Article

Toward Long-Term Aquatic Science Products from Heritage Landsat Missions

Nima Pahlevan 1,2,*, Sundarabalan V. Balasubramanian 1,3, Sudipta Sarkar 1,2 and Bryan A. Franz 1

1 NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA; sundarabalan.velaudhaperumalbalasubramanian@nasa.gov (S.V.B.); sudipta.sarkar@nasa.gov (S.S.); bryan.a.franz@nasa.gov (B.A.F.)
2 Science Systems and Applications, Inc., 10210 Greenbelt Road, Suite 600 Lanham, MD 20706, USA
3 Department of Geographical Sciences, University of Maryland, College Park, MD 20740, USA
* Correspondence: nima.pahlevan@nasa.gov; Tel.: +1-301-614-6684

Received: 29 January 2018; Accepted: 15 August 2018; Published: 22 August 2018

Abstract: This paper aims at generating a long-term consistent record of Landsat-derived remote sensing reflectance (\(R_{rs}\)) products, which are central for producing downstream aquatic science products (e.g., concentrations of total suspended solids). The products are derived from Landsat-5 and Landsat-7 observations leading to Landsat-8 era to enable retrospective analyses of inland and nearshore coastal waters. In doing so, the data processing was built into the SeaWiFS Data Analysis System (SeaDAS) followed by vicariously calibrating Landsat-7 and -5 data using reference in situ measurements and near-concurrent ocean color products, respectively. The derived \(R_{rs}\) products are then validated using (a) matchups using the Aerosol Robotic Network (AERONET) data measured by in situ radiometers, i.e., AERONET-OC, and (b) ocean color products at select sites in North America. Following the vicarious calibration adjustments, it is found that the overall biases in \(R_{rs}\) products are significantly reduced. The root-mean-square errors (RMSE), however, indicate noticeable uncertainties due to random and systematic noise. Long-term (since 1984) seasonal \(R_{rs}\) composites over 12 coastal and inland systems are further evaluated to explore the utility of Landsat archive processed via SeaDAS. With all the qualitative and quantitative assessments, it is concluded that with careful algorithm developments, it is possible to discern natural variability in historic water quality conditions using heritage Landsat missions. This requires the changes in \(R_{rs}\) exceed maximum expected uncertainties, i.e., 0.0015 [1/sr], estimated from mean RMSEs associated with the matchups and intercomparison analyses. It is also anticipated that Landsat-5 products will be less susceptible to uncertainties in turbid waters with \(R_{rs}(660) > 0.004 [1/sr]\), which is equivalent of ~1.2% reflectance. Overall, end-users may utilize heritage \(R_{rs}\) products with “fitness-for-purpose” concept in mind, i.e., products could be valuable for one application but may not be viable for another. Further research should be dedicated to enhancing atmospheric correction to account for non-negligible near-infrared reflectance in CDOM-rich and extremely turbid waters.

Keywords: Landsat; coastal/inland waters; atmospheric correction; vicarious calibration; validation; water quality; time-series applications

1. Introduction

The Landsat program is among the longest Earth-observing programs on record, with continuous operations dating from 1972. The primary goal for the Landsat series is to monitor Earth resources from space using best-available technology. Following successful launches and operations of Landsat 1, 2, and 3, which carried analogue cameras [1], Landsat-4 and -5 carried the Thematic Mapper (TM),
which, at the time, had been considered a technological breakthrough, i.e., 30 m spatial sampling, six spectral bands, 8-bit radiometric resolution, and a single thermal channel. A similar technology with some improvements [2] was utilized for the Enhanced Thematic Mapper Plus (ETM+) launched in 1999. The TM and ETM+ were whiskbroom imagers with 16 detectors laid out in the along-track direction [3]; thus, with each mirror scan, ~480 m is swept on the ground. The short dwell-time together with 8-bit radiometric resolution, however, yields low signal-to-noise ratios (SNR) for dark surfaces such as water bodies [4,5]. The two instruments carried stimulation lamps to allow for onboard radiance calibrations [6–8]. In addition to onboard calibration observations, multiple desert sites have historically been used to monitor long-term stability of these missions. For reference, the relative spectral responses (RSR) for TM and ETM+ in the visible and the near infrared (NIR) regions alongside those of the most recent Landsat sensor, i.e., the Operational Land Imager (OLI), are illustrated in Figure 1.

Currently available ocean color satellites are not suitable for long-term monitoring of estuarine environments. First, well-calibrated, dedicated ocean color sensors (e.g., the Sea-viewing Wide Field-of-view Sensor (SeaWiFS)), have only become available since the late 1990s, so longer term historical studies [9] are not possible. Second, ocean color missions are generally equipped with 1 km nadir ground sampling distance [10]; therefore, lack sufficient spatial resolution to capture the spatial variability necessary to observe critical features, such as upwelling, small-scale eddies/blooms, river plumes, tidal effects, and local gradients in inland and nearshore waters. Third, due to steep gradients and high variability in optical properties as well as interfering optical signals from high concentrations of particles, chlorophyll-a (Chl), and colored dissolved organic matter (CDOM) in these regions, a clear delineation of these water quality properties is difficult without careful calibration and validation with in situ data [11]. Moreover, although standard, operational ocean color algorithms are well-suited for ocean color imagers (e.g., SeaWiFS) for relatively clear ocean waters, to examine inland and nearshore coastal waters, the development of algorithms customized to specific aquatic systems of interest is necessary for Landsat-type (broad-band) sensors.

The long record of Landsat observations provides unprecedented synoptic views of Earth resources; a record that has been substantially under-utilized in aquatic studies. Recognizing that the TM and ETM+ sensors (a) were equipped with only three visible bands (Figure 1), (b) lack sufficient SNR [5,12], and (c) suffer from artifacts (e.g., memory effects, striping) [13], they still have been found useful in some aquatic studies. These studies have primarily revolved around quantifying concentrations of total suspended solids (TSS) [14–23]. Attempts have also been made to quantify Chl [24–27] and the absorption by CDOM [28–30].

![Figure 1](image_url)  
**Figure 1.** The relative spectral response (RSR) for the Thematic Mapper (TM), the Enhanced Thematic Mapper Plus (ETM+), and the Operational Land Imager (OLI) with their corresponding nominal center wavelength. The largest difference between TM/ETM+ and OLI is in the near infrared (NIR) channel where OLI band was designed narrow to avoid water-vapor absorption in the atmosphere.
Further applications include analyzing bottom properties and bathymetry [31–35]. Many studies have also processed and examined time-series of images to assess feasibility of using Landsat for monitoring applications [9,27,36–45]. Although the utilization of the Landsat archive has recently increased (since 2008 when the entire archive was made available to public at no cost [46]), the invaluable historic image data have not yet been fully exploited. To date, every retrospective study has followed a different strategy for data processing and/or atmospheric corrections [9,47]. Some studies have not even applied atmospheric corrections and used Level-1 (top-of-atmosphere; TOA) radiance ($L_t$) to evaluate water quality [36]. In general, over-water TM and ETM+ images are expected to contain relatively large uncertainties (compared to those of ocean color missions, like the Moderate Resolution Imaging Spectroradiometer (MODIS) [48–50]); however, with a common processing system and vicarious corrections to Level-1B data, we anticipate that not only can retrospective studies be conducted but also current and future research can be traced back to the past.

Owing to the successful integrations of Landsat-8 [51] and Sentinel-2 [52] processing in the SeaWiFS Data Analysis System (SeaDAS), we intend to expand this capability to Landsat-5 and -7 (to the extent possible within the limitations of instrument radiometric performances). More specifically, we implement a physics-based and extensively validated atmospheric correction method to generate Landsat products for the science community; enabling long-term studies of water quality in inland and nearshore coastal areas. Therefore, we first implement the atmospheric correction method (Section 2.1) through which non-negligible water-leaving radiance ($L_w$) in the near-infrared (NIR) band [53] is accounted for using a slightly different strategy than what is currently built into SeaDAS for Landsat-8 and Sentinel-2 [54]. Then, to reduce existing calibration biases in TOA observations of TM and ETM+, we perform vicarious calibrations (Section 2.2) using reference ocean color data [55,56]. Due to uncertainties in the vicarious calibrations, intercomparisons/cross-calibrations are further conducted (Section 2.3) to minimize residual biases and ensure consistency in Level-2 products, i.e., remote sensing reflectance ($R_{rs}$), with those derived from more recent missions. Sections 3.1 and 3.2 provide qualitative and quantitative assessments of TM- and ETM+-derived $R_{rs}$ products. The quantitative evaluation is carried out using the Aerosol Robotic Network (AERONET) data measured by in situ ocean color radiometers, i.e., AERONET-OC. Here, the remote sensing reflectance is defined as the ratio of $L_w$ to the total downwelling irradiance just above the water surface. Since AERONET-OC sites do not adequately represent aquatic/atmospheric regimes in inland areas, the heritage $R_{rs}$ products are further cross-validated against those derived from ocean color sensors over relatively large inland waters. Further, time-series of $R_{rs}$ products over 12 aquatic systems across the continental United States are analyzed (Section 3.3) to examine the utility of products derived from the three Landsat missions, i.e., TM, ETM+, and OLI. Further details on the limitations of heritage Landsat missions for retrospective water quality studies in inland and nearshore coastal areas are provided in Section 4.

2. Materials and Methods

Although the atmospheric correction methodology [57,58] and its integration [51] have been described in previous studies, we provide a brief description of how $L_w(NIR) \neq 0$ is accounted for in the TM and ETM+ processing (Section 2.1). This section continues with computing vicarious calibration gains for TM and ETM+ data. This is an essential step towards minimizing biases in products over bodies of water [55,56], which has proven to enhance the overall quality of Landsat-8 and Sentinel-2A $R_{rs}$ products [52,59]. The derived gains are further slightly adjusted using cross-calibrations against ocean color products derived from OLI [59] and MODIS [48–50].

2.1. Atmospheric Correction

Similar to the implementations for Landsat-8 and Sentinel-2, per-pixel geometry data, i.e., view zenith/azimuth angles and solar zenith/azimuth angles, were utilized in the processing. The current methodology to correct for $L_{w0}(NIR) \neq 0$ is based on a bio-optical model, which uses $R_{rs}(443)$ data [54].
More specifically, this model utilizes $r_{rs}(443)/r_{rs}(561)$ to estimate the spectral slope ($\eta$) of particulate backscattering ($\theta_{bp}$), which is applied in the retrieval of $R_{\text{rs}}(\text{NIR})$, where $r_{rs}$ is the subsurface remote sensing reflectance [60]. The empirical method for estimating $\eta$ (referred to as $\eta_{443/561}$) is as follows

$$
\eta_{443/561} = 2 \left( 1 - 1.2 \exp \left( -0.9 \frac{r_{rs}(443)}{r_{rs}(561)} \right) \right)
$$

The TM and ETM+ spectral band centers are tabulated in Table 1. Throughout this manuscript, we refer to their visible bands as blue and green bands unless specific bands for each instrument are discussed. Since the 443 nm channel is not available for TM and ETM+, we developed a slightly different strategy. To do so, in situ radiometric data collected over various inland waters [61] were used to examine the correlations between $[r_{rs}(443)/r_{rs}(561)]$ and $[r_{rs}(482)/r_{rs}(561)]$ for OLI. We found a strong relationship between the two ratios (Figure 2 (left)). Hence, the ratio in Equation (1) can potentially be replaced by $[r_{rs}(482)/r_{rs}(561)]$, which results in the equation below

$$
\eta_{482/561} = 2 - 2.2 \times \exp \left( -0.9 \times 0.7554 \left[ \frac{r_{rs}(482)}{r_{rs}(561)} \right]^{1.3994} \right)
$$

Figure 2. The strong relationship (left) between the band ratios of subsurface remote sensing reflectance in the blue and green channels, i.e., $r_{rs}(482)/r_{rs}(561)$ and $r_{rs}(443)/r_{rs}(561)$. The relationship follows the power-law function ($A = 0.7554$, $B = 1.3994$) and provides a high correlation coefficient, i.e., $R^2 = 0.987$. The scatterplot (right) shows the cross-validation of $\eta$ derived using two different band ratios, i.e., $r_{rs}(482)/r_{rs}(561)$ and $r_{rs}(443)/r_{rs}(561)$.

Table 1. The OLI, ETM+ and TM spectral band centers (nm) computed in SeaDAS.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OLI</td>
<td>482</td>
<td>561</td>
<td>655</td>
<td>865</td>
<td>1609</td>
<td>2201</td>
</tr>
<tr>
<td>ETM+</td>
<td>478</td>
<td>560</td>
<td>661</td>
<td>835</td>
<td>1648</td>
<td>2205</td>
</tr>
<tr>
<td>TM</td>
<td>485</td>
<td>569</td>
<td>660</td>
<td>840</td>
<td>1676</td>
<td>2223</td>
</tr>
</tbody>
</table>

The performance of the newly estimated $\eta_{482/561}$ was tested against $\eta_{443/561}$ in Figure 2 (right). Similar relationships were then built for TM and ETM+. This assessment further suggests that Equation (2) can be utilized to estimate $R_{\text{rs}}(\text{NIR})$ in an iterative fashion for processing TM and ETM+ as implemented for OLI.

2.2. Vicarious Calibration

In this exercise, the TM and ETM+ data were vicariously calibrated using different reference data to reduce systematic biases in the three visible bands (Table 1). Theoretically, all the Earth-observing
Remote Sens. 2018, 10, 1337

satellites passing over one of the existing vicarious calibration systems, such as the Marine Optical Buoy (MOBY) [62] or BOUSSOLE [63], can be vicariously calibrated. The ETM+ data, with regular image acquisitions over both sites, can thus be vicariously calibrated. However, due to the limitations in the downlink capacity of Landsat-5, TM images were acquired neither over the MOBY nor over the BOUSSOLE site, which became operational in 2003. Hence, we employed MODIS data to vicariously calibrate, i.e., cross-calibrate, TM data.

2.2.1. Landsat-7 (ETM+)

The time-series of all USGS ETM+ (Collection-1) images over the MOBY and BOUSSOLE sites were obtained and processed using SeaDAS to compute the vicarious gains [56]. The calibration is performed in two steps. In Step I, the NIR and the Short-wave Infrared (SWIR) bands are calibrated. This is carried out by selecting scenes with low aerosol loading, i.e., aerosol optical thicknesses in the NIR (AOT(835)) < 0.15. The goal is to minimize propagations of systematic errors (e.g., calibration biases) in aerosol path radiance estimates for these channels when generating $R_{rs}$ products [64].

With the calibration gains derived for these bands, the vicarious gains for the three visible bands were derived in Step II. Overall, 58 cloud-free images over the MOBY ($n = 30$) and BOUSSOLE ($n = 28$) sites were utilized in the analyses. While MOBY hyperspectral $L_{w}$ resampled to ETM+ multispectral bands were used, the BOUSSOLE narrow-band, multispectral $L_{w}$ were converted to ETM+ multispectral bands using a trained deep neural network [61]. The vicarious gains were initially produced using three different band combinations to account for aerosol contributions. The band combinations for ETM+ were 835–1648, 835–2205, and 1648–2205 (Table 1).

Similar to the results in Reference [59], preliminary analyses of the $R_{rs}$ retrievals at AERONET-OC [65] sites indicated that the most robust matchups were produced with 835 and 1648 nm bands (Table 1); therefore, we only present results derived with this band pair. The median gains within $5 \times 5$-element windows centered over the exact coordinates are utilized to represent spectral gains for each ETM+ image. The median statistics were chosen to suppress image artifacts, including striping. The median statistics were found to be nearly identical ($\sim 10^{-4}$) to the average computed within the interquartile range [66]. Furthermore, a sensitivity analysis on the window size, i.e., deriving gains for $3 \times 3$, $7 \times 7$, $9 \times 9$, and $11 \times 11$-element kernels, indicated minimal differences, i.e., $<10^{-3}$, in the derived gains as a function of the window size. The time-series of vicarious gains for the visible bands derived using the 835–1648 combination are shown in Figure 3. The median gains and one standard deviation (in round brackets) are reported in Table 2. The MOBY and BOUSSOLE median gains in the blue, green, and red channels were 1.033, 1.011, 1.021 and 1.027, 1.009, 1.033, respectively. The gains listed in Table 2 are the average of the two. The relatively large difference in the red channel can be attributed to uncertainties in the spectral band adjustments for the multispectral field observations at the BOUSSOLE site. While the standard deviations ($\sigma = 2.3\%$) were found comparable to the gains derived from BOUSSOLE data for SeaWiFS, they are twice as large as those derived from MOBY data [66]. The high uncertainties suggest that there may be residual errors in the gains that can be identified and accounted for using further cross-calibration (see Section 2.3). Therefore, we refer to these gains as preliminary. Radiometric artifacts (e.g., detector striping) and low SNRs in the NIR and SWIR bands contribute to large uncertainties in the gains [64], as do environmental factors like cirrus clouds or surface waves.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Blue</th>
<th>Green</th>
<th>Red</th>
<th>NIR</th>
<th>SWIR-I</th>
<th>SWIR-II</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETM+</td>
<td>1.030 (0.023)</td>
<td>1.01 (0.021)</td>
<td>1.0278 (0.022)</td>
<td>1.002 (0.008)</td>
<td>0.964 (0.12)</td>
<td>0.991 (0.15)</td>
</tr>
<tr>
<td>TM</td>
<td>1.047 (0.02)</td>
<td>1.011 (0.024)</td>
<td>1.0581 (0.032)</td>
<td>0.956 (0.011)</td>
<td>1.095 (0.13)</td>
<td>1.0151 (0.072)</td>
</tr>
</tbody>
</table>

Table 2. The vicarious gains for the ETM+ and TM bands. One-standard deviations are presented in round brackets. Note that these gains are derived from the NIR-SWIR-I band combination. The gains for these two bands are derived in step I.
Due to the absence of TM images over the calibration sites, we used normalized $L_{\text{sw}} (nL_{\text{sw}})$ products derived from the MODIS observations collected onboard both Aqua and Terra platforms since 1999. For this exercise, near-concurrent (defined as image-pair collections on the same day) TM and MODIS images over a number of preselected locations were analyzed. The selected locations were those considered favorable sites for potential future vicarious calibration systems, due to factors such as frequency of cloud cover and temporal and spatial stability of atmosphere and water optical properties [67]. In effect, we cross-calibrated TM data with MODIS products at these stable locations. First, the cloud- and glint-free USGS TM (Collection-1) images were visually identified. The corresponding Level-2 MODIS products were obtained from NASA’s web portal: https://oceancolor.gsfc.nasa.gov/. Note that these standard products have undergone very conservative pre-processing procedures. Areas with low Chl, i.e., a median of 0.25 mg/m$^3$, AOT(869) < 0.15, and within ±60° viewing zenith angles, were selected to perform the vicarious calibration. These products are assumed to carry minimal uncertainties in the atmospheric correction process [66].

The median $R_{\text{rs}}$ spectra were extracted from MODIS 3 x 3-element windows. The sites include West Australia ($n = 60$), US West Coast ($n = 65$), São Miguel Island ($n = 25$), and the Northern Mediterranean Sea ($n = 46$). Note that the Mediterranean site is centered around the existing BOUSSOLE site. The TM data over the other suggested calibration sites (e.g., Puerto Rico) were scarce and thus not used. A total of 196 matchups were identified for this vicarious calibration exercise. The MODIS-derived $R_{\text{rs}}$ were then adjusted for differences in spectral bands to create TM-equivalent MODIS data ($R_{\text{rs MOD-TM}}$) [61].

The TM images were processed in the vicarious calibration mode in SeaDAS by supplying $R_{\text{rs MOD-TM}}$ spectra. Given MODIS viewing zenith angles, TM image products were spatially sampled such that they accurately represent the extent “viewed” by MODIS. Similar to the ETM+ processing, the gain calculations were carried out in two steps (see Section 2.2.1) and we only present gains derived from the NIR and SWIR-I, i.e., 840–1676. Here, we assume a black ocean in these two bands, which is valid for these relatively clear waters. Figure 4 illustrates time-series of gains derived for the TM data. It should also be noted that Terra-MODIS data collected only prior to 2003 were used in the vicarious calibration [68]. The derived gains are color-coded differently for the four sites. Table 2 contains median gains and (one) standard deviations. Similar to the ETM+ gain derivations, the median statistics were...
very similar to the mean value calculated within the interquartile range. It can be inferred that the TM radiometric responses in the visible bands are biased low. Analyses of gains computed at each site indicated noticeable differences in the blue gains, i.e., 1.031, 1.046, 1.054, and 1.056 for the four sites. Furthermore, TM images contain more systematic errors (e.g., nonlinearity in instrument response) than those of ETM+. Uncertainty in spectral band adjustments is another source of error in gain computations. Nevertheless, the vicarious gains should largely minimize the biases in the long-term Landsat-5 data record (see Section 3).

![Figure 4](image-url). The vicarious gains for TM visible bands. Each data point corresponds to a TM-MODIS near-concurrent matchup, i.e., same-day overpasses, over various calibration sites. The temporally median gains and (one) standard deviation applied to TM data are provided in Table 2. Note that the gains are derived after vicariously calibrating NIR and SWIR-I bands.

### 2.3. Gain Adjustments

Considering the uncertainties (i.e., instrument artifacts and environmental noise), there remain residual biases in both TM and ETM+ gains. Therefore, these gains only provide initial estimates for the biases and slight adjustments may be inevitable. In this section, we will further examine the robustness of these gains by evaluating TM- and ETM+-derived $R_{rs}$ products against those measured/derived from AERONET-OC and OLI-derived $R_{rs}$ products and fine-tune them as needed. Because of their orbit configurations, Landsat-5 and -7 and Landsat-7 and -8 do not create opportunities for direct intercomparisons (of TOA products) for which near-simultaneous overpasses are necessary [69–71]. Inherently, the satellite pairs observe a target at 8-day intervals. It is thus imperative to choose sites with optically stable waters within this time window and assess the $R_{rs}$ products (Level-2). We chose Crater Lake and Lake Tahoe representing oligotrophic and mesotrophic water bodies, respectively [69]. Here, the OLI data products (in operation since 2013 and in good agreement with MODIS [59]) are regarded as references for evaluating ETM+ products, which are then utilized to evaluate TM products within the 1999–2012 period. These sites, however, allow only for intercomparisons over a limited signal range in the blue, i.e., 0.002 < $R_{rs}$ (485) < 0.008 [1/sr], and even smaller in the green and red channels. To extend that range, we also performed intercomparisons at AERONET-OC sites. Note that in both analyses, uncertainties in (a) the atmospheric correction and (b) OLI products are unavoidable.

To begin, TM and ETM+ data over these sites (i.e., all available matchups at AERONET-OC stations, Crater and Tahoe lakes) were processed using the computed vicarious gains (Table 2). Initial
analyses of the data revealed that the TM- and ETM+-derived \( R_{rs} \) products are slightly biased low (i.e., <0.0016 \([1/sr]\)) in the blue and red channels. The difference was found largest in the TM’s blue band.

As an example, Figure 5 illustrates the corresponding scatterplots for the OLI-ETM+ and TM-ETM+ intercomparisons in the blue bands. The root-mean-square differences (RMSD), the median relative differences (MRD), and median biases (Bias) for the OLI-ETM+ matchups can be computed to gauge the relative differences:

\[
\text{RMSD}(\lambda) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (R_{rs(i)}^{\text{ETM+}(\lambda)} - R_{rs(i)}^{\text{OLI}(\lambda)})^2}
\]

\[
\text{MRD}(\lambda) = \text{median} \left\{ \frac{R_{rs(i)}^{\text{ETM+}(\lambda)} - R_{rs(i)}^{\text{OLI}(\lambda)}}{R_{rs(i)}^{\text{AERONET}(\lambda)}} \right\} \times 100 \% 
\]

\[
\text{Bias}(\lambda) = \text{median} \left\{ \frac{R_{rs(i)}^{\text{ETM+}(\lambda)} - R_{rs(i)}^{\text{OLI}(\lambda)}}{R_{rs(i)}^{\text{AERONET}(\lambda)}} \right\} \text{ [1/sr]}
\]

The RMSD, Bias, and the slope for the linear fit are reported in Figure 5. After running a few experiments to minimize the differences in TM and ETM+ products with respects to those of AERONET-OC and OLI, final vicarious adjustments were determined. Table 3 contains these final gains. It should be noted that due to large uncertainties in the gain computations for the SWIR-I band experiments to minimize the differences in TM and ETM+ products with respects to those of Aqua-MODIS (MODISA) and Terra-MODIS (MODIST) over a handful of sites across North America (see Section 3.2.2).

Table 3. The final vicarious gains for after slight adjustments made to gains enlisted in Table 2. The adjustments were minimal for the green bands. Note that no change was made to the SWIR-II channels.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Blue</th>
<th>Green</th>
<th>Red</th>
<th>NIR</th>
<th>SWIR-I</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETM+</td>
<td>1.034</td>
<td>1.012</td>
<td>1.039</td>
<td>0.992</td>
<td>1.12</td>
</tr>
<tr>
<td>TM</td>
<td>1.039</td>
<td>1.011</td>
<td>1.058</td>
<td>0.931</td>
<td>1.21</td>
</tr>
</tbody>
</table>

\[ \text{Figure 5.} \text{ The scatterplots obtained from intercomparisons of OLI-ETM+ and TM-ETM+ } R_{rs} \text{ products over optically stable waters of Crater and Tahoe lakes. The gains tabulated in Table 2 were supplied to SeaDAS to generate these products. Final gains are tabulated in Table 3. Note that the image pairs were acquired 8 days apart.} \]
3. Results

After implementing the vicarious gains (Table 3), the generated $R_{rs}$ products are first evaluated for qualitative and quantitative rigor followed by analyzing the time-series products.

3.1. Qualitative Assessment of $R_{rs}$ Products

Two examples of products processed via SeaDAS are illustrated in Figures 6 and 7. The goal is to qualitatively demonstrate the TM and ETM+ products and visually compare them to those obtained from OLI. Note that these products are associated with different seasons/times; therefore, this analysis is not suitable to gauge consistency, which will be discussed in Sections 3.2.2 and 3.3. Note that a $7 \times 7$-element (uniform) averaging filter is applied in the aerosol removal process. Figure 6 illustrates the $R_{rs}$ products derived from TM (left), ETM+ (middle), and OLI (right) over the turbid and shallow northern San Francisco Bay area (i.e., San Pablo Bay). The natural variability (e.g., gradients in optical regimes) in water-column is clearly captured in TM and ETM+ images. Small eddies, local resuspensions, and upwelling events can visually be discerned in the TM- and ETM+-derived $R_{rs}$ products. Nevertheless, artifacts are pronounced in TM and ETM+ products when compared to those of OLI. The artifacts in TM products (in particular in the blue band) are due to memory effect, coherent noise, and striping [13]. Due to OLI’s higher SNR and absence of major artifacts, the quality of OLI products is noticeably better than those of TM and ETM+. The scan-line corrector issue, i.e., SLC-off, [3] associated with ETM+ retrievals are masked, i.e., white stripes. Furthermore, it should be noted that, although TM and ETM+ measure spectral radiances in three broad bands, these bands capture relatively distinct spectral signatures of near-surface constituents and their optical properties.

![Figure 6](image_url)

Figure 6. The $R_{rs}$ products derived from TM (left), ETM+ (middle), and OLI (right) over the northern San Francisco Bay area. The band centers shown to the right refer to TM bands. The scan-line corrector issues (i.e., SLC-off) associated with ETM+ retrievals are masked, i.e., white stripes. Varieties of physical phenomena, including upwelling events and sediment resuspensions are captured by heritage Landsat missions. Note that a $7 \times 7$-element (uniform) averaging filter is applied in the aerosol removal process.

Another demonstration of the natural variability in inland and nearshore coastal waters captured in $R_{rs}(569)$ and $R_{rs}(560)$ products derived from TM and ETM+ is illustrated in Figure 7. The images...
were acquired eight days apart in September 2006 over the eastern North Carolina inland waters and shorelines. The scan-line corrector issues (i.e., SLC-off) associated with ETM+ retrievals are masked (shown with white stripes). The TM and ETM+ clearly reveal various water types, such as relatively clear waters in the coastal shelf area, nearshore turbid waters, and CDOM-rich river waters denoted with arrows, i.e., Tar/Pamlico River. Resuspension events can also be inferred in shallower waters. These demonstrations suggest the remarkable value of archived Landsat data and their utility enabled via SeaDAS. The negative retrievals, i.e., failure of the atmospheric correction, in CDOM-rich river waters are a known problem requiring further study and investigations into the future.

![Landsat-5 (TM) and Landsat-7 (ETM+)](image)

**Figure 7.** Scene-wide \( R_{rs}(569) \) and \( R_{rs}(560) \) products derived from TM (left) and ETM+ (right) over the North Carolina shorelines and Pamlico Sound (USA). The images were acquired eight days apart on 18 September (TM) and 26 September 2006 at about 15:30 GMT. Large gradients, i.e., relatively clear coastal shelf waters, eddies, resuspensions, CDOM-dominated waters and bottom features, are captured in the two images. The CDOM-rich waters of Tar/Pamlico Rivers are indicated with arrows. Note that negative retrievals do exist in such challenging waters in the ETM+ product [72] and a \( 7 \times 7 \)-element (uniform) averaging filter was applied in the aerosol removal process.

### 3.2. Validation

For a quantitative assessment of heritage Landsat products, we utilized two data sources; (a) AERONET-OC data and (b) \( R_{rs} \) products obtained from MODIS. This was to provide as much evidence as possible to ensure that these products were, on average, in reasonable agreement with standard products.

#### 3.2.1. AERONET-OC

Since the early 2000s, the AERONET-OC data (https://aeronet.gsfc.nasa.gov) have been used as standard datasets to evaluate the quality of ocean color data products. Although the sites do not represent global in-water conditions, they provide adequate and easily accessible matchups to evaluate the performance of both atmospheric correction and instrument’s radiometric characteristics. The scatterplots for ETM+ and TM data for all the bands are shown in Figure 8. For the matchups, hazy images and areas affected by partial sunglint were removed from the analyses. A total of 216 and 79 matchups were found for ETM+ and TM data, respectively. For TM and ETM+, >90% and >65% of the matchups correspond to moderately turbid waters of the Venise (located in Northern Adriatic Sea) and Martha’s Vineyard (situated in southern coastal waters of Massachusetts) sites. The statistics were generated in a similar fashion as in Equations (3)–(5).
The computed statistics after implementing gains (Table 3) are provided in Table 4. For reference, MRD and root-mean-square error (RMSE) computed with unity gains, i.e., prior to implementing the gains, are given in round brackets. Note that RMSE was computed as in Equation (3); yet, here, we treat the AERONET-OC data as reference standards. Comparing MRD and RMSE statistics indicate that the vicarious gains, on average, have significantly reduced median biases in products (in particular for TM); thereby enhancing the overall consistency in long-term Landsat-derived products. The slope and intercepts for TM matchups allude to potential non-linearity at low radiance levels. Collectively, the statistics suggest that TM products should be handled with care when accurate/precise products (e.g., Chl) are desired. The best performances are, in general, found in the green channels. The relatively large slopes in ETM+ analyses implicate potential nonlinearity in instrument responses. We further computed the median \( R_{rs} \) spectra, i.e., \( \overline{R_{rs}} \), from AERONET-OC data to estimate percent RMSE (\( RMSE/\overline{R_{rs}} \)) in retrievals. The \( \overline{R_{rs}} \) for ETM+ and TM data were 0.0049, 0.00519, 0.00146 [1/sr] and 0.0050, 0.00371, 0.0013 [1/sr], respectively. The percent RMSEs were estimated as 36%, 28%, 62%, and 46%, 35%, 84% for ETM+ and TM’s blue, green, and red bands, respectively. Although the quality of ETM+ is adequately evaluated here using AERONET-OC data, there is still a need to better assess the quality of TM products require further scrutiny (see Section 3.2.2).

### Table 4. The statistics associated with matchup analyses of Landsat-7 and -5 at AERONET-OC sites.

The values in round brackets refer to statistics derived prior to implementing the gains listed in Table 3. The large RMSEs and deviations of slopes from unity are associated with random and systematic errors in TM and ETM+ observations. The Landsat-5 matchups are available only for the last eight years of the mission.

<table>
<thead>
<tr>
<th>Band Center [nm]</th>
<th>MRD [%]</th>
<th>RMSE [1/sr]</th>
<th>Slope</th>
<th>Intercept</th>
<th>Bias [1/sr]</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-7 (ETM+)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>478</td>
<td>9.81 (−33.8)</td>
<td>0.0018 (0.0022)</td>
<td>1.25</td>
<td>−0.000016</td>
<td>0.0006</td>
<td>0.79</td>
</tr>
<tr>
<td>560</td>
<td>3.4 (−7.9)</td>
<td>0.0013 (0.0015)</td>
<td>1.10</td>
<td>−0.000024</td>
<td>0.000099</td>
<td>0.81</td>
</tr>
<tr>
<td>661</td>
<td>1.07 (−46.5)</td>
<td>0.00076 (0.0011)</td>
<td>1.18</td>
<td>−0.000014</td>
<td>0.000111</td>
<td>0.64</td>
</tr>
<tr>
<td>Landsat-5 (TM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>485</td>
<td>13.6 (−85.5)</td>
<td>0.0020 (0.0006)</td>
<td>1.17</td>
<td>0.000013</td>
<td>0.00072</td>
<td>0.56</td>
</tr>
<tr>
<td>569</td>
<td>11.9 (−14.7)</td>
<td>0.0016 (0.0002)</td>
<td>1.04</td>
<td>0.00101</td>
<td>0.00082</td>
<td>0.85</td>
</tr>
<tr>
<td>660</td>
<td>10.1 (−116.2)</td>
<td>0.00101 (0.0003)</td>
<td>1.14</td>
<td>0.0002</td>
<td>0.00011</td>
<td>0.59</td>
</tr>
</tbody>
</table>
3.2.2. MODIS-Derived $R_{rs}$ Products

Since the AERONET-OC data do not provide a full representation of aquatic conditions in inland and nearshore coastal waters, we further evaluated TM/ETM+ products against (same-day) MODISA and MODIST products over the central basin of Lake Mead ($n = 100/101$), San Francisco (SF) Bay ($n = 63/75$), San Pablo Bay ($n = 66/69$), and the eastern basin of Lake Erie ($n = 19/29$) in North America for which a rich archive of TM data is available (limitations in the data downlink during periods of TM operations preclude a robust global analysis). The number of matchups (in round brackets) refer to the number of MODIST-ETM+ and MODISA-TM matchups, respectively. These sites collectively enable a fair assessment of TM and ETM+ products over a wider range of aquatic conditions (e.g., areas with high CDOM and/or sediment loads) and are sufficiently spatially extensive for valid ocean color products from MODIS observations.

Landsat-5, Landsat-7, and Terra were/are all morning satellites, while Aqua is in an afternoon orbit. Our initial assessments indicated noticeable uncertainties in intercomparisons due to the large time differences between ETM+ and MODISA overpasses. This was, in particular, observed over SF Bay influenced by significant tidal forcing. Therefore, for cross-validating ETM+ products, MODIST products were applied. In addition, because the two missions are in very similar orbits, near-nadir MODIST observations/products are available for cross-validating ETM+ products. On the other hand, Landsat-5 and Terra’s orbits are different such that only MODIST observations at the edge of the swaths are available for valid, same-day intercomparisons with those of TM. Such products are represented with large footprint sizes; therefore, larger uncertainties in products are expected. It was then decided to utilize MODISA products for investigating TM products, recognizing that there is two-to-three-hour time gap between data collections, which increase uncertainties in intercomparisons. The Level-1B MODIS data were processed to Level-2 products using a NIR-SWIR band combination [73]. The rest of the parameter-setting was identical to the current NASA Ocean Biology Processing Group’s standard processing. The spatial sampling for Aqua-MODIS $R_{rs}$ products were conducted by taking the median values within $3 \times 3$-element windows to avoid contaminations from land. The equivalent sample area for Landsat products (i.e., median within $3 \times 3$ km regions) were adjusted according to MODIS viewing geometries. Further, the differences in the spectral bands were accounted for [61]. Figure 9 shows the scatterplots with the statistical analyses tabulated in Table 5.

Overall, there is a reasonable correlation in MODIST-ETM+ intercomparisons. The scattered matchups correspond to artifacts in TM and ETM+ products, uncertainties in the atmospheric correction, the spectral band adjustments, and possible changes in water-column conditions (in particular for TM-MODISA intercomparisons over SF Bay and San Pablo Bay influenced by strong tides). The uncertainties due to the time difference in overpasses were also found over Lake Mead and Lake Erie when both MODIST-ETM+ and MODISA-ETM+ intercomparisons were examined. Further site-specific analyses suggested larger uncertainties in TM-derived blue and red bands in Lake Mead and Lake Erie where $R_{rs}(485)$ and $R_{rs}(660)$ were, on average, 0.0034 [1/sr] and 0.0022 [1/sr], respectively.

Table 5. The statistics associated with TM-MODIS ($n = 274$) and MODIS-ETM+ ($n = 249$) over select sites across North America. Uncertainties associated with the intercomparisons are attributed to atmospheric correction and spectral band adjustments.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-7 (ETM+)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>478</td>
<td>−9.56</td>
<td>0.00152</td>
<td>0.923</td>
<td>−0.000017</td>
<td>−0.00044</td>
<td>0.91</td>
</tr>
<tr>
<td>560</td>
<td>−3.29</td>
<td>0.00103</td>
<td>1.007</td>
<td>−0.00021</td>
<td>−0.0002</td>
<td>0.97</td>
</tr>
<tr>
<td>661</td>
<td>−8.15</td>
<td>0.00102</td>
<td>0.939</td>
<td>−0.00057</td>
<td>−0.00009</td>
<td>0.98</td>
</tr>
<tr>
<td>Landsat-5 (TM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>485</td>
<td>−7.11</td>
<td>0.002283</td>
<td>0.984</td>
<td>−0.00025</td>
<td>−0.00033</td>
<td>0.76</td>
</tr>
<tr>
<td>569</td>
<td>−1.83</td>
<td>0.001708</td>
<td>1.012</td>
<td>−0.000048</td>
<td>−0.00017</td>
<td>0.93</td>
</tr>
<tr>
<td>660</td>
<td>1.97</td>
<td>0.001905</td>
<td>0.981</td>
<td>0.000228</td>
<td>0.0000361</td>
<td>0.91</td>
</tr>
</tbody>
</table>
which were produced to minimize noise in the analyses. The three plots shown in Figures 10–12 were derived from locally processed Landsat data over 12 sites, three of which are presented in Remote Sens. 2018, 10, x FOR PEER REVIEW 13 of 23

Figure 9. The scatterplots for the TM-MODISA and MODIST-ETM+ intercomparisons. Considering the dynamic range in the signal and the uncertainties in the intercomparisons, TM and ETM+ products are in good agreement with those of MODIS. The largest scatter is found in TM-derived Rs(485) products. The larger uncertainties in the TM-MODISA analyses have aroused from two-to-three-hour time difference in their respective overpass times.

3.3. Time-Series Applications

In order to examine the long-term record of Landsat-derived Rs products, we analyzed seasonal composites generated for 12 aquatic systems across the continental USA. The seasonal composite plots were derived from locally processed Landsat data over 12 sites, three of which are presented in this manuscript. The median Rs computed in each season represents seasonal composite products, which were produced to minimize noise in the analyses. The three plots shown in Figures 10–12 correspond to the Indian River Lagoon (IRL) (27.8355N, 80.4515W), Pamlico Sound (35.1963N, 76.2670W), and San Joaquin River (38.09182N, 121.6297W). The other sites were Boston Harbor (MA), Suisan Bay (CA), Neuse River (NC), Upper Chesapeake Bay (MD), Potomac River (MD), San Francisco Bay (CA), Onondaga Lake (NY), Lake Tahoe (CA), and Crater Lake. The only products evaluated were Rs for the three visible bands (Table 1). Figure 10 shows these products for the IRL in the vicinity of Sebastian Inlet State Park (FL).

Our approach developed for compensating for non-zero Lw(NIR) (Section 2.1) significantly increased the number of valid retrievals, i.e., more than 40% of the retrievals were negative in Rs(485) products. Since data were collected eight days apart, potentially yielding biases in seasonal composites, it is difficult to infer consistency in the data products. In addition, seawater-intrusion and tidal exchanges may complicate such analyses. Seasonal Rs composites over Pamlico Sound, North Carolina, are illustrated in Figure 11. Some of the retrievals may be associated with the noise in the NIR and SWIR-I bands that propagate through the atmospheric correction [64]. Note that prior to accounting for Lw(NIR) the entire record in the blue band was invalid. The time-series plot for the San Joaquin River (California) is shown in Figure 12. For this riverine system dominated with scattering particles, Lw(NIR) should be accounted for to allow for valid and/or accurate retrievals [54]. Similar results were found for a downstream system, i.e., Suisan Bay. The impact of the proposed methodology was most noticeable in such CDOM-dominated waters, i.e., frequent negative retrievals in the blue and green were observed prior to its implementation.
Figure 10. The seasonal composites of $R_{rs}$ products derived from Landsat-5 (solid circles), Landsat-7 (triangle), and Landsat-8 (hollow circles) over the Indian River Lagoon (27.8355N, 80.4515W), Florida. The composites were obtained by taking the median of retrievals in each season. The images were processed via SeaDAS. The methodology for accounting for non-negligible $L_w$(NIR) significantly improved the number of valid (non-negative) retrievals.

Figure 11. Seasonal composites of $R_{rs}$ products derived from Landsat-5 (solid circles), Landsat-7 (triangle), and Landsat-8 (hollow circles) in Pamlico Sound, North Carolina (35.1963N, 76.2670W). The composites were obtained by taking the median of retrievals in each season. The images processed via SeaDAS. The product consistency can be inferred in the overlapping periods indicating the utility of long-term Landsat data record for aquatic science studies. Note that because of the differences in the temporal sampling (i.e., eight days) and considering various tidal stages, a robust analysis of product consistency is not possible.
was processed by introducing various noise levels to simulate TM, ETM+, and OLI performances [5,12].

We demonstrated that following the cross-calibration and vicarious calibration of TM performance [5] representing a high-fidelity ocean color sensor. To generate various scenarios, noise (Table 6) was processed by introducing various noise levels to simulate TM, ETM+, and OLI performances [5,12]. For each sensor, noise ($\sigma_{R_{\text{rs}}}$) is estimated via $L_{\text{i}}/\text{SNR}$. The median $L_{\text{i}}$ and $R_{\text{rs}}$ (derived from OLI subset) are tabulated in Table 6. The SNRs are adopted from References [5,74].

Although we showed that TM- and ETM+-derived $R_{\text{rs}}$ products agree well, on average, with those of OLI [59], OLI’s SNR is three-to-four times higher than those of TM and ETM+ [5,12]. To better understand the implications of these differences in $R_{\text{rs}}$ products, a brief sensitivity analysis is carried out. Here, a small subset (10 x 10) of a cloud-free OLI image (ID:LC80220402014227LGN00, solar zenith angle of 26.3°) with low aerosol load, i.e., AOT(865) = 0.08, and over moderately turbid waters (Table 6) was processed by introducing various noise levels to simulate TM, ETM+, and OLI performances [5,12]. For each sensor, noise ($n_{\text{s}}$) is estimated via $L_{\text{i}}/\text{SNR}$. The median $L_{\text{i}}$ and $R_{\text{rs}}$ (derived from OLI subset) are tabulated in Table 6. The SNRs are adopted from References [5,74].

Note that OLI’s SNR from [74] were scaled to match the SNRs derived from typical radiances reported in Reference [5] for TM, ETM+, and MODIS. The analyses were conducted via Monte Carlo simulations for all the bands (Table 1) except the SWIR-II band, which was not used in the aerosol removal. To provide a reference, the OLI subset was also processed with the MODIS noise performance [5] representing a high-fidelity ocean color sensor. To generate various scenarios, $L_{\text{i}}$ were allowed to vary within the $L_{\text{i}} \pm n_{\text{s}}$ range with $n_{\text{s}}$ having a normal distribution. All the permutations resulted in ~7000 SeaDAS runs, which were repeated four times for TM, ETM+, OLI, and MODIS.

The uncertainties ($\sigma_{R_{\text{rs}}}$) provided in Table 4 represent one standard deviation derived from the resultant $R_{\text{rs}}$ distributions. The derived uncertainties provide guidance on expected random

Figure 12. Seasonal $R_{\text{rs}}$ composites derived from Landsat-5 (solid circles), Landsat-7 (triangle), and Landsat-8 (hollow circles) in San Joaquin River (38.09182N, 121.6297W), California. The composites were obtained by taking the median of retrievals in each season. The images were processed via SeaDAS. The product consistency can be inferred in the cross-mission overlapping periods. Few negative retrievals are attributed to imperfect corrections for the contributing water-leaving radiance in the NIR channel.

4. Discussions: Product Quality

The analyses thus far have highlighted some of the potentials and limitations of Landsat archive for long-term aquatic science studies. Now, with a rigorous, physics-based processing system in place, the science community can further explore the utility of the Landsat data record in science studies, such as examining impacts of human-induced activities (e.g., landuse/landcover change) on water quality. We demonstrated that following the cross-calibration and vicarious calibration of TM and ETM+ data, the biases in $R_{\text{rs}}$ products are significantly reduced. In this section, we will further elaborate on the limitations of TM and ETM+ $R_{\text{rs}}$ products with regard to noise contributions and processing flaws. A brief discussion on adjacency effects is also provided.

4.1. Signal-to-Noise Ratio (SNR)

The analyses thus far have highlighted some of the potentials and limitations of Landsat archive for long-term aquatic science studies. Now, with a rigorous, physics-based processing system in place, the science community can further explore the utility of the Landsat data record in science studies, such as examining impacts of human-induced activities (e.g., landuse/landcover change) on water quality. We demonstrated that following the cross-calibration and vicarious calibration of TM and ETM+ data, the biases in $R_{\text{rs}}$ products are significantly reduced. In this section, we will further elaborate on the limitations of TM and ETM+ $R_{\text{rs}}$ products with regard to noise contributions and processing flaws. A brief discussion on adjacency effects is also provided.
variability in heritage Landsat products compared to those derived from OLI and MODIS. In general, the uncertainties in TM are two-to-five times higher than those of OLI. Given the magnitude of the $R_{rs}$ spectrum, uncertainties in TM-, ETM+-, OLI-, and MODIS-derived $R_{rs}$ products are, on average, ~25%, ~27%, ~8%, and ~2%, in the blue and green bands, respectively. This implies that subtle changes in $R_{rs}$, i.e., < $\sigma_{R_{rs}}$, leading to small changes in water quality parameters, may not be detected. In inland and nearshore coastal waters (e.g., bays, estuaries) where concentrations of TSS and Chl are often much higher than those in coastal ocean waters, highly precise concentrations (at levels expected from ocean color missions) may not be required. Therefore, end-users may utilize heritage $R_{rs}$ products with the “fit-for-purpose” concept in mind, i.e., products could be valuable for one application but may not be viable for another.

In addition, the choice of algorithm may alter how the noise is propagated to end products. For example, band-ratio algorithms [75,76] amplify the uncertainties. It is, thus, critical that users apply smoothing techniques, such as spatial filtering, to reduce the uncertainties. Assuming a 7 x 7-element (uniform) averaging window, $\sigma_{R_{rs}}$ in Table 6 are reduced by a factor of 1/7. This analysis was carried out for TOA radiance. In practice, $n_s$ increases with increase in $L_t$ leading to variable uncertainties across a scene. Note, however, that the median $L_t$ considered in this analysis is higher than typical radiances encountered in Landsat scenes [74]. Hence, the computed values in Table 6 provide an upper bound for (random component of) uncertainties.

It should also be noted that although ETM+’s SNRs in the visible bands are higher than those of TM, uncertainties in TM-derived $R_{rs}$ are slightly smaller. This is attributed to TM’s higher SNR in the NIR band employed in the aerosol removal [64].

### 4.2. Image Artifacts and Processing Flaws

Heritage Landsat data, in particular TM images, often contain artifacts (e.g., striping, coherent noise, memory effects) suggesting that $R_{rs}$ products should always be interpreted with care. Some of these artifacts are evident in Figure 6. Further, we found anomalies in TM products within the 1990–94 interval. This is highlighted in Figure 13 where time-series of $R_{rs}$ products over fairly optically stable inland waters [59], i.e., Crater Lake and Lake Tahoe, are illustrated. Further analyses of $L_t$ and other intermediate products (e.g., $L_t$) indicated unusually high $L_t$ in the red, NIR, and SWIR-I bands. These high radiometric responses yield invalid $R_{rs}$ products [64]. No such anomalies were identified for ETM+ time-series, i.e., fewer artifacts present in ETM+ data. Nonetheless, occasional noisy retrievals were observed in time-series plots. Another observation was nonlinearity in response when adjusting vicarious gains in Section 2.3. This was, in particular, confirmed in the ETM+ matchup analysis (Table 5). Overall, the products may be used with caution in oligotrophic (e.g., Lake Tahoe), mesotrophic (e.g., Lake Tahoe), and CDOM-rich waters.

<table>
<thead>
<tr>
<th>Band (nm)</th>
<th>485</th>
<th>569</th>
<th>660</th>
<th>840</th>
<th>1676</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SNR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TM</td>
<td>71.8</td>
<td>9.1 $\times 10^{-4}$</td>
<td>42.3</td>
<td>7.2 $\times 10^{-4}$</td>
<td>29.0</td>
</tr>
<tr>
<td>ETM+</td>
<td>77.9</td>
<td>1.0 $\times 10^{-3}$</td>
<td>69.7</td>
<td>8.5 $\times 10^{-4}$</td>
<td>40.7</td>
</tr>
<tr>
<td>OLI</td>
<td>382.8</td>
<td>3.0 $\times 10^{-4}$</td>
<td>256.9</td>
<td>2.4 $\times 10^{-4}$</td>
<td>135.1</td>
</tr>
<tr>
<td>MODIS</td>
<td>2209.2</td>
<td>1.0 $\times 10^{-4}$</td>
<td>2401.8</td>
<td>6.9 $\times 10^{-5}$</td>
<td>1422.0</td>
</tr>
<tr>
<td>$L_t$</td>
<td>54.2</td>
<td>29.9</td>
<td>15.2</td>
<td>4.3</td>
<td>0.30</td>
</tr>
<tr>
<td>$R_{rs}$</td>
<td>0.0031</td>
<td>0.0028</td>
<td>0.0028</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 6. The SNR, TOA radiance ($L_t$), and the uncertainties in $R_{rs}$ ($\sigma_{R_{rs}}$) associated with TM, ETM+, OLI, and MODIS. $\sigma_{R_{rs}}$ represents one standard deviation corresponding to the $R_{rs}$ distributions derived from many SeaDAS runs for a small OLI subset. Note that the aerosol removal was conducted using the 840–1676 band combination and OLI’s SNR from [74] were scaled to provide comparable SNRs from many SeaDAS runs for a small OLI subset. Note that the aerosol removal was conducted using the 840–1676 band combination and OLI’s SNR from [74] were scaled to provide comparable SNRs.
To infer uncertainties in products, we quadratically added up systematic and random errors. It is, in general, difficult to fully characterize systematic errors for heritage sensors/products. However, according to the matchup analyses (Table 4), intercomparisons with MODIS (Table 5), and SNR assessments (Table 6), it is possible to speculate a maximum uncertainty estimate of 0.0019, 0.0018, and 0.0012 [1/sr] in the blue, green, and red channels of TM and ETM+ products, respectively. These estimates are calculated by taking the average of RMSEs in Tables 4 and 5. Also, note that these maximum uncertainties are estimated recognizing (a) implementation of smoothing filters for reducing random noise contribution and (b) uncertainties in the matchup/intercomparison analyses. The changes and/or anomalies larger than the above threshold are expected to be detected in the heritage products. According to the matchup analyses (Section 3.2), the quality of the TM-derived Rs(660) were found to be reliable for relatively highly backscattering waters where Rs(660) > 0.004 [1/sr]. Our processing system may also contribute to uncertainties. For instance, it is likely that the retrievals over Crater and Tahoe lakes are affected by imperfect corrections of water-vapor transmission. Bear in mind that TM and ETM+’s NIR channels (Figure 1), i.e., ~140 nm wide, contain a water absorption feature centered around ~820 nm, whereas OLI’s NIR band is relatively narrow (Figure 1). Furthermore, some of the existing noise in the analyses may be attributed to failure in cloud-screening/-masking conducted using thresholds adopted from ocean color data processing. There are also inaccurate retrievals in periods where CDOM loads are high leading to the underestimation of Rs. More research must be dedicated to these topics to enhance these historic Rs products.

4.3. Adjacency Effects

The adjacency effect (AE) is another prime topic when using Landsat data in inland waters. Such signal complicating Rs retrievals were pronounced over Onondaga Lake (NY). To qualitatively evaluate presence/absence of adjacency effects, the Rayleigh-corrected ($L_{rc}$) data over three sites, including Onondaga Lake (NY), Chowan River (NC), and Boston Harbor (MA) were analyzed. The exact locations for these sites are (43.0919N, 76.2145W), (36.1763N, 76.7378W), and (42.3401, 70.9824W), respectively. To do so, below index (AE_Ind) was computed

$$AE_{\text{Ind}} = \frac{L_{rc}(\text{NIR}) - L_{rc}(\text{Red})}{L_{rc}(\text{NIR}) + L_{rc}(\text{Red})} \quad (6)$$

Figure 13. The time-series of Rs products in the blue bands over optically stable waters of Crater and Tahoe lakes. The overall signal level is very similar in time. A potential anomaly found for TM products is highlighted in the Crater Lake time-series.
Although AE_Ind (ranging from $-1$ to 1) includes aerosol contributions, it provides insights into whether the $L_{rc}$ spectrum deviates from expected trends, i.e., a monotonically decreasing one towards longer wavelengths. Therefore, the higher the AE_Ind, the stronger the NIR signal is, potentially alluding to the presence of adjacency effects. The time-series plots of AE_Ind extracted from Boston Harbor, Chowan River, and Onondaga Lake are illustrated in Figure 14. From the analyses of the $R_{rs}$ data products (not shown here) over the Boston Harbor and the corresponding time-series shown in Figure 14, it appears that there is negligible adjacency effect. Considering the plots associated with Boston Harbor as the reference, it is found that there is AE in some instances of Chowan River retrievals. The impact of AE is much more pronounced over Onondaga Lake where AE_Ind is frequently very close to zero and even positive for some instances. Since the topography around Onondaga Lake and Chowan River is fairly flat and the two systems are comparable in their morphology, i.e., <2 km wide, it can be concluded that environmental conditions (e.g., wind, humidity, turbulence at land-atmosphere interface) play a major role in the magnitude of adjacency signal. Note further that the increased in the 1992–1996 period (marked with ovals) verifies a potential issue in TM products (as in Figure 13), which requires further investigation. Since products are not corrected for aerosol contributions, irregularities in Landsat-5 orbit may explain such anomalous trends in time-series analyses that do not accounted for variations in solar angles/radiation [77].

While with only six spectral measurements of TM and ETM+, it is nearly impossible to remove adjacency effects, it is certainly possible to minimize such effects with prior knowledge of the spectral characteristics of surrounding land targets, dominant environmental conditions and topography. Therefore, it is cautioned that the adjacency signals may not be assumed for all inland water sites (simply because they are small or narrow water bodies). For instance, land signals were not identified over San Joaquin River (CA), which is narrower than Onondaga Lake or Chowan River. Extensive research is required to identify water bodies susceptible to these effects, understand the dominant environmental conditions, and minimize them for the past and existing multispectral missions. With future hyperspectral missions planned/proposed for the next decade, it may be possible to more accurately account for these unwanted signals.

Figure 14. The AE_Ind computed for three sites to analyze presence/absence of adjacency effects (AE). The higher the index, the larger the magnitude of AE is. The Onondaga Lake has the highest and most frequent presence of AE whereas the effect is deemed to be negligible or absent over Boston Harbor. Note that the NIR bands of TM and ETM+ are much wider than that of OLI. This has likely resulted in smaller OLI-derived AE_Ind. The higher-than-normal AE_Ind highlighted with ovals may be due to abnormalities in Landsat-5 orbit in early/mid 90’s [77].
5. Conclusions

This work described the potentials and limitations of heritage Landsat missions for use in aquatic studies. Here, we demonstrated that the Landsat-5 and Landsat-7 data records, when corrected for atmospheric effects and vicariously calibrated via the algorithms and software available in the SeaWiFS Data Analysis System (SeaDAS), yield remote-sensing reflectance ($R_{rs}$) products that, on average, agree with those derived from in situ measurements and/or other ocean color products. We further proposed a slightly different strategy to account for $L_w$(NIR) for TM and ETM+, which were/are missing the 443 nm channel. With careful regional or global algorithm developments, the processing system has enabled the creation of a long-term record of aquatic science products since 1984. Nevertheless, larger uncertainties in TM and ETM+ due to random and systematic noise, along with the availability of only three visible spectral bands, limit the choice of algorithms. Hence, algorithm developments for estimating turbidity and concentrations of TSS may be conducted with care. Following our analyses of several water bodies in North America, the TM and ETM+ data products are recommended to be developed and used in moderately to highly turbid waters (i.e., $R_{rs}(660) > 0.004 \, [1/\text{sr}]$), where instrument artifacts are less noticeable. Overall, although the uncertainties in TM and ETM+ products are higher than those of OLI, they are still valuable assets for retrospective studies. This may include assessing changes in water quality of lakes, reservoirs, water supplies, and coastal receiving waters due to short-/long-term landuse/landcover changes as well as examining climate variability (i.e., weather patterns) and its impacts on water quality (e.g., location, frequency, and persistence in historic harmful algal blooms). There is also a need for the community to measure and quantify optical properties of water constituents in these poorly known aquatic systems. A more extensive database of in-water optical properties will enable developing regional or global long-term aquatic science products. Future research should focus on (a) developing algorithms that are less prone to random noise, (b) enhancing atmospheric correction to account for non-negligible near-infrared reflectance in extremely turbid or CDOM-rich waters, and (c) characterizing adjacency effects under various environmental/climatic conditions.

Author Contributions: The analyses were all performed by S.V.B., N.P. led and wrote the manuscript, B.A.F. provided the necessary look-up-tables, and S.S. integrated the processing in SeaDAS.

Funding: Nima Pahlevan has been funded under NASA ROSES #NNX16AI16G, USGS Landsat Science Team Award #140G0118C0011, and the Massachusetts Institute of Technology SeaGrant Award # 2015-R/RC-140.

Acknowledgments: The authors are sincerely grateful to Wesley Moses for his comments and commitment to enhancing the quality of this manuscript. The reviewers’ efforts which helped enrich the discussions on product uncertainties are appreciated as well. We also would like to thank Landsat-8 project scientist Jim Irons, with NASA Goddard Space Flight Center, for his encouragements and support of this study. The authors also thank the MOBY Operations Team and the BOUSSOLE project for maintaining and disseminating in situ radiometric data. We also thank three anonymous reviewers for their very constructive comments.

Conflicts of Interest: The authors declare no conflicts of interest.

References


69. Pahlevan, N.; Schott, J.R. Characterizing the relative calibration of Landsat-7 (ETM+) visible bands with Terra (MODIS) over clear waters: The implications for monitoring water resources. Remote Sens. Environ. 2012, 125, 167–180. [CrossRef]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).