Article

A 30 m Resolution Surface Water Mask Including Estimation of Positional and Thematic Differences Using Landsat 8, SRTM and OpenStreetMap: A Case Study in the Murray-Darling Basin, Australia

Gennadii Donchyts [1,2,*], Jaap Schellekens [2], Hessel Winsemius [2], Elmar Eisemann [3] and Nick van de Giesen [2]

[1] Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands

[2] Deltares, Boussinesqweg 1, 2629 HV Delft, The Netherlands; jaap.schellekens@deltares.nl (J.S.); hessel.winsemius@deltares.nl (H.W.); n.c.vandegiesen@tudelft.nl (N.G.)

[3] Computer Graphics and Visualization, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands; e.eisemann@tudelft.nl

* Correspondence: g.donchyts@tudelft.nl; Tel.: +31-88-335-8273

Abstract: Accurate maps of surface water are essential for many environmental applications. Surface water maps can be generated by combining measurements from multiple sources. Precise estimation of surface water using satellite imagery remains a challenging task due to the sensor limitations, complex land cover, topography, and atmospheric conditions. As a complementary dataset, in the case of hilly landscapes, a drainage network can be extracted from high-resolution digital elevation models. Additionally, Volunteered Geographic Information (VGI) initiatives, such as OpenStreetMap, can also be used to produce high-resolution surface water masks. In this study, we derive a high-resolution water mask using Landsat 8 imagery and OpenStreetMap as well as (potential) a drainage network using 30 m SRTM. Our approach to derive a surface water mask from Landsat 8 imagery comprises the use of a lower 15% percentile of Landsat 8 Top of Atmosphere (TOA) reflectance from 2013 to 2015. We introduce a new non-parametric unsupervised method based on the Canny edge filter and Otsu thresholding to detect water in flat areas. For hilly areas, the method is extended with an additional supervised classification step used to refine the water mask. We applied the method across the Murray-Darling basin, Australia. Differences between our new Landsat-based water mask and the OpenStreetMap water mask regarding positional differences along the rivers and overall coverage were analyzed. Our results show that about 50% of the OpenStreetMap linear water features can be confirmed using the water mask extracted from Landsat 8 imagery and the drainage network derived from SRTM. We also show that the observed distances between river features derived from OpenStreetMap and Landsat 8 are mostly smaller than 60 m. The differences between the new water mask and SRTM-based linear features and hilly areas are slightly larger (110 m). The overall agreement between OpenStreetMap and Landsat 8 water masks is about 30%.

Keywords: water mask; rivers; Landsat 8; MNDWI; Canny edge filter; Otsu thresholding; SRTM; HAND; OpenStreetMap; CART; eartH2Observe

1. Introduction

High-resolution measurements of geometry and dynamics of surface water are essential parameters required for many environmental applications, such as flood forecasting and warning,

agricultural and urban water management, and simulation of transport of pollutants in water bodies [1]. While there are many applications available focusing on detailed modeling of different rivers, lakes, and coastal areas, a consistent and high-resolution (less than 100 m) global dataset on surface water geometry and dynamics at high-resolution and planetary scale is not yet available. An important use of global datasets of surface water is in global hydrological and hydraulic models that, in many cases, include a fraction of surface water per grid cell in their parameterization (see for example, [2]). At the same time, these models are moving away from their typical resolution of around 0.5 degrees to the hyper-resolution domain [3,4] and include more elaborate routing schemes, thus increasing the need for accurate surface water mapping.

Most methods to detect water from multispectral (satellite) imagery use the fact that water significantly absorbs most radiation at near-infrared wavelengths and beyond. This fact makes it easy to detect clear water by introducing a spectral index, such as the Normalized Difference Water Index (NDWI), representing the slope of the water spectral curve [5,6]. The Modified Normalized Difference Water Index (MNDWI) [7], appears to be more sensitive, due to the use of a shortwave infrared band instead of the near-infrared one in NDWI, resulting in better water feature detection. While detection of clear water features appears to be trivial, many factors make it more difficult in real world situations. Clouds, snow, and ice cause many errors. Additionally, errors of commission (false positive detection of water) can be observed in areas with shadows due to topographic conditions or the presence of cloud shadows. Methods for cloud, cloud shadow, and snow detection are described by [8,9]. These methods make use of information on satellite sensor view angle and solar zenith/azimuth and are also used to produce Surface Reflectance Landsat products [10]. These parameters, combined with elevation data, can also be used to detect hill shadows [11]. An alternative approach to exclude clouds and shadows is the use of average reflectance composites instead of instantaneous images [12,13].

The fact that water is almost never clear in the real world results in changes in its spectral curve, making it difficult to use spectral indices with a single threshold to separate water pixels from non-water pixels. Typical variations of threshold values for different spectral water indices can be found in [14]. One of the approaches to overcome this problem is to use methods that allow detecting threshold values based on a histogram of all NDWI values in a given area. One such method is Otsu thresholding [15,16]. This method is very similar to the k-means method applied to the histogram of spectral index values [17].

Significant improvements in water detection can be achieved by combining optical satellite imagery with elevation data, assuming that most of the permanent surface water can be found in valleys. The use of elevation data to extract a drainage network is also a common practice in hydrological applications [18,19]. A drainage network can be used to construct a Height Above the Nearest Drainage (HAND) map [20], that can be used to separate areas where water can occur from those where it is very unlikely that surface water occurs. A contoured HAND map can also be used to estimate possible flooded areas [21] similar to other geomorphological methods [22]. Many efforts to map open water focus on raw information coming from remote sensing, such as optical multispectral imagery, synthetic-aperture radar (SAR) imagery, or radar altimetry [23]. The combination of remote sensing with existing GIS vector datasets that contain information on water occurrence, gets much less attention, even less so when combined with geomorphological information (*i.e.*, HAND). GIS vector datasets are frequently measured using high-resolution GPS devices or by manually digitizing aerial or satellite imagery. As a result, these datasets usually demonstrate much better precision and contain semantic information, such as feature names, types, and other attributes. Their quality and completeness are not uniform around the globe. One key global vector dataset containing water information is OpenStreetMap (OSM) [24], initiated in 2004 and currently including more than 3 billion objects. From these, more than 8 million objects relate to water [25]. Many environmental applications use OSM, including the extraction of paved area and surface water coverage for hydrological applications [26].

The volunteered nature of OSM is the main factor that makes it less adopted by GIS professionals [27] stressing the importance of the development of automated methods and tools to validate its quality in comparison to other datasets. A good example of how OSM quality can be assessed can be found in [28], showing how positional differences between linear and polygonal features can be addressed. Furthermore, the "increasing buffer" method [29] can be used to estimate the quality of linear features [24]. An excellent overview of papers focusing on OSM quality can be found in [30]. The latter also suggests using elements of ISO 19113:2002 [31] (such as completeness, an error of commission/omission, or positional differences) to evaluate the OSM quality. Unfortunately, to our knowledge, no academic literature exists focusing solely on the quality of water features present in OSM and using global or nearly global remote sensing datasets.

The main reason why OSM was chosen in the current study over local, authoritative datasets, is that it provides a global coverage, even though its local coverage and quality may vary. Additional research would be required to perform a detailed comparison of the datasets presented in this paper to the local Australian authoritative datasets, such as Surface Hydrology [32], Water Observations from Space [33], or 5 m Digital Elevation Model (DEM) of Australia [34].

We derive a high-resolution water mask using Landsat 8 imagery and OpenStreetMap, as well as a (potential) drainage network using 30 m SRTM. Extracting a water mask from OpenStreetMap data is relatively straightforward (Section 2.2), but the other data sources require a specialized workflow. Our approach to derive a surface water mask from Landsat 8 imagery is described in Section 2.4, involving the steps to compute cloud-free average reflectance composites in Section 2.4.1. Additionally, we introduce a new non-parametric unsupervised method to detect water in flat areas (Section 2.4.2). We also propose a supervised classification step to refine the water mask in hill areas (Section 2.4.3).

We make use of several open geospatial and remote sensing datasets to construct an open water map. Section 2 provides an overview of the main input datasets utilized in the study, as well as methods applied or developed to detect water, and to compare the resulting water masks. The main input datasets include (1) images acquired by the NASA Landsat 8 mission [35]; (2) a new revision of a nearly-global 30 m DEM, measured by the SRTM mission [36]; and (3) OSM data for the Murray-Darling basin in Australia.

2. Methods and Study Location

2.1. Study Site: Murray-Darling River Basin

The Murray-Darling Basin [37], named after the two main rivers in the basin, is located in semi-arid and arid climate zones, covering 1,059,000 square kilometers, an equivalent of 14% of Australia's total surface area. OSM water features include 10,106 linear and 6708 aerial hydrographic features representing both natural and manmade water features, such as rivers, lakes, and canals (Figure 1). The features also include water bodies, which are only partially covered by water during the year. Figure 1 shows both hydrographic features extracted from OSM and the (potential) river network obtained from HydroSHEDS [18].

The Murray-Darling Basin receives most of its rainfall from a very small percentage of the basin area; mainly along the southern and eastern parts. The rest of the basin is flat and low-lying, contributing very little or no run-off to the rivers. Relevant for this study, many tributaries within the basin are intermittent streams, with highly variable flows dependent on the wetness of the year.
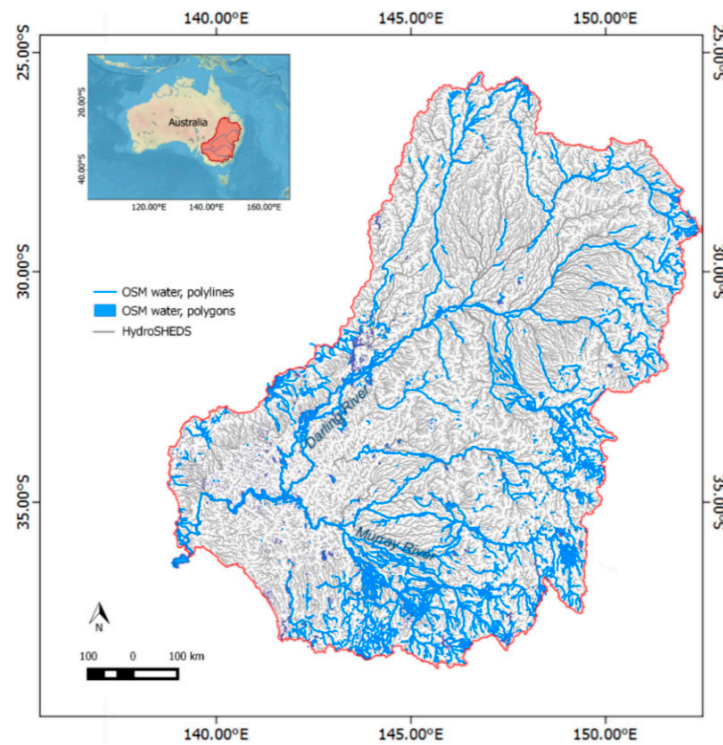
Figure 1. Overview of the study area, Murray-Darling River Basin, and vector water mask dataset extracted from OSM. The inset map is based on Natural Earth 1 raster dataset [38].

2.2. Input Datasets Used to Extract the Water Mask

Even though HydroSHEDS provides a much better coverage, its resolution is limited to 15 arcseconds (~450 m), which is insufficient to resolve all of the small water features required for detailed water-related applications. Furthermore, it was based on an older, 90 m revision of SRTM, which is less accurate when comparing to a recently-released 30 m version. Therefore, to develop a method to generate the water mask, we have used four raw input datasets (see Table 1 and Figure 2). The first dataset, Landsat 8, combines 2743 scenes of optical multi-spectral satellite imagery acquired during 2013–2015 over the study area. The Landsat 8 mission has already operated for more than two years and, therefore, provides a reasonably long period of information to estimate water dynamics on a global scale. The second dataset is based on the new 30 m revision of the SRTM digital elevation model. The last data source consists of all water features, extracted from a recent version of OSM Planet File [39]. Additionally, HydroBASINS [40] was used as a supplementary dataset to simplify the extraction of a drainage network from SRTM.

Extraction of water features from OSM is a relatively trivial task, mainly consisting of a careful selection of proper filters and data conversion tools. We have used data conversion tools provided by OSM and a set of programs based on GDAL [41] to select only relevant features representing surface water.

To derive a high-resolution drainage network from SRTM, we have used the D8 method [42], implemented using PCRaster [43,44]. The stream delineation step was applied for every catchment of the Murray-Darling basin. The HydroBASINS level 8 catchment geometries were used for this purpose. This step was required due to the large size of the catchment making it difficult to perform processing in a single step.

Table 1. Input datasets used in the study.

Dataset	Type	Resolution	Notes
Landsat 8 TOA	Multispectral Imagery	15 m, 30 m, 60 m	2743 scenes were used, acquired during 2013–2015, top of atmosphere (TOA) reflectance.
SRTM	Elevation Imagery	30 m	Effective resolution is lower due to the presence of high-frequency noise
OpenStreetMap	Polyline and Polygon Vector	1–100 m	Planet file from August 2015, the following tags query was used to indicate water features: *natural = water or natural = spring or waterway = or landuse = basin or landuse = reservoir or barrier = ditch or landuse = saltpond*
HydroBASINS	Polygons	~450 m	Level 8 basins were used to delineate HAND using 30 m version of SRTM
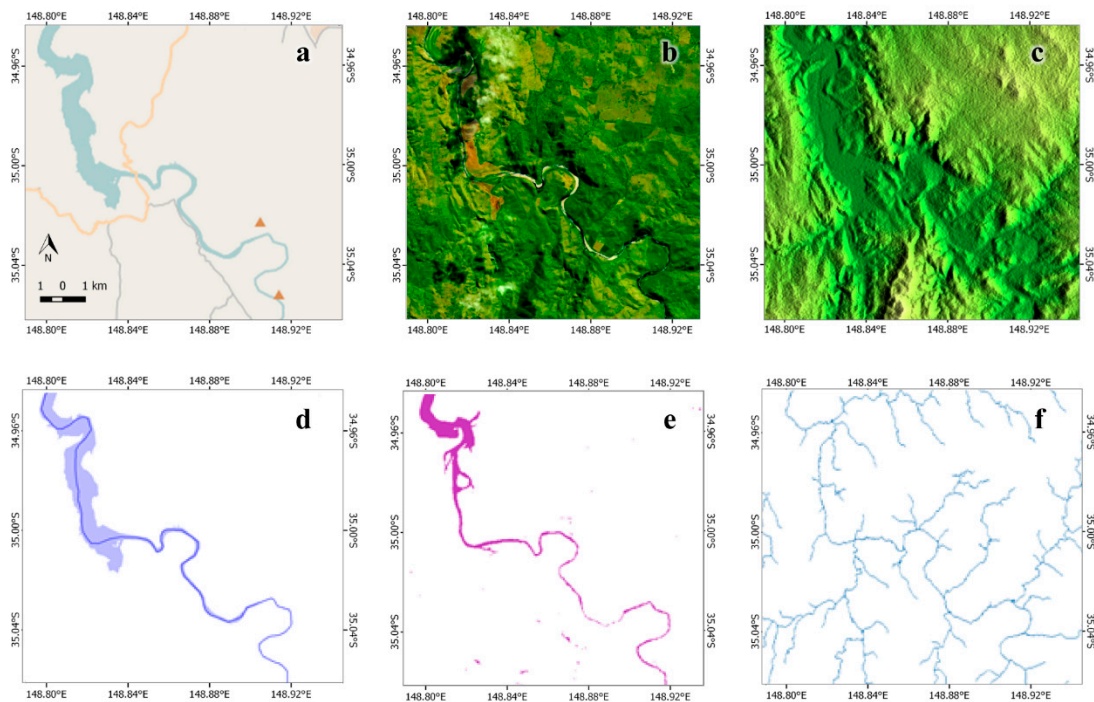
Figure 2. The datasets: OpenStreetMap (**a**); Landsat 8 (**b**) and 30m SRTM (**c**); and their corresponding water masks (**d**–**f**). Murrumbidgee River, north of Canberra.

Extraction of the water mask using Landsat 8 imagery was the most challenging task and required the development of a new method to derive a water mask from multi-spectral imagery. The new method combines the use of the MNDWI spectral index (Equation (2)) extended with a non-parametric detection of a local threshold to improve the accuracy of water detection. The MNDWI is very similar to the NDWI spectral index (Equation (1)) but uses a shortwave infrared band instead of near-infrared. Additional steps include the use of the Normalized Difference Vegetation Index (NDVI) [45] (Equation (3)) with a high threshold value (0.3) to exclude false water detection in very dark vegetated areas.

$$NDWI = \frac{\rho_{green} - \rho_{nir}}{\rho_{green} + \rho_{nir}} \tag{1}$$

$$MNDWI = \frac{\rho_{green} - \rho_{swir1}}{\rho_{green} + \rho_{swir1}} \qquad (2)$$

$$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}} \qquad (3)$$

where ρ_{green}, ρ_{swir1}, ρ_{nir}, and ρ_{red} represent TOA reflectance for corresponding Landsat 8 bands.

We have included an additional classification step to refine the water mask for hilly areas, where the results of automated classification led to high commission errors. We have used a supervised classification based on Classification And Regression Tree (CART) [46], which was trained using a manually digitized training set to distinguish between water and land pixels.

Our selection of the input datasets was based on the assumption that the accuracy of all three datasets is similar. However, estimation of the actual errors of OSM would be difficult, mainly because OSM features are usually based on different measurement methods (GPS traces, or manual digitizing using medium- or high-resolution imagery of a bulk import from other databases of varying nature). For Landsat 8 and SRTM, the main limitations are well known [35,47]. Horizontal accuracy of Landsat 8 is known to be better than 12 m in 90% of the 30 m resolution images of the Operational Land Imager (OLI). For SRTM, vertical relative and absolute errors can be explained by its radar nature and are in the order of 10 and 16 meters for 90% of the data, with an absolute geolocation error below 13 m.

2.3. Derivation of Hydrological Variables: Drainage Network and HAND

In addition to the drainage network, a number of other hydrological parameters were derived during SRTM analysis, such as local drainage direction, flow accumulation, and HAND (see Figure 3).
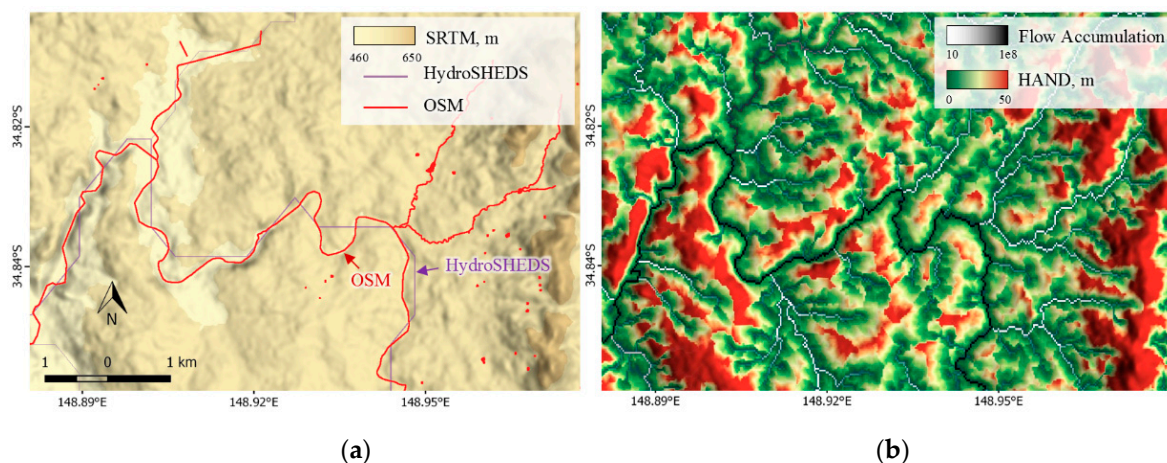
Figure 3. Digital elevation model, SRTM 30 m (**a**); and height above the nearest drainage (**b**).

A (potential) water mask was generated by thresholding HAND values using 1 m elevation above the nearest drain. We found this approach to work better in estimating water masks in flat areas compared to the methods based on flow accumulation.

All hydrological variables were derived using the following steps: (1) clip DEM using a Google Earth Engine script and one of 1494 HydroBASIN polygons and download it to Google Compute Engine (GCE); (2) compute slopes, local drainage directions (LDD, includes pit-removal step), flow accumulations, and HAND; and (3) upload results to Google Earth Engine [48] and Google Fusion Table [49] so they can be used in parallel-processing scripts for further analysis. To compute HAND, a drainage network had to be estimated by thresholding the flow accumulation. We have used threshold equal to 100 upstream cells, which was sufficient to detect most of the potential rivers.

The resulting HAND can be used for the estimation of potential flood areas, but also to detect pixels where potential errors occur due to hill shadows. These areas were estimated by generating

a binary mask based on a variation of HAND values. The mask showing potential hilly areas was computed by marking pixels as hilly in a case where HAND values larger or equal to 30 m were found in the 300 m radius neighborhood.

2.4. Method of Water Detection Using Landsat 8

Our method of water detection from multi-spectral multi-temporal imagery is based on a step-wise approach combining unsupervised and supervised classification steps (Figure 4). The unsupervised step was applied first to detect the initial water mask using percentile images of reflectance, resulting in minor omission errors in flat areas. However, we obtain very high commission errors in hilly areas due to terrain shadow. Therefore, an additional step was required to refine the water mask using supervised classification.
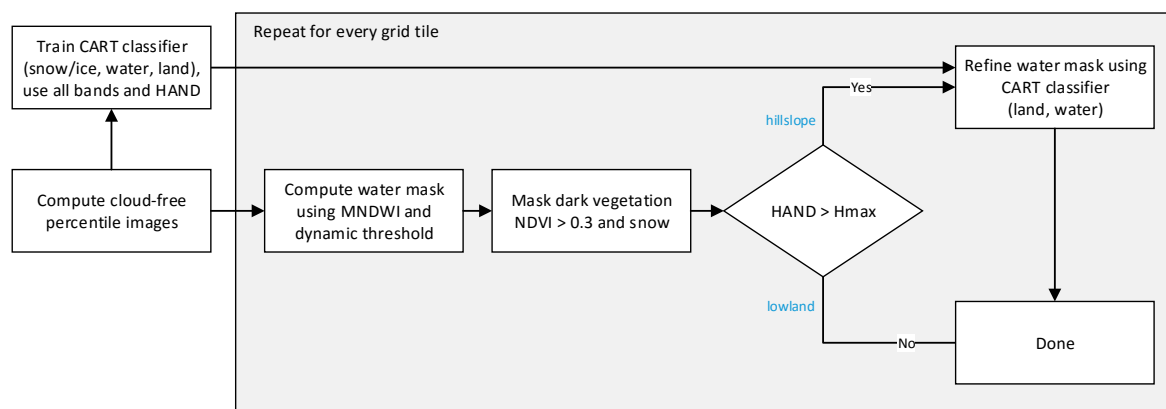

Figure 4. Water detection processing pipeline.

The unsupervised classification step is based on the local adaptive threshold detection method presented in Section 2.4.2. The supervised classification step, based on CART, was introduced to reduce commission errors found in the case of shadows and snow/ice pixels. It was performed only for hilly areas where misclassified pixels were observed. Hilly areas were detected using a threshold of Hmax, representing the minimum HAND value (30 m in our case). In this way, we could keep omission errors low for flat areas and ensure low commission errors for hilly areas. The final error of water detection was very low (less than 1%), mainly due to a presence of mixed pixels or incomplete training data for hilly areas. An additional step was required to exclude very dark vegetated areas, resulting in high MNDWI values and, therefore, misclassified as water. These errors were removed by eliminating pixels with NDVI values greater than 0.3.

To compute the final water mask, the study area was divided into regular grid tiles of 0.2×0.2 degrees in size (Figure A1). This step was needed to make sure that dynamic MNDWI threshold is estimated for every tile, but also to parallelize the processing. Finally, the above workflow was applied to every tile.

2.4.1. Cloud-Free Landsat 8 Percentile Images

Our method to exclude clouds and shadows in the satellite imagery is based on the use of percentile images extracted from an image collection, spanning a two year period (2013–2015), instead of the original images. The percentile images were computed on a per-pixel and per-band basis using 50–130 top of atmosphere (TOA) intensity values. The larger number of images comes from a higher revisit frequency due to an overlap in the satellite swath.

The percentile images (Figure 5) appeared to describe the water dynamics in a better way when compared to the interval mean images used in the other studies [12,13]. Even though we have confirmed this result only by visual inspection, the reasoning comes from the fact that the water

surface area may change sharply depending on local topographic conditions. This water area change results in sharp changes between water masks present in different percentiles (see Figure 6).
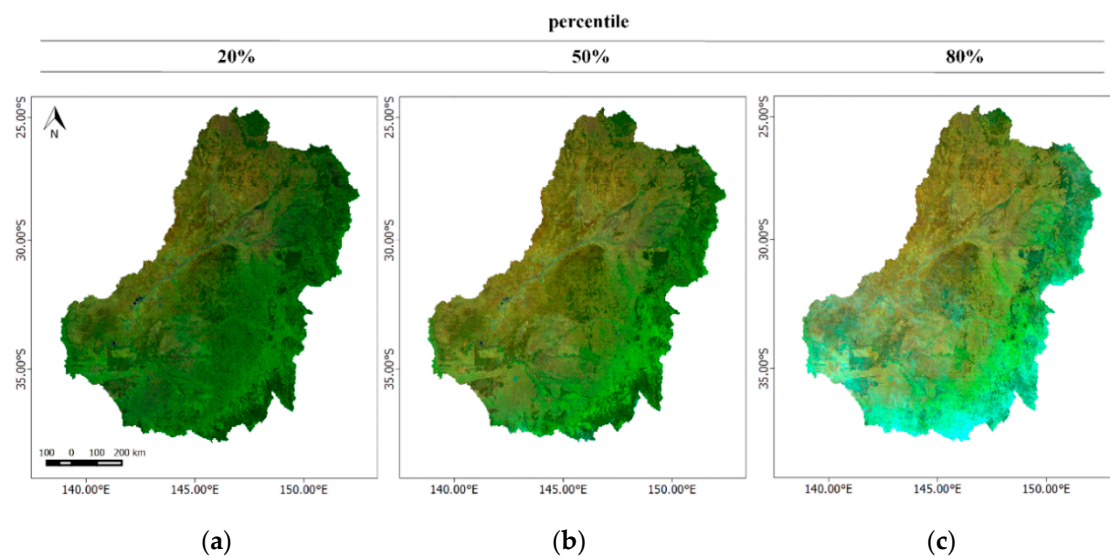
Figure 5. Intensity percentile false-color image (SWIR1, NIR, green) based on 2743 Landsat 8 images for 2013–2015, 20% (**a**); 50% (**b**); and 80% (**c**).
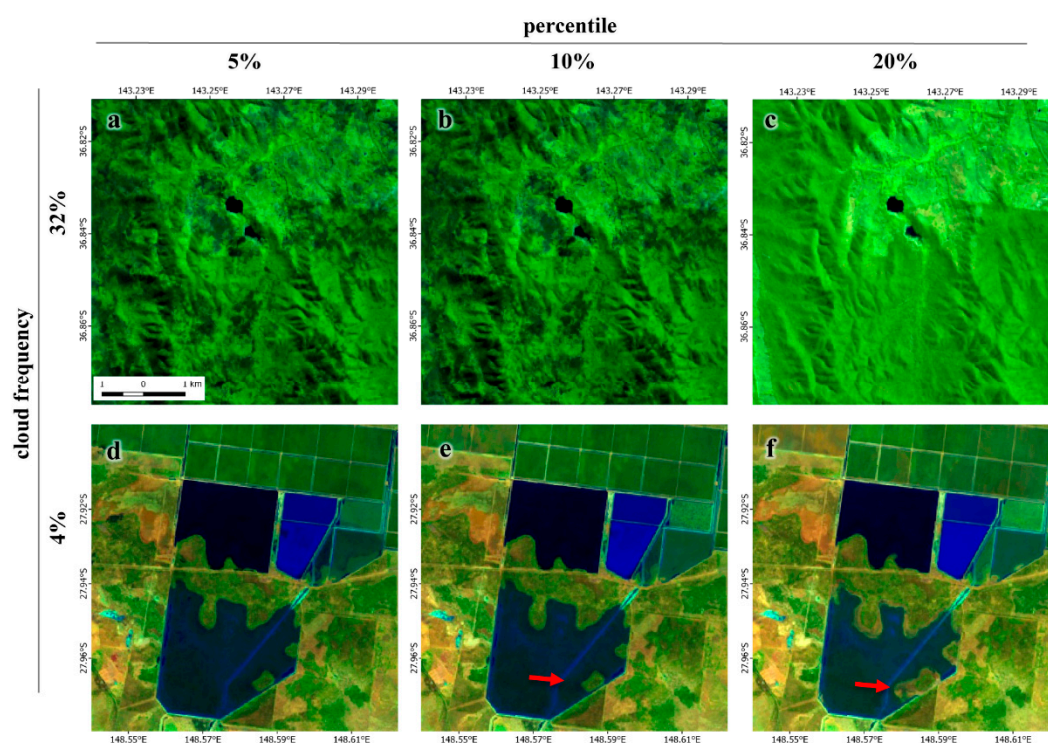

Figure 6. Intensity percentile (5%, 10%, 20%) false-color images (SWIR1, NIR, green) for areas with an average cloud cover of 32% (**a–c**) and 4% (**d–f**).

Over the Murray-Darling Basin, the percentile range of 15%–55% of all TOA intensities was empirically found to be suitable for permanent water detection. However, for semi-arid and flat areas, where cloud frequency is very low, a larger range of percentiles (up to 90%) could be used as well. Average images generated for very low percentiles usually result in too many artifacts present in the images due to cloud and hill shadows, making them difficult to interpret automatically (Figure 6a,b).

At the same time, the use of lower percentiles has a higher chance to represent a larger amount of surface water, present during floods and wet seasons, and a smaller amount of surface water observed during dry seasons (Figure 6d–f). For a more detailed analysis, the choice of the upper and lower percentiles can be estimated by taking cloud frequency and topographic conditions into account.

2.4.2. Adaptive Threshold Detection Using MNDWI, Canny Edge Filter, and Otsu Thresholding

As was previously mentioned, varying thresholds of spectral water indices usually limit their applicability, requiring manual threshold adjustment because the spectral properties of open water are not the same everywhere. In many cases, surface water constitutes only a small fraction of the overall land cover (Figure 7), making it harder to detect with threshold-based methods. Large local errors may be introduced when a constant threshold of 0.0 is used to distinguish between water and land (Figure 8). The challenge is to establish a varying threshold, that can be derived automatically.

(a)　　　　　　　　　　　　　(b)　　　　　　　　　　　　　(c)
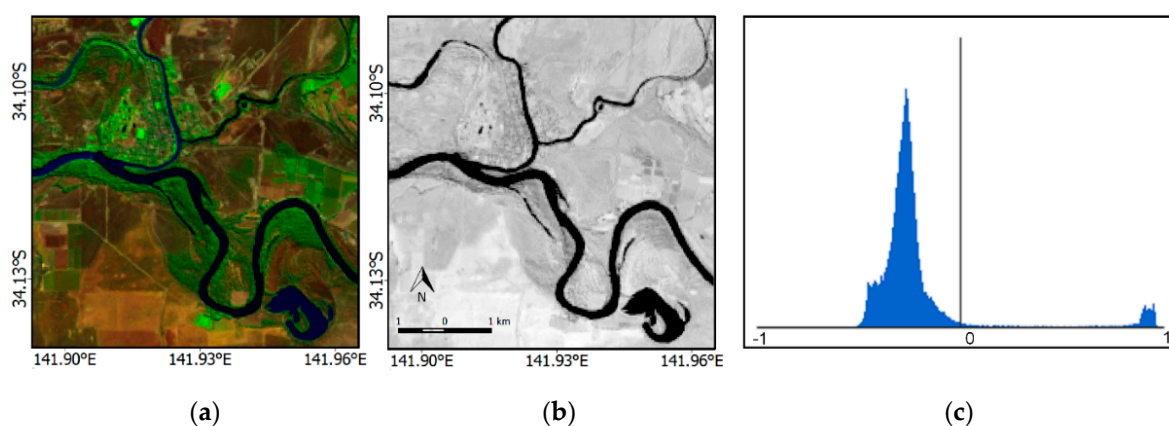
Figure 7. False-color composite image (SWIR1, NIR and green) (**a**); MNDWI index values (**b**); and its histogram (**c**).

(a)　　　　　　　　　　　　　　　　　(b)
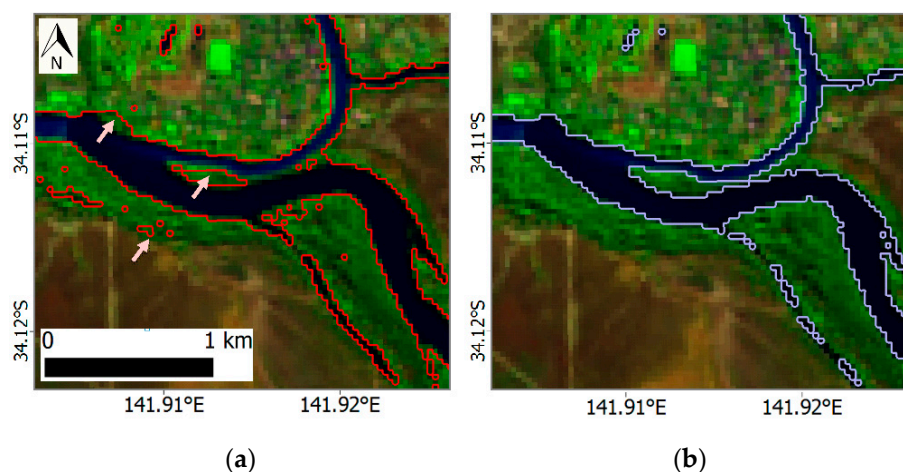
Figure 8. Water boundary classified using threshold = 0.0 (**a**) and threshold = 0.32 automatically detected for this location using the new method (**b**).

The use of non-parametric threshold detection methods, such as histogram-based Otsu thresholding helps to overcome some of the problems [15,16], but this approach does not work when the fraction of water pixels is small.

The main principle of our method is an extension of the Otsu-based methods by a Canny edge filter to reduce the number of input pixels only to those located near water-land edges. The use of the

Canny edge filter with a very high threshold applied to the spectral index image reveals only edges located near sharp value changes. In the case of water spectral indices, this usually takes place when the near-infrared band abruptly changes in value from one pixel to the next.

Potential water and land pixels located near water are then computed using morphological dilation applied to the detected edges. It is important to note that this approach can result in a skewed distribution in the case of thin, single pixel wide water bodies (canals). To overcome this problem we have used a buffer size (dilation) equal to half of the pixel in Step 3. In an ideal situation, the resulting distribution should look bimodal (Figure 9) so that a clear distinction between land and water can be made from this distribution.
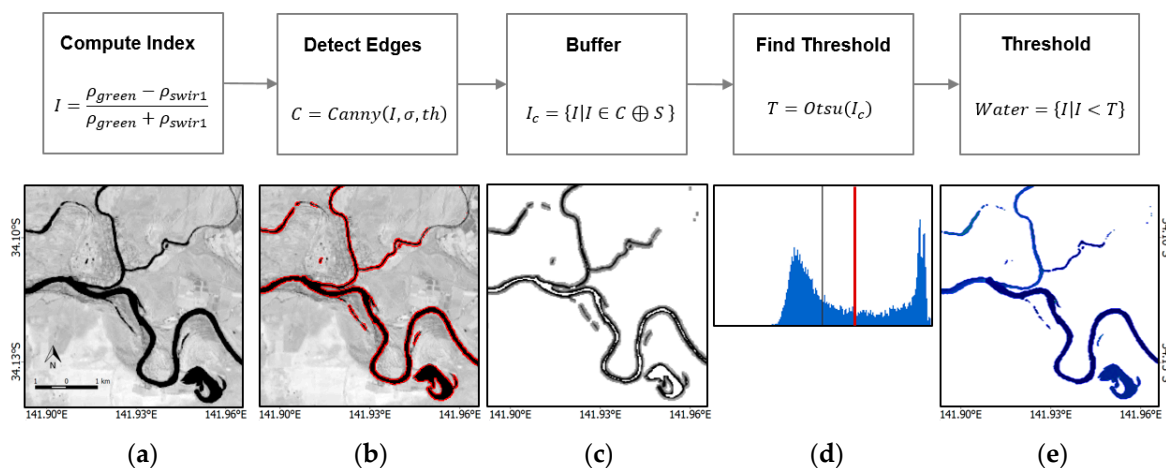

Figure 9. Method of water detection, consisting of the following steps: (**a**) compute water index; (**b**) compute edges using a Canny edge detector; (**c**) create a buffer around the edges using dilation; (**d**) compute threshold value using Otsu method, using only the pixels in the buffer (**e**) threshold index values.

The proposed method was applied to the cloud-free percentile images. In a case of two classes in the grid tile, we were able to get an almost perfect detection of water pixels using the following parameters: $\sigma = 0.7$, $th = 0.99$ for the Canny edge filter, and a structuring element with the size 15 m $\times$ 15 m to dilate the edges in Step (c) and create a surrounding buffer region. The σ and th parameters are used to define the standard deviation of the Gaussian smoothing kernel and the threshold used to define sensitivity of the filter, respectively.

2.4.3. Refining Water Detection Using Supervised Classification Based on CART and HAND

The method was applied over 1725 spatial boxes of 20 km $\times$ 20 km covering the Murray-Darling basin. The 20 km $\times$ 20 km area was chosen arbitrarily, and it is assumed that the threshold values are the same within each spatial box area. The resulting MNDWI threshold values have a range of -0.25 to 0.4 (Figure 10), which clearly highlights the need for varying the MNDWI threshold, spatially.

The use of simple water spectral indices in hilly areas is usually not sufficient, as they frequently result in very high commission errors due to false detection of water pixels. This misclassification is especially true for MNDWI, which is more sensitive to hill shadows than NDWI. The reason is the spectral signature of hill shadows, which looks similar to the one corresponding to water, resulting in large MNDWI values. To remove these errors, we have trained a CART classifier (Figure 11) using a manually-created training set, all Landsat 8 bands, and HAND. The training set was created only to include those pixels that were misclassified during the unsupervised step, which appeared to be true only in hilly areas. The HAND values used to train the classifier were included only when they were greater than 10 m. This constraint was introduced to reduce the influence of SRTM errors that get higher near water bodies.

(**a**)   (**b**)
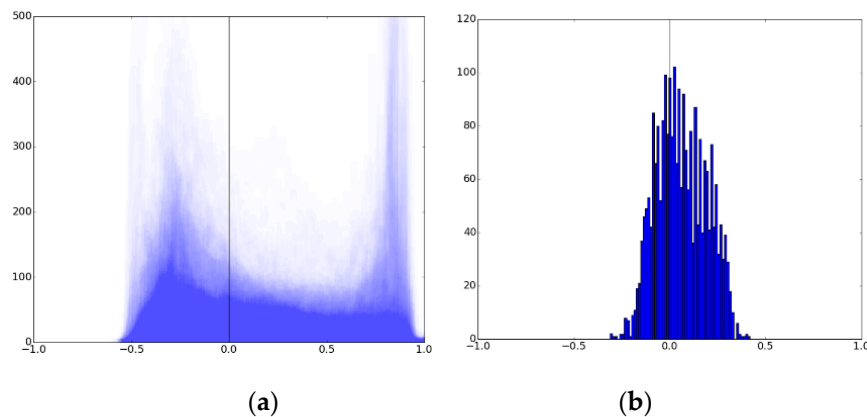
Figure 10. Histograms of MNDWI values within 15 m of detected edges (**a**) and a histogram of detected threshold values (**b**).
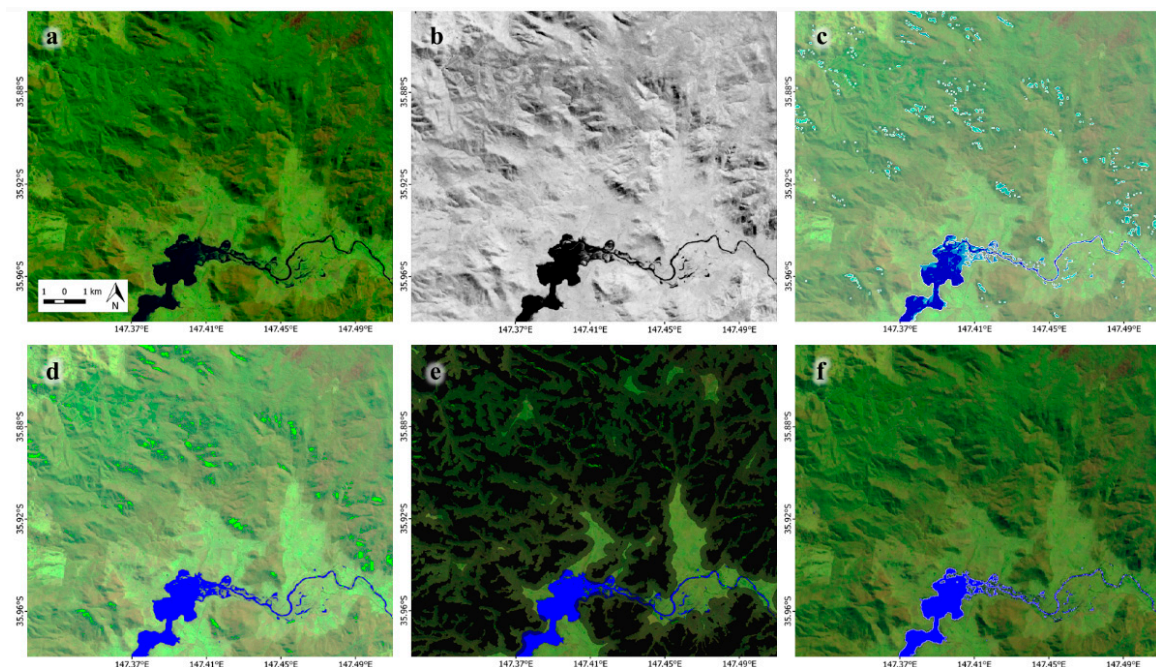

Figure 11. Steps performed during supervised classification post-processing step for hillslope areas. (tile 2847). False color 15% percentile composite image (**a**); MNDWI (**b**); water detected using adaptive threshold method (**c**); results of the application of the CART classifier (**d**); 300 m buffer areas around HAND > 30 m used to select final water mask (**e**); and the final water mask (**f**).

The classification was performed using a Google Earth Engine implementation of CART with a tree depth increased from 10 (default value) to 20. The larger tree depth was required to avoid overfitting and because our study basin covers a relatively large area, resulting in a large variation of water and land use types. Overfitting was detected by observing the confusion matrix generated after training the classifier. In the final training set, the confusion error was very close to zero.

The final training set contains about 500 polygons created manually and iteratively by training the classifier for one set of tiles and then validating it for all other tiles where supervised classification was required. This step was required mainly for the grid tiles located in the southern part of the catchment (hilly landscapes).

Since Landsat can reveal water features better than SRTM, we have also decided to analyze the HAND values of all pixels used during the training stage (Figure 12). The results reveal that HAND values can take values up to 40 m even though visual inspection shows that these pixels belong to water. A low SRTM accuracy can explain this mismatch near water. Another confusing result comes from a comparison of NDWI and MNDWI. Theoretically, they should be strongly correlated. However, substantial differences were observed in multiple locations. These differences occur because many pixels requiring manual classification are mixed pixels, partially covered by land. Further research would be necessary to explain these results.
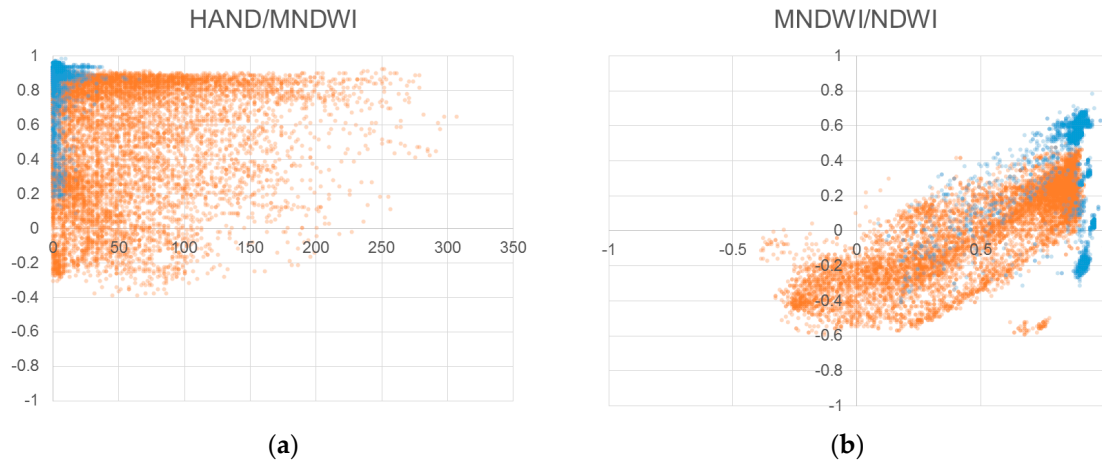
Figure 12. Values of HAND *vs.* MNDWI (**a**) and MNDWI *vs.* NDWI (**b**) used during training of the CART classifier. Orange and blue dots represent land and water, correspondingly.

2.5. River Centerline Estimation from Landsat 8 Water Mask

An additional skeletonization step was required before computing positional differences between OSM linear water features and water masks detected from Landsat imagery (or HAND). We have used the method of mathematical morphology [50] by applying an iterative thinning operator applying a hit-or-miss transform to a binary water mask image. The actual steps include:

$$S(W) = S^K \left(S^{K-1} \left(\ldots \left(S^0(W) \right) \right) \right) \tag{4}$$

$$S^k(W) = \left(W \otimes B^1 \right) \otimes B^2 \tag{5}$$

where S denotes a skeleton image constructed from W—original binary water mask, K—number of iterations for morphological skeletonization operator defined by S^k, B^1 and B^2—are the two composite structuring elements, and $\otimes$ is a binary thinning operator defined as:

$$W \otimes B = W - (W \odot B) \tag{6}$$

where $\odot$ is a hit-or-miss transform operator defined as:

$$W \odot B = (W \ominus B_1) \cap (W^c \ominus B_2) \tag{7}$$

where W^c is the set complement of W, and $\ominus$ is erosion operator, B_1 represents background and B_2—foreground parts of the composite structured element $B = (B_1, B_2)$.

The structuring elements B^1 and B^2 (Figure 13) allow reconstruction of river skeletons even without the need to introduce pruning (the removal of small branches), which is usually required

during such processing. The actual implementation performs hit-or-miss transform using four rotated versions of the structuring elements applied during every thinning iteration.
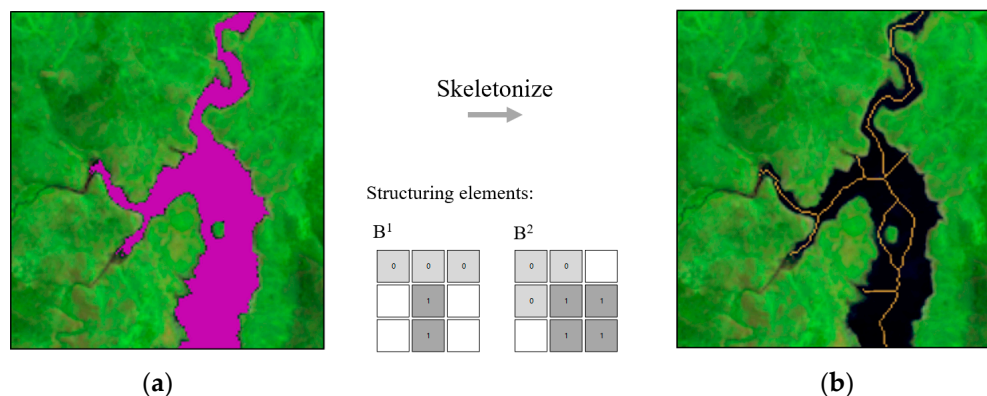
Figure 13. Input water mask (**a**) and results (**b**) of medial axis detection algorithm using mathematical morphology. Structuring elements (kernels) shown in the middle. Four rotated version of the kernels were applied sequentially.

Additionally, morphological smoothing was performed before the skeletonization step, to make sure fewer branches were present in the final centerline.

To implement most algorithms, we have used the Google Earth Engine (GEE) parallel computing platform [48]. Some of the computations, where GEE was less suitable, were performed using Google Compute Engine (GCE) running Ubuntu 15.04 (Vivid Vervet). The use of a dedicated machine was more suitable for the computation of hydrological parameters from the DEM, for which no algorithms are available yet in the GEE environment. JavaScript and Python were used as programming languages, as well as some open source tools and libraries (GDAL [41], Fiona [51], Shapely [52], ee-runner [53], and PCRaster [44]). We aimed at automating most of the steps to make sure that they would easily scale to a planetary scale, and that the results of the research can be reproduced when updated versions of the underlying datasets will be released.

3. Results

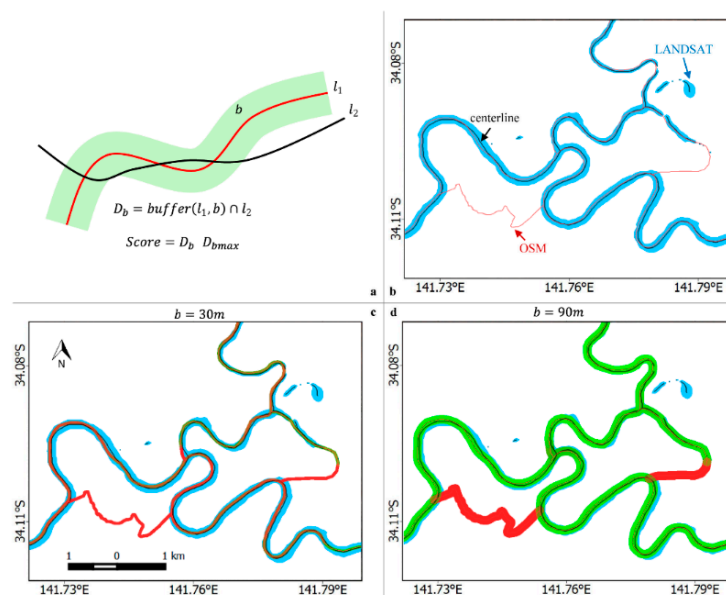### 3.1. Estimation of Positional Differences between Rivers

After the centerline is computed, positional differences can be easily computed using Goodchild's method of increasing overlay polygons.

The above method works best when the lengths of the segments and the distance between segments are significantly larger than the maximum buffer size D_{bmax} used during dilation.

3.2. Positional Differences between OpenStreetMap, Landsat, and SRTM

Goodchild's method to estimate positional differences was applied for every line segment of the OSM dataset and two water mask raster datasets: (1) the drainage network estimated using HAND and (2) water centerline estimated using the Landsat 8 water mask. The final differences were estimated using *Score = 0.85* as a threshold, which corresponds to a distance where 85% of the pixels from the second dataset (black line in Figure 14) are covered by a dilation via this distance. The frequency and cumulative histograms of the resulting differences are presented in Figure 15.

Figure 14. Method used to estimate positional differences between vector and raster datasets. Increasing overlay buffer method (**a**); example of OSM water polyline and Landsat mask (**b**); and results for two different buffer sizes (**c,d**).
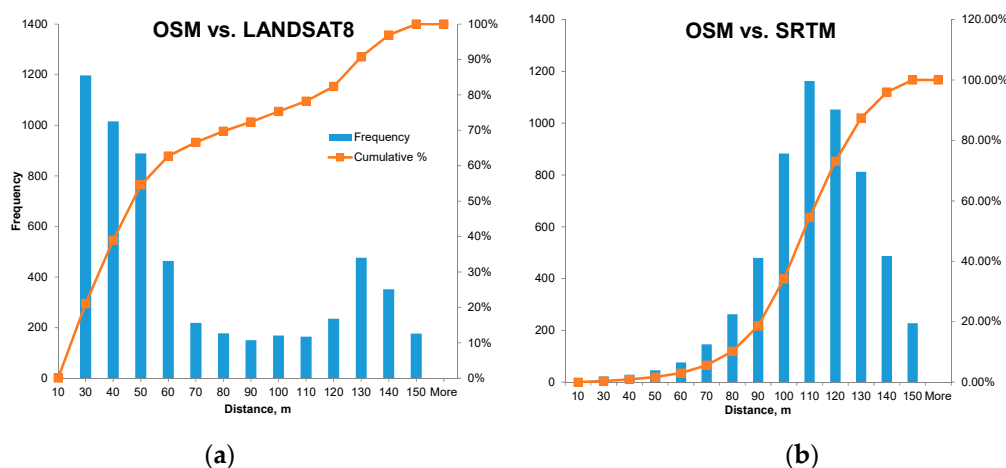
Figure 15. Frequency histogram of distances between OSM river segments, Landsat 8 centerline (**a**); and the SRTM (30 m) drainage network (**b**).

The second peak of the left histogram (distance > 100 m) occurs mainly due to wider water bodies, where the centerline cannot be reached within 20 thinning steps, but also for water bodies that start to overlap within the maximum size dilation, such as oxbows or meandering rivers located close to each other. Positional differences between Landsat centerlines and OSM become larger when a river width becomes larger. These large differences are observed because OSM polylines frequently represent the thalweg (the line of lowest elevation within a river) instead of a medial axis.

It can be seen that the distance between OSM water polylines and the centerlines computed using the Landsat 8 water mask falls in the range of 30–60 m for 60% of the water features. We were able to confirm this for about 17% (*N* = 5687) of the linear OSM water features. The length of the river segments was selected to be about 0.02 degree (~2.2 km).

In the case of SRTM, about 30% (*N* = 9887) of the OSM segments could be compared, mainly located in the southern, hilly part of the catchment (see Figure 16). The distances are slightly larger than those for Landsat and also appear normally distributed with a mean value of 110 m. The differences

can be explained by a systematic shift between the OSM and SRTM datasets, combined with other factors, like the fact that smaller river meanders are not resolved well by SRTM dataset.
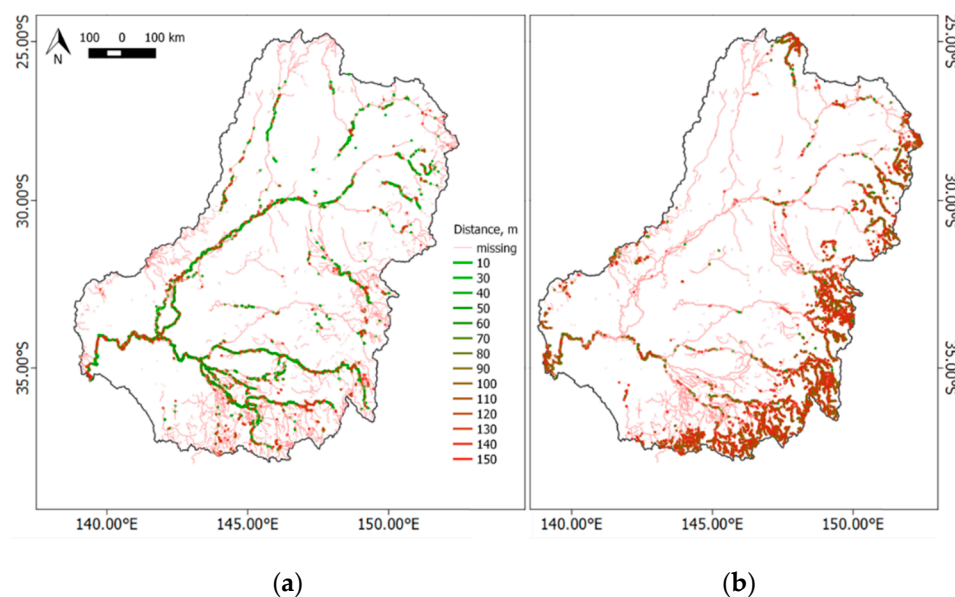
Figure 16. Positional differences between linear water features extracted from OSM, water mask centerline extracted from Landsat (**a**); and drainage network extracted from SRTM (**b**).

Spatial representation of the distances between OSM river segments and centerlines of the Landsat water mask, as well as the distances between OSM river segments and drainage network cells, are shown in Figure 16. It can be clearly seen that, in the case of Landsat, mainly large rivers could be compared due to the 30 m resolution limitation of the OLI sensor. At the same time, a much larger number of the OSM segments could be confirmed using the drainage network derived from 30 m SRTM, even though positional differences are not as good as in the case of Landsat. On the other hand, the drainage network derived from SRTM seems to be very inaccurate for flatter landscapes. Therefore, we have excluded drainage network pixels for the cases where HAND values are smaller than 30 m in the 300 m radius neighborhood. These parameters were identified empirically after careful visual inspection.

3.3. Goodness of Fit between OpenStreetMap and Landsat Water Masks

To estimate the overall overlap between water masks extracted from both OSM and Landsat, we have aggregated the overlap using a regular grid for a better understanding of the results. The analysis was performed using the actual water mask (polygonal and linear features of the OSM). For every grid cell, a total surface area of water mask was computed. For linear features, a dilation buffer of 15 m was applied using a square kernel to make sure it matches, at least, one Landsat grid cell. Then, the resulting thematic differences were computed (Figure 17).

While the centerline analysis reveals a good fit between Landsat and OSM, the surface area analysis demonstrates quite a large mismatch. There can be several explanations for this mismatch. Firstly, OSM (as well as Google Maps) misses a large number of small agricultural reservoirs in this area. Secondly, many of the large reservoirs are intermittent and were mostly dry during 2013–2015. Thirdly, riverbank information frequently does not match between two datasets; in some cases it is missing in OSM, in other instances Landsat misses small rivers (W < 30 m) or the water bodies are partially covered by vegetation.

The surface water area of both OSM and Landsat constitutes about 0.85% of the total catchment area. Only about one-third (32%) of a total surface water area can be observed using both datasets. The rest of the surface water can be seen using only Landsat or only OSM (Table 2).
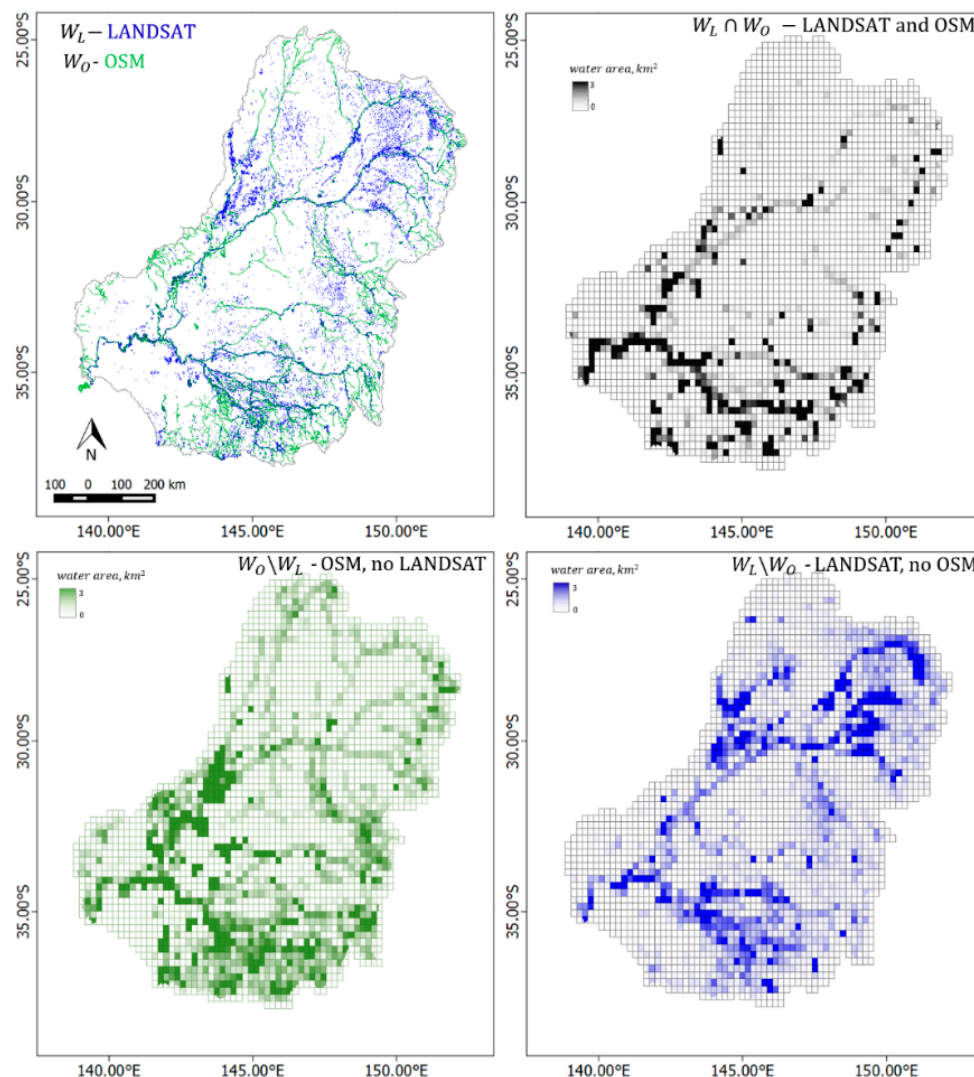
Figure 17. Thematic differences between water surface area provided by OSM and detected using Landsat 8.

Table 2. Overlap between surface water detected using OSM and Landsat.

Variable	Area, km^2	Ratio, %		
Total Water	8981	100%		
W_O	7073	79%		
W_L	4799	53%		
$W_L \cap W_O$	2891	32%		
$W_O \backslash W_L$	4182	47%		
$W_L \backslash W_O$	1908	21%		

The results of the research, including the newly derived permanent water mask, hydrological datasets, source code required to reproduce the analysis and a supporting website, can be found in supplementary materials accompanying the paper (Table A1, Figure A2).

4. Discussion

The research demonstrates clear benefits of the use of the new imagery acquired by Landsat 8 to detect water bodies when combined with water masks derived from other sources. We have also found the SRTM to be an excellent complementary dataset enabling improvement of the water mask detection method for hilly areas, after its transformation into HAND.

However, none of the three water masks were found to be perfect regarding positional differences and completeness. The main issue of the water masks derived from Landsat 8 is its shortcomings in detecting small water features, such as small rivers or man-made canals and detecting water bodies (partially) covered by riparian or surface water vegetation. The main limitation with the drainage network derived from SRTM is its inability to detect river features for flat terrain conditions. The latter constituted a major part of our study area. However, this might not be the case when applied elsewhere. An additional challenge is related to the presence of high-frequency noise and a relatively poor quality of SRTM near water bodies. The noise can be explained by the radar origin of the dataset.

One of the next logical steps of the present research could be the development of a data fusion algorithm using benefits from all three datasets. Such development would require the introduction of objective criteria regarding confidence of every water mask depending on topographic and other conditions. Another step might be to perform the same analysis globally. However, performing global analysis would require significantly larger computational efforts and includes both detection of the water mask and estimation of HAND at 30 m resolution. Additional validation of the OSM and its fusion with the datasets produced by the local governmental agencies (Surface Hydrology, Water Observations from Space, local high-resolution elevation models) will help in harmonizing existing vector and raster water mask datasets.

Possible improvements to the method of water detection might include utilization of the panchromatic band and entropy-based methods, in addition to the spectral methods. Additional significant improvements can be achieved through the use of the other medium or high-resolution satellite missions such as Sentinel 2, PlanetLabs, and SkyBox. The use of higher resolution imagery would allow detection of much smaller (width < 30 m) river features, resulting in improved coverage.

The method of water detection can be easily extended to use Landsat 7 or any other multi-spectral imagery, for example, to generate inter-annual water masks or to study water dynamics.

The proposed method of water detection might face difficulties in the areas where an insufficient number of cloud-free observations are available, for example, in very wet or cold climates. In this case, it might be difficult to determine a correct range of the cloud-free percentiles to be used for water detection.

5. Conclusions

The new method allows detection of a high-resolution water mask using optical multispectral satellite imagery. The resulting water mask, together with the newly-generated 30 m-resolution drainage network derived from SRTM, was compared to the water mask extracted from OSM.

Our results reveal that the best surface water mask should be generated by combining all three datasets that were analyzed in the current study; water masks extracted from OSM, optical satellite imagery and the drainage network derived from high-resolution digital elevation models for hilly areas.

We found a good agreement, concerning positional accuracy, between river water features from OSM and water masks derived from Landsat 8. However, only 32% of the total OSM and Landsat 8 water mask matches when analyzing the actual intersecting surface area.

The newly generated Landsat 8 water mask reveals many new water bodies previously not present in OSM or any other vector dataset we have explored. A large part constitutes a large number of small agricultural reservoirs located in the northern and southern parts of the catchment.

Supplementary Materials: The following are available online at www.mdpi.com/2072-4292/8/5/386/s1.

Author Contributions: Most of the work in the current study was done by Gennadii Donchyts as a part of his Ph.D. research. The other authors contributed by regularly participating in the work meetings, discussing intermediate results and contributing to the actual writing and review of the paper.

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

Appendix

All scripts used in the research, including the training set used for water mask refinement in hilly areas, can be found in the following GitHub repository [54]. The repository contains all scripts which allow one to: (1) extract water features from an OpenStreetMap planet file; (2) download SRTM DEM files clipped by HydroBASIN catchments and perform calculation of HAND using the Python version of PCRaster tools; (3) Google Earth Engine JavaScript scripts used to generate regular grids for processing, detect water masks for a given grid tile, as well as scripts required to skeletonize that water mask and compute distances between OSM segments and another river raster dataset (drainage network derived from HAND or Landsat water mask centerline). Additional supplementary scripts are available in a form of Jupyter Notebooks. These scripts were used to clean up HydroBASINS catchments, generate tiled version of HAND from basin-based images, split OpenStreetMap water geometries into smaller segments to perform local spatial analysis, as well as scripts used to generate histograms from GeoJSON files produced by water detection scripts.

A1. Grids Used During Analysis

The following two grids were used during analysis: (1) HydroBASINS level 8 catchment polygons and (2) regular ~20 × 20 km grids (Figure A1). The first one was used to parallelize the generation of hydrological parameters. The latter one was used to perform water detection and to visualize the results in an aggregated form, and also to parallelize the water detection analysis using Landsat imagery.
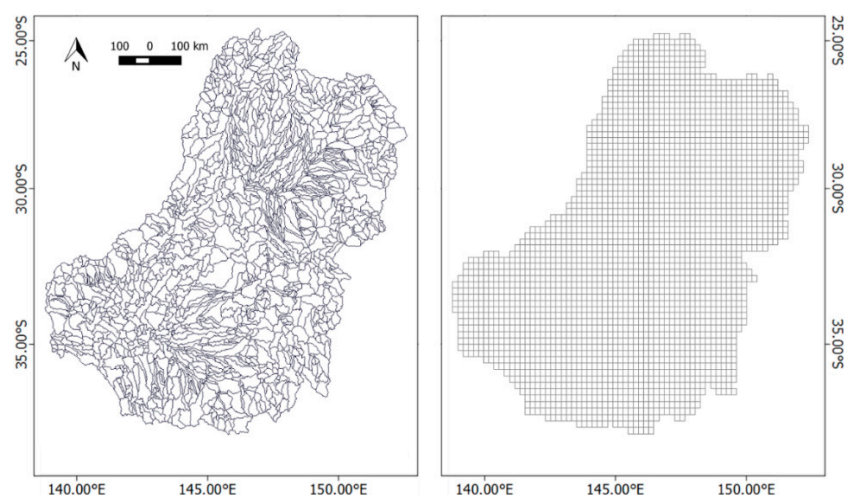
Figure A1. HydroBASINS Pfafstetter level 8 catchments and regular ~20 × 20 km grid.

A2. Results as Raster and Vector Datasets

Some of the datasets used during processing were uploaded and shared in a form of Google Fusion Tables and Google Earth Engine Raster Assets (Table A1).

Table A1. List of vector and raster datasets available online.

Name	Type	Link
OpenStreetMap water features for Australia	Fusion Table	[55]
HydroBASIN catchments for Australia	Fusion Table	[55]
Landsat water mask (raster)	EE Asset	users/gena/AU_Murray_Darling/MNDWI_15_water_WGS
Height above the nearest drainage (raster)	EE Asset	users/gena/AU_Murray_Darling/SRTM_30_Murray_Darling_hand
Local flow accumulation (raster)	EE Asset	users/gena/AU_Murray_Darling/SRTM_30_Murray_Darling_flow_accumulation
Distance to the nearest drainage (raster)	EE Assets	users/gena/AU_Murray_Darling/SRTM_30_Murray_Darling_dist
Google Earth Engine script	JavaScript	[56]

A3. Website

Most of the datasets produced by this research can be accessed using a website [57] dedicated to this study (Figure A2). The website is hosted in a Google Cloud and uses Google App Engine infrastructure, providing integration with Google Earth Engine where most of the datasets produced in this study can be found.
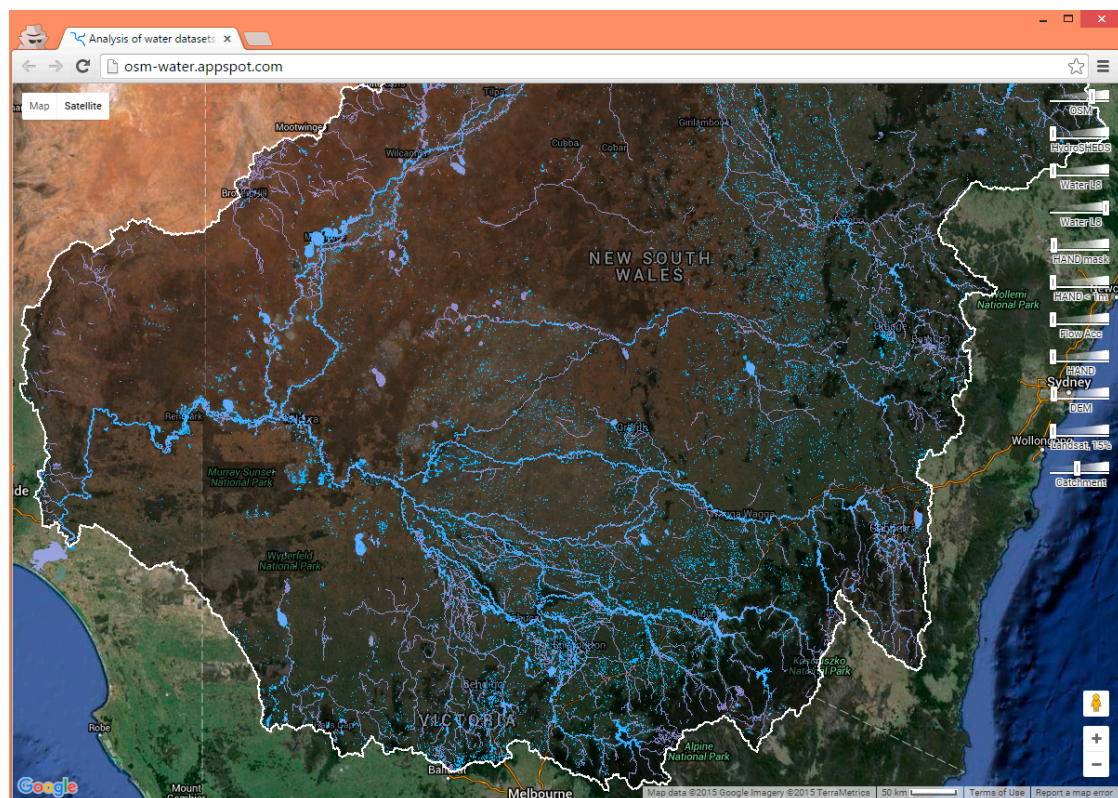
Figure A2. A screenshot of the website demonstrating the water mask derived from Landsat 8, OpenStreetMap, and 30 m SRTM.

References

1.   Maidment, D. *Arc Hydro: GIS for Water Resources*; ESRI Press: Redlands, CA, USA, 2002.

2.   Van Beek, L.; Bierkens, M. The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification. Available online: http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf (accessed on 10 April 2016).

3.   Wood, E.F.; Roundy, J.K.; Troy, T.J.; van Beek, L.P.H.; Bierkens, M.F.P.; Blyth, E.; de Roo, A.; Döll, P.; Ek, M.; Famiglietti, J.; *et al.* Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resour. Res.* **2011**, *47*. [CrossRef]

4.   Bierkens, M.F.P. Global hydrology 2015: State, trends, and directions. *Water Resour. Res.* **2015**, *51*, 4923–4947. [CrossRef]

5.   Gao, B.C. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [CrossRef]

6.   McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]

7.   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]

8.   Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [CrossRef]

9.   Zhu, Z.; Woodcock, C.E. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sens. Environ.* **2014**, *152*, 217–234. [CrossRef]

10.   Landsat Higher Level Science Data Products. Available online: http://landsat.usgs.gov/CDR_ECV.php (accessed on 1 January 2016).

11.   Tan, B.; Masek, J.G.; Wolfe, R.; Gao, F.; Huang, C.; Vermote, E.F.; Sexton, J.O.; Ederer, G. Improved forest change detection with terrain illumination corrected Landsat images. *Remote Sens. Environ.* **2013**, *136*, 469–483. [CrossRef]

12.   Potapov, P.V.; Turubanova, S.A.; Hansen, M.C.; Adusei, B.; Broich, M.; Altstatt, A.; Mane, L.; Justice, C.O. Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM+ data. *Remote Sens. Environ.* **2012**, *122*, 106–116. [CrossRef]

13.   Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; *et al.* High-resolution global maps of 21st-century forest cover change. *Science* **2013**, *342*, 850–853. [CrossRef] [PubMed]

14.   Ji, L.; Zhang, L.; Wylie, B. Analysis of dynamic thresholds for the normalized difference water index. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 1307–1317. [CrossRef]

15.   Li, W.; Du, Z.; Ling, F.; Zhou, D.; Wang, H.; Gui, Y.; Sun, B.; Zhang, X. A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. *Remote Sens.* **2013**, *5*, 5530–5549. [CrossRef]

16.   Yang, K.; Li, M.; Liu, Y.; Cheng, L.; Duan, Y.; Zhou, M. River delineation from remotely sensed imagery using a multi-scale classification approach. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4726–4737. [CrossRef]

17.   Liu, D.; Yu, J. Otsu method and K-means. In Proceedings of the HIS '09. Ninth International Conference onHybrid Intelligent Systems, Shenyang, China, 12–14 August 2009.

18.   Lehner, B.; Verdin, K.; Jarvis, A. New global hydrography derived from spaceborne elevation data. *EOS* **2008**, *89*, 93–94. [CrossRef]

19.   Tarboton, D.G. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resour. Res.* **1997**, *33*, 309. [CrossRef]

20.   Rennó, C.D.; Nobre, A.D.; Cuartas, L.A.; Soares, J.V.; Hodnett, M.G.; Tomasella, J.; Waterloo, M.J. HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. *Remote Sens. Environ.* **2008**, *112*, 3469–3481. [CrossRef]

21.   Nobre, A.D.; Cuartas, L.A.; Momo, M.R.; Severo, D.L.; Pinheiro, A.; Nobre, C.A. HAND contour: A new proxy predictor of inundation extent. *Hydrol. Process.* **2016**, *30*, 320–333. [CrossRef]

22.   Manfreda, S.; Nardi, F.; Samela, C.; Grimaldi, S.; Taramasso, A.C.; Roth, G.; Sole, A. Investigation on the use of geomorphic approaches for the delineation of flood prone areas. *J. Hydrol.* **2014**, *517*, 863–876. [CrossRef]

23.   Eilander, D.; Annor, F.O.; Iannini, L.; van de Giesen, N. Remotely sensed monitoring of small reservoir dynamics: A Bayesian approach. *Remote Sens.* **2014**, *6*, 1191–1210. [CrossRef]

24. Haklay, M. How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. *Environ. Plan. B Plan. Des.* **2010**, *37*, 682–703. [CrossRef]

25. OpenStreetMap Taginfo. Available online: http://taginfo.openstreetmap.org/tags (accessed on 24 November 2015).

26. Schellekens, J.; Brolsma, R.J.; Dahm, R.J.; Donchyts, G.V.; Winsemius, H.C. Rapid setup of hydrological and hydraulic models using OpenStreetMap and the SRTM derived digital elevation model. *Environ. Model. Softw.* **2014**, *61*, 98–105. [CrossRef]

27. Mooney, P.; Corcoran, P.; Winstanley, A.C. Towards quality metrics for OpenStreetMap. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information, New York, NY, USA, 5–7 September 2010.

28. Girres, J.F.; Touya, G. Quality Assessment of the French OpenStreetMap Dataset. *Trans. GIS* **2010**, *14*, 435–459. [CrossRef]

29. Goodchild, M.F.; Hunter, G.J. A simple positional accuracy measure for linear features. *Int. J. Geogr. Inf. Sci.* **1997**, *11*, 299–306. [CrossRef]

30. Barron, C.; Neis, P.; Zipf, A. A comprehensive framework for intrinsic OpenStreetMap quality analysis. *Trans. GIS* **2014**, *18*, 877–895. [CrossRef]

31. International Organization for Standardization: ISO 19113:2002: Geographic Information—Quality Principles. Available online: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26018 (accessed on 10 April 2016).

32. Crossman, S.; Li, O. *Surface Hydrology Lines (Regional)*, 2nd ed.; Australia. Available online: http://www.ga.gov.au/metadata-gateway/metadata/record/83130 (accessed on 5 April 2016).

33. Mueller, N.; Lewis, A.; Roberts, D.; Ring, S.; Melrose, R.; Sixsmith, J.; Lymburner, L.; McIntyre, A.; Tan, P.; Curnow, S.; *et al.* Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sens. Environ.* **2015**, *174*, 341–352. [CrossRef]

34. Digital Elevation Model (DEM) of Australia Derived from LiDAR 5 Metre Grid. Geoscience Australia, Canberra. Available online: http://dx.doi.org/10.4225/25/5652419862E23 (accessed on 5 April 2016).

35. 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]

36. NASA Shuttle Radar Topography Mission. Available online: http://www2.jpl.nasa.gov/srtm/ (accessed on 24 November 2015).

37. Murray-Darling Basin Authority. Available online: http://www.mdba.gov.au/about-basin (accessed on 1 January 2016).

38. Natural Earth Map Dataset. Available online: http://www.naturalearthdata.com/ (accessed on 5 April 2016).

39. OpenStreetMap Planet File. Available online: http://planet.osm.org/ (accessed on 24 November 2015).

40. Lehner, B.; Grill, G. Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems. *Hydrol. Process.* **2013**, *27*, 2171–2186. [CrossRef]

41. Warmerdam, F. The Geospatial Data Abstraction Library. Available online: http://link.springer.com/chapter/10.1007/978-3-540-74831-1_5 (accessed on 15 January 2016).

42. O'Callaghan, J.; Mark, D. The extraction of drainage networks from digital elevation data. *Comput. Vis. Graph. Image Process.* **1984**, *28*, 323–344. [CrossRef]

43. Karssenberg, D.; Schmitz, O. A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. *Environ. Model. Softw.* **2010**, *25*, 489–502. [CrossRef]

44. PCRaster Web Page. Available online: http://pcraster.geo.uu.nl/ (accessed on 1 May 2016).

45. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]

46. Breiman, L.; Friedman, J.; Stone, C.; Olshen, R. *Classification and Regression Trees*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 1984.

47. Rodriguez, E.; Morris, C.; Belz, J. An assessment of the SRTM topographic products. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 249–260.

48. Gorelick, N. Google Earth Engine. Available online: http://adsabs.harvard.edu/abs/2012AGUFM.U31A.04G (accessed on 28 April 2016).

49. Google Fusion Table Web Page. Available online: http://tables.googlelabs.com (accessed on 1 May 2016).

50. Serra, J. *Image Analysis and Mathematical Morphology*; Academic Press, Inc.: Orlando, FL, USA, 1982; Volume 1.

51. Fiona GitHub Web Page. Available online: https://github.com/Toblerity/Fiona (accessed on 28 March 2016).

52. Shapely GitHub Web Page. Available online: https://github.com/Toblerity/Shapely (accessed on 1 January 2016).

53. Ee-Runner GitHub Web Page. Available online: https://github.com/gena/ee-runner (accessed on 1 January 2016).

54. Supplementary GitHub Repository Web Page. Available online: http://github.com/gena/paper-osm-2015 (accessed on 1 May 2016).

55. Supplementary Google Fusion Tables Web Page. Available online: http://bit.ly/paper-osm-2016-fusion-tables (accessed on 1 May 2016).

56. Supplementary Google Earth Engine Script Web Page. Available online: http://bit.ly/paper-osm-2016-gee-assets (accessed on 1 May 2016).

57. Supplementary Project Web Page. Available online: http://osm-water.appspot.com (accessed on 1 May 2016).