

Article

# A Global MODIS Water Vapor Database for the Operational Atmospheric Correction of Historic and Recent Landsat Imagery

David Frantz <sup>1,\*</sup>, Marion Stellmes <sup>2</sup> and Patrick Hostert <sup>1,3</sup>

<sup>1</sup> Geomatics Lab, Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany; patrick.hostert@geo.hu-berlin.de

<sup>2</sup> Remote Sensing and Geoinformatics, Freie Universität Berlin, Malteserstr. 74-100, 12249 Berlin, Germany; marion.stellmes@fu-berlin.de

<sup>3</sup> Integrated Research Institute on Transformations of Human Environment Systems (IRI THESys), Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

\* Correspondence: david.frantz@geo.hu-berlin.de

Received: 7 January 2019; Accepted: 26 January 2019; Published: 28 January 2019



**Abstract:** Analysis Ready Data (ARD) have undergone the most relevant pre-processing steps to satisfy most user demands. The freely available software FORCE (Framework for Operational Radiometric Correction for Environmental monitoring) is capable of generating Landsat ARD. An essential step of generating ARD is atmospheric correction, which requires water vapor data. FORCE relies on a water vapor database obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS). However, two major drawbacks arise from this strategy: (1) The database has to be compiled for each study area prior to generating ARD; and (2) MODIS and Landsat commissioning dates are not well aligned. We have therefore compiled an application-ready global water vapor database to significantly increase the operational readiness of ARD production. The free dataset comprises daily water vapor data for February 2000 to July 2018 as well as a monthly climatology that is used if no daily value is available. We systematically assessed the impact of using this climatology on surface reflectance outputs. A global random sample of Landsat 5/7/8 imagery was processed twice (i) using daily water vapor (reference) and (ii) using the climatology (estimate), followed by computing accuracy, precision, and uncertainty (APU) metrics. All APU measures were well below specification, thus the fallback usage of the climatology is generally a sound strategy. Still, the tests revealed that some considerations need to be taken into account to help quantify which sensor, band, climate, and season are most or least affected by using a fallback climatology. The highest uncertainty and bias is found for Landsat 5, with progressive improvements towards newer sensors. The bias increases from dry to humid climates, whereas uncertainty increases from dry and tropic to temperate climates. Uncertainty is smallest during seasons with low variability, and is highest when atmospheric conditions progress from a dry to a wet season (and vice versa).

**Keywords:** atmospheric correction; analysis ready data; global; Landsat; water vapor

## 1. Introduction

Landsat data are one of the most important Earth Observation datasets ever acquired due to their long-term availability on regional to global scales [1–3], data continuity between sensors [4], high quality of (inter-)calibration (e.g., [5,6]), free and open data access [7], and adequate spatial, temporal, and spectral resolutions to monitor many processes on the landscape level [8]. However, most analyses require a significant amount of pre-processing which needs to be increasingly automated to enable both large area [9] and time series analyses [10]. Necessary processing steps include (1) cloud and

cloud shadow detection on a per-pixel basis, also including quality flags that indicate only partially or not usable observations from other factors than clouds or cloud shadows like saturation, snow, illumination conditions etc.; (2) radiometric normalization and transformation of top-of-atmosphere reflectance to surface reflectance, which includes at least atmospheric correction (AC), but other corrections like topographic or adjacency effect correction may be required as well; and (3) mapping the data onto a uniform spatial grid (often referred to as data cubing), which is highly advantageous for data-intensive applications. Data that have undergone all these pre-processing steps are currently referred to as Analysis Ready Data (ARD, [11], <http://ceos.org/ard>).

The freely available software FORCE (Framework for Operational Radiometric Correction for Environmental monitoring, available from <http://force.feut.de>) is a toolkit for generating ARD as well as for deriving higher-level products from the generated ARD (such as pixel-based composites, spectral variability metrics, time series derivatives like trend and change indicators, Land Surface Phenology metrics, etc.). The AC included in FORCE [12] uses radiative transfer theory [13] to transform top-of-atmosphere reflectance to surface reflectance [14]. The quality of the FORCE AC was recently assessed in the Atmospheric Correction Inter-Comparison Exercise (ACIX) [15].

AC of Landsat data requires knowledge on the amount of atmospheric water vapor. Water vapor absorbs a significant amount of radiation in specific regions of the electromagnetic spectrum. Unlike other atmospheric gases, water vapor is spatially and temporally variable and can only be directly estimated from satellite data that have a designated water vapor absorption channel [16], which, for example, provides highly accurate measurements for Sentinel-2 [15]. On the contrary, the Landsat sensors are not equipped with spectral bands that enable a reliable estimation of atmospheric water vapor, therefore auxiliary data are necessary to perform the necessary corrections.

For this purpose, FORCE relies on a water vapor database (WVDB) obtained from the Moderate Resolution Imaging Spectroradiometer MODIS Terra/Aqua water vapor products [14]. If available, nearly concurrent, daily values are used. If not, a monthly climatology is used as a surrogate. This climatology is derived from a statistical analysis of all daily estimates and represents the climatological seasonal water vapor phenology. The major drawback is that the climatology has to be compiled for each study area prior to AC. This is computationally expensive, time consuming, and most likely redundant as diverse users and applications require similar pre-processing. Therefore, this extra processing step is not well aligned with the application-readiness required for large-scale ARD production. The objectives of this paper are as follows:

- To compile, provide and describe an open and free global water vapor database that can be simply ingested into the FORCE AC to significantly increase readiness of ARD production;
- To provide a comprehensive and extended analysis of the actual impact of falling back to a water vapor climatology vs. correcting with daily values for (i) the different Landsat sensors, (ii) different climatic regions of the world, and (iii) different seasons.

## 2. Data

### 2.1. MODIS Water Vapor

MODIS water vapor data are used for the physically-based correction of water vapor absorption in the earth's atmosphere [14]. We used the MOD05 and MYD05 collection 6.1 products as well as the MOD03 and MYD03 geolocation tables, which were automatically downloaded from the Level 1 and Atmosphere Archive and Distribution System (LAADS) at NASA's Goddard Space Flight Center. We used data that were derived from the near infrared (NIR) water vapor algorithm [17], which relies on water vapor attenuation of the NIR radiation. The column water vapor amount was determined from radiative transfer theory based on ratios between bands affected by water vapor absorption and bands in unaffected atmospheric windows. The product was obtained at 1 km spatial resolution and the temporal resolution was up to one day, considering both Aqua and Terra observations.

## 2.2. Landsat

Based on an analysis of the water vapor database, a large number of globally distributed Landsat images (4320) was randomly selected for further analysis of the uncertainty in using a water vapor climatology vs. having daily values for the atmospheric correction of Landsat imagery. We used Landsat 7 and Landsat 8 Collection 1 Level 1 Tier 1 L1TP data for the year 2015 with cloud cover <70% as well as Landsat 5 for the year 2010. The images were acquired from the U.S. Geological Survey (USGS). Note that the failure of Landsat 7's Scan Line Corrector does not influence this analysis: Radiometric data quality was not affected by the failure [18]. Apart from the cloud cover constraint and the constraint that a daily value needed to be available in the water vapor database, the sampling was completely random, thus the datasets covered all land cover classes.

## 3. Methods

### 3.1. Building the Water Vapor Database

The global MODIS water vapor database was generated with the designated FORCE module WVDB [12], which was implemented for automatic data acquisition and processing of Terra/Aqua MODIS water vapor data; for details about this methodology see Frantz et al. (2016) [14].

For each day between March 2000 and July 2018, all daytime observations were downloaded. For each day, the mean atmospheric water vapor for an area of  $\pm 0.5^\circ$  around the central coordinate of each land-intersecting, descending Worldwide Reference System 2 (WRS-2) scene was computed (13,281 coordinates). The  $1^\circ$  box covers the better part of a Landsat WRS-2 scene, and thus allows the extraction of one representative average for the atmospheric correction of a near-concurrent Landsat image. As the Sun-cloud-sensor path is substantially shorter than the Sun-surface-sensor path, column water vapor retrievals over clouds are substantially lower than in a clear atmosphere [17]. In order to use the water vapor retrievals for converting top-of-atmosphere reflectance to surface reflectance, it is important to filter for clouds. Sun-glint was filtered as well. As the MODIS cloud detection agrees with an independent dataset in 85% of cases [19], we opted to use a rather conservative threshold for obtaining a daily value: If there were less than 10% valid pixels per site and date, the average was expected to be unreliable, and no daily value was derived. If there were estimates from both sensors, the one with more valid data was selected.

To allow AC of Landsat data at times with no concurrent MODIS data (including the pre-MODIS era, coverage gaps between the swaths, sensor outages, etc.), we additionally generated a water vapor climatology for each location as a surrogate value. This climatology was derived from a statistical analysis of the complete MODIS water vapor time series, i.e., the daily values were aggregated to long-term (2000–2018) averages per month (January to December, i.e., 12 values per coordinate).

The full water vapor database was released under the terms of the Creative Commons Attribution 3.0 Unported and can be freely accessed at <https://doi.pangaea.de/10.1594/PANGAEA.893109> [20]. The database is ready-to-use with the AC implemented in FORCE and can also be used as an input for other atmospheric corrections or generally for other purposes.

### 3.2. Impact of Water Vapor Climatology on Landsat Data

#### 3.2.1. Sampling

To systematically explore the impact of using a water vapor climatology vs. having daily values for the AC of Landsat imagery as well as to evaluate how this approach performs across the globe and throughout seasons, we used a global stratification to sample Landsat test data. We clustered the water vapor climatology to derive regions with similar water vapor seasonality and monthly variability. The monthly water vapor averages and the monthly standard deviation were z-transformed, respectively, and combined into a set of 24 features. This set was partitioned into 12 clusters around medoids, which is a more robust version of K-means clustering [21,22].

We randomly selected Landsat images with <70% cloud cover on dates where daily water vapor values were available. We sampled 10 images per cluster, month, and sensor. This resulted in a total of 4320 images. This sampling allowed us to test how large the effect of using the climatology is (i) for the different sensors, (ii) in different climatic regions of the world, and (iii) between seasons.

### 3.2.2. Atmospheric Correction

All acquired Landsat images were processed twice using the FORCE AC [14]. All settings but the use of the water vapor database were held constant such that the resulting difference in surface reflectance could solely be attributed to the use of the water vapor climatology. The usage of daily water vapor values from MODIS is common practice for AC of Landsat (e.g., [23]), which is enabled by the good accuracy of the MODIS water vapor retrieval (5–10% errors) [17] and facilitated by the orbital similarity between Landsat 7 and the Terra spacecraft [24]. In the first set, the images were corrected with daily water vapor values. The resulting surface reflectance dataset served as a reference dataset for the following analyses. It should be noted that the scope of this manuscript is not to provide an absolute measure of surface reflectance accuracy (for such information, see [15], where surface reflectance generated with daily water vapor was independently validated in the ACIX experiment using an AERONET (Aerosol RObotic NETwork [25])-derived surface reflectance reference) but rather to explore the feasibility of using a climatological surrogate value as a backup method when daily values are not available. In the second set, the images were corrected with the climatological values, thus simulating the worst case scenario where no concurrent water vapor estimates from MODIS were available, as, for example, is the case for all Landsat images acquired before the commissioning of MODIS in March 2000.

### 3.2.3. Surface Reflectance Evaluation

Following Vermote and Kotchenova [26], the two surface reflectance sets were compared in terms of accuracy, precision, and uncertainty (APU analysis) as follows:

$$A = \frac{1}{n} \cdot \sum_{i=1}^n \varepsilon_i, \quad (1)$$

$$P = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (\varepsilon_i - A)^2}, \quad (2)$$

$$U = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n \varepsilon_i^2}. \quad (3)$$

APU analyses are an established evaluation of surface reflectance estimates against a well characterized reference dataset and have been extensively used to evaluate MODIS, Landsat, and Sentinel-2 surface reflectance products [15,23,26,27]. The surface reflectance generated with daily water vapor served as the reference  $\rho_{r,i}$ , and the surface reflectance generated with the water vapor climatology as the estimate  $\rho_{e,i}$ . The difference  $\varepsilon_i$  was defined as the estimate minus the reference, where  $n$  is the number of samples  $i$  and  $M$  is the mean value of the reference. Accuracy, precision, and uncertainty are measures of the mean bias, repeatability, and statistical deviation (i.e. RMSE). To address whether uncertainty is high or not, values need to be evaluated against the surface reflectance specification, which we adapted from Vermote and Kotchenova [26] and Claverie et al. [27] as follows:

$$S = 0.05\rho + 0.005. \quad (4)$$

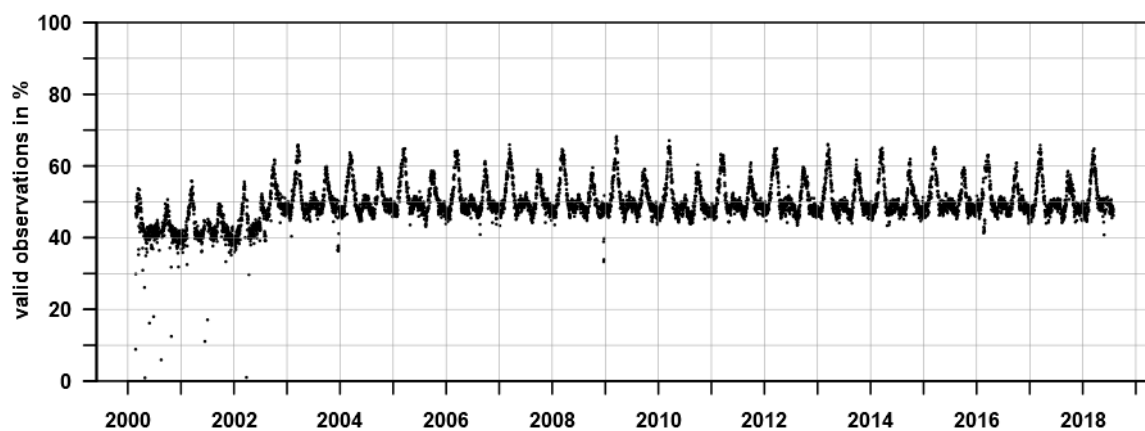
The APU analyses were performed for each band and for each of the testing sets outlined above. Pixels flagged as cloud, cloud shadow, snow, or saturated were removed prior to the analysis.

## 4. Results and Discussion

### 4.1. Water Vapor Database

#### 4.1.1. Data Availability

The daily water vapor database consists of 6,675 daily tables, starting from 24 February 2000 to 31 July 2018. The daily availability before commissioning of the Aqua platform in 2002 was 40–55%, and after commissioning was 45–65% (Figure 1). Note that this does not imply that only 40–65% of Landsat scenes have daily values as Landsat has no daily coverage and is well aligned with MODIS orbits. During the early years of MODIS operation (2000–2002), sensor outages were more frequent [28,29], thus data availability is lower on some days. There is some seasonality in data availability, which seemingly is related to the climatic seasons and large scale atmospheric circulations (causing persistent cloud cover).



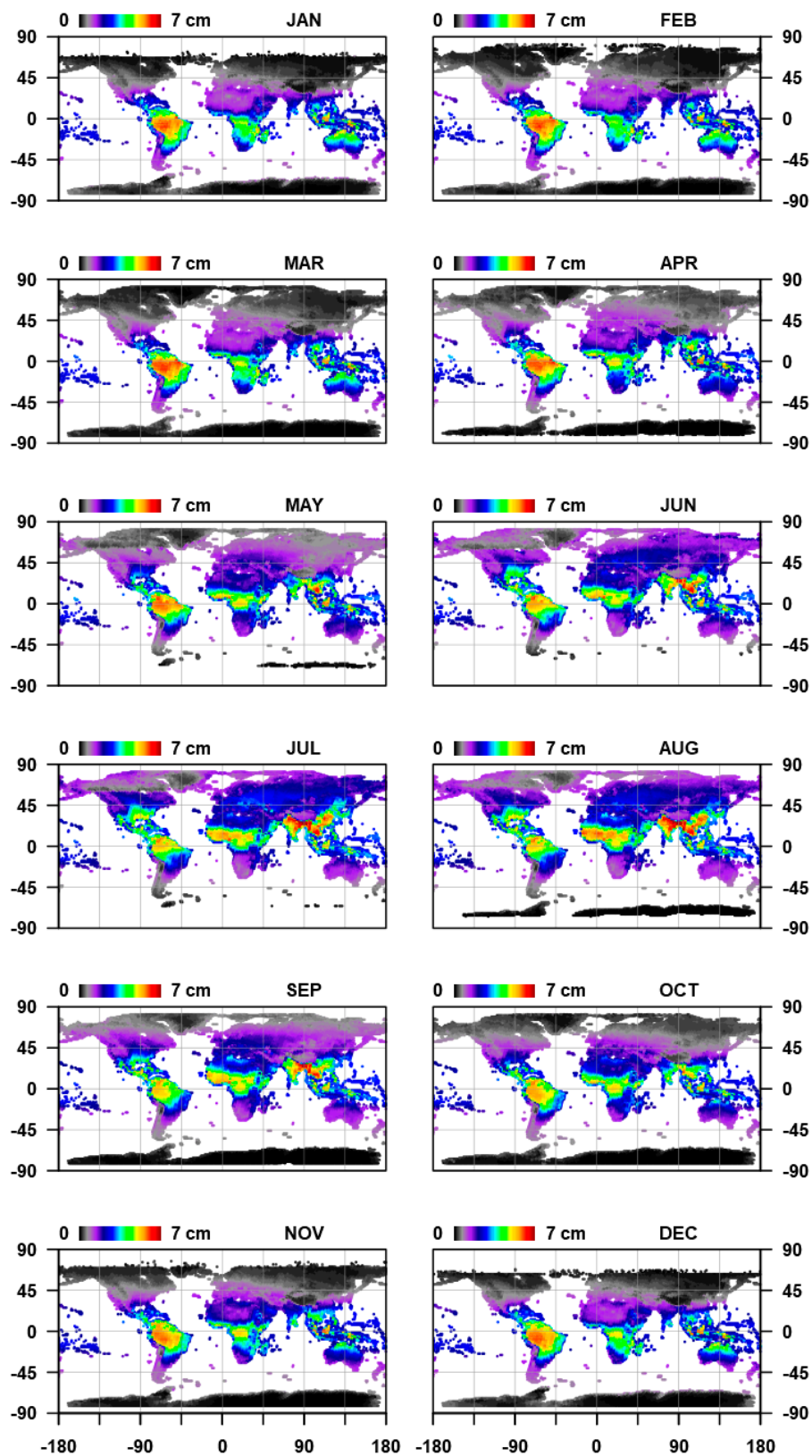
**Figure 1.** Daily availability of water vapor retrievals. The daily availability shows the percentage of the 13,281 coordinates with valid water vapor retrievals.

#### 4.1.2. Water Vapor Climatology

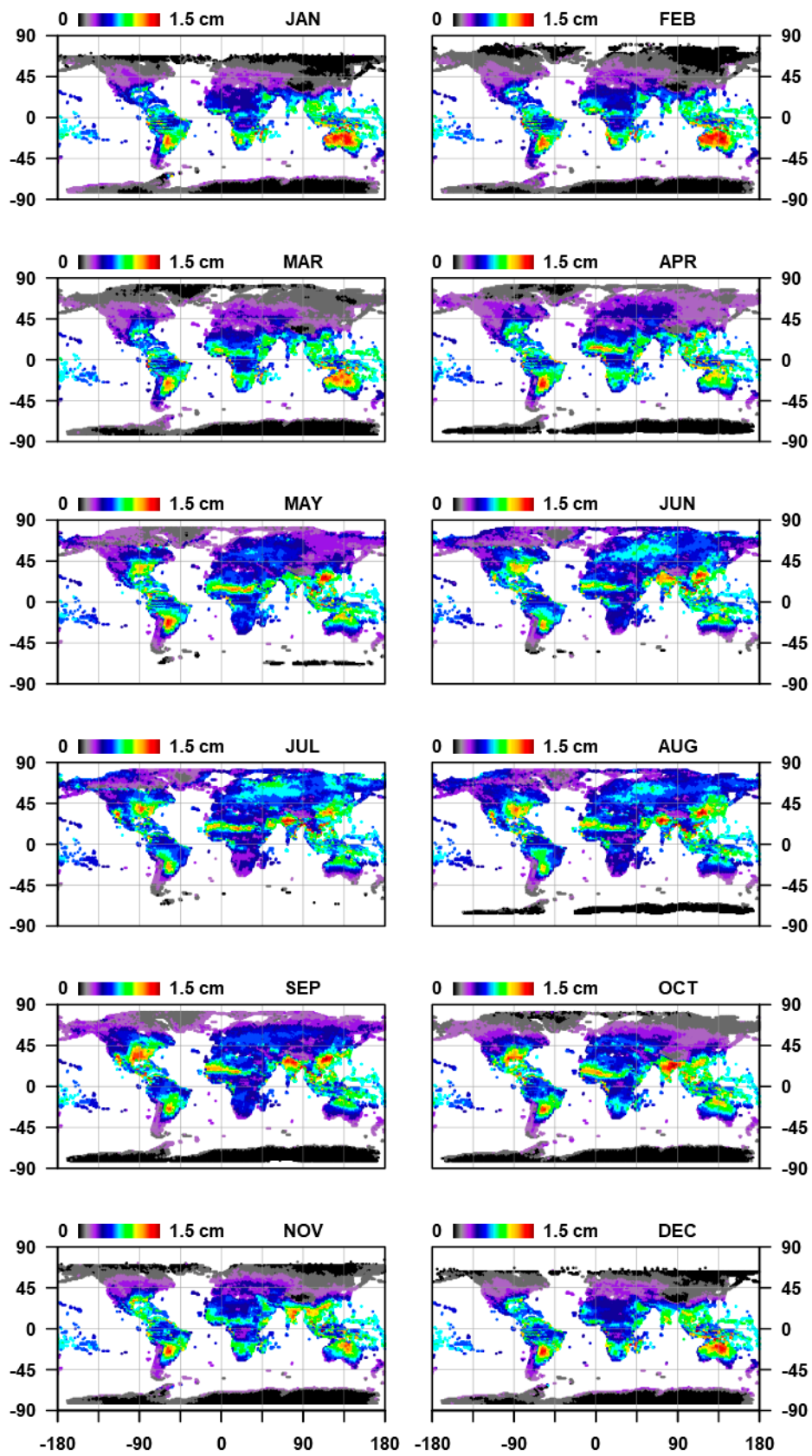
The water vapor climatology was computed on basis of the daily water vapor retrievals. For each coordinate, the long-term monthly average was computed (Figure 2). The climatology readily reveals the world's climate zones as well as their seasonality, with high water vapor in the tropics and small values in the polar regions, and the latitudinal movement of the Intertropical Convergence Zone or the progression of the monsoon on the Indian subcontinent. Consistent missing values in the Arctic/Antarctic are due to the polar night. Figure 3 shows the monthly standard deviation as a measure of variability from the climatological mean. The highest variability can be found year-round in Southern America and Northern Australia and in wet season climates throughout the world, e.g., in Summer (Winter) in Northern (Southern) Africa, whereas the tropics are characterized by a fairly stable water vapor climatology. Long-term trends of annual water vapor averages per month reveal that some significant changes in water vapor have occurred during the past 18 years (Figure 4). As an example, South-Eastern Asia has become significantly wetter. Thus, accounting for the trend might give better performance than using a simple long-term average. This will likely be incorporated into a future version of the algorithm and database. However, this should only be considered for interpolating between 2000 to 2018 as the trend cannot be reliably extrapolated into the past with this methodology. In addition, in a previous study [14], we assessed that less than 0.4% of all Landsat images after 2002 could not have been paired with a daily water vapor value (using all Landsat data in a significant portion of Southern Africa with cloud coverage <70%). This is mainly because a high cloud coverage in MODIS (resulting in no data in the water vapor database) is also commonly paired with high cloud coverage in the near-concurrent Landsat image.



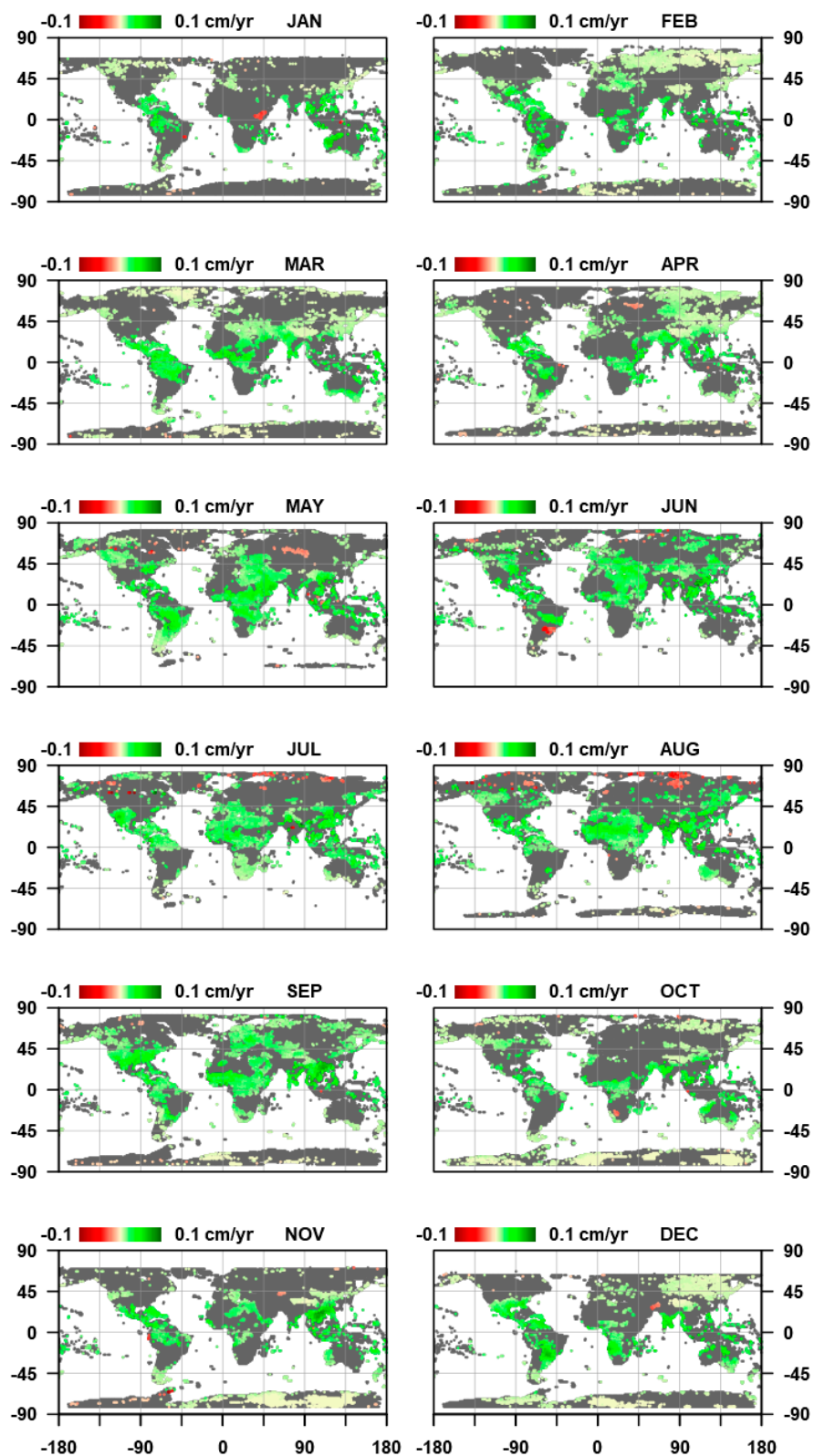
Animated versions of the monthly mean (Video S1), standard deviation (Video S2), and long-term trend of monthly averages (Video S3) can be found in the Supplementary Materials.



**Figure 2.** Long-term mean water vapor per month, averaged from February 2000–July 2018 based on daily values.



**Figure 3.** Long-term variability of water vapor per month, standard deviation from February 2000–July 2018 based on daily values.



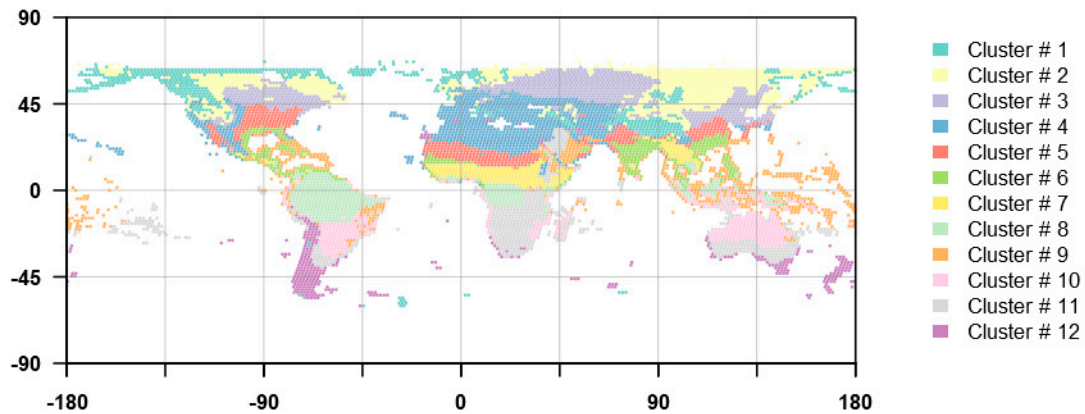
**Figure 4.** Long-term trend of annual water vapor averages per month (2000–2018). Only significant trends ( $\alpha = 0.05$ ) are shown.

#### 4.1.3. Clustering

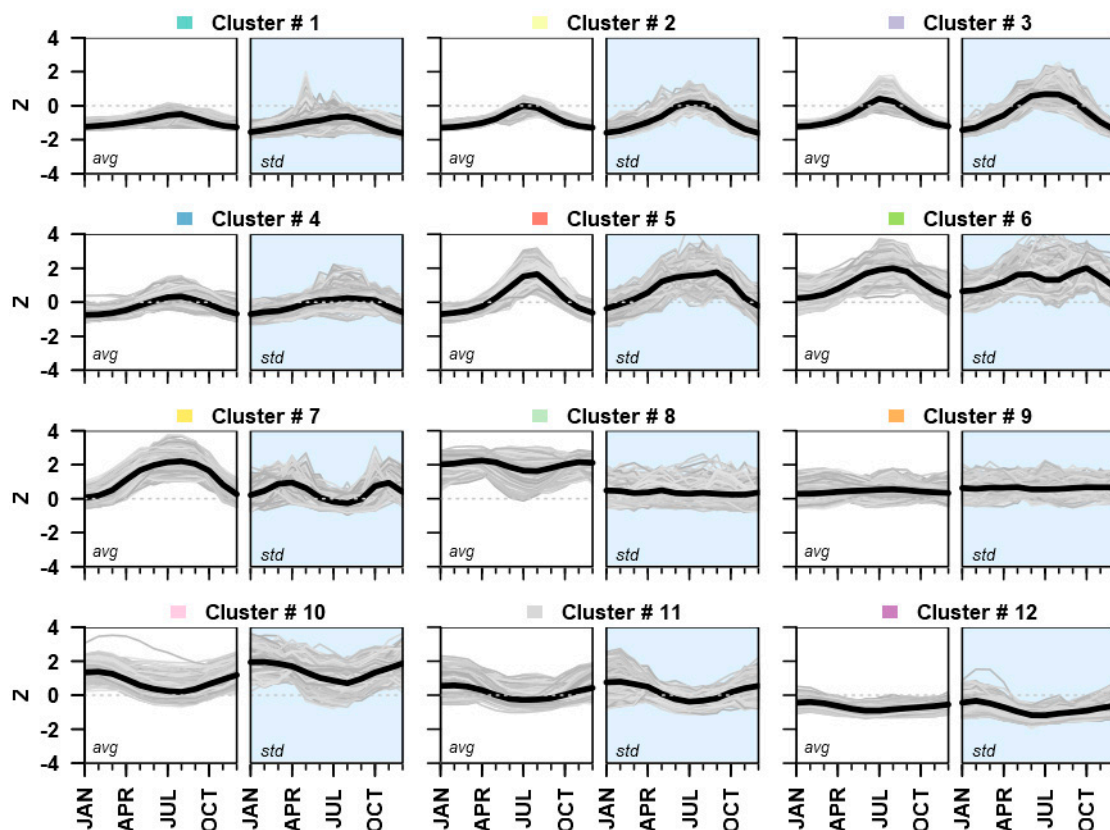
To assess the effect of falling back to climatological values in the atmospheric correction of Landsat imagery, we clustered (Figure 5) the water vapor climatology into broad clusters for sampling purposes.



The average climatology and variability per cluster are shown in Figure 6. Values greater (smaller) than zero show mean water vapor and/or variability that is higher (smaller) than the global average. As an example, Cluster #5 is characterized by a dry atmosphere with small variability during winter and a wet atmosphere with much variability during summer. This climate is, for example, dominant in Mid-Eastern North America and in the Sahel zone.



**Figure 5.** Cluster map, derived from a partitioning of the water vapor climatology and monthly variability of each location around medoids.



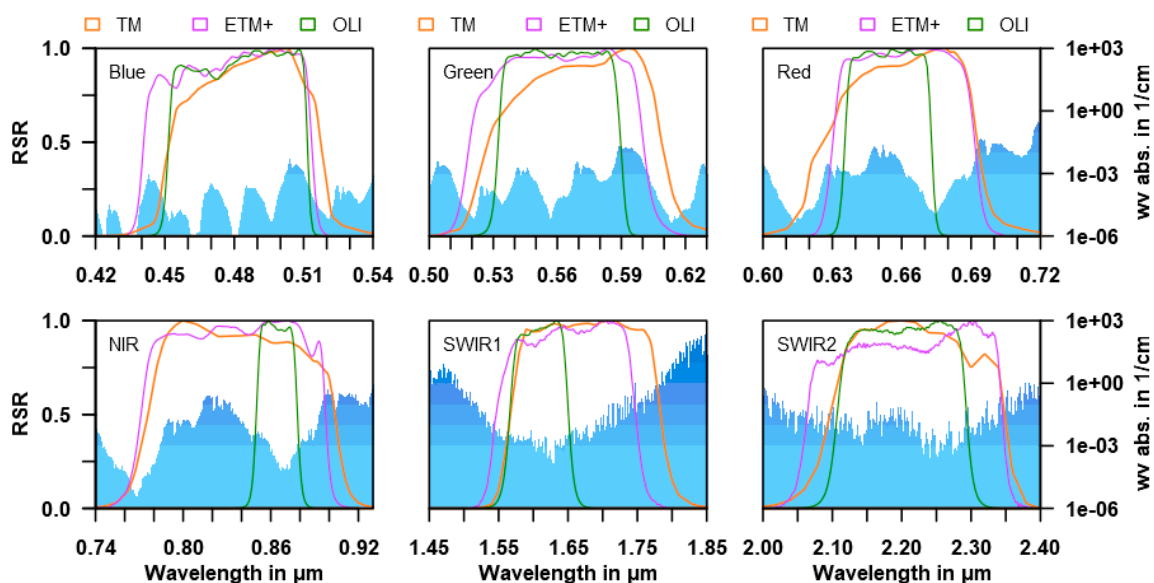
**Figure 6.** Water vapor seasonality and variability per cluster (Figure 5). The left panel (white background) shows the mean seasonality; the right panel (blue background) shows monthly variability, i.e., standard deviation. The grey curves represent the seasonality and variability of each coordinate in the cluster. The black curve represents the mean seasonality and variability of the cluster. Values were z-transformed.

#### 4.2. Effect of Water Vapor Climatology on Landsat Reflectance

To systematically explore the impact on surface reflectance when using a water vapor climatology vs. having daily values for the atmospheric correction of Landsat imagery, we have processed each Landsat test image twice (i) using daily water vapor values, i.e., the reference, and (ii) using the water vapor climatology. APU metrics were generated for testing the effect of using the climatology (i) for the different Landsat sensors, (ii) in the different clusters, and (iii) between seasons.

##### 4.2.1. Sensor

Band placement, width, and spectral response are different in Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper plus (ETM+), and the Landsat 8 Operational Land Imager (OLI) for each Landsat legacy band [30] as depicted by the Relative Spectral Response (RSR) functions (Figure 7) in relation to water vapor absorption coefficients [31]. Only small changes from one sensor generation to the other affect the visible bands where water vapor absorption is also low. The most significant changes affected the NIR and SWIR1 bands. The TM and ETM+ NIR RSR functions are fairly similar, but the band width was substantially reduced for OLI, and the band was centered on an atmospheric window. Similarly, the OLI SWIR1 was narrowed and moved to an atmospheric window while the broader TM and ETM+ bands are influenced by the right wing of a strong absorption band. Although the band width of the OLI SWIR2 was slightly reduced, the improvement seems to be smaller than for NIR and SWIR1. In the following, we assess how the different RSR functions affect surface reflectance when using climatological water vapor estimates.

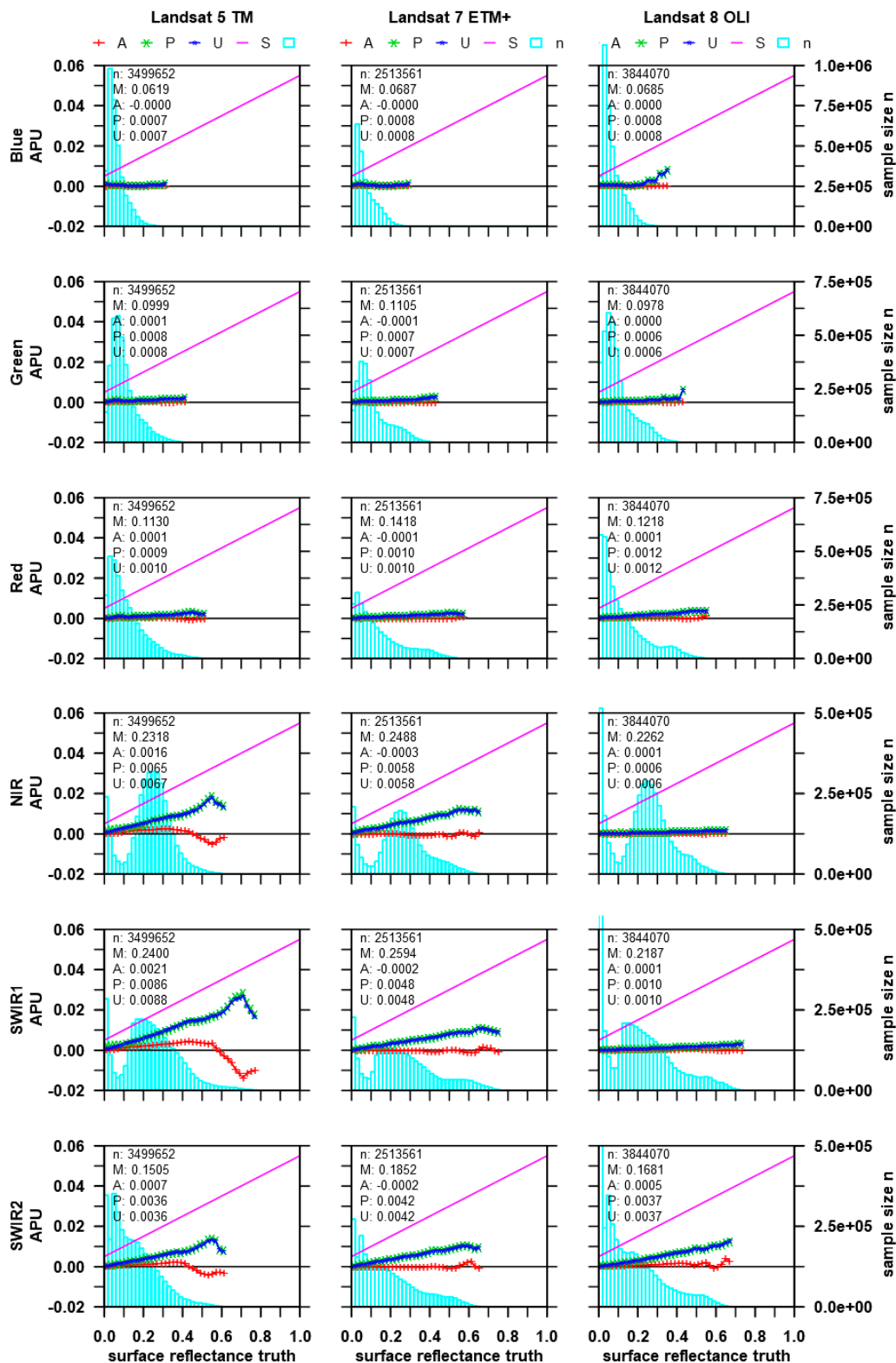


**Figure 7.** Band-wise Relative Spectral Response (RSR) per sensor (colored lines) and water vapor absorption coefficients (blue polygon). The water vapor absorption coefficients are plotted with logarithmic scale and color bar.

Figure 8 presents the unambiguous APU analysis for Landsat 5, Landsat 7, and Landsat 8. The specified uncertainty is based on a conservative error budget derived for MODIS [23,32]. Accuracy, precision, and uncertainty never exceed the specification. All measures are close to zero for the visible bands in each sensor, which is due to low absorption (see Figure 7). The statistics for NIR and SWIR bands show larger values. Uncertainty and precision are fairly proportional with reflectance in all cases, as is specification. This confirms that the effect of water vapor is albedo-driven as assessed in Frantz et al. [14]. Uncertainty and precision are almost identical, which means that the bias is low in relation to the absolute value of the surface reflectance truth. For Landsat 5, the largest bias (accuracy, 2.1‰ reflectance) and uncertainty (8.8‰ reflectance) are found in SWIR1, which is substantially

affected by the right wing of a strong water vapor absorption band (Figure 7). Overall, the bias is positive, which means that the reflectance is overestimated when using the climatology. A bias is present for each brightness level, but there is a strong increase in absolute bias above 55% reflectance. This is likely related to the decreasing sample size as Landsat 5 TM saturates at this reflectance level, which causes the descriptive statistics to become less reliable. For Landsat 7, the effect is visible as minor oscillation. However, it is less pronounced compared to Landsat 5 as the ETM+ has the capability to adjust offset and gain settings in response to expected reflectance range within the footprint. Landsat 8 is unaffected, as improved radiometric quantization and sensitivity prevent saturation. Landsat 5 NIR and SWIR2 are also characterized by a bias of 1.6‰ and 0.7‰ reflectance, respectively. Uncertainty has decreased and precision has improved with Landsat 7 for the NIR and SWIR1 bands, whereas no improvement is evident in the SWIR2. This reflects that the NIR and SWIR1 were narrowed while SWIR2 was not substantially changed (Figure 7). In all bands, the bias largely vanished (all bands <0.5‰ reflectance). Further improvements have been made with Landsat 8: There is virtually no bias in most bands except SWIR2 (0.5‰ reflectance). Uncertainty and precision have improved considerably for the NIR and SWIR1 (<1‰ reflectance). The performance of the SWIR2 has remained the same as with Landsat 5 and 7. This is in accordance with the spectral setup, as NIR and SWIR1 were narrowed substantially and both positioned in an atmospheric window, while no substantial change was done for the SWIR2 (Figure 7).

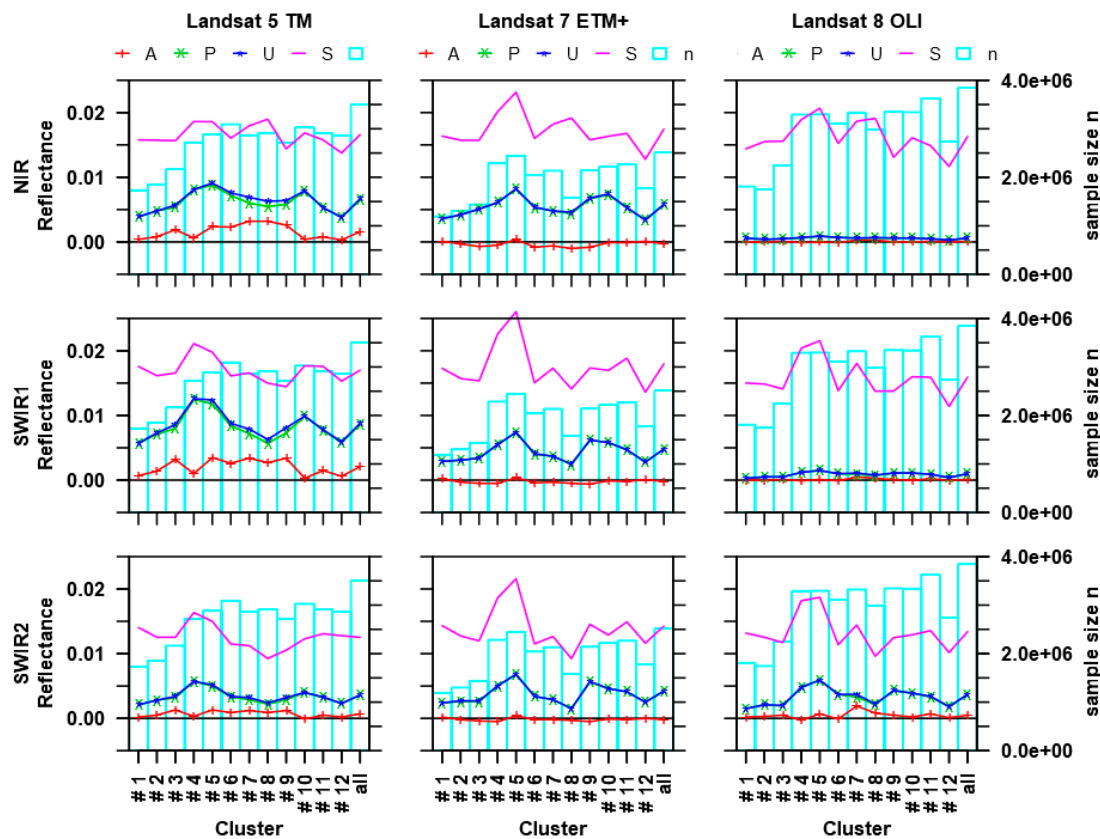
To summarize, all sensors and bands are well below specification, which means that the fallback usage of the climatology is generally a sound strategy. However, it is noted that the uncertainty of the climatology adds to the combined uncertainty of all error sources and processing steps (calibration, AC in general, etc.) and thus might push combined uncertainty above the specification. Nevertheless, the use is uncritical for the visible bands whereas some considerations have to be taken into account for the NIR and SWIR bands. Uncertainty in SWIR2 is the same for all Landsat sensors, which means that special care is appropriate when using SWIR2 information in quantitative analyses, probably even for Landsat 8. Thus, it is important to routinely update the water vapor dataset [20] and not simply use the fallback climatology for all upcoming Landsat 8 acquisitions. For NIR and SWIR1, the use of the climatology seems to be fairly uncritical for Landsat 8. Highest bias and uncertainty were assessed for Landsat 5's NIR and SWIR bands, which is of special importance as daily MODIS acquisitions are unavailable for many Landsat 5 acquisitions (due to the 1984–2013 lifetime that substantially extends the availability of Terra and Aqua MODIS data since 2000 and 2002, respectively). This has important implications for the continuity of long-term analyses where a sudden reduction of uncertainty after the commissioning of MODIS (and thus availability of daily water vapor estimations) needs to be factored in. Compared to Landsat 5, the uncertainty and bias for Landsat 7's NIR and SWIR1 was considerably improved. Even though some uncertainty remains, we do not consider this problematic in practice because (1) Landsat 7 was commissioned in 1999 and thus only data related to a short time period cannot be aligned with daily MODIS data; and (2) Landsat 7 and Terra are in the same orbit with only a 30 minute observation delay [33], which significantly increases the chance of nearly concurrent observations under similar cloudiness.



**Figure 8.** Accuracy, precision, and uncertainty (APU, left y-axis) for all Landsat sensors and bands. Sensors are arranged in columns, and the 6 Landsat heritage bands in rows. The reference reflectance (truth) was obtained using daily water vapor values, and the estimated reflectance was derived using the water vapor climatology. The cyan bars are the number of samples for each 2% reflectance bin (right y-axis). M is the average surface reflectance truth. A (red line), P (green), and U (blue) were calculated for each bin (lines) and for all samples in total (legend). APU should not exceed the theoretical specification, i.e., they should remain under the magenta line.

#### 4.2.2. Cluster

Due to the sensor-specific findings, APU analyses for Landsat 5–8 relating to different clusters are presented for the NIR and SWIR bands only (Figure 9). The sensor-specific findings of the previous section can be confirmed with this test. In addition, regional differences do exist, and they are fairly consistent between sensors and bands.



**Figure 9.** Accuracy, precision, and uncertainty (APU, left y-axis) for each cluster (Figure 5). Landsat sensors are arranged in columns, and NIR and SWIR bands are arranged in rows. For each case (e.g., Cluster #1, Landsat 5, NIR), an APU plot was generated, and visualized here are the overall APU statistics (which are shown as text in the legend of Figure 8). Thus, the ‘all’ Clusters are identical to the values presented in Figure 8, and are included in Figure 9 for interpreting the regional differences with regards to the global assessment. Colors and signatures are as in Figure 8. The cyan bars are the number of samples in each cluster (right y-axis). The magenta line represents the specified theoretical uncertainty based on the mean surface reflectance truth.

The smallest effect when using the water vapor climatology relates to the driest world regions, which are represented in clusters #1–3 and #11–12 in the Northern and Southern hemispheres, respectively. As a general rule, the bias increases from the most Northern/Southern clusters towards the equator, with highest bias in the tropical and sub-tropical climates. This is likely an effect of increasing humidity from North (South) towards the equator where larger absolute deviations between the climatology and the daily values cause a larger bias when calculating surface reflectance. As an exception, the bias is small in the temperate clusters #4 and #10, which may relate to cluster heterogeneity.

The regional patterns of uncertainty and precision are mostly similar to the bias. Lowest uncertainty is found in the most Northern and Southern clusters. With increasing humidity towards the South (North), uncertainty mostly increases (Clusters #1 to 5, and Clusters #12 to 10). However, the pattern is interrupted in the tropical and sub-tropical climates (Clusters #6–9), where uncertainty

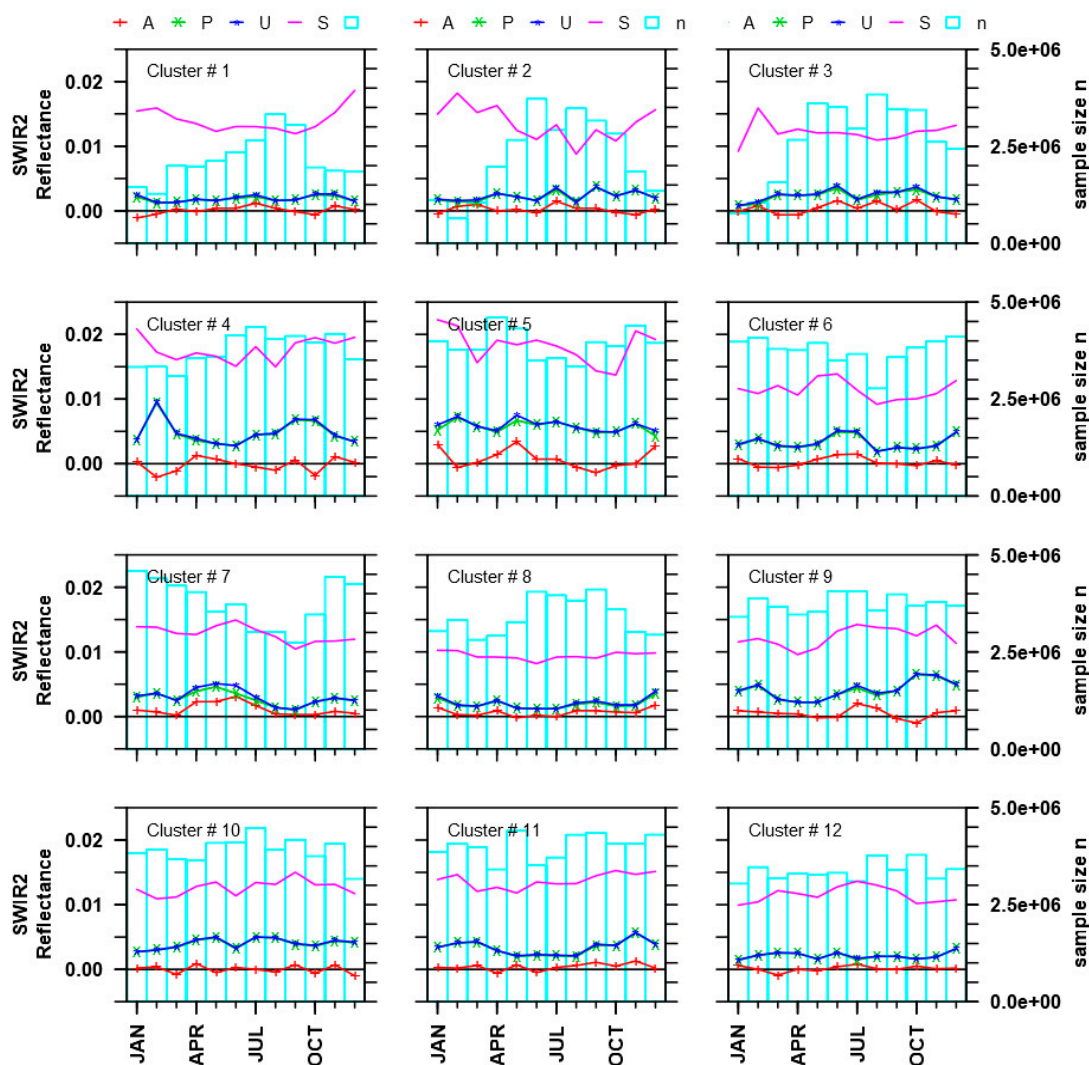


becomes smaller again. This might indicate that the effect is not simply a function of overall humidity but also of seasonality and variability. As an example, a fairly low uncertainty is found in tropical cluster #8, where humidity is high but seasonality and seasonal variability is low (compare with Figure 6). This will be examined further using the next test.

To summarize the findings of this test, the usage of the water vapor climatology is most accurate in dry climates. Towards wetter climates, the bias of the surface reflectance retrieval increases and is highest in the tropics. The uncertainty of using climatological values also increases towards wetter climates but is comparatively smaller for tropical and sub-tropical regions than for temperate climate zones. However, in temperate climates, the bias is low.

#### 4.2.3. Season

The test across the different clusters was most consistent for the SWIR2 band across different sensors. In order to increase sample size to 30 Landsat images per stratum, the per-season test was only performed for the SWIR2 band by pooling data from all Landsat sensors (Figure 10).



**Figure 10.** Accuracy, precision, and uncertainty (APU, left y-axis) for each season (x-axis) in each cluster (Figure 5) for the pooled Landsat 5–8 SWIR2 band. For each case (e.g., Cluster #1, January) an APU plot was generated, and visualized here are the overall APU statistics (which are shown as text in the legend of Figure 8). Colors and signatures are as in Figure 8. The cyan bars are the number of samples in each stratum (right y-axis). The magenta line represents the specified theoretical uncertainty based on the mean surface reflectance truth.

Our previous observation that bias is related to the absolute amount of water vapor seems to hold true in many cases (e.g., April–July, Cluster #7 in Figure 10). However, as accuracy in SWIR2 did vary between sensors, our description and discussion will focus on uncertainty.

As observed in our analysis along clusters, uncertainty is related to the overall humidity, but other factors seem to play a role as well. It now becomes evident that uncertainty is more closely related to the seasonal variability (Figure 10). If variability is small, uncertainty is small as well (e.g., all-year round in Clusters #1, 8, and 12), whereas high uncertainty is related to high variability (e.g., the seasonality in Cluster #7). A linear regression between uncertainty and variability confirms this:  $R^2 = 0.167$ , with a highly significant slope coefficient ( $p < 3.73 \times 10^{-7}$ ). This model does not have a very high predictive power, but the test nonetheless indicates an unambiguous correlation. The fairly low coefficient of determination suggests that the sample size of this test was probably still too low (30 Landsat images per stratum), which may have caused some unexplainable irregularities in the uncertainty assessment (e.g., the sudden spike in February in Cluster #4).

However, some general conclusions can nonetheless be drawn from this test. Using the water vapor climatology is most accurate in climates and seasons where the monthly variability is low. If variability of water vapor increases, the uncertainty of surface reflectance retrieval increases likewise. Our test accordingly disentangles where and when the caution of using a water vapor climatology is appropriate. In addition, it helps to suggest algorithm improvements: It appears that uncertainty is specifically high in months where the atmospheric conditions progress from a dry to a wet season (and vice versa, e.g., Clusters #4, 7, and 11). In these cases, the climatological mean is the average of values from wet and dry conditions. Therefore, we will consider narrowing the temporal binning for calculating the climatology, for example, to use bi-weekly, weekly, or daily long-term averages.

## 5. Conclusion

We have compiled a global water vapor database from the Terra and Aqua MODIS Total Precipitable Water Vapor products. This water vapor database is freely available from <https://doi.pangaea.de/10.1594/PANGAEA.893109> and can readily be ingested for the atmospheric correction of Landsat imagery with FORCE (freely available from <http://force.feut.de>). If available, daily water vapor values are used for the atmospheric correction. If not, a monthly, long-term water vapor climatology is used instead.

We systematically assessed the impact on surface reflectance when using this water vapor climatology. A global random sample of Landsat imagery was processed twice (i) using daily water vapor values, i.e., the reference, and (ii) using the water vapor climatology, i.e., the estimate. APU metrics were generated for testing the effect of using the climatology (i) for the different Landsat sensors, (ii) in different parts of the world, and (iii) between seasons.

- (1) Results varied greatly between sensors and bands. Whereas visual bands showed only small deviations, larger differences were observed for the NIR and SWIR bands. However, all sensors and bands were well below specification, which means that the fallback usage of the climatology is generally a sound strategy. Still, some considerations need to be taken into account:
  - Highest uncertainty and bias are found for Landsat 5, which has important implications for the continuity of long-term analyses where the sudden availability of daily water vapor needs to be factored in;
  - The uncertainty and bias were progressively reduced from Landsat 5 over Landsat 7 to Landsat 8 (with the exception of SWIR2 uncertainty) up to the point that the use of the climatology only marginally influences surface reflectance for Landsat 8's NIR and SWIR1;
  - Uncertainty in SWIR2 remains similar for all Landsat sensors, which implies that the water vapor database still needs to be updated regularly instead of using a static fallback climatology for all upcoming Landsat 8 acquisitions.
- (2) Some general conclusions were drawn from the test between climate zones:

- The use of the water vapor climatology is most accurate and hence largely uncritical in dry climates;
  - With the exception of the temperate zone, bias gradually increases from the most Northern/Southern dry zones to the humid tropics;
  - Uncertainty increases from dry to temperate climates but decreases towards the tropics.
- (3) Results of the test between seasons indicate that:
- The uncertainty of using the water vapor climatology is smallest during seasons with low variability in atmospheric water vapor, e.g., all-year round in the tropics;
  - Uncertainty is highest in months where the atmospheric conditions progress from a dry to a wet season (and vice versa), e.g., during the onset of the monsoon in India. To mitigate this, it might be useful to use weekly or daily long-term averages instead of monthly data and to factor in trends for interpolating.

In summary, the generation and public distribution of the water vapor database greatly increases the readiness in generating global, long-term ARD from Landsat. The tests performed in this study help in the understanding and quantification of (i) which sensor, (ii) which band, (iii) which climate, and/or (iv) which season are most or least affected by the strategy of using a MODIS-derived water vapor database with a fallback climatology.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2072-4292/11/3/257/s1>, Video S1: Long-term mean water vapor per month; Video S2: Long-term variability of water vapor per month; Video S3: Long-term trend of monthly averages.

**Author Contributions:** Conceptualization, D.F.; data curation, D.F., and M.S.; investigation, D.F., M.S., and P.H.; methodology, D.F.; resources, M.S. and P.H.; software, D.F.; validation, D.F.; writing—original draft, D.F.; writing—review and editing, M.S. and P.H.

**Funding:** This publication was financially supported by Geo.X, the Research Network for Geosciences in Berlin and Potsdam under grant number SO\_087\_GeoX.

**Acknowledgments:** Landsat and SRTM data courtesy of the U.S. Geological Survey. The Terra and Aqua MODIS Total Precipitable Water Vapor 5-Min L2 Swath 1km and 5km data sets were acquired from the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC), located in the Goddard Space Flight Center in Greenbelt, Maryland (<https://ladsweb.nascom.nasa.gov/>). This research contributes to the Landsat Science Team 2018–2023. We acknowledge support by the German Research Foundation (DFG) and the Open Access Publication Fund of Humboldt-Universität zu Berlin.

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

## References

1. Vogelmann, J.E.; Gallant, A.L.; Shi, H.; Zhu, Z. Perspectives on monitoring gradual change across the continuity of Landsat sensors using time-series data. *Remote Sens. Environ.* **2016**, *185*, 258–270. [[CrossRef](#)]
2. 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. [[CrossRef](#)] [[PubMed](#)]
3. Zhang, G.; Yao, T.; Chen, W.; Zheng, G.; Shum, C.K.; Yang, K.; Piao, S.; Sheng, Y.; Yi, S.; Li, J.; et al. Regional differences of lake evolution across China during 1960s–2015 and its natural and anthropogenic causes. *Remote Sens. Environ.* **2019**, *221*, 386–404. [[CrossRef](#)]
4. Irons, J.R.; Dwyer, J.L.; Barsi, J.A. The next Landsat satellite: The Landsat Data Continuity Mission. *Remote Sens. Environ.* **2012**, *122*, 11–21. [[CrossRef](#)]
5. Mishra, N.; Haque, M.; Leigh, L.; Aaron, D.; Helder, D.; Markham, B. Radiometric Cross Calibration of Landsat 8 Operational Land Imager (OLI) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+). *Remote Sens.* **2014**, *6*, 12619–12638. [[CrossRef](#)]
6. Helder, D.; Thome, K.J.; Mishra, N.; Chander, G.; Xiaoxiong, X.; Angal, A.; Taeyoung, C. Absolute Radiometric Calibration of Landsat Using a Pseudo Invariant Calibration Site. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1360–1369. [[CrossRef](#)]

7. Woodcock, C.E.; Allen, R.; Anderson, M.; Belward, A.; Bindschadler, R.; Cohen, W.; Gao, F.; Goward, S.N.; Helder, D.; Helmer, E. Free Access to Landsat Imagery. *Science* **2008**, *320*, 1011a. [[CrossRef](#)]
8. Danaher, T.; Scarth, P.; Armston, J.; Collett, L.; Kitchen, J.; Gillingham, S. Remote Sensing of Tree–Grass Systems: The Eastern Australian Woodlands. In *Ecosystem Function in Savannas: Measurement and Modeling at Landscape to Global Scales*; Hill, M.J., Hanan, N.P., Eds.; CRC Press: Boca Raton, FL, USA, 2010; pp. 175–194.
9. Hansen, M.C.; Loveland, T.R. A Review of Large Area Monitoring of Land Cover Change using Landsat Data. *Remote Sens. Environ.* **2012**, *122*, 66–74. [[CrossRef](#)]
10. Kennedy, R.E.; Yang, Z.; Cohen, W.B. Detecting Trends in Forest Disturbance and Recovery using Yearly Landsat Time Series: 1. LandTrendr—Temporal Segmentation Algorithms. *Remote Sens. Environ.* **2010**, *114*, 2897–2910. [[CrossRef](#)]
11. Dwyer, J.; Roy, D.; Sauer, B.; Jenkerson, C.; Zhang, H.; Lyburner, L. Analysis Ready Data: Enabling Analysis of the Landsat Archive. *Remote Sens.* **2018**, *10*, 1363. [[CrossRef](#)]
12. Frantz, D. *FORCE v. 2.0—Technical User Guide*. Available online: [https://www.researchgate.net/publication/328094593\\_FORCE\\_v\\_20\\_-\\_Technical\\_User\\_Guide](https://www.researchgate.net/publication/328094593_FORCE_v_20_-_Technical_User_Guide) (accessed on 28 January 2018).
13. Tanré, D.; Herman, M.; Deschamps, P.Y.; de Leffe, A. Atmospheric Modeling for Space Measurements of Ground Reflectances, Including Bidirectional Properties. *Appl. Opt.* **1979**, *18*, 3587–3594. [[CrossRef](#)] [[PubMed](#)]
14. Frantz, D.; Röder, A.; Stellmes, M.; Hill, J. An Operational Radiometric Landsat Preprocessing Framework for Large-Area Time Series Applications. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3928–3943. [[CrossRef](#)]
15. Doxani, G.; Vermote, E.; Roger, J.-C.; Gascon, F.; Adriaensen, S.; Frantz, D.; Hagolle, O.; Hollstein, A.; Kirches, G.; Li, F.; et al. Atmospheric Correction Inter-Comparison Exercise. *Remote Sens.* **2018**, *10*, 352. [[CrossRef](#)]
16. Schläpfer, D.; Borel, C.C.; Keller, J.; Itten, K.I. Atmospheric Precorrected Differential Absorption Technique to Retrieve Columnar Water Vapor. *Remote Sens. Environ.* **1998**, *65*, 353–366. [[CrossRef](#)]
17. Gao, B.-C.; Kaufman, Y.J. Water Vapor Retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) Near-Infrared Channels. *J. Geophys. Res. Atmos.* **2003**, *108*. [[CrossRef](#)]
18. 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](#)]
19. Ackerman, S.A.; Holz, R.E.; Frey, R.; Eloranta, E.W.; Maddux, B.C.; McGill, M. Cloud Detection with MODIS. Part II: Validation. *J. Atmos. Ocean. Technol.* **2008**, *25*, 1073–1086. [[CrossRef](#)]
20. Frantz, D.; Stellmes, M. Water vapor database for atmospheric correction of Landsat imagery. *PANGAEA* **2018**. [[CrossRef](#)]
21. Reynolds, A.P.; Richards, G.; de la Iglesia, B.; Rayward-Smith, V.J. Clustering Rules: A Comparison of Partitioning and Hierarchical Clustering Algorithms. *J. Math. Model. Algorithms* **2006**, *5*, 475–504. [[CrossRef](#)]
22. Maechler, M.; Rousseeuw, P.; Struyf, A.; Hubert, M.; Hornik, K.; Studer, M.; Roudier, P.; Gonzalez, J.; Kozłowski, K. *Cluster: Cluster Analysis Basics and Extensions*. Available online: <https://cran.r-project.org/web/packages/cluster/cluster.pdf> (accessed on 28 January 2018).
23. 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](#)]
24. Justice, C.O.; Vermote, E.; Townshend, J.R.G.; DeFries, R.; Roy, D.P.; Hall, D.K.; Salomonson, V.V.; Privette, J.L.; Riggs, G.; Strahler, A.; et al. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land Remote Sensing for Global Change Research. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1228–1249. [[CrossRef](#)]
25. Holben, B.N.; Eck, T.F.; Slutsker, I.; Tanré, D.; Buis, J.P.; Setzer, A.; Vermote, E.; Reagan, J.A.; Kaufman, Y.J.; Nakajima, T.; et al. AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization. *Remote Sens. Environ.* **1998**, *66*, 1–16. [[CrossRef](#)]
26. Vermote, E.F.; Kotchenova, S. Atmospheric correction for the monitoring of land surfaces. *J. Geophys. Res. Atmos.* **2008**, *113*. [[CrossRef](#)]
27. Claverie, M.; Vermote, E.F.; Franch, B.; Masek, J.G. Evaluation of the Landsat-5 TM and Landsat-7 ETM+ surface reflectance products. *Remote Sens. Environ.* **2015**, *169*, 390–403. [[CrossRef](#)]
28. NASA. MODIS/Terra Data Outages. Available online: [https://modaps.modaps.eosdis.nasa.gov/services/production/outages\\_terra.html](https://modaps.modaps.eosdis.nasa.gov/services/production/outages_terra.html) (accessed on 5 October 2018).
29. NASA. MODIS/Aqua Data Outages. Available online: [https://modaps.modaps.eosdis.nasa.gov/services/production/outages\\_aqua.html](https://modaps.modaps.eosdis.nasa.gov/services/production/outages_aqua.html) (accessed on 5 October 2018).

30. USGS. Spectral Characteristics Viewer. Available online: <https://landsat.usgs.gov/spectral-characteristics-viewer> (accessed on 5 October 2018).
31. Gordon, I.E.; Rothman, L.S.; Hill, C.; Kochanov, R.V.; Tan, Y.; Bernath, P.F.; Birk, M.; Boudon, V.; Campargue, A.; Chance, K.V.; et al. The HITRAN2016 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **2017**, *203*, 3–69. [[CrossRef](#)]
32. Vermote, E.F.; Saleous, N. Operational Atmospheric Correction of MODIS Visible to Middle Infrared Land Surface Data in the Case of an Infinite Lambertian Target. In *Earth Science Satellite Remote Sensing: Vol. 1: Science and Instruments*; Qu, J.J., Gao, W., Kafatos, M., Murphy, R.E., Salomonson, V.V., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 123–153.
33. Gao, F.; Masek, J.; Schwaller, M.; Hall, F. On the Blending of the Landsat and MODIS Surface Reflectance: Predicting Daily Landsat Surface Reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218. [[CrossRef](#)]



© 2019 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/>).