



## Article

# A Multi-Stage Approach Combining Very High-Resolution Satellite Image, GIS Database and Post-Classification Modification Rules for Habitat Mapping in Hong Kong

Ivan H. Y. Kwong <sup>1,2,\*</sup> , Frankie K. K. Wong <sup>1</sup> , Tung Fung <sup>1,2</sup> , Eric K. Y. Liu <sup>3</sup>, Roger H. Lee <sup>3</sup> and Terence P. T. Ng <sup>3</sup>

<sup>1</sup> Department of Geography and Resource Management, The Chinese University of Hong Kong, Hong Kong; kkit@cuhk.edu.hk (F.K.K.W.); tungfung@cuhk.edu.hk (T.F.)

<sup>2</sup> Institute of Future Cities, The Chinese University of Hong Kong, Hong Kong

<sup>3</sup> Agriculture, Fisheries and Conservation Department, 5/F, Cheung Sha Wan Government Offices, 303 Cheung Sha Wan Road, Kowloon, Hong Kong; eric\_ky\_liu@afcd.gov.hk (E.K.Y.L.); roger\_h\_lee@afcd.gov.hk (R.H.L.); pun\_tung\_ng@afcd.gov.hk (T.P.T.N.)

\* Correspondence: khoyinivan@link.cuhk.edu.hk

**Abstract:** Identification and mapping of various habitats with sufficient spatial details are essential to support environmental planning and management. Considering the complexity of diverse habitat types in a heterogeneous landscape, a context-dependent mapping framework is expected to be superior to traditional classification techniques. With the aim to produce a territory-wide habitat map in Hong Kong, a three-stage mapping procedure was developed to identify 21 habitats by combining very-high-resolution satellite images, geographic information system (GIS) layers and knowledge-based modification rules. In stage 1, several classification methods were tested to produce initial results with 11 classes from a WorldView-2/3 image mosaic using a combination of spectral, textural, topographic and geometric variables. In stage 2, modification rules were applied to refine the classification results based on contextual properties and ancillary data layers. Evaluation of the classified maps showed that the highest overall accuracy was obtained from pixel-based random forest classification (84.0%) and the implementation of modification rules led to an average 8.8% increase in the accuracy. In stage 3, the classification scheme was expanded to all 21 habitats through the adoption of additional rules. The resulting habitat map achieved >80% accuracy for most of the evaluated classes and >70% accuracy for the mixed habitats when validated using field-collected points. The proposed mapping framework was able to utilize different information sources in a systematic and controllable workflow. While transitional mixed habitats were mapped using class membership probabilities and a soft classification method, the identification of other habitats benefited from the hybrid use of remote-sensing classification and ancillary data. Adaptive implementation of classification procedures, development of appropriate rules and combination with spatial data are recommended when producing an integrated and accurate map.

**Keywords:** Multi-stage approach; post-classification modification; habitat mapping



**Citation:** Kwong, I.H.Y.; Wong, F.K.K.; Fung, T.; Liu, E.K.Y.; Lee, R.H.; Ng, T.P.T. A Multi-Stage Approach Combining Very High-Resolution Satellite Image, GIS Database and Post-Classification Modification Rules for Habitat Mapping in Hong Kong. *Remote Sens.* **2022**, *14*, 67. <https://doi.org/10.3390/rs14010067>

Academic Editors: Cristina Tarantino and Javier Marcello

Received: 8 November 2021

Accepted: 21 December 2021

Published: 24 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Identification and mapping of natural and artificial habitats can serve as the basis for assessments of biodiversity and ecosystem services, thus supporting environmental planning and management [1,2]. By reflecting changing ecological patterns at different spatial and temporal scales, habitat mapping provides baseline data to understand potential anthropogenic pressures and establish conservation policies [3,4]. Compared to traditional field surveys, remote sensing offers a cost-effective, rapid and repeatable option for habitat mapping, as it provides a synoptic view of phenomena on the ground continuously and consistently from a wide range of sensors with various spatial and spectral resolutions [5,6]. Medium-resolution imageries, such as those from Landsat satellites at a spatial resolution of

30 m, have been used to produce land cover products on regional and national scales [4,7]. By contrast, these datasets were argued to be lacking sufficient spatial and thematic details for effective monitoring by local governments and communities [2,8]. Very high-resolution (VHR) imageries (pixel size <5 m) potentially enable fine-scale mapping of small habitat patches in spatially heterogeneous landscapes that can meet the demand of end-users [9,10].

Despite the well-established benefits of VHR satellite images [11], their applications to map habitats over large geographical areas are uncommon [12,13]. In addition to the time and cost requirements, several challenges to extract information from multiple scenes were discussed in [14,15], including image acquisition, pre-processing, spatial diversity and temporal heterogeneity. Nagendra et al. [8] added that provision of too much detail in VHR images can decrease the accuracy in their classification. Furthermore, current applications of VHR images often focused on distinctions within specific physiognomic types, such as tree species [16], forest health [17] and grassland mapping [18,19]. In Hong Kong, existing habitat mapping exercises using VHR images also concentrated on small areas in a country park [20,21] or wetlands in a nature reserve [22,23]. A territory-wide and seamless characterization of all major habitat types with sufficient spatial details, providing an integrated view of the entire mosaic for local environmental management [24], is lacking in many cities.

To map land cover patterns from satellite observations, various classification methods have been gaining considerable attention including several relatively mature machine learning algorithms [25,26]. For instance, support vector machine (SVM) [27] and random forest (RF) [28] are well-known classifiers that achieve promising results in the literature [29,30]. Object-based classification also emerged as a popular approach using segmented image objects as the classification unit, which is expected to be of most benefit when a habitat area is divided into many pixels in VHR images [12,31]. Nevertheless, there is no clear consensus on the performance of classification methods for all purposes. While comparative studies have shown the best performances could be obtained from different classifiers when applied to different datasets [25,32], they argued that the choice of optimal algorithm is case-specific, and evaluation of multiple methods is recommended [33]. Besides the classification algorithms, the results can be dependent on several crucial steps in the mapping process such as characteristics of the study area, classification scheme, characteristics of data sources, use of different types of variables and ancillary data [34,35]. Specific to the scope of this study, it remains unknown which mapping procedure and classifier can be efficient for large data volume, robust to small pixel size in VHR image and able to distinguish subtle habitat differences.

Compared to the use of different elements in a single classification process, a multi-stage mapping approach is examined to a lesser extent. By separating different land cover types into a series of sub-classifications, the final results were found to be better than direct classifications of all classes [36,37]. Some scholars adopted a similar idea of a hierarchical classification framework, which usually identifies several main classes in the first level and gradually distinguishes different subtypes of classes within each main class [7,38]. Another attractive but also less investigated way to improve the mapping results is to apply post-classification modification rules with thematic layers in a geographic information system (GIS), such as terrain, land use, climatic and geological data [39,40]. The combination of remotely sensed and ancillary data through knowledge-based rules was demonstrated to provide contextual information that could enhance the classification accuracies [41–43].

In the context of habitat mapping, it was argued that habitat mapping is less straightforward and much harder to undertake compared to the delineation of land cover classes [8,44], but it was also suggested that decision rules can be developed to translate land cover maps to habitat categories based on different criteria [10,45]. Considering the complexity of diverse habitat types which makes it difficult for automated image classification to be optimal, a dedicated and context-dependent classification framework is expected to be superior to traditional classification techniques [46]. To date, there has been little work on the combination of a multi-stage mapping approach with a local GIS database for a habitat mapping



exercise. In particular, although it is possible to infer habitat and ecological properties indirectly from land cover categorization and spatially referenced information [3,14,45], a multi-stage framework from VHR satellite images to end products in city-scale is not yet available.

The major objective of this study is to produce a high-resolution territory-wide terrestrial habitat map for Hong Kong, a city with heterogeneous landscapes and high biodiversity. A multi-stage approach was developed to facilitate effective mapping of diverse habitat classes through the integration of VHR satellite image, GIS database and post-classification rules. In the first stage, initial classification was performed on a 2 m WorldView-2/3 image to map several classes with a variety of variables and classification methods. In the second and third stages, modification procedures were adopted with GIS data and spatial relationships to identify further habitats and produce the final results. Specifically, this study aimed to evaluate (i) the performance of different classification methods, (ii) the potential improvement provided by post-classification rules compared to initial results, and (iii) the effectiveness of the proposed multi-stage approach especially in identifying complex habitats such as transitional ecotones and those related to human activities. This suite of processing techniques was incorporated to provide the best solution in this specific study context. Through collaboration with local ecologists and policymakers, the mapping products were also believed to be of practical significance for future planning of the city.

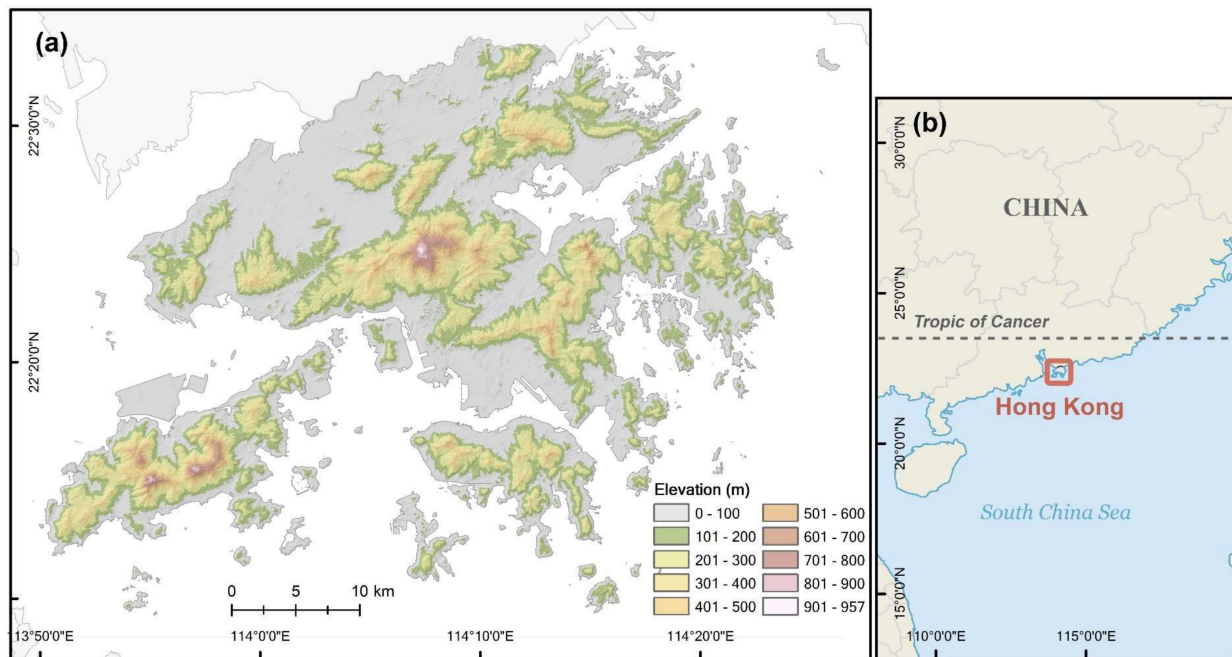
## 2. Materials and Methods

### 2.1. Study Area

Hong Kong, a city with around 1110 km<sup>2</sup> land area, lies at the northern limits of the Asian tropics between latitudes 22°08' N and 22°35' N and longitudes 113°49' E and 114°31' E. The climate is subtropical, with hot wet summer and cool dry winter. Located on the coast of South China Sea, the city has more than 700 km of coastline and more than 200 offshore islands. The terrain is mountainous and rugged, with the landscape rising from sandy beaches and rocky foreshores to the highest point of 957 m at Tai Mo Shan in the New Territories. Roughly 60% of the land areas are covered by natural terrain and about 40% of land is designated as protected areas.

Despite the small territory size and the densely populated urban environment, the topography and subtropical climate nurture a wide range of habitats in Hong Kong to support rich biodiversity with more than 3300 species of vascular plants [47]. While local studies suggested that natural succession and afforestation projects have brought rapid changes in the woodland–shrubland–grassland continuum in past years [21], urbanization and anthropogenic activities were identified as threats to some priority habitats [47]. In view of the conservation challenges, the Hong Kong government has recently formulated the first city-level Biodiversity Strategy and Action Plan [48] and one of the specific actions is to improve our knowledge by compiling an updated territorial habitat map.

The study area comprises the entire 1110-km<sup>2</sup> terrestrial area in Hong Kong (Figure 1). The climax vegetation belongs to evergreen broadleaf forest of the subtropical flora, but due to massive clearance during the Second World War, the majority of the existing vegetation including secondary forest is developed from tree planting and natural succession in the latter half of the 20th century [49]. The major types of vegetation in Hong Kong are woodland, shrubland and grassland, while other habitats are found in relation to freshwater and coastal environments.



**Figure 1.** (a) The study area of this study; (b) location of Hong Kong in East Asia.

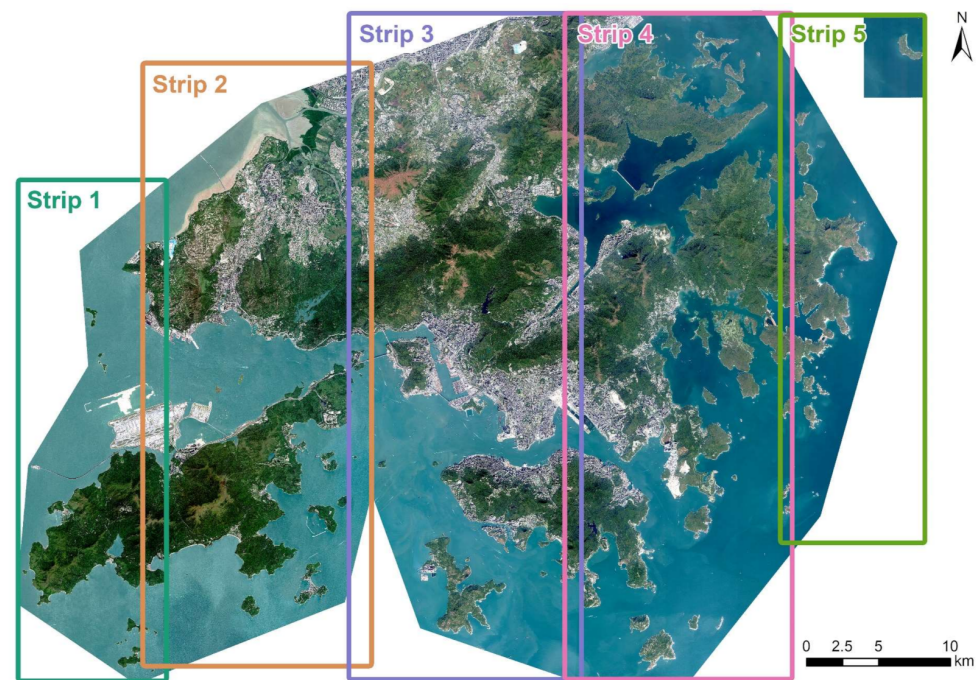
## 2.2. Data

### 2.2.1. WorldView-2 and -3 Satellite Image

VHR images from WorldView-2 and -3 satellites with eight multispectral bands at 2 m spatial resolution were used as the major data source. The high spatial resolution provides sufficient resolving power for relatively small features in the study area. In addition to the four conventional bands (Blue, Green, Red, Near-infrared), WorldView-2 and -3 provide four new spectral bands (Coastal blue, Yellow, Red Edge, Near-infrared 2) which were proved to be effective in vegetation studies [16]. Two strips of WorldView-3 imagery acquired on 22 September 2019 and three strips of WorldView-2 imagery acquired on 14 December 2019 were combined to achieve complete cloud-free coverage of the study area (Figure 2). The two WorldView satellites bear strong resemblances to each other and the similar acquisition dates ensure consistency in illumination and surface conditions. Table 1 shows the scene ID and acquisition parameters of each image. The raw images were ortho-rectified using ground control points with sub-pixel accuracies and converted into surface reflectance values using ATCOR-3 (Atmospheric and Topographic Correction) model [50] in PCI Geomatica 2018, followed by mosaicking into a single image.

**Table 1.** Information of the WorldView-2 (WV-2) and -3 (WV-3) images acquired for this study.

Strip	Satellite	ID	Date	Off Nadir (°)	Target Azimuth (°)	Coverage Area (km <sup>2</sup> )
1	WV-3	10400100528D3800	22 September 2019	23.0	128.4	236
2	WV-3	1040010052065F00	22 September 2019	24.3	114.0	547
3	WV-2	10300100A19B0600	14 December 2019	7.8	327.0	651
4	WV-2	103001009D694A00	14 December 2019	12.6	353.8	697
5	WV-2	103001009C0BA500	14 December 2019	17.3	2.2	180



**Figure 2.** Mosaic of the satellite image consisted of five strips of WorldView-2 and -3 images used in this study.

### 2.2.2. Field Survey

To collect reference data for training and accuracy assessment, 45 days of field surveys were carried out from 13 January 2020 to 1 February 2021. Considering the heterogeneous landscape in Hong Kong, the routes were planned with reference to an existing land utilization map [51] to facilitate selection of points covering various habitat types. Stratified random sampling was first performed to identify regions of interest from the existing map according to different land uses, followed by formulation of survey routes that could include diverse habitats within a feasible distance and cover different parts of the territory with various altitudes.

The actual survey points were carefully selected during the survey, satisfying criteria including uniform spatial coverage and high representativeness of specific habitat types and sufficient distance from other survey points. A  $10 \times 10$  m area was adopted as the mapping unit in the field [52]. The position of each survey point was recorded using a survey-grade Trimble R10 GNSS system with sub-meter accuracy. Local plant experts were included in the survey team to visually examine the sites and identify the habitat types according to the classification scheme (Section 2.3). Supplementary information related to each site such as vegetation condition and species composition was also investigated and recorded.

A total of 938 points were collected in all field surveys. Although the survey points were selected along accessible routes [52], each habitat type was represented by adequate numbers of points with large ranges of spatial and structural variabilities over the study area. Spatial distribution and summary of the survey points are shown in the supplementary materials (Figures S1 and S2; Tables S1 and S2). We used 283 survey points as training data while the remaining 655 points were used as validation data.

### 2.2.3. Geographic Information System (GIS) Database

In addition to the satellite image, ancillary layers from existing local GIS databases (Table 2) were included to provide supplementary information in the mapping procedures. For example, a high-resolution digital elevation model (DEM) was obtained from an airborne LiDAR survey covering the whole territory in 2011 [53]. Some of the GIS layers,

such as coastline, cultivated land and urban parks, were extracted as shapefiles from the iB5000 digital topographic map maintained continuously by the Survey and Mapping Office under the Lands Department of the Hong Kong government [54]. A few datasets were also provided by the Agriculture, Fisheries and Conservation Department of the Hong Kong government, including tree planting records and locations of seagrasses, which were both gathered from long-term monitoring programmes. The GIS databases had been updated within 1 year of the WorldView image acquisition and could facilitate integration for temporal analysis, except for the LiDAR data for which negligible changes in the terrain were assumed. The use of these GIS layers would be described in later sections.

**Table 2.** List of ancillary local geographic information system (GIS) data applied in this study. The reference dates refer to dates of light detection and ranging (LiDAR) data collection (for DEM), image acquisition (for artificial hard shoreline) and last update dates of the GIS database (for other layers).

Layer	Source	Reference Date
Digital elevation model (DEM)	Airborne LiDAR survey (Civil Engineering and Development Department of Hong Kong Government) [53]	1 December 2010–8 January 2011
Coastline	iB5000 digital topographic map (Survey and Mapping Office, Lands Department of Hong Kong Government) [54]	23 January 2019
Cultivated land		
Urban park		
Pond		
Reservoir		
Tree planting record	Agriculture, Fisheries and Conservation Department of Hong Kong Government	30 April 2019
Seagrass		
Building shadow	In-house computation from building height and solar angle at image acquisition time	23 January 2019
Artificial hard shoreline	Manual digitization from satellite image	22 September 2019–14 December 2019

### 2.3. Classification Scheme

Previous studies in Hong Kong have suggested that classification schemes for subtropical regions are not well developed and the selection of classes would need to be study area-specific [20]. Habitat mapping in Hong Kong in earlier periods have developed a classification scheme with 34 different categories, which were digitized through visual interpretation of aerial photographs [55]. Later updates of the map simplified the scheme to 24 habitat classes to facilitate mapping using satellite images and automated classification [56]. This study followed the classification scheme used in [56] and further revised it to 21 habitat categories (Table 3) based on multi-disciplinary expertise from local ecologists and remote-sensing experts. Similar to other studies in this region [20,21], habitats were classified mainly based on structural characteristics, since current vegetation in Hong Kong has been developed mainly through structural succession [47]. Habitats located in the intertidal zone were also included in the classification scheme and mapped in this study according to the satellite observation.

**Table 3.** Habitat classification scheme and definitions adopted in this study.

Habitat	Definition
Woodland	Rural lands mainly covered by tree species.
Shrubland	Rural lands mainly covered by shrub species.
Grassland	Rural lands mainly covered by grass species.

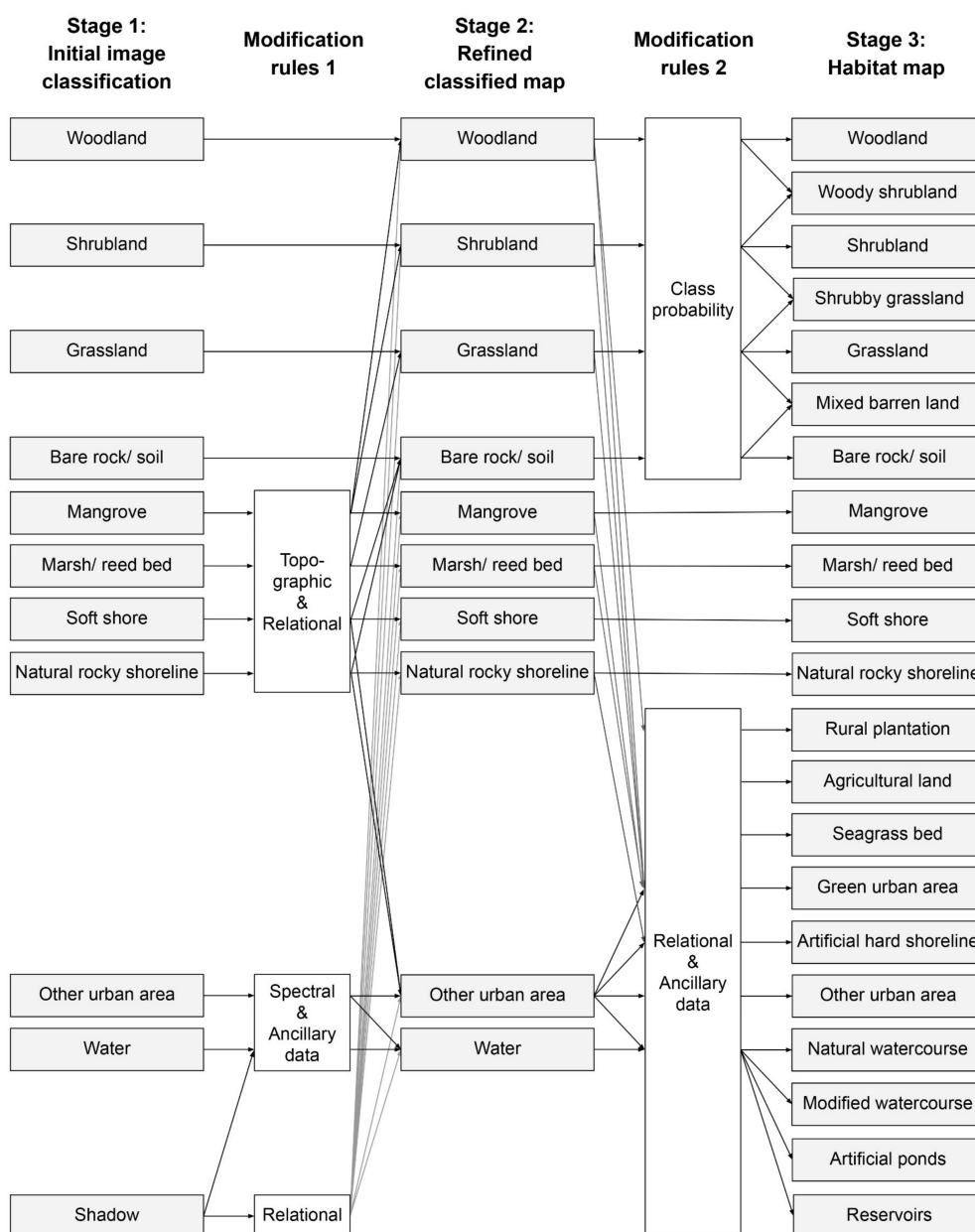


Table 3. Cont.

Habitat	Definition
Rural plantation	Rural lands mainly covered by woody plants and the top canopy is dominated by manually planted species in an organized and systematic way.
Marsh/reed bed	Lands, including abandoned agricultural land, covered with shallow waters and dominated by hydrophytes seasonally or all year round.
Mangrove	Coastal lands covered by true mangrove plant species.
Seagrass bed	Coastal lands covered by seagrass species.
Soft shore	Coastal lands of fine-grained sediment (i.e. sand, silt or finer particles) between high and low tide marks.
Natural rocky shoreline	Coastal lands of rocks between high and low tide marks.
Bare rock/soil	Natural open rock faces or disturbed lands, or "badlands" denuded of vegetation.
Natural watercourse	Rivers and streams experiencing natural flow patterns in unchanneled watercourse beds and banks.
Modified watercourse	Channelized rivers and streams, which are often without natural banks and beds, and are not subject to natural flow patterns (e.g., drainage channels and nullahs).
Reservoirs	Artificial lake used as a source of water supply.
Artificial hard shoreline	Man-made intertidal hard shore habitats (e.g., seawalls, jetties, groins and piers).
Artificial ponds	Small artificial water bodies constructed for the aquaculture purpose (e.g., gei wai and fishponds).
Agricultural land	Lands currently under cultivation, and lands not currently under land cultivation and yet to transform into other habitats such as marsh/reed bed.
Green urban area	Urban lands undergone artificial greening for various purposes (e.g., golf area courses, urban parks, and vegetation on the roadside).
Other urban area	Lands occupied by urban, other highly modified habitats (e.g., quarry, landfill) or industrial storage/containers.
Woody shrubland	Rural lands covered by a mixture of wood and shrub species and each of them occupies at least 1/3 of the coverage.
Shrubby grassland	Rural lands covered by a mixture of shrub and grass species and each of them occupies at least 1/3 of the coverage.
Mixed barren land	Rural lands covered by a mixture of grass and bare rock/ soil and each of them occupies at least 1/3 of the coverage.

#### 2.4. Multi-Stage Mapping Approach

In this study, a three-stage procedure was designed to map 21 habitats in separate stages based on their characteristics (Figure 3). In the first stage, initial classification results were produced from the WorldView-2/3 image mosaic with 11 distinctive categories which are shown in the first column of Figure 3. The initial results were modified into a classification map with 10 classes in the second stage and further transformed into a habitat map with 21 classes in the last stage, through modification procedures with GIS data and spatial relationships. The modification rules included spectral, topographic, relational, class probability and ancillary data rules, which would be explained in corresponding stages (Sections 2.6 and 2.7). A minimum mapping unit (MMU) of 100 m<sup>2</sup> (25 pixels) was also defined by considering the level of image details and size of habitats observed in the field.



**Figure 3.** The three-stage classification framework adopted in this study.

### 2.5. Stage 1: Initial Image Classification

In the first stage, both pixel-based and object-based supervised classification approaches were used [31] to identify the optimal classification result. In object-based approach, large-scale mean-shift segmentation in Orfeo Toolbox [57] was applied to segment the image. Two segment sizes were adopted in this step to derive information in different scales as suggested in [58], including a size of 20 pixels to delineate fine-scale features similar to the defined MMU and a coarser size of 80 pixels to depict larger habitats.

#### 2.5.1. Variables

Besides the spectral reflectance values from the WorldView-2/3 bands, another set of variables was generated from spectral and spatial domains as classification inputs.

Spectral indices—many spectral indices can be calculated by combining the reflectance at two or more wavelengths. In this study, several indices (Table 4) which are able to quantify vegetation characteristics and adopted in similar studies [59,60] were selected,

including the Normalized Difference Vegetation Index (NDVI) [61], Enhanced Vegetation Index (EVI) [62] and Green Normalized Difference Vegetation Index (GNDVI) [63]. Two indices aimed to utilize the availability of red-edge band in WorldView-2/3 images [59], including the Red Edge Normalized Difference Vegetation Index (RENDVI) [63] as well as the Modified Chlorophyll Absorption in Reflectance Index (MCARI) [64].

**Table 4.** Spectral indices used as classification variables.

Variable	Equation	Reference
Normalized Difference Vegetation Index (NDVI)	$\frac{NIR1-Red}{NIR1+Red}$	[61]
Enhanced Vegetation Index (EVI)	$2.5 \times \frac{NIR1-Red}{NIR1+6 \times Red-7.5 \times Blue+1}$	[62]
Green Normalized Difference Vegetation Index (GNDVI)	$\frac{NIR1-Green}{NIR1+Green}$	[63]
Red Edge Normalized Difference Vegetation Index (RENDVI)	$\frac{RedEdge-Red}{RedEdge+Red}$	[63]
Modified Chlorophyll Absorption in Reflectance Index (MCARI)	$[(RedEdge - Red) - 0.2 \times (RedEdge - Green)] \times \frac{RedEdge}{Red}$	[64]

Textures—grey level co-occurrence matrix (GLCM) is a second-order metric computing the local variation of pixel values with surrounding pixels. Principal component analysis was first performed on the image to extract the first principal component as the basis to compute GLCM [65]. Since GLCM can be computed by moving a window size in four directions (0°, 45°, 90°, 135°), an average of all directions was used to accommodate features with varying shapes [66]. Eight GLCM statistics can be calculated in ENVI 5.5 software. To identify suitable statistics and window sizes, a preliminary analysis was performed by evaluating the importance scores of all statistics from 5 × 5 to 15 × 15 window sizes using cross-validations of field survey data. Based on the results, a total of 10 GLCM features were adopted in this study, including mean (13 × 13, 15 × 15), contrast (11 × 11, 13 × 13, 15 × 15), dissimilarity (15 × 15), entropy (15 × 15) and correlation (7 × 7, 9 × 9, 15 × 15).

Terrain—topographic variables were computed from the 2 m DEM using ArcMap 10.5.1. The variables included slope and aspect, both in degree units.

Geometry (object-based only)—geometric properties were also computed for each segmented object in object-based classification, including area, compactness and rectangularity. Compactness and rectangularity were calculated as the ratios of segment areas to the areas of minimum bounding circle and rectangle [67], which had the largest values when the object was a circle and a rectangle, respectively.

For object-based classification, spectral statistics for each segmented object were generated from the underlying pixels. The variables included mean and standard deviation, characterizing average values and variations within the objects. The numbers of variables included in pixel- and object-based classification were 25 and 47, respectively (Table 5).

**Table 5.** List of variables used in pixel- and object-based classifications.

Classification	Type	Description	Number of Variables
Pixel-based classification	Spectral bands	WorldView-2/3 bands (Coastal blue, Blue, Green, Yellow, Red, Red-edge, Near-infrared [NIR]-1, NIR-2)	8
	Spectral indices	NDVI, EVI, GNDVI, RENDVI, MCARI	5
	Textures	Grey level co-occurrence matrix (GLCM) features	10
	Terrain	Slope, Aspect	2
			Total: 25

Table 5. Cont.

Classification	Type	Description	Number of Variables
Object-based classification	Spectral band statistics	Means and standard deviations of eight bands at 20 segmentation scale	16
		Means and standard deviations of eight bands at 80 segmentation scale	16
	Spectral indices	Means and standard deviations of five spectral indices	10
	Terrain	Slope, Aspect	2
	Geometry	Area, Compactness, Rectangularity	3
			Total: 47

### 2.5.2. Training Data

Training data used in the classification process were obtained from part of the field survey described in the previous section as well as visual interpretation of the satellite image. While the field surveys focused more on vegetated habitats which could be difficult to accurately select by viewing the satellite image, visual interpretation served as an efficient way to add a complementary set of spatially well-distributed training data especially in inaccessible areas and create balanced data for all classes.

The same set of training sites was applied to both pixel- and object-based classifications to ensure consistency for comparison. Underlying pixels were used as training data in a pixel-based approach while intersected segmented objects were used in an object-based approach. The training dataset included 16,035 pixels (10,135 from field survey and 5900 from visual selection) for pixel-based and 836 objects (485 from field survey and 351 from visual selection) for object-based classifications. The difference in numbers was due to the transformation of each site (usually  $10 \times 10$  m) to dozens of pixels with only a few objects. Detailed distribution of the training dataset in different classes are provided in Table S3 and Figure S3.

### 2.5.3. Classification Algorithms

For both pixel- and object-based classifications, two promising machine learning algorithms, SVM and RF, were applied to obtain the mapping results. SVM is a non-parametric classification algorithm with no assumption on the data distribution and the objective of SVM is to create hyperplanes to separate the dataset into classes [27]. SVM implemented in *e1071* package in R 3.6.3 [68] was adopted in this study with radial basis function kernel. The package provides a *tune.svm* tool to find the best gamma and costs parameters through cross-validation. A 10-fold cross-validation was applied to find the optimal set of *gamma* from 0.015625 ( $2^{-6}$ ) to 4 ( $2^2$ ) and *cost* from 0.5 ( $2^{-1}$ ) to 128 ( $2^7$ ). Based on the results, *gamma* = 0.5 and *cost* = 8, as well as *gamma* = 0.03125 and *cost* = 16, were chosen to compute the pixel- and object-based SVM models, respectively.

RF is an ensemble classification technique, which combines hundreds of decision trees and decides the final output class by the majority vote [28]. Randomization is involved in RF models including the construction of each decision tree with part of the training samples and a random subset of predictor variables to determine the tree split conditions [69]. RF implemented in *randomForest* package in R [68] was adopted. The required parameters included the number of classification trees (*ntree*) and the number of predictor variables used in each split (*mtry*). The default *ntree* of 500 trees was applied. For *mtry*, a 10-fold cross-validation was used to select the optimal value from 1 to 15. The selected *mtry* values were 8 and 11 for the pixel- and object-based RF models, respectively.

### 2.6. Stage 2: Rectification of Misclassified Pixels

The initial classification results obtained from the above procedures were then modified using modification rules (Table 6). Similar to the practice in [41–43], this set of rules



aimed to refine misclassified pixels according to the contextual and spatial relationships with other classes and ancillary data layers. Four types of knowledge-based rules, including spectral, topographic, relational (e.g., distance from the sea) and ancillary data rules (e.g., overlap with vector) [5], were developed from known ecological characteristics of the habitats. To identify the expected areas of divergence, the rectification of some classes could comprise multiple rules spanning several categories.

**Table 6.** Modification rules used in stage 2 to rectify misclassified pixels.

Rule	From Class	Type	Criteria	To Class	Objective
1	Natural rocky shoreline OR Soft shore	Topographic, relational	Distance from coastline > 50 m OR Terrain height > 5 m	Bare rock/ soil OR Other urban area	Merge rocky/ soft shore regions located in highland to adjacent bare or urban area
2	Mangrove	Topographic, relational	Distance from coastline > 2000 m OR Terrain height > 5 m	Woodland OR Shrubland	Merge mangrove regions located in highland to adjacent woodland or shrubland
3	Marsh/ reed bed	Topographic	Terrain height > 5 m	Grassland	Rectify marsh/ reed bed regions located in highland to grassland
4	Water	Spectral, ancillary data	GNDVI > 0.3 OR Intersect with building shadow layer	Shadow	Rectify water pixels to shadow based on spectral index
5	Other urban area OR Shadow	Ancillary data	Not located inside coastline layer	Water	Rectify pixels outside land area to water
6	Shadow	Relational	All	Class of the nearest neighbour	Rectify shadow pixels (including those generated in Rule 4) to nearby classes
7	All	Relational	Area < 100 m <sup>2</sup> (25 pixels)	Class of the nearest neighbour	Eliminate regions with areas smaller than minimum mapping unit (MMU)

Specifically, Rules 1–3 removed coastal habitats located in the highland area according to coastline layer and terrain height. Thresholds were set based on field survey observations that all these habitats appeared in the lowland area with terrain height below 5 m (Table S1). The threshold distances from the coastline were also determined by analysing the survey site locations. Rules 4–5 aimed to reduce confusion between shadow and water pixels using the building shadow layer, coastline layer and a spectral threshold set by trial and error, followed by Rule 6 to replace shadows with neighbouring classes. Rule 7 was related to the MMU and aimed to remove the salt-and-pepper effects by merging isolated pixels with neighbouring classes. Figure 4 illustrates a sub-area of the classification map before and after applying this set of rules. It should also be noted that while the explicit rules and thresholds were customized for this context, local knowledge and data would be necessary to develop appropriate rules when applied in other study areas.

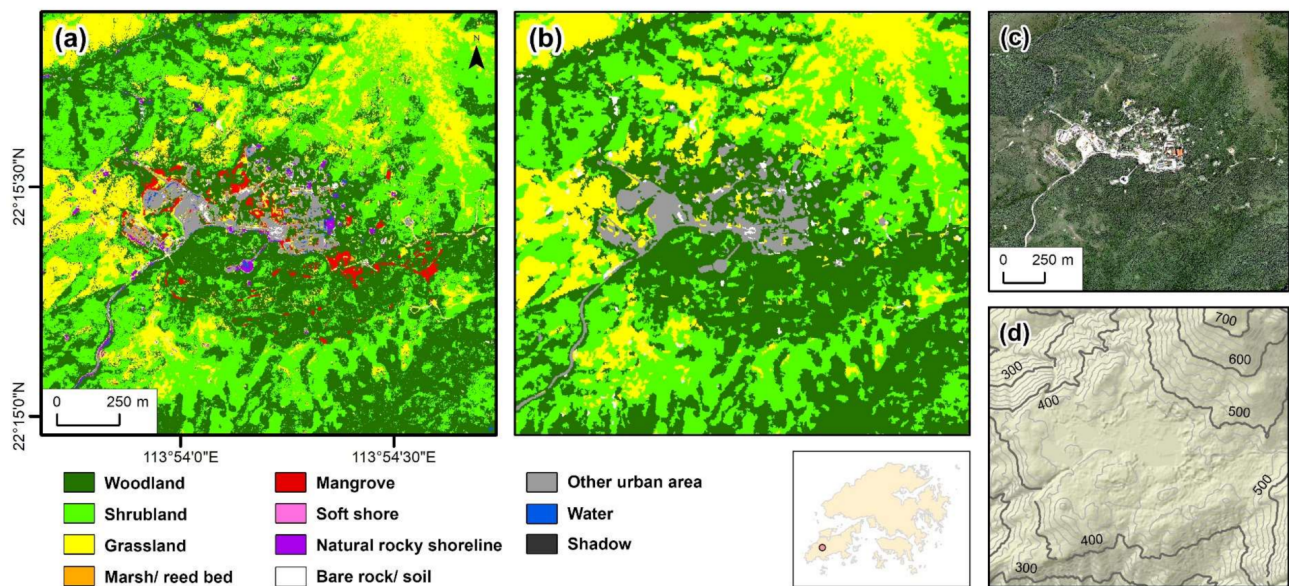
### 2.7. Stage 3: Production of Habitat Map

While the above process generated classification maps with only 10 classes, post-classification methods here aimed to expand the scheme to all 21 defined habitats.

#### 2.7.1. Generation of Mixed Habitat Classes

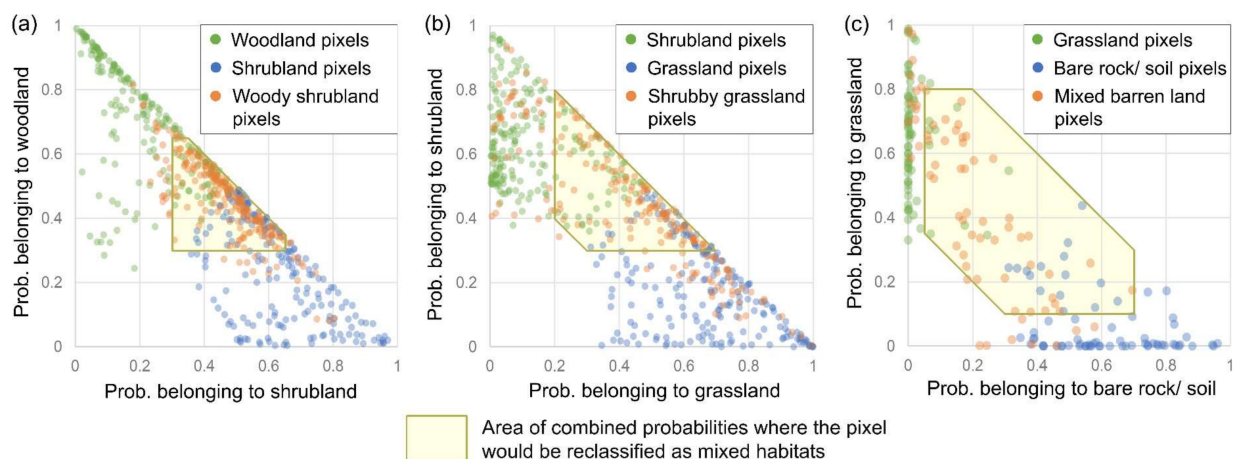
The classification scheme used in this study included three mixed habitats, namely woody shrubland, shrubby grassland and mixed barren land. A method to generate these classes was developed based on a fuzzy set combining the probabilities belonging to different classes for the target pixels [70]. Since the RF model provided probabilities of class membership for each pixel indicated by the frequency of decision tree votes [16,71] and the

mixed habitats were likely the transitional zone between two corresponding habitats, it was expected that they could be identified using a soft classification approach [72].



**Figure 4.** A selected area of classification map before and after applying the modification rules in stage 2; (a) initial classification result in stage 1 produced using pixel-based RF method; (b) refined classification map in stage 2; (c) true-colour composite of the WorldView image; (d) topographic map of the same area.

Based on this presumption, the underlying RF probability values were analysed using the training pixels which were identified as mixed habitats obtained in the field surveys. The mixed habitat pixels (orange points in Figure 5) were compared to random selections of similar numbers of pixels which were classified as each of the two corresponding habitats (green and blue points in Figure 5). The randomly selected habitat pixels served as pseudo-absence background points [73], which were used to identify the optimal thresholds of probability combinations. The threshold values were selected by analysing the resulting accuracies through trial-and-error methods (yellow bounding box in Figure 5) and transformed to Rules 1–3 in Table 7. The class of a pixel was modified if the probabilities of class membership satisfied the defined criteria.



**Figure 5.** Illustration of development of the mixed habitat classes based on probabilities belonging to

the pure classes as revealed by the random forest (RF) classification model; (a) classification of woody shrubland based on woodland and shrubland probabilities; (b) classification of shrubby grassland based on shrubland and grassland probabilities; (c) classification of mixed barren land based on grassland and bare rock/ soil probabilities. The thresholds of probability combination were selected by analysing the resulting accuracies through trial-and-error, i.e., maximizing the inclusion of mixed habitat pixels (orange points) and minimizing the inclusion of corresponding habitats (green and blue points).

**Table 7.** Modification rules used in stage 3 to expand from 10 to 21 habitat classes.

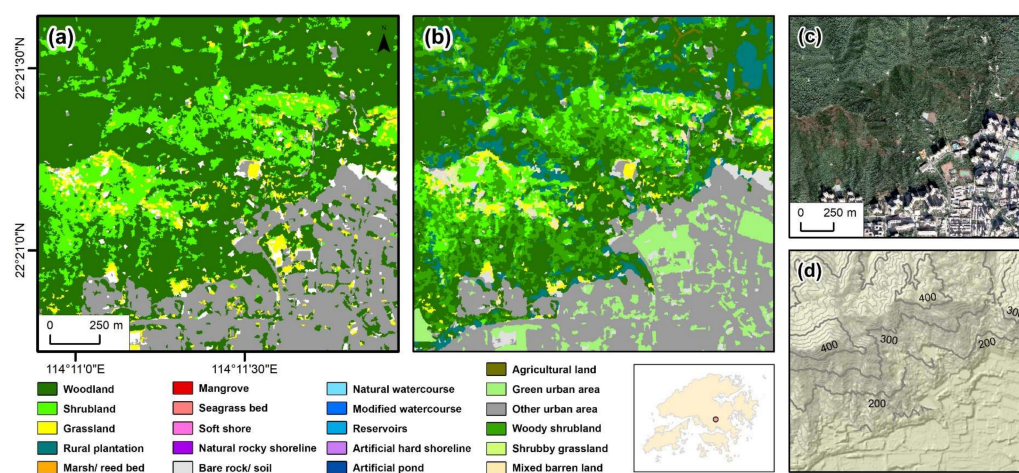
Rule	From Class	Criteria	To Class	Objective
1	Woodland OR Shrubland	$0.3 \leq P(\text{Woodland}) \leq 0.65$ AND $0.3 \leq P(\text{Shrubland}) \leq 0.65$	Woody shrubland	
2	Shrubland OR Grassland	$0.3 \leq P(\text{Shrubland}) \leq 0.8$ AND $0.2 \leq P(\text{Grassland}) \leq 0.7$ AND $P(\text{Shrubland}) + P(\text{Grassland}) \geq 0.6$	Shrubby grassland	Create mixed habitats by combining class membership probabilities
3	Grassland OR Bare rock/ soil	$0.1 \leq P(\text{Grassland}) \leq 0.8$ AND $0.05 \leq P(\text{Bare rock/ soil}) \leq 0.7$ AND $P(\text{Grassland}) + P(\text{Bare rock/ soil}) \geq 0.4$	Mixed barren land	
4	Woodland	Random forest classification based on field survey data	Rural plantation	Discriminate rural plantation from woodland
5	Woodland OR Woody shrubland	Intersect with tree planting record layer	Rural plantation	Create habitats based on ancillary layers
6	Vegetation-related	Intersect with urban park layer	Green urban area	
7	Vegetation-related	Surrounded by other urban area	Green urban area	Create habitats based on relational rules
8	Vegetation-related	Intersect with cultivated land layer	Agricultural land	
9	All	Intersect with seagrass layer	Seagrass bed	
10	Other urban area OR Natural rocky shoreline	Intersect with artificial hard shoreline layer	Artificial hard shoreline	Create habitats based on ancillary layers
11	Water	Intersect with pond layer	Artificial ponds	
12	Water	Intersect with reservoir layer	Reservoirs	
13	Water	Surrounded by other urban area	Modified watercourse	Create habitats based on relational rules
14	Water	Not satisfying Rule 11–13	Natural watercourse	Modify remaining water pixels
15	Water	Located outside the coastline layer	No data	Remove sea area

### 2.7.2. Expansion of Habitat Classification

Besides the mixed habitats described above, the second set of modification rules aimed to create other new classes based on ancillary layers and expand the classification scheme beyond the remote-sensing output (Table 7). For instance, for rural plantation habitat, while existing GIS data covered some potential sites in the territory (Rule 5), additional areas were identified using the satellite image with the aid of field-collected training data. Observations in the field showed that rural plantation often differed from woodland class in terms of species composition, hence a binary classification was chosen to separate the rural plantation from the woodland class, using the same classification method and variables as in stage 1 (Rule 4). A similar idea was also applied to green urban area habitat, where urban



parcs were mapped by GIS data (Rule 6) and street trees were identified by intersecting vegetation-related class with urban areas (Rule 7). For other habitats such as agricultural land, seagrass bed and artificial hard shoreline, corresponding ancillary layers were used to extract these habitats from related classes (Rules 8–10). Water class in the classified map was also separated into various habitats according to available GIS data and spatial relationships (Rules 11–15). Figure 6 shows a sub-area of the map before and after applying this set of modification rules.



**Figure 6.** A selected area of habitat map showing the results before and after applying the modification rules in stage 3; (a) classification map with 10 classes in stage 2 produced using pixel-based RF method; (b) habitat map with 21 classes in stage 3; (c) true-colour composite of the WorldView image; (d) topographic map of the same area.

### 2.7.3. Accuracy Assessment

The mapping accuracies in different stages were evaluated using two sets of assessment points. For the four classified maps in stage 1 and 2, stratified random sampling was used to generate 770 points (Figure S4). Each point was then assigned to one of the 10 classes based on visual interpretation of the satellite image as well as supplementary information such as aerial photographs and GIS maps, by an analyst with local knowledge and an ecological background. To investigate the effectiveness of the post-classification process, the randomly sampled points were applied to both classification results before (stage 1) and after (stage 2) applying the first set of modification rules. For the habitat map produced in stage 3, the accuracy was assessed using the validation set of field survey points consisting of most of the habitats.

Several commonly used statistics were computed to report the performances, including producer's accuracies (PA) and user's accuracies (UA) for each class and overall accuracies (OA) of the map [74]. The Kappa coefficient was also calculated as a widely adopted indicator of classification performance eliminating bias from chance agreement [75]. Considering the sampling variability caused by the small and uneven distribution of assessment points for some habitats, a confusion matrix of estimated area proportions was constructed to statistically calculate the standard error associated with each class, and then quantify the 95% confidence intervals of PA, UA, OA and area estimates using the error-adjusted area estimator, according to the methods suggested by [76,77].

## 3. Results

### 3.1. Classification Maps and Accuracies

Classification results obtained from the four classification methods achieved OA from 76.0% to 84.0% after the first set of post-classification modification rules were applied, with corresponding Kappa statistics ranging from 0.73 to 0.82 (Table 8). The highest OA was

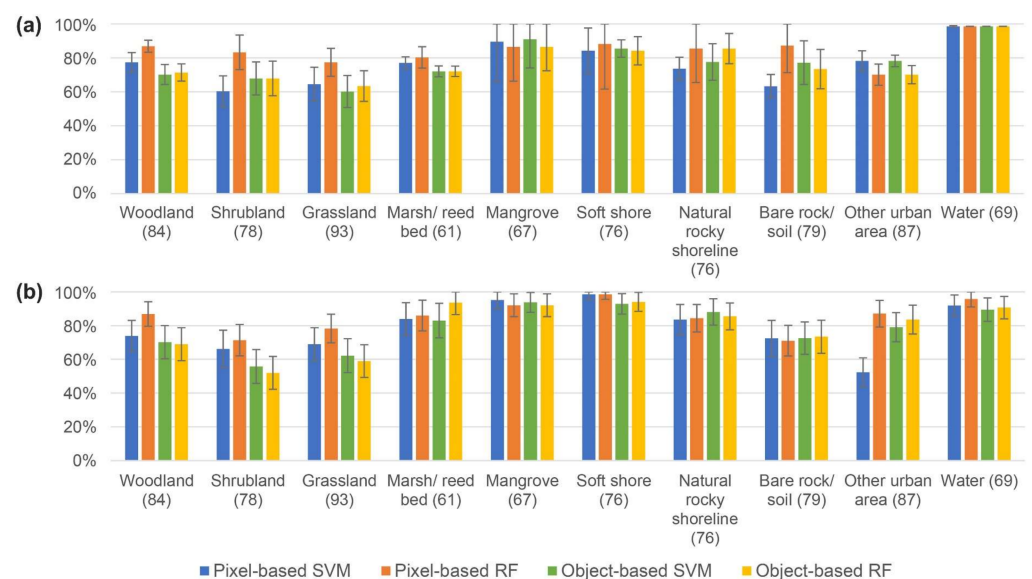


obtained from pixel-based RF classification (84.0% and 0.82 Kappa). While object-based SVM and RF classification methods obtained similar OA (76.6% and 77.1% respectively), pixel-based SVM classification performed the worst with 76.0% accuracy. Before the application of modification rules, the accuracies were between 67.7% and 73.1%, which suggested that this set of modification rules were able to rectify the misclassified pixels and resulted in an average increase of 8.8% in OA.

**Table 8.** Overall accuracies (OA) and Kappa statistics of the four classification methods, including pixel-based support vector machine (SVM), pixel-based random forest (RF), object-based SVM and object-based RF. The values inside the parentheses indicate the 95% confidence intervals of the OA.

Classification Accuracy	Pixel-based SVM	Pixel-based RF	Object-based SVM	Object-based RF
OA	76.0% ( $\pm 3.9\%$ )	84.0% ( $\pm 3.1\%$ )	77.1% ( $\pm 4.2\%$ )	76.6% ( $\pm 4.1\%$ )
Kappa	0.73	0.82	0.75	0.74
OA (before rules)	67.7% ( $\pm 3.6\%$ )	73.1% ( $\pm 3.5\%$ )	69.0% ( $\pm 4.8\%$ )	68.6% ( $\pm 4.6\%$ )

When accuracies of individual classes were considered (Figure 7), woodland, shrubland and grassland tended to perform poorer compared to other classes and sometimes had accuracies lower than 70%. In particular, the lower accuracies for shrubland (55.8–83.3%) might indicate the difficulty of accurately identifying these areas covered with intermediate levels of vegetation. For these classes, pixel-based RF classification was the only method that could obtain all PA and UA higher than 70%, and lower accuracies were generally obtained from object-based (52.0–71.4%) compared to pixel-based approach (60.3–86.9%) regardless of the algorithms. Classification of coastal habitats including marsh/reed bed, mangrove, soft shore and natural rocky shoreline was encouraging, with the majority of accuracies higher than 80% in all methods. Satisfactory levels of accuracies, mainly 70–80%, were obtained for bare rock/soil and other urban area classes, except in pixel-based SVM classification where confusion between these two classes was a major source of errors. As expected, the water class provided high PA (98.6%) and UA (89.5–95.8%) with narrow confidence intervals, especially when shadows had been eliminated using the modification rules. The classification maps and confusion matrices are provided in Figure S5 and Tables S4–S7.



**Figure 7.** Comparison of (a) producer's accuracies and (b) user's accuracies of individual class using the four classification methods obtained in stage 2. The error bars indicate the 95% confidence intervals of the accuracy measures. The numbers inside the parentheses indicate the numbers of assessment points.

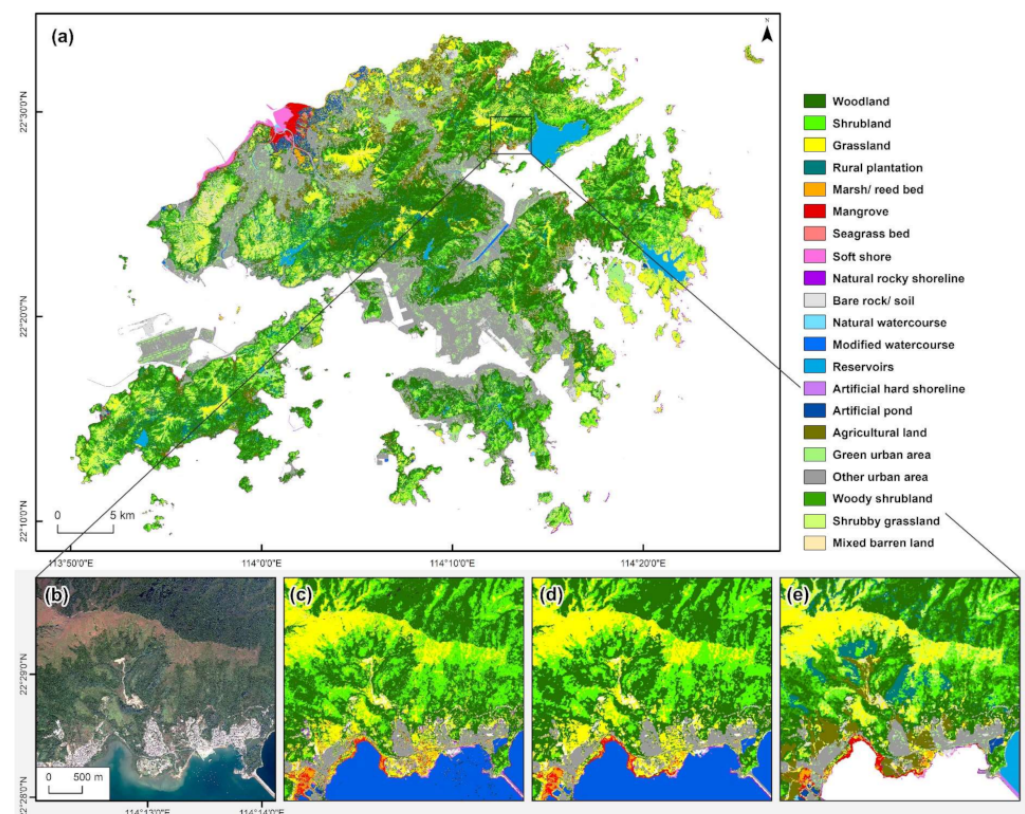
### 3.2. Habitat Map and Accuracies

Since the highest OA was obtained from pixel-based RF classification with superior PA and UA compared to other methods for most classes, pixel-based RF classification result was selected as the primary dataset for further processing. The habitat map produced in this study, consisting of all 21 habitat categories, is displayed in Figure 8, with a sub-area shown to visualize the mapping procedures and habitat distributions.

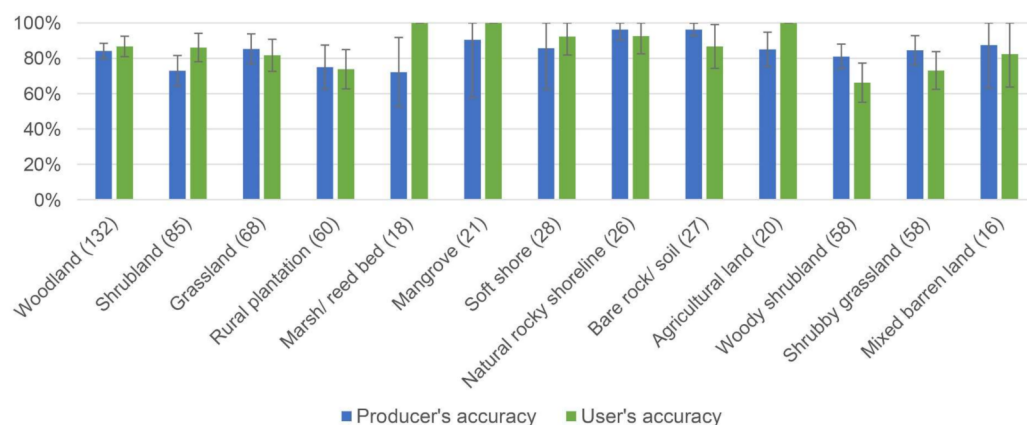
The habitat map produced after the three-stage classification procedure was assessed using field survey points. Considering the imbalanced numbers of survey points which could bias the computation of OA to dominating classes, individual PA and UA of habitats with more than 15 points were investigated (Figure 9). The complete confusion matrix is given in Table S8.

For woodland, the PA was 84.1% and the UA was 86.7%. Shrubland had similar UA (86.1%) but lower PA (72.9%), which was mainly caused by confusion with woody shrubland and shrubby grassland; 85.3% PA and 81.7% UA were produced for grassland. While the rural plantation class had slightly lower PA (75.0%) and UA (73.8%) compared to other vegetation habitats, the accuracy was satisfactory in view of the complicated nature of the plantation, which is illustrated in Section 4.5.

Marsh/ reed bed, mangrove and agricultural land had 100% UA and slightly lower PA (72.2–90.5%), and misclassification mainly occurred with woodland and grassland. Although 100% accuracies were reported, the values might not represent the actual performance considering the limited number of reference points and the fact that survey points were often collected in areas with uniform coverage. Further visual interpretation showed that commission errors of these habitats happened especially in isolated patches or edges.



**Figure 8.** (a) The habitat map produced in this study consisting of 21 categories; (b–e) selected area for visualizing the results at different stages of the mapping procedures; (b) true-colour composite of the WorldView image; (c) initial classification with 11 classes in stage 1 produced using pixel-based RF method; (d) refined classification map with 10 classes in stage 2; (e) habitat map with 21 classes in stage 3.



**Figure 9.** Producer's and user's accuracies of the habitat classes in stage 3 evaluated using field survey points. The error bars indicate the 95% confidence intervals of the accuracy measures. The numbers inside the parentheses indicate the numbers of assessment points.

Soft shore and natural rocky shoreline were found to have relatively high PA (85.7–96.2%) and UA (92.3–92.6%), despite the occasional confusion that appeared between these two habitats. Finally, the three mixed habitats had 81.0–87.5% PA and 66.2–82.4% UA. Most of the misclassifications occurred with the adjacent vegetation types. The resulting accuracies were slightly lower than those obtained for other vegetation classes, suggesting higher difficulty to identify these mixed habitats from the satellite image.

Considering the uncertainty caused by sampling variability, the 95% confidence intervals for larger habitats including woodland, shrubland and grassland were within 4.4–9.1% of estimated accuracies. The confidence intervals of PA of marsh/ reed bed, mangrove and soft shore were relatively wide, which could be attributable to the smaller numbers of assessment points, limited coverages of these habitats in the territory and the confusion found with other major habitats. Larger uncertainties were also found for rural plantations and the three mixed habitats (95% confidence intervals =  $\pm 7.0$ –24.4%). Besides the confusion with other vegetation classes, this could be further influenced by the number of assessment points, especially for mixed barren land. Nevertheless, the lower bounds of confidence intervals of these habitats were all higher than 60%, proving the mapping effectiveness when the sampling variability was taken into consideration.

Overall, this study was able to achieve higher than 80% accuracy for most of the evaluated classes and higher than 70% accuracy for the more-challenging habitats. Figure 9 evaluates 13 habitats out of the total 21 mapped classes. For the remaining habitat categories, most of them covered only less than 1% of the area of the territory according to the map produced (Table S9), posing obstacles for the deployment of sufficient field survey points and quantitative analysis of the results. In spite of this limitation, we are confident in the mapping accuracies of these habitats, since they were produced from reliable databases and a set of logical rules. For instance, the water-related habitats were derived from the water class in stage 2, which was found to be accurate in previous evaluation. Further visual interpretation supported the assertion that the coverages of these small habitats were determined largely by the ancillary layers and modification rules compared to remote-sensing classification.

#### 4. Discussion

A systematic three-stage classification framework was developed in this study, which combined remotely sensed VHR satellite images, GIS layers in existing databases and two sets of post-classification modification rules. This study has achieved the following: (i) it translated the habitat mapping process into a straightforward and controllable workflow with enhanced accuracies; (ii) it revealed varying performances on specific habitats using different classification methods which could be context-dependent; (iii) it exploited many-

to-many relationships between habitat categories through geographical data and contextual knowledge in the post-classification phase; (iv) it presented a method to identify mixed habitats by combining the soft probability outputs from the RF classification model; and (v) it demonstrated the benefits of integrating remote-sensing classification and GIS data to extract particular habitats.

#### 4.1. Three-Stage Mapping Procedure

One of the aims of this framework was to compensate for the limitations brought by direct attribution of spectral signatures to all habitat classes [8]. This is especially challenging when landscapes become more heterogeneous and the numbers of classes increase [6,78], as experienced in this study. The multi-stage classification implemented in [79] produced an average OA of 93.2% compared to the direct use of SVM (84.4%) and RF (83.8%) in classifying coastal wetlands, while the methods in [37] increased the accuracy of land use classification from 85% to 94%. Similarly, this study attempted to map 21 habitats and was able to achieve higher than 80% accuracy for most of the evaluated classes and higher than 70% accuracy for the more-challenging habitats.

Instead of a single uniform classification method, the proposed framework separated mapping procedures into different stages with input from ecological specialists. By translating the ecological properties into simple GIS rule-models, this semi-automated approach allows adaptive control in different nodes of the mapping process [3,7]. Starting with some observable and basic habitat classes in the initial classification, the procedure gradually expanded the classes to other habitats which were further distinguished using probability combinations, overlay operations, morphological characteristics and spatial relationships. Compared to fully automated classifiers, this knowledge-based method incorporated expert rules from ecologists and could enhance the understanding of habitats in the map products [19].

#### 4.2. Selection of Algorithms during Classification Process

A key consideration in stage 1 of this framework was the selection of the optimal classification method that was suitable for the study context. In this study, two machine learning algorithms, SVM and RF, were tested with both pixel- and object-based approaches to produce different initial results. While investigation on the variable importance (Figure S6) illustrated that the adopted spectral, textural, topographic and geometric variables could contribute to the identification of different habitats, the contrasts in variable importance scores among methods and the classification accuracies presented above also indicated the intrinsic difference of the implemented algorithms.

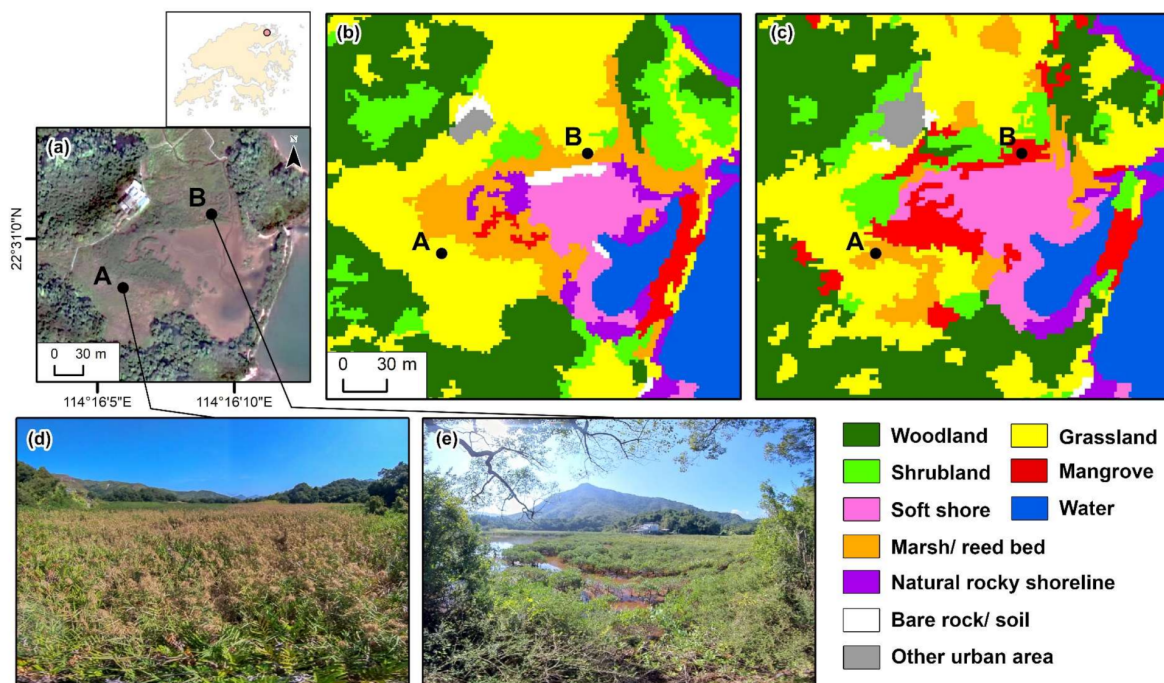
The accuracy assessment results showed that pixel-based RF classification performed the best in this study context. Object-based classification was generally found to yield higher accuracies compared to the pixel-based counterpart in existing literature [31,80], especially when applied to VHR images. However, the relatively low accuracy found in this study was possibly due to the lower number of training objects, uncertainty emerged during the additional segmentation step [81], the ability of pixel-based texture variables to capture contextual relationships similar to object-based statistics [69], and the flexibility to compute textures and represent habitats with different vertical and horizontal heterogeneity [82] as evidenced by the importance of texture variables obtained with multiple window sizes in pixel-based models (Figure S6).

Comparing the classification algorithms, SVM produced slightly higher accuracy than RF when an object-based approach was adopted, which was possibly due to the lower number of training observations in object-based models and the higher generalization capacity of SVM [27]. Despite the promising results obtained from pixel-based RF classification for most classes, slightly higher accuracies for mangrove and marsh/reed bed classes were obtained from object-based SVM classification. Further visual interpretation supported the assertion that object-based SVM classification was superior in distinguishing these two habitats, especially in coastal areas, probably due to the irregular shapes and



small habitat patch sizes (Figure 10). Textures in pixel-based models were generated in rectangular windows while spectral statistics in object-based models were computed from irregularly segmented objects in each band [69]. This might affect the classification performance of specific habitats which had irregular coverage and possessed spectral variations in specific bands.

The investigation above indicated that the optimal algorithm can be case-dependent and requires specific evaluation with respect to the project [25,32]. As demonstrated in this study, selecting classification variables from different aspects and comparing multiple classification methods could be an effective approach. This was also vital in the habitat mapping framework since the best classification map would be used as the primary dataset for further processing into the finalized habitat map. Besides the overall superiority, this study also found subtle mapping differences for particular habitats. Further combination of these classification results in the post-classification stage, such as overlay operations [83], could be a possibility to exploit the benefits of both maps.



**Figure 10.** A selected wetland site including both marsh/ reed bed (Point A) and mangrove (Point B) habitats; (a) true-colour composite of the WorldView image; (b) classification result using pixel-based random forest (RF) method; (c) classification result using object-based support vector machine (SVM) method; (d) photo of point A taken in the field, which is covered mainly by *Phragmites australis*; (e) photo of point B taken in the field, which is covered mainly by *Kandelia obovata*.

#### 4.3. Use of Information Layers and Modification Rules to Enhance Mapping Accuracies and Expand the Classification Scheme

In the post-classification phase, expert knowledge and ancillary reference data were first utilized to develop modification rules and refine the initial classification outcomes (stage 2). Similarly, Manandhar et al. [39] developed knowledge-based rules by incorporating data such as land use, DEM, spatial texture and NDVI value, while Rapinel et al. [43] attempted to reclassify misclassified objects according to context, shape and texture criteria. In this study, spectral (band and index), landscape (texture and shape), environmental features (terrain) and vector layers were also jointly used in designing the mapping procedures to address the characteristics of various habitat features [40,84]. Such a decision support system underpinned the suggestion of [42] to let geographical data have a stronger voice rather than relying on the classification parameters.

The implementation of modification rules in this study successfully improved the OA from 67.7–73.1% to 76.0–84.0%. This was in line with other land cover mapping studies applying post-classification corrections with ancillary data, which resulted in improvements in OA from 72–79% to 87–91% [39] and from 67.8–71.9% to 82.8–87.4% [40] respectively. Since the thematic layers were gathered from reliable sources and independent from the satellite images, they are able to provide robust contextual information and be applied to different parts of the study area under various observation conditions.

Apart from refining misclassified areas, another major function of post-classification rules was to expand the classification scheme to all habitat classes (stage 3). The initial classification map was combined with contextual knowledge to obtain the desired additional habitat information. It was suggested that classes defined in different mapping domains can be translated through one-to-many and many-to-many relationships [45]. For instance, the water class in stage 2 was separated into several habitats in stage 3, while agricultural land and green urban area were derived from various vegetation-related classes. The overall process increased the number of classes from 10 in stage 2 to 21 in stage 3. Harris and Ventura [85] described this procedure as improving the specificity of mapping results, such as the increase of the number of urban classes from 5 to 13 by adding population and zoning information in their study.

#### 4.4. Soft Classification Method to Identify Mixed Habitats

Mixed habitat classes were considered as many-to-many relationships that could be resolved through modification rules in this study. Although mixed classes were also defined in other mapping products, they were arguably challenging to accurately map, due to the variations in thresholds and sub-pixel heterogeneity [86]. Furthermore, the mixed classes defined in this study were mainly ecotones representing transitional zones from grass to forest, which are important landscape structures with distinctive ecological functions and wildlife importance [47,70,87]. This study presented a novel method to identify the locations of mixed habitats by combining the soft probability outputs from the RF classification model. Instead of a traditional hard classifier, fuzzy logic was adopted to soften the decision boundaries by allowing mixed observations to have memberships in corresponding classes. This was followed by a defuzzification process to assign the mixed classes according to the classification scheme and optimized probability thresholds [14,70].

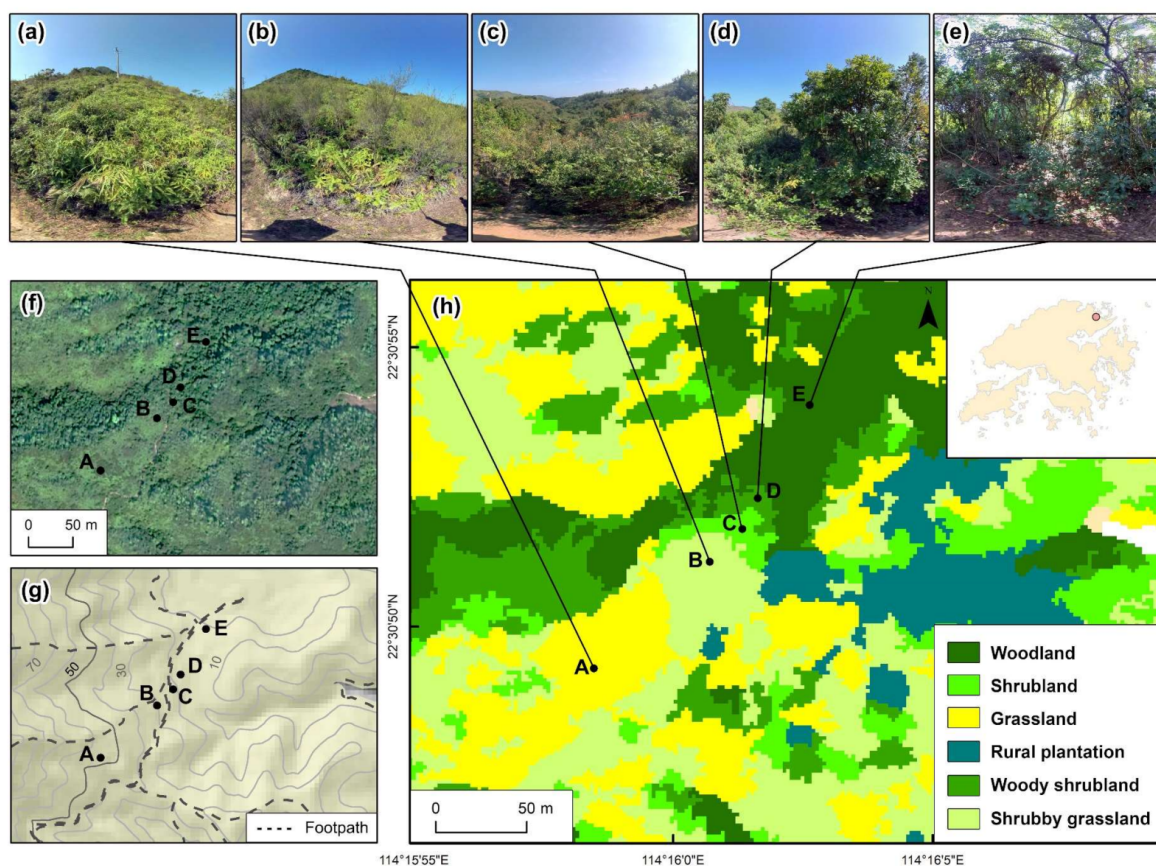
By utilizing probability outputs from a modern machine learning algorithm, the methodology adopted here extended a previous study in Hong Kong, which placed fuzzy boundaries between shrubs and grass using simple spectral thresholds [20], and another study that considered the co-occurrence property in ratio maps interpolated from tree species data [88]. The proposed defuzzification step was also more suitable for thematic map production compared to similar use of RF probabilities to estimate sub-pixel fractional abundance in [72]. Figure 11 illustrates a sub-area showing rapid changes of vegetation structures from grassland through shrubby grassland, shrubland and woody shrubland to woodland in a 500 m distance along a hiking trail. The produced habitat map successfully revealed the transitional patterns and mapped the mixed habitats as spatial intergrades between classes. As evidenced by the low accuracies in the direct classification approach, these mixed habitats could be easily confused with other vegetation in a single classification model, thus this soft classification method was believed to be a better method to discern the spatial dynamics along ecological gradients.

#### 4.5. Hybrid Approach to Identify Rural Plantation Habitats

Remote-sensing data offer direct land cover observation but it can be hard to determine land use information without the support of external data sources. For example, rural plantation habitat focuses on the formation process of vegetation areas by humans. Although the planting of pioneer species such as *Acacia confusa* and *Lophostemon confertus* has taken place throughout the territory [49], natural succession has added more native trees and made many plantations indistinguishable [20]. Recent plantation programmes

and enrichment schemes also tended to adopt a higher portion of native seedlings and increase species diversity. To facilitate correct classification of the rural plantation habitat, two sources of information were combined, namely remote-sensing classification and GIS data (planting records) provided by the government department.

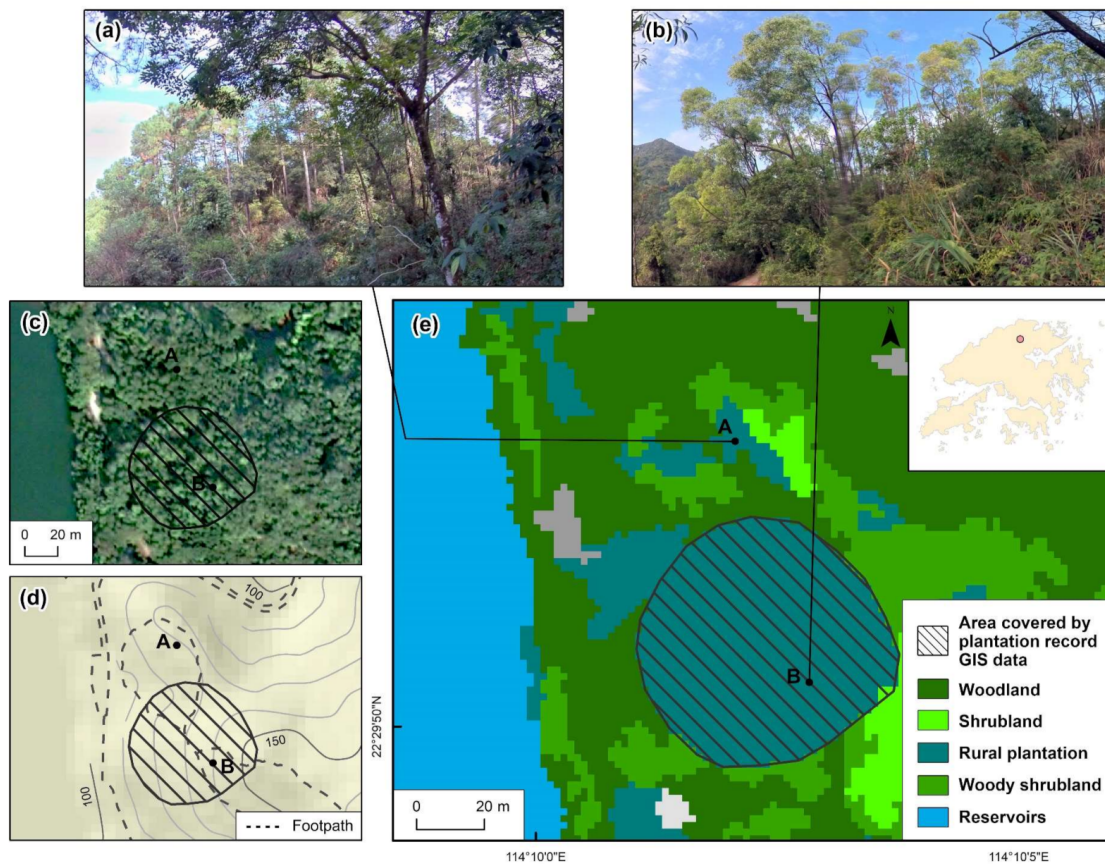
This hybrid approach led to 75.0% PA and 73.8% UA for rural plantation habitat when validated using field-collected points. Among the correctly identified points, two-thirds were obtained from the classification and the remaining one-third were contributed from the use of GIS data. Figure 12 illustrates two selected rural plantation areas that were successfully mapped using the two sources of information. The planting record data were found to focus mainly on large plantation areas in recent years especially those inside the designated country parks. As the data were managed by the government, these contained all varieties of plantations including mixtures of native trees, and were believed to be accurate. In contrast, the classification model relied on field-collected training data and image characteristics to identify potential plantation areas. Further investigation showed that the model was able to extract some areas covered by a few particular tree species such as *Acacia confusa*, *Pinus massoniana*, *Melaleuca cajuputi* and *Eucalyptus* spp. in all regions in Hong Kong, probably due to their dominance in the canopy layer and the distribution of training data.



**Figure 11.** A selected site showing the transition from grassland (Point A) to woodland (Point E) within a short distance; (a) photo of point A taken in the field. Species included *Dicranopteris pedata*, *Blechnum orientale*; (b) photo of point B taken in the field. Species included *Baeckea frutescens*, *Dicranopteris pedata*; (c) photo of point C taken in the field. Species included *Rhaphiolepis indica*, *Baeckea frutescens*; (d) photo of point D taken in the field. Species included *Schefflera heptaphylla*, *Rhodomyrtus tomentosa*; (e) photo of point E taken in the field. Species included *Cinnamomum camphora*, *Aporosa dioica*; (f) true-colour composite of the WorldView image; (g) topographic map; (h) habitat map produced in this study.



It is worth noting that besides describing the data products, this study also aimed to present the procedures to reliably produce these maps. It has been a common practice to intersect multiple datasets to construct a desired habitat map [89]. As explained in [41], the actual definition of rules in one study was created from expert knowledge by observing systematic error patterns in the maps and classification of other areas would require a different set of rules. However, this rule-based approach has advantages in terms of its transparency, efficiency and relative simplicity, which facilitate understanding by people without a background in remote sensing [41]. Since the development of appropriate rules also focused on identifying the ecological characteristics of target habitats [5,85], the general process can be transferable to other applications where local knowledge and data are available.



**Figure 12.** Selected sites of rural plantation identified by RF classification model (Point A) and plantation record GIS data (Point B) respectively; (a) photo of point A taken in the field, which is a plantation of *Pinus massoniana*; (b) photo of point B taken in the field, which is a plantation of *Acacia auriculiformis*; (c) true-colour composite of the WorldView image; (d) topographic map; (e) habitat map produced in this study.

## 5. Conclusions

The three-stage mapping procedures demonstrated in this study effectively utilized the characteristics of various sources of information in a systematic workflow to produce promising results, with the additional advantages of being straightforward, controllable and easy to understand. As illustrated with examples of sub-area mapping results, transitional patterns of mixed habitats were successfully mapped using a soft classification method and the identification of some habitats benefited from the hybrid use of remote-sensing classification and GIS data. Nevertheless, the proposed method was also limited by the availability and quality of useful ancillary data, as well as the efforts required to customize modification rules for specific contexts. Adaptive implementation of classifi-



cation procedures, development of appropriate rules and availability of reliable GIS data were all believed to be vital for the production of a high-quality habitat map. In addition to the optical satellite image, future studies can explore the combined use of other novel data sources to enhance habitat information in multiple dimensions, such as three-dimensional structures from LiDAR and unmanned aerial vehicles [22,90]. Overall, through collaboration with local ecologists and policymakers, the three-stage approach of this study has increased mapping accuracy for diverse and transitional habitats. We also recommend the procedures developed are adopted in future map updates to facilitate long-term monitoring of habitat changes. The resulting habitat map can be an informative resource for supporting environmental planning and managing cities with heterogeneous landscapes.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs14010067/s1>, Figure S1: Distribution of survey points used for training data. Figure S2: Distribution of survey points used for validation data. Table S1: Summary of terrain heights of each habitat according to field survey points. Table S2: Examples of species with the highest occurrences in vegetation-related habitats during field survey. Table S3: Numbers of training pixels/objects used in pixel-based and object-based classification. Figure S3: Distribution of training data over the study area. Figure S4: Distribution of randomly sampled points used for accuracy assessment in stage 2. Figure S5: Refined classification maps with 10 classes in stage 2 obtained using the four classification methods. Table S4: Confusion matrix of pixel-based SVM classification result. Table S5: Confusion matrix of pixel-based RF classification result. Table S6: Confusion matrix of object-based SVM classification result. Table S7: Confusion matrix of object-based RF classification result. Table S8: Confusion matrix of habitat map against field survey points. Table S9: Land coverage of each habitat type mapped in this study. Figure S6: Top 10 classification variables with the highest importance using different evaluation metrics.

**Author Contributions:** Conceptualization, F.K.K.W., E.K.Y.L., R.H.L., T.P.T.N.; Methodology, I.H.Y.K., F.K.K.W.; Investigation, I.H.Y.K.; Data Curation, I.H.Y.K., F.K.K.W.; Writing—Original Draft, I.H.Y.K.; Writing—Review and Editing, F.K.K.W., T.F., E.K.Y.L., R.H.L., T.P.T.N.; Project administration, T.F.; Funding Acquisition, F.K.K.W., T.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Agriculture, Fisheries and Conservation Department of the HKSAR Government.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding authors.

**Acknowledgments:** We thank LM Chu (School of Life Sciences, CUHK), David Lau, Joyce Siu (Shiu-Ying Hu Herbarium, CUHK), Leo Ng, Martin Lee (Earth System Science Programme, CUHK) and all helpers for their assistance in field data collection. This work was supported by the Agriculture, Fisheries and Conservation Department and the Planning Department of the HKSAR Government. We thank the editors and anonymous reviewers for their helpful comments on this manuscript.

**Conflicts of Interest:** I.H.Y.K., F.K.K.W. and T.F. received funding support from the Agriculture, Fisheries and Conservation Department of the HKSAR Government to work on a consultancy project that is directly related to this work. E.K.Y.L., R.H.L. and T.P.T.N. are employees of the funding body.

## References

1. Bunce, R.G.H.; Bogers, M.M.B.; Evans, D.; Halada, L.; Jongman, R.H.G.; Mucher, C.A.; Bauch, B.; de Blust, G.; Parr, T.W.; Olsvig-Whittaker, L. The significance of habitats as indicators of biodiversity and their links to species. *Ecol. Indic.* **2013**, *33*, 19–25. [[CrossRef](#)]
2. Buchanan, G.M.; Brink, A.B.; Leidner, A.K.; Rose, R.; Wegmann, M. Advancing terrestrial conservation through remote sensing. *Ecol. Inform.* **2015**, *30*, 318–321. [[CrossRef](#)]
3. Weiers, S.; Bock, M.; Wissen, M.; Rossner, G. Mapping and indicator approaches for the assessment of habitats at different scales using remote sensing and GIS methods. *Landsc. Urban Plan* **2004**, *67*, 43–65. [[CrossRef](#)]

4. Pettorelli, N.; Laurance, W.F.; O'Brien, T.G.; Wegmann, M.; Nagendra, H.; Turner, W. Satellite remote sensing for applied ecologists: Opportunities and challenges. *J. Appl. Ecol.* **2014**, *51*, 839–848. [\[CrossRef\]](#)
5. Bell, G.; Neal, S.; Medcalf, K. Use of remote sensing to produce a habitat map of Norfolk. *Ecol. Inform.* **2015**, *30*, 293–299. [\[CrossRef\]](#)
6. Corbane, C.; Lang, S.; Pipkins, K.; Alleaume, S.; Deshayes, M.; García Millán, V.E.; Strasser, T.; Vanden Borre, J.; Toon, S.; Michael, F. Remote sensing for mapping natural habitats and their conservation status—New opportunities and challenges. *Int. J. Appl. Earth Obs.* **2015**, *37*, 7–16. [\[CrossRef\]](#)
7. Mao, D.; Wang, Z.; Du, B.; Li, L.; Tian, Y.; Jia, M.; Zeng, Y.; Song, K.; Jiang, M.; Wang, Y. National wetland mapping in China: A new product resulting from object-based and hierarchical classification of Landsat 8 OLI images. *ISPRS J. Photogramm.* **2020**, *164*, 11–25. [\[CrossRef\]](#)
8. Nagendra, H.; Lucas, R.; Honrado, J.P.; Jongman, R.H.G.; Tarantino, C.; Adamo, M.; Mairota, P. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecol. Indic.* **2013**, *33*, 45–59. [\[CrossRef\]](#)
9. Feilhauer, H.; Dahlke, C.; Doktor, D.; Lausch, A.; Schmidlein, S.; Schulz, G.; Stenzel, S. Mapping the local variability of Natura 2000 habitats with remote sensing. *Appl. Veg. Sci.* **2014**, *17*, 765–779. [\[CrossRef\]](#)
10. Múcher, C.A.; Roupioz, L.; Kramer, H.; Bogers, M.M.B.; Jongman, R.H.G.; Lucas, R.M.; Kosmidou, V.E.; Petrou, Z.; Manakos, I.; Padoa-Schioppa, E.; et al. Synergy of airborne LiDAR and Worldview-2 satellite imagery for land cover and habitat mapping: A BIO\_SOS-EODHaM case study for the Netherlands. *Int. J. Appl. Earth Obs.* **2015**, *37*, 48–55. [\[CrossRef\]](#)
11. Räsänen, A.; Virtanen, T. Data and resolution requirements in mapping vegetation in spatially heterogeneous landscapes. *Remote Sens. Environ.* **2019**, *230*, 111207. [\[CrossRef\]](#)
12. Ma, L.; Li, M.; Ma, X.; Cheng, L.; Du, P.; Liu, Y. A review of supervised object-based land-cover image classification. *ISPRS J. Photogramm.* **2017**, *130*, 277–293. [\[CrossRef\]](#)
13. McCarthy, M.J.; Radabaugh, K.R.; Moyer, R.P.; Muller-Karger, F.E. Enabling efficient, large-scale high-spatial resolution wetland mapping using satellites. *Remote Sens. Environ.* **2018**, *208*, 189–201. [\[CrossRef\]](#)
14. McDermid, G.J.; Franklin, S.E.; LeDrew, E.F. Remote sensing for large-area habitat mapping. *Prog. Phys. Geog.* **2005**, *29*, 449–474. [\[CrossRef\]](#)
15. Inglada, J.; Vincent, A.; Arias, M.; Tardy, B.; Morin, D.; Rodes, I. Operational high resolution land cover map production at the country scale using satellite image time series. *Remote Sens.* **2017**, *9*, 95. [\[CrossRef\]](#)
16. Immitzer, M.; Atzberger, C.; Koukal, T. Tree species classification with random forest using very high spatial resolution 8-Band WorldView-2 satellite data. *Remote Sens.* **2012**, *4*, 2661–2693. [\[CrossRef\]](#)
17. Lottering, R.; Mutanga, O.; Peerbhay, K.; Ismail, R. Detecting and mapping *Gonipterus scutellatus* induced vegetation defoliation using WorldView-2 pan-sharpened image texture combinations and an artificial neural network. *J. Appl. Remote Sens.* **2019**, *13*, 014513. [\[CrossRef\]](#)
18. Sibanda, M.; Mutanga, O.; Rouget, M. Testing the capabilities of the new WorldView-3 space-borne sensor's red-edge spectral band in discriminating and mapping complex grassland management treatments. *Int. J. Remote Sens.* **2017**, *38*, 1–22. [\[CrossRef\]](#)
19. Adamo, M.; Tomaselli, V.; Tarantino, C.; Vicario, S.; Veronico, G.; Lucas, R.; Blonda, P. Knowledge-based classification of grassland ecosystem based on multi-temporal WorldView-2 data and FAO-LCCS taxonomy. *Remote Sens.* **2020**, *12*, 1447. [\[CrossRef\]](#)
20. Nichol, J.; Wong, M.S. Habitat mapping in rugged terrain Using multispectral Ikonos Images. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 1325–1334. [\[CrossRef\]](#)
21. Abbas, S.; Nichol, J.E.; Fischer, G.A. A 70-year perspective on tropical forest regeneration. *Sci. Total Environ.* **2016**, *544*, 544–552. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Li, Q.; Wong, F.K.K.; Fung, T. Classification of mangrove species using combined WorldView-3 and LiDAR data in Mai Po nature reserve, Hong Kong. *Remote Sens.* **2019**, *11*, 2114. [\[CrossRef\]](#)
23. Wan, L.; Lin, Y.; Zhang, H.; Wang, F.; Liu, M.; Lin, H. GF-5 hyperspectral data for species mapping of mangrove in Mai Po, Hong Kong. *Remote Sens.* **2020**, *12*, 656. [\[CrossRef\]](#)
24. Erdős, L.; Kröel-Dulay, G.; Bátori, Z.; Kovács, B.; Németh, C.; Kiss, P.J.; Tölgyesi, C. Habitat heterogeneity as a key to high conservation value in forest-grassland mosaics. *Biol. Conserv.* **2018**, *226*, 72–80. [\[CrossRef\]](#)
25. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [\[CrossRef\]](#)
26. Talukdar, S.; Singha, P.; Mahato, S.; Shahfahad; Pal, S.; Liou, Y.-A.; Rahman, A. Land-use land-cover classification by machine learning classifiers for satellite observations—A review. *Remote Sens.* **2020**, *12*, 1135. [\[CrossRef\]](#)
27. Huang, C.; Davis, L.S.; Townshend, J.R.G. An assessment of support vector machines for land cover classification. *Int. J. Remote Sens.* **2002**, *23*, 725–749. [\[CrossRef\]](#)
28. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
29. Pouteau, R.; Meyer, J.-Y.; Taputuarai, R.; Stoll, B. Support vector machines to map rare and endangered native plants in Pacific islands forests. *Ecol. Inform.* **2012**, *9*, 37–46. [\[CrossRef\]](#)
30. Sabat-Tomala, A.; Raczko, E.; Zagajewski, B. Comparison of support vector machine and random forest algorithms for invasive and expansive species classification using airborne hyperspectral data. *Remote Sens.* **2020**, *12*, 516. [\[CrossRef\]](#)

31. Whiteside, T.G.; Boggs, G.S.; Maier, S.W. Comparing object-based and pixel-based classifications for mapping savannas. *Int. J. Appl. Earth Obs.* **2011**, *13*, 884–893. [[CrossRef](#)]
32. Lawrence, R.L.; Moran, C.J. The AmericaView classification methods accuracy comparison project: A rigorous approach for model selection. *Remote Sens. Environ.* **2015**, *170*, 115–120. [[CrossRef](#)]
33. Schmidt, M.A.R.; Bressiani, J.X.; Dos Reis, P.A.; Salla, M.R. Evaluation of the performance of image classification methods in the identification of vegetation. *J. Urban Environ. Eng.* **2016**, *10*, 62–71. [[CrossRef](#)]
34. Lu, D.; Weng, Q. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.* **2007**, *28*, 823–870. [[CrossRef](#)]
35. Khatami, R.; Mountrakis, G.; Stehman, S.V. A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. *Remote Sens. Environ.* **2016**, *177*, 89–100. [[CrossRef](#)]
36. San Miguel-Ayanz, J.; Biging, G.S. Comparison of single-stage and multi-stage classification approaches for cover type mapping with TM and SPOT data. *Remote Sens. Environ.* **1997**, *59*, 92–104. [[CrossRef](#)]
37. Abou El-Magd, I.; Tanton, T.W. Improvements in land use mapping for irrigated agriculture from satellite sensor data using a multi-stage maximum likelihood classification. *Int. J. Remote Sens.* **2003**, *24*, 4197–4206. [[CrossRef](#)]
38. Dibs, H.; Idrees, M.O.; Alsalhin, G.B.A. Hierarchical classification approach for mapping rubber tree growth using per-pixel and object-oriented classifiers with SPOT-5 imagery. *Egypt. J. Remote Sens. Space Sci.* **2017**, *20*, 21–30. [[CrossRef](#)]
39. Manandhar, R.; Odeh, I.O.A.; Ancev, T. Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. *Remote Sens.* **2009**, *1*, 330–344. [[CrossRef](#)]
40. Thakkar, A.K.; Desai, V.R.; Patel, A.; Potdar, M.B. Post-classification corrections in improving the classification of Land Use/Land Cover of arid region using RS and GIS: The case of Arjuni watershed, Gujarat, India. *Egypt. J. Remote Sens. Space Sci.* **2017**, *20*, 79–89. [[CrossRef](#)]
41. Van De Voorde, T.; De Genst, W.; Canters, F. Improving pixel-based VHR land-cover classifications of urban areas with post-classification techniques. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 1017–1027.
42. Rozenstein, O.; Karnieli, A. Comparison of methods for land-use classification incorporating remote sensing and GIS inputs. *Appl. Geogr.* **2011**, *31*, 533–544. [[CrossRef](#)]
43. Rapinel, S.; Clément, B.; Magnanon, S.; Sellin, V.; Hubert-Moy, L. Identification and mapping of natural vegetation on a coastal site using a Worldview-2 satellite image. *J. Environ. Manage* **2014**, *144*, 236–246. [[CrossRef](#)] [[PubMed](#)]
44. Kosmidou, V.; Petrou, Z.; Bunce, R.G.H.; Múcher, C.A.; Jongman, R.H.G.; Bogers, M.M.B.; Lucas, R.M.; Tomaselli, V.; Blonda, P.; Padoa-Schioppa, E.; et al. Harmonization of the Land Cover Classification System (LCCS) with the General Habitat Categories (GHC) classification system. *Ecol. Indic.* **2014**, *36*, 290–300. [[CrossRef](#)]
45. Adamo, M.; Tarantino, C.; Tomaselli, V.; Kosmidou, V.; Petrou, Z.; Manakos, I.; Lucas, R.M.; Múcher, C.A.; Veronico, G.; Marangi, C.; et al. Expert knowledge for translating land cover/use maps to General Habitat Categories (GHC). *Landsc. Ecol.* **2014**, *29*, 1045–1067. [[CrossRef](#)]
46. Thoonen, G.; Spanhove, T.; Vanden Borre, J.; Scheunders, P. Classification of heathland vegetation in a hierarchical contextual framework. *Int. J. Remote Sens.* **2013**, *34*, 96–111. [[CrossRef](#)]
47. Dudgeon, D.; Corlett, R. *The Ecology and Biodiversity of Hong Kong, Revised Edition*; Lion Nature Education Foundation: Hong Kong, China, 2011.
48. Environmental Bureau. *Hong Kong Biodiversity Strategy and Action Plan 2016–2021*; The Government of the Hong Kong Special Administrative Region: Hong Kong, China, 2016.
49. Corlett, R.T. Environmental forestry in Hong Kong: 1871–1997. *For. Ecol. Manag.* **1999**, *116*, 93–105. [[CrossRef](#)]
50. Balthazar, V.; Vanacker, V.; Lambin, E.F. Evaluation and parameterization of ATCOR3 topographic correction method for forest cover mapping in mountain areas. *Int. J. Appl. Earth Obs.* **2012**, *18*, 436–450. [[CrossRef](#)]
51. Planning Department of Hong Kong Government. Land Utilization in Hong Kong. Available online: [https://www.pland.gov.hk/pland\\_en/info\\_serv/open\\_data/landu/index.html](https://www.pland.gov.hk/pland_en/info_serv/open_data/landu/index.html) (accessed on 18 October 2021).
52. McCoy, R.M. *Field Methods in Remote Sensing*; Guilford Press: New York, NY, USA, 2005.
53. Lai, A.C.S.; So, A.C.T.; Ng, S.K.C.; Jonas, D. The territory-wide airborne light detection and ranging survey for the Hong Kong Special Administrative Region. In Proceedings of the 33rd Asian Conference on Remote Sensing, Pattaya, Thailand, 26–30 November 2012; pp. 1682–1690.
54. Lands Department of Hong Kong Government. Digital Topographic Map. Available online: <https://www.landsd.gov.hk/en/survey-mapping/mapping/multi-scale-topographic-mapping/digital-map.html> (accessed on 18 October 2021).
55. Ashworth, J.M.; Corlett, R.T.; Dudgeon, D.; Melville, D.S.; Tang, W.S.M. *Hong Kong Flora and Fauna: Computing Conservation*; World Wide Fund for Nature Hong Kong: Hong Kong, 1993.
56. Environmental Resources Management. *2008 Update of Terrestrial Habitat Mapping and Ranking Based on Conservation Value*; Final Report to Sustainable Development Division, Hong Kong Special Administrative Region Government: Hong Kong, China, 2010.
57. Grizonnet, M.; Michel, J.; Poughon, V.; Inglada, J.; Savinaud, M.; Cresson, R. Orfeo ToolBox: Open source processing of remote sensing images. *Open Geospat. Data Softw. Stand.* **2017**, *2*, 15. [[CrossRef](#)]

58. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [[CrossRef](#)]
59. Zhu, Y.; Liu, K.; Liu, L.; Myint, S.W.; Wang, S.; Liu, H.; He, Z. Exploring the potential of WorldView-2 red-edge band-based vegetation indices for estimation of mangrove leaf area index with machine learning algorithms. *Remote Sens.* **2017**, *9*, 1060. [[CrossRef](#)]
60. Solano, F.; Di Fazio, S.; Modica, G. A methodology based on GEOBIA and WorldView-3 imagery to derive vegetation indices at tree crown detail in olive orchards. *Int. J. Appl. Earth Obs.* **2019**, *83*, 101912. [[CrossRef](#)]
61. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *J. Sens.* **2017**, *2017*, 1353691. [[CrossRef](#)]
62. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
63. Gitelson, A.A.; Merzlyak, M.N. Remote sensing of chlorophyll concentration in higher plant leaves. *Adv. Space Res.* **1998**, *22*, 689–692. [[CrossRef](#)]
64. Daughtry, C.S.T.; Walthall, C.L.; Kim, M.S.; de Colstoun, E.B.; McMurtrey, J.E. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [[CrossRef](#)]
65. Kupidura, P. The comparison of different methods of texture analysis for their efficacy for land use classification in satellite imagery. *Remote Sens.* **2019**, *11*, 1233. [[CrossRef](#)]
66. Zhou, J.; Yan Guo, R.; Sun, M.; Di, T.T.; Wang, S.; Zhai, J.; Zhao, Z. The Effects of GLCM parameters on LAI estimation using texture values from Quickbird Satellite Imagery. *Sci. Rep.* **2017**, *7*, 7366. [[CrossRef](#)]
67. Oksanen, T. Shape-describing indices for agricultural field plots and their relationship to operational efficiency. *Comput. Electron. Agr.* **2013**, *98*, 252–259. [[CrossRef](#)]
68. R Core Team. *R. A Language and Environment for Statistical Computing*, 3.6.3; R Foundation for Statistical Computing: Vienna, Austria, 2020.
69. Feng, Q.; Liu, J.; Gong, J. UAV remote sensing for urban vegetation mapping using random forest and texture analysis. *Remote Sens.* **2015**, *7*, 1074–1094. [[CrossRef](#)]
70. Fisher, P.; Arnot, C.; Wadsworth, R.; Wellens, J. Detecting change in vague interpretations of landscapes. *Ecol. Inform.* **2006**, *1*, 163–178. [[CrossRef](#)]
71. Loosvelt, L.; Peters, J.; Skriver, H.; Lievens, H.; Van Coillie, F.M.B.; De Baets, B.; Verhoest, N.E.C. Random Forests as a tool for estimating uncertainty at pixel-level in SAR image classification. *Int. J. Appl. Earth Obs.* **2012**, *19*, 173–184. [[CrossRef](#)]
72. Yang, Z.; D’Alpaos, A.; Marani, M.; Silvestri, S. Assessing the fractional abundance of highly mixed salt-marsh vegetation using random forest soft classification. *Remote Sens.* **2020**, *12*, 3224. [[CrossRef](#)]
73. Phillips, S.J.; Dudík, M.; Elith, J.; Graham, C.H.; Lehmann, A.; Leathwick, J.; Ferrier, S. Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecol. Appl.* **2009**, *19*, 181–197. [[CrossRef](#)] [[PubMed](#)]
74. Foody, G.M. Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 627–633. [[CrossRef](#)]
75. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [[CrossRef](#)]
76. Olofsson, P.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **2013**, *129*, 122–131. [[CrossRef](#)]
77. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [[CrossRef](#)]
78. Bradter, U.; O’Connell, J.; Kunin, W.E.; Boffey, C.W.H.; Ellis, R.J.; Benton, T.G. Classifying grass-dominated habitats from remotely sensed data: The influence of spectral resolution, acquisition time and the vegetation classification system on accuracy and thematic resolution. *Sci. Total Environ.* **2020**, *711*, 134584. [[CrossRef](#)] [[PubMed](#)]
79. Jiao, L.; Sun, W.; Yang, G.; Ren, G.; Liu, Y. A hierarchical classification framework of satellite multispectral/hyperspectral images for mapping coastal wetlands. *Remote Sens.* **2019**, *11*, 2238. [[CrossRef](#)]
80. Ghosh, A.; Joshi, P.K. A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. *Int. J. Appl. Earth Obs.* **2014**, *26*, 298–311. [[CrossRef](#)]
81. Myint, S.W.; Gober, P.; Brazel, A.; Grossman-Clarke, S.; Weng, Q. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens. Environ.* **2011**, *115*, 1145–1161. [[CrossRef](#)]
82. Wood, E.M.; Pidgeon, A.M.; Radeloff, V.C.; Keuler, N.S. Image texture as a remotely sensed measure of vegetation structure. *Remote Sens. Environ.* **2012**, *121*, 516–526. [[CrossRef](#)]
83. Du, P.; Xia, J.; Zhang, W.; Tan, K.; Liu, Y.; Liu, S. Multiple classifier system for remote sensing image classification: A review. *Sensors* **2012**, *12*, 4764–4792. [[CrossRef](#)]
84. Buck, O.; Millán, V.E.G.; Klink, A.; Pakzad, K. Using information layers for mapping grassland habitat distribution at local to regional scales. *Int. J. Appl. Earth Obs.* **2015**, *37*, 83–89. [[CrossRef](#)]
85. Harris, P.M.; Ventura, S.J. The integration of geographic data with remotely sensed imagery to improve classification in an urban area. *Photogramm. Eng. Remote Sens.* **1995**, *61*, 993–998.



86. Sulla-Menashe, D.; Gray, J.M.; Abercrombie, S.P.; Friedl, M.A. Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product. *Remote Sens. Environ.* **2019**, *222*, 183–194. [[CrossRef](#)]
87. Somay, L.; Szigeti, V.; Boros, G.; Ádám, R.; Báldi, A. Wood pastures: A transitional habitat between forests and pastures for dung beetle assemblages. *Forests* **2021**, *12*, 25. [[CrossRef](#)]
88. Shea, M.E.; Clayton, M.K.; Townsend, P.A.; Berg, S.; Elza, H.; Mladenoff, D.J. Identifying ecotone location using the co-occurrence property. *J. Veg. Sci.* **2021**, *32*, e12929. [[CrossRef](#)]
89. Jung, M.; Dahal, P.R.; Butchart, S.H.M.; Donald, P.F.; De Lamo, X.; Lesiv, M.; Kapos, V.; Rondinini, C.; Visconti, P. A global map of terrestrial habitat types. *Sci. Data* **2020**, *7*, 256. [[CrossRef](#)]
90. Kwong, I.H.Y.; Fung, T. Tree height mapping and crown delineation using LiDAR, large format aerial photographs, and unmanned aerial vehicle photogrammetry in subtropical urban forest. *Int. J. Remote Sens.* **2020**, *41*, 5228–5256. [[CrossRef](#)]