

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/303096538>

The impact of improved signal-to-noise ratios on algorithm performance: Case studies for Landsat class instruments

Article in *Remote Sensing of Environment* · May 2016

DOI: 10.1016/j.rse.2016.04.015

CITATIONS

20

READS

218

7 authors, including:



Aaron Gerace

Rochester Institute of Technology

42 PUBLICATIONS 536 CITATIONS

[SEE PROFILE](#)



Curtis E. Woodcock

Boston University

275 PUBLICATIONS 32,432 CITATIONS

[SEE PROFILE](#)



Zhe Zhu

University of Connecticut

55 PUBLICATIONS 7,779 CITATIONS

[SEE PROFILE](#)



Randolph Wynne

Virginia Polytechnic Institute and State University

175 PUBLICATIONS 10,219 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Mapping and Characterizing Human Activity Changes using NASA Black Marble Product Suite [View project](#)



EWMACD [View project](#)



Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

The impact of improved signal-to-noise ratios on algorithm performance: Case studies for Landsat class instruments

John R. Schott^a, Aaron Gerace^{a,*}, Curtis E. Woodcock^{a,b}, Shixiong Wang^{a,b}, Zhe Zhu^{a,b}, Randolph H. Wynne^{a,c}, Christine E. Blinn^{a,c}

^a Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Dr, Rochester, NY 14623, United States

^b Department of Earth and Environment and Center for Remote Sensing, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, United States

^c Department of Forestry, Virginia Polytechnic Institute and State University, 319 Cheatham Hall, Blacksburg, VA 24061, United States

ARTICLE INFO

Article history:

Received 11 May 2015

Received in revised form 6 April 2016

Accepted 26 April 2016

Available online xxx

ABSTRACT

The Landsat Operational Land Imager (OLI) has 5 to 10 times better signal-to-noise ratios (SNRs) in all spectral bands than previous Landsat instruments. SNR performance has long been recognized as a value in instrument design, however, the impact on algorithm performance for earth science applications is poorly documented. Since SNR performance may drive design/cost tradeoffs on future missions, a set of experiments were designed to evaluate the impact of various SNR levels on algorithms applied to different science applications. The application areas studied spanned a wide range including water quality, land cover and forestry. The experiments involved producing data sets with a range of signal-dependent SNR values ranging from Landsat 7 ETM+ levels to OLI levels. Algorithms were then run on these otherwise identical data sets and evaluation metrics applied to evaluate the relative performance versus SNR. In all cases, performance was shown to be a strong function of SNR with substantial increase in performance as SNR increased (e.g. constituent retrieval errors reduced by a factor of three). However, in some cases, the rate of increase slowed at higher SNR levels. Regrettably, the point of diminishing returns was not the same for all applications leaving significant burden on design teams to decide which application's needs could be fully met in terms of SNR requirements.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

The Operational Land Imager (OLI) onboard Landsat 8 represents a marked improvement in medium resolution earth imaging (Irons, Dwyer, & Barsi, 2012). Following in the Landsat Worldwide reference system (WRS) II orbit with 30 m ground sample distance (GSD), the data look very similar to Landsat's 5 and 7. However, unlike the TM instruments the OLI is a push-broom instrument with approximately 6000 across-track detectors. As a result the detectors can dwell on each spot on the earth longer, significantly improving signal-to-noise (Schott, 2007, Chapter 6). To take advantage of this improved signal-to-noise ratio (SNR), OLI has an increased bit depth of 12 compared to the Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) which had 8-bit data and the Landsat MSS sensor which had 6-bit data. OLI also has two additional bands and slightly narrower spectral bandwidths as seen in Fig. 1.

The correlation between higher SNR's and image quality is well recognized. In the remote sensing community this is perhaps most clearly expressed in the General Image Quality Equation (GIQE). The GIQE is designed to relate image metrics such as pixel size on the ground,

edge sharpness, and SNR to the interpretability of black and white remotely sensed images (Leachtenauer et al., 1997). The GIQE suggests that while image interpretability is a function of SNR it is nowhere near as large a factor as pixel size or image sharpness and that after a point increasing SNR will not have an observed impact for photo interpretation purposes. Fiete and Tantaló (2001) took this further in terms of the noise equivalent change in reflectance (NE Δ R), showing that interpretability was linearly related to NE Δ R.

There is not as clear a body of literature relating SNR to the utility of multispectral data. In part, this is because it typically requires very high SNR data that can then be degraded so that images, whose only change is in SNR, can be compared through some analytical application algorithm. Williams et al. (1984) did just this after Landsat switched from the 6-bit MSS sensors to the 8-bit TM sensors. They showed that degrading 8-bit TM data to 6-bit data (i.e. roughly a factor of 4 in SNR) resulted in a loss of approximately 5 percentage points in land cover classification accuracy. Other studies have shown that improved SNR can improve preprocessing algorithms for hyper-spectral data. Seidel et al. (2008), showed that aerosol optical depth retrieval improved with higher SNR and Chen (2006) shows that hyperspectral end member selection algorithms perform better at higher SNR. Kruse (2000) looked at AVIRIS imaging spectrometer data and found that the classification accuracy and number of minerals mapped improved with SNR.

* Corresponding author.

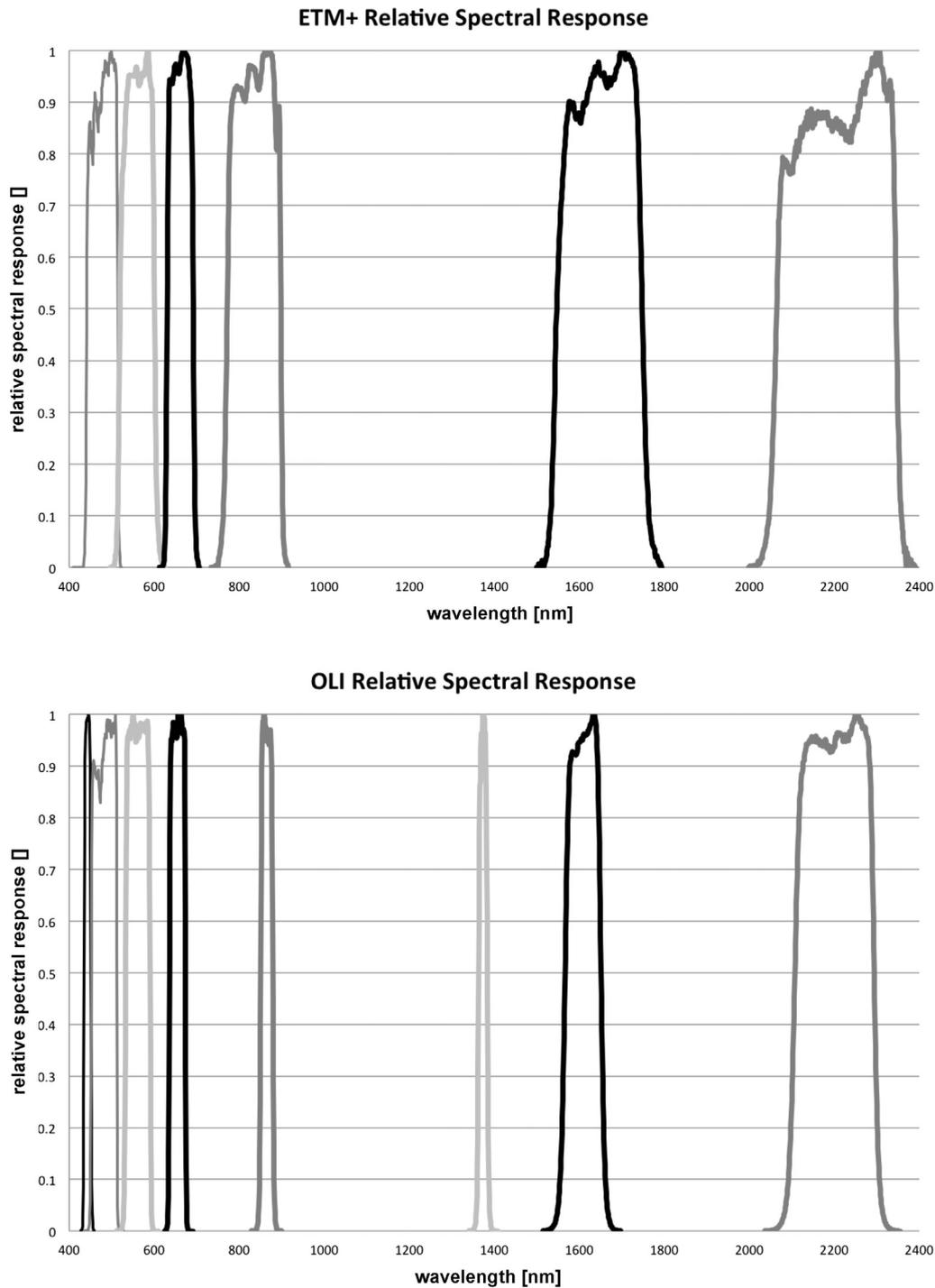


Fig. 1. Landsat ETM+ a (top) and OLI b (bottom) spectral response functions.

This was based on the roughly 20-times improvement in AVIRIS SNR due to a number of instrument improvements from 1987 to 1998. However, no quantitative values for the improvements were stated. Green (2000), using AVIRIS data showed that the number of Minimum Noise Fraction (MNF) images with eigen values above a meaningful information threshold was directly related to SNR. However, the implication of this for specific applications was not demonstrated. Swayze, Clark, Goetz, Chrien, and Gorelick (2003) using simulated AVIRIS spectra showed, using the Tetracorder algorithm, that the classification accuracy of individual spectra improved with SNR. The relationship of the improvement to SNR being a function of the absorption spectrum of the mineral. In general for Landsat class instruments (medium

resolution–multispectral) there is only limited data relating SNR to algorithm performance across any range of applications. The high SNR achieved by OLI provided a potential source of data to support an investigation of the role SNR plays in algorithm performance.

In prelaunch testing the OLI instrument dramatically exceeded its specified SNR and on orbit testing has verified that these high SNR values are being maintained in operational use (Knight & Kvaran, 2014). SNR or radiometric resolution is almost always an important parameter in instrument design trade studies (Schott, 2007, Chapters 6 & 13). Generally speaking, improved SNR comes at the cost of poorer spatial resolution, lower spectral resolution, less coverage, larger optics (size and weight) and higher data rates. As we look to the next

generation of Landsat class sensors, the value of SNR will need to be traded against many other competing parameters to arrive at acceptable cost/performance points. In an effort to better characterize the value of SNR for medium resolution (10–100 m) data, from an image data utility standpoint, the authors, as a subset of the Landsat Science Team (LST), undertook the studies reported here.

These studies took Landsat OLI data as a high quality, high SNR, reference point and then generated degraded image data to provide a number of sample points where the only differences were in SNR values. Recognizing that the role of SNR relative to image utility is likely to vary with application/algorithm, three distinct applications were chosen where existing algorithms were in use studying OLI data. The applications include: water constituent retrieval, land cover classification, and assessment of forest condition as characterized by Leaf Area Index (LAI). These applications run a gamut from water resource assessment, which we expect to be sensitive to SNR, because the target is so dark, to the widely used land cover classification algorithms, which have been used since the early Landsat MSS days (1972) with 6-bit data (Swain & Davis, 1978). While not comprehensive in scope, we believe the three rather disparate algorithms span the applications range that characterizes much of the current use of Landsat data.

2. Data generation

In order to study the impact of SNR on algorithm performance, it was necessary to select a reasonable range of SNR values. In addition, since SNR is a function of illumination levels and instrument response, we expect SNR to vary with wavelength (Fig. 2). We therefore choose to bracket the study with estimated SNR data characteristic of two sensors. OLI represents the current state of the art of a push-broom technology instrument built with relatively large size, mass and cost budget (Roy et al., 2014). In addition, OLI as built significantly outperformed on its SNR requirements (Knight & Kvaran, 2014). Therefore, OLI was chosen as the high end of our SNR study. This was also attractive since the OLI data archive provides us with real data to use as the starting point for subsequent studies of data with degraded SNR. For the lower SNR boundary we chose the Landsat 7 ETM+ SNR values. Since the ETM+ uses a whisk broom scanner and a 20 year old design we felt it represents a lower end for future SNR values.

Fig. 2 shows a family of SNR curves used for this study for the typical radiance (L_{typ}) levels specified for the OLI instrument. Morfitt et al. (2015). The top curve is used as the OLI reference (or OLI_{100}). It is based on the on orbit assessment of OLI's performance (Morfitt et al. (2015). The lowest curve reflects the ETM+ performance (low gain) at approximately 12% of OLI_{100} . The next lowest level represents the required SNR that was set in the OLI contract document (i.e. a level that any vendor was expected to meet (Irons & Masek, 2006)) at about

37% of OLI_{100} . Two additional SNR levels between the requirement and the OLI_{100} were included at 59% and 86% of OLI_{100} to provide additional design points. To simplify a naming convention, the spectral average of the percentage of the value of each curve relative to the maximum curve (OLI_{100}) was calculated over the blue to SWIR range (6 bands) and is used to describe the aggregate SNR performance. Note this is just a naming convention and visible through SWIR bands were used as required for each algorithm. As seen in Fig. 2 this provides us with approximately an order of magnitude change in SNR over the various noise level “designs” to be studied.

Neither the noise level nor the SNR are constants within an image. Rather as shown in Fig. 3 the SNR is a function of the scene radiance level (i.e. the radiance of each pixel). For OLI the noise level was characterized prelaunch for several radiance levels yielding the top curve in Fig. 3 Morfitt et al. (2015). The remaining curves were generated by calculating the ratio of the SNR at L_{typ} for each “design” to the SNR at L_{typ} for OLI. This ratio was used to approximate the radiance dependent SNR at the rest of the OLI test levels to produce Fig. 3. This process was repeated for each band. At this point we can estimate the SNR expected for each radiance value in each band of each design to be studied. Note; OLI band 1 was not available on ETM+ and it is not used in the land cover, LAI, or Landsat 7 water retrieval algorithms described below.

In order to utilize these SNR values a method to generate signal dependent noise and appropriately introduce it to modeled or observed data is required. This can be accomplished in a straightforward fashion for modeled data, which is noise free. In this case the radiance with noise (L_n) can be expressed in terms of the modeled (noise free) radiance L as:

$$L_n = L + Rnd(0.1)N(L) \tag{1}$$

where $Rnd(0.1)$ is a Gaussian random number generator with zero mean and standard deviation of 1, $N(L)$ is found by calculating the SNR from the signal dependent curves shown in Fig. 3 for the radiance level L and the design of interest, where we recognize that.

$$N(L) = L/SNR(L) \tag{2}$$

When dealing with actual image data we need to adjust for the fact that the data will already have some signal dependent noise, which for OLI can be expressed as $N_{OLI}(L)$. $N_{OLI}(L)$ is obtained using the OLI_{100} curves for each band as shown in Fig. 3 and Eq. (2). The desired noise level in the output, noise degraded, image $N_i(L)$ can also be found using the appropriate design curve from Fig. 3 for radiance L and Eq. 2. The noise level ($N_d(L)$) that needs to be added to the OLI data to degrade

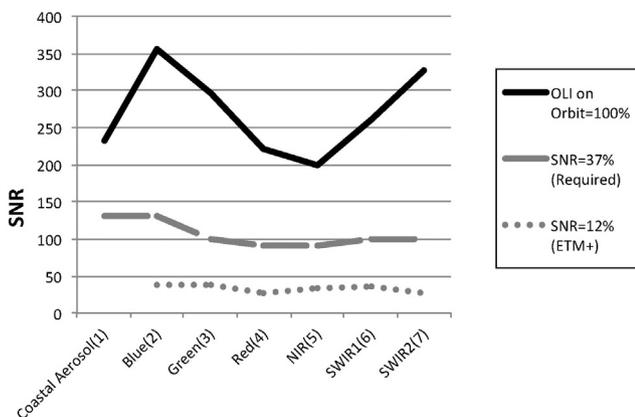


Fig. 2. Plot of wavelength (Band #) vs. SNR for the various designs at typical radiance (L_{typ}).

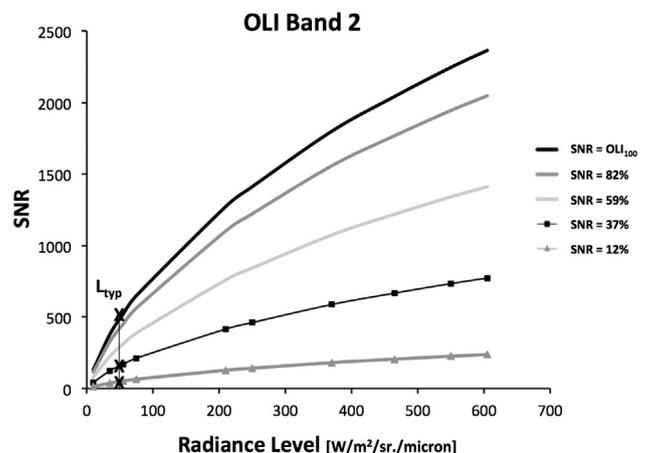


Fig. 3. Plot of radiance level vs. SNR for the design cases used in this study (Band 2).

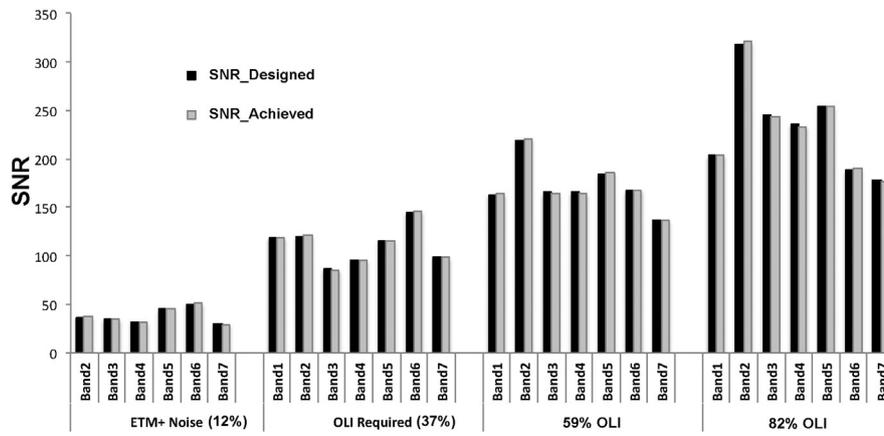


Fig. 4. Comparison of noise levels in noise degraded images to the design values to verify proper implementation of the algorithms. (Note these are for a specific dark target not L_{yp}).

the SNR to the desired noise level $N_i(L)$ can be expressed as:

$$N_d(L) = \left[(N_i(L))^2 - (N_{OLI}(L))^2 \right] \tag{3}$$

where we recognize that

$$N_i(L) = \left[(N_d(L))^2 + (N_{OLI}(L))^2 \right]^{1/2} \tag{4}$$

and take advantage of the fact that noise adds in quadrature (i.e. the noise squared is the sum of the squared values of each independent noise source). Thus, the radiance for each pixel of the degraded image L_i can be expressed in terms of the radiance of the input OLI pixel (L_{OLI}) according to.

$$L_d = Rnd(0.1)N_d(L) + L_{OLI} \tag{5}$$

In practice, the curves associated with Fig. 3 are produced for each band and piece wise linearly interpolated to generate the appropriate SNR values to use with Eq. 2 to calculate the noise level. The mean signal dependent noise levels are calculated using Eq. 3 and the sampled noise for a given pixel obtained and added to the original OLI image pixel according to Eq. 5.

Fig. 4 shows the results of a study to verify that the noise addition was being performed correctly (i.e. that Eqs. 2, 3 & 4 were implemented correctly). This study involved injecting noise into a uniform region of an image and then comparing the observed SNR to the desired SNR. Fig. 5 shows a natural color representation of a simulated scene consisting of a water surface to which noise levels associated with the various designs have been added. This provides a visual representation of noise levels from the nearly indiscernible levels for OLI 100 to the obvious noise characteristics in the ETM + data (12%).

With the tools described in this section the various noise models (designs) introduced in Fig. 2 can be used to generate noise to be added to fully synthetic data (as in the water quality studies in Section 3) or to calculate the appropriate noise to add to actual OLI

data to achieve the desired noise levels. The impact of noise on algorithm performance in different applications is discussed in the next three sections.

3. Water constituent retrieval

Water is a dark target across all bands. As a result, detailed multi-parameter analysis of water constituents has been a difficult task using Landsat 5 and 7. Modeling results of expected OLI performance showed promise due largely to the improved SNR (Gerace, Schott, & Nevins, 2013). In this study, we attempt to show how the performance of a particular water constituent retrieval algorithm is specifically related to SNR.

The algorithm used is designed to simultaneously estimate the concentrations of chlorophyll [CHL] and suspended materials [SM], and the absorption of colored dissolved organic material [CDOM]. These are the three primary coloring agents in water and their values and change in value across space and time can tell us a great deal about the condition of a water body. The algorithm used generates a three-dimensional look up table (LUT) using the Hydrolight in water radiative transfer model (Mobley, 1998). The model's inputs are the inherent optical properties of water and the three constituents, as well as, illumination, viewing conditions and the roughness of the water surface (as defined by wind speed). The output is a spectral reflectance curve of the water surface for each triplet of coloring agents modeled ([CHL], [SM], [CDOM]) as seen in Fig. 6. The spectral reflectance curves are then convolved with the relative spectral response (RSR) functions (Fig. 1a) for each band of the sensor of interest (OLI in this case) to generate sampled reflectance spectra (Note only the VNIR bands were used for water which has near zero reflectance in the SWIR). Given an image derived reflectance spectrum, the LUT can be searched using a nonlinear least squares optimization to find the values of the water constituents that would generate a best match. The match uses the root mean square error (RMSE) between the curves as the function to minimize (see Fig. 6). For this portion of the study only modeled data without an atmosphere

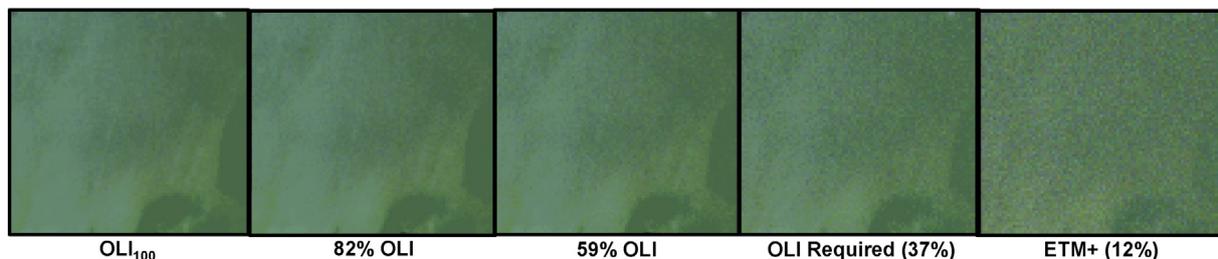


Fig. 5. Simulated natural color image of a water surface to which various noise levels have been added.

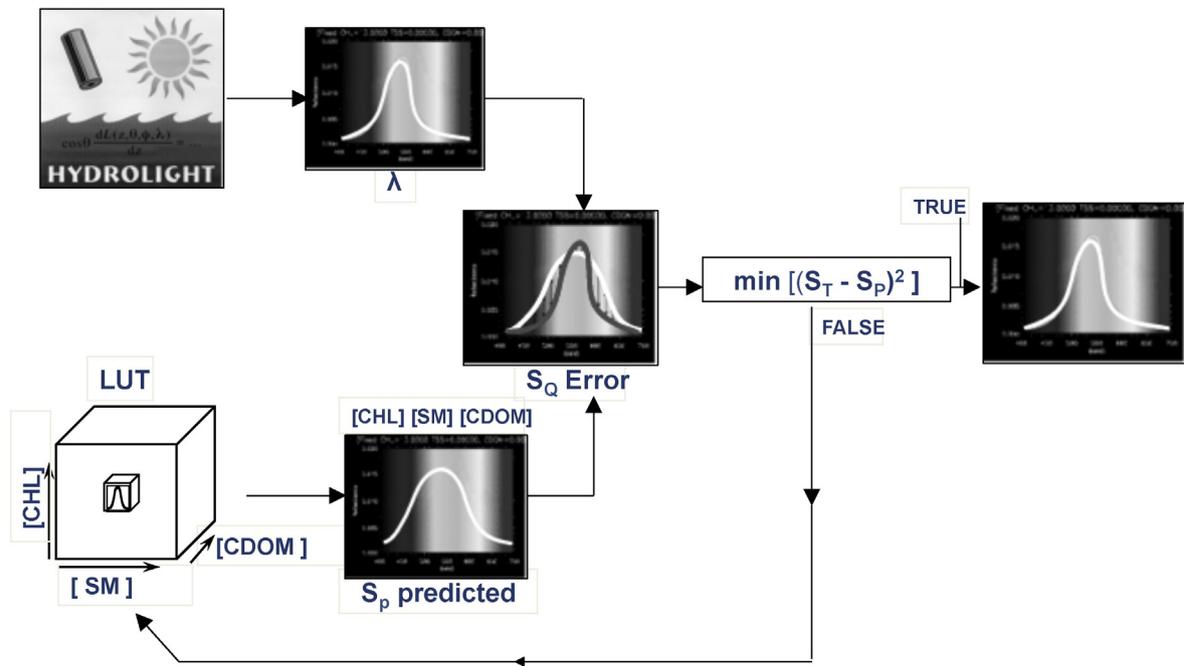


Fig. 6. Illustration of the water quality constituent retrieval algorithm.

were used which yields results in surface leaving reflectance. In this way, all of the performance variation can be attributed to SNR.

A random spectrum generated by Hydrolight with appropriate noise added to it is compared to the look up table (LUT) and the concentration values from the LUT for the best match are assigned to that sample.

The data for this study consisted of reflectance curves that were produced by Hydrolight for triplets of [CHL], [SM], [CDOM] concentration/absorption values which were chosen by a random number generator. The reflectance curves were then convolved with the appropriate RSR, had noise added as described in Section 2 and were then quantized to 12 bits using a quantizer designed to span the OLI signal range. These resampled spectra were treated as simulated image data and fed into the constituent retrieval algorithm, which returned a triplet of concentration/absorption values for each sample spectrum.

The error metric for this process is simply the difference between the constituent values input to Hydrolight that were used to generate the simulated image reflectance curve and the constituent values retrieved by the algorithm. For each constituent, the RMSE of all the simulated

image pixels is calculated and this RMSE value is then expressed as the percentage of the range of the constituent values included in the study. Separate errors are computed for each constituent. The results are plotted in Fig. 7 vs. the average SNR relative to the OLI₁₀₀ SNR (i.e. the designs being tested). Note that the ETM+ design (SNR12%) did not include the coastal aerosol band in the analysis. Also shown in Fig. 7 is a set of designs, which includes all the OLI bands plus a red edge band at 708 nm (a MERIS band was used), which is designed to improve chlorophyll retrieval (Rast, Bezy, & Bruzzi, 1999).

If we first consider just the designs using the Landsat bands, we see that the error in constituent retrievals decreases as SNR increases. In addition, the rate of improvement tends to decrease as SNR increases suggesting that we may reach a point of diminishing returns. Recognizing that these results represent a best-case scenario, with perfect atmospheric compensation, a reasonable goal would be to set desired errors at 5 to 10% of the range. These results suggest that the 10% value could be achieved for [SM] and [CDOM] with the OLI required design (37% in Fig. 7) but that a considerably higher SNR (82% of OLI₁₀₀) would be

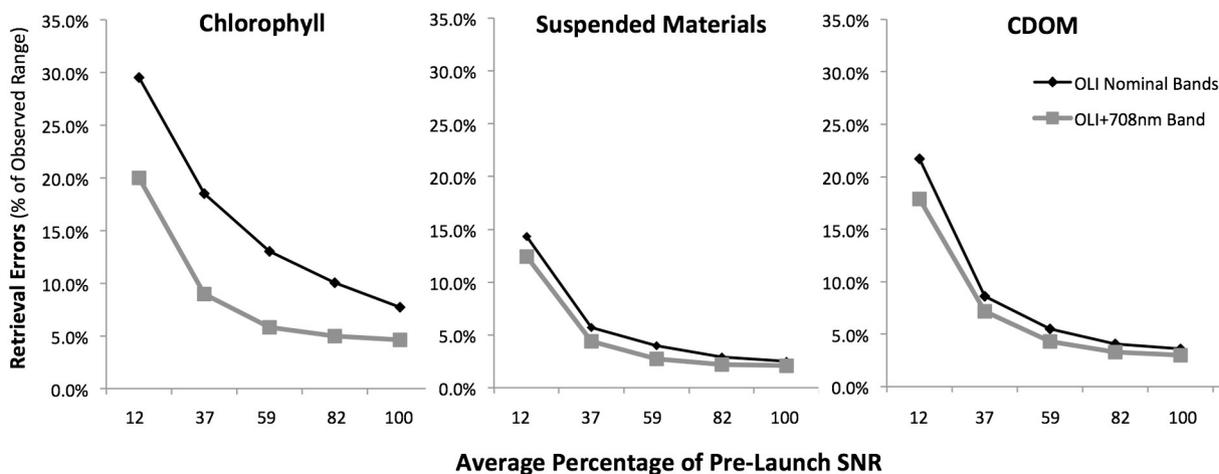


Fig. 7. Plot of constituent retrieval error vs. design for each constituent.

Table 1
Land cover classes used in the different classification tests.

	New Orleans	Boston	Rockwood
1	Brown Wetland	Water	Water
2	Green Wetland	Wetland	Wetland
3	Forest	Forest	Forest
4	Barren	Herbaceous	Herbaceous
5	Herbaceous	Barren	Barren
6	Water	Low Density Built	Low Density Built
7	Low Density Built	High Density Built	High Density Built
8	High Density Built	Cropland	Cropland
9	Cropland	Other	
10	Mined Field		

required for [CHL]. If we require the error to be less than 5% of the range in constituent values, we would need to achieve the 58% of OLI₁₀₀ SNR design for [SM] and [CDOM] and could not reach the 5% goal for [CHL] even with the OLI₁₀₀ SNR.

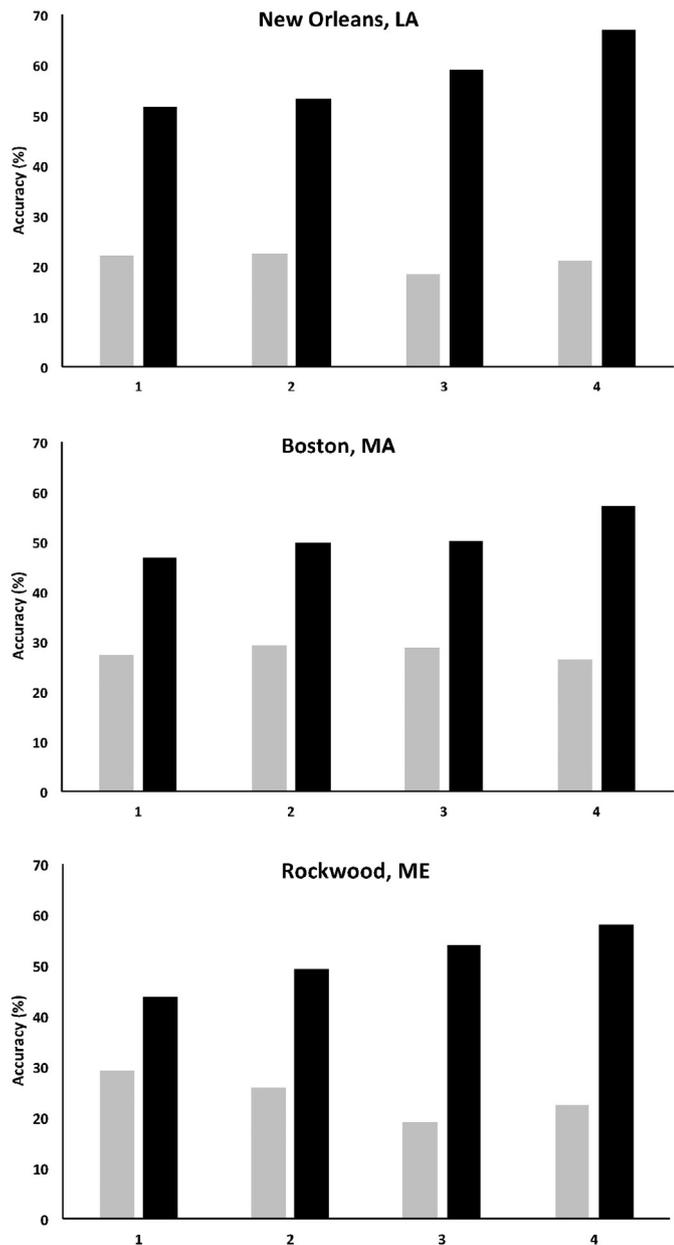


Fig. 8. Results for 12 pairwise comparisons of image classification results (3 locations—top is New Orleans, LA, middle is Boston, MA, and bottom is Rockland, ME × 4 levels of image SNR (1 is 37%, 2 is 59%, 3 is 82% and 4 is 100%).

However, the second set of designs in Fig. 7 shows that in some cases there may be alternate methods to address a problem. In this case, we repeated the SNR study with one additional band included at 708 nm that is often included in instruments designed for water quality studies (which has not been a primary Landsat objective). We used the RSR for the 708 nm MERIS band and interpolated SNRs from adjacent bands to estimate what the performance might be for our various designs. The results clearly show that such a band would greatly improve the performance for [CHL] with the goal of errors less than 10% of the range in constituent values obtainable with SNRs associated with the OLI design specification (37% in Fig. 7) and the goal of 5% of the range being achievable with OLI₁₀₀ SNRs (100% in Fig. 7).

4. Land cover: classification

Image classification accuracies vary with the number of classes and their spectral separability as well as image quality. Many studies have compared classification algorithms (Carpenter, Gjaja, Gopal, & Woodcock, 1997; Rogan et al., 2008), and many studies evaluate the contributions of various kinds of data as input to image classification (Carrão et al., 2008; Li and Chen, 2005; Corcoran et al., 2012; Zhu, Woodcock, Rogan, & Kellndorfer, 2012), but it is more difficult to find studies that compare image data from the same sensor but with different SNRs. The basic idea is that one cause of error in image classification is noise, and as the SNR improves, errors in classification results should decline.

The three images used for this experiment are located near New Orleans, LA (path 22, row 39), Boston, MA (path 12, row 31) and Rockwood, ME (path 12, row 28). Subsets (3000 by 3000 pixels) from the original scenes with 0% cloud cover were used. Training data for each of the three regions were collected for each of the classes listed in Table 1. The classes used in this test are a relatively simple set and typical of those used in land cover classification. Given the wide number of uses of land cover information, there are not standardized sets of classes or definitions of individual classes. For this test, these classes are intended to be examples of the kinds of classes many users typically desire.

Images were created for the different SNRs as described in Section 2. The random forest classification algorithm was used for this experiment (Brieman, 2001). The same training data were given to the classifier such that the only factor that differed between trials was the SNR of the input data. To evaluate the results, we made pairwise comparisons between the classified maps generated from the image simulated to have the SNR of ETM+ (SNR12%) and all other SNR levels. In this pairwise manner, we first identify the pixels where the classifications differ, as it is only where they differ that we can evaluate the effect of image quality on classification accuracy. If the accuracies of the pixels

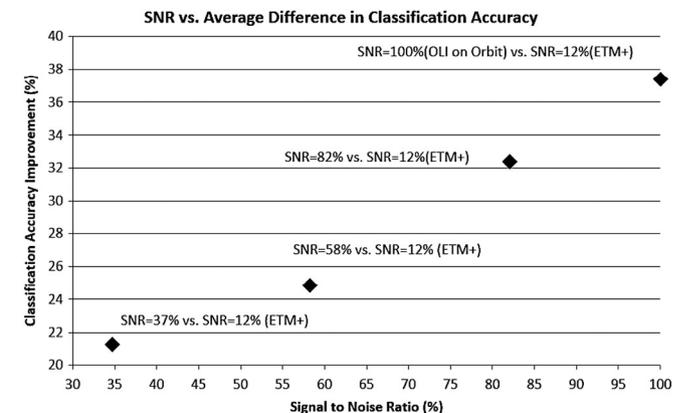


Fig. 9. Average difference in classification accuracy results for the three tests sites as a function of SNR.

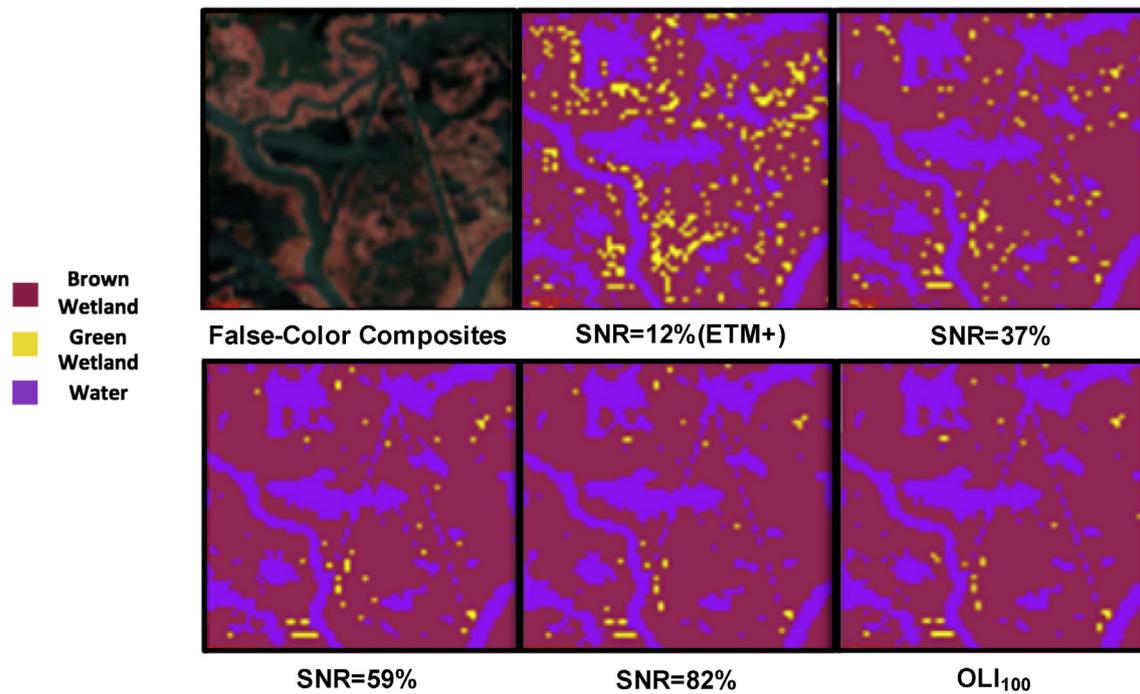


Fig. 10. Classification results for a small area for all the different SNRs. Note that the amount of “speckle” decreases as the SNR increases.

where they differ are similar between trials, then improved SNR is not improving classification results. However, if the accuracy for the trials with higher SNR is higher than for the ETM + equivalent SNR, then we can conclude that improved SNR improves image classification.

About 85% of the pixels in each comparison are classified as the same in all trials and these pixels are not used in the comparison. A random sample (205 pixels for each comparison) of the pixels where the classifications differed was selected. For each of these pixels the true land cover class was identified through careful examination of very high resolution imagery in GoogleEarth. These “ground truth” pixels were then compared with the classification results to determine whether the classification from the ETM + simulated image (SNR12%) or the higher SNR alternative was correct, or neither were correct.

The results are plotted in Fig. 8. In total there are twelve pairwise comparisons presented, each based on a different random sample, as there is a different set of pixels where the classifications differ for each pairwise comparison. In each pairwise comparison, the accuracies of the ETM + simulated (SNR12%) and the higher SNR alternatives are plotted next to each other. The left bar in each comparison is the percentage of the pixels where the answer differs and the ETM + answer

is correct, and the right bar is the percentage of those same pixels that differ where the alternative SNR level is correct. In these graphs, it is important to remember that the combined accuracy (the sum of the two accuracies) can only be as high as 100% and if there were no effect related to the SNR of the input images, then the accuracies would be roughly equal. Where the sum of the two accuracies is less than 100%, some pixels were wrong in both classifications. In all twelve cases (3 sites times 4 different SNR levels), the images with improved SNRs have higher classification accuracies among the pixels where the classifications differed. Also, as the SNR increases through the four alternatives at each of the 3 sites, the accuracy of the classifications also increases. However, note that the inverse is not always true, meaning the accuracy for the ETM + results does not always decline as the SNR of the alternative images increases. In general, the ETM + accuracies decline, but not in every case. This result illustrates the stochastic nature of image classification and the reason we included three different sites in the experiment to ensure that our findings would be representative.

The first conclusion is that image quality as measured by SNRs influences classification accuracy. This is seen in Fig. 9, which shows the average improvement in classification accuracy between the scenes with

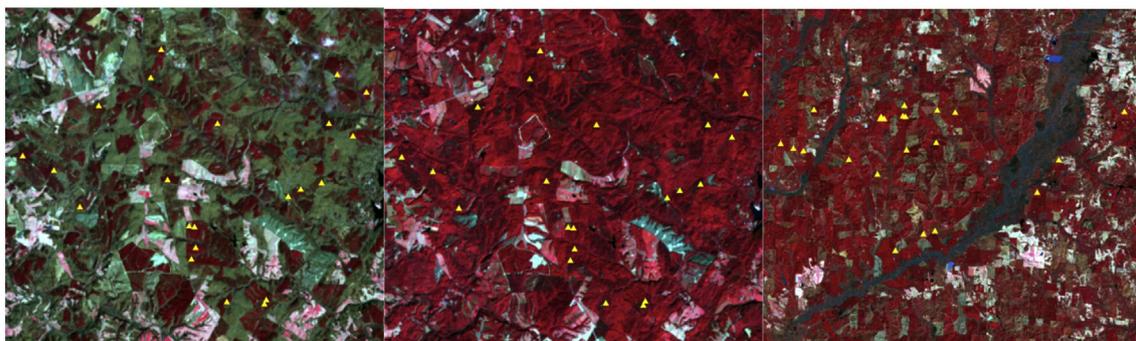


Fig. 11. Ground LAI plot locations shown as yellow triangles over Landsat 8 imagery displayed with the color-infrared band combination 5, 4, 3 as RGB. a) Path 16/Row 34 on 3/28/13, b) Path 16 Row 34 on 10/5/13 and c) Path 21 Row 37 on 2/13/14.

Table 2
Regression results between ground measured LAI and Landsat 8 OLI SNR image based simple ratio vegetation index. SNR increases from left to right from ETM+ to Landsat 8 OLI 100%.

Path/row time period	Landsat date	ETM + 12% R-sq (RMSE)	OLI required 37% R-sq (RMSE)	OLI 59% R-sq (RMSE)	OLI 82% R-sq (RMSE)	OLI 100% R-sq (RMSE)	Number of plots
16/34 2013 Min	3/28/13	84.1 (0.260)	84.9 (0.253)	84.9 (0.254)	84.8 (0.254)	84.7 (0.255)	19
16/34 2013 Peak	10/5/13	75.6 (0.660)	88.7 (0.449)	91.0 (0.401)	92.1 (0.376)	93.0 (0.355)	8
21/37 2014 Min	2/13/14	72.0 (0.571)	70.4 (0.587)	69.8 (0.594)	69.4 (0.597)	69.1 (0.601)	21

ETM+ noise levels, and the design noise levels (of the pixels that change) versus the mean SNR. This plot shows a consistent and steady improvement in accuracy with SNR suggesting that even better results might be possible if higher SNR data were available.

An additional observation was that there appeared to be less speckle effect in the classification results as image SNR increased (Fig. 10). Image classification results generally suffer from high uncertainties at the scale of individual pixels and it appears that the improved SNR of OLI images may reduce this undesirable effect.

5. Leaf area index prediction

Ground estimates of leaf area index (LAI) were collected using the LI-COR LAI-2200C Plant Canopy Analyzer at locations in loblolly pine plantations at two study sites near Appomattox, VA and Tuscaloosa, AL (see Fig. 11) on thirty by thirty meter plots. These estimates were used as the dependent variable in regression models with the simple ratio vegetation index (SR) calculated from Landsat images with varying degrees of SNR. SR values were extracted for each pixel containing a LAI plot center according to:

$$SR = NIR/RED \quad (6)$$

Where NIR and RED are the near infrared and red image radiance values respectively.

Regression model fits improved with the increase in SNR from ETM+ imagery to OLI imagery for the two 16/34 images (Table 2), but not for the 21/37 image. SNR appears to be more important at peak LAI than at minimum LAI given the larger increase in model accuracy for the 10/5/2013 image as SNR increased. The variability within ground based LAI estimates may play a bigger role than SNR in this application.

For the October data (peak LAI) the increase in correlation (R^2 value) with SNR is over 15 points (Fig. 12). For this case where the vegetation index and ground based LAI measurements are highly correlated, the increases in correlation with SNR is most dramatic with the step from ETM+ (12%) to OLI specified (37%) levels, with continuing but smaller improvements as SNR continues to increase.

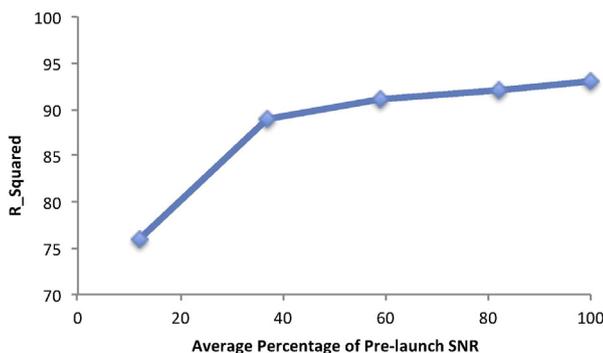


Fig. 12. Plot of correlation coefficient (R^2) vs. SNR design for the near maximum LAI study date (10/5/13).

6. Conclusions & recommendation

As illustrated in Figs. 7, 9 and 11, increases in SNR, from the ETM+ to the OLI levels, show improvements at each step up in SNR for all the applications studied (Note this was not the case for the near minimum LAI images). More importantly, the increases were not only significant (up to 37% increase in classification accuracy of the changed pixels and change in R^2 from 76 to 93% for the correlation with LAI (Figs. 9 & 11) but in some cases (e.g. water quality) have made sufficient improvement to make an otherwise unacceptable approach viable (e.g. chlorophyll errors reduced from 30% to less than 10% of the range [Fig. 7]). For the small sample studied here, it appears the OLI as built SNR may be approaching the useful limit of SNR as we see the improvements in performance beginning to taper off (at least for these algorithms and applications). On the other hand, for the land cover classification case (see Fig. 9), there still appears to be room for improvement at higher SNR (i.e. the step up in classification accuracy is not substantially diminishing with a step up in SNR even at the OLI₁₀₀ level).

In conclusion, sizable improvements in algorithm performance tied directly to SNR improvements were shown for the three different algorithms in three different application areas studied here. On the other hand, for the LAI study on near minimum LAI images little or no improvement with SNR was shown. This suggests that, in many applications where models of the phenomenology are well captured by the algorithms such that instrument noise is a significant source of residual variation in the model (i.e. algorithm performance), SNR increase will improve performance. However, in cases where algorithm performance is limited by variations in phenomenology not captured in the algorithm then SNR improvements will not be significant. Response to SNR variation is not a way in which we have traditionally characterized or evaluated our algorithms since SNR was not a variable to which we have had ready access. With the high SNR available from Landsat 8 OLI the methodology presented here can be readily applied to a wide range of algorithms/applications to evaluate the response to SNR changes. Our limited study suggests, that there are likely to be numerous algorithms that improve in performance as we increase SNR, particularly by nearly an order of magnitude from ETM+ to OLI.

We recommend that similar studies be conducted to provide a published data base to enable instrumentation engineers to adequately factor in SNR performance to trade space studies for the next generation of instruments. There is no question that there will be a push to improve spatial resolution and spectral resolution (band width and number of bands). These improvements often come at the cost of SNR (Schott, 2007, Chapter 13). As these studies clearly point out, SNR should be an important part of the design trade space. However, it can only be effectively traded if we know the impact of SNR on a range of applications/algorithms. In closing we note that, the dramatic improvements in algorithm performance with SNR shown here are available now to the entire remote sensing community thanks to the outstanding SNR levels that Landsat 8's OLI has achieved.

Acknowledgements

The authors would like to acknowledge the United States Geological Survey (USGS) and the National Aeronautics and Space Administration (NASA) for supporting this effort with grant #G12PC00065 and grant #NNX14AP40G, respectively.

References

- Carpenter, G. A., Gajda, M. N., Gopal, S., & Woodcock, C. E. (1997). ART neural networks for remote sensing: Vegetation classification from Landsat TM and terrain data. *IEEE Transactions on Geoscience and Remote Sensing*, 35(2), 308–325.
- Carrão, H., Gonçalves, P., & Caetano, M. (2008). Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sensing of Environment*, 112(3), 986–997.
- Chen, C.H. (2006). Signal and image processing for remote sensing. CRC Press.
- Corcoran, J., Knight, J., Brisco, B., Kaya, S., Cull, A., & Murnaghan, K. (2012). The integration of optical, topographic, and radar data for wetland mapping in northern Minnesota. *Canadian Journal of Remote Sensing*, 37(5), 564–582.
- Fiete, R. D., & Tantaló, T. (2001). Comparison of SNR image quality metrics for remote sensing systems. *Optical Engineering*, 40(4), 574–585.
- Gerace, A. D., Schott, J. R., & Nevins, R. (2013). Increased potential to monitor water quality in the near-shore environment with Landsat's next-generation satellite. *Journal of Applied Remote Sensing*, 7(1), 073558.
- Green, R. O. (2000). Exploration of the relationship between information content and signal-to-noise ratio and spatial resolution in AVIRIS spectral data. *Proceedings, 9th JPL airborne Earth science workshop*.
- Irons, J. R., & Masek, J. G. (2006). Requirements for a Landsat data continuity mission. *PE&RS*, 1102–1108.
- Irons, J. R., Dwyer, J. L., & Barsi, J. A. (2012). The next Landsat satellite: The Landsat data continuity mission. *Remote Sensing of Environment*, 122, 11–21.
- Knight, E. J., & Kvaran, G. (2014). Landsat-8 operational land imager design, characterization and performance. *Remote Sensing*, 6(11), 10286–10305.
- Kruse, F. A. (2000). The effects of spatial resolution, spectral resolution, and SNR on geologic mapping using hyperspectral data, northern Grapevine Mountains, Nevada. *Proceedings, 9th JPL airborne Earth science workshop*.
- Leachtenauer, J. C., Malila, W., Irvine, J., Colburn, L., & Salvaggio, N. (1997). General image-quality equation: GIQE. *Applied Optics*, 36(32), 8322.
- Li, J., & Chen, W. (2005). A rule-based method for mapping Canada's wetlands using optical, radar and DEM data. *International Journal of Remote Sensing*, 26(22), 5051–5069.
- Mobley, Curtis D (1998). Hydrolight 4.0 users guide. Sequoia Scientific Inc., Mercer Island, WA.
- Morfitt, R., Barsi, J., Levy, R., Markham, B., Micijevic, E., Ong, L., ... Vanderwerff, K. (2015). Landsat-8 operational land imager (OLI) radiometric performance on-orbit. *Remote Sensing*, 7(2), 2208–2237.
- Rast, M., Bezy, J. L., & Bruzzi, S. (1999). The ESA medium resolution imaging spectrometer MERIS a review of the instrument and its mission. *International Journal of Remote Sensing*, 20(9), 1681–1702.
- Rogan, J., Franklin, J., Stow, D., Miller, J., Woodcock, C. E., & Roberts, D. (2008). Mapping land cover modifications over large areas: A comparison of machine learning algorithms. *Remote Sensing of Environment*, 112(5), 2272–2283.
- Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., Scambos, T. A., Schaaf, C. B., et al. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, 145, 154–172.
- Schott, J. R. (2007). *Remote sensing, the image chain approach* (2nd ed.). Oxford University Press.
- Seidel, F., Schlapfer, D., Nieke, J., & Itten, K. I. (2008). Sensors performance requirements for the retrieval of atmospheric aerosols by airborne optical remote sensing. *Sensors, Switzerland*, 8(3), 1901–1914.
- Swain, P. H., & Davis, S. M. (1978). *Remote sensing: The quantitative approach*. New York, NY: McGraw-Hill.
- Swayze, G. A., Clark, R. N., Goetz, A., F. H., Chrien, T. G., & Gorelick, N. S. (2003). Effects of spectrometer band pass, sampling, and signal-to-noise ratio on spectrometer identification using the Tetracorder algorithm. *Journal of Geophysical Research*, 108, E9.
- Williams, D.L., Irons, J.R., Markham, B.L., Nelson, R.F., Toll, D.L., Latty, R.S., Stauffer, M.L., (1984). Geoscience and remote sensing, IEEE transactions, Vol: GE-22, Issue: 3, 294–302.
- Zhu, Z., Woodcock, C. E., Rogan, J., & Kelndorfer, J. (2012). Assessment of the spectral, polarimetric, temporal and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data. *Remote Sensing of Environment*, 117(15), 72–82.