remote sensing

MDPI

Editorial

Science of Landsat Analysis Ready Data

Zhe Zhu

Department of Natural Resources and the Environment, University of Connecticut, Storrs, CT 06269, USA; zhe@uconn.edu; Tel.: +01-860-486-6885

Received: 15 September 2019; Accepted: 16 September 2019; Published: 18 September 2019

Abstract: The free and open policy of Landsat data in 2008 completely changed the way that Landsat data was analyzed and used, particularly for applications such as time series analysis. Nine years later, the United States Geological Survey (USGS) released the first version of Landsat Analysis Ready Data (ARD) for the United States, which was another milestone in Landsat history. The Landsat time series is so convenient and easy to use and has triggered science that was not possible a few decades ago. In this Editorial, we review the current status of Landsat ARD, introduce scientific studies of Landsat ARD from this special issue, and discuss global Landsat ARD.

Keywords: Landsat; USGS; ARD; Time Series

---

1. Introduction

The series of Landsats 1–8 has been the gold standard for satellite remote sensing, mainly because of its long history of relatively high radiometric, spatial, temporal, and spectral resolutions [1–4]. The free and open policy of Landsat data, started in 2008, has made Landsat data even more beneficial to society, government, the private sector, and research institutions [5]. One of the major challenges of using the dense Landsat time series was the lack of standard and operational algorithms to create consistent cloud, cloud shadow, and snow free surface reflectance observations. This was no longer an issue when the United States Geological Survey (USGS) started to operationally produce Landsat surface reflectance with quality assessment (QA) bands, in which LEDAPS [6] and LaSRC [7] were used for atmospheric correction, and a C version of Fmask (CFmask) was used to create the QA bands [8–10]. However, because of slight orbital drift, these Landsat surface reflectance products still cannot be compared directly for detecting landscape change or used for time series analysis (i.e., images from the same Landsat path/row are not exactly the same size and location). Therefore, the USGS made another step forward by providing Landsat Analysis Ready Data (ARD), which provides Landsat data for the conterminous United States (CONUS), Alaska, and Hawaii in formats that can be directly compared for change and time series analysis.

The current Landsat ARD are created from Landsat 4–5 Thematic Mapper (TM) Tier 1 data, Landsat 7 Enhanced Thematic Mapper Plus (ETM) Tier 1 data, and Landsat 8 Operational Land Imager (OLI)/Thermal Infrared Sensor (TIRS) Tier 1/Tier 2 data. Landsat 1–5 Multispectral Sensor (MSS) ARD will also be included in Landsat ARD products in the future. Landsat ARD are consistently processed to the highest scientific standards and level of processing required for time series analysis. For each Landsat ARD tile, there are five major layers including: (1) Top of atmosphere (TOA) reflectance; (2) TOA brightness temperature (BT); (3) surface reflectance (SR); (4) provisional surface temperature (PST); and (5) pixel QA band.

2. Science of Landsat ARD in This special Issue

The release of Landsat ARD makes the application of Landsat data to time series analysis a lot easier and opens doors for many scientific applications. This special issue only introduces a small proportion of (but some of the first) scientific work generated from Landsat ARD, and a brief summary of the seven papers published in this special issue are provided below.

Dwyer et al. [11] provided a comprehensive explanation of the Landsat ARD. They first introduced the ARD processing and tile structure, in which the specific inputs used for created Landsat ARD, the detailed ARD projection information, and the method for tiling were explained and illustrated. They also introduced the major contents provided within Landsat ARD, which include TOA reflectance and BT, viewing and solar geometry data, SR, QA bands, and filename convention, format, metadata, and documentation. Considering the fact that many of the authors are also the creators of Landsat ARD products, this paper provides firsthand information on Landsat ARD.

Shi et al. [12] evaluated and improved the Landsat ARD data and made Landsat time series more consistent with each other. In this study, four major processing streamlines for improving data consistency (in the temporal domain) were explored, which included data resampling, cloud/cloud shadow detection, bidirectional reflectance distribution (BRDF) correction, and topographic correction. They made four major observations as follows: First, the single resampled data (ARD) are more consistent than double-resampled Collection 1 data, though the improvement is small. Second, improved cloud and cloud shadow detection methods (Fmask 4.0 [13] vs Fmask 3.3) can moderately improve data consistency. Third, BRDF correction contributed the most to improving data consistency. Finally, topographic correction cannot improve data consistency, and it may even have negative impacts on the consistency of Landsat time series. Therefore, to make Landsat ARD more consistent, better screening of cloud and cloud shadow (e.g., Fmask 4.0) and BRDF correction are highly recommended.

Egorov et al. [14] analyzed Landsats 4–7 ARD coverage from 1982 to 2017 over the CONUS. In this study, the average annual number of non-cloudy observations (including cirrus clouds, cloud shadow, and snow) within each ARD tile varied from 0.5 to 16.8 for Landsat 4 TM, 11.1 to 22.8 for Landsat 5 TM, 9.7 to 21.7 for Landsat 7 ETM+, and 14.2 to 20.1 if all three sensors were used. Moreover, when all three sensors were combined, the most frequent number of observations was 18 to 20 for 23% of the tiles. They also identified Landsat ARD tiles with the most and least frequently observed ARD and admitted that the reported results may have overestimated the number of good observations because they did not consider cirrus clouds and shadows.

Ernst et al. [15] evaluated the impacts of cloud detection and saturation for 31 years of Landsat ARD within Digital Earth Australia. They found cloud commission errors (from Fmask 3.2), and sensor saturation rates showed specific characteristics for different targets, which can lead to an imbalanced data density for different land surface types. They concluded that the level of detail in pixel quality flags are pivotal for better use of Landsat ARD.

Brooks et al. [16] applied a window regression method that was originally developed to impute low-quality moderate resolution imaging spectroradiometer (MODIS) to Landsat ARD between 2014 and 2016 to gap-fill Landsat ETM+ SLC-off data. For the five study areas, root mean square error derived from the observed reflectance ranged from 3.7%–7.6% for different spectral bands, and it outperformed other algorithms such as the neighborhood similar pixel interpolator gap-filling algorithm. This method has the potential to provide Landsat ARD that are free from any artifacts or missing values.

Yan and Roy [17] proposed another large-area gap-filling method called spectral-angle-mapper based spatio-temporal similarity (SAMSTS). The SAMSTS algorithm was demonstrated using half a year's Landsat 8 surface reflectance over three areas in California, Minnesota, and Kansas that contained a variety of land cover types. The gap-filling accuracy of this approach was higher than the simple closest temporal pixel substitution gap-filling approach and the sinusoidal harmonic model-based approach. The root mean square differences were less than 0.02, which was at a similar magnitude to the OLI reflectance calibration accuracy.

Egorov et al. [18] demonstrated mapping percent tree cover using Landsat ARD and evaluated its sensitivity with respect to Landsat ARD processing levels. In this study, five years of Landsat 5 and Landsat 7 ARD at 12 tiles were tested. Four processing levels of ARD were evaluated independently, including: (1) TOA ARD; (2) SR ARD; (3) BRDF adjusted SR ARD; and (4) weekly composited BRDF adjusted SR ARD. The SR ARD provided the highest mapping accuracies, but the differences between the four scenarios were small when modest amounts of training data were used. On the other hand, the TOA ARD provided the most accurate maps when only a small amount of training data were used for establishing the bagged regression tree. This study demonstrated the importance of including TOA reflectance in Landsat ARD, as some researchers may rely more on TOA reflectance than SR.

3. Global Landsat ARD

The US Landsat ARD have already demonstrated their value, and it is imperative to expand this dataset to other parts of the world. The USGS is planning to generate global Landsat ARD based on a newer Landsat collection (Collection 2) and will include images from Landsat 9 if it is successfully launched in December 2020. However, there is no unanimous preference in global Landsat ARD projection among the Landsat science team and Landsat International Cooperator Ground Station managers. The USGS scientists and Landsat science team are currently evaluating different projection frameworks, such as (1) a single global projection (e.g., Sinusoidal); (2) a series of continental-scale projects (e.g., optimal equal-area projection for each continent); (3) UTM-based military reference grid system (MGRS) used by ESA for Sentinel-2; and (4) existing scene-based UTM. No matter what projection framework is selected, global Landsat ARD will play a major role in future remote sensing science and applications.

4. Conclusions

The release of Landsat ARD has made time series analysis a lot easier. Their scientific data layers, such as TOA reflectance, TOA brightness temperature, surface reflectance, provisional surface temperature, and pixel QA band, will facilitate many remote sensing applications and trigger new scientific discoveries.

Author Contributions: Z.Z. designed and wrote the paper.

Funding: This research was supported by the USGS–NASA Landsat Science Team (LST) Program for Toward Near Real-Time Monitoring and Characterization of Land Surface Change for the Conterminous US (#140G0119C0008).

Acknowledgments: I would like to acknowledge the major contribution from USGS that made Landsat data free of charge and provided high-level science products such as Landsat ARD. The content of this document does not necessarily represent the views or policies of the Department of the Interior, nor does mention of trade names, commercial products or organizations imply endorsement by the U.S. Government.

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

References

1. Wulder, M.A.; Loveland, T.R.; Roy, D.P.; Crawford, C.J.; Masek, J.G.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Belward, A.S.; Cohen, W.B.; et al. Current status of Landsat program, science, and applications. *Remote Sens. Environ.* **2019**, *225*, 127–147. [CrossRef]
2. Wulder, M.A.; White, J.C.; Loveland, T.R.; Woodcock, C.E.; Belward, A.S.; Cohen, W.B.; Fosnight, E.A.; Shaw, J.; Masek, J.G.; Roy, D.P. The global Landsat archive: Status, consolidation, and direction. *Remote Sens. Environ.* **2016**, *185*, 271–283. [CrossRef]
3. 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.; et al. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172. [CrossRef]
4. Schott, J.R.; Gerace, A.; Woodcock, C.E.; Wang, S.; Zhu, Z.; Wynne, R.H.; Blinn, C.E. The impact of improved signal-to-noise ratios on algorithm performance: Case studies for Landsat class instruments. *Remote Sens. Environ.* **2016**, *185*, 37–45. [CrossRef]
5. Zhu, Z.; Wulder, M.A.; Roy, D.P.; Woodcock, C.E.; Hansen, M.C.; Radeloff, V.C.; Healey, S.P.; Schaaf, C.; Hostert, P.; Strobl, P.; et al. Benefits of the free and open Landsat data policy. *Remote Sens. Environ.* **2019**, *224*, 382–385. [CrossRef]

6. Masek, J.G.; Vermote, E.F.; Saleous, N.E.; Wolfe, R.; Hall, F.G.; Huemmrich, K.F.; Gao, F.; Kutler, J.; Lim, T.K. A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 68–72. [CrossRef]
7. Vermote, E.; Justice, C.; Claverie, M.; Franch, B. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sens. Environ.* **2016**, *185*, 46–56. [CrossRef]
8. Foga, S.; Scaramuzza, P.L.; Guo, S.; Zhu, Z.; Dilley, R.D.; Beckmann, T.; Schmidt, G.L.; Dwyer, J.L.; Joseph Hughes, M.; Laue, B. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sens. Environ.* **2017**, *194*, 379–390. [CrossRef]
9. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [CrossRef]
10. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [CrossRef]
11. Dwyer, J.L.; Roy, D.P.; Sauer, B.; Jenkerson, C.B.; Zhang, H.K.; Lymburner, L. Analysis ready data: Enabling analysis of the Landsat archive. *Remote Sens.* **2018**, *10*, 1363.
12. Qiu, S.; Lin, Y.; Shang, R.; Zhang, J.; Ma, L.; Zhu, Z.; Qiu, S.; Lin, Y.; Shang, R.; Zhang, J.; et al. Making landsat time series consistent: Evaluating and improving Landsat analysis ready data. *Remote Sens.* **2019**, *11*, 51. [CrossRef]
13. Qiu, S.; Zhu, Z.; He, B. Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery. *Remote Sens. Environ.* **2019**, *231*, 111205. [CrossRef]
14. Egorov, A.V.; Roy, D.P.; Zhang, H.K.; Li, Z.; Yan, L.; Huang, H. Landsat 4, 5 and 7 (1982 to 2017) Analysis Ready Data (ARD) observation coverage over the conterminous United States and implications for terrestrial monitoring. *Remote Sens.* **2019**, *11*, 447. [CrossRef]
15. Ernst, S.; Lymburner, L.; Sixsmith, J. Implications of pixel quality flags on the observation density of a continental Landsat archive. *Remote Sens.* **2018**, *10*, 1570. [CrossRef]
16. Brooks, E.B.; Wynne, R.H.; Thomas, V.A. Using window regression to gap-fill Landsat ETM+ post SLC-Off data. *Remote Sens.* **2018**, *10*, 1502. [CrossRef]
17. Yan, L.; Roy, D.P. Large-area gap filling of Landsat reflectance time series by spectral-angle-mapper based spatio-temporal similarity (SAMSTS). *Remote Sens.* **2018**, *10*, 609. [CrossRef]
18. Egorov, A.V.; Roy, D.P.; Zhang, H.K.; Hansen, M.C.; Kommareddy, A. Demonstration of percent tree cover mapping using Landsat Analysis Ready Data (ARD) and sensitivity with respect to Landsat ARD processing level. *Remote Sens.* **2018**, *10*, 209. [CrossRef]

© 2019 by the author. 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/).