

Article

Landsat 4, 5 and 7 (1982 to 2017) Analysis Ready Data (ARD) Observation Coverage over the Conterminous United States and Implications for Terrestrial Monitoring

Alexey V. Egorov ¹, David P. Roy ^{2,*}, Hankui K. Zhang ¹, Zhongbin Li ¹, Lin Yan ¹
and Haiyan Huang ¹

¹ Geospatial Sciences Center of Excellence, South Dakota State University, Brookings, SD 57007, USA; alexey.egorov@sdstate.edu (A.V.E.); hankui.zhang@sdstate.edu (H.K.Z.); zhongbin.li@sdstate.edu (Z.L.); lin.yan@sdstate.edu (L.Y.); haiyan.huang@sdstate.edu (H.H.)

² Department of Geography, Environment & Spatial Sciences and Center for Global Change and Earth Observations, Michigan State University, East Lansing, MI 48824, USA

* Correspondence: roydavi1@msu.edu; Tel.: +1-517-355-3898

Received: 28 January 2019; Accepted: 17 February 2019; Published: 21 February 2019



Abstract: The Landsat Analysis Ready Data (ARD) are designed to make the U.S. Landsat archive straightforward to use. In this paper, the availability of the Landsat 4 and 5 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) ARD over the conterminous United States (CONUS) are quantified for a 36-year period (1 January 1982 to 31 December 2017). Complex patterns of ARD availability occur due to the satellite orbit and sensor geometry, cloud, sensor acquisition and health issues and because of changing relative orientation of the ARD tiles with respect to the Landsat orbit paths. Quantitative per-pixel and summary ARD tile results are reported. Within the CONUS, the average annual number of non-cloudy observations in each 150 × 150 km ARD tile varies from 0.53 to 16.80 (Landsat 4 TM), 11.08 to 22.83 (Landsat 5 TM), 9.73 to 21.72 (Landsat 7 ETM+) and 14.23 to 30.07 (all three sensors). The annual number was most frequently only 2 to 4 Landsat 4 TM observations (36% of the CONUS tiles), increasing to 14 to 16 Landsat 5 TM observations (26% of tiles), 12 to 14 Landsat 7 ETM+ observations (31% of tiles) and 18 to 20 observations (23% of tiles) when considering all three sensors. The most frequently observed ARD tiles were in the arid south-west and in certain mountain rain shadow regions and the least observed tiles were in the north-east, around the Great Lakes and along parts of the north-west coast. The quality of time series algorithm results is expected to be reduced at ARD tiles with low reported availability. The smallest annual number of cloud-free observations for the Landsat 5 TM are over ARD tile h28v04 (northern New York state), for Landsat 7 ETM+ are over tile h25v07 (Ohio and Pennsylvania) and for Landsat 4 TM are over tile h22v08 (northern Indiana). The greatest annual number of cloud-free observations for the Landsat 5 TM and 7 ETM+ ARD are over southern California ARD tile h04v11 and for the Landsat 4 TM are over southern Arizona tile h06v13. The reported results likely overestimate the number of good surface observations because shadows and cirrus clouds were not considered. Implications of the findings for terrestrial monitoring and future ARD research are discussed.

Keywords: Landsat; analysis ready data; observation coverage; big data

1. Introduction

The advent of free and open Landsat data, increasing computational capabilities and analysis ready data (ARD), have stimulated the uptake of Landsat data and a growth in Landsat time series

research and applications [1]. Recently, the United States Geological Survey (USGS) generated Landsat ARD over the conterminous United States (CONUS), Alaska and Hawaii [2]. The ARD are available for all the Landsat sensors except the Multispectral Scanner (MSS) [3]. The ARD are designed to be straightforward to use and include atmospherically corrected data and are provided in fixed non-overlapping tiles that are “stackable” for time series analysis [2]. The ARD are stored with band specific scaling factors and fill values to denote no data was observed at an ARD tile pixel location.

The availability of satellite observations influences surface monitoring capabilities. Higher temporal resolution observations enable more reliable change detection and surface monitoring. However, the availability of Landsat surface observations changes geographically because of cloud cover at the time of Landsat overpass, the increasing spatial overlap of adjacent Landsat orbits further poleward and because not every sunlit land image is acquired due to a variety of mission and operational constraints [4–6].

The purpose of this study is to quantify the availability of the CONUS ARD data. Previous Landsat availability analyses have been undertaken over large areas considering Landsat image geographic and cloud metadata information [4,5,7–9], satellite-sensor orbit models [10,11] and Landsat per-pixel image data [12] but for limited time periods. In this study, the Landsat temporal data availability is quantified in detail considering *every* CONUS ARD 30 m pixel location, the cloud state of each ARD pixel and 36 years of data. The Landsat 4 and 5 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) ARD for a 36-year period (1 January 1982 to 31 December 2017) are considered.

The number of times that each ARD pixel was observed and the average annual number of non-fill non-cloudy observations at each pixel are derived considering each Landsat sensor independently and the three sensors together. The average annual numbers of non-fill non-cloudy observations for each ARD tile are derived. CONUS ARD tiles with particularly low and high numbers of observations are documented to identify regions that are sensitive to time series data availability issues and to support algorithm developers who wish to test algorithms using data rich and data sparse ARD tiles.

2. Landsat Data and ARD Study Data

The CONUS Landsat 4, 5 and 7 ARD were used. Landsat 4 was launched July 16, 1982 and functioned until December 14, 1993 when the satellite lost its ability to transmit data. Landsat 4 suffered several technical problems during its mission life with no systemic acquisition plan over the CONUS [13]. Landsat 5 was launched March 1, 1984 and observed the Earth until 15 January 2013, when Landsat 5's instruments were powered off and the satellite was moved into a lower orbit. The Landsat 5 TM acquired every available image over the CONUS, even during its commercialization period from 1985 to 2001 [13]. Landsat 7 was launched April 15, 1999 and continues to acquire every available image over the CONUS [5]. The Landsat 4, 5 and 7 satellites have approximately 710 km sun-synchronous circular 98.2° inclined near-polar orbits and, with 15° sensor fields-of-view, provide a 16 day repeat cycle [6]. Nominally, each Landsat satellite can sense the same earth location 22 or 23 times per year, depending on the first January overpass date [8]. The orbits are phased relative to each other to provide an 8-day full earth coverage for both Landsat 4 and 5 and for both Landsat 5 and 7 [14,15].

The Landsat ARD are discussed in detail in [2]. They are derived using the most recent Landsat Collection 1 processing algorithms applied to all Landsat images in the U.S. Landsat archive that can be corrected to ≤ 12 m Root Mean Square Error (RMSE) geodetic accuracy. The ARD are provided as top of atmosphere and atmospherically corrected products in the Albers equal area projection, in fixed non-overlapping geo-registered tiles, with spatially explicit quality assessment and cloud mask information. Each ARD tile is composed of 5000×5000 30 m pixels (i.e., covers 150×150 km). The CONUS ARD archive contains 512 tiles that are referenced by horizontal and vertical tile coordinates (Figure 1).

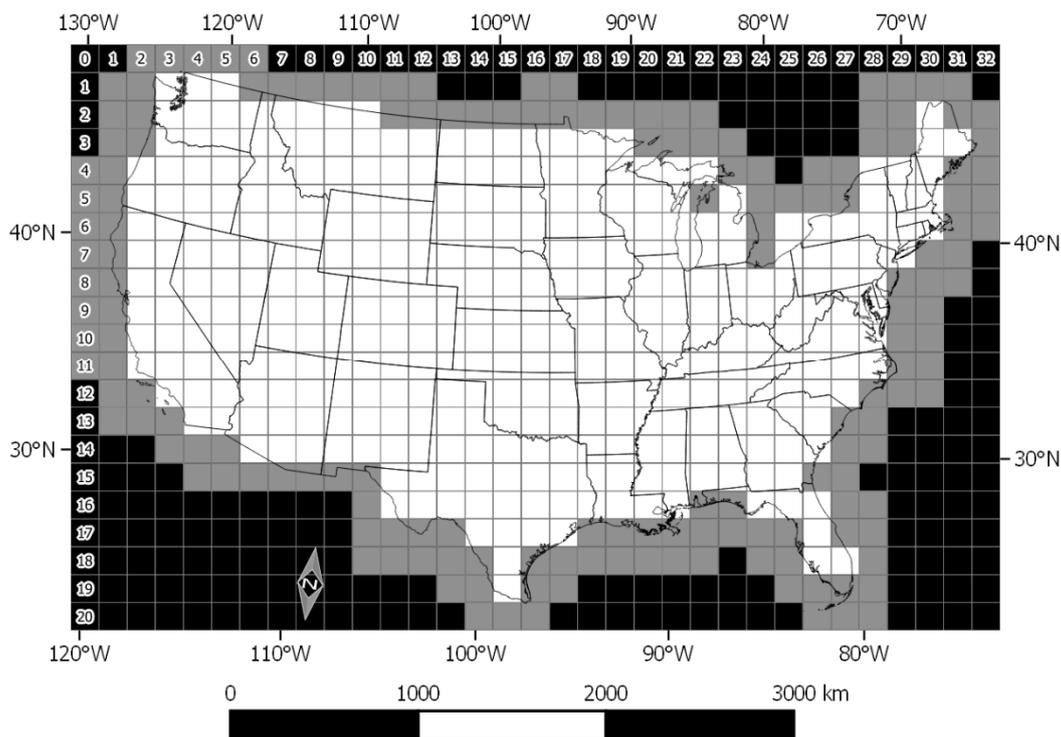


Figure 1. CONUS ARD tile locations with horizontal ($h = 0$ to 32) and vertical ($v = 0$ to 20) tile coordinates. The ARD tile boundaries are shown by grey lines and are defined by 5000×5000 30 m pixels in the Albers equal area projection (see [2] for the projection parameters); national and state boundaries are shown as black lines. Tiles where at least 50% of their coverage is within the CONUS (defined by vector data, [16]) are shown white, otherwise shown grey.

Each Landsat sensor orbit that overpasses each ARD tile is stored as a separate file [2]. In this paper, we refer to each ARD file as a granule. Each ARD granule is available as a “compressed tarball” which must be uncompressed and untarred before the bands are readable. After the tarball is uncompressed the ARD bands are stored with internal compression to save disk space. All of the available Landsat 4, 5 and 7 ARD granules over the 512 CONUS ARD tiles spanning a 36-year period from 1 January 1982 to 31 December 2017 were considered - a total of 1,027,231 (7,374, 562,576 and 457,281 Landsat 4 TM, 5 TM and 7 ETM+, respectively) granules.

3. Methods

3.1. Sensor Summary Information

The *annual number of Landsat ARD granules* over the CONUS was first derived by counting the total number of CONUS ARD granules for each year and sensor over the 36-year study period (1 January 1982 to 31 December 2017). The *seasonal number of Landsat ARD granules* over the CONUS was also derived using the standard seasonal definition adopted by the climate modeling community where winter is defined by the months December, January and February and for a given reporting year the December count data are from the previous year. Two other metrics were derived to better document the discontinuous acquisition availability over the 36-year period. The *sensor lifetime* was defined as the number days from the first date with one or more CONUS ARD granules to the last date with one or more CONUS ARD granules. As there may be no CONUS sensor acquisitions on certain days over the sensor lifetime, the *number of sensing days* was also derived as the total number of days with one or more CONUS ARD granules.

These metrics and the ones described in the following sections, were calculated for each sensor independently and considering the three sensors together. They were derived using in-house C-code and shell scripts implemented in the Linux operating system.

3.2. Pixel-Level Summary Information

Only pixels that were not labelled as “fill” in the ARD pixel quality assessment band (PIXELQA) were considered. This is because the ARD fill value denotes no Landsat data was observed [2]. At each ARD 30 m pixel location, the total number of times that the location had non-fill non-cloudy data was counted. Pixels were considered cloudy if they were labelled as medium or high confidence cloud in the ARD pixel quality assessment band. Pixels labelled as snow or water in the ARD pixel quality assessment band were also counted as in many CONUS regions snow and water may not be permanent over the year [17–19] and Landsat data are used for snow and water monitoring [20–22]. The counting was undertaken at each ARD pixel over each sensor lifetime independently - to provide at each ARD pixel location counts for Landsat 4 TM, Landsat 5 TM and Landsat 7 ETM+, denoted n_{pixel}^{L4} , n_{pixel}^{L5} , n_{pixel}^{L7} respectively – and also considering all three Landsat sensors together, to provide a combined count denoted n_{pixel}^{L457} .

At each ARD 30 m pixel location, the average annual number of non-fill non-cloudy observations was derived, for each sensor and for the three sensors combined, as:

$$\mu_{pixel}^{sensor} = \frac{n_{pixel}^{sensor}}{(\text{number of sensing days}^{sensor} / 365)} \quad (1)$$

where μ_{pixel}^{sensor} is the average annual number of non-fill non-cloudy observations at a 30 m ARD pixel location, n_{pixel}^{sensor} is the total number of times that the location had non-fill non-cloudy data (defined as described above) for $sensor \in \{L4, L5, L7, L457\}$ and the *number of sensing days* is defined in Section 3.1. The *number of sensing days* is converted to number of sensing years in the Equation (1) denominator by dividing by 365 as only a minority (9 of the 36 study period years) are leap years composed of 366 instead of 365 days.

At each ARD 30 m pixel location, the annual variation in the number of non-fill non-cloudy observations was derived. The number of non-fill non-cloudy observations in each calendar year (from January 1st to December 31st) was counted over the *sensor lifetime*. For example, 29 annual counts were derived at each ARD pixel location for Landsat 5 TM from 1984 to 2012. The annual count data were not normally distributed and at some locations for certain years and sensors there were no observations. Therefore, non-parametric statistics, namely, the 25th and 75th percentiles (lower and upper quartiles respectively) were derived from the annual counts. The interquartile range (IQR) was then derived:

$$IQR_{pixel}^{sensor} = Q3_{pixel}^{sensor} - Q1_{pixel}^{sensor} \quad (2)$$

where IQR_{pixel}^{sensor} is the interquartile range and $Q3_{pixel}^{sensor}$ is the upper quartile (75th percentile) and $Q1_{pixel}^{sensor}$ is the lower quartile (25th percentile) derived from the number of non-fill non-cloudy observations in each calendar year over the *sensor lifetime*. The IQR_{pixel}^{sensor} provides a non-parametric measure of the annual variation in the number of non-fill non-cloudy observations among years at each ARD 30 m pixel location.

3.3. ARD Tile-Level Summary Information

For each CONUS ARD tile, the average annual number of non-fill non-cloudy observations across the tile was derived, for each sensor and for the three sensors combined, as:

$$\mu_{tile}^{sensor} = \frac{\sum_{pixel=1}^N \mu_{pixel}^{sensor}}{N} \quad (3)$$

where μ_{tile}^{sensor} is the average annual number of non-fill non-cloudy observations across the tile, μ_{pixel}^{sensor} is defined as Equation (1) and N is the number of ARD tile pixels that have $n_{pixel}^{L457} > 0$. Usually $N = 25,000,000$ (as there are 5000×5000 30 m pixels in each ARD tile) but border ARD tiles that contain Landsat images acquired only over Canada or Mexico are not included in the CONUS ARD processing (for example, see Figure 1 in Reference [23]).

The ten CONUS ARD tiles with the highest and lowest μ_{tile} values were identified. For ranking purposes, only tiles within at least 50% of the CONUS land mass (shown shaded white in Figure 1) were considered as some tiles contain only a minority of land pixels, for example, coastal tiles and because some ARD tiles only partially cover the CONUS.

4. Results

4.1. Sensor Summary Information

Figure 2 illustrates the annual number of Landsat ARD granules over the CONUS for each sensor over the 36-year study period. From 1999 to 2011 there were significantly more Landsat ARD granules over the CONUS than in other years because both the Landsat 5 TM and Landsat 7 ETM+ sensors were operating in this period. The annual maximum number of ARD granules (48,511) was in 2001. The annual minimum number of granules (1,517) was in 1983 when only Landsat 4 was on orbit. In Figure 2 a small number of Landsat 4 TM granules in 1984 (52) and Landsat 5 TM granules in 2012 (70) are not apparent because of the barplot y-axis scale. The Landsat 4 TM data were temporally quite discontinuous. In 1983, Landsat 4 lost the use of two of its four solar panels and both of its direct downlink transmitters, which caused a gap in observations until 1987, when the Tracking and Data Relay Satellite System (TDRSS) became operational and was used to relay Landsat 4 TM data [6].

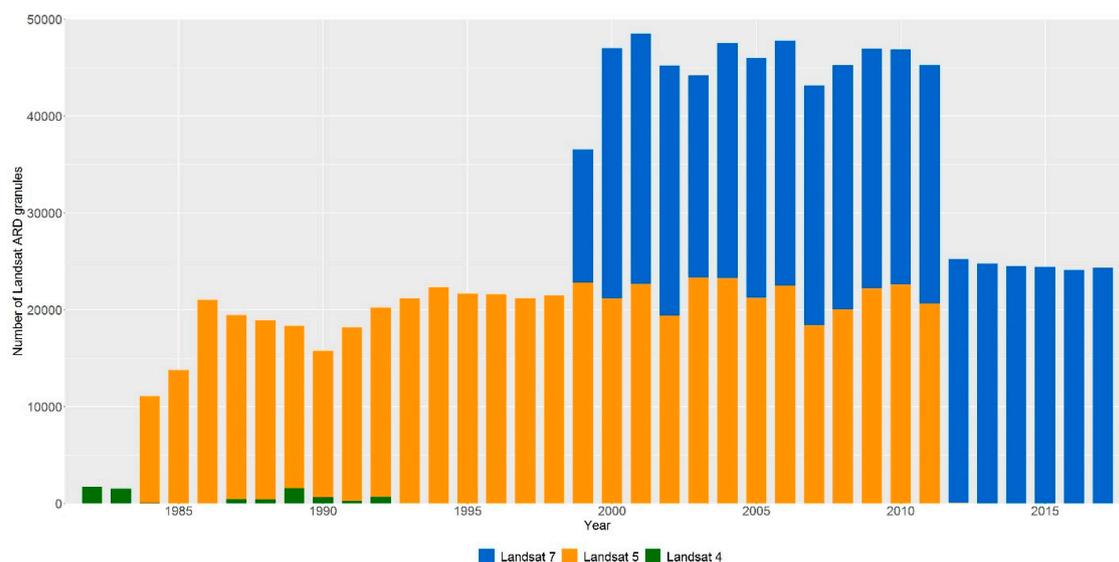


Figure 2. Annual number of Landsat ARD granules over the CONUS for each sensor over the 36-year study period (1 January 1982 to 31 December 2017). The sensor count data (colored bars) are shown stacked on top of each other in periods with overlapping sensor data.

Figure 3 illustrates the seasonal number of Landsat ARD granules over the CONUS for each sensor over the 36-year study period. More ARD granules occur typically in the summer followed by the spring, fall and winter seasons. The cause of this seasonality is not immediately obvious as, nominally, every Landsat 5 TM and Landsat 7 ETM+ overpass of the CONUS is acquired. It is due to the ARD geolocation accuracy requirement and CONUS cloud seasonality. Only Landsat images that can be geometrically corrected to a geodetic accuracy ≤ 12 m RMSE are used to generate the ARD [2]. At the time of Landsat overpass the CONUS is on average cloudier in the winter and progressively less

cloudy in the fall, spring and summer [8]. In the cloudy seasons there are a reduced relative number of ground control points used in the Landsat geolocation [24,25] and so fewer ARD granules. There are some exceptions to this seasonal pattern because of Landsat sensor and/or acquisition abnormalities. For example, the May 2003 Landsat 7 ETM+ Scan Line Corrector (SLC) failure [26] resulted in fewer Landsat acquisitions in the following 2003 summer months.

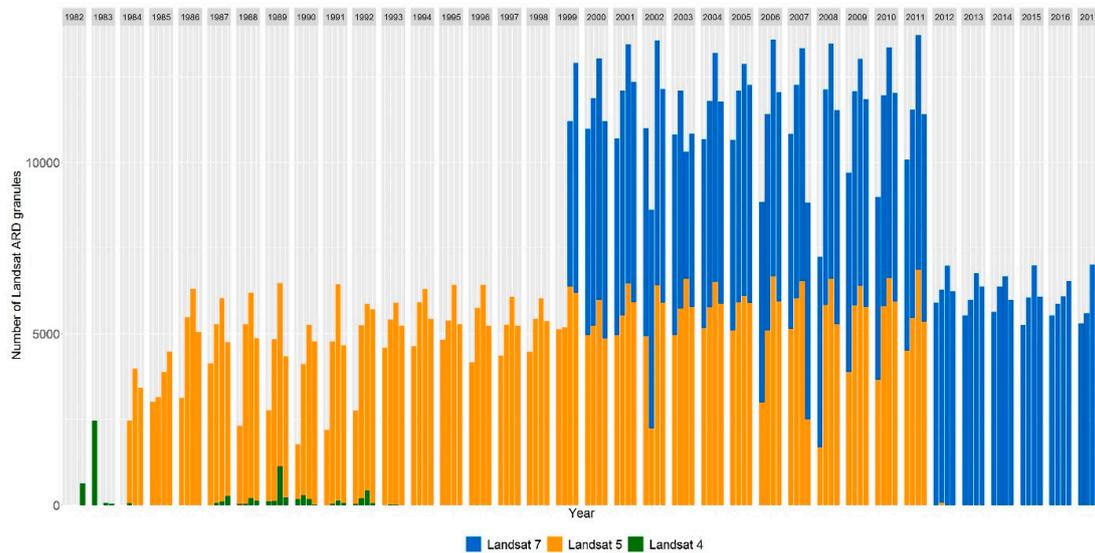


Figure 3. Seasonal number of Landsat ARD granules over the CONUS over the 36-year study period shown in order of winter (December to February), spring (March to May), summer (June to August) and autumn (September to November) for each year. The sensor count data (colored bars) are shown stacked on top of each other in periods with overlapping sensor data.

Table 1 summarizes, for each sensor, the dates of the first and last days with one or more CONUS ARD granules and the *sensor lifetime* and *number of sensing days*, over the 36-year study period. The Landsat 5 TM provided the longest operating Earth remote sensing satellite mission in history [27] with a 28.156 year *sensor lifetime*. In Table 1 Landsat 7 ETM+ had an 18.518 year *sensor lifetime* but continued to sense beyond the end of the 31 December 2017 study period. Landsat 4 TM had a 10.685 *sensor life time* but as noted above, was not always operational. The discontinuous acquisition and availability of Landsat data over the CONUS, particularly for Landsat 4 TM, is evident by the difference between the *sensor lifetime* and the *number of sensing days*. In particular, the Landsat 4 TM sensor was in orbit for 3,900 days but acquired CONUS data to make ARD for only 387 days, that is, for only 9.92% of the *sensor lifetime*. In contrast, the Landsat 5 TM and 7 ETM+ sensors acquired data to make ARD for 93.17% and 98.15% of their *sensor lifetime* values respectively. Over the 36 year study period there were 12,191 days (>33 years) when there was at least one sensor acquiring CONUS data that were used to make ARD.

Table 1. CONUS ARD sensor summary information over the 36 year study period (1 January 1982 to 31 December 2017). The *sensor lifetime* and *number of sensing days* are shown with units of days and years.

Sensor	First CONUS ARD Granule Acquisition Date	Last CONUS ARD Granule Acquisition Date	Sensor Lifetime days years	Number of Sensing days years
Landsat 4	1982/11/11	1993/07/16	3,900 10.685	387 1.060
Landsat 5	1984/03/16	2012/05/05	10,277 28.156	9,575 26.233
Landsat 7	1999/06/30	2017/12/31	6,759 18.518	6,634 18.175
Landsat 4, 5 and 7	1982/11/11	2017/12/31	12,834 35.162	12,191 33.400

4.2. Pixel-Level Summary Information

Figure 4 shows, for the three sensors together, the total number of (a) non-fill observations, (b) non-fill and non-cloudy observations and (c) non-fill, non-cloudy and non-shadow observations, at each CONUS ARD 30 m pixel location over the 36 year study period. Many of the CONUS border ARD tiles contain pixels with $n_{pixel}^{L457} = 0$ (shown as black). Discarding pixel locations with $n_{pixel}^{L457} = 0$, the mean CONUS total number of observations over the 36 year study period derived from Figure 4a–c is 936.2, 658.4 and 628.0 respectively, demonstrating the effect of cloud and shadow in reducing the number of surface observations.

Across the CONUS there is a regional variation in the number of observations. Notably, without consideration of the cloud or shadow status (Figure 4a) there are fewer observations in the north-east, north-west and also in certain ARD tiles that are predominantly over water. This is due to the ARD geolocation accuracy requirement – only Landsat images that can be geometrically corrected to a geodetic accuracy ≤ 12 m RMSE are used to generate the ARD [2]. Consequently, in coastal and cloudy regions, which typically have a reduced relative number of ground control points used in the Landsat geolocation [24,25], there are fewer ARD granules.

Figure 4b shows the total number of non-fill non-cloudy observations, that is, it shows the same information as Figure 4a but without the cloudy observations. There are fewer observations due to cloud cover at the time of Landsat overpass, which previous CONUS Landsat cloud studies have shown predominates in the north-east and north-west, with more observations in the arid south-west [4,8,12]. The ARD cloud mask is imperfect and even at national scale, some cloud detection issues are evident. For example, the extensive highly reflective sands of White Sands, New Mexico, have a reduced number of observations (blue spot near the center of ARD tile h10v14) which is a previously documented Landsat cloud mask commission error [12].

Figure 4c shows the total number of non-fill, non-cloudy and non-shadow observations, that is, it shows the same information as Figure 4b but without the shadow observations. There are fewer observations due to shadows across the CONUS. This is particularly evident over mountainous regions, for example, over the Rocky mountains, from Colorado to Idaho, where shadows due to relief and not just clouds are common. Pixels were considered shadow contaminated if they were labelled as Cloud Shadow in the ARD PIXELQA band [2]. However, at local scale, the ARD shadow mask can be quite unreliable. For example, we have found examples where it fails to mask shadows or falsely labels non-shadow observations as shadow. Efforts by other researchers are underway to improve the shadow mask [28]. Therefore, in the remainder of the paper we do not consider the shadow state as it is not reliably labelled in the current ARD version.

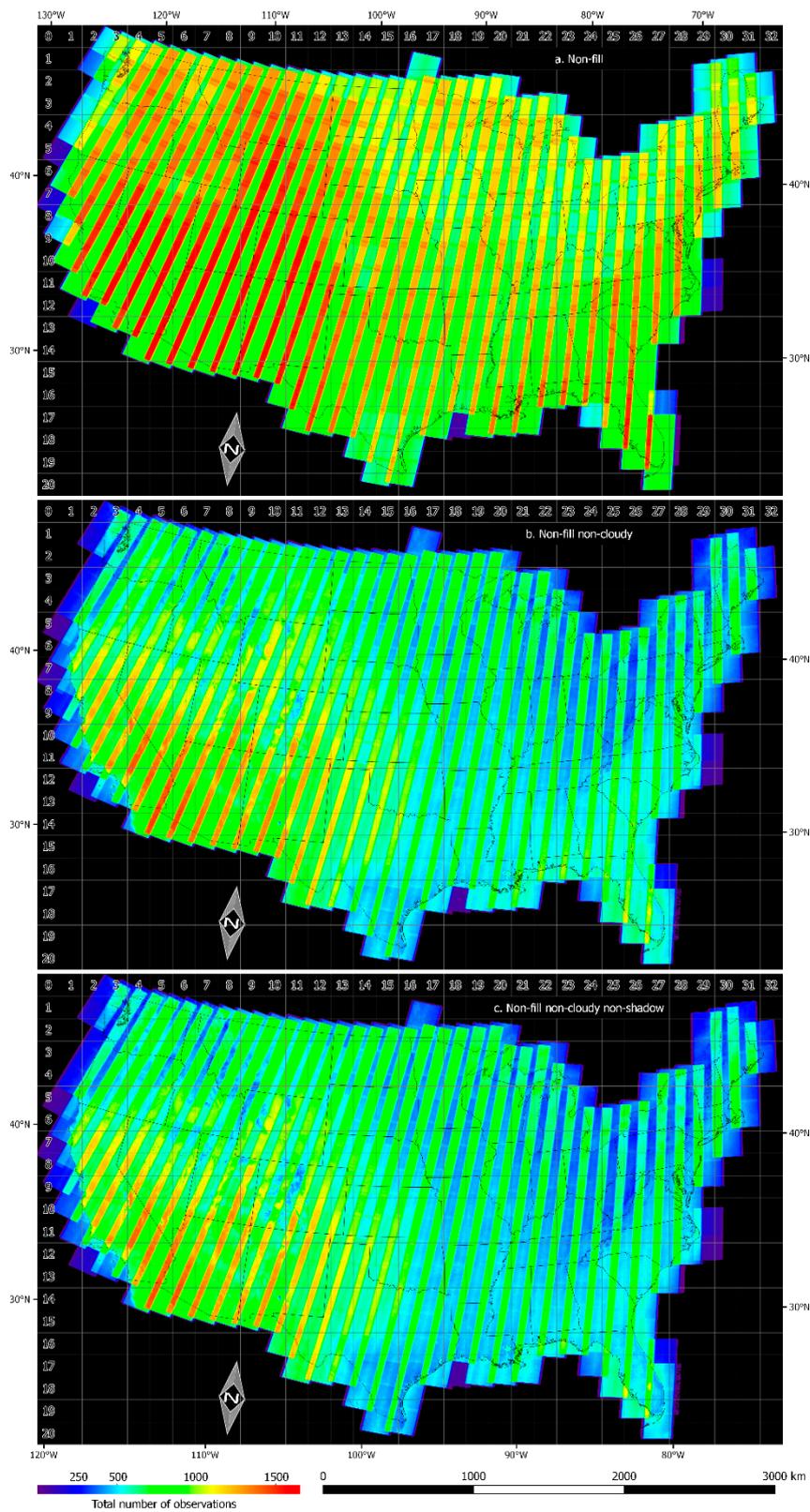


Figure 4. Total number (see color scale) of (a) non-fill, (b) non-fill non-cloudy and (c) non-fill non-cloudy non-shadow Landsat 4 TM, 5 TM and 7 ETM+ observations at each CONUS ARD 30 m pixel location over the 36 year study period.

The approximately north-south oriented stripes in Figure 4 occur because spatially adjacent Landsat orbits overlap in the across-scan direction, with a greater number of observations (red tones) evident in the overlap regions. The overlap regions are wider in the north than in the south because the Landsat orbit paths converge further poleward [4]. The ARD tiles are defined in the Albers equal area projection. Consequently, the change in the relative orientation of the ARD tile boundaries to the fixed Landsat orbit inclination is evident, with an approximately top to bottom relative orientation in the eastern CONUS and a more diagonal relative orientation in the western CONUS. As a result, many tiles in the western CONUS contain a larger number of overlapping Landsat observations than in the eastern CONUS.

The Landsat orbits drift slightly at annual scale, for example, considering three years of global Landsat 5 TM and Landsat 7 ETM+ data the mean drift was several kilometers east-west and less than a kilometer north-south direction [4]. Over the satellite lifetime, the Landsat orbits are maintained by station keeping maneuvers (orbit burns) but for Landsat 5, the orbit was allowed to drift considerably when the satellite was operated commercially from 1985 to 2001 [13,27]. The orbits drift spatially, which blurs the number of observations over long time periods and this is apparent when the Figure 4 data are examined in detail. As an example, Figure 5 shows a spatial subset of Figure 4a for six ARD tiles.

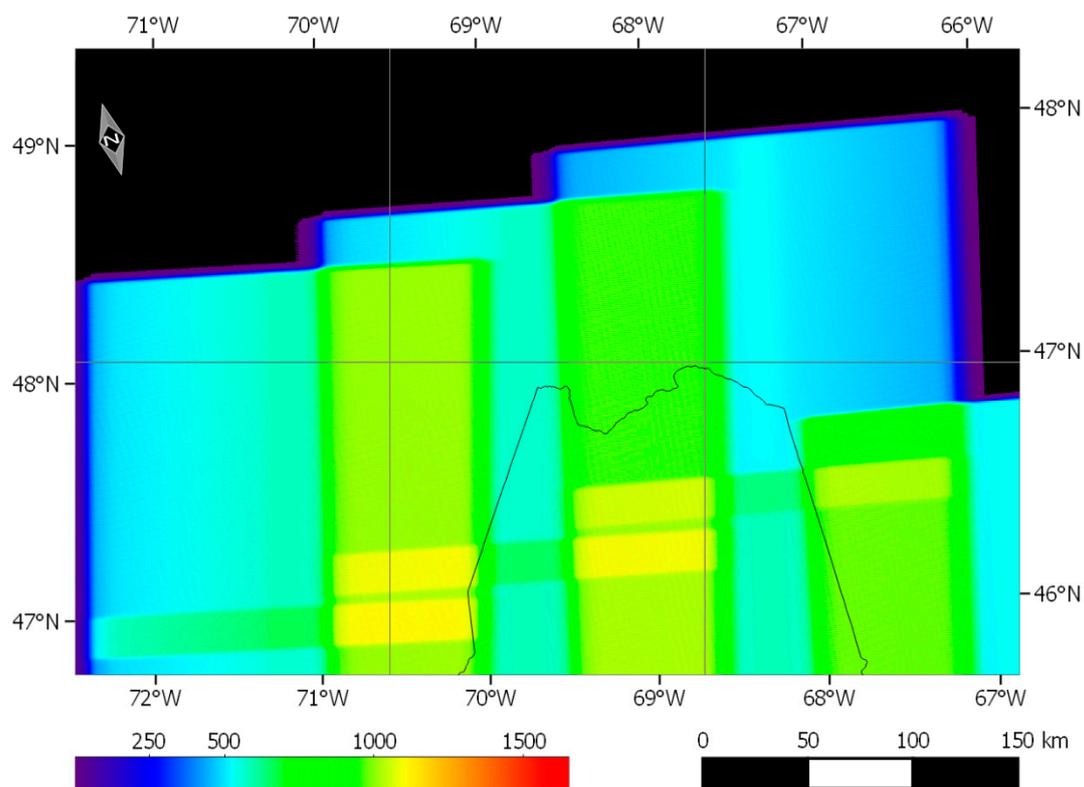


Figure 5. Example detailed spatial subset of Figure 4a over the northern tip of Maine (state boundary in black) showing the total number of non-fill Landsat 4 TM, 5 TM and 7 ETM+ observations over the 36 year study period for six ARD tiles (h29 to h31, v01 to v02).

Figure 6 shows the average annual numbers of non-cloudy and non-fill observations (μ_{pixel}^{sensor} , Equation (1)) at each 30 m ARD pixel location. The same geographic patterns as Figure 4b are apparent, including the stripes of greater observations where spatially adjacent Landsat orbits overlap. As described above, nominally, there are a maximum of 22 or 23 Landsat overpasses per year per sensor and much of the less cloudy CONUS regions have μ_{pixel}^{L5} values and μ_{pixel}^{L7} similar to this maximum. The μ_{pixel}^{L7} values are often smaller than the μ_{pixel}^{L5} values, particularly in the across-scan orbit overlap regions. This is because of the May 2003 Landsat 7 ETM+ Scan Line Corrector (SLC) failure that

resulted in a 22% data loss occurring in a wedge-shaped pattern increasing in width at scan angles further from nadir [26]. The μ_{pixel}^{L4} have the smallest values (Figure 6d) and even with 387 sensing days (Table 1), full CONUS Landsat 4 TM coverage did not occur. In Figure 6d, there are regions with no observations and there are ARD pixels with significantly reduced μ_{pixel}^{L4} values due to clouds. The CONUS-wide mean μ_{pixel} values, considering only pixels with $n_{pixel}^{L457} > 0$ and pixels in ARD tiles covering at least 50% of the CONUS, are 4.85 (Landsat 4 TM), 16.41 (Landsat 5 TM), 15.03 (Landsat 7 ETM+) and 21.22 (all three Landsat sensors).

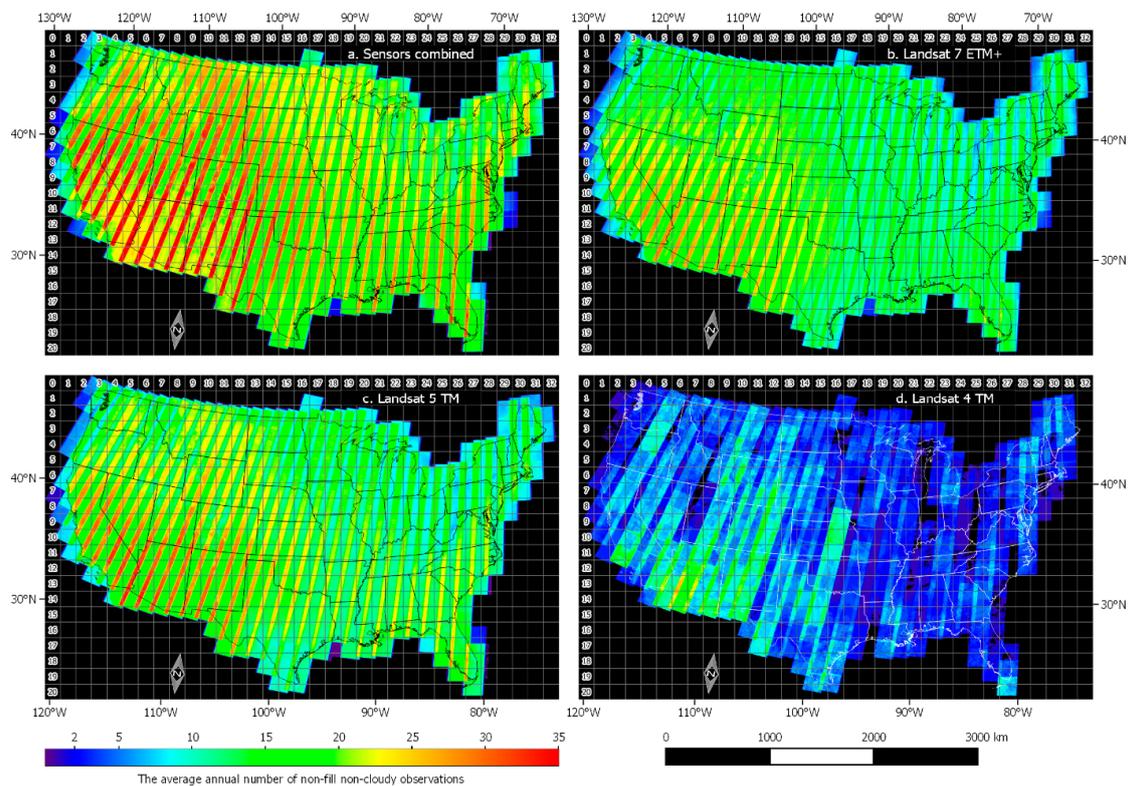


Figure 6. The average annual number (μ_{pixel}^{sensor} Equation (1)) of non-fill non-cloudy observations at each CONUS ARD 30 m pixel location over the 36 year study period for (a) Landsat 4 TM, 5 TM and 7 ETM+ combined, (b) Landsat 7 ETM+, (c) Landsat 5 TM, (d) Landsat 4 TM.

Figure 7 shows the interquartile range (IQR_{pixel}^{sensor} , Equation (2)) of the annual number of non-fill non-cloudy observations at each CONUS ARD 30 m pixel location. It provides a non-parametric measure of the annual variation in the number of non-fill non-cloudy observations among years. It is zero valued (shown grey) if the lower and upper quartiles are the same, for example, if the annual number of non-fill non-cloudy observations is the same in every year. The combined sensor values (IQR^{L457} , Figure 7a) are generally greater, that is, there is more variation among years, than for the individual sensors. This is expected because 23 of the 36 years had simultaneously two sensors acquiring data (Figure 2) and this results in a greater annual variation over the 36 years than over the lifetimes of the individual sensors. The IQR^{L5} values (Figure 7c) are generally greater than the IQR^{L7} values (Figure 7b). This is because the Landsat 5 TM acquisition strategy was less consistently maintained, particularly in the commercial operational period from 1985 to 2001 [13,27]. The IQR^{L4} values (Figure 7d) are generally lower than for the other sensors because the annual number of Landsat 4 TM observations are low (Figure 6) and because for several years there were no or few Landsat 4 TM ARD granules (Figure 2). The stripes of high interquartile range values evident across the CONUS are expected as they occur where the adjacent orbits overlap (Figures 4–6) resulting in a greater absolute variation in the number of observations among years.

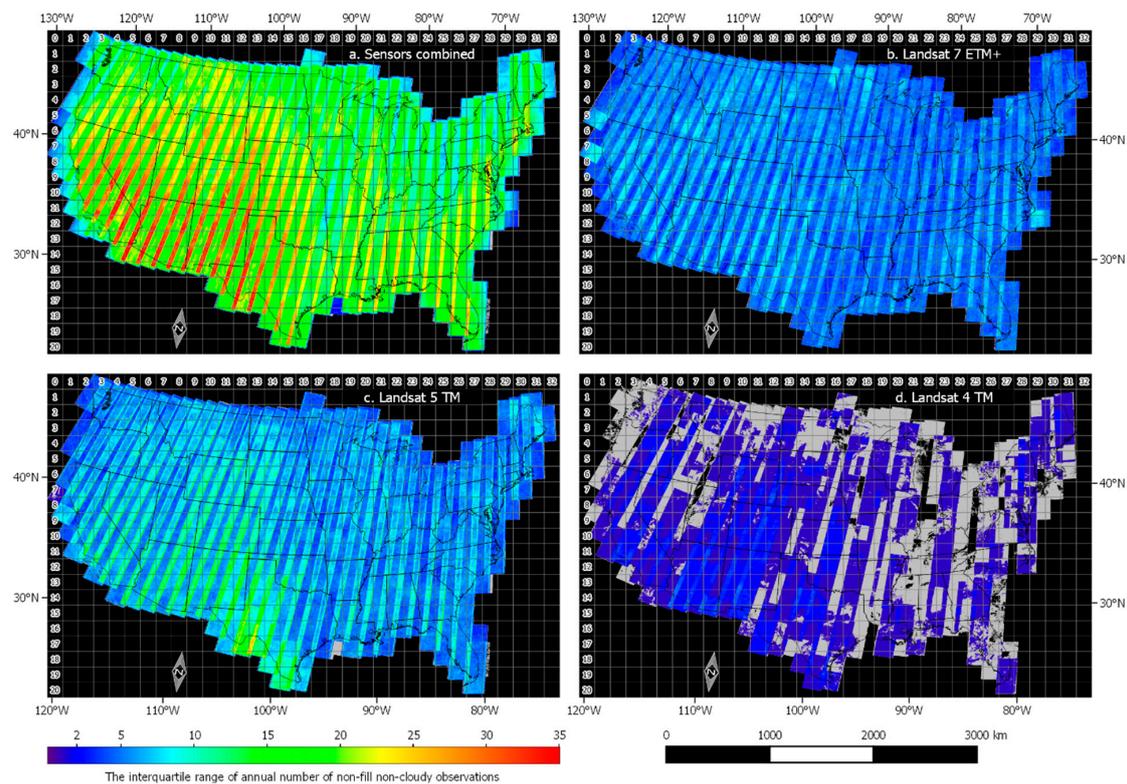


Figure 7. The interquartile range (IQR_{pixel}^{sensor} Equation (2)) of the annual number of non-fill non-cloudy observations at each CONUS ARD 30 m pixel location for (a) Landsat 4 TM, 5 TM and 7 ETM+ combined, (b) Landsat 7 ETM+, (c) Landsat 5 TM, (d) Landsat 4 TM. The maximum CONUS ARD IQR_{pixel}^{sensor} values are 55 (a), 27 (b), 31 (c) and 6 (d). Light grey show locations where there were observations over the sensor lifetime but where the interquartile range is zero, for example, where the annual number of non-fill non-cloudy observations was the same among the years.

Figure 8 shows histograms of the temporal difference between consecutive observations of a single ARD pixel location. The results are shown for the most observed CONUS 30 m ARD pixel, that is, the pixel location with the greatest number of non-fill and non-cloudy Landsat 4 TM, 5 TM and 7 ETM+ observations over the 36 year study period. The most observed CONUS ARD pixel location is in a region of southern California where adjacent Landsat orbits overlap at $32^{\circ}47'34.09''$ N, $114^{\circ}55'47.60''$ W (pixel row 3656 and column 2154 in ARD tile h05v13). At this pixel location there were 24, 940 and 570 (total 1534), non-fill non-cloudy Landsat 4 TM, 5 TM and 7 ETM+ observations over the 36 years, respectively. The results illustrate that when considering each sensor independently (Figure 8b–d), the most common temporal differences between consecutive pixel observations are 7, 9 and 16 days. The 7 and 9 day differences reflect how adjacent orbits laterally overlap at the Landsat swath edges and the 16 day difference is due to the 16 day Landsat repeat cycle. Other temporal differences are due to cloud obscuration between consecutive observations and periods of missing observations due to sensor and/or acquisition abnormalities. The maximum consecutive temporal differences (not plotted in Figure 8) are 57, 176 and 1975 for the Landsat 7 ETM+, 5 TM and 4 TM observations respectively. Similar results are evident considering the three sensors combined (Figure 8a) but with a greater variety of differences and notably one and eight day differences that can occur between successive Landsat 7 ETM+ and 5 TM observations [29,30].

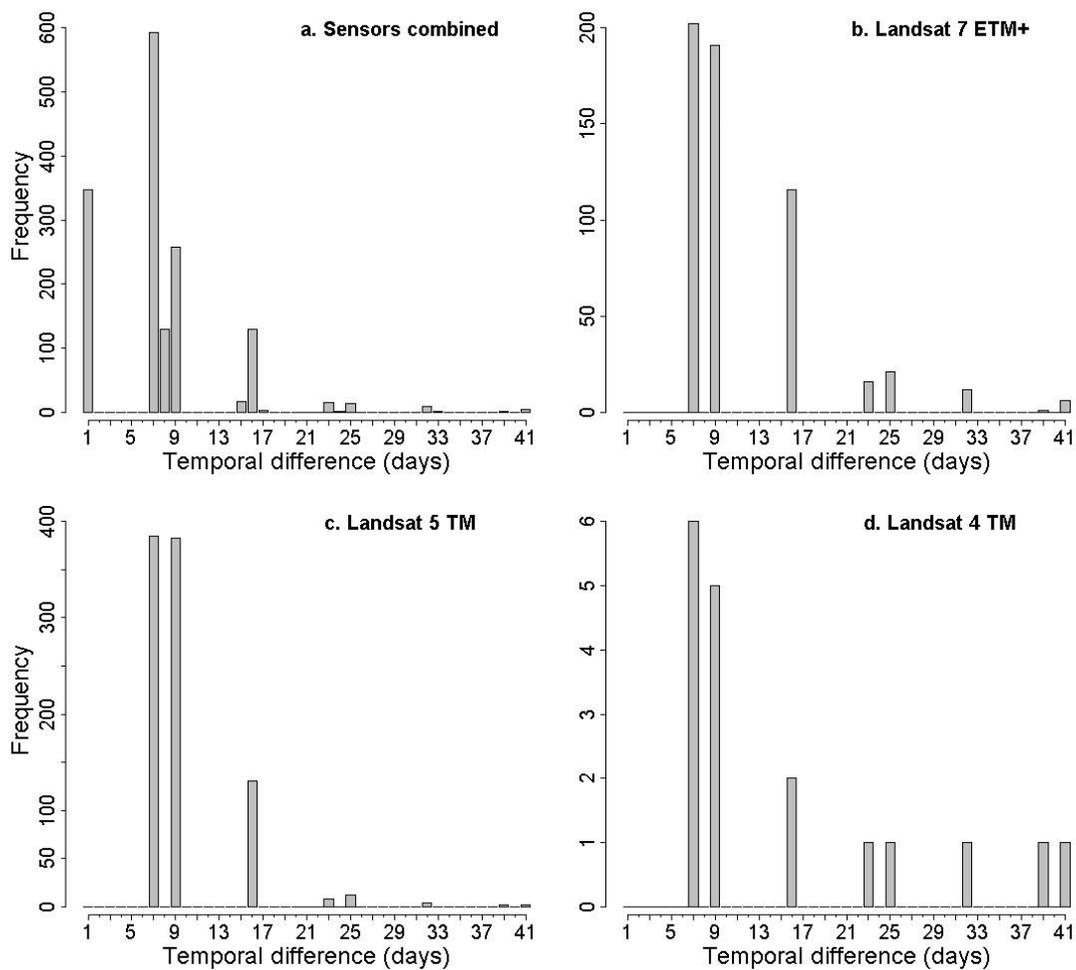


Figure 8. Histogram of the temporal differences between consecutive non-fill non-cloudy Landsat observations over the 36 year study period at the most observed non-fill non-cloudy CONUS ARD pixel location (located at $32^{\circ}47'34.09''$ N $114^{\circ}55'47.60''$ W, southern California, ARD tile h05v13) for (a) Landsat 4 TM, 5 TM and 7 ETM+ combined, (b) Landsat 7 ETM+, (c) Landsat 5 TM, (d) Landsat 4 TM.

The results in Figure 8 are for the single most observed non-fill non-cloudy CONUS ARD pixel. At other ARD pixel locations greater temporal differences between consecutive non-cloudy Landsat observations occur more frequently. These differences are not straightforward to summarize at every CONUS ARD pixel because they vary considerably. In certain years (Figure 2) and in the non-summer months (Figure 3), there may be relatively fewer observations of an ARD pixel and at some pixel locations there may be no or only a small number of non-fill non-cloudy observations over the sensor lifetime (Figure 6).

4.3. ARD Tile-Level Summary Information

Figure 9 shows the per tile annual mean number of non-fill non-cloudy observations (μ_{tile}^{sensor} , Equation (3)). Considering each sensor separately and all three sensors combined, reveals similar patterns. The most frequently observed tiles are in the south-west, particularly over the Mojave and Sonora deserts and also in mountain rain shadow regions of the Cascade Range (Columbia basin and High Desert), Sierra Nevada Range (Great Basin Desert), Coast Range (Central Valley) and the Rocky Mountains (Great Plains). The least observed tiles tend to be in the north-east, around the Great Lakes and along parts of the north-west coast. As noted above, the coastal ARD tiles have low μ_{tile}^{sensor} values as there are fewer ARD granules for the geolocation processing reasons discussed at the beginning of Section 4.2.

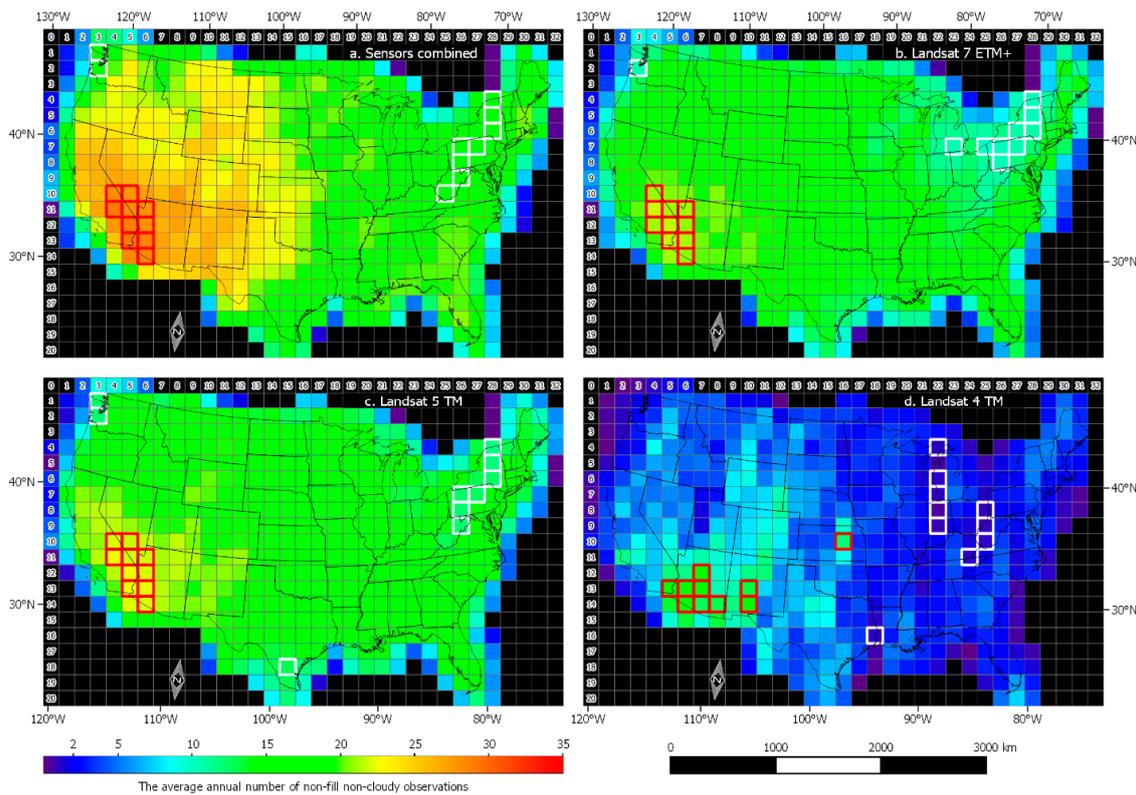


Figure 9. The average annual number (μ_{tile}^{sensor} , Equation (3)) of non-fill non-cloudy observations across each CONUS ARD tile over the 36 year study period for (a) Landsat 4 TM, 5 TM and 7 ETM+ combined, (b) Landsat 7 ETM+, (c) Landsat 5 TM, (d) Landsat 4 TM. The ten *most* (red outlined) and the ten *least* (white outlined) observed CONUS ARD tiles (white tiles in Figure 1) are shown and are summarized in Tables 2 and 3.

Table 2. The ten *most* observed CONUS cloud-free ARD tiles over the 36 year study period (red outlined in Figure 9).

Landsat 4,5,7		Landsat 7		Landsat 5		Landsat 4	
Tile	μ_{tile}	Tile	μ_{tile}	Tile	μ_{tile}	Tile	μ_{tile}
h05v13	30.067	h04v11	21.718	h04v11	22.828	h06v13	16.794
h04v11	29.974	h05v13	21.621	h05v13	22.727	h06v14	16.419
h05v12	29.876	h05v12	21.522	h05v12	22.669	h07v13	14.637
h06v13	29.740	h06v13	21.001	h06v13	22.635	h07v12	14.413
h06v14	29.217	h04v12	20.937	h04v10	22.303	h05v13	14.211
h04v10	29.024	h05v11	20.900	h06v14	22.280	h10v14	12.974
h05v11	28.982	h06v11	20.761	h06v12	22.064	h10v13	12.292
h06v12	28.941	h04v10	20.746	h05v11	22.035	h08v14	12.128
h06v11	28.749	h06v12	20.695	h06v11	21.855	h07v14	11.790
h05v10	28.512	h06v14	20.576	h05v10	21.804	h16v10	11.441

Table 3. The ten *least* observed CONUS cloud-free ARD tiles over the 36 year study period (white outlined in Figure 9).

Landsat 4,5,7		Landsat 7		Landsat 5		Landsat 4	
Tile	μ_{tile}	Tile	μ_{tile}	Tile	μ_{tile}	Tile	μ_{tile}
h28v04	14.231	h25v07	9.728	h28v04	11.077	h22v08	0.526
h28v05	14.273	h28v05	9.792	h28v05	11.177	h22v09	0.789
h03v02	14.546	h28v04	9.905	h03v02	11.453	h25v09	0.840
h26v07	14.582	h26v07	9.916	h26v09	11.535	h25v08	0.859
h26v08	14.589	h26v08	10.031	h26v08	11.547	h18v16	0.995
h26v09	14.864	h03v02	10.069	h26v07	11.587	h22v07	1.019
h27v07	14.905	h23v07	10.105	h03v01	11.652	h22v04	1.056
h03v01	14.946	h27v07	10.217	h28v06	11.726	h25v10	1.076
h28v06	14.957	h28v06	10.262	h27v07	11.789	h24v11	1.087
h25v10	15.144	h27v06	10.293	h15v18	11.983	h22v06	1.104

In Figure 9 the pixel-level Landsat orbit swath overlap patterns (Figures 4 and 6) are less evident because of the tile level averaging. However, the change in the relative orientation of the ARD tile locations to the Landsat orbits is still apparent, with more tiles in the western CONUS having greater μ_{tile}^{sensor} values than in the eastern CONUS due to this phenomenon and also due to reduced cloudiness.

Figure 10 shows histograms of the Figure 9 data. The histograms are approximately normally distributed, except for the Landsat 4 TM histogram that is positively skewed (i.e., a greater frequency of small values). Across the CONUS, the most common μ_{tile} values are 18 to 20 (23.4% of tiles) for all three sensors, 12 to 14 (30.7% of tiles) for the Landsat 7 ETM+, 14 to 16 (26.2% of tiles) for the Landsat 5 TM and 2 to 4 (36.5% of tiles) for the Landsat 4 TM.

The ten CONUS ARD tiles with the highest μ_{tile} values are summarized in Table 2 and highlighted in Figure 9. The tiles are clustered over the south-west and with the exception of Landsat 4 TM, have $\mu_{tile} > 20$. Considering Landsat 7 ETM+, Landsat 5 TM and the three sensors combined, eight of the ten most frequency observed ARD tiles are common. Southern California ARD tiles h05v13 (centered 33°06′56.45″ N, 114°53′38.52″ W) and h04v11 (centered 35°27′32.82″ N, 117°11′13.44″ W) have the highest μ_{tile} values for these sensors. The top ten Landsat 4 TM ARD tiles have a different geographic distribution compared to the other sensors, due to the spatially irregular number of observations (Figure 6). Southern Arizona ARD tile h06v13 (centered 33°22′12.09″ N, 113°18′11.39″ W) has the highest μ_{tile} value for the Landsat 4 TM.

The ten CONUS ARD tiles with the lowest μ_{tile} values are summarized in Table 3 and highlighted in Figure 9. They are located predominantly in north-west coastal regions and in tiles surrounding the Great Lakes. For all three sensors and also Landsat 5 TM, the least observed tile is h28v04 that encompasses northern New York state and parts of southern Canada (centered 44°58′01.15″ N, 74°06′31.86″ W). For Landsat 7 ETM+, the least observed tile is h25v07 that is over the state boundary of Ohio and Pennsylvania (centered 41°47′42.71″ N, 80°40′9.54″ W). Notably, these least observed tiles have approximately half the μ_{tile} values compared to the most observed tile μ_{tile} values for these sensors (Table 2). The least observed Landsat 4 TM ARD tile is h22v08 that is over northern Indiana (centered 40°59′40.72″ N, 86°17′1.36″ W) and has a very low μ_{tile} (0.526) that is, in this tile ARD pixels have on average less than one non-fill non-cloudy observation per year.

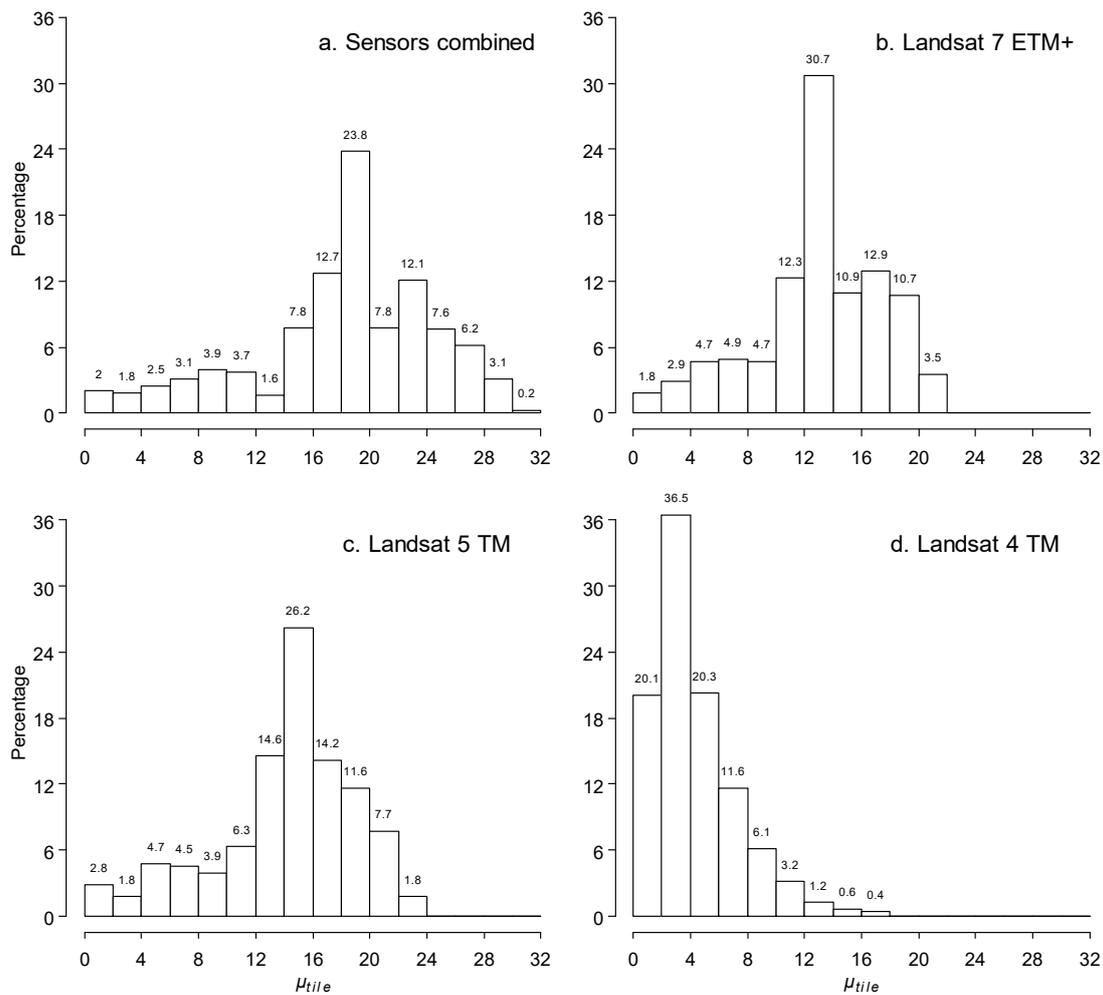


Figure 10. CONUS histograms of the average annual number (μ_{tile} , Equation (3)) of non-fill non-cloudy observations across each ARD tile for the 36 year study period, for (a) Landsat 4 TM, 5 TM and 7 ETM+ combined, (b) Landsat 7 ETM+, (c) Landsat 5 TM, (d) Landsat 4 TM. The histogram bins are defined as $<2, 2 - <4, 4 - <6, \dots, 30 - <32$. All results derived from the Figure 9 data.

5. Discussion

It is well established that the availability of satellite observations influences terrestrial monitoring capabilities. More observations in a given time period mean that users have more opportunities to select single cloud free images over an area and time period of interest. More observations also enable more reliable time series fitting [31–33] and improved change detection through more precise definition of the timing of the change and characterization of the change type [34,35]. More observations also mean that temporally composited products have a higher likelihood of being generated with cloud-free data and that shorter temporally compositing periods can be used which is useful if the surface phenomena are changing rapidly [36–38].

The results reported in this paper illustrate that the Landsat 4 TM, 5 TM and 7 ETM+ ARD observation record has different availability across the CONUS, at annual and seasonal scales and that the availability can vary spatially quite significantly. For example, for Landsat 5 TM the μ_{tile} values vary from 11.08 (tile h28v04, northern New York state) to 22.83 (h04v11, Southern California). This uneven distribution poses challenges for CONUS-wide terrestrial applications that use the historical Landsat ARD. Complex patterns of data availability occur across the CONUS due to different sensor data acquisition frequency, cloud, the satellite orbit and sensor geometry and because of the relative orientations of the ARD tiles with respect to the Landsat orbit paths. This may be a concern for

certain applications. For example, Landsat ARD analyses results generated over the arid south-west, where there are the most cloud-free ARD observations, will likely be more reliable than results in the north-east, around the Great Lakes and along parts of the north-west coast where significantly fewer observations are available. It is recommended that the ARD tiles with particularly high and low numbers of observations (described in Tables 2 and 3) be considered by algorithm developers who wish to evaluate algorithms at data rich and data sparse ARD tile locations. The quality of time series algorithm results is expected to be reduced at tiles with low reported temporal ARD availability.

The results reported in this paper also illustrate the utility of combining different Landsat sensor data together to take advantage of their different acquisition patterns to improve temporal data availability. The CONUS average annual number of non-cloudy observations at each 30 m ARD pixel location is 4.85 (Landsat 4 TM), 16.41 (Landsat 5 TM), 15.03 (Landsat 7 ETM+) and 21.22 (all three Landsat sensors). The greatest temporal availability of ARD pixel observations occurs in cloud-free regions where spatially adjacent Landsat orbits overlap across-scan. These locations offer potential for new Landsat-scale remote sensing studies where higher temporal resolution observations are needed. For example, Landsat reflectance bi-directional effects have been documented by taking advantage of the one day overpass difference between Landsat 5 TM and Landsat 7 ETM+ at orbit overlap locations [29].

For systematic monitoring, satellite observations should ideally occur uniformly in time with at least twice the observation frequency of the surface variations being monitored [33]. This is not the case for Landsat data. For example, at the most observed CONUS 30 m ARD pixel there is considerable variety in the temporal difference between consecutive pixel observations (Figure 8). At other CONUS ARD pixel locations greater temporal differences between consecutive non-cloudy Landsat observations occur more frequently. Future research on the temporal difference between consecutive pixel observations in the ARD record is recommended, including the minimum temporal difference as near coincident observations can be used for among-sensor comparison and characterization activities [39]. This is beyond the scope of the present study however.

The reported results likely overestimate the number of “good” surface observations because shadows and cirrus clouds were not considered. Clouds and shadows reduced the CONUS mean total number of Landsat 4 TM, 5 TM and 7 ETM+ observations for the 36 years by 32.9% (comparing Figure 4a with Figure 4c, whereas clouds reduced the number by 29.7% (comparing Figure 4a with Figure 4b). However, at local scale the ARD shadow mask is not reliable and consequently it was not used in this study except to generate Figure 4c. Reliable shadow and cloud masking are difficult to implement over all surface and atmospheric conditions and the ARD cloud-mask and shadow-mask will be improved in future ARD versions [2,28,40]. Cirrus clouds were also not considered in this study as there is no Landsat 4 TM, 5 TM and 7 ETM+ cirrus mask. A recent analysis of one year of Landsat 8 Operational Land Imager (OLI) imagery, that includes a new Landsat band dedicated to cirrus cloud detection, indicated that the historical CONUS Landsat archive is about 7% cirrus cloud contaminated [12].

The reported analysis was straightforward to undertake computationally because the ARD are stored in fixed geolocated ARD tiles. The ARD filenames include the tile coordinates and so are amenable to scripting. In addition, computationally expensive image reprojection and resampling needed to align Landsat images [25] was not required because this was already undertaken in the ARD generation. The majority of the computational overhead was on reading the ARD. The CONUS ARD volume is considerable. In this study, a total of 1,027,231 Landsat 4 TM, 5 TM and 7 ETM+ ARD granules were considered that required 119 TB of dedicated disk storage after tarball decompression.

Research to assess and consider more sophisticated processing using the ARD is ongoing, including correction for bi-directional reflectance and topographic effects, gap-filling and improved cloud masking [23,28,41,42]. A recommendation of this study is to improve the Landsat geolocation processing. The ARD are generated using only Landsat images corrected to a geodetic accuracy ≤ 12 m RMSE [2]. More images in the U.S. Landsat archive could be processed as ARD if the Landsat

geometric processing was improved, for example, by using an orbit based methodology [43,44] rather than one that relies on the collection of ground control from individual images [24].

6. Conclusions

The release of the Landsat ARD reduces the magnitude of pre-processing required to produce large-area Landsat derived data products or to perform long term terrestrial monitoring [2]. In this study, the availability of the CONUS Landsat 4 TM, 5 TM and 7 ETM+ ARD was examined for each sensor separately and combined for a 36 year period (1 January 1982 to 31 December 2017). This is important as the availability of satellite observations influences surface monitoring capabilities. The main quantitative paper findings are summarized below.

Over the 36 years there were 12,191 days (>33 years) on which at least one Landsat sensor (Landsat 4 TM, 5 TM, 7 ETM+) acquired imagery that was used to make CONUS ARD. Examination of the number of days with CONUS ARD revealed that Landsat 4 TM, Landsat 5 TM and Landsat 7 ETM+ acquired images used to make the CONUS ARD for 9.92%, 93.17% and 98.15% of their sensor lifetimes in the 36 year period respectively.

The single ARD 30 m pixel location with the greatest number of non-cloudy Landsat 4 TM, 5 TM and 7 ETM+ observations is in southern California ARD tile h05v13 (column 5, row 13; specifically, 32°47′34.09″ N 114°55′47.60″ W). The CONUS average annual number of non-cloudy observations at each 30 m ARD pixel location is 4.85 (Landsat 4 TM), 16.41 (Landsat 5 TM), 15.03 (Landsat 7 ETM+) and 21.22 (all three Landsat sensors).

Considering the average annual number of non-cloudy observations in each 150 × 150 km ARD tile (μ_{tile}), the most common CONUS μ_{tile} values are 2 to < 4 (Landsat 4 TM), 14 to < 16 (Landsat 5 TM), 12 to < 14 (Landsat 7 ETM+) and 18 to < 20 (all three Landsat sensors). The ten CONUS ARD tiles with the highest μ_{tile} values are clustered over the south-west; and the highest values are at ARD tiles h04v11 (Landsat 5 TM and 7 ETM+), h06v13 (Landsat 4 TM) and h05v13 (all three sensors). The ten CONUS ARD tiles with the lowest μ_{tile} values are located predominantly in north-west coastal regions and in tiles surrounding the Great Lakes; the ARD tiles with the lowest values are h28v04 (Landsat 5 TM and all three sensors), h25v07 (Landsat 7 ETM+), and h22v08 (Landsat 4 TM).

Author Contributions: D.P.R., A.V.E. and H.K.Z., conceived the experiments; A.V.E. processed the Landsat ARD data with assistance from H.K.Z. and Z.L.; A.V.E. generated the graphics assisted by H.K.Z. and D.P.R.; D.P.R. structured and drafted the manuscript with assistance from A.V.E., H.K.Z., Z.L., L.Y. and H.H.

Funding: This research was funded by the NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) program (cooperative agreement NNX13AJ24A) and by the U.S. Geological Survey Landsat science team (grant G12PC00069).

Acknowledgments: The USGS Landsat program management and staff, are thanked for the free provision of the Landsat ARD.

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

References

1. Wulder, M.A.; Loveland, T.R.; Roy, D.P.; Crawford, C.; 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**. In Press.
2. Dwyer, J.L.; Roy, D.P.; Sauer, B.; Jenkerson, C.B.; Zhang, H.K.; Lyburner, L. Analysis Ready Data: Enabling analysis of the Landsat archive. *Remote Sens.* **2018**, *10*, 1363.
3. U.S. Landsat Analysis Ready Data. Available online: <https://landsat.usgs.gov/ard> (accessed on 14 December 2018).
4. Kovalskyy, V.; Roy, D.P. The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for global 30m Landsat data product generation. *Remote Sens. Environ.* **2013**, *130*, 280–293. [CrossRef]

5. 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](#)]
6. Goward, S.; Williams, D.; Arvidson, T.; Rocchio, L.; Irons, J.R.; Russell, C.; Johnston, S. *Landsat's Enduring Legacy: Pioneering Global Land Observations from Space*; American Society for Photogrammetry and Remote Sensing: Bethesda, MD, USA, 2017; p. 586.
7. Asner, G.P. Cloud cover in Landsat observations of the Brazilian Amazon. *Int. J. Remote Sens.* **2001**, *22*, 3855–3862. [[CrossRef](#)]
8. Ju, J.; Roy, D.P. The Availability of Cloud-free Landsat ETM+ data over the Conterminous United States and Globally. *Remote Sens. Environ.* **2008**, *112*, 1196–1211. [[CrossRef](#)]
9. Tolnai, M.; Nagy, J.G.; Bakó, G. Spatiotemporal distribution of Landsat imagery of Europe using cloud cover-weighted metadata. *J. Maps* **2016**, *12*, 1084–1088. [[CrossRef](#)]
10. Whitcraft, A.K.; Becker-Reshef, I.; Killough, B.D.; Justice, C.O. Meeting earth observation requirements for global agricultural monitoring: An evaluation of the revisit capabilities of current and planned moderate resolution optical earth observing missions. *Remote Sens.* **2015**, *7*, 1482–1503. [[CrossRef](#)]
11. Li, J.; Roy, D.P. A global analysis of Sentinel-2A, Sentinel-2B and Landsat-8 data revisit intervals and implications for terrestrial monitoring. *Remote Sens.* **2017**, *9*, 902.
12. Kovalskyy, V.; Roy, D.P. A one year Landsat 8 conterminous United States study of cirrus and non-cirrus clouds. *Remote Sens.* **2015**, *7*, 564–578. [[CrossRef](#)]
13. Goward, S.; Arvidson, T.; Williams, D.; Faundeen, J.; Irons, J.; Franks, S. Historical record of Landsat global coverage. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1155–1169. [[CrossRef](#)]
14. Arvidson, T.; Goward, S.N.; Gasch, J.; Williams, D. Landsat-7 long-term acquisition plan: Development and validation. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1137–1146. [[CrossRef](#)]
15. Teillet, P.M.; Barker, J.L.; Markham, B.L.; Irish, R.R.; Fedosejevs, G.; Storey, J.C. Radiometric cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM sensors based on tandem data sets. *Remote Sens. Environ.* **2001**, *78*, 39–54. [[CrossRef](#)]
16. Food and Agriculture Organization of the United Nations, Global Administrative Unit Layers. Available online: <http://www.fao.org/geonetwork/srv/en/main.home> (accessed on 14 December 2018).
17. Krimmel, R.M.; Key, C.H.; Fagre, D.B.; Menicke, R.K. Glaciers of the conterminous United States—Glaciers of the western United States. In *Satellite Images of Glaciers of the World*; Professional Paper 1386–J–2; Williams, R.S., Ferrigno, J.G., Eds.; USGS, US Government Printing: Reston, VA, USA, 2002; pp. 329–381.
18. Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [[CrossRef](#)] [[PubMed](#)]
19. Egorov, A.V.; Hansen, M.C.; Roy, D.P.; Kommareddy, A.; Potapov, P.V. Image interpretation-guided supervised classification using nested segmentation. *Remote Sens. Environ.* **2015**, *165*, 135–147. [[CrossRef](#)]
20. Martinec, J.; Rango, A. Areal distribution of snow water equivalent evaluated by snow cover monitoring. *Water Resour. Res.* **1981**, *17*, 1480–1488. [[CrossRef](#)]
21. Olmanson, L.G.; Bauer, M.E.; Brezonik, P.L. A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Remote Sens. Environ.* **2008**, *112*, 4086–4097. [[CrossRef](#)]
22. 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](#)]
23. Egorov, A.V.; Roy, D.P.; Zhang, H.K.; Hansen, M.C.; Kommareddy, A. Demonstration of percent tree cover classification using Landsat analysis ready data (ARD) and sensitivity analysis with respect to Landsat ARD processing level. *Remote Sens.* **2018**, *10*, 209. [[CrossRef](#)]
24. Lee, D.S.; Storey, J.C.; Choate, M.J.; Hayes, R. Four years of Landsat-7 on-orbit geometric calibration and performance. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 2786–2795. [[CrossRef](#)]
25. Roy, D.P.; Ju, J.; Kline, K.; Scaramuzza, P.L.; Kovalskyy, V.; Hansen, M.C.; Loveland, T.R.; Vermote, E.F.; Zhang, C. Web-enabled Landsat Data (WELD): Landsat ETM+ Compositing Mosaics of the Conterminous United States. *Remote Sens. Environ.* **2010**, *114*, 35–49. [[CrossRef](#)]
26. Markham, B.L.; Storey, J.C.; Williams, D.L.; Irons, J.R. Landsat sensor performance: History and current status. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 2691–2694. [[CrossRef](#)]

27. Zhang, H.K.; Roy, D.P. Landsat 5 Thematic Mapper reflectance and NDVI 27-year time series inconsistencies due to satellite orbit change. *Remote Sens. Environ.* **2016**, *186*, 217–233. [[CrossRef](#)]
28. Qiu, S.; Lin, Y.; Shang, R.; Zhang, J.; Ma, L.; Zhu, Z. Making Landsat Time Series Consistent: Evaluating and Improving Landsat Analysis Ready Data. *Remote Sens.* **2019**, *11*, 51. [[CrossRef](#)]
29. Roy, D.P.; Zhang, H.K.; Ju, J.; Gomez-Dans, J.L.; Lewis, P.E.; Schaaf, C.B.; Sun, Q.; Li, J.; Huang, H.; Kovalsky, V. A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. *Remote Sens. Environ.* **2016**, *176*, 255–271. [[CrossRef](#)]
30. Ju, J.; Masek, J.G. The vegetation greenness trend in Canada and US Alaska from 1984–2012 Landsat data. *Remote Sens. Environ.* **2016**, *176*, 1–16. [[CrossRef](#)]
31. Hermance, J.F. Stabilizing high-order, non-classical harmonic analysis of NDVI data for average annual models by damping model roughness. *Int. J. Remote Sens.* **2007**, *28*, 2801–2819. [[CrossRef](#)]
32. Brooks, E.B.; Thomas, V.A.; Wynne, R.H.; Coulston, J.W. Fitting the multitemporal curve: A Fourier series approach to the missing data problem in remote sensing analysis. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3340–3353. [[CrossRef](#)]
33. Roy, D.P.; Yan, L. Robust Landsat-based crop time series modelling. *Remote Sens. Environ.* **2018**. [[CrossRef](#)]
34. Lu, D.; Mausel, P.; Brondizio, E.; Moran, E. Change detection techniques. *Int. J. Remote Sens.* **2004**, *25*, 2365–2401. [[CrossRef](#)]
35. Zhu, Z. Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 370–384. [[CrossRef](#)]
36. Maisongrande, P.; Duchemin, B.; Dedieu, G. VEGETATION/SPOT: An operational mission for the Earth monitoring; presentation of new standard products. *Int. J. Remote Sens.* **2004**, *25*, 9–14. [[CrossRef](#)]
37. Roy, D.P.; Lewis, P.; Schaaf, C.; Devadiga, S.; Boschetti, L. The global impact of clouds on the production of MODIS bidirectional reflectance model-based composites for terrestrial monitoring. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 452–456. [[CrossRef](#)]
38. Griffiths, P.; van der Linden, S.; Kuemmerle, T.; Hostert, P. A Pixel-Based Landsat Compositing Algorithm for Large Area Land Cover Mapping. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **2013**, *6*, 2088–2101. [[CrossRef](#)]
39. Helder, D.; Markham, B.; Morfitt, R.; Storey, J.; Barsi, J.; Gascon, F.; Clerc, S.; LaFrance, B.; Masek, J.; Roy, D.P.; et al. Observations and recommendations for the calibration of Landsat 8 OLI and Sentinel 2 MSI for improved data interoperability. *Remote Sens.* **2018**, *10*, 1340. [[CrossRef](#)]
40. Foga, S.; Scaramuzza, P.L.; Guo, S.; Zhu, Z.; Dille, 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](#)]
41. 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](#)]
42. Brooks, E.; Wynne, R.; Thomas, V. Using Window Regression to Gap-Fill Landsat ETM+ Post SLC-Off Data. *Remote Sens.* **2018**, *10*, 1502. [[CrossRef](#)]
43. Wolfe, R.; Nishihama, M.; Fleig, A.; Kuyper, J.; Roy, D.; Storey, J.; Patt, F. Achieving sub-pixel geolocation accuracy in support of MODIS land science. *Remote Sens. Environ.* **2002**, *83*, 31–49. [[CrossRef](#)]
44. Yan, L.; Roy, D.P.; Li, Z.; Zhang, H.K.; Huang, H. Sentinel-2A multi-temporal misregistration characterization and an orbit-based sub-pixel registration methodology. *Remote Sens. Environ.* **2018**, *215*, 495–506. [[CrossRef](#)]

