0.82</td>
</tr>
</tbody>
</table>

(a) Broad-bands at NIR and red; (b) 11-bit data
(a) Broad-bands at NIR and red; (b) 16-bit data
(a) Narrow-bands at NIR and red; (b) 8-bit data

U.S. Geological Survey
U.S. Department of the Interior
Satellite Sensor Data Normalization Issues

Inter-sensor comparisons so that we can use multiple-sensor data in analysis

IKONOS: Feb. 5, 2002 (hyper-spatial)

ETM+: March 18, 2001 (multi-spectral)

ALI: Feb. 5, 2002 (multi-spectral)

Hyperion: March 21, 2002 (hyper-spectral)
ETM+ NDVI = 0.8694 * IKONOS NDVI - 0.1908

$R^2 = 0.68$

Sudan savanna
Derived savanna
Humid forests
All three ecoregions

Linear (all three ecoregions)

Kayawa village, Northern Guinea Savanna, Nigeria (Cyan in plots below)
Eglime, Derived Savanna, Benin (green in plots below)

Akok village, Humid Forests, Cameroon (magenta in plots below)

Eco-regions from which the Data for plots is taken

Satellite Sensor Data Normalization Issues
Inter-sensor relationships: ETM+ vs. IKONOS acquired on same Dates in Different Eco-regions

U.S. Geological Survey
U.S. Department of Interior
1. at-sensor reflectance

is a must as a minimum for all future Landsat and/or other satellite sensor data delivery;

2. Surface reflectance

will be ideal..... But there are issues that needs to be discussed before we take this route. How reliable is it?............this maybe acceptable route to take, if we have ground calibration and validation (but is that feasible?);

3. Mosaics

We should consider delivering Landsat data as mosaics (e.g., country, state);

4. Metadata

should include precise locations of time-invariant sites, darkest object, brightest object?.
Data normalization should be more holistic...we should think of not Landsat sensor alone, but all sensor data...but Landsat could set the standards...this will enable user to use data from multiple sensors for their applications with true understanding of inter-sensor relationships....