

Article

Preliminary Evaluation of the Consistency of Landsat 8 and Sentinel-2 Time Series Products in An Urban Area—An Example in Beijing, China

Zhen Nie ¹ , Karen Kie Yan Chan ² and Bing Xu ^{2,3,4,*}

¹ State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, China; niezhen9876@mail.bnu.edu.cn

² Ministry of Education Key Laboratory for Earth System Modelling, Center for Earth System Science, Tsinghua University, Beijing 100084, China; cqel15@mails.tsinghua.edu.cn

³ Joint Center for Global Change Studies, Beijing 100875, China

⁴ Center for Healthy Cities, Institute for China Sustainable Urbanization, Tsinghua University, Beijing 100085, China

* Correspondence: bingxu@tsinghua.edu.cn; Tel.: +86-010-6279-0189

Received: 8 November 2019; Accepted: 5 December 2019; Published: 10 December 2019



Abstract: Global urbanization is occurring rapidly, and numerous moderate resolution remote sensing data are being used to monitor this process. Landsat 8 OLI and Sentinel-2 MSI data are combined in many applications but few studies have focused on either urban change or consistency between these two data in time series. To evaluate the varying correlation between the two sensors in a time series, the correlation coefficient (R) and root-mean-square deviation (RMSD) of seven band pairs and three indices (NDVI, NDBI, and MNDWI) were calculated in this study and the results of the built-up area identified by IBI derived from the above three indices were compared. It was found that the correlation between the two sensors ($R > 0.8534$, $p < 0.0001$) was good in most bands but not as good for indices (in half of the results, $R < 0.9$). Meanwhile, the correlation of the two sensors of both bands and indices fluctuated between seasons and the comparative results of built-up area identification between the two data are relative to this variation. Therefore, when the OLI and MSI data are used in future collaboration applications, the data and threshold selection should consider the consistency and the fluctuation between the two data, especially in both time series studies and urban detection.

Keywords: Landsat 8 OLI; Sentinel-2 MSI; comparison; urban; time-series

1. Introduction

Urban areas, though occupying a small percentage of the Earth's land surface, have an important impact on economic contribution, population services, energy consumption, and global environmental change [1]. With rapid urbanization worldwide, natural vegetation inside and around cities has been replaced by buildings and impervious surfaces [2]. This unprecedented progress has brought many challenges in many fields, such as public health [3], the urban heat island effect and climate change contribution [4,5], and phenology impact and ecological services [6,7]. Fortunately, the development of remote sensing technology gives us a better way to observe urban development and many different data, such as luminous data [8,9], optical remote sensing data [2–10], and synthetic aperture radar (SAR) data [11,12] were used alone or in combination [4,13,14] to detect this change.

A spatial scale of 10–70 m was proposed as appropriate for urban research [15], so Landsat series satellites and SPOT satellites have always been important optical remote sensing data in urban detection [13,14]. Landsat 8 (L8) and Sentinel-2 (S2) are the successors of these two important series [16,17], and an average revisit time of 2.9 days may be achieved through the combination of Landsat 8

and Sentinel-2A/B data [18]. Furthermore, due to the free availability and similar spectral bands (Figure 1), the radiometric data products of Landsat 8 Operational Land Imager (OLI) and Sentinel-2 Multi-Spectral Instrument (MSI) data are being combined into many synergistic applications, such as the mapping of crops [19,20], the measurement of aquatic systems and water [21,22] and the detection of plastic-covered greenhouses [23].

Many studies have compared the consistency of OLI and MSI products and varying results have been obtained in the many comparisons of the two sensors. Their results of the slope of the MSI-to-OLI regression function have ranged from 0.9912 to 1.0524 for the surface reflectance in the green band, while correlation coefficients have been obtained as $R > 0.9916$ ($R^2 > 0.9832$) in five study areas [24] and $R = 0.9577$ ($R^2 = 0.9180$) [25]. For top of atmosphere (TOA) reflectance in the red band [22–26], the root-mean-square difference (RMSD) has ranged from 0.0012 to 0.0163. Other research also found Sentinel-2A matched with Landsat 7 and 8 for surface reflectance data with small systematic differences ranging from 1% to 9% on different bands [27]. Furthermore, the comparison between OLI and MSI data was also processed when these two data were combined in various research fields. For example, the reflection difference between OLI and MSI bands was less than 0.1 overall and the NDVI and NDWI of these two data were generally consistent in the field of agriculture [28]. A study showed that Sentinel-2A and Landsat 8 are consistent in TOA reflectance products and remote-sensing reflectance (Rrs) products within 1% and 6%, respectively in monitoring water quality [22]. Although the results showed minor differences, almost all studies have observed good agreement between the products.

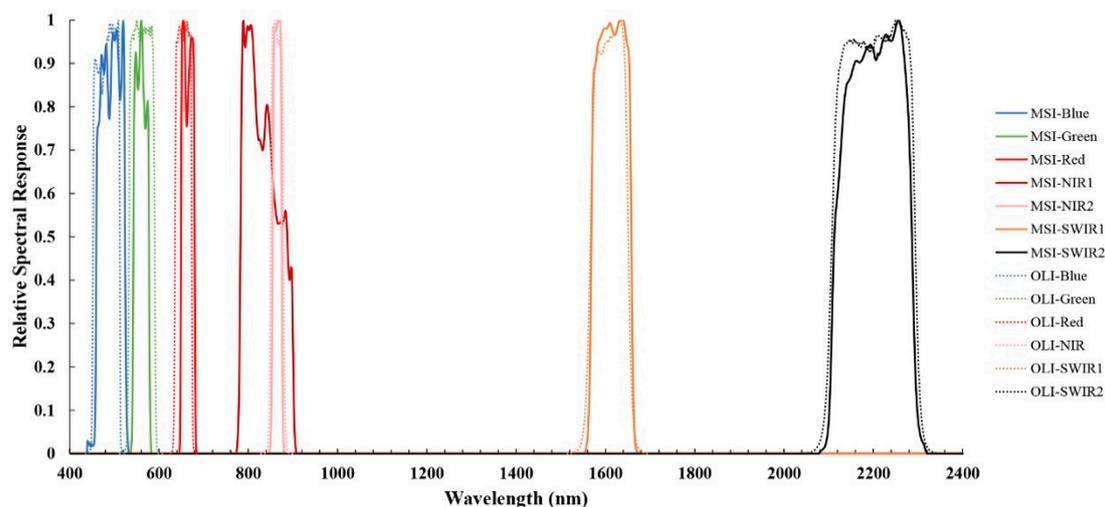


Figure 1. Relative spectral response functions of corresponding bands of Landsat 8 OLI [29] and Sentinel-2 MSI [30] in this study.

Sentinel-2 and Landsat 8 data have been well combined to observe the urban area [31] and the temporal consistency of these two data has been evaluated for the Igneada Longos Forest [32]. However, there is still a lack of focus on the temporal consistency between these two datasets when applied in an urban area. These deficiencies suggest a need for additional research on data consistency, since many applications are based on the time series data [33,34] and land surface changes in an urban area can be rapid [13]. Therefore, the surface reflectance from Band 2 to Band 7 (Landsat heritage bands of the OLI) and three common indices in urban study (normalized difference vegetation index (NDVI), normalized difference built-up index (NDBI), and modified normalized difference water index (MNDWI)) with the corresponding MSI bands and indices were investigated.

2. Materials and Methods

2.1. Data Source

Landsat 8 carrying the OLI was launched in 2013 to obtain multispectral data in eight bands with a 30 m resolution and one panchromatic band with a 15 m resolution with a revisiting time of 16 days. Its high-quality data allow the continuation of the Landsat Earth observation program [35]. The Sentinel-2 mission comprises of two satellites equipped with identical MSIs; i.e., Sentinel-2A launched in 2015, and Sentinel-2B launched in 2017. The mission obtains data in 13 bands at different spatial resolutions (i.e., 10, 20, and 60 m) with a revisiting time of 5 days for both satellites. Sentinel-2 data are used throughout the paper to refer to both Sentinel-2A and 2B as they carry identical onboard instruments [17].

To eliminate atmospheric effects [16], the surface reflectance (SR) product of OLI band 2 to 7 and MSI band 2–4, 8, 8a, 11, and 12 were tested in this study (Table 1). Four OLI SR data and the two latest MSI SR were downloaded from Google Earth Engine (GEE) because previous MSI SR data is only uploaded to the GEE dataset (https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR). The two early MSI SR data were produced from the L1C product downloaded from the Copernicus Open Access Hub (SciHub) by its exclusive atmospheric correction module Sen2Cor (<http://step.esa.int/main/third-party-plugins-2/sen2cor/>).

Table 1. Information of spectral bands considered in the comparison between Sentinel-2 MSI and Landsat-8 OLI [26]. CW means Central Wavelength and the unit is nm.

	Sentinel-2 MSI			Landsat 8 OLI		
	Band (Resolution)	CW	Bandwidth (nm)	Band (Resolution)	CW	Bandwidth (nm)
Blue	B2 (10 m)	490	65	B2 (30 m)	482	65
Green	B3 (10 m)	560	35	B3 (30 m)	561	60
Red	B4 (10 m)	665	30	B4 (30 m)	655	40
NIR1	B8 (10 m)	842	115			
NIR2	B8a (20 m)	865	20	B5 (30 m)	865	30
SWIR1	B11 (20 m)	1610	90	B6 (30 m)	1609	90
SWIR2	B12 (20 m)	2190	180	B7 (30 m)	2200	175

2.2. Study Area

Beijing is not only the capital of China but also an international metropolis with rapid urban expansion and population growth [2]. The megacity, located on the northern edge of the north China plain, is affected by temperate monsoon climate, with the majority of the annual precipitation being concentrated in summer [36]. This climatic characteristic influenced the data pair in summer, as two data one day apart were chosen for this season (Table 2).

Table 2. Information of image pairs in the study area. The scene center time was converted to the local time—China Standard Time (GMT+8 hours).

Seasons	Date	Product	Data Sources	Scene Center Time	Cloud Cover
Spring	20180408	L8 L2	GEE	10:53	1.93
		S2B L2A	Sen2Cor	11:11	0.0081
Summer	20190817	L8 L2	GEE	10:53	3.84
		S2A L2A		11:14	6.6162
Autumn	20171030	L8 L2	GEE	10:53	0.07
		S2B L2A	Sen2Cor	11:08	0.3537
Early Winter	20181204	L8 L2	GEE	10:53	1.65
		S2B L2A		11:11	5.4285

The Region of interest (ROI) is the overlap between L8 OLI and S2 MSI divided into $60 \text{ km} \times 60 \text{ km}$ square (Figure 2A). Figure 2B shows that this area covers the most built-up areas of Beijing including the entire district of Dongcheng, Xicheng, Chaoyang, most of Haidian, Shijingshan, Fengtai, and the dense Changping, Shunyi, Tongzhou, and Daxing. The details of the four data pairs are presented in Figure 2 and Table 2. In addition, the dates of the four pairs roughly fall in spring, summer, autumn, and early winter. In order to highlight the consistency changes due to the progression of the seasons, all the tables and figures were ordered in accordance to the seasons and not in accordance with the date of data acquisition.

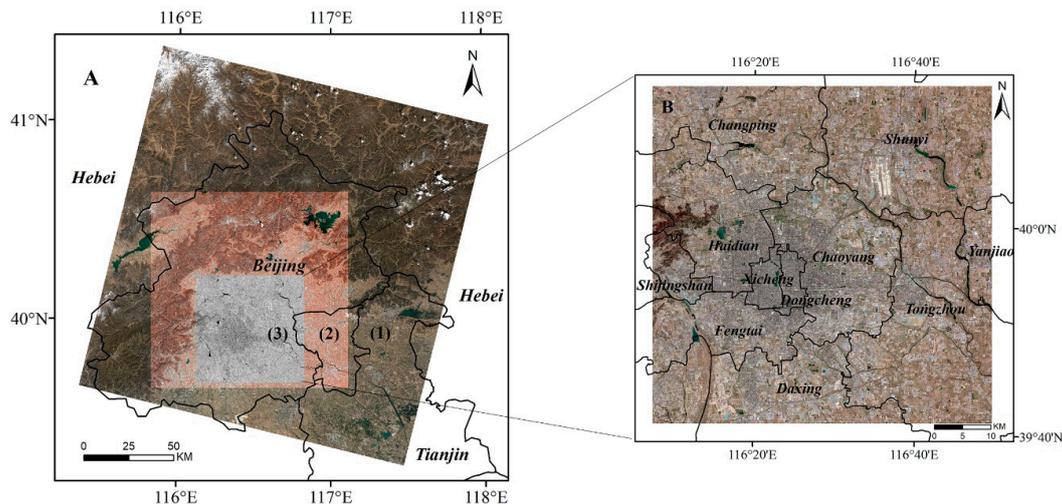


Figure 2. Location of the study areas and illustrative scene pairs from 8 April 2018. In Part A, (1) is Landsat 8 OLI data shown by true color (red, green, blue), (2) is Sentinel-2 MSI data shown by false color (NIR, green, blue), and (3) is the test area shown by the blue band. Part B shows the administrative region covered by ROI.

2.3. Data Pre-processing

First pre-processing is cloud and cloud shadow removal. This process was completed before the Landsat 8 SR data were uploaded on the GEE dataset (https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C01_T1_SR). But the cloud mask product of Sentinel-2A L1C is not currently reliable and Sen2Cor (Version 2.5.0) was used to produce the cloud and shadow mask for Sentinel-2 MSI data [37].

Geometric co-registration was performed in the second step to rectify any misregistration of Landsat 8 OLI and Sentinel-2 MSI imagery [38]. An automated precise registration and orthorectification package (AROP) was used to complete the geometric co-registration between OLI and MSI data [39]. The package reports the root-mean-square deviations of 0.217, 0.283, 0.246, and 0.255 OLI pixel for spring, summer, autumn, and winter.

Finally, MSI images (with resolutions of 10 and 20 m) were downsampled to the same resolution of OLI data (30 m) using a weighted average algorithm. The MSI 30 m pixel is derived by averaging nine 10 m values within a $30 \times 30 \text{ m}$ square, while averaging the four 20 m pixels overlapping the intended 30 m pixel with area-based weights of 4/9, 2/9, 2/9, and 1/9 [24].

2.4. Statistical Analysis

Built-up areas were an important object in urban surface observation [13] and three indices were frequently used for this purpose, including normalized difference vegetation index (NDVI), normalized difference built-up index (NDBI), and modification normalized difference built-up index (MNDWI) [2–40]. NDVI was an important index to detect vegetation and has also been widely used for OLI and MSI multispectral consistency analysis [24,25]. NDBI was first introduced by Zha et al. [41]

and MNDWI is an index improved from NDWI to detect water [42]. These three indices can be calculated by Equations (1)–(3):

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} \quad (1)$$

$$NDBI = \frac{\rho_{MIR} - \rho_{NIR}}{\rho_{MIR} + \rho_{NIR}} \quad (2)$$

$$MNDWI = \frac{\rho_{Green} - \rho_{MIR}}{\rho_{Green} + \rho_{MIR}} \quad (3)$$

where ρ_i is the reflectance of i band. The corresponding band can be found in Table 1, as Band 2 of OLI and MSI belong Green Band, Band 3 belong Red and Band 11 of MSI and Band 6 belong MIR Band. There are two Sentinel-2 MSI NIR bands and the Sentinel-2A NDVI was derived using the narrower MSI NIR band 8A (855–875 nm) throughout the paper as it is more comparable with the Landsat-8 OLI NIR band 5 (851–879 nm) [37].

To assess the consistency of surface reflectance data between the Landsat 8 OLI and Sentinel-2 MSI, ordinary least-squares (OLS) regression functions between each band and index pair, the Pearson correlation coefficient (R with a p -value) and the root-mean-square deviation (RMSD) were calculated in Matlab 2018a. The present paper calculates the distribution probability at 0.1 intervals for NDVI, NDBI, and MNDWI in order to detect these differences between OLI and MSI in each season. For evaluating the magnitude of difference between OLI indices and MSI indices in each range, the relative difference $\underline{\Delta}$ of each range was calculated for each image and obtained as Equations (4) and (5):

$$P_{Sat}(i) = \frac{N_{Sat}(i)}{SUM_{Sat}} \quad (4)$$

$$\underline{\Delta} = \frac{|P_{OLI}(i) - P_{MSI}(i)|}{MIN(P_{OLI}(i), P_{MSI}(i))} \quad (5)$$

where i is located at every 0.1 interval from -1.0 to 0.9 , ($i, i+0.1$] is the interval range, and Sat is the satellite product. For each range, $P_{Sat}(i)$ and $N_{Sat}(i)$ are the percentage and the number of NDVI in the range ($i, i+0.1$] respectively, and SUM_{Sat} is the total number of NDVI in range $[-1,1]$ in the image. $P_{OLI}(i)$ and $P_{MSI}(i)$ are, respectively, the percentage of Landsat 8 and Sentinel-2 indices in the range ($i, i+0.1$], MIN is used to found the minimum of one. $\underline{\Delta}$ is a relative index used to determine the relative difference for each range.

In order to understand the consistency of L8 and S2 in practical application more intuitively, the index-based built-up index (IBI) was used in the study to extract built-up areas and conduct a comparative analysis. IBI is a built-up index derived from NDBI, MNDWI, and soil-adjusted vegetation index (SAVI), although SAVI could be replaced by with NDVI [2]. IBI is designed to identify built-up and non-built-up regions by thresholding image segment, so there is no pixel-by-pixel comparison and consistency verification of data pairs; instead, IBI partition results are used directly. The values of NDVI, NDBI, and MNDWI were added to 1 before calculating the IBI using Equation (6):

$$IBI = \frac{NDBI - (NDVI + MNDWI)/2}{NDBI + (NDVI + MNDWI)/2} \quad (6)$$

3. Results

3.1. Comparison Between Landsat 8 and Sentinel-2

Figures 3 and 4 show scatter plots of surface reflectance for each corresponding spectral band (Table 2) and three indices between the Landsat 8 OLI and Sentinel-2 MSI data. To be specific, the scenes (seasons) are arranged into columns while the bands are arranged into rows and the X-axis and Y-axis represents OLI data and MSI data, respectively, in every subplot. A 1:1 line is drawn in each plot to

describe the perfect condition of no difference between OLI and MSI, while the red dashes and green dashes represent the linear regression of the MSI data against the OLI data and the OLI data against the MSI data, respectively. In the linear equation of OLI data to MSI data, correlation coefficient R and RMSD of these two data are marked in each subplot.

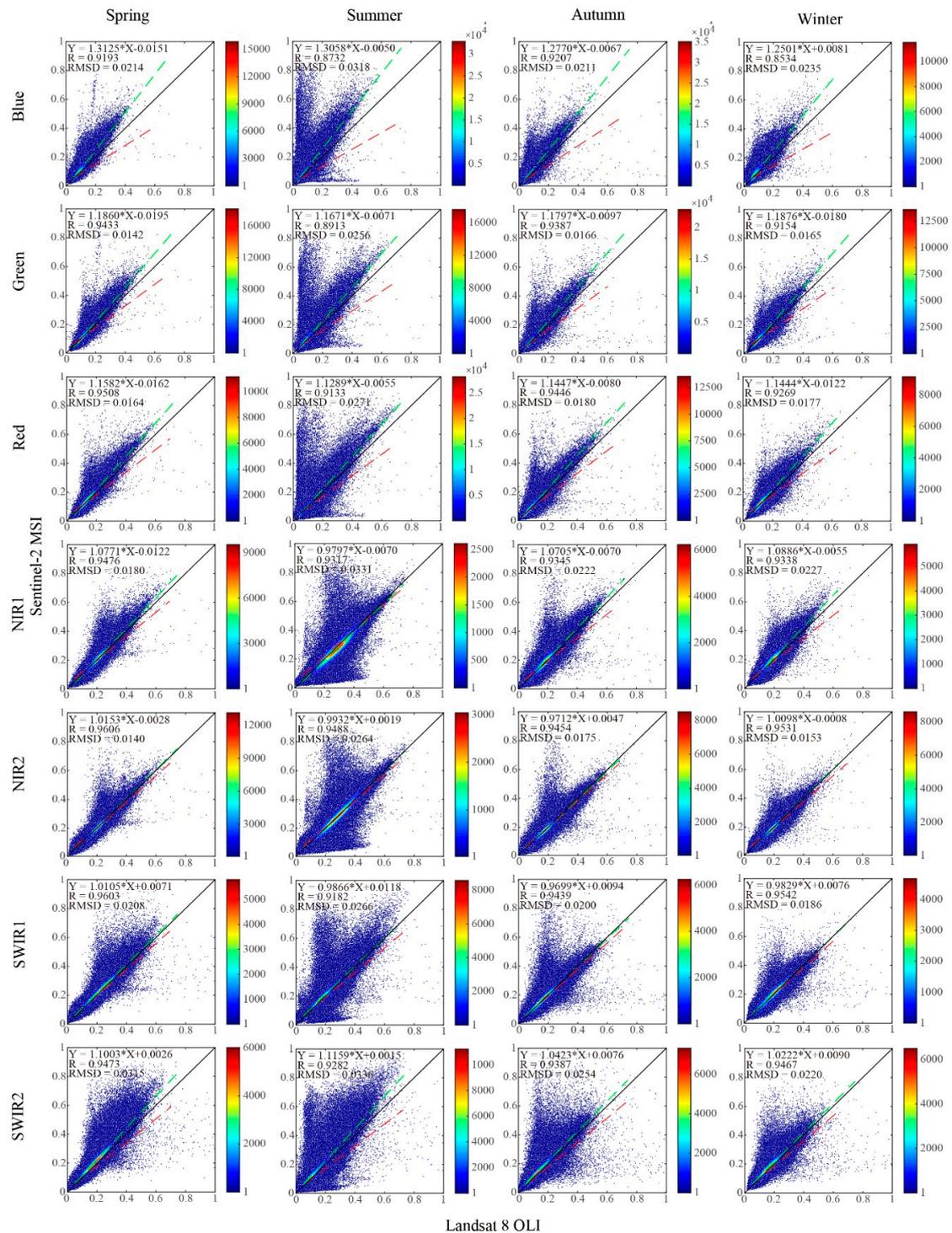


Figure 3. All scatter plots of the comparison between OLI bands and MSI bands arranged in horizontal as four seasonal result of each band and in the vertical direction as all bands in one season. The P-value for all Pearson coefficient R is <0.0001. The ordinary least squares (OLS) regression of the OLI against the MSI data are plotted in green dotted lines, the OLS regression of the MSI data against the OLI data in red dotted lines, and the identity lines (y = x) in black lines.

In Figure 3, the slopes of green dashed lines are larger than 1 in most subplots, which indicates that the value of OLI data is overall lower than the MSI data. As a key coefficient to indicate the consistency between the two data, all R are no smaller than 0.9133 (best is 0.9606), except for three pairs (0.8534 of blue bands in winter and 0.8732 of blue band and 0.8913 of the Green Band in summer). At a glance, the dispersion degree of summer data pairs is generally higher in seasonal contrast and SWIR2 band data pairs have the greatest dispersion of all the bands. This large dispersion could be quantified by RMSD, as large RMSD could be found in summer (0.0256–0.0336) and for SWIR2 (0.0220–0.0336). Only four data pairs have a RMSD larger than 0.03; of the remainder, 12 are between 0.02 and 0.03 and all the other RMSD are smaller than 0.02. Although the dispersion displayed on each subplot varies, the central position of the intensity scatter is concentrated close to the 1:1 line. With the exception of SWIR2, the density for each season within a band is different; for example, NIR1 and NIR2 are sensitive to vegetation change, so their scatter is quite different in summer than in other seasons.

To explore the comparison of OLI and MSI data in different ways, the mean R and mean RMSD from seasons or bands were calculated in Table 3. Seasonally, the highest R appears in spring and the lowest in summer, while the highest RMSD is in summer and the lowest in spring and winter. Spectrally, the NIR2 band has the highest R, while the blue band has the lowest R. The highest RMSD appears in SWIR2 and the lowest in the green band.

Table 3. Mean correlation coefficient (R) and mean root-mean-square deviation (RMSD) of bands comparison by season or bands.

	Season				Bands						
	Spring	Summer	Autumn	Winter	Blue	Green	Red	NIR1	NIR2	SWIR1	SWIR2
R	0.9470	0.9150	0.9380	0.9262	0.8917	0.9222	0.9339	0.9369	0.9520	0.9442	0.9402
RMSD	0.0195	0.0292	0.0201	0.0195	0.0245	0.0182	0.0198	0.0240	0.0183	0.0215	0.0280

3.2. NDVI, NDBI, and MNDWI Comparison

The compared results of three indices (NDVI, NDBI, and MNDWI) commonly used in urban classification for Landsat 8 OLI and Sentinel-2 MSI data are plotted in Figure 4. The linear equation of OLI to MSI indices, correlation coefficient R and RMSD of these two data are showed at the bottom right corner of each subplot. It can be found that the slopes of the green lines are larger than red lines in NDVI and MNDWI and slightly smaller than red lines in NDBI in Figure 4, which indicates that the value of OLI NDVI and MNDWI are overall smaller than the MSI ones, with the opposite result in NDBI. Generally speaking, in the comparison of the three indexes, the dispersion of scatter plots are relatively large, especially in NDVI and MNDWI. This large degree of dispersion is also reflected in RMSD and R, as the value of RMSD is relatively large (0.0350–0.0812) and only half of the correlation coefficients (R) are larger than 0.9 which shows that the consistency of indices between OLI and MSI was not as good as band comparison.

When the densities of the scatter plots are compared in Figure 4, the variations in density for the different seasons are not as high as those for the different spectrums. Meanwhile, the major usage of the indices could be illustrated in the distributional characteristics of the scatter, such as the concentration of the NDVI scatter in the high-value area in summer and near 0 in winter. The specific index distribution was calculated and is shown in Figure 5.

Like Table 3, Table 4 tabulated the mean R and mean RMSD for all the seasons and indices. From the perspective of seasons, the largest RMSD occurred in the summer and the smallest in spring. Seasonally, the largest R was in summer and the smallest R in winter. In terms of the indices, the R of NDVI is the highest, while the R of MNDWI is the lowest. MNDWI has the smallest RMSD, while MNDWI has the largest RMSD. The average relative difference (ARD) was calculated for each compared pair to illustrate the overall situation of the difference between the OLI and MSI indexes. The specific values of all ARD are shown in the Figure 5 and the mean ARD for the seasons and indices are calculated in Table 4.

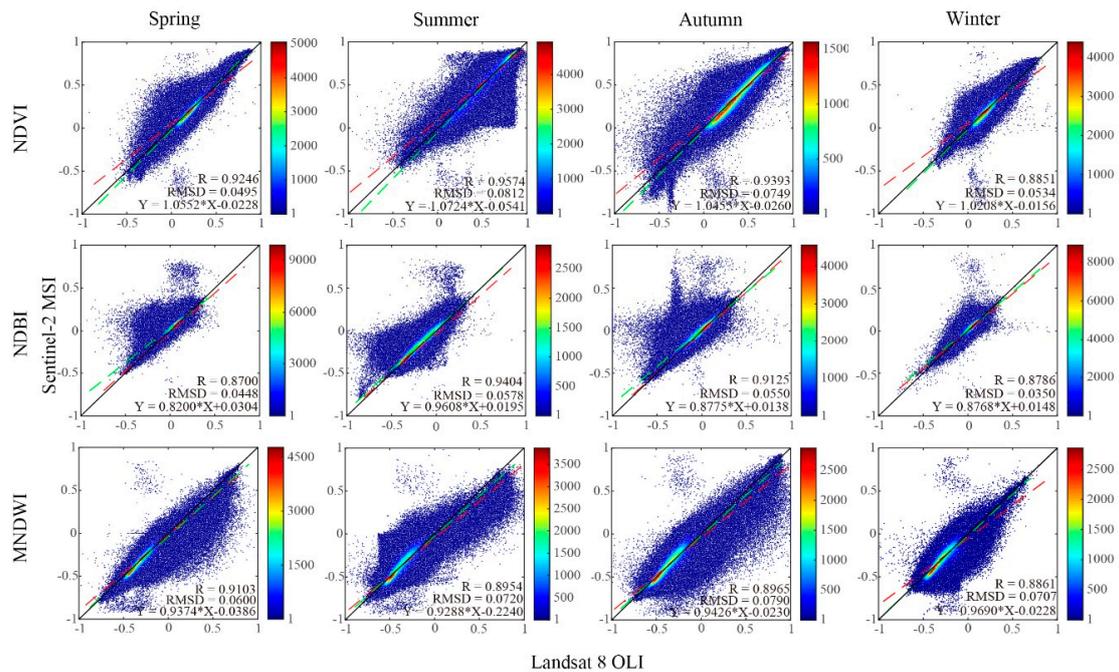


Figure 4. All the scatter plots of the comparison between OLI indices and MSI indices arranged in horizontal as four seasonal result of three indices (NDVI, NDBI, and MNDWI) and in the vertical direction as three indices in one season. The P-value for all Pearson coefficient R is <0.0001. The ordinary least squares (OLS) regression of the OLI indices against the MSI indices are plotted in the green dotted lines, the OLS regression of the MSI indices against the OLI indices in the red dotted lines, and the identity lines ($y = x$) in black lines.

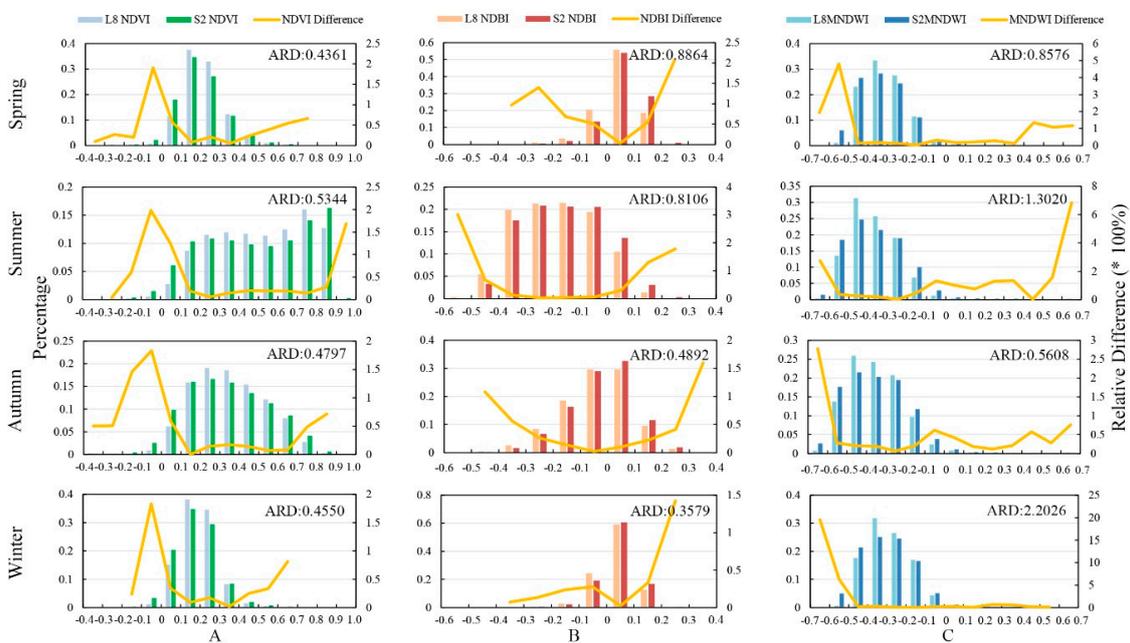


Figure 5. The percentage of frequency in equal interval index ranges. A, B and C indicate NDVI, NDBI, and MNDWI. The relative difference is calculated except abnormal value. Every statistic range includes the upper bound but not the lower one, for example, the count of 0.5–0.6 included NDVI = 0.6 pixels and no NDVI = 0.5 pixels.

Table 4. Mean R, mean RMSD and mean average relative difference (ARD) for all seasons and indices.

	Season				Indices		
	Spring	Summer	Autumn	Winter	NDVI	NDBI	MNDWI
R	0.9016	0.9311	0.9161	0.8833	0.9266	0.9004	0.8971
RMSD	0.0514	0.0703	0.0696	0.0530	0.0648	0.0482	0.0704
ARD	0.7267	0.8823	0.5099	1.0052	0.4763	0.6360	1.2310

For estimating the correlation and statistics for all R and RMSD (Figures 3 and 4, Tables 3 and 4), Table 5 highlights the R, RMSD, and Rank of the data pairs with the two highest R and the two lowest R in addition to the highest and lowest mean R from the various seasons and bands. It is evident that the statistics in spectral bands showed high correlation coefficient R are always associated with low RMSD and low R with high RMSD, as the data pairs for the highest R and the highest mean R of season both have the highest R and the lowest RMSD within their class. However, in the indices statistics, there does not appear to be an association between R and RMSD.

Table 5. Statistical highlights from Figures 3 and 4 and Tables 3 and 4. Rank is an order from low to high in RMSD.

Spectral Bands								
Pair	Two Highest R		Two Lowest R		Mean R of Season		Mean R of bands	
					Highest	Lowest	Highest	Lowest
	NIR2 in spring	SWIR1 in spring	blue in winter	blue in summer	spring	winter	NIR2	blue
R	0.9606	0.9603	0.8534	0.8732	0.9470	0.9150	0.9520	0.8917
RMSD	0.0140	0.0208	0.0235	0.0318	0.0195	0.0292	0.0183	0.2450
Rank	1/28	13/28	19/28	26/28	1/4	4/4	2/7	6/7
Indices								
Pair	Two Highest R		Two Lowest R		Mean R of Season		Mean R of indices	
					Highest	Lowest	Highest	Lowest
	NDVI in summer	NDBI in summer	NDBI in spring	blue in summer	summer	winter	NDVI	MNDWI
R	0.9574	0.9404	0.8700	0.8786	0.9311	0.8833	0.9266	0.8971
RMSD	0.0812	0.0578	0.0448	0.0350	0.0703	0.0530	0.0648	0.0704
Rank	12/12	6/12	2/12	1/12	3/4	2/4	2/3	3/3

In order to investigate the difference of NDVI, NDBI, and MNDWI between Landsat 8 and Sentinel-2 data, histograms were used as a compared statistical method. As shown in Figure 5, the column represented the percentage of frequency in equal interval ranges of NDVI, NDBI, and MNDWI, while the relative difference is shown by the yellow line and ARD is marked in the top right corner of each subplot. Obviously, MNDWI has the most stable distribution of the three indexes and its distribution results are very similar in the four seasons. Although the distribution of NDVI and NDBI were quite similar, the changes in distribution with seasonal change was found in Figure 5. Both NDBI and MNDWI errors exist at the maximum and minimum values and the relative differences in the intermediate range are small. In the statistical comparison of NDVI in four seasons, the error also appeared at the maximum value and the value range from -0.1 to 0.1 . In terms of seasons, ARD was the largest in summer and the smallest in autumn. In terms of indices, MNDWI had the largest ARD and NDVI had the smallest.

3.3. Build-up Area Identification by IBI Between OLI and MSI

Table 6 shows the built-up area identification and mutual comparison results of Landsat 8 and Sentinel-2 data in four seasons under the same threshold (0.013) [2]. Because the primary purpose of

the article is to compare the consistency of the two data, the statistics of the extraction results is showed in percentage of total area. The variations in the percentage of built-up areas can be found between different seasons of the same dataset and between different datasets in the same season. For example, the maximum percentage extracted from Landsat 8 data is 43.79% in winter with a minimum of 31.75% in summer, while the maximum from Sentinel-2 data is 54.50% in winter and the minimum is 36.77% in summer. In each season, the area of built-up area identified from MSI data is larger than that identified from OLI data, and the maximum difference is 12.57% in spring and the minimum is 5.02% in summer. Although the extraction results of Landsat 8 and Sentinel-2 were different in different seasons, the total extraction results of the same built-up areas and same non-built-up areas were still relatively high, with the highest at 91.54% in summer and the worst at 84.81% in winter. It is the case that the built-up area in S2 is identified as non-built area in L8 more than the reverse.

Table 6. Statistics of the built-up area extraction by OLI IBI and MSI IBI in four seasons. Same built-up and same non-built-up means that this area was identified as identical classes in both OLI and MSI data respectively. Built-up to non-built-up means that this area was just identified as built-up in this data. All data are a percentage and multiplied by 100%.

Season	Data	Built-Up	Non-Built-Up	Same Built-Up	Same Non-Built-Up	Total Same	Built-up to Non-Built-Up
Spring	L8 OLI	41.63	38.37	40.49	44.66	85.15	1.14
	S2 MSI	54.20	45.80				13.71
Summer	L8 OLI	31.75	69.25	30.03	61.51	91.54	1.72
	S2 MSI	36.77	63.23				6.74
Autumn	L8 OLI	39.35	60.65	36.64	53.16	89.80	2.70
	S2 MSI	44.13	55.87				7.49
Winter	L8 OLI	43.79	56.21	40.05	44.76	84.81	3.74
	S2 MSI	51.50	48.50				11.45

In order to test the robustness and sensitivity of this threshold, IBI thresholds from 0 to 0.025 with an interval of 0.005 were used to identify built-up areas and the result is plotted in Figure 6. Obviously, the percentage of identical classes is stable with a slight increase in summer, autumn and steadily decreased in spring and winter throughout the threshold range. However, the similarity between OLI and MSI data in summer was always superior to the other seasons.

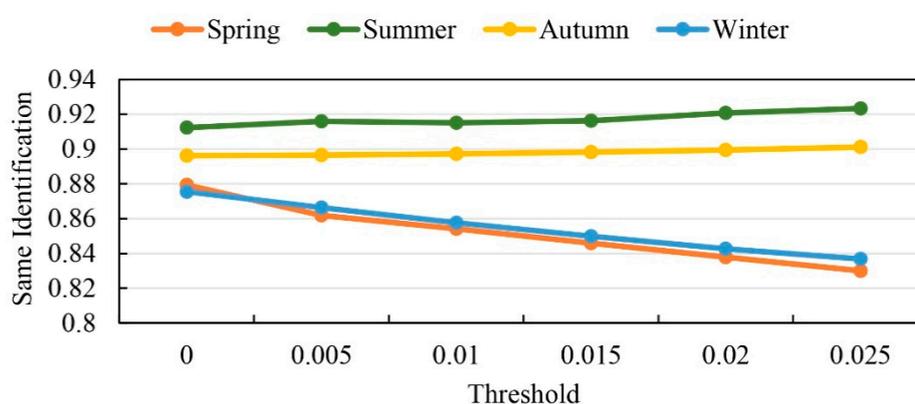


Figure 6. The percentage of same identification of built-up and non-built-up area at different thresholds.

To locate the differences between L8 and S2 on built-up areas, the similarities and differences of the built-up extraction between Landsat 8 and Sentinel-2 in summer data pairs are shown in Figure 7. Figure 7A is the whole scene and the identical built-up and non-built-up area are shown in black and white, respectively. The built-up area in S2 data that was classified as non-built-up areas in L8 is marked as green and the opposing as red. For a better clarity, Figure 7B (1–4) enlarged four selected

locations: Beijing Capital International Airport (impervious area), Beijing Olympic Park, and Temple of Heaven (green spaces) and Beijing Economic-Technological Development Area (multi-surface).

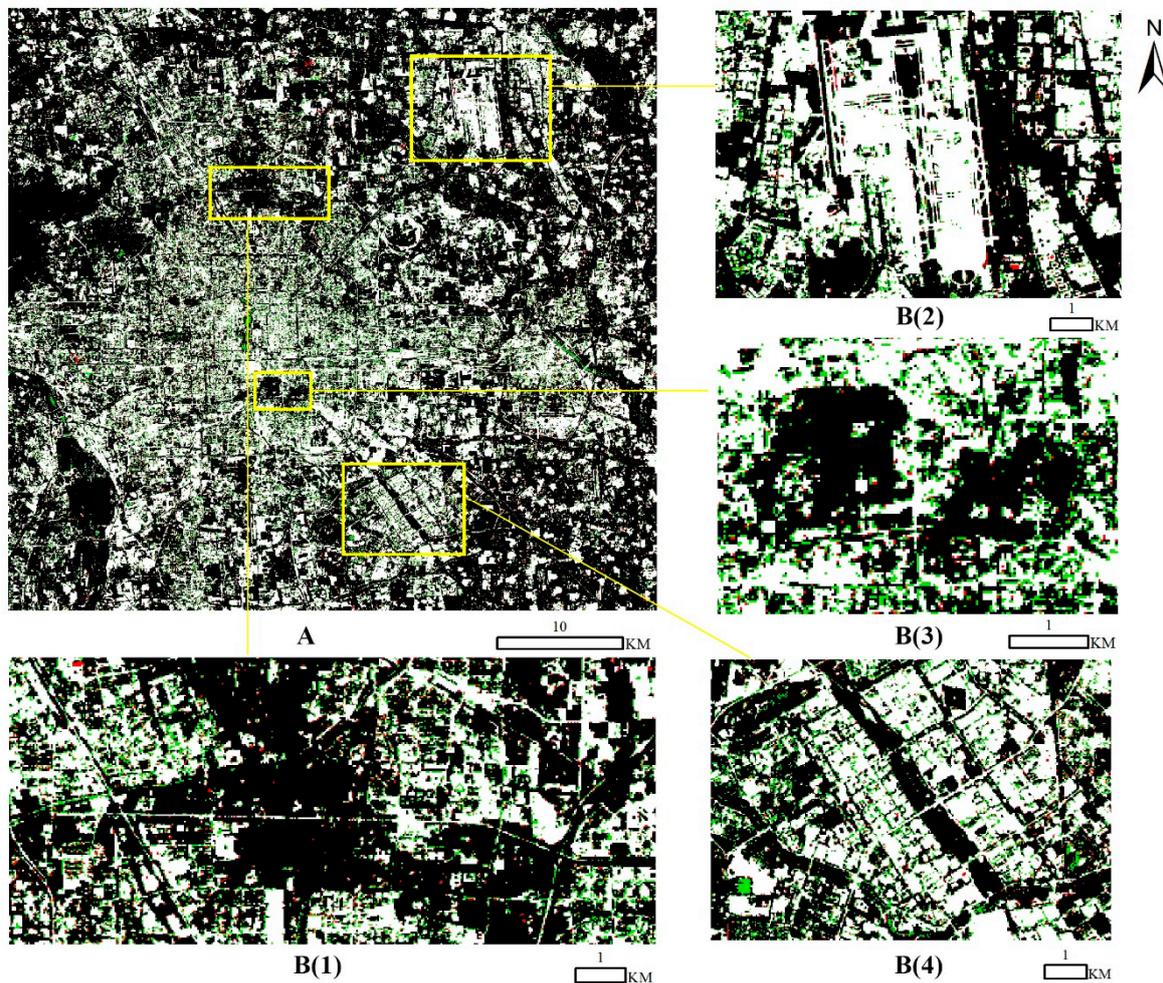


Figure 7. Comparison of built-up identification between Landsat 8 and Sentinel-2 in 30 m data. A is the whole scene while Bs are the enlarged extracts. The extracts comprise of the Beijing Olympic Park (1), the Beijing Capital International Airport (2), the Temple of Heaven (3), and the Beijing economic-technological development area (4).

It can be found that the built-up area in S2 classification identified as non-built-up in L8 is more common than the opposing situation. These differences usually occur in the marginal area of built-up area, and rarely exist on the inside.

4. Discussion

4.1. Analysis on the Consistency of Bands and Indices Between OLI and MSI

The result analysis of the comparison between Landsat 8 OLI and Sentinel-2 MSI shows that these two data have good consistency while the value of R is little lower than previous studies in both bands and indices pairs [32–37]. The difference in the complex land surface in the urban area might have been caused by the influence from heterogeneity on the correlation of the two data [24]. Furthermore, there are other similar results like the low R of blue band and large RMSD of NIR1 (band 8); the former might have been caused from aerosol scattering in the shortest wavelength [43] and the later may have resulted from the inclusion of the water absorption wavelength in the broader NIR band which was purposefully avoided in the OLI NIR [35].

However, the consistency characteristics of L8 and S2 are obvious and diverse: a) most value of OLI bands are smaller than the value of corresponding bands of MSI, b) except for individual bands, R ranged from 0.9133 to 0.9606 and c) the distribution diagram between OLI and MSI are all consistent in NDVI, NDBI, and MNDWI. There is a good consistency between RMSD and R in the band comparison, for example, when spring's R is the highest with the minimum RMSD, while summer's R is the lowest with maximum RMSD. However, in the comparison between the various indices, neither the RMSD nor dispersion has any influence on R, for example, while RMSD = 0.0703 then R = 0.9311 in summer and while RMSD = 0.0530 then R = 0.8833 in winter.

The fluctuating consistency across seasons between the performance of OLI and MSI could be illustrated by different seasonal R, RMSD, and ARD, although their fluctuation may differ between the bands and indices comparison. For example, the highest R occurred in spring (0.9470) in band statistics but in summer (0.9311) during the indices comparison. This characteristic is bound to have an impact on land cover detection that relies on time series data for identification and extraction, such as farmland [34] and mangrove forests [44]. Therefore, this should be noted in joint applications of long time series and work to explore this issue has already started in Longos forest [32]. Meanwhile, the observation capability of S2 10 m data and L8 30 m data is different [31], which is reflected in Figure 5, in which error is located mostly at both ends of the data range in NDBI and MNDWI and $[-0.1, 0.1]$ in NDVI which is the mixed zone of water, impervious surface and bare land [45]. Therefore, even though the MSI data are resampled to 30 m for the comparison, the features of the original data can still affect the data of S2 30 m because a new pixel was generated from the features of original data in the process of fusion and resampling [46].

4.2. Comparison of Roughly Built-Up Area Identification Between OLI and MSI and Future Work

Since this study focused on a preliminary comparison between Landsat 8 OLI data and Sentinel-2 MSI data, the extraction accuracy was not verified and these built-up extraction results were only used to test the similar results of the two data when applied. Meanwhile, the results of identical classification between OLI and MSI in different thresholds were also plotted to demonstrate the similarity within seasons.

The comparison of built-up area identification between OLI and MSI by IBI shows that summer has the highest identical classification. On one hand, the phenological pattern of vegetation should be considered. During the growing season (summer), there could be more features to assist in land cover classification [6]. On the other hand, NDVI, NDBI, and MNDWI determined the IBI and the data consistency of these indexes is the highest in summer (R is highest), which might lead to a greater similarity in the identification between OLI and MSI data. The second reason appears to be more convincing as the compared results of these two data are positively correlated with the seasonal mean R for all indices. The different statistics such as R, RMSD, and ARD might indicate the difference between OLI and MSI when only one index was applied, but in the case of multi indices, the R could be a key indicator of the consistence, as we discussed above.

When the data consistency between Landsat 8 and Sentinel-2 was tested on the identification of built-up areas in an urban setting, the majority of the differences occurred at the boundaries of the built-up areas. The comparison between Figure 7B (2) with (3) shows that the more complex the land surface of an urban area is, the more boundaries of different land cover exist and the greater the differences identified. These inconsistencies mostly resulted from the larger built-up area classified by S2 in comparison to those identified by L8. This might be caused by the more powerful observation capability of S2 10 m data than L8 30 m data on the urban areas [31]. Furthermore, considering the formula for calculating IBI (Equation (6)) and the rough estimation of the data value of NDVI, NDBI, and MNDWI between OLI and MSI in Section 3.2, the MSI IBI is overall smaller than OLI IBI. When using the same threshold to segment images, it could be predictable to get more built-up area in MSI data. Therefore, to set different threshold for OLI and MSI data should be considered in collaboration applications of the two data.

Although many studies utilized various applications with different dates data and showed that a difference of less than three days does not lead to significant changes in the correlation coefficient [28,32,47], the compared result of the only one pair from the different dates in this study shows the lowest correlation and largest scatter and RMSD among all comparisons. This may be due to the presence of thin clouds which were not completely excluded by cloud removal or atmospheric correction [48] and influenced the comparison of these data pairs in summer [37]. Therefore, it is hoped that more quality data will be available in the future so that the summer data can be used to find comparable data obtained on the same date.

5. Conclusions

In summary, the present research not only tested the overall consistency of Landsat 8 OLI and Sentinel-2 MSI time series products by comparing the surface reflectance in corresponding bands and data value of three indices, it also focused on the change in synergetic behavior between the two sensors. Although the consistency of Landsat 8 OLI data and Sentinel-2 MSI data as demonstrated by the comparative result of spectral bands and indices, differs from previous works, the results are still similar (R range from 0.8534 to 0.9606 in band and from 0.8700 to 0.9574 in indices). The differences in data consistencies may have been caused by variations in the heterogeneity and complexity of urban surfaces and should be considered when multiple datasets are combined to observe urban change.

It is easy to see that the consistency of OLI and MSI varies in different seasons in either spectrums or indices, and the fluctuations of seasonal consistency differs in spectral bands and indices. However, the results indicate that when multiple indices from a combination of OLI and MSI data were used, the compared result is clearly affected by the dataset consistency. Hence, it is vital to verify and evaluate the data consistency during a study period, and pick out the most relevant data and threshold when a combination of both OLI data and MSI data are used.

Author Contributions: Conceptualization, Z.N. and K.K.Y.C; methodology, software and formal analysis Z.N. and B.X.; writing—original draft preparation and review and editing, Z.N., K.K.Y.C and B.X.; project administration and funding acquisition, B.X.

Funding: This research was partially supported by the National Research Program of the Ministry of Science and Technology of the People's Republic of China (2016YFA0600104), the youth talent support program of College of Global Change and Earth System Science (312231103) and donations from Delos Living LLC and the Cyrus Tang Foundation to Tsinghua University.

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

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

References

1. Lahariya, C. The State of the World Population 2007: Unleashing the potential of urban growth. *Indian Pediatr.* **2008**, *45*, 481–482. [[PubMed](#)]
2. Xu, H. A new index for delineating built-up land features in satellite imagery. *Int. J. Remote Sens.* **2008**, *29*, 4269–4276. [[CrossRef](#)]
3. Gong, P.; Liang, S.; Carlton, E.J.; Jiang, Q.; Wu, J.; Wang, L.; Remais, J.V. Urbanisation and health in China. *Lancet* **2012**, *379*, 843–852. [[CrossRef](#)]
4. Wei, C.; Zhang, Y.; Pengwang, C.; Gao, W. Evaluation of urbanization dynamics and its impacts on surface heat islands: A case study of Beijing, China. *Remote Sens.* **2017**, *9*, 453.
5. Kaufmann, R.K.; Seto, K.C.; Schneider, A.; Liu, Z.; Zhou, L.; Wang, W. Climate response to rapid urban growth: Evidence of a human-induced precipitation deficit. *J. Clim.* **2007**, *20*, 2299–2306. [[CrossRef](#)]
6. Li, X.; Zhou, Y.; Asrar, G.R.; Mao, J.; Li, X.; Li, W. Response of vegetation phenology to urbanization in the conterminous United States. *Glob. Chang. Biol.* **2017**, *23*, 2818–2830. [[CrossRef](#)]
7. Yao, R.; Wang, L.; Huang, X.; Guo, X.; Niu, Z.; Liu, H. Investigation of urbanization effects on land surface phenology in Northeast China during 2001–2015. *Remote Sens.* **2017**, *9*, 66. [[CrossRef](#)]

8. Zhang, Q.; Seto, K.C. Can night-time light data identify typologies of urbanization? A global assessment of successes and failures. *Remote Sens.* **2013**, *5*, 3476–3494. [[CrossRef](#)]
9. Xie, Y.; Weng, Q. Spatiotemporally enhancing time-series DMSP/OLS nighttime light imagery for assessing large-scale urban dynamics. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 1–15. [[CrossRef](#)]
10. Zhou, Y.; Yang, G.; Wang, S.; Wang, L.; Wang, F.; Liu, X. A new index for mapping built-up and bare land areas from Landsat-8 OLI data. *Remote Sens. Lett.* **2014**, *5*, 862–871. [[CrossRef](#)]
11. Zhang, H.; Lin, H.; Wang, Y. A new scheme for urban impervious surface classification from SAR images. *ISPRS J. Photogramm. Remote Sens.* **2018**, *139*, 103–118. [[CrossRef](#)]
12. Solari, L.; Ciampalini, A.; Raspini, F.; Bianchini, S.; Moretti, S. PSInSAR analysis in the pisa urban area (Italy): A case study of subsidence related to stratigraphical factors and urbanization. *Remote Sens.* **2016**, *8*, 120. [[CrossRef](#)]
13. Li, X.; Zhou, Y.; Zhu, Z.; Liang, L.; Yu, B.; Cao, W. Mapping annual urban dynamics (1985–2015) using time series of Landsat data. *Remote Sens. Environ.* **2018**, *216*, 674–683. [[CrossRef](#)]
14. Cao, X.; Chen, J.; Imura, H.; Higashi, O. A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. *Remote Sens. Environ.* **2009**, *113*, 2205–2209. [[CrossRef](#)]
15. Small, C. High spatial resolution spectral mixture analysis of urban reflectance. *Remote Sens. Environ.* **2003**, *88*, 170–186. [[CrossRef](#)]
16. 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](#)]
17. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [[CrossRef](#)]
18. 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.
19. Xiong, J.; Thenkabail, P.; Tilton, J.; Gumma, M.; Teluguntla, P.; Oliphant, A.; Congalton, R.; Yadav, K.; Gorelick, N. Nominal 30-m Cropland Extent Map of Continental Africa by Integrating Pixel-Based and Object-Based Algorithms Using Sentinel-2 and Landsat-8 Data on Google Earth Engine. *Remote Sens.* **2017**, *9*, 1065. [[CrossRef](#)]
20. Inglada, J.; Arias, M.; Tardy, B.; Hagolle, O.; Valero, S.; Morin, D.; Dedieu, G.; Sepulcre, G.; Bontemps, S.; Defourny, P.; et al. Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. *Remote Sens.* **2015**, *7*, 12356–12379. [[CrossRef](#)]
21. Brezonik, P.L.; Olmanson, L.G.; Finlay, J.C.; Bauer, M.E. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. *Remote Sens. Environ.* **2015**, *157*, 199–215. [[CrossRef](#)]
22. Pahlevan, N.; Chittimalli, S.K.; Balasubramanian, S.V.; Vellucci, V. Sentinel-2/Landsat-8 product consistency and implications for monitoring aquatic systems. *Remote Sens. Environ.* **2019**, *220*, 19–29. [[CrossRef](#)]
23. Novelli, A.; Aguilar, M.A.; Nemmaoui, A.; Aguilar, F.J.; Tarantino, E. Performance evaluation of object based greenhouse detection from Sentinel-2 MSI and Landsat 8 OLI data: A case study from Almería (Spain). *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 403–411. [[CrossRef](#)]
24. Mandanici, E.; Bitelli, G. Preliminary comparison of sentinel-2 and landsat 8 imagery for a combined use. *Remote Sens.* **2016**, *8*, 1014. [[CrossRef](#)]
25. Roy, D.P.; Kovalsky, V.; Zhang, H.K.; Vermote, E.F.; Yan, L.; Kumar, S.S.; Egorov, A. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sens. Environ.* **2016**, *185*, 57–70. [[CrossRef](#)]
26. Chastain, R.; Housman, I.; Goldstein, J.; Finco, M. Empirical cross sensor comparison of Sentinel-2A and 2B MSI, Landsat-8 OLI, and Landsat-7 ETM+ top of atmosphere spectral characteristics over the conterminous United States. *Remote Sens. Environ.* **2019**, *221*, 274–285. [[CrossRef](#)]
27. Flood, N. Comparing Sentinel-2A and Landsat 7 and 8 using surface reflectance over Australia. *Remote Sens.* **2017**, *9*, 659. [[CrossRef](#)]
28. Lessio, A.; Fissore, V.; Borgogno-Mondino, E. Preliminary tests and results concerning integration of sentinel-2 and Landsat-8 OLI for crop monitoring. *J. Imaging* **2017**, *3*, 49. [[CrossRef](#)]
29. European Space Agency. Available online: <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-1c/algorithm> (accessed on 8 April 2019).

30. National Aeronautics and Space Administration. Available online: http://landsat.gsfc.nasa.gov/wp-content/uploads/2014/09/Ball_BA_RSR.v1.2.xlsx (accessed on 28 November 2018).
31. Lefebvre, A.; Sannier, C.; Corpetti, T. Monitoring urban areas with Sentinel-2A data: Application to the update of the Copernicus High Resolution Layer Imperviousness Degree. *Remote Sens.* **2016**, *8*, 606. [CrossRef]
32. Arekhi, M.; Goksel, C.; Balik Sanli, F.; Senel, G. Comparative Evaluation of the Spectral and Spatial Consistency of Sentinel-2 and Landsat-8 OLI Data for Igneada Longos Forest. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 56. [CrossRef]
33. Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sens. Environ.* **2004**, *91*, 332–344. [CrossRef]
34. Xu, Y.; Yu, L.; Zhao, F.R.; Cai, X.; Zhao, J.; Lu, H.; Gong, P. Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: Experiments from three sites in Africa. *Remote Sens. Environ.* **2018**, *218*, 13–31. [CrossRef]
35. 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]
36. Lu, C.; Fan, L. Winter wheat yield potentials and yield gaps in the North China Plain. *Field Crop. Res.* **2013**, *143*, 98–105. [CrossRef]
37. Zhang, H.K.; Roy, D.P.; Yan, L.; Li, Z.; Huang, H.; Vermote, E.; Skakun, S.; Roger, J.C. Characterization of Sentinel-2A and Landsat-8 top of atmosphere, surface, and nadir BRDF adjusted reflectance and NDVI differences. *Remote Sens. Environ.* **2018**, *215*, 482–494. [CrossRef]
38. Storey, J.; Roy, D.P.; Masek, J.; Gascon, F.; Dwyer, J.; Choate, M. A note on the temporary misregistration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) imagery. *Remote Sens. Environ.* **2016**, *186*, 121–122. [CrossRef]
39. Gao, F.; Masek, J.G.; Wolfe, R.E. Automated registration and orthorectification package for Landsat and Landsat-like data processing. *J. Appl. Remote Sens.* **2009**, *3*, 033515.
40. Xu, R.; Liu, J.; Xu, J. Extraction of high-precision urban impervious surfaces from sentinel-2 multispectral imagery via modified linear spectral mixture analysis. *Sensors* **2018**, *18*, 2873. [CrossRef]
41. Zha, Y.; Gao, J.; Ni, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* **2003**, *24*, 583–594. [CrossRef]
42. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033. [CrossRef]
43. 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]
44. Hu, L.; Li, W.; Xu, B. Monitoring mangrove forest change in China from 1990 to 2015 using Landsat-derived spectral-temporal variability metrics. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 88–98. [CrossRef]
45. Chen, B.; Nie, Z.; Chen, Z.; Xu, B. Quantitative estimation of 21st-century urban greenspace changes in Chinese populous cities. *Sci. Total Environ.* **2017**, *609*, 956–965. [CrossRef] [PubMed]
46. Chen, B.; Xu, B. A novel method for measuring landscape heterogeneity changes. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 567–571. [CrossRef]
47. Skakun, S.; Vermote, E.; Roger, J.-C.; Franch, B. Combined Use of Landsat-8 and Sentinel-2A Images for Winter Crop Mapping and Winter Wheat Yield Assessment at Regional Scale. *AIMS Geosci.* **2017**, *3*, 163–186. [CrossRef] [PubMed]
48. Stillinger, T.; Roberts, D.A.; Collar, N.M.; Dozier, J. Cloud Masking for Landsat 8 and MODIS Terra Over Snow-Covered Terrain: Error Analysis and Spectral Similarity Between Snow and Cloud. *Water Resour. Res.* **2019**, *55*, 6169–6184. [CrossRef]

