Enabling Regular Map Updates and Identification of Impervious Surfaces Through Satellite Data Fusion, Machine Learning and Cloud Platforms

Melchior Vitalis Shukuru[1]*, Stephen McCarron[1], Guillermo Castro Camba[2], Conor Cahalane[1]

[1] Maynooth University, Department of Geography, Maynooth, Co. Kildare, Ireland
[2] Tailte Éireann, Irish Life Centre, Abbey Street Lower, Co. Dublin, Ireland
*Corresponding author: melchior.shukuru.2025@mu.ie

Keywords: Data Fusion, Machine Learning, Copernicus Data, Change Detection, Remote Sensing, Feature Importance.

Abstract

Frequent cloud cover is a common impediment deterring many countries from employing optical earth observation data for the purposes of national map updates. A decision-level data fusion approach allows the inclusion of SAR satellite imagery in such locations and therefore has potential to assist in this task. In this study we test a combination of cloud penetrating Sentinel-1 and multispectral Sentinel-2 to enhance the delineation of impervious surfaces from other land cover types, impervious surfaces being a key component of hydro-climatological models in urban and semi-urbanised areas. Using machine learning techniques and leveraging the Copernicus archive in the Google Earth Engine (GEE) platform, a post-classification change detection approach was explored to assess impervious surface expansion between 2017 and 2023 across the urban centre of Dublin, Ireland. Image classification, conducted using a random forest classifier, achieved overall accuracies of 93% and 91% and kappa coefficients of 0.91 and 0.89 for 2017 and 2023 data, respectively for key classes. The potential of multispectral and RADAR indices such as NDVI, NDBI and PRISI was tested and proved generally effective, but showed limitations in areas adjacent to the coast and inland water bodies, with indications of confusion between land cover types. The inclusion of NDWI in the fused dataset was shown to help differentiate waterbodies from impervious surfaces, highlighting the importance of integrating a water-specific index. NDVI outperformed other indices in feature importance, though PRISI was shown to helpfully cluster impervious surfaces.

1. Introduction

The expansion of impervious surface (IS) extents in urban settings has significant environmental impacts, making it highly relevant to public authorities and environmental agencies reliant on spatial data from National Mapping Agencies (NMAs) for various urban or infrastructure studies. Direct impacts of IS expansion include increases in urban heat island (UHI) effects, increased and accelerated runoff, water quality degradation and the loss of biodiversity (Tian *et al.*, 2018). Furthermore, IS negatively affects carbon storage, sequestration, and the hydrological properties of sub-IS sediment (Zhang *et al.*, 2020).

Remote sensing, through multi-source and multitemporal satellite data, has been effectively used to estimate and monitor changes in similar landcover classifications at different scales in an urban context (Zhang *et al.*, 2020). In terms of open-source satellite data, Sentinel-2 (S-2) imagery is becoming more commonly used for landcover classification by NMAs in rural areas (Lydon *et al.*, 2018). However, for many urban areas, Sentinel imagery may not be sufficient to identify all landcover changes, primarily due to a comparatively lower spatial resolution (10m for most S-2 spectral bands – whereas NMA orthophotography is generally sub-metre). Some urban features such as narrow roads (less than 10 m width) and small buildings of less than 100 square meters may be misclassified (Kuc and Chormański, 2019). Previous studies have attempted to improve performance by employing methods such as the Normalized Difference Built-up Index (NDBI) for IS mapping using both the shortwave infrared (SWIR) and near-infrared (NIR) bands to better distinguish urban areas from other land cover types (Osgouei *et al.*, 2019; Kebede, Hailu and Suryabhagavan, 2022). The Normalized Difference Vegetation Index (NDVI) can categorise IS where index values between 0.2 and -1.0 can imply a built-up environment (Samadi Todar, Attarchi and Osati, 2021).

The suitability of S-2 imagery as a tool for NMA programs tasked with regular map updates on fixed schedules is therefore still an active research question and further exacerbated by the acknowledged weakness of optical satellites to image the earth surface during periods of cloud cover. Land cover types such as bare soils, water and dry vegetation also exhibit similar spectral reflectance signatures in certain bands, leading to misclassification (Kebede, Hailu and Suryabhagavan, 2022). Optical indices such as the NDVI and NDBI also lack pre-defined thresholds for the classification of specific IS and land cover types that are transferrable to all countries, and incorporating the S-2 SWIR band with a spatial resolution of 20m requires resampling to ensure compatibility with the NIR band (10m) during computation of ratios such as the NDBI (Osgouei *et al.*, 2019). However, S-2 imagery for IS identification has traditionally produced higher accuracy results than Sentinel-1 (S-1) data, primarily due to its broader spectral coverage (Samadi Todar, Attarchi and Osati, 2021). S-1 synthetic aperture RADAR (SAR) imagery does however offer the opportunity to increase the temporal frequency of data collection and analyses, potentially reducing the cycles required for full NMA mapping revisions, although its suitability for mapping IS has not been fully explored (Szigarski *et al.*, 2018; Ding *et al.*, 2022).

Single and dual-polarization S-1 SAR data, specifically the vertical transmit/vertical receive (VV) and the vertical transmit/horizontal receive (VH) data in Ground Range Detected (GRD) format (calibrated, orthorectified and stored in decibel values) allows the mapping of IS through an analysis of surface scattering (Shao *et al.*, 2016; Liu *et al.*, 2023). Due to its stronger response in urban areas, VV backscatter values ranging between -5 and -10 dB are generally considered more suitable for the identification of IS relative to VH data (Liu *et al.*, 2023). One recently proposed index that demonstrates potential for mapping IS is the Piecewise Radar Impervious Surface Index (PRISI). PRISI combines the VV and VH polarisations in a thresholding approach to better distinguish IS from other land cover types

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-345-2026

(Ding *et al.*, 2022). This method relies primarily on the relatively strong scattering behaviour of urban surfaces, which results in relatively dominant VV returns compared to VH values (Liu *et al.*, 2023). It is ‘piecewise’ because it makes use of different logical expressions to identify IS from polarization combinations. Optimised thresholding determines the final accuracy of the delineated IS, which, similar to optical data, can be challenging to define accurately in varied urban spaces and may result in loss of detail (He *et al.*, 2025).

In addition, S-1 SAR derived VV polarization when used in isolation has shown promise in urban landcover mapping. Despite these promising results, high backscatter variations associated with complex scattering behaviour of urban environments (e.g speckle, data shadows, and corner reflections from walls and structures) make the accurate extraction of IS from either VH and/or VV challenging to implement (Dermosinoglou and Petropoulos, 2024). Because of these limitations, a fusion method that leverages the strengths of both SAR and multispectral imagery is justified (He *et al.*, 2025).

One applied pixel-based machine learning (ML) classifier that can successfully incorporate and fuse SAR and optical data and has consistently achieved high classification accuracy is the random forest (RF) classifier (Liu, Wang and Zhang, 2012). The RF can leverage multiple images, independently classified datasets and derived-indices as predictors (El-Hattab, 2016). This capability facilitates the development of a methodology that reflects the contemporary operational practice, characterized by transparency, computational efficiency and ease of implementation within NMAs. RF classifiers are preferred for a fusion approach as they are an ensemble learning algorithm which uses a technique of bagging to create a number of decision trees. Each tree is trained on a random subset of the training data, so that at every node a random subset of variables is applied to reduce overfitting. The classification output is determined by a majority vote of the decision trees within a forest (Rodriguez-Galiano *et al.*, 2012). The RF classifier is trained using a number of samples for each dataset in the fused whole, representing known different land cover types. Once trained the model computes the classification to determine the class of every pixel in the input image based on aggregated predictions from the ensemble (Shrestha, Stephen and Ahmad, 2021). Although deep learning and foundational models represent the state-of-the-art in IS mapping (Zhao *et al.*, 2023), they were not applied in this study. Their routine application in NMAs remains constrained by several operational requirements, including massive annotated samples, significant computational resources and specialized expertise (Vali, Comai and Matteucci, 2020).

The incorporation of merged indices in RF classifiers and change detection has also resulted in higher accuracies in land cover classification, particularly when distinguishing IS, vegetation and water bodies (Osgouei *et al.*, 2019). These indices (predictors) improve the differentiation of land cover types by capturing complementary information across multiple spectral bands, improving the classification metrics (Nasar *et al.*, 2025). The relevance of post-classification change detection comes from the fact that it not only quantifies the extent, patterns and distribution of a change but importantly also provides insights into how other land cover types such as vegetation and water contribute to those observed changes (El-Hattab, 2016).

The primary objective of this study was therefore to explore the fusion of spectral and radar-based indices to delineate IS changes in the urban area of Dublin, Ireland between the years 2017 and 2023 using S-1 (SAR) and S-2 (multispectral) imagery, leveraging the strengths of a cloud platform such as Google Earth Engine (GEE). Indices used include NDVI and NDBI (Kebede, Hailu and Suryabhagavan, 2022), derived from S-2 imagery. From S-1, the VV polarization band was incorporated, alongside the PRISI which leverages both polarizations.

In addition, the study examined incorporation of the Normalized Difference Water Index (NDWI) as an additional index to improve the classification metrics of 2017 and 2023 outputs (Lee, Acharya and Lee, 2018) focusing on three land cover types of particular interest to NMAs for urban mapping and to end users in hydro-climatology. This also facilitated exploration of an earth observation approach to provide information to augments more traditional aerial survey methods.

To delineate and monitor changes in IS between 2017 and 2023, a post-classification change detection analysis was performed. This approach was used to quantify transitions between land-cover types, such as vegetation changing to IS and IS changing to vegetation, providing insights into the magnitude and spatial patterns of change over time (Lu *et al.*, 2004). RF classifier was employed for image classification (Yuan *et al.*, 2023) while image differencing and post-classification comparison were applied for change detection (Lu *et al.*, 2004) (Yuan *et al.*, 2023).

1.1 Study Area – Dublin City, eastern Ireland

The urbanised area of Dublin city (Fig. 1) was used as a case study site of approximately 115 km^2 to evaluate the efficacy of data fusion and machine learning methods in the classification of IS in a complex and heterogenous urbanised setting. In the last decade, Dublin city has been experiencing rapid infrastructural development rates related to continued high levels of economic development leading to strong demand for infrastructure expansion and improvement and the provision of accommodation, all of which have directly contributed to the expansion of urban IS between the period of the tests of 2017 and 2023 (Dublin City Council, 2022).

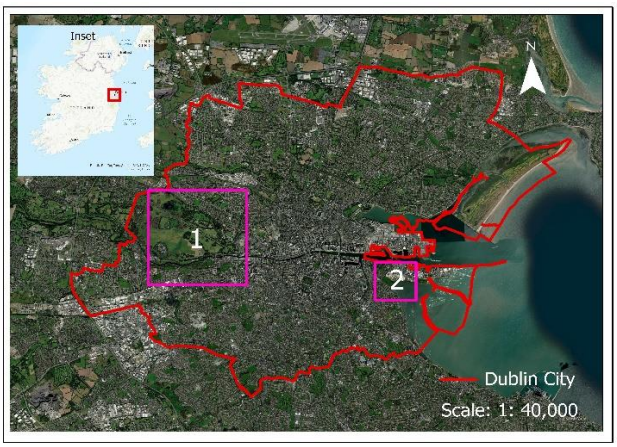

Figure 1. Location map of the study area – Dublin, Ireland. Phoenix Park, the initial inspection and classification test site, is the large green area immediately west of Dublin city centre (Fig. 3 & 4). The other test site is the Sean Moore Park (Fig. 6 & 8).

2. Data and Methods

2.1 Data

The generally low-rise, low-density urban fabric of Dublin city (Fig. 1) includes vegetated (predominately green) areas, <5 storey buildings, pavements and roads, all of which can be

mapped with reasonable accuracy using 10m NIR and red bands, along with the 20m SWIR band all from S-2 (Dermosinoglou and Petropoulos, 2024) but not of the accuracy generally required/necessary for NMA updates. Imagery for the area was accessed, queried, inspected and selected via GEE, a popular cloud platform for data processing with clear potential for incorporation in NMA workflows using large satellite datasets (Mensah, Barrett and Cahalane, 2025).

Satellite images (S-1 and S-2) for the years 2017 and 2023 were used to calculate the indices, generate the classified outputs and for the post-classification change detection. A single S-2A (satellite) Level-2A image (where Level-*2A* refers to atmospherically corrected data providing Bottom-of-Atmosphere [BOA] reflectance) with less than 10% cloud cover was selected, together with S-1 image corrected for both speckle and topographic/radiometric effects converted to σ^0. Image compositing was deliberately avoided due to the persistent cloud cover characteristic of Dublin and the specific temporal requirements of the study.

2.2 Methods

The workflow undertaken in this study from satellite image acquisition through image pre-processing/processing, data fusion, calculation of priority classes and finally validation is summarised in Fig. 2.

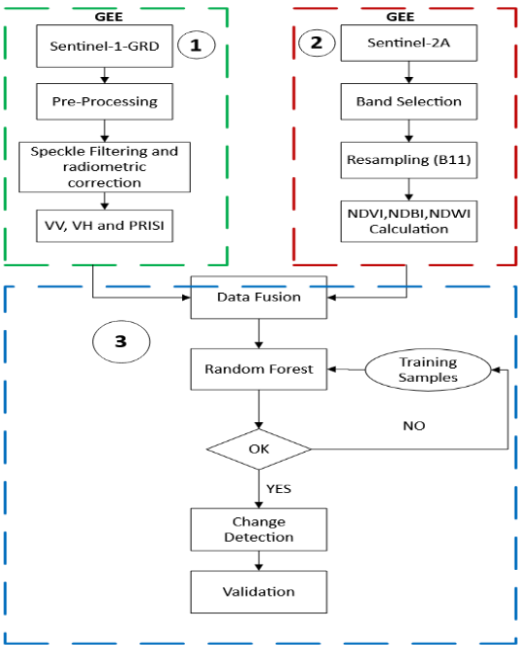

Figure 2. Workflow of processing steps for S-1 (green dotted box) and S-2 (red dotted box) datasets, followed by data fusion, classification and analysis (blue dotted box).

2.2.1 Impervious Surface Mapping—Sentinel-1 and 2

An NDVI for the study area was computed using the in-built GEE function '***ee.normalizedDifference***' which applies the S-2 NIR (B8) and red (B4) bands (Dermosinoglou and Petropoulos, 2024). An NDBI was calculated after resampling band 11 (SWIR) from 20 to 10m resolution to allow compatibility with the NIR band (Osgouei *et al.*, 2019). The final outputs were reprojected from WGS84 to EPSG:2157 – IRENET95/Irish Transverse Mercator for comparison with NMA datasets in the validation stage.

For the S-1 data, images were inspected and selected from the Interferometric Wide (IW) Swath catalogue with ascending orbit properties and VV polarization. For reasons previously outlined, VV was expected to provide a clearer contrast between IS and other land cover types in urban areas. This follows best practice suggested by the ESA S-1 observation strategy for monitoring changes in land cover (Shrestha, Stephen and Ahmad, 2021). GEE stores S-1 backscatter in dB (σ^0 in dB), and converting to a linear scale gives σ^0 in linear units, which is considered more suitable for quantitative analysis applications such as land cover classification, particularly IS extraction (Wu *et al.*, 2023).

Radiometric calibration was performed to reduce the effect of the topography on the backscatter data (Mullissa *et al.*, 2021). To assess the suitability of this method for a wide range of NMAs, and therefore independent of the quality of any available national elevation datasets, the freely available Copernicus DEM was used to perform radiometric correction of the two bands. A 3 × 3 Lee speckle filter was subsequently applied to the corrected data. This filter considers contributions from neighbouring pixels to reduce speckle noise while preserving edges (Mullissa *et al.*, 2021). This was applied to both the VV and VV/VH inputs to the PRISI datasets (Dermosinoglou and Petropoulos, 2024).

Thresholds defined as optimal for classification of IS from a previous EO study (He *et al.*, 2025) were applied. Specifically, VV backscatter σ^0 values between 0.1 and 0.5 (linear units) and a VV/VH ratio below 0.5 were linearly scaled, while VV σ^0 values greater than 0.5 and VV/VH ratio values above 0.5 were classified as impervious. The determination of the optimal thresholds to use in PRISI was the most challenging as these thresholds directly influence the contribution of the S-1 images in classification outputs from the fused dataset (He *et al.*, 2025). In addition, thresholding approaches can potentially lead to incorrect merging of other land cover types and eventually affect the final classification and measures of its influence in feature/predictor importance (Ding *et al.*, 2022).

2.2.2 Classification and Feature Importance

Image classification inputs included the three indices; NDBI, NDVI, PRISI along with the VV polarization derived in Section 2.2.1. Prior to data fusion, the VV polarization band was normalized to a stretched scale of 0 to 1 for compatibility with other layers. Samples representing the three priority land cover classes namely water, vegetation, and IS were generated using a stratified random sampling approach. A minimum separation distance of 100 m was applied between sample points to reduce spatial autocorrelation, which could otherwise bias classification results. An initial set of 200 samples was tested for model training and validation; however, preliminary results indicated signs of overfitting and class overgeneralization. To mitigate this, the sample size was increased by an additional 100 samples, resulting in a total of 300 samples. These were eventually divided into training (80%) and validation (20%) subsets for implementation within RF classifier. Following similar optical-only studies such as that of Osgouei et al, (2019), 300 training samples were deemed sufficient when fusing the data using a four-band layer stack for classification of the 3 key land cover types. For quantitative comparison, S-2 derived indices were initially fused independently and subsequently integrated with S-1 data.

Change detection was accomplished using the function ***'ee.Classifier.smileRandomForest'*** from the Statistical Machine Intelligence & Learning Engine (SMILE) library in GEE, with a forest of 100 decision trees, a commonly used value that provides a balance between computational time and model performance

(Shrestha, Stephen and Ahmad, 2021). The feature importance for each of the indices was also computed using the function ***‘classifier.explain()’*** enabling exploration of each predictor’s contribution to the model.

2.2.3 Change Detection and Validation

Validation of the outputs was performed using a combination of vector and raster datasets. Orthophotography from Tailte Éireann (the Irish NMA) and existing building and road data downloaded from OpenStreetMap were used to validate the four calculated indices and the changes identified during the change detection stage. In addition, high-resolution imagery available in GEE was used to verify post-classification change detection results occurring outside the NMA revision cycle.

3. Results and Discussion

3.1 Impervious Surface (IS) Mapping

Phoenix Park is a mature public park of >1500 ha defining the western margin of Dublin city centre along the northern fringe of the River Liffey (Site 1, Fig. 1). It is an easily identified area of relatively stable land cover types on the urban fringe, bordering a major permanent watercourse (River Liffey) with extensive vegetated areas traversed by well-defined roads e.g. Chesterfield Avenue, pavement and cycle track surfaces and containing long-standing isolated, large-scale infrastructure e.g. Dublin Zoo.

The Phoenix Park therefore contains examples of each of the principal land cover types to be differentiated in this methodology, namely IS, water and vegetation (labelled 1, 2, 3 respectively in Fig. 3) enabling an evaluation of the spectral, spatial and backscatter variations over multiple image acquisition periods. The images used in this study were all acquired in April (springtime) to reduce likely seasonal variation in deciduous vegetation cover and soil moisture conditions.

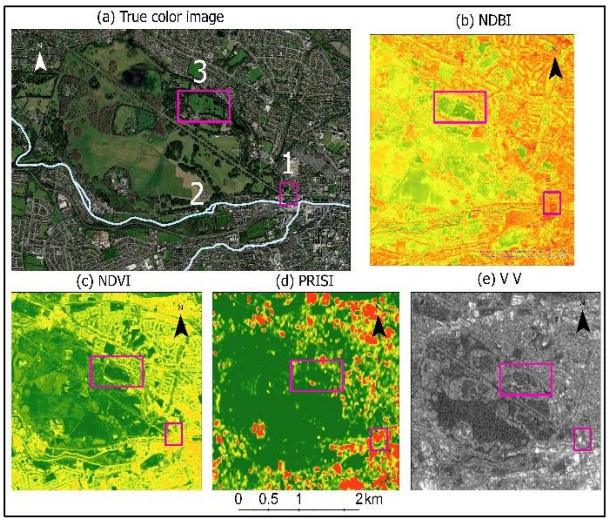

Figure 3. 2017 imagery of the Phoenix Park test site (west-central Dublin). Land cover classes: (1) IS, (2) water, (3) vegetation. (a) True colour; (b) NDBI and (c) NDVI, with orange–red indicating IS and green representing vegetation and water; (d) PRISI highlighting IS clusters; (e) VV polarization in grayscale showing backscatter variation.

Initial visual inspections indicated that the NDBI (Fig. 3b) outperformed the other indices and VV backscatter when delineating the three land cover types (Osgouei *et al.*, 2019). Features within the marked boxes of 3(b) were more distinguishable than those in 3(d) and 3(e). Overall, S-2 indices, particularly NDVI (3(c)), provided greater differentiation of land cover types due to higher spectral separability relative to S-1 SAR-derived indices (3 d & e).

Visual inspection also tended to confirm the assumption that the application of predefined thresholds to NDVI values make it an effective method of differentiating IS from other land cover types (Kuc and Chormański, 2019). However, a limitation of NDVI demonstrated for feature 2 in Figure 3 (c) is the differentiation of water from IS where both may appear to have low reflectance in NIR and may require the application of other thresholds to differentiate them. Since water absorbs NIR, it becomes difficult to differentiate water from the IS especially when they are spatially adjacent to each other (Valdiviezo-N *et al.*, 2018). As a result, the addition of a Normalized Difference Water Index (NDWI) was proposed and included in the fused dataset to improve the delineation of IS from water bodies and eliminate confusion with other land cover types (Lee, Acharya and Lee, 2018). Although ancillary datasets such as CORINE land cover could be employed as a water mask, such products may omit dynamically changing water bodies, including sections of the River Liffey. Additionally, the study was intentionally designed to remain independent of supplementary datasets in order to ensure broader transferability for NMAs.

The thresholding approach in PRISI suggested similar confusion, with the result that certain IS were incorrectly mapped as larger clusters, with a loss of important spatial details such as River Liffey and the Chesterfield Avenue (He *et al.*, 2025). This appears to be confirmed by Figure 3 (d), where PRISI proved better for identifying the clusters of IS rather than a detailed differentiation of land cover in the input image as a whole. The suitability of VV backscatter for differentiating between high-rise and low-rise impervious structures, due to corner reflections and look angle effects (Holtgrave *et al.*, 2020), was also evident in visual inspection. Because both VV and PRISI capture the surface scattering characteristics of these structures, they potentially facilitate a more accurate identification in the S-1 imagery when compared to indices derived solely from S-2. VV offered a better differentiation compared to PRISI, potentially because of the application of thresholds in PRISI coupled with low backscatter consistency, making it difficult to accurately differentiate different types of IS. Site 1 in Figure 3 (e) highlights a multi-story building, these Criminal Courts of Justice appear brighter than nearby buildings due to high scattering associated with double bounce/corners of buildings. Bright areas represented the other areas of IS while dark ones may include vegetated areas or water bodies. Although the S-1 VV-backscatter is not a derived index like NDVI, NDBI, PRISI and NDWI, it clearly provides complementary information about the surface and warrants inclusion in the RF classifier. This was particularly significant in differentiating the IS from vegetation and water as the VV responds strongly to the roughness of the surface and soil moisture.

Initial comparisons were also conducted between the fusion of S-2–derived indices alone and the combined fusion of S-1 and S-2 data. The results indicated that integrating S-1 with S-2 significantly improved the delineation of IS. Among the advantages of S-1 were its ability to reduce misclassified outliers present in the S-2–only results, particularly the confusion between bare soils and IS and its contribution of structural and physical properties of the surface. The integration of this SAR-derived information with spectral indices enhanced class separability and improved the robustness of land cover

distinction. Figure 4 illustrates five examples of this improvement.

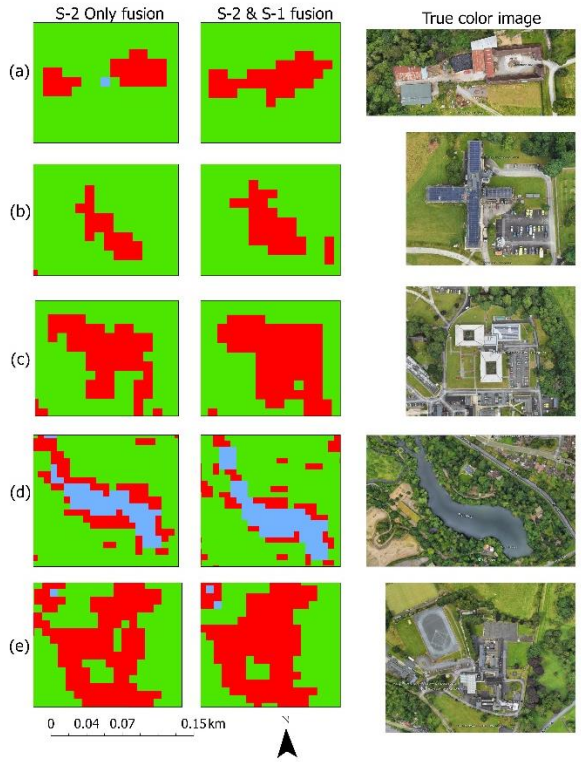

Figure 4. Comparison of results derived from S-2 data fusion alone and from the fusion of S-1 and S-2 data for Test Site 1 (Fig. 1). The sub-figures include: (a) Knockmaroon Farm; (b–c) sections of St. Mary’s Hospital, Phoenix Park; (d) the fish pond adjacent to Dublin Zoo; and (e) Mount Sackville Nursing Home. Improved delineation of IS is observed in (a), (b), (c), and (e), while (d) demonstrates improved water body delineation.

3.2 Image Classification—Accuracy Assessment

Using the remaining 20% of the points for validation, the RF classification achieved an overall accuracy of 93% and a Kappa coefficient of 0.91 for the 2017 image, and 91% and 0.89, respectively, for the 2023 image (Table 1). The high accuracy was partially attributed to a high percentage of training samples (80%) and the relatively low number of priority classes used, indicating a high level of correspondence between the training samples and the three output classes. The commission and omission errors for all classes were less than 10% except for IS, which had high commission errors (12%) for both the 2017 and 2023 images. The omission errors of 7% and 8% for vegetation for 2017 and 2023 respectively suggested a confusion between vegetation and IS. These classification errors are likely related to mixed pixels along land cover edges, particularly where IS was adjacent to water bodies and vegetation. Figure 5 illustrates an example from two classified images for 2017 with the highlights of improvement in IS delineation along River Liffey upon the addition of NDWI.

Year	Overall Accuracy (%)	K. Coefficient	Class	C. Error (%)	O. Error (%)
2017	93	0.91	IS	12	8
			Vegetation	4	7
			Water	3	4
2023	91	0.89	IS	12	10
			Vegetation	5	8
			Water	4	5

Table 1. Accuracy assessment results summarizing the overall accuracy and the class-specific commission and omission errors of the 2017 and 2023 classification.

The occurrence of high omission and commission errors is a common challenge in the mapping and monitoring IS (Zhang *et al.*, 2020). This can be attributed to the high spectral heterogeneity of urban materials such as concrete, asphalt, and rooftops, which often exhibit spectral characteristics similar to those of bare soil and dry vegetation, thereby causing spectral confusion between classes (Su *et al.*, 2022). Moreover, the spatial resolution of S-2 imagery increases the likelihood of mixed pixels containing both impervious, water and vegetated elements, which also contribute to misclassification (Li *et al.*, 2021). For S-1, the backscatter behavior of urban features particularly double bounce from structures can introduce additional complexities impacting land cover differentiation. Shadowing effects and geometric distortions such as layover can also alter the backscatter recorded by the sensor, leading to inaccuracies in class assignment and consequently increasing omission and commission errors. The inclusion of S-1 data in densely built-up areas with complex structural details may therefore introduce uncertainties during classification (Nasar *et al.*, 2025). Volumetric scattering and the vegetation backscatter in S-1 were also identified as potentially significant sources of errors in IS delineation (Liu *et al.*, 2023). These factors collectively make the accurate delineation of IS considerably more difficult than that of spectrally distinct classes such as vegetation and water. The application of object-based image analysis (OBIA) and SAR ratio features other than PRISI can be used to mitigate the mixed-pixel impacts while post-classification filters may also be used to improve the overall classification accuracy and reduce the misclassification of IS in future studies (Feng and Fan, 2021).

Addition of the NDWI has been shown to improve classification accuracy in optical-only studies, particularly in images containing water bodies adjacent to IS (Lee, Acharya and Lee, 2018). The addition of the NDWI resulted in only minimal improvements in the commission and omission errors by 2%, enabling more accurate differentiation of the land covers especially those adjacent to water bodies. It did however improve the separation of different features and importantly enhanced differences between water and the IS as seen in Figure 5. This application of NDWI and NDBI improved the analysis to some extent, allowing for a clearer distinction between water and IS (Saleh, 2022). Additionally, the integration of a S-1 index in the fused dataset provided complementary backscatter information, further enhancing the classification of IS and water features (Shrestha, Ahmad and Stephen, 2021; Nasar *et al.*, 2025).

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-345-2026

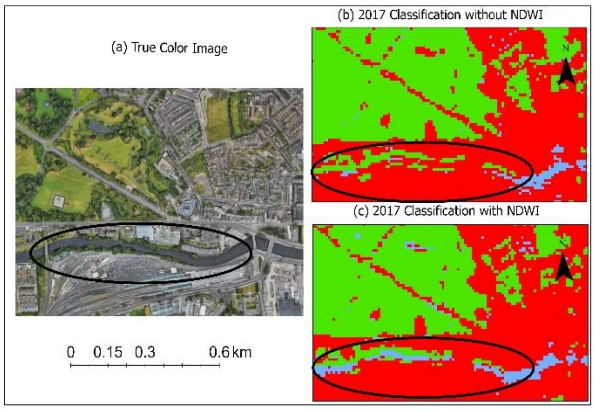

Figure 5. Improvements in IS delineation along the River Liffey (Site 1, Fig. 1). Areas initially misclassified as (IS) were correctly reclassified as water following the inclusion of NDWI.

3.3 Transition Between Land Cover Types

The post-classification change detection stage was a comprehensive way of understanding different transitions between classes. Figure 6 illustrates the output of post-classification comparison both with and without the inclusion of NDWI. Areas represented in red show transition from vegetation to IS. Confusion between urban areas and bare/barren soils was identified as another problem inherent in extracting IS from S-2 using spectral indices only, resulting in misclassification, as also reported by Valdiviezo-N et al. (2018). This was evident in several areas (represented in dark pink) that transitioned from IS to vegetation suggesting spectral confusion between vegetation and IS. This was further associated with commission errors of greater than 10% in IS for both images in Table 1.

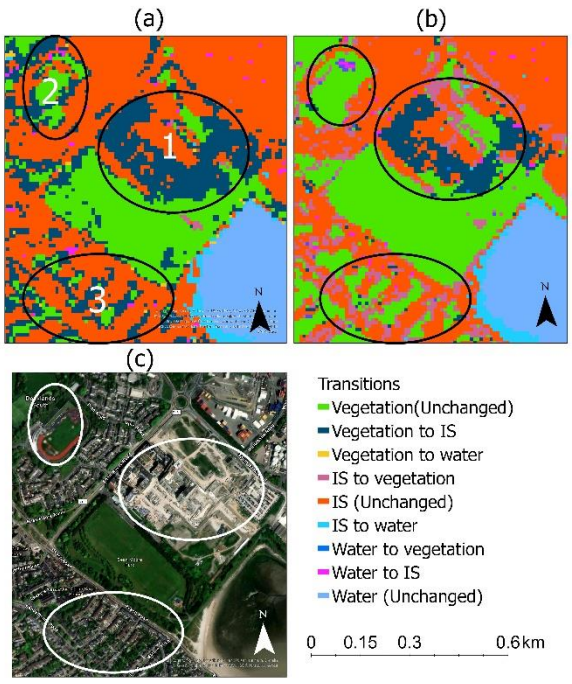

Figure 6. Post-classification comparison illustrating the impact of including NDWI as an additional index, showing enhanced detection of changes from vegetation to IS. (a) Categorical change without NDWI, (b) Categorical change with NDWI and (c) 2023 True color image.

3.4 Analysis of Feature Importance

Feature importance (Figure 7) was also calculated to determine the overall influence of each of the individual layers used for image classification and specifically in the makeup of the Random Forest classification model. Although NDBI is a specific index for mapping urban areas, it did not emerge as the most influential index in image classification for these tests, possibly due inherent issues stemming from resampling of the 20m resolution SWIR band or the differences in spectral response of artificial urban surfaces (Rouibah, 2025). According to Rouibah (2025), NDBI often struggles to accurately distinguish IS from other spectrally similar features such as bare soil or dry vegetation due to spectral heterogeneity of those artificial surfaces. NDVI played the most significant role in the fused dataset in accurately classifying the three land cover types, whereas S-1 PRISI had the lowest contribution and was the least important in this study. NDWI ranked just behind NDBI which supports the assumption around the importance of a water index such as NDWI as a supplementary input in mapping and monitoring IS in areas containing water bodies (Saleh, 2022).

Although PRISI, VV, and NDWI show importance values close to each other, the differences between them are marginal and may not be statistically significant in practical terms. Overall, the range 18.99 – 22.60% indicated that the RF model does not solely rely on a single variable. Instead, all predictors contributed to the model performance supporting the robustness of data fusion approach. Alternative indices for IS mapping, such as Built-up Indiex (BI) and Built-up Area Extraction Index (BAEI), do not directly use SWIR bands and can significantly improve mapping accuracy from Sentinel-2 imagery (Osgouei *et al.*, 2019).

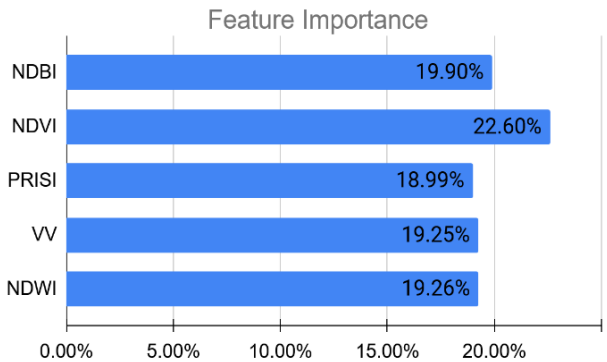

Figure 7. Feature importance (%) of each of the four applied indices and the VV band used in image classification.

3.5 Value of Cloud-based Image Processing

Exploring the implementation of this method in GEE greatly influenced the efficiency and scalability of mapping and monitoring IS. The transition from simple quantitative fusion of S-2–derived indices to the integration of S-1 and S-2 datasets resulted in both increased processing complexity and improved IS mapping performance. Through cloud-based processing, GEE allowed for rapid selection, querying, inspection and application of large multi-temporal datasets for both S-1 and S-2, with significant time saving from not having to manually download and store data locally.

The average cloud processing time for index computation was less than 3 seconds per 115 km² area. In comparison, processing S-2–only indices required approximately 10 seconds under similar cloud-based conditions. Performing the same procedure,

including computing a single index and conducting S-2–only fusion in a typical desktop-based environment such as ERDAS Imagine, can take approximately 30–60 seconds (Sharma *et al.*, 2024). Beyond improved processing time, cloud-based environments support access to extensive satellite archives, enabling efficient monitoring of IS expansion and supporting mapping updates by NMAs. Through the use of GEE built-in packages and libraries for image processing, including classification and change detection, computational demands were substantially reduced in terms of processing time, storage requirements, reproducibility, automation, scalability, and flexibility. This makes GEE particularly suitable for multi-temporal image analysis, enabling NMAs to undertake routine monitoring of urban growth and update land-cover datasets using available long-term satellite archives.

3.6 Validation of Classification against Imagery

This study utilised high-resolution Google Earth images (Shrestha, Stephen and Ahmad, 2021) together with Tailte Éireann orthophotography to explore IS changes over time. Figure 8 shows an example for a site at Sean Moore Park in Dublin, which changed between 2017 and 2023 and was successfully detected in the binary change map. This site was then examined further for validation using its corresponding location in the orthophoto. However, due to the spatial resolution of the images used and the complex morphology of different land cover types within the validation site, certain features could not be identified in the classified images.

The high-resolution datasets enabled more robust validation of the changes detected between the two fused satellite images. However, as shown in Section 3.2, errors in image classification may have contributed to discrepancies between the change maps and the reference imagery shown in Figure 8. Despite these limitations, the combination of high-resolution datasets substantially increased confidence in interpreting and verifying land cover transitions.

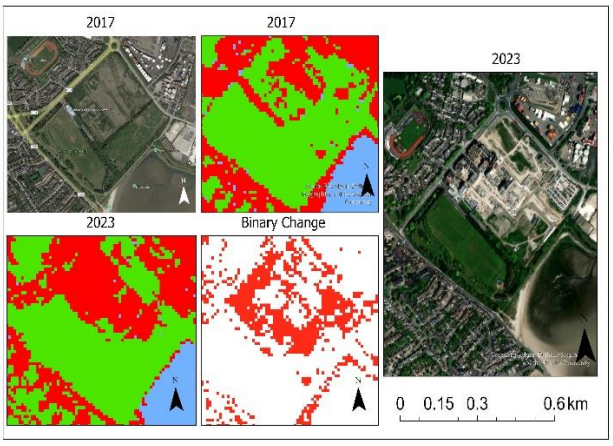

Figure 8. Sean Moore Park in Dublin that was shown as vegetated in 2017, changed to IS in 2023 and confirmed as having changed to IS in 2023 using high resolution imagery.

4. Conclusion

This study aimed to demonstrate an Earth Observation (EO) only data fusion methodology for mapping changes in IS using open source and free to use multispectral satellite imagery and Synthetic Aperture Radar (SAR) data. It was found that SAR imagery can augment optical data, but with limitations. SAR-derived data including the PRISI index and VV backscatter ranked lowest in terms of feature importance detection, however they still have the potential of being used to differentiate IS from other similar optical properties land covers. This is largely due to the complementary structural information provided by SAR data. In particular, the VV polarization band shows sensitivity to IS and contributes to their delineation. The PRISI performs well when the IS are spatially clustered but less accurate when the concern is to perform a detailed feature differentiation. The integration of these different data types can be further improved using a post-classification change detection approach. Given the frequent cloud cover restrictions restricting the use of multispectral data for regular mapping, advantages are demonstrated in the data fusion for the detection of built urban fabric change that could feed into NMA mapping programmes.

The comparison between S-2-only results and the combined S-1 and S-2 fusion demonstrates the clear benefits of data integration. While S-2 alone provided important spectral information in distinguishing IS from other land cover types, incorporating S-1 data improved delineation accuracy, reducing misclassifications and minimizing outliers present in S-2-only results. This enhancement reflects the complementary structural information provided by S-1, emphasizing the value of a fusion approach in monitoring urban development. These findings provide practical insights for NMAs and other planning authorities seeking to track urban change more accurately. Future work will explore the application of deep learning and foundational models, as well as additional radar-based indices, to further strengthen the mapping and monitoring of IS.

Acknowledgement

The authors gratefully acknowledge the use of Google Earth Engine (GEE), open-access satellite data from the Copernicus Programme, orthophotography provided by Tailte Éireann, and the Copernicus DEM used in this study.

References

Dermosinoglou, A. and Petropoulos, G.P. (2024) ‘Exploring long term Impervious Surface Areas (ISA) dynamics using Landsat imagery, Machine Learning and GEE: The case of Attica, Greece’, *Remote Sensing Applications: Society and Environment*, 36(February), p. 101338. https://doi.org/10.1016/j.rsase.2024.101338.

Ding, Y. *et al.* (2022) ‘PRISI: A novel piecewise radar impervious surface index for urban monitoring using Sentinel-1 data’, *International Journal of Applied Earth Observation and Geoinformation*, 114(September), p. 103033. https://doi.org/10.1016/j.jag.2022.103033.

Dublin City Council (2022) ‘Chapter 14 : Land-use Zoning’, pp. 525–548. https://www.dublincity.ie/dublin-city-development-plan-2022-2028/written-statement/chapter-14-land-use-zoning.

El-Hattab, M.M. (2016) ‘Applying post classification change detection technique to monitor an Egyptian coastal zone (Abu Qir Bay)’, *Egyptian Journal of Remote Sensing and Space Science*, 19(1), pp. 23–36. https://doi.org/10.1016/j.ejrs.2016.02.002.

Feng, S. and Fan, F. (2021) ‘Impervious surface extraction based on different methods from multiple spatial resolution images : a comprehensive comparison’, 8947. https://doi.org/10.1080/17538947.2021.1936227.

He, S. *et al.* (2025) ‘A comparative analysis of machine learning-based methods for impervious surface mapping using SAR and optical data’, *Geocarto International*, 40(1), p. https://doi.org/10.1080/10106049.2025.2521833.

Holtgrave, A.K. *et al.* (2020) ‘Comparing Sentinel-1 and -2 data and indices for agricultural land use monitoring’, *Remote Sensing*, 12(18). https://doi.org/10.3390/RS12182919.

Kebede, T.A., Hailu, B.T. and Suryabhagavan, K.V. (2022) ‘Evaluation of spectral built-up indices for impervious surface extraction using Sentinel-2A MSI imageries: A case of Addis Ababa city, Ethiopia’, *Environmental Challenges*, 8(June), p. 100568. https://doi.org/10.1016/j.envc.2022.100568.

Kuc, G. and Chormański, J. (2019) ‘Sentinel-2 imagery for mapping and monitoring imperviousness in urban areas’, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(1/W2), pp. 43–47. https://doi.org/10.5194/isprs-archives-XLII-1-W2-43-2019.

Lee, J.K., Acharya, T.D. and Lee, D.H. (2018) ‘Exploring land cover classification accuracy of Landsat 8 image using spectral index layer stacking in hilly region of South Korea’, *Sensors and Materials*, 30(12), pp. 2927–2941. https://doi.org/10.18494/SAM.2018.1934.

Li, C. *et al.* (2021) ‘A Comparative Analysis of Index-Based Methods for Impervious Surface Mapping Using Multiseasonal Sentinel-2 Satellite Data’, 14, pp. 3682–3694. https://doi.org/10.1109/JSTARS.2021.3067325.

Liu, J. *et al.* (2023) ‘Large-Scale Impervious Surface Area Mapping and Pattern Evolution of the Yellow River Delta Using Sentinel-1/2 on the GEE’, *Remote Sensing*, 15(1). https://doi.org/10.3390/rs15010136.

Liu, Y., Wang, Y. and Zhang, J. (2012) ‘New Machine Learning Algorithm : Random Forest’, pp. 246–247. https://link.springer.com/chapter/10.1007/978-3-642-34062-8_32.

Lu, D. *et al.* (2004) ‘Change detection techniques’, *International Journal of Remote Sensing*, 25(12), pp. 2365–2401. https://doi.org/10.1080/0143116031000139863.

Lydon, K. *et al.* (2018) ‘National Land Cover Map of Ireland 2018’. https://tailte.ie/wp-content/uploads/2025/03/National-Land-Cover-Map-Final-Report-2018.pdf.

Mensah, I., Barrett, B. and Cahalane, C. (2025) ‘Assessing change point detection methods to enable robust detection of early stage Artisanal and Small-Scale mining (ASM) in the tropics using Sentinel-1 time series data’, *International Journal of Applied Earth Observation and Geoinformation*, 139(April), p. 104525. https://doi.org/10.1016/j.jag.2025.104525.

Mullissa, A. *et al.* (2021) ‘Sentinel-1 sar backscatter analysis ready data preparation in google earth engine’, *Remote Sensing*, 13(10), pp. 5–11. https://doi.org/10.3390/rs13101954.

Nasar, M. *et al.* (2025) ‘Enhanced urban impervious surface land use mapping using a novel multi - sensor feature fusion method and remote sensing data’. https://link.springer.com/article/10.1007/s12665-025-12217-0.

Osgouei, P.E. *et al.* (2019) ‘Separating built-up areas from bare land in mediterranean cities using Sentinel-2A imagery’, *Remote Sensing*, 11(3), pp. 1–24. https://doi.org/10.3390/rs11030345.

Rodriguez-Galiano, V.F. *et al.* (2012) ‘An assessment of the effectiveness of a random forest classifier for land-cover classification’, *ISPRS Journal of Photogrammetry and Remote Sensing*, 67(1), pp. 93–104. https://doi.org/10.1016/j.isprsjprs.2011.11.002.

Rouibah, K. (2025) ‘A new combination of spectral indices derived from Sentinel - 2 to enhance built - up mapping accuracy of cities in semi - arid land’, *Arabian Journal of Geosciences* [Preprint]. https://doi.org/10.1007/s12517-025-12225-1.

Saleh, A.M. (2022) ‘Improving the Accuracy of Land Cover Classification using Sentinel 2 Data and Knowledge Based Classification System in the West of Amara City, Iraq’, *Journal of the Indian Society of Soil Science*, 70(1), pp. 1–9. https://doi.org/10.5958/0974-0228.2022.00001.9.

Samadi Todar, S.A., Attarchi, S. and Osati, K. (2021) ‘Investigation the seasonality effect on impervious surface detection from Sentinel-1 and Sentinel-2 images using Google Earth engine’, *Advances in Space Research*, 68(3), pp. 1356–1365. https://doi.org/10.1016/j.asr.2021.03.039.

Shao, Z. *et al.* (2016) ‘Mapping urban impervious surface by fusing optical and SAR data at the decision level’, *Remote Sensing*, 8(11), pp. 1–21. https://doi.org/10.3390/rs8110945.

Sharma, S. *et al.* (2024) ‘Remote Sensing and GIS in Natural Resource Management : Comparing Tools and Emphasizing the Importance of In-Situ Data’, (Ml). https://www.mdpi.com/2072-4292/16/22/4161.

Shrestha, B., Ahmad, S. and Stephen, H. (2021) ‘Fusion of Sentinel-1 and Sentinel-2 data in mapping the impervious surfaces at city scale’, *Environmental Monitoring and Assessment*, 193(9). https://doi.org/10.1007/s10661-021-09321-6.

Shrestha, B., Stephen, H. and Ahmad, S. (2021) ‘Impervious surfaces mapping at city scale by fusion of radar and optical data through a random forest classifier’, *Remote Sensing*, 13(15). https://doi.org/10.3390/rs13153040.

Su, S. *et al.* (2022) ‘An Impervious Surface Spectral Index on Multispectral Imagery Using Visible and Near-Infrared Bands’, *Remote Sensing*, 14(14), pp. 1–23. https://doi.org/10.3390/rs14143391.

Szigarski, C. *et al.* (2018) ‘Analysis of the Radar Vegetation Index and potential improvements’, *Remote Sensing*, 10(11), pp. 1–15. https://doi.org/10.3390/rs10111776.

Tian, Y. *et al.* (2018) ‘A novel index for impervious surface area mapping: Development and validation’, *Remote Sensing*, 10(10). https://doi.org/10.3390/rs10101521.

Valdiviezo-N, J.C. *et al.* (2018) ‘Built-up index methods and their applications for urban extraction from Sentinel 2A satellite data: discussion’, *Journal of the Optical Society of America A*, 35(1), p. 35. https://doi.org/10.1364/josaa.35.000035.

Vali, A., Comai, S. and Matteucci, M. (2020) ‘Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral

and Multispectral Earth Observation Data : A Review’. https://doi.org/10.3390/rs12152495.

Wu, W. *et al.* (2023) ‘Urban Impervious Surface Extraction Based on Deep Convolutional Networks Using Intensity, Polarimetric Scattering and Interferometric Coherence Information from Sentinel-1 SAR Images’, *Remote Sensing*, 15(5). https://doi.org/10.3390/rs15051431.

Yuan, X. *et al.* (2023) ‘Feature Importance Ranking of Random Forest-Based End-to-End Learning Algorithm’, *Remote Sensing*, 15(21), pp. 1–20. https://doi.org/10.3390/rs15215203.

Zhang, X. *et al.* (2020) ‘Development of a global 30m impervious surface map using multisource and multitemporal remote sensing datasets with the Google Earth Engine platform’, *Earth System Science Data*, 12(3), pp. 1625–1648. https://doi.org/10.5194/essd-12-1625-2020.

Zhao, S. *et al.* (2023) ‘Land Use and Land Cover Classification Meets Deep Learning ’: https://www.mdpi.com/1424-8220/23/21/8966.