Review

Change Detection Techniques Based on Multispectral Images for Investigating Land Cover Dynamics

Dyah R. Panuju 1,2,* , David J. Paull 2 and Amy L. Griffin 3

1 Department of Soil Science and Land Resource, Bogor Agricultural University, Bogor 16680, Indonesia
2 School of Science, UNSW Canberra, Northcott Drive, Campbell, ACT 2600, Australia; d.paull@adfa.edu.au
3 School of Science, RMIT University, Melbourne, VIC 3001, Australia; amy.griffin@rmit.edu.au
* Correspondence: panuju@apps.ipb.ac.id

Received: 30 March 2020; Accepted: 27 May 2020; Published: 1 June 2020

Abstract: Satellite images provide an accurate, continuous, and synoptic view of seamless global extent. Within the fields of remote sensing and image processing, land surface change detection (CD) has been amongst the most discussed topics. This article reviews advances in bitemporal and multitemporal two-dimensional CD with a focus on multispectral images. In addition, it reviews some CD techniques used for synthetic aperture radar (SAR). The importance of data selection and preprocessing for CD provides a starting point for the discussion. CD techniques are, then, grouped based on the change analysis products they can generate to assist users in identifying suitable procedures for their applications. The discussion allows users to estimate the resources needed for analysis and interpretation, while selecting the most suitable technique for generating the desired information such as binary changes, direction or magnitude of changes, “from-to” information of changes, probability of changes, temporal pattern, and prediction of changes. The review shows that essential and innovative improvements are being made in analytical processes for multispectral images. Advantages, limitations, challenges, and opportunities are identified for understanding the context of improvements, and this will guide the future development of bitemporal and multitemporal CD methods and techniques for understanding land cover dynamics.

Keywords: bitemporal; change detection; multitemporal; multispectral images; land cover; time series; trajectory

1. Introduction

Remote sensing change detection (CD) is commonly defined as a process to identify differences in geographical surface phenomena over time [1,2]. CD is also defined as a process to identify significant differences in sequential pixel appearances due to the emergence, disappearance, movement, or shape alteration of objects [3]. The detection process includes the observation and evaluation of differences over time to document the spectral and temporal progression of biophysical and physical phenomena [4].

The spectral differences between two observations due to land surface alteration can be used to identify a change when the magnitude of spectral differences is greater than any distortions [5] that arise from internal image-production-related factors or external factors. Acquisition geometry and sensor characteristics are among the internal factors [6], whereas atmospheric absorption and scattering differences associated with water vapor or aerosol concentrations, temporal variations in the solar zenith or azimuth angles, and inconsistencies of sensor calibration are amongst the external factors [7].

Several reviews of CD methods have discussed diverse aspects including approaches, algorithms, and applications. Nelson [8] and Singh [1] classified CD techniques into (1) comparative analyses based
on post-classified data and (2) simultaneous analyses of multitemporal images. Later, Mouat, Mahin and Lancaster [4] focused on the use of CD to investigate ecosystem change by employing several sensors, including Advanced Very High Resolution Radiometer (AVHRR), Landsat Thematic Mapper (TM) and Multispectral Scanner (MSS), Satellite pour l’Observation de la Terre (SPOT), and aerial photographs. Lu, et al. [9] and Coppin, Jonckheere, Nackaerts, Muys and Lambin [7] also investigated CD for ecological applications. Lu, Mauel, Brondizio and Moran [9] grouped ecosystem explorations into ten classes, including general or specific aspects of land use, environment, forest, urban, and agricultural topics. Coppin, Jonckheere, Nackaerts, Muys and Lambin [7] comprehensively discussed CD from three viewpoints, i.e., ecosystem properties, preprocessing procedures, and CD techniques, then, Radke, et al. [3] focused on CD algorithms. Hussain, et al. [10] differentiated CD techniques into pixel-based and object-based methods while Bovolo and Bruzzone [11] discussed CD from the perspective of data fusion at the feature and decision levels.

In general, these reviews indicate that optical sensors such as AVHRR, Landsat, Moderate Resolution Imaging Spectroradiometer (MODIS), and SPOT are the main datasets providing historically frequent observations. Meanwhile, recent European constellations, such as the Sentinels should play an important role in the future. Some researchers have also employed synthetic aperture radar (SAR) data to deal with the challenge of cloud cover in tropical regions or to derive information that could not be generated by using optical images. However, none of the reviews discussed CD from the perspective of users wishing to have an overview of possible data and suitable techniques for generating required change information. This paper, therefore, reviews CD techniques that use pixel-wise and region-based techniques to conduct bitemporal or trajectory analysis to multitemporal and time series datasets that mainly employ multispectral images. This review should assist users to collect suitable datasets and to select the most appropriate CD techniques to support the users’ objectives. The discussion of what is offered by each CD technique and its advantages and drawbacks provides insights about the technique’s time, cost, and knowledge requirements. The review starts by considering the impact of data selection and preprocessing on the detection of valid changes for the different approaches and various CD products. Then, constraints, challenges, and future directions of CD techniques are discussed.

2. Data Selection and Its Impacts on Change Detection

2.1. The Requirements and the Challenges of Change Analysis

The definition of change guides important steps for producing a valid CD map. A valid map must ensure that spectral difference is a result of phenomena on the ground rather than distortions, errors, or noise. For change analysis, data pairs should be spectrally, spatially, and temporally comparable. Thus, the analyst needs to know how to prepare conditionally comparable image pairs through preprocessing before implementing CD.

CD has primarily employed optical images, which parallels the development of airborne and spaceborne sensors and their long-term availability and recurring observations [12]. AVHRR and SPOT vegetation archives have facilitated remotely-sensed monitoring at a global scale [13]. Earth observation satellites and their systematic products support this function. AVHRR and SPOT now provide continuous observations for historical research spanning more than four decades.

Landsat images have also contributed to the monitoring of land surfaces with evolving sensors beginning with the multispectral scanner (MSS), thematic mapper (TM), enhanced thematic mapper (ETM) [14], and the latest sensor, the Operational Land Imager (OLI) [15]. The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) was collaboratively constructed by the Ministry of International Trade and Industry (MITI) Japan and National Aeronautics and Space Administration (NASA) [16] to provide an alternative source of images that are comparable with Landsat. The Advanced Land Observing Satellite (ALOS) Advanced Visible and Near Infrared (AVNIR), a product of the Japan Aerospace Exploration Agency (JAXA), is another option that has similar spatial resolution to Landsat.
and ASTER [17]. Unfortunately, the production of AVNIR was decommissioned, in 2011, due to the end of the satellite’s lifespan [18]. Very high spatial resolution images such as World-View, Quickbird, and IKONOS have accommodated ground truthing [19,20] and detailed, object-based CD such as applications focused on tree-crown detection [21,22], building change detection [23], and damage assessment [24]. Later, a Copernicus program designed by the European Space Agency (ESA) launched a constellation of Sentinel satellites to provide continuous monitoring of Earth surfaces, generating satellite images and in-situ data, as well as delivering several information services, i.e., atmosphere, marine, land, climate change, security, and emergency [25,26]. These image sources have facilitated CD from a global to local scales. However, not much research has compared or combined several optical sensors for simultaneous exploration of land cover change due to difficulties in ensuring spectral, spatial, and temporal consistencies among sensors [27].

Although optical imagery is beneficial for monitoring the global extent, cloud cover limits its usability for monitoring change at some locations such as in tropical montane environments. Figure 1 shows all available Landsat 8 acquisitions in 2014, for a mountainous location in West Java, Indonesia; many acquisitions are affected by this problem. Employing active sensors such as Synthetic Aperture Radar (SAR) may be an option for investigating areas with severe cloud problems due to its ability to penetrate cloud. SAR has been explored for CD applications such as monitoring deforestation in tropical regimes [28], ice cover [29], or hazards such as floods [30,31]. However, until the launch of the Copernicus Program, in 2014, by the ESA, which is capable of generating five-day revisit periods through Sentinel-1, the temporal frequency of SAR images was not comparable to optical images.

![Figure 1. Visual condition of all available Landsat 8 images for Path 122 Row 065, during 2014. Cloud cover affects many of the images.](https://earthexplorer.usgs.gov/)

2.2. The Role of Ancillary Data in Change Analysis

ESRI’s GIS Dictionary defines ancillary data in terms of image processing as “data from sources other than remote sensing, used to assist in analysis and classification or to populate metadata” [32]. Ancillary data are needed due to the limitations of spectral information for describing change [33].
Ancillary data can be used for the following: (a) pre-classification, (b) post-classification, or (c) direct inclusion during classification [34]. Hayes and Sader [35] demonstrated the use of ancillary data to assist with the selection of training and testing samples before classification and reducing the subjectivity of and assisting in the interpretation of change phenomena following image classification. Oetter, et al. [36] reconfirmed these roles of ancillary data. Ancillary data can be used to define land cover change trajectories and to refine classification as they can complement or even substitute for ground truth data [37]. Their use can improve CD accuracy by about 5%–10% [38] or even as much as 15% [34]. Moreover, Röder, et al. [39] used ancillary data to mask out unwanted land cover types post-classification.

Ancillary data can be grouped based on their source into the following groups: (1) derived images, (2) maps, and (3) documents. Statistical parameters such as the slope and intercept of statistical models of change occurrence and various physical and anthropogenic factors enrich change information through correlation analysis [40], while digital elevation models have been used to investigate forest change [41]. Thematic maps of topography, forest inventory, forest stands, land cover, or aerial photographs are commonly employed ancillary data for change analysis. Documents such as master plans have been used to comprehend spatial patterns of urban growth [42]. Désclée, Bogaert and Defourny [41] selectively employed multiple ancillary data sources for different purposes, such as the following: elevation for differentiating change occurrences in lowlands and highlands, slope for minimizing the effect of shadowing in mountainous areas, aspect for indicating the suitability of an area for wildlife, and vegetation maps for categorising vegetation types.

3. Preprocessing and Its Impact on Change Detection

Spectral, spatial, and temporal consistency and comparability of images are essential in CD. Bias or false detections should not overshadow real phenomena for revealing accurate change. Preprocessing procedures are essential to generate consistent data and to avoid such errors. Borrowing the concept from data mining, preprocessing involves data integration, cleaning, normalization, and transformation [43]. Data integration is an important step to identify redundancies and inconsistencies while matching and arranging datasets. Integrating remote sensing data relates to aligning images through rectification and co-registration using the same reference and control points. In addition, it involves reprojecting and resampling images from various formats into a common format to allow a comparison. Data cleaning is an essential procedure for reducing noise and for handling missing data and it includes atmospheric correction and cloud masking for optical images or filtering speckles for microwave data. Various strategies and techniques related to preprocessing are used to examine image quality [44], to deal with error correction [6], and to tackle noise that causes missing data or complicates visualization and interpretation. Normalizing images is necessary to produce consistent yet comparable datasets, particularly, when multitemporal or time series observations are involved [45,46]. Meanwhile, transformation is used to enhance data quality or to highlight features by reducing dimensions or combining layers of multispectral images to represent biophysical properties [47–49]. Key issues related to CD are discussed in the following sections.

3.1. Integrating Data through Geometric Correction and Co-Registration

Two or more images of the same location taken from different acquisitions need to be integrated through co-registration preceding CD. According to Kennedy and Cohen [50], aligning two images requires image tie points (ITPs) to identify candidate objects, then, performing a calculation based on ITPs to transform each image’s coordinates into another image’s spatial coordinates, and finally, resampling the transformed image. A semi-automated procedure of geometric correction based on polynomial transformation has successfully been applied in a forest change study [51]. Processing multitemporal or massive numbers of images has been facilitated by automated co-registration systems such as COSI-corr [52] and GeoCDX [53]. A new technique called Patch-Wise Co-Registration (PWCR)
uses different sensors and view angles to support long temporal period change analysis employing several sensors [54].

Co-registration can enrich information and improve image quality in various applications. A ground deformation study, for example, benefited from employing co-registration to enrich information through integrating several data sources including a DEM, film, and scan artefacts [52,55]. Bias in elevation data, used for detecting glacier elevation changes, was reduced by integrating ancillary data for co-registration [56]. Moreover, co-registration assisted in detecting anomalous change and discriminated between real anomalies and incidental differences through a small adjustment [57].

The importance of geometric correction has been demonstrated by assessing the errors that resulted from misregistration. Misregistration, misalignment, or any positional error results in bias [58] and in turn affects the accuracy of CD. According to Townshend, et al. [59], misregistration can contribute up to 50% of CD error. Dai and Khorram [60] found that in order to obtain CD error of less than 10%, registration error should be at most 20%. Moreover, misregistration resulted in missing true changes that were distributed away from image edges. Roy [61] highlighted that errors have greater impacts for composite images than non-composite images due to greater variability of orbits, view zenith angles, and degrees of misregistration. In addition, Verbyla and Boles [58] demonstrated that false detection of land cover change (LCC) was greater with a larger number of land cover classes.

Several factors contribute to sensitivity to misregistration within a set of images, for instance the spatial heterogeneity of areas [59] or the noise removal approach being applied for preprocessing [62]. Moreover, the spatial and temporal properties of spectral bands differ in their sensitivity to misregistration; Dai and Khorram [60], for example, demonstrated that TM Band 4 was more sensitive to misregistration as compared with other spectral bands.

3.2. Data Cleaning and Normalizing Images

In a remote sensing context, data cleaning and normalization can be highly interrelated, with both aiming to minimize noise. Normalization removes noise or unexpected effects that alter spectral characteristics for reasons other than actual land surface changes [63]. Either sensor-related or scene-related factors may cause noise [64]. Sensor- and system-calibration differences can be single or co-contributors to sensor-related noise [63]. Sensor aging can lead to bias, and hence necessitates renewed system calibration [64]. Differences in illumination and viewing angles, the variability of atmospheric conditions and their effects, reflectance anisotropy, uneven topography, as well as change on the ground, are among scene-related factors [64]. Johnson and Kasischke [65] highlighted that normalization was essential to correct for differences due to sensor calibration, atmospheric conditions, and phenological variability. According to Hayes and Sader [35], radiometric normalization is practical for generating consistent data for CD when sufficient information is not available for absolute correction. The procedure is considered mandatory for CD, and facilitates discrimination between artefacts and real changes [66].

Another essential reason for performing normalization is that the radiometric response of the same target observed at different times is likely to be different. In this respect, radiometric normalization or calibration requires information regarding atmospheric and sensor conditions, as highlighted by Song, et al. [67]. According to Johnson and Kasischke [65], solar illumination normalization is vital for change vector analysis (CVA). Normalized images generated from transforming the radiance value to an image with a common radiometric scale or to surface reflectance [64,67] is a standard data processing step that results in reliable change information.

A survey of the literature shows that there are two types of radiometric calibration approaches, namely, absolute and relative approaches [64,68,69]. By comparing absolute and relative normalization, it was shown that absolute normalization that generated surface reflectance was superior in forming consistent temporal data for normalized reflectance trajectories [69]. Absolute calibration, however, requires complex atmospheric parameters, such as the surface radiance of atmospheric optical depth and sensor calibration data [46]. In situ measurement of atmospheric parameters is generally not available for many parts of the world; therefore, relative calibration is often a more practical option.
Considering its minimal requirement for atmospheric parameters and due to the importance of calibration for aligning data pairs, relative radiometric calibration is a popular procedure for change analysis. According to [45], relative radiometric calibration is meant to find inconsistencies between pairs of images through rectifying subject images to reference images [46]. Relative normalization reduces the dependencies to non-scene changes caused by the differences of observation angle, the response of the sensor, and atmospheric conditions between paired images [70].

A popular strategy for relative radiometric calibration is to select unchanged features called pseudo invariant features (PIFs). PIFs should have a relatively constant reflectivity and remain unchanged during different times of observation; they are generally human-constructed features such as roads, buildings, airport runways, and dams [70]. Principal component analysis and quality control have been used to select PIFs for temporal image normalization prior to change analysis [71]. Alternatively, regression of homologous reference invariant data can also be used for PIF selection [72]. On the basis of the PIF principle, multivariate alteration detection (MAD) has been successfully applied to normalize several sensors, such as Landsat, ASTER, and MODIS [70]. Iterative slow feature analysis (ISFA), which minimizes radiometric variance, is an alternative method for normalization [73].

Despite the benefits of radiometric normalization, there are limitations and drawbacks to performing the procedure. Rogan, et al. [74] demonstrated the ineffectiveness of normalization for removing spatial heterogeneity of haze. Furthermore, Lamparelli, et al. [75] reported that normalization distorted temporal patterns when monitoring crop growth.

3.3. Image Transformation Preceding Change Analysis

Converting digital numbers into surface reflectance is a type of transformation, but it may also be considered to be an image calibration. Transformation can characterize physical properties from fewer spectral images by converting multispectral images into a smaller number of components with minimum information loss [47]. Principal component analysis (PCA) is a spectral volume-reduction method that transforms the original data into linear combinations of uncorrelated spectral information [76]. The technique has been used to integrate two observations as a bundle of datasets, and then changes are identified from the generated components with smaller variances. PCA can be efficiently computed, however, the derived components present considerable problems with regard to the interpretation of physical features [77].

Another common transformation type, spectral indices, combine two or more bands to represent biophysical properties such as vegetation greenness. Jensen [78] explained that selection of spectral bands to represent the biophysical characteristics of an object should be taken carefully to maximize contrast between different states.

Vegetation indices, particularly the Normalized Difference Vegetation Index (NDVI), have been the most popular indices for various applications [79–81]. The applications include identification of the expansion or shrinkage of vegetated areas such as agriculture [82] or forest [83], and within a built-up environment [84]. Other indices include the Normalized Difference Built-up Index (NDBI), which represents built-up areas [85], Normalized Difference Impervious Surface Index (NDISI) [86], which identifies impervious surfaces, and Normalized Difference Water Index (NDWI), used to characterize water areas [87].

Tasseled Cap (TC) is another type of spectral transformation that describes the relationships among bands and extracts’ biophysical characteristics [88,89]. TC generates three features that represent brightness, greenness, and wetness [88]. TC coefficients have been developed using various sensors, including Landsat TM [89], Landsat 7 ETM+ [90], Landsat 8 OLI [91], MODIS [92], and Sentinel-2 [93]. The applications of TC include the study of plant succession [94], the development of annual land cover maps [95], and the spatiotemporal quantification of urban environments [96]. While having the advantages of reducing the number of spectral bands and being easy to relate to biophysical features, TC is reported to depend on sensor properties [77].
Transformed images can be used to explore various conditions affected by different drivers of change. Despite many transformations having been presented in the literature, there are few published papers comparing their use for time series analysis.

3.4. Essential Strategies to Suppress Errors

Various strategies have been implemented to optimize CD accuracy and to suppress errors, such as selecting the most suitable image type and temporal recurrence. The errors due to differences in external conditions can be minimized by employing the same sensor [9]. This strategy minimizes problems of differing spatial or radiometric resolution, but it does not guarantee comparable radiometric properties. Employing the same sensor can yield better radiometric corrections as different sensors complicate the interpretation of change analysis [63]. However, long-term temporal monitoring likely requires different sensors due to finite satellite and sensor lifetimes or missing periodical observations. The use of multiple sensors, therefore, produces differences in spectral properties, which necessitate preprocessing.

Acquiring anniversary date image sets, when available, is another strategy for selecting observations in order to minimize the effects of seasonal differences [97–99]. The selection of anniversary dates can alleviate undesirable variability [65,100] while obtaining high fidelity acquisition of multispectral images for bitemporal change detection [101]. However, anniversary date observations are often unavailable in tropical environments that frequently experience atmospheric disturbances such as cloud cover [102].

Compositing images, an approach that provides consistent datasets at a specific time interval, and potentially reduces atmospheric disturbance, was introduced by Holben [103]. Numerous compositing algorithms help in reducing noise and provide consistent yet adjusted data. Compositing algorithms can assist in observing natural phenomena, such as the use of temporal bidirectional composites, which resemble crop conditions [104]. Composite images generated through temporal synthesis can be less contaminated by cloud, systematic noise, or random errors, than a single image [105,106]. Compositing has been beneficial in the production of regional to global scale mapping, which is less likely to be composed from a single observation, by reducing data volumes and noise [107]. Longer temporal compositing has been suggested as a strategy for improving maps of forest cover [108].

However, composite images are not suitable for detecting highly dynamic changes occurring over hours, such as those caused by meteorological events [109]. Composite images containing a single parameter, such as a minimum or maximum, could record noise instead of real data [107]. Compositing has also been reported to contain residual artifacts that limit capabilities for temporal comparison [106]. Moreover, composite images with a large view angle do not always result in cloud-free data, particularly, when the maximum value composite algorithm is employed [110].

4. Change Detection Techniques and Synthesis Product of Changes

Following successful preprocessing, various synthesis products can be yielded through numerous CD techniques depending upon the context and purpose of the study; for instance, composition, configuration, and probability of changes [111]. A survey of literature reveals that CD can generate maps of binary change, change types, magnitude and direction of changes, probability of changes, or temporal trajectories. In addition, users can use change information to simulate future land cover. Each possible CD product and the techniques used to derive the product are discussed in the following subsections.

4.1. Binary Change Maps

Binary change maps are synthesis products of bitemporal CD that highlight the location of changes, and therefore are standard information for decision-making systems. Various techniques have been applied to create them in diverse ecosystems, including simple differencing [112] or ratioing [113] of either original multispectral images or their derivatives, such as indices and principle component analysis (PCA) or MAD [114], followed by thresholding [115].
Image differencing examines the dissimilarity of image properties on a per-pixel basis by simply subtracting one digital image value from another to yield a numerical difference between pixel pairs [112]. Differencing is generally performed on spectrally and atmospherically preprocessed images to minimize errors [35,112]. The approach has often been used with indices instead of original spectral bands for better association with biophysical properties. Another simple technique for CD is ratioing, which is dividing pixel values in one digital image by those of another [116].

The result of differencing and ratioing cannot directly be interpreted as changed or unchanged as thresholding is needed to differentiate the binary classes. Thresholding should consider image content and the selection of values [117]. Sahoo, et al. [118] reviewed and classified thresholding methods as follows: (1) point dependent techniques, such as Otsu method [119], entropic method, and minimum error method [120]; (2) region-dependent techniques, for example, histogram transformation method, and relaxation method [121–123]; (3) local thresholding; and (4) multi-thresholding methods, for example, amplitude segmentation method, and uniform contrast method [124,125]. The determination of the threshold can be based on various criteria, i.e., (a) a normal model, (b) local intensity distribution, (c) the Poisson distribution, and (d) a stable number of regions [117]. The challenge of differencing and ratioing when multispectral bands are employed is to isolate changed from unchanged pixels. Ridd and Liu [126] proposed a new transformation to develop a single-band change image from multispectral differencing and ratioing by using the Chi-square principle.

Principal component analysis for bitemporal CD uses either a covariance or correlation matrix [47,127]. As a redundancy reduction technique, PCA transforms the original matrix into new linear forms, and it has been used to detect change from Landsat data [128]. The change information is mainly interpreted from the third and fourth principal components, which represent unique phenomena having small variances, whereas the first and second components generally represent variances of unchanged pixels, which are larger than those of the changed pixels [115]. Comparing standardized and non-standardized PCA, Guirguis, et al. [129] demonstrated that the standardized version had a better ability to detect changes. PCA has been used as either a stand-alone technique [115,128,130,131] or in combination with various classifiers or techniques [76,132–134].

Another popular technique that produces bitemporal change detection is MAD [114] and its modification, Iteratively Reweighted MAD (IRMAD), which relies on canonical correlation and has been demonstrated to be robust for CD [135]. MAD is a well-established method for bitemporal change detection [136] that combines registration, spectral normalization, and spatial autocorrelation [114]. The method is based on Hotteling’s canonical correlation in which two sets of vector images of the same place, acquired at two time points, M = (M1, . . . , Mk) and N = (N1, . . . , Nk), are transformed into new images Q = aT M and R = bT N. From p spectral bands in the original images, two images can be generated where each new image is composed of p MAD-variates. The MAD-variates are orthogonal to each other and are ordered by descending variance. Standardizing values by computing their correlation instead of their covariance has been suggested to be able to account for different scales due to different gain factors and atmospheric conditions [135,137]. IRMAD’s iterative processing is performed to improve separation among classes by adding more weight to no-change probabilities during the process [135]. The method has been applied to various datasets such as Bands 1–7 of Landsat [137], Channel 1 and Channel 2 of AVHRR [138], Bands 1–9 of ASTER [137], and synthetically derived images from 242 bands of Hyperion [139]. However, few studies have specifically compared the application of MAD or IRMAD to differing sensors or types of transformation.

Slow feature analysis (SFA) is another potential technique that separates changed from unchanged pixels by suppressing the invariant components of multitemporal images [140]. SFA is based on the principle that apparent change processes occur slowly on the unchanged pixels when they are observed temporally, indicated by sequential variance changes. This happens when the radiometric values of unchanged pixels are not zero, due to various factors such as atmospheric, sensor, and automated calibration effects. Nevertheless, they are comparatively smaller than those of changed pixels, allowing
unchanged pixels to be identified [73,140]. The technique was tested on Landsat data and was demonstrated to be superior to change vector analysis, PCA, and MAD [140].

Deep learning has recently been explored, mainly for object detection employing various high spatial resolution optical or SAR images. Deep-learning-based techniques have been implemented, particularly, for unsupervised change detection. The techniques include feature level change detection [141], novel difference image creation [142], discriminative feature learning [143], superpixel-based difference learning [144], and saliency analysis [145]. An improved version of slow feature analysis (SFA) based on deep learning is another contemporary technique for CD [146]. This combination employed a deep learning network to project two images taken from two time points, whereas SFA was used to suppress unchanged pixels and to highlight the changed pixels. Moreover, a combination of deep Siamese convolutional and multiple-layers recurrent neural networks was introduced to process CD in multisource VHR images [147]. Most techniques were tested on areas with sizes ranging from about 300 × 300 pixels to around 1500 × 1100 pixels, except a few test areas such as the one from Cao, Wang, Xavier, Yang and Southworth [142] that used a size of 6407 × 5521 pixels. Several advantages were claimed for deep-learning-based techniques, such as the capability to reduce noise and redundancy [141,142], robustness to register error [144], and ability to detect changes from heterogenous images [143]. Some studies have asserted that implementing deep learning requires a large number of training datasets [148,149], though others demonstrated that a pair of multispectral images could be employed to train networks for bitemporal CD [150].

Table 1 presents a summary of techniques used to derive binary change maps employing various data types applied to a range of ecosystems. The applications vary from monitoring forest to urban areas and employ either optical or radar datasets. The derived binary change maps vary in their spatial resolution from 0.5 meters (Quickbird) to 30 meters (Landsat).

<table>
<thead>
<tr>
<th>Methods</th>
<th>Data types</th>
<th>Application</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differencing</td>
<td>Landsat</td>
<td>Wetlands</td>
<td>[151]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Savanna</td>
<td>[152]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Forest</td>
<td>[153]</td>
</tr>
<tr>
<td>VI differencing</td>
<td>Landsat</td>
<td>Land cover</td>
<td>[154]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urban</td>
<td>[155]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Forest</td>
<td>[156]</td>
</tr>
<tr>
<td>Differencing combined with other techniques (regression, PCA, genetic algorithm)</td>
<td>Landsat</td>
<td>Forest</td>
<td>[157,158]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Land cover</td>
<td>[159]</td>
</tr>
<tr>
<td>Rationing</td>
<td>ERS-1 SAR</td>
<td>Forest</td>
<td>[113]</td>
</tr>
<tr>
<td></td>
<td>Shuttle Imaging Radar</td>
<td>Agriculture</td>
<td>[116]</td>
</tr>
<tr>
<td></td>
<td>ERS, ASAR and PALSAR</td>
<td>Urban</td>
<td>[160]</td>
</tr>
<tr>
<td>Principal Component Analysis</td>
<td>Landsat</td>
<td>Land cover</td>
<td>[115,128]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urban</td>
<td>[161]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wetland</td>
<td>[130]</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2</td>
<td>Insect infestation</td>
<td>[162]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Landslide</td>
<td>[163]</td>
</tr>
<tr>
<td></td>
<td>Landsat and SPOT HRV</td>
<td>Urban</td>
<td>[164]</td>
</tr>
<tr>
<td>Multivariate Alteration Detection</td>
<td>AVHRR and Landsat</td>
<td>Urban</td>
<td>[114]</td>
</tr>
<tr>
<td></td>
<td>Landsat</td>
<td>Forest</td>
<td>[69]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reservoir</td>
<td>[165]</td>
</tr>
<tr>
<td></td>
<td>Quickbird</td>
<td>Urban</td>
<td>[166]</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Data types</th>
<th>Application</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep-learning-based techniques</td>
<td>SPOT</td>
<td>Urban</td>
<td>[114]</td>
</tr>
<tr>
<td></td>
<td>Worldview</td>
<td>Wind turbine</td>
<td>[149]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rural</td>
<td>[148]</td>
</tr>
<tr>
<td>Gaofen (GF1)</td>
<td>Lake, river</td>
<td>[141]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Buildings</td>
<td>[145]</td>
<td></td>
</tr>
<tr>
<td>Aerial photograph</td>
<td>Forest</td>
<td>[167]</td>
<td></td>
</tr>
</tbody>
</table>

4.2. Types of Change: “From-To” Information

Type of change and the composition of change are typical products of post-classification CD that relate to land cover types assigned to a unit area, such as a pixel, patch, or administrative region. Typically, post-classification change analysis generates “from-to” information, which adds the detail of change type to a binary change map [168] and can be used to understand the spatial pattern of landscape change. To obtain accurate CD, it is important to optimize classification so that error is minimized, because CD error is propagated through each classification process [169]. Meanwhile, the performance of classifiers is constrained by image type and quality, ground observations, and computational capacity [170–172]. Unsupervised or supervised pixel-based techniques can generate accurate CD products [171]. This section avoids discussing advances in classification algorithms themselves but focuses instead on analytical techniques and applications that generate types and composition of change.

Apart from post-classification change analysis, “from-to” information can be derived from local feature- or region-based approaches [173]. The region-based approaches have been shown to be superior for CD as compared with pixel-based approaches in terms of the following: (a) their ability to distinguish different land cover types, (b) their insensitivity to misregistration, and (c) the change of feature representation that may be well differentiated by combining texture metrics [174]. Recently developed region-based approaches include the regularized level set approach [175], faster region-based convolutional neural network (Faster R-CNN) [176], and the combined use of pixel-based and region-based approaches [177,178].

Table 2 summarizes relevant published research according to contextual methods, data sources, and ecosystem applications used to generate change type and composition of change information. Post-classification CD is the main technique for generating type of change. The combination of post-classification techniques with other methods has improved CD sensitivity, and has improved classification accuracy [179,180]. Various classification techniques, ranging from standard unsupervised and supervised classification methods, such as k-means clustering and maximum likelihood, respectively, or more advanced methods, such as artificial intelligence or support vector machines, have been employed to achieve greater classification accuracies preceding change analysis [181,182]. Various image sources have been assessed including multispectral images with medium or high spatial resolutions and hyperspectral images from airborne or satellite platforms [40,183]. CD using SAR data has also been implemented to detect land alterations in urban and forest areas [182,184].

Numerous strategies to improve accuracy when deriving “from-to” information have been studied through employing the latest classification algorithms or by combining two or more techniques. An Iterative Compound Classification was tested in an agricultural region, and generated higher accuracy compared to post-classification analysis [185]. Although computationally expensive, other research combined change vector analysis, PCA, and temporal signature description to improve change detection accuracy for urban land cover [186]. The combination of classification techniques with PCA or with geostatistical analysis has been implemented to enhance classification accuracy and in turn increase CD accuracy [184,187].
Table 2. A non-exhaustive list of methods, data source, and ecosystem applications of research that generates type or composition of change information.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Data types</th>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-classification</td>
<td>CASI</td>
<td>Invasive species</td>
<td>[133]</td>
</tr>
<tr>
<td>(various algorithms)</td>
<td>Landsat</td>
<td>Land cover</td>
<td>[181]</td>
</tr>
<tr>
<td></td>
<td>Landsat and MODIS</td>
<td>Wetland</td>
<td>[189]</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2 and</td>
<td>Forest</td>
<td>[190]</td>
</tr>
<tr>
<td></td>
<td>Orthophoto-map</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCA and classification</td>
<td>Landsat</td>
<td>Forest</td>
<td>[191]</td>
</tr>
<tr>
<td></td>
<td>Landsat and SPOT</td>
<td>Forest</td>
<td>[192]</td>
</tr>
<tr>
<td></td>
<td>Landsat and Quickbird</td>
<td>Urban</td>
<td>[180]</td>
</tr>
<tr>
<td>Post-classification with other techniques</td>
<td>Quickbird and WorldView</td>
<td>Tree mortality</td>
<td>[179]</td>
</tr>
<tr>
<td></td>
<td>Landsat</td>
<td>Savannah</td>
<td>[187]</td>
</tr>
<tr>
<td></td>
<td>ADAR Hyperspectral</td>
<td>Urban</td>
<td>[40]</td>
</tr>
<tr>
<td>Region-based approaches</td>
<td>SPOT-5</td>
<td>Urbanisation</td>
<td>[193]</td>
</tr>
<tr>
<td></td>
<td>Cartosat-1</td>
<td>Building and forest</td>
<td>[194]</td>
</tr>
<tr>
<td>Regularised level set region-based approach</td>
<td>Landsat</td>
<td>Various test site (urban, watershed)</td>
<td>[175]</td>
</tr>
<tr>
<td>Faster region-based convolutional neural networks</td>
<td>IKONOS – UAV and Gaofen-2 – Pleiades 1B</td>
<td>Agriculture and rural areas</td>
<td>[176]</td>
</tr>
<tr>
<td>Combining pixel- (classification) and region-based approaches</td>
<td>KVR</td>
<td>Mining area</td>
<td>[177]</td>
</tr>
<tr>
<td></td>
<td>Landsat</td>
<td>Land cover (heterogenous area)</td>
<td>[178]</td>
</tr>
</tbody>
</table>

4.3. Magnitude and Direction of Change

Other products of CD analysis are the magnitude and direction of change information, which are typically generated from change vector analysis (CVA) [195]. The products sum the information to be used for differentiating change types, which is a common reason to favor this technique. In general, the direction of change can be identified by employing two bands while the magnitude of change uses more than two spectral bands. A graphical representation of CVA was introduced by Wegmann, et al. [196] and Figure 2 presents an illustration of CVA.

Figure 2. Graphical illustration of change vector analysis.
Three essential prerequisites for CVA include the following: (a) precise registration, (b) preprocessing to reduce noise and stabilize performance, and (c) determining a minimum area of interest to avoid the “salt and pepper” effect [65]. According to Kontoes [197], a change vector can be defined as the angle of change.

Original image spectra or their transformations, such as indices, have been analyzed by using CVA. Transformed data such as TC [48,65,198] or vegetation indices [199] have also been used to undertake CVA. Schoppmann and Tyler [48] concluded that performing CVA on transformed images allows change process inference, particularly when obtaining ground truth information is challenging. By employing different types of data, including vegetation indices and surface reflectance, Lambin and Strahler [200] highlighted that different indicators identified different change processes. Combined application of raw spectral and transformed data was demonstrated for investigating a mountainous area by Nordberg and Evertson [158], who formulated the direction of change from the combination of Bands 3 and 4 of Landsat data, while the magnitude was determined from the three components of the TC transformation, i.e., greenness, wetness, and brightness.

CVA best suits CD within land cover classes; hence, CVA alone cannot provide the desirable “from-to” information. Johnson and Kasischke [65] highlighted the limitation of CVA for change labeling as follows: Change vectors describe change dynamics, but not states (i.e., class membership at the start and end of vectors). Defining thresholds to delineate changed from unchanged pixels is another challenge for CVA [201], particularly when dealing with multispectral images. Multispectral images are likely to be heterogeneous in their spectral bands, thus, applying a single threshold could be insufficient [202]. This issue can be overcome by normalizing the data and transforming the result into a posterior probability [201].

CVA can be used to compare the difference of biophysical indicators along a temporal trajectory, from the combination of the length of the vector, representing different magnitudes, and change direction, which indicates the nature of the change [203]. CVA is a multivariate technique since it can analyze at least two spectra. The direction of change can be used to classify change types at 2n change directions, where n equals the number of spectral bands [65], while the magnitude of change can form 3n+2 feature spaces to filter meaningful change areas [197]. CVA is most advantageous when spectral changes are unknown or when spectral variabilities are difficult to comprehend [65]. An improved version of CVA, i.e., one based on deep learning, is capable of spatially modeling correlated pixels, and has been tested for CD with three VHR images, i.e., Worldview, Quickbird, and Pleiades [204].

Various investigations have employed CVA including, but not limited to, forest areas, post-hazard change phenomena, and urban expansion. Table 3 describes examples of studies of diverse ecosystems that have applied CVA as a single technique or in combination with other approaches.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Data</th>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVA on image spectra or TC</td>
<td>Landsat</td>
<td>Forest and land cover</td>
<td>[195,198,205]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Polluted areas</td>
<td>[48]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arid vegetation</td>
<td>[206]</td>
</tr>
<tr>
<td></td>
<td>Landsat, Sentinel-1, Sentinel-2</td>
<td>Forest</td>
<td>[207]</td>
</tr>
<tr>
<td>CVA with various classification algorithms</td>
<td>Landsat</td>
<td>Wetland</td>
<td>[208]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Land cover</td>
<td>[197,202]</td>
</tr>
<tr>
<td></td>
<td>AVIRIS</td>
<td>Conserved areas</td>
<td>[49]</td>
</tr>
<tr>
<td>CVA with multivariate analysis</td>
<td>AVHRR</td>
<td>Desert</td>
<td>[203]</td>
</tr>
<tr>
<td></td>
<td>MODIS</td>
<td>Desert, city</td>
<td>[199,209]</td>
</tr>
<tr>
<td>CVA with level set evolution or continuity rules</td>
<td>Landsat</td>
<td>Mountainous area</td>
<td>[210,211]</td>
</tr>
<tr>
<td>CVA based on deep learning</td>
<td>Worldview, Quickbird, Pleiades</td>
<td>Urban</td>
<td>[204]</td>
</tr>
</tbody>
</table>

Table 3. Research that has applied change vector analysis (CVA) as a single technique or in combination with other methods to produce magnitude and direction of change information for various ecosystems.
4.4. The Probability of Change

Change probability is an estimated value representing the likelihood of change in an area and is usually generated from temporal analysis of historical datasets. The identification of the probability of change enables forecasting future conditions, which is important for planning-related research.

The change probability has generally been produced using different statistical techniques following, for example, Chi-square, logistic, or Gaussian distributions. Shi and Ding [212], for instance, employed the Chi-square distribution to estimate the probability of change. The probability was derived by utilizing Bayesian probability from temporally stacked images or from binary images of differing acquisitions [213]. Examining two observations, a modified form of CVA was used to generate change probabilities [201]. They called this technique CVA in posterior probability space (CVAPS). It has the advantage of scaling the change magnitudes to a consistent range (that of posterior probabilities, zero to one), and thus requires only one threshold across classes, making optimizing the threshold more efficient as compared with CVA, which needs one threshold per class.

Estimating change probability has been implemented for multiple applications. A likelihood ratio has been used to identify the probability of flood-affected areas and vehicle distributions by using ERS-2 SAR and airborne SAR, respectively [214]. The change probability of biologically complex ecosystems was estimated, based on the proportional area of changed and unchanged lands by using multiband image difference (MID) [154]. Probability of urbanization was identified by using a Bayesian P-spline to highlight changes across time and location [215], while Wang, et al. [216] employed the scale invariant features transform (SIFT) method to calculate the probability of changed land. Finally, coherent change detection was implemented to map the probability of volcanic ash distribution [217] and the probability of damage due to natural disaster [218].

Some researchers have demonstrated the use of change probabilities while identifying driving forces of change. Employing various regression techniques, the likelihood of change can be estimated by considering the contribution of determinant factors from multiple variables. An example of this employed generalized linear models (GLMs) and semi-variograms to judge the importance of change metrics and produced probability of change and variability of change images [219]. Logistic regression is a frequently used technique for determining driving forces of change; for instance, the identification of change determinants while estimating change likelihood of a pixel [220], the determination of reforestation probability [221], and quantification of forest degradation probability based on biophysical and socioeconomic datasets [222].

The probability of change can also inform error detection rates including the error and false alarm probabilities [223]. Despite being able to highlight the most probable locations of false alarms, local estimates of error probability were not calculated in that article. Other research by Cartus, et al. [224] investigated the change detection error probability during an investigation of forest cover loss by using L-band backscatter.

Integrating evidential information can improve the accuracy of change information or likelihood of CD. The Dempster–Shafer theory is generally the basis of evidential fusion [225], which was implemented to integrate images from several optical satellites [226,227], airborne laser scanners [228], and SAR images or a combination of optical and SAR images [229]. A modified version of Dempster–Shafer, the dynamic evidential reasoning method, was proposed for sequential image fusion to deal with the changing frame of different images [230]. The theory has been implemented for studies with a range of aims, for example, to update a building map [229], to detect landslides [231], and to detect newly built-up areas [232].

4.5. Temporal Change Trajectories

A land use and cover change (LUCC) trajectory can be defined as the periodically observed temporal sequence of a pixel’s land cover that is integrated in serial datasets [111]. The LUCC trajectory change detection (TCD) can also be seen as a temporal development curve that could be generated from multitemporal images [203]. The use of multitemporal datasets improves sensitivity to seasonal,
gradual, or subtle changes. Then, trends can be identified from the sequence of change to highlight the dynamic character of change [233,234].

The availability of freely accessible image archives and the insensitivity of bitemporal analysis to describe the dynamics of sub-annual changes have driven advances in trajectory change analysis. Space agencies have administered programs to support monitoring global land cover by providing free data from NASA and USGS’s Landsat and through the ESA’s Sentinel program. Freely accessible data such as AVHRR, MODIS, SPOT, and PROBA accelerated the development of techniques and algorithms, as well as applications that allow further investigation of global temporal LUCC dynamics [235].

Change trajectories can be analyzed from either periodical or nonperiodical datasets. Frequent periodical observations have distinctive advantages, such as the capability to indicate either temporary or permanent change processes and the start and end of growing seasons [236,237]. This sort of information is crucial for monitoring management practices, to investigate the intensity of land use or climatic conditions, or changes in the way that land is being used. The use of remote sensing indices for TCD was introduced by Lambin and Strahler [203], who employed vegetation indices and land surface temperatures. Temporal trajectories of indices indicate the cause of change processes and record alterations over time [4]. Temporal patterns can also be used to estimate disturbances [238]. Trends assist in understanding influential factors of LUCC [239] or understanding seasonal dynamics [240], which cannot be portrayed in bitemporal observations. In most cases, the study of temporal trajectories has employed statistical analysis, particularly, time series analysis.

Descriptive or statistical time series analysis (TSA) can be implemented for TCD. Descriptive analyses employ any number of observations while statistical TSA strictly requires frequent, periodical, observable facts [241]. Indeed, there are various interpretations of TSA applied in remote sensing. For instance, according to Giri, et al. [242], TSA refers to multitemporal analysis of images to represent two seasons, i.e. summer and harvest, whereas for Jonsson and Eklundh [237], TSA implies a statistical time series approach to studying seasonal dynamics. The use of statistical time series in remote sensing aids in understanding general trends and seasonal dynamics of change from sub-annual observations. The following sections discuss techniques focusing on change detection particularly on trend and seasonal dynamics.

4.5.1. Temporal Trend

Trend is a parameter that can be derived either from temporally periodical or nonperiodical observations. Trends from nonperiodical observations are generally obtained from aggregate information such as annual or any periodically aggregated observations. In land-cover-related research, a change in trend can indicate disturbances [243]. Curve fitting is an alternative to nonperiodical techniques to analyze trends. A hypothetical curve pattern is often used to evaluate the trend of LCC phenomena [233,244]. Hüttich, et al. [245] employed annual data derived from SPOT mosaics to investigate the trend of vegetation green-up in the boreal forest zone, while Zhou [246] quantified trajectory changes using landscape metrics. Common statistical techniques such as regression, can be utilized to estimate trend parameters. Xue, et al. [247] implemented curve fitting to investigate the trajectory of change before ingesting images into classification algorithms.

Trend parameters can be gathered from seasonal observations through deseasonalizing data or by employing stationary or nonstationary decomposition techniques [243,248,249]. The assumption of stationary temporal patterns can limit implementing these techniques to phenomena with particular temporal change characteristics. Abrupt changes that drastically deviate from a temporal curve of biophysical properties, such as those induced by natural events such as volcanic explosions, floods or fires, do not meet the stationarity criterion. Table 4 summarizes the major decomposition techniques used to detect LCC and provides a short description about the techniques, sorted from older to newer applications in remote sensing-related research.
Table 4. Major time series decomposition techniques for LCC-related studies.

<table>
<thead>
<tr>
<th>Method</th>
<th>Abbreviation</th>
<th>Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seasonal Time Series Analysis based on loess regression</td>
<td>STL</td>
<td>Decomposing time series of spectral bands into three components, i.e., trend, seasonal and remainder.</td>
<td>[250,251]</td>
</tr>
<tr>
<td>Fourier transform</td>
<td>FT</td>
<td>Decomposing time series of spectral bands by using superposition of sine and cosine functions.</td>
<td>[252,253]</td>
</tr>
<tr>
<td>Singular Spectrum Analysis</td>
<td>SSA</td>
<td>Decomposing time series of spectral bands into different oscillatory components without fitting.</td>
<td>[254,255]</td>
</tr>
<tr>
<td>Wavelet transform</td>
<td>WT</td>
<td>Decomposing time series into linear combinations of wavelet function.</td>
<td>[256,257]</td>
</tr>
<tr>
<td>Seasonal trend analysis</td>
<td>STA</td>
<td>Harmonic analysis to extract annual and semi-annual time series harmonics.</td>
<td>[258]</td>
</tr>
<tr>
<td>Landsat-based detection of trends in disturbances and recovery</td>
<td>LandTrendR</td>
<td>Extracting spectral trajectories of land surface change from yearly Landsat time series stacks.</td>
<td>[259]</td>
</tr>
<tr>
<td>Break for Additive Season and Trend</td>
<td>BFAST</td>
<td>Integrating decomposition of time series data with the capacity to detect change within the time series.</td>
<td>[243]</td>
</tr>
<tr>
<td>Auto-regressive Integrated Moving Average</td>
<td>ARIMA</td>
<td>Univariate time series that may transform nonstationary data series into stationary models through differencing.</td>
<td>[260,261]</td>
</tr>
<tr>
<td>Seasonal Autoregressive Integrated Moving Average</td>
<td>SARIMA</td>
<td>Similar model to ARIMA but involving backshift of seasonal period.</td>
<td>[262]</td>
</tr>
<tr>
<td>X12-Autoregressive Integrated Moving Average</td>
<td>X12-ARIMA</td>
<td>Seasonal adjustment of univariate time series models that decomposes spectral bands into seasonal-trend-irregular components.</td>
<td>[263–265]</td>
</tr>
<tr>
<td>Ensemble Empirical Mode Decomposition</td>
<td>EEMD</td>
<td>Decomposing satellite image time series into intrinsic mode functions (IMFs) and selecting relevant IMFs to separate trend and seasonal components.</td>
<td>[266–268]</td>
</tr>
<tr>
<td>–Detecting Breakpoints and estimating segment in trend</td>
<td>DBEST</td>
<td>A segmentation algorithm that simplifies the trend into linear segments employing three parameters including a generalisation threshold, the largest changes and the magnitude of change.</td>
<td>[269]</td>
</tr>
<tr>
<td>Sub-annual Change Detection</td>
<td>SCD</td>
<td>Detecting abrupt land cover change from sub-annual scale.</td>
<td>[270]</td>
</tr>
<tr>
<td>Change, After Effect and Trend</td>
<td>CAT</td>
<td>Suppressing seasonal effects to identify change, after effect, and trend.</td>
<td>[271]</td>
</tr>
</tbody>
</table>

Break for additive season and trend (BFAST) is the most widely employed technique for remote sensing time series analysis amongst the decomposition techniques. BFAST is a modified technique for seasonal time series analysis based on loess regression, and was introduced by Cleveland, Cleveland and Terpenning [250]. BFAST combines a harmonic seasonal model and ordinary least squares analysis based on a moving sum (OLS-MOSUM) test which can detect the breakpoints occurring within a serial dataset through an additive decomposition model [243].

Table 5 presents the implementation of BFAST for different remote sensing-based analyses. The table shows that various periodical observations have been applied in different studies. Most of these studies employed MODIS and Landsat data, while a few studies used SPOT or other platforms. The studies mostly employed vegetation indices, such as NDVI and Enhanced Vegetation Index (EVI). Moreover, the majority of BFAST implementations were directed to understanding deforestation and forest disturbances.
Table 5. Varying image revisit times and observation targets implementing BFAST.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Periodical/Number of Observations (Scenes)</th>
<th>Observation Targets</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat</td>
<td>16-day</td>
<td>Combining time series data for LUCC</td>
<td>[272]</td>
</tr>
<tr>
<td></td>
<td>Irregular/55</td>
<td>Small-scale forest disturbances</td>
<td>[273]</td>
</tr>
<tr>
<td></td>
<td>Irregular/270</td>
<td>Forest disturbances</td>
<td>[274]</td>
</tr>
<tr>
<td></td>
<td>Irregular/(135, 104, 94)</td>
<td>Error generation of forest change</td>
<td>[275]</td>
</tr>
<tr>
<td></td>
<td>Irregular/94</td>
<td>Forest disturbances</td>
<td>[276]</td>
</tr>
<tr>
<td>MODIS vegetation</td>
<td>16-day</td>
<td>Exploring change within near observed data/deforestation</td>
<td>[277,278]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Response of vegetation to flood and fire</td>
<td>[279]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phenology-based classification</td>
<td>[247]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Characterising vegetation dynamics</td>
<td>[280,281]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The impact of confining elephants in a forest on land use change</td>
<td>[282]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Biomass estimation</td>
<td>[283]</td>
</tr>
<tr>
<td>MODIS reflectance</td>
<td>8-day</td>
<td>Identifying LCC that corresponded to urban expansion</td>
<td>[284]</td>
</tr>
<tr>
<td>SPOT vegetation</td>
<td>10-day</td>
<td>Estimating water balance variables</td>
<td>[285]</td>
</tr>
<tr>
<td>Sentinel-2</td>
<td>5-day</td>
<td>Monitoring agricultural parcels</td>
<td>[286]</td>
</tr>
<tr>
<td>The combination of several images</td>
<td>unstated</td>
<td>Fire-induced changes</td>
<td>[223]</td>
</tr>
<tr>
<td>Simulated data</td>
<td>unstated</td>
<td>Comparing time series models</td>
<td>[288]</td>
</tr>
</tbody>
</table>

4.5.2. Seasonal Pattern

In general, time series methods decompose a signal into three components, i.e., trend, seasonal, and irregular, to reveal several change phenomena. Seasonal trend decomposition based on loess (STL) and seasonal trend analysis (STA) are amongst the decomposition techniques that have been implemented in remote sensing studies [250]. Many investigations of temporal trajectories focus on detecting seasonal patterns of change dynamics [237,243,258]. Since indices serve as the primary data source, selection of proper indices influences the capabilities of STA.

The incorporation of a seasonality adjustment should be considered, since most time series remote sensing applications employ reflectance or indices that are affected by season. Seasonality is defined as the amplitude of the temporal profile of indices to indicate photosynthetic activity of vegetation [289]. Seasonality metrics can be estimated by using statistical time series such as the beginning and end of seasons, the number of growing seasons, or the rates of growth and decline [237]. The metrics have been employed to discriminate the phenology of different vegetation types and the intensity of agricultural cultivation within a year.

STL is capable of differentiating areas with annual vegetation from those with perennial vegetation [251]. The idea was first introduced by Cleveland, Cleveland and Terpenning [250] and was applied to understand air quality based on remote sensing observations [250] and to map seasonal variation of vegetation [290]. A later version of STL, namely STA, uses harmonic equations capable of extracting annual and semi-annual harmonics [258]. Carrao, et al. [291] showed that STA could be used to estimate the near future of earth-surface characteristics, such as vegetation greenness or temperature. The combination of STA to extract seasonal parameters and PCA to contrast changed and unchanged places was found to be beneficial for characterizing many types of land cover transition [292]. A further extension of STA, seasonal trend analysis and segmentation to detect land cover (SDTC), allows multiple variables to be combined [293]. The significance of temporal trend was tested by using
the Mann–Kendall procedure and it appeared to be valuable in detecting and characterizing land cover change.

TIMESAT is a popular tool for investigating seasonal patterns. With optional distributions of asymmetric Gaussian, Savitsky–Golay, double logistic function, or nonlinear least square fits, this tool was used to define seasonality parameters, including intensity of growth, the beginning and end of a season, and growth rate [237]. The tool provides smoothed, meaningful data derived from remote sensing images [294]; applications of this tool are summarized in Table 6. In general, the table indicates a variety of data types, temporal revisit intervals, and objectives or ecosystem applications that have applied TIMESAT. The data sources include medium to coarse spatial resolution optical images, such as AVHRR, MODIS, MERIS, SPOT vegetation, and PROBA, as well as coarse microwave images such as AMSR-E. Combining images has also been used to improve the interpretation of seasonal changes and their drivers.

**Table 6. Remote sensing investigations employing TIMESAT.**

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Temporal Revisit</th>
<th>Objectives</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVHRR</td>
<td>10-day</td>
<td>Seasonality metrics extraction</td>
<td>[237]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The trend of VI</td>
<td>[295]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phenological change</td>
<td>[296,297]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exploring anthropogenic and climate effects on vegetation</td>
<td>[298]</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>Phenological metrics and productivity extraction</td>
<td>[299]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assessing land degradation/recovery</td>
<td>[300]</td>
</tr>
<tr>
<td>MODIS Vegetation Index</td>
<td>16-day</td>
<td>The extraction of phenological metrics for:</td>
<td>[301]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- rice crops</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- forest</td>
<td>[302]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- vegetation affected by insect infestation</td>
<td>[303]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- abandoned agriculture</td>
<td>[304]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- complex ecosystems</td>
<td>[305]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- spring ephemerals</td>
<td>[306]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mapping seasonality metrics</td>
<td>[307]</td>
</tr>
<tr>
<td>MODIS Surface Reflectance</td>
<td>8-day</td>
<td>Improving ecosystem classification</td>
<td>[302]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mapping fractional forest cover</td>
<td>[308]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tracking seasonal change in coniferous forest</td>
<td>[309]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monitoring cotton stages</td>
<td>[310]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delineating vegetation phenology</td>
<td>[311]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ecosystem assessment</td>
<td>[312]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validating spring green-up</td>
<td>[313]</td>
</tr>
<tr>
<td>MODIS LAI</td>
<td>8-day</td>
<td>Smoothing and gap-filling data</td>
<td>[314]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modelling land surface and climate</td>
<td>[315]</td>
</tr>
<tr>
<td></td>
<td>10-day</td>
<td>Estimating start of season (SOS)</td>
<td>[316]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extracting phenological information for deciduous forest</td>
<td>[317]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Estimating SOS and ecoregion variability</td>
<td>[318]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Estimating crop phenology</td>
<td>[319]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mapping peanut cultivation</td>
<td>[320]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Scheduling irrigation</td>
<td>[321]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mapping macrophyte phenology at a lake</td>
<td>[322]</td>
</tr>
<tr>
<td></td>
<td>8-day</td>
<td>Studying deciduous forest carbon flux</td>
<td>[323]</td>
</tr>
<tr>
<td></td>
<td>16-day</td>
<td>Impact of illumination on seasonal metrics</td>
<td>[324]</td>
</tr>
<tr>
<td></td>
<td>5-day</td>
<td>Exploring macrophyte seasonal dynamics</td>
<td>[322]</td>
</tr>
</tbody>
</table>
4.6. Dynamic Simulation of Changes

Simulation of future land cover changes has been an interest of many remote sensing users for land and ecosystem monitoring, management, and planning. The topic has been discussed widely, often by integrating the concepts of remote sensing and GIS or spatial analysis [325]. Temporal interpolation is usually employed in a simulation, whereas spatial interpolation has been applied with stricter assumptions. The integration of temporal and spatial models allows the understanding of complex systems and their spatiotemporal dynamics [326].

Various techniques for dynamic simulation have been developed in the areas of geoscience and ecosystem modeling. The Conversion of Land Use and its Effect (CLUE) model [327] and cellular automata principles [328], have been implemented for dynamic modeling of various ecosystems contexts. In another case, Bacani, et al. [329] combined a Markov chain model and cellular automata to generate spatiotemporal change probability maps for wetlands. A supervised classification using a multilayer perceptron (MLP), a type of Markov chain model, was implemented to predict land cover change in a semiarid area [330]. Markov chain models, however, have been criticized regarding their ability to detect errors since the drivers of change are omitted. This drawback can be minimized by employing binary logistic regression to determine the drivers of change for future simulations by artificial neural networks (ANN) [331]. Comparing parametric and non-parametric models for simulating land use change, Tayyebi, et al. [332] demonstrated that ANNs outperformed two other non-parametric models, classification and regression tree (CART) and multivariate adaptive regression splines (MARS), to estimate short time intervals of coarse spatial resolution data.

5. Accuracy Assessment of Change Detection Techniques

Accuracy assessment is required for affirming reliable information and for evaluating the robustness of classification or segmentation techniques. Common accuracy assessment is based on an error matrix or confusion table [333], which identifies overall accuracy, false positive/commission errors, and false negative/omission errors. A few references described false positives as false alarms and false negatives as missed alarms/detections [157,334]. Some approaches have also been proposed to assess change detection accuracy, such as the change detection error matrix [169], area-based accuracy [335], accuracy assessment curve [336], and rule-based rationality [333]. The manageability of the error matrix was questioned when dealing with more than two images for change analyses, therefore Li and Zhou [337] proposed a trajectory error matrix for change analysis involving more than two images. Since not all of the assessments have been examined in various settings, the robustness of the techniques within differing environments needs to be investigated in the future.

Accuracy has been reported to be the result of many factors, including image properties and the seasonal condition affecting the quality of images [338], preprocessing being involved for aligning or correcting the quality of the images [60]; correlation among spectral bands and among image pairs; sampling strategies [339]; and techniques for generating change information. When classification or segmentation techniques were involved, the accuracy of change detection depended upon single date accuracies and was commonly calculated as the product of these [340]. Nonetheless, Congalton and Green [341] argued that correlation among layers contributed to the accuracy, where the sum product would be relevant for independent classified layers, whereas correlated layers would relate to single-date accuracy.

6. Challenges, Comparisons, and Future Directions for CD Techniques

The primary challenge for applying bitemporal techniques lies in the decision strategy used to separate changed from unchanged land. A common strategy is to employ the mean and standard deviation of features to determine thresholds for differencing, ratioing, and PCA with various ranges of constants [126,154,342–344]. Meanwhile, a hypothetical range was applied for thresholding by determining 10% of the proportion as change [129], employing Bayes theory [345,346], or the Wishart
distribution [347] for automatic thresholding. It appears that advances in binary change detection go beyond simple algebra techniques. The integration of relative radiometric corrections in a simultaneous manner, such as in MAD, offers a robust yet efficient procedure to generate binary change maps. Similarly, ISFA offers normalization while applying unsupervised change detection. Meanwhile, various deep-learning-based techniques have recently been developed for unsupervised change detection and tested in several settings. Nonetheless, the effectiveness of these methods for monitoring regional or global extent needs to be explored.

A different challenge applies for post-classification change analysis. The selection of suitable classifiers is essential to optimize accuracy, meanwhile, particular sampling strategies such as cross-validation or tuning parameters can improve the performance of classifiers [348]. Various strategies have been proposed for cross-validation for simultaneous change detection, such as by using a generative adversarial network to optimize change detection based on min-max game theory [349] or using a learning algorithm called iterative active learning to define multitemporal training sets [350]. Subtle changes related to spatial resampling, land management, and vegetation species or vegetation growth stages are understudied. However, a few such examples were found, for instance, a combination between IRMAD and random forest classification of amended SAR images with polarimetric decomposition layers to detect subtle change related to vegetation growth stages in cloud-prone regions [351].

Employing multiple sensors and spatial resolutions for long temporal analysis is the consequence of sensors’ limited lifetimes. The consequence is well known; yet, a paucity of research related to sensor compatibility is evident in the literature. Meanwhile, trajectory change detection employing time series datasets needs to adapt to different properties of datasets. The exploration of techniques for relative radiometric corrections becomes more important for applications employing massive and temporally continuous observations.

Detecting changes at the subpixel level can now be undertaken, allowing the investigation of detailed aspects of land cover when access to high resolution imagery is limited. Employing subpixel techniques outperformed traditional techniques in terms of analyzing images with different spatial resolutions [213] or inspecting subpixel image features [352].

The development of techniques, algorithms, and tools has facilitated the investigation of change for bitemporal and trajectory analyses, and simulation of future changes. R-based algorithms enable such analyses using a number of libraries for change detection, ranging from preprocessing such as geometric, radiometric, and atmospheric corrections; bitemporal change detection methods such as CVA; to seasonal time series analysis such as BFAST, DBEST, STA, STL, ARIMA, and X12-ARIMA [353]. Tools implemented in software packages other than R are also available from respected researchers, for instance, the IDL-based tools such as IRMAD for bitemporal CD [354] or TiSeG for MODIS quality assessment [44,355], and TIMESAT for seasonal time series analysis based on Fortran/MATLAB [294]. Meanwhile, more available image archives and the recent Copernicus program are likely to motivate more participation in and development of tools for global monitoring of land cover changes.

In addition to the development of novel algorithms, the combination of available techniques is being considered to improve CD accuracy. Silveira, Mello, Acerbi Júnior and Carvalho [187], for instance, combined differencing and geo-statistical parameters for segmentation preceding CD. Differencing, masking, and multivariate clustering were combined to detect tree mortality in mixed-species woodlands [179]. An improved Kappa has been obtained by combining CVA, PCA, and temporal signature description [186]. Moreover, an advanced version of CVA, called compressed CVA, can perform change probability estimation [356]. Another strategy that combines multiple sources of evidence by employing Dempster–Shafer Theory could be an option for analyzing changes utilizing various images [227].

Characterization and attribution of changes have been implemented to improve interpretation of change processes [357]. Spectral mixture analysis (SMA) is a model to separate a variable
into wavelength-independent and -dependent components [358]. The model has been applied to discriminate soil, rock, shade, and the secondary illumination effects using Viking Lander imagery [358].

To sum up, various land cover change processes require suitable techniques for synoptic yet seamless monitoring. Remote sensing is a fast-moving field, and new techniques keep emerging and these topics are a key area for global monitoring. Considering the need to understand interrelations among ecosystem components, understanding complex models is undoubtedly be even more relevant in the future. Beyond 2D change detection, the 3D detections will also be valuable. Contextual research on land cover change and its effects should include drivers of change, and the prediction of location and quantity of changes [359]. The use of ancillary data from various sources, instead of merely relying on remotely sensed data, has been practical for assisting the interpretation of complex change processes and indicating relevant stressors. It appears that integrating several methods and datasets should become the standard of analysis in the future.

The acceleration of change globally driven by naturogenic and anthropogenic factors creates greater variability of change processes. Hence, bitemporal, multitemporal, and time series techniques are needed to enable the investigation of heterogenous change types, intensities, and process durations, as well as to suit the diverse purposes of studies. The era of freely accessible data in parallel with the growth of non-proprietary toolboxes should propagate doubly through remote sensing communities, as well as users.

**Author Contributions:** D.R.P. conceptualized and designed the search of references and structured literature collections into the original draft; D.J.P. and A.L.G. assisted by commenting on the manuscript and providing editorial oversight. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Australian Department of Foreign Affairs and Trade, Australia Awards Scholarship. Thanks to MDPI for APC’s discount to returning authors.

**Acknowledgments:** The authors would like to thank the Australian Department of Foreign Affairs and Trade for the opportunity to pursue a PhD research program. Special thanks to Bambang Trisasongko for the discussion. We are grateful to three anonymous reviewers for their constructive comments leading to an improved manuscript.

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

**References**


44. Panuju, D.R.; Paull, D.J.; Griffith, J.J.; Paull, D.J.; Paull, D.J.; Griffith, J.J.; Paull, D.J.; Griffith, J.J. Spatio-temporal quality distribution of MODIS vegetation indices 5 and 6: Implications for forest-non-forest separability. *Int. J. Remote Sens.* 2020, 41, 373–397. [CrossRef]
51. Song, D.X.; Huang, C.; Sexton, J.O.; Channan, S.; Feng, M.; Townshend, J.R. Use of Landsat and CORONA data for mapping forest cover change from the mid-1960s to 2000s: Case studies from the eastern United States and central Brazil. *ISPRS J. Photogramm. Remote Sens.* 2014, 103, 81–92. [CrossRef]
56. Nuth, C.; Kääb, A. Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change. *Cryosphere* 2011, 5, 271–290. [CrossRef]


64. Teillet, P. Image correction for radiometric effects in remote sensing. *Int. J. Remote Sens.* 1986, 7, 1637–1651. [CrossRef]


92. Lobser, S.E.; Cohen, W.B. MODIS tasselled cap: Land cover characteristics expressed through transformed MODIS data. *Int. J. Remote Sens.* 2007, 28, 5079–5101. [CrossRef]


158. Nordberg, M.L.; Evertson, J. Vegetation index differencing and linear regression for change detection in a Swedish mountain range using Landsat TM® and ETM+® imagery. Land Degrad. Dev. 2005, 16, 139–149. [CrossRef]
164. Fung, T. Land use and land cover change detection with landsat MSS and SPOT HRV data in Hongkong. Geocarto Int. 1992, 7, 33–40. [CrossRef]
165. Bernardo, N.; Watanabe, F.; Rodrigues, T.; Alcântara, E. An investigation into the effectiveness of relative and absolute atmospheric correction for retrieval the TSM concentration in inland waters. Model. Earth Syst. Environ. 2016, 2, 114. [CrossRef]


194. He, C.Y.; Zhao, Y.Y.; Tian, J.; Shi, P.J.; Huang, Q.X. Improving change vector analysis by cross-correlogram spectral matching for accurate detection of land-cover conversion. *Int. J. Remote Sens.* **2013**, *34*, 1127–1145. [CrossRef]

195. Lambin, E.F.; Strahler, A.H. Indicators of land-cover change for change-vector analysis in multitemporal space at coarse spatial scales. *Int. J. Remote Sens.* **1994**, *15*, 2099–2119. [CrossRef]

211. Savage, L.S.; Lawrence, L.R.; Squires, R.J.; Holbrook, D.J.; Olson, E.L.; Braaten, D.J.; Cohen, B.W. Shifts in drivers of forest degradation in tropical dry forests: A case study in Western Mexico. Land Use Policy 2014, 36, 2703–2721. [CrossRef]


236. Feng, H.; Liu, Y. Trajectory based detection of forest-change impacts on surface soil moisture at a basin scale [Poyang Lake Basin, China]. J. Hydrol. 2014, 514, 337–346. [CrossRef]


244. Lupo, F.; Linderman, M.; Vanacker, V.; Bartholome, E.; Lambin, E.F. Categorization of land-cover change processes based on phenological indicators extracted from time series of vegetation index data. Int. J. Remote Sens. 2007, 28, 2469–2483. [CrossRef]


284. Tsutsumida, N.; Saizen, I.; Matsuoka, M.; Ishii, R. Land cover change detection in Ulaanbaatar using the breaks for additive seasonal and trend method. *Land* 2013, 2, 534–549. [CrossRef]


357. Alaibakhsh, M.; Emelyanova, I.; Barron, O.; Mohyeddin, A.; Khadiani, M. Multivariate detection and attribution of land-cover changes in the Central Pilbara, Western Australia. Int. J. Remote Sens. 2015, 36, 2599–2621. [CrossRef]


© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).