IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE GEOSCIENCE AND REMOTE SENSING SOCIETY SECTION

Received 10 May 2024, accepted 27 May 2024, date of publication 3 June 2024, date of current version 18 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3408805

RESEARCH ARTICLE

Automated Pixel Purification for Delineating Pervious and Impervious Surfaces in a City Using Advanced Hyperspectral Imagery Techniques

PAYAM SAJADI[1], MEHDI GHOLAMNIA[1], STEFANIA BONAFONI[2], (Senior Member, IEEE), GERALD MILLS[3], YAN-FANG SANG[4,5], ZETING LI[3], SALMAN KHAN[1], JIAZHENG HAN[6], AND FRANCESCO PILLA[1]

[1]School of Architecture, Planning and Environmental Policy, University College Dublin, Dublin 4, D04 V1W8 Ireland
[2]Department of Engineering, University of Perugia, 06125 Perugia, Italy
[3]School of Geography, University College Dublin, Dublin 4, D04 V1W8 Ireland
[4]Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100190, China
[5]Key Laboratory of Compound and Chained Natural Hazards, Ministry of Emergency Management of China, Beijing 100731, China
[6]School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China

Corresponding author: Payam Sajadi (payam.sajadi@ucd.ie)

This work was supported by the National Challenge Fund, Science Foundation Ireland (SFI), under Grant 22/NCF/OT/11263.

ABSTRACT Conventional urbanization transforms natural into paved landscapes, posing a significant environmental challenge. Detecting the changes in (im)pervious surfaces in cities, where patches are small and intermingled, is particularly challenging. This study introduces a novel approach to these changes by integrating Coupled Non-negative Matrix Factorization (CNMF) image fusion with an automatic pixel purification algorithm. By fusing low-resolution hyperspectral (30m) with high-resolution panchromatic (5m) PRISMA imagery, we achieved enhanced spatial resolution, crucial for accurate land use and land cover (LULC) classification. We introduced automatic pixel purification as a key innovation method to improve LULC mapping accuracy, sensitive to training pixel selection and mixed pixels. This method, which is tested in Dublin City area, enhanced/ refined spectral signatures and clarity across major LULC classes including bare soil, industrial roofs, grasslands, trees, residential roofs/asphalts, and water bodies, significantly improving classification accuracy by removing outliers and ensuring spectral consistency. The Random Forest (RF) algorithm, applied before and after pixel purification, showed substantial increases in overall accuracy (from 94.04% to 96.69%,) and Kappa coefficient (from 92.60% to 95.91%) for 2021, with similar improvements in 2022. This method enabled accurate differential analysis of (im)pervious surfaces, revealing a 4.08% decrease in pervious (from 33.29 km^2 to 28.08 km^2) and a 4.09% increase in impervious surfaces (from 79.96 km^2 to 82.92 km^2) over one year, highlighting the rapid urbanization's impact on Dublin's landscape permeability. This study significantly advances LULC classification and urban monitoring, offering valuable insights for sustainable urban development and advocating for its integration into future remote sensing and urban planning initiatives.

INDEX TERMS Hyperspectral image fusion, impervious-pervious surface, pixel purification, land cover classification.

I. INTRODUCTION

Land Use and Land Cover (LULC) mapping in cities is paramount for sustainable urban planning, environmental

The associate editor coordinating the review of this manuscript and approving it for publication was Vineetha Menon.

management, and decision making. Assessing the current and changing status of urban expansion and intensification (particularly impervious surface cover) informs a wide range of urban environmental studies, including those on climate change, the urban heat island and hydrology. The accurate classification and mapping of urban LULC provides essential data for evaluating environmental impacts, guiding land use policies, and developing strategies for mitigating adverse effects on ecosystems and human health. As urban areas continue to expand globally, the need for accurate, up-to-date,and appropriate LULC information becomes increasingly critical for resilience planning, especially in the context of global climate change. Satellite remote sensing technology, including multispectral imaging (MSI) can provide some of this information and advance the study of LULC patterns on local and global scales [1], [2].

However, the spectral resolution of MSI limits its capability and accuracy for LULC classifications, particularly when distinguishing between classes with closely similar spectral signatures, in complex landscapes, such as urban areas [3]. Hence, hyperspectral imaging (HSI) can make a significant contribution to LULC mapping in cities where landcover types can be intermingled at small scales; Its unmatched capability to capture detailed spectral information enhances the discrimination between different land covers and surface materials, thus facilitating a nuanced understanding of urban dynamics. This technology has been instrumental in augmenting traditional land mapping techniques by offering a spectral richness that allows for a more precise classification and monitoring of urban land covers [4], [5], [6]. The importance of hyperspectral sensing in urban areas extends beyond simple mapping, as it is vital for monitoring impervious and pervious surfaces such as green spaces, industrial and residential buildings, roads and car parks, etc. By providing detailed spectral signatures using hundreds of narrow (less than 10 *nm*) spectral bands typically covering the visible and near infra-red (NIR) domains, hyperspectral data enhances the accuracy of classification with complex classes, and complex surface areas such as impervious surface detection, offering insights into urban flooding [7], [8], [9]. Its application has proven more effective than traditional methods in extracting impervious surfaces, especially when integrated with advanced machine learning algorithms [9], [10], [11]. While hyperspectral remote sensing advances LULC mapping, its high spectral resolution comes at the expense of spatial resolution [12], [13], hindering its application in fine-scale urban monitoring [11], [14], [15]. This trade-off also introduces a ''curse of dimensionality,'' where the increased number of spectral bands does not necessarily improve classification accuracy [14], [16], [17], [18]. To overcome this trade-off, image fusion techniques have become essential [13], [17], [19]. These image fusion algorithms combine multiple data sources to preserve hyperspectral imagery's spectral details while enhancing spatial resolution with data from other sensors like MSI and PAN images [9], [17], [20]. Numerous researchers have developed different fusion algorithms based on different assumptions to enhance the spatial resolution of HSI while preserving the integrity of spectral resolution, eventually enhancing LULC classification accuracy [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31]. While the abovementioned advanced image fusion techniques are indeed promising in enhancing the spatial resolution and preserving the spectral integrity of the data, the process of LULC classification, particularly with supervised learning algorithms such as Support Vector Machines (SVM), Random Forest (RF), Deep Neural Networks (DNN), etc., is intrinsically linked to the quality and selection of training pixels. The accuracy of these models in distinguishing various land cover types depends not only on the algorithm's sophistication but also on the representativeness and purity of the training data used. Inaccuracies in training pixel selection can lead to substantial errors in classification outcomes, underscoring the critical need for high-quality training datasets [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42]. Such contamination not only distorts the representation of land cover classes but also hampers the ability to effectively monitor changes over time. Moreover, the challenge of obtaining a sufficiently diverse and extensive training dataset becomes more pronounced with the high dimensionality of hyperspectral images. This necessitates the use of sophisticated techniques to ensure the training data is both manageable for the algorithms and representative of the complex spectral signatures encountered in such datasets. Several methods, such as spectral unmixing, noise reduction, and spatial-spectral feature extraction, are designed to improve the fidelity of training pixels by isolating pure spectral signatures from mixed ones and reducing background spectral noise [8], [43], [44], [45]. In addition, machine learning-based techniques for feature selection and dimensionality reduction, such as autoencoders and manifold learning, can further refine the dataset by highlighting the most informative features and reducing the dimensionality of the data, thereby making it more manageable for classification algorithms [46], [47], [48], [49]. However, the deployment of these sophisticated data preprocessing and enhancement techniques comes with its own set of challenges. Many of these methods are computationally intensive and time-consuming, necessitating significant processing power and expert knowledge to implement effectively. This can make them less desirable for practical applications where resources are limited, or rapid processing is required. The complexity of these approaches often requires specialized computational infrastructure and software, as well as a deep understanding of both the algorithms and the data being processed, which can be a barrier to their widespread adoption in certain contexts.

Addressing the complexities inherent in LULC monitoring in urban environments requires innovative approaches that moves beyond the traditional, complex and computationally intensive methodologies. In this context, we propose an automated, straightforward pixel-based purification algorithm which utilizes a spectral similarity angle and spectral distance

in n-dimensional feature space to precisely identify and separate mixed pixels, outliers, and erroneously selected pixels during the algorithm training phase of LULC classifier. Such innovative and automated algorithm can significantly enhance the LULC accuracy of hyperspectral imageries like PRISMA (Hyperspectral Precursor of the Application Mission) hyperspectral imaging of the Italian Space Agency [15], [50], [51]. This advanced purification methodology is crucial for enhancing the quality of the training dataset, thereby ensuring higher classification accuracy. A distinctive feature of our algorithm is its ability to derive mean spectral signatures for each target class within the urban landscape. By aligning the derived mean spectral signatures with the 'endmember' profiles of the respective classes, our method can significantly improve classification accuracy in complex urban settings. The main contribution of our automated pixel purification algorithm can be summarized as: 1) Implementation of an advanced CNMF image fusion technique to enhance the spatial resolution of PRISMA hyperspectral images. 2) Leveraging the innovative concept of automated pixel-wise quality improvement, we produced highly reliable training pixels, thereby mitigating the influence of mixed pixels and outliers, which significantly improves the accuracy of the LULC classification process. 3) Demonstration of the proposed automated method's effectiveness in pixel-based impervious and pervious surface change detection in highly complex urban environments, precisely analyzing permeable surface reduction in these areas.

This sophisticated approach to training pixel selection and purification described here is a straightforward and powerful tool that can be deployed alongside other approaches for improving LULC classification outcomes. The specific objectives of this study are:

i To obtain high-spatial resolution (5m) PRISMA hyperspectral imagery from the native 30m
ii To enhance LULC mapping accuracy using automated pixel purification algorithm
iii To monitor impervious-pervious surfaces in highly urbanized and complex areas

The method introduced here is needed for conducting precise environmental monitoring and assessing urban impacts. For example, detailed spatial knowledge of impervious surface cover can be used to assess the magnitude of surface runoff, the capacity of storm sewer systems and the potential for localised flooding. Improved LULC mapping will enhance our understanding of intra-urban hazards and the role of urban landscape management to create more resilient and sustainable environments.

The remainder of this paper is structured as follows: Section II details the presentation of materials and methods, covering the hyperspectral image fusion algorithm, our automatic pixel-purification algorithm to enhance training data quality, LULC mapping with the RF algorithm, and the strategies for mapping impervious and pervious surfaces, alongside change detection methods. Implementing this innovative and automatic approach, particularly in complex urban areas, holds the promise of distinctly improving urban water resource management, informing more accurate flood risk assessments, and guiding the development of sustainable urban policies. Section III delivers the results, including the efficacy of image fusion, the impact of pixel purification on LULC classification accuracy, and changes in impervious versus pervious surface areas from 2021 to 2022. Section IV discusses the algorithm's performance, evaluates the precision of pixel purified LULC maps, and outlines the study's key findings. Section V concludes the paper, summarizing the research contributions and potential future directions.

II. MATERIALS AND METHODS

A. PRISMA HYPERSPECTRAL SENSOR AND STUDY AREA

Our study employs the advanced capabilities of the Precision Imaging Spectrometer for Advanced Hyperspectral Monitoring (PRISMA) sensor, a pivotal advancement in remote sensing launched by the Italian Space Agency on March 22, 2019 [52], [53]. Recognized for revolutionizing remote sensing, PRISMA offers unparalleled hyperspectral imagery with a resolution across 233 spectral bands, ranging from the visible near-infrared (VNIR) to the short-wave infrared (SWIR) spectrum (400-2500 nm), achieving a spectral resolution finer than 12 nm. Its orbit at 615 kilometers facilitates a synthesis of hyperspectral sensing with high-definition panchromatic imaging, covering an extensive 200,000 square meters daily, and providing ground sample distances of 30 meters (hyperspectral) and 5 meters (panchromatic), underpinned by a 29-day revisit cycle for thorough terrestrial monitoring [54], [55].

Dublin City (Ireland, Figure 1) was selected as the site for the application of our method. Its climate is moderate (3-19 °C) thoughout the year and it receives between 750-1000 mm annually, which is evenly distributed throughout the year. Its geographic extent of 117.8 km2 includes a diverse and spatially heterogenous landscape and provides a suitable challenge for our approach. While the city boundary is fixed, the urban landscape is dynamic and is undergoing rapid intensification in response to population increases, economic development and policies to increase built density to meet climate goals. Many of these changes take place at building plot scales as new buildings are constructed and older buildings are extended, increasing impervious surface cover. The net effect of these changes is to displace more rainwater, which can result in localised pluvial flooding. Managing this hazard at neighbourhood scales requires monitoring of LULC at a resolution that matches the scale of these 'micro' landcover changes, that in aggregate affect urban water management.

We analyzed PRISMA imagery in July 2021 and August 2022 to study land use and cover changes, focusing on transitions between pervious and impervious surfaces. The PRISMA-L2D dataset, cloud-free offers an orthorectified and atmospherically corrected 30 × 30 kilometer view of Dublin, positioned at 53° 21' 0.5040" N latitude and 6° 15'

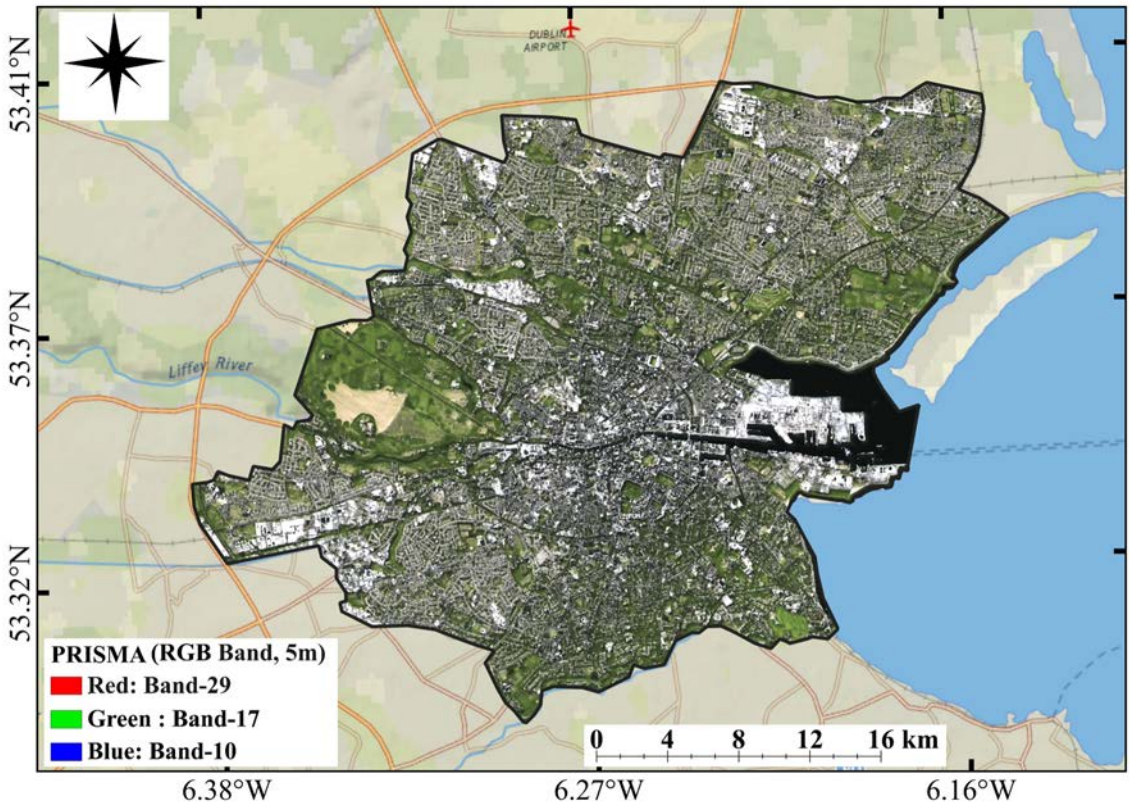

FIGURE 1. Geographical location map of the study area in Dublin, Ireland (PRISMA, August 2022).

58.1580" W longitude (Figure 1). This dataset facilitates a detailed comparison over the one-year period, shedding light on urban growth and its effects on environmental sustainability and resource management. Dublin's diverse land cover includes bare soil, grass, industrial and residential roofs, trees, and water bodies, providing a comprehensive baseline for our analysis [56].

B. METHODOLOGICAL FRAMEWORK FOR LULC CHANGE DETECTION

To identify and analyze land use and land cover (LULC) changes, particularly between impervious and pervious surfaces from 2021 to 2022, we followed a structured Four-step framework:

1. **Preprocessing and Noise Reduction of HSI Data:** The initial step involves cleaning the hyperspectral data using advanced noise reduction techniques, crucial for removing artifacts and ensuring data integrity for further processing.
2. **PRISMA Hyperspectral Image Fusion:** We fused the PRISMA hyperspectral imagery with the high-resolution panchromatic image using the Coupled Non-negative Matrix Factorization (CNMF) algorithm. This step is key for merging the detailed spatial resolution of the panchromatic data with the comprehensive spectral information of the hyperspectral imagery.
3. **LULC Classification and Accuracy Enhancement Using Automatic Pixel Purification Algorithm:** This step involves classifying the imagery to identify different land cover types, using an automatic pixel purification algorithm to improve LULC accuracy, and mixed pixels identification and outliers' removal.
4. **Detection of Changes in Impervious and Pervious Surfaces:** Finally, we focus on detecting and quantifying the changes in land cover over the one-year period, using advanced algorithms to highlight urban development trends and environmental impacts.

This methodological framework, leveraging PRISMA's hyperspectral sensor capabilities and a strategic temporal comparison, provides a comprehensive understanding of land use dynamics in Dublin. A comprehensive flowchart (presented in Figure. 2) visually outlines this methodological sequence, serving as a guide through the intricate process of LULC change detection.

C. PRISMA IMAGE PREPROCESSING AND NOISE REDUCTION

To ensure the integrity and accuracy of the data used in our classification procedure, PRISMA imagery is subjected to a series of essential preprocessing steps. Initially, hyperspectral (HS) data is converted to a format suitable for detailed

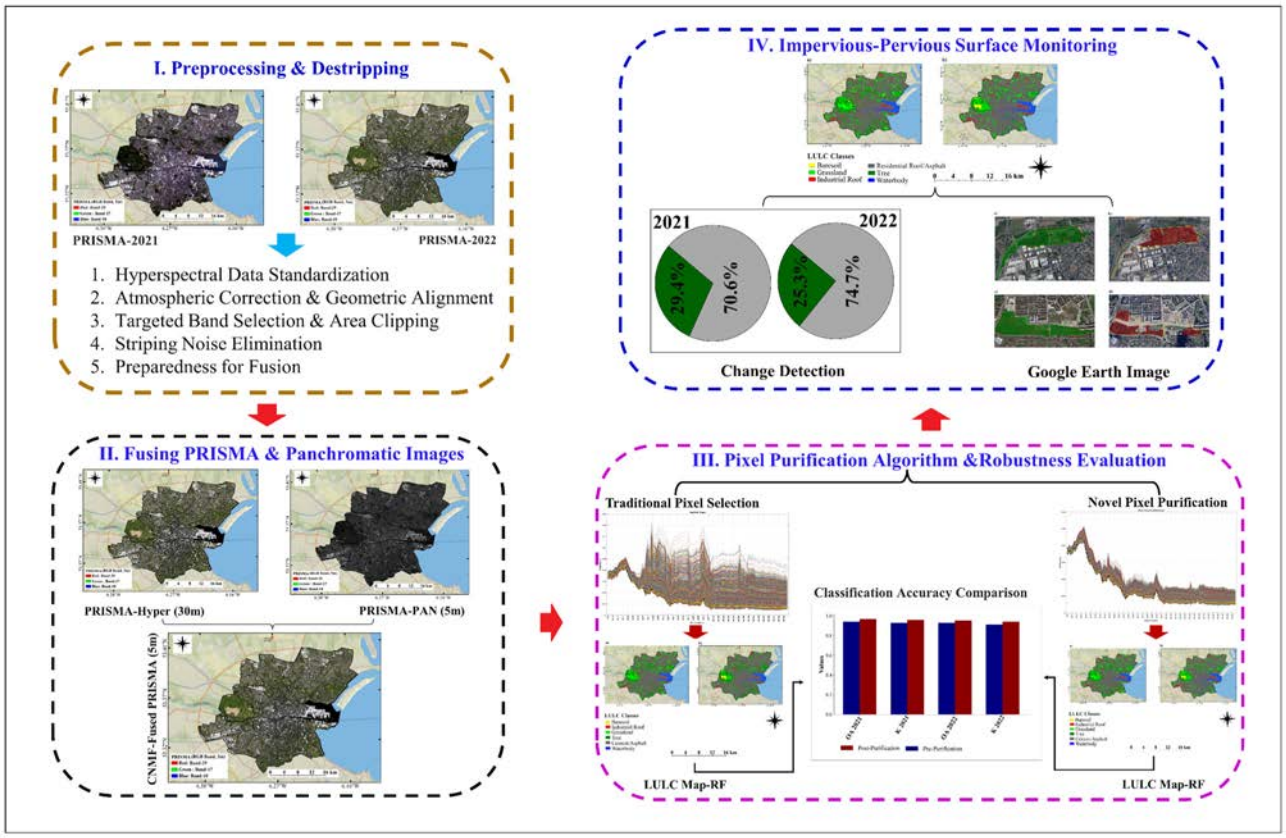

FIGURE 2. Schematic of adopted methodological framework for LULC and permeability change detection in the study area.

analysis. This conversion is facilitated by the ''prismaread'' software package, designed specifically for importing and transforming hyperspectral data cubes into analyzable formats [57], [58]. Subsequently, a critical selection process eliminates spectral bands affected by atmospheric absorption, especially those bands impacted by water vapor. This step is crucial for minimizing potential analysis distortions, ensuring that the data used for modeling accurately reflects the observed landscapes [53], [58], [59].

The next phase involves the identification and removal of striping noise from the bands. Striping noise, which presents vertical or horizontal inconsistencies in brightness across the image, can significantly degrade the quality and interpretability of the imagery. Through the application of a spectral-based regression model, these disruptions are effectively neutralized, thereby enhancing the HS data's overall quality in preparation for subsequent fusion. The criteria for band selection and the approach to noise mitigation are informed by the detailed methodologies documented in the literature, ensuring a thorough and effective preprocessing routine [60], [61], [62]. This structured approach to preprocessing PRISMA imagery is pivotal in maintaining the high standards of data quality required for accurate land cover and land use classification, ultimately contributing to the reliability of our study's findings.

D. FUSION OF PRISMA HYPERSPECTRAL IMAGERY WITH PANCHROMATIC IMAGE USING CNMF ALGORITHM

HSI is characterized by its high spectral resolution, capturing a wide range of the electromagnetic spectrum. This capability enables the precise and accurate identification of different LULC classes and surface objects [63], [64], [65]. However, HSI is limited by a trade-off between spectral and spatial resolution and the challenges posed by high-dimensional data management [13], [66], [67]. To address these limitations, we employ the Coupled Non-negative Matrix Factorization (CNMF) algorithm, a sophisticated unmixing technique for its efficiency in fusing PRSIMA hyperspectral imagery (HIS; 30m) and high-resolution panchromatic image (PAN; 5m) for years 2021 and 2022 [68]. This fusion is performed using the MATLAB-based ''Hyperspectral and multispectral data fusion toolbox'' [31]. The main concept behind CNMF approach is the principles of spectral unmixing and endmember extraction, crucial for remote sensing data processing [22], [68]. Nonnegative Matrix Factorization (NMF), the foundation of CNMF, decomposes a nonnegative matrix V into two nonnegative matrices, i.e. basis matrix (W) and coefficient matrix (H), where $V \approx WH$.This decomposition process is instrumental in separating mixed pixel spectra into distinct spectral signatures (endmembers) and quantifying their abundances [22], [30], [31], [69]. Furthermore, the

Vertex Component Analysis (VCA) facilitates this process by ensuring that the spectral unmixing adheres to the non-negativity constraint, thus enhancing interpretability [30], [70], [71], [72]. The comprehensive documentation of CNMF fusion algorithm is provided by [31].

E. LULC MAPPING USING RF

For the supervised classification of fused images $PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$, we employed the Random Forest (RF) algorithm, renowned for its efficacy and robustness in classification tasks. RF, an ensemble learning method, builds upon the concept of decision trees to enhance classification accuracy. By constructing multiple decision trees during training and adopting the mode of the classes of these trees for predictions, RF substantially improves upon the performance of a single decision tree [73], [74]. The fundamental appeal of RF lies in its use of multiple trees to achieve greater classification accuracy and stability. This is achieved through principles of bagging (Bootstrap Aggregating) and feature randomness split (*Mtry*), where RF introduces variety by training each tree on random data subsets and considering a random subset of features at each node for splitting. This methodology effectively reduces tree correlation, leading to a decrease in variance without a significant bias increase [33], [74]. One of RF's key strengths is its resilience against overfitting, courtesy of its mechanism that generates diverse trees using random subsets of observations and features. Additionally, RF's ability to handle high-dimensional data with correlated features makes it particularly adept for the complex datasets typical of LULC classification tasks. Critical hyperparameters influencing RF's performance include the number of trees (*Ntree*) and the number of features considered for splitting at each node (*Mtry*), with optimal values typically found to be between 100 to 500 for *Ntree* and the square root of the variable set for *Mtry* [33], [73], [75], [76]. Moreover, RF offers an internal validation mechanism through Out-of-Bag (OOB) error estimation, providing an unbiased prediction of model generalization error. This aspect is particularly beneficial for LULC classification, where the model's ability to accurately predict new results, previously unseen data, generalization, is crucial [75]. For this study, pixels were randomly selected from the fused images ($PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$), for the RF classification process to generate LULC maps from the fused images. The selected pixels were divided into training (70%) and validation/test dataset (30%). Classification was conducted across six land cover classes: residential roofs/asphalt, bare soil, grass, industrial roofs, trees, and water bodies. The RF model was implemented using Python with the hyperparameters set to $Ntree = 100$ and $Mtry = 12$ for $PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$ respectively.

F. AUTOMATIC PIXEL PURIFICATION ALGORITHM

The accuracy of LULC for accurate change detection, especially in the context of change detection, is dependent on the purity, integrity, and quality of the ML algorithm's training data [77], [78]. This type of process is necessary, as inaccurately chosen pixels, contaminated with mix-pixels and outliers, can substantially affect the classification's overall accuracy. Such contamination distorts the representation of land cover classes and hampers the ability to monitor changes effectively over time. By implementing a sophisticated approach to training pixel selection, this algorithm stands as a cornerstone for enhancing LULC classification outcomes across varied landscapes [79]. Introducing our automatic pixel purification algorithm significantly addresses these challenges, ensuring the selection of high-quality training pixels that correctly represent their respective classes. The purification algorithm of the training procedure can be explained in the following steps: (1) The initial step involves selecting representative pixels for each target class, such as bare soil, industrial roofs, grass, residential roof/asphalt, trees, and water bodies, from fused images ($PRISMA_{Fused,2021}$, $PRISMA_{Fused,2022}$). This careful selection aims to capture the broad spectral diversity inherent in each class across different imaging conditions, laying a foundational dataset for subsequent purification processes. (2) Following selection, we extract the spectral signatures for each pixel within the classes from the fused images. (3) At the core of our algorithm is utilizing the feature space and spectral similarity space by calculating Euclidean Distance (ED) and Spectral Angle Measure (SAM) for all pixel classes. ED is a measure between k-dimensional spectral space pixels and can be derived from the equation reported in Table 1 [80]. ED is calculated as a geometric vector-distance between pixels in k-dimensional spectral space compared to the pixel class centroid (average pixel class/endmember) [80], [81]. This distance can enable us identify pixels that are significantly different from typical class signatures, suggesting possible outliers or mixed pixels. The ED value was calculated from spectral reflectance of image product from fused images ($PRISMA_{Fused,2021}$, $PRISMA_{Fused,2022}$), which has values ranging from 0-1, which can be considered as normalize.

Then we calculate the SAM for these six classes using the equation in Table 1 [80], [82], [83]. SAM index, which is a widely used method to quantify the preservation of spectral information of each pixel, measures the spectral similarity/dissimilarity (or distance) by calculating the angle between two vectors, representing the pure-spectra (endmember) [80], [83]. A smaller angle (SAM) indicates high spectral similarity, crucial for confirming that pixels belong to their assigned classes. This angle measurement complements the ED by focusing on the angular difference rather than spatial separation, offering a nuanced view of spectral data integrity. Both metrics are crucial in our automatic approach for pixel purification, allowing us to identify, isolate and remove non-representative pixels from given dataset accurately and precisely. It is important to note that SAM is less sensitive to absolute pixel values (shadow effects, water vapor, etc.). On the other ED is highly sensitive to absolute

TABLE 1. Mathematical formulation for automatic pixel purification algorithm and LULC accuracy.

No	Metric	Equation	References
		Pixel-Purification	
1	ED	$ED_{S_iS_m} = \|S_i - S_m\|^2$	[80, 83]
2.	SAM	$SAM_{S_iS_m} = cos^{-1}\left(\frac{\langle S_i, S_m\rangle}{\|S_i\|^2\|S_m\|^2}\right)$	[83, 115, 116]
	where S_i is spectral signature for each specific pixel class and S_m is mean spectral signature of the given class.		
		Classifications Accuracy Assessment Criteria	
5.	K	$\kappa = \frac{T\sum_{i=1}^{c} l_{ii} - \sum_{i=1}^{c} l_{i+}l_{+1}}{T^2 - \sum_{i=1}^{c} l_{i+}\, l_{+1}}$	[93, 95, 97]
4.	OA	$OA = \frac{\sum_{i=1}^{c} l_{ii}}{T}$	[34, 93, 95, 117]
6.	UA	$UA = \frac{l_{ii}}{l_{+i}}$	[92, 93, 95]
	Where c is number of classes, l_{ii} is values along the diagonal of the confusion matrix (number of pixels of class i predicted in class i), i.e., pixel accurately classified, l_{i+} is the total number of truth values belonging to class i (total pixel in row i), l_{+i} is the total number of predicted values belonging to class i (total pixel in column i), and T is total test pixel number.		

spectral distance in k-dimensional spectral space. This automatic approach enhances the training dataset quality, crucial for accurate LULC classification. Further theoretical background of ED and SAM can be found in [80], [82], and [83], and the mathematical formulation of these metrics is provided in Table 1.

(4) Next, we implement a Gaussian thresholding function (G) defined by equation 1, employing a 95% confidence interval as the optimal interval for both ED and SAM [84], [85], [86], [87]. Pixel values that simultaneously fall outside this confidence interval for both ED and SAM metrics are considered potential outliers, mixed pixels and pixels thar are mistakenly selected. These outliers are excluded from both the training and test datasets. This method meticulously identifies pixels that are outside the n-dimensional feature space, representing a higher spectral similarity angle (SAM) with a higher spectral distance (ED), confirming that these pixels do not belong to the particular class and should be considered as outliers or mixed pixels. By effectively eliminating these outliers, our algorithm purifies the training dataset, making it indispensable in the preparation of training data for machine learning models. The exclusion of these non-representative pixels from the training set ensures that the machine learning algorithms are trained on the most accurate and representative data possible, enhancing the overall accuracy and reliability of the LULC classification outcomes.

$$G(\mathrm{X} : \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mathrm{X}-\mu}{\sigma}\right)^2} \quad (1)$$

where X is the variable, μ and o are the mean value and standard deviations of SAM and ED for individual class.

Moreover, a critical feature of our algorithm is the derivation of the mean spectral signature for each target class (μ), facilitating the selection and spectral analysis of pixels from fused PRSIMA images ($PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$). This derived mean spectral signatures (μ), representing the 'endmember or pure spectral profiles of the respective classes enhances the accuracy of delineating each land cover type. Endmembers refer to the pure spectral signatures characteristic of specific materials or objects, essential for the precise identification and classification of land cover types [88]. Our pixel purification algorithm, leveraging the concept of pixel-wise training, not only mitigates the influence of mixed pixels and outliers but also significantly improves the classification process's accuracy. By concentrating on the spectral signatures of training pixels, this methodical approach ensures minimal impact from common classification challenges, marking a substantial enhancement in supervised classification accuracy [89], [90].

G. PIXEL PURIFICATION-BASED LULC CLASSIFICATION ASSESSMENT

To validate the effectiveness of our pixel purification algorithm in enhancing land use and land cover (LULC) classification accuracy, we conducted a comprehensive assessment utilizing the RF algorithm across datasets from $PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$. This evaluation focused on comparing classification accuracies derived from

purified and original fused images, underscoring the robustness of our approach in various landscapes and datasets [91], [92], [93], [94]. After the implementation of the pixel-purification algorithm, approximately 8.6% of the total pixels were removed from the PRISMA Fused 2021 dataset, and 7.4% from the PRISMA Fused 2022, and verified /matched by very high-resolution (VHR) Google Earth imagery. Of the remaining pixels, 70% were allocated for the training set, and 30% were designated for validation. This distribution ensures a robust training process while providing a representative validation set to assess the algorithm's performance accurately. The RF algorithm was again performed on purified tarin pixels and LULC maps generated from both purified and non-purified pixels underwent validation using a validation dataset.

(code available: https://github.com/PayamSJ/Automated-Pixel-Purification-). The results were compared to underscore the robustness of our algorithm. Utilizing the test dataset, an error matrix, commonly known as a confusion matrix, was constructed to enable a comparative analysis between the classification outputs and the actual test data [93], [94], [95]. Following the construction of the error matrix, various accuracy metrics were calculated to quantitatively assess the classification performance, including Overall Accuracy (OA), Kappa Coefficient (κ), and User Accuracy (UA) [93], [95], [96]. OA refers to the proportion of correctly classified pixels across all categories, calculated by dividing the sum of the diagonal elements by the total number of pixels in the reference data. This metric offers a general measure of the classifier's performance across all classes [92], [93], [95]. UA, measures the accuracy from the map user's perspective, calculated for each class as the number of correctly classified pixels in a category divided by the total number of pixels classified into that category. It reflects the probability that a pixel classified into a given category accurately represents that category on the ground [93], [95]. Finally, κ, is a statistic that measures overall accuracy while accounting for chance agreement. It compares the observed accuracy with the expected accuracy if classifications were made randomly, offering insight into the precision of the classification beyond random chance [92], [93], [95], [97]. The mathematical formulations for these accuracy assessment metrics are detailed in Table 1.

H. PERVIOUS AND IMPERVIOUS SURFACE CHANGE DETECTION

To address the pivotal challenge of quantifying LULC transformations, especially the critical transition between pervious and impervious surfaces, our study implements a differential image analysis technique. This approach, based on pixel-by-pixel comparison, accurately compares corresponding pixels of $PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$ reclassified images. This differential method, based on subtraction, illuminated areas of change, thus enabling the quantification of LULC transformations, with each pixel value indicating the magnitude of change. Ensuring precise alignment of images was a preliminary step, confirming that each pixel across the temporal spectrum corresponded to its exact geographical location. This technique is celebrated for its effectiveness in revealing temporal changes, offering a precise means to observe and measure land cover alterations [98], [99], [100]. The extraction and categorization of impervious and pervious surfaces were based on the initial six LULC classes, categorizing industrial and residential roofs/asphalts as impervious surfaces due to their water-resistant nature, which contributes to surface runoff and reduced water infiltration. Conversely, bare soil, grasslands, and trees were recognized as pervious surfaces, indicative of their ability to absorb and filter water, thus supporting natural hydrological processes [101]. Water bodies are excluded from this binary classification due to their distinct characteristics and our study's focus on permeability changes resulting from urban development [102]. By performing a pixel-based comparison within the differential image, we quantified the change from previous to impervious surfaces, calculating the net change in surface permeability across the study area.

A thresholding procedure followed image differencing to distinguish significant from negligible changes, concentrating on transformations indicative of a conversion from previous to impervious surfaces. The pixel-purified based methodology, i.e. the novel contribution of our study, refines change detection accuracy by reducing the inclusion of mixed pixels that often obscure the true scale of land cover transformations [102]: Let's assume the derived six classes from classification are:

$$L = \{l_1, l_2, l_3, l_4, l_5, l_6\} \tag{2}$$

where, l_1 = bare soil, l_2 = industrial roofs, l_3 = grasslands, l_4 = trees, l_5 = residential roofs/asphalts, l_6 = water Bodies, and the classification derived from $PRISMA_{Fused,2021}$, and $PRISMA_{Fused,2022}$ are C_{2021} and C_{2022}, respectively.

Reclassifying C_{2021} and C_{2022} into pervious (P) and impervious (I) surfaces, excluding water bodies l_6 generates R_{2021} and R_{2022}:

$$R_{year} = \{P, I\}, \quad \text{and} \tag{3}$$

$$P = \{l_1, l_3, l_4\}, \quad I = \{l_2, l_5\} \tag{4}$$

Performing the differential analysis (D) between R_{2021} and R_{2022} identifies and quantifies changes from pervious to impervious surfaces :

$$D(R_{2021}, R_{2022}) = \Delta R \tag{5}$$

where, ΔR represents the net change in land cover from pervious to impervious surfaces.

Finally, quantifying the extent of change $(P \rightarrow I)$, from 2021 to 2022, assesses the transformation of permeability.

To complement and validate the quantitative analysis, we conducted a visual inspection using high-resolution Google Earth imagery from 2021 and 2022. Area's indicative

of substantial change, as identified through differential analysis, were visually examined to confirm land cover transformations. This step involved overlaying computed changes onto Google Earth images, providing a direct, intuitive evaluation of urban expansion and its impact on surface permeability.

III. RESULTS

A. RESULTS FROM FUSION IMPLEMENTATION AND VISUAL EVALUATION OF ALGORITHM

The implementation of the CNMF algorithm for image fusion has yielded significant enhancements in the spatial resolution of PRISMA hyperspectral images for the years 2021 and 2022, transforming them into high-resolution imagery at 5m resolution ($PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$). In addition, the CNMF fusion algorithm has demonstrated its ability to significantly enhance spatial resolution without sacrificing spectral fidelity of the original image. The primary focus of employing CNMF was to elevate the quality of hyperspectral images, thus offering an improved tool for analyzing land cover changes, particularly in distinguishing between pervious and impervious surfaces within Dublin's urban landscape. The impact of the CNMF fusion process is visually articulated through Figures 3 (a-d) and 4 (ad), providing a comprehensive before-and-after comparison. Providing a broad view of the study area, Figure 3 (a-d) illustrates the extensive impact of the fusion process. Figure 3a and Figure 3c display the $HSI_{PRISMA,2021}$, $HSI_{PRISMA,2022}$, respectively, showcasing the baseline spectral detail pre-fusion. In contrast, Figures 3b and 3d highlight the enhanced fused images ($PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$) for the corresponding years, showcasing a noticeable improvement in spatial resolution. The fused images, as evidenced in Figures 3b and 3d, demonstrate a significant improvement in spatial resolution, retaining the spectral information vital for LULC analysis but with heightened clarity and definition at a finer 5m scale. This improvement is particularly notable in urban areas, where the precise delineation of impervious surfaces like buildings, roads, and industrial complexes is crucial. The enhanced resolution allows for a more accurate identification of land cover changes over the study period. In areas characterized by dense vegetation, the enhanced resolution of the fused images reveals greater detail in trees and grasslands.

Zooming into specific locales for an in-depth analysis, Figure 4 (a-d) provides a detailed visual comparison that emphasizes the CNMF algorithm's effectiveness in enhancing image quality. The figures illustrate a marked contrast between the pre-fusion (Figure 4a and Figure 4c) and post-fusion (Figure 4b and Figure 4c) images, with the latter exhibiting sharper urban features and more distinct delineation of land cover types due to the spatial resolution enhancements. The spatial clarity in the fused images reveals distinct outlines of buildings, greatly enhancing the differentiation of built-up areas (Urban Core: yellow box in figure 4). The fusion algorithm's impact is vividly illustrated along the water channels and Liffey River, where the boundaries of these water bodies are captured with enhanced quality (water bodies: blue box in figure 4). This distinction is crucial for monitoring waterbody dynamics and assessing urban encroachment along riverbanks, facilitating targeted environmental conservation efforts. Furthermore, the fusion algorithm's efficacy is highlighted in the refined representation of vegetative cover (grassland and trees: green box). The enhanced detail aids in accurately classifying pervious surfaces, offering insights into the role of urban green spaces in maintaining ecological balance within the cityscape.

B. LULC MAPPING USING UN-PURIFIED PIXELS AND RF ALGORITHM

LULC classification was performed on the derived fused images ($PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$) employing the RF algorithm, known for its effectiveness in managing complex multispectral data and mitigating overfitting. The classifier, initially using un-purified training data, aimed to identify six critical land cover classes linked to urban permeability: bare soil, industrial roofs, grasslands, trees, residential/asphalt roofs, and water bodies. This process, executed with carefully chosen hyperparameters, resulted in the temporal LULC maps shown in Figure 5.

Figure 5 illustrates the initial LULC distribution within Dublin for the respective years, revealing instances of mixed pixels indicative of the need for further purification. The LULC map (Figure 5a; C_{2021}), indicates well-represented land cover classes but with notable misclassifications along water bodies, where pixels of water, grassland, and residential roofs/asphalt intermingle, possibly due to spectral similarities or shadows effect. Such misclassifications underscore the presence of mixed pixels, which may confuse our understating of the built environment. The 2022 map (Figure 5b; C_{2022}), exhibits similar challenges, with certain areas reflecting a mix of land cover types, suggesting the presence of mixed pixels. In densely built-up areas, blue pixels (water) erroneously intersperse within gray regions (residential roofs/asphalt), while confusion between red (industrial roofs) and gray areas highlight potential spectral confusion due to the intricate interaction of light, shadows, and urban structures. While these initial classifications have effectively highlighted the primary LULC classes within Dublin's urban landscape, the identified misclassifications point to the critical need for pixel purification to refine these outcomes. This refinement process is designed to minimize the impact of mixed pixels and enhance the delineation of land covers, which is particularly valuable in complex urban settings where proximity and spectral diversity can lead to spectral mixing.

C. LULC MAPPING USING UN_PURIFIED PIXELS AND RF ALGORITHM PIXEL PURIFICATION AND LULC MAPS FROM PURIFIED PIXELS

As described in the methodology section (2.2.4), we extracted spectral signatures from selected pixels within six landcover

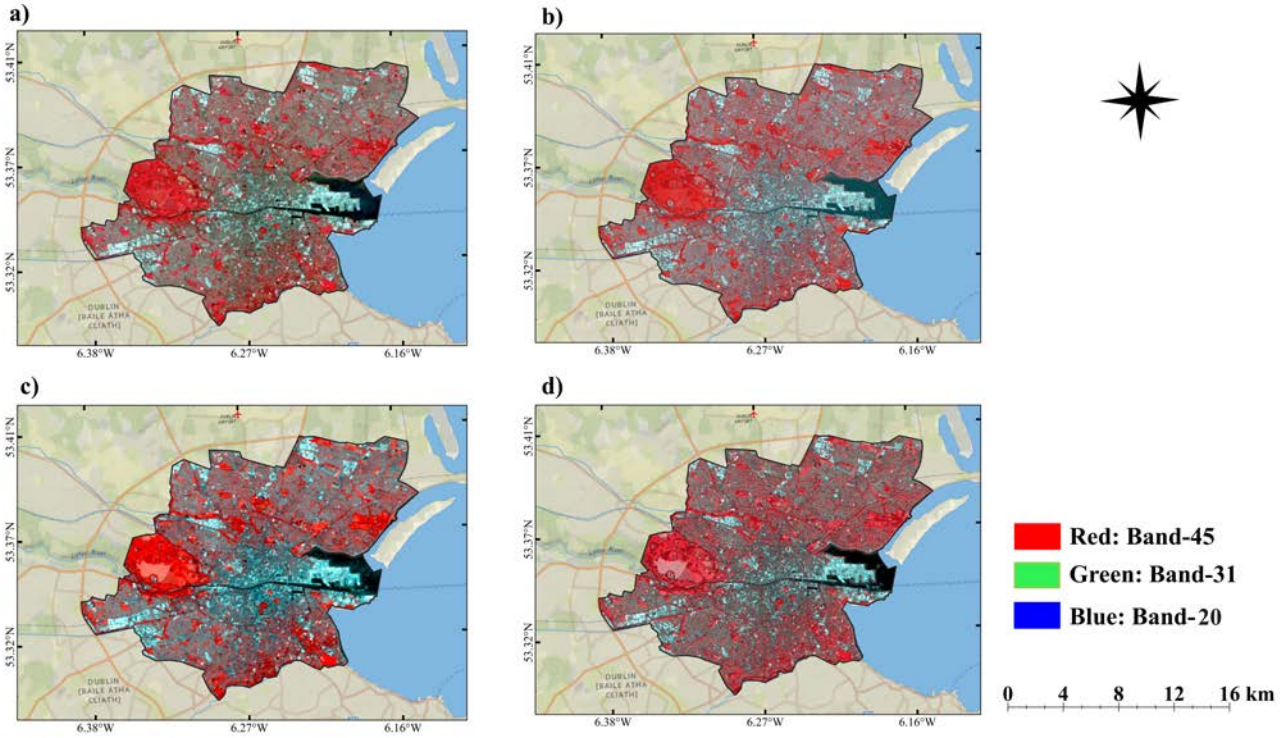

FIGURE 3. CNMF image fusion performed on $HSI_{PRISMA,2021}$, $HSI_{PRISMA,2022}$ of Dublin, showcasing the enhanced resolution of urban and natural landscapes for the years 2021 (first row) and 2022 (second row). Panels a and c represent the original hyperspectral images before fusion, capturing the baseline spatial detail. Panels b and d illustrate the results post-fusion ($PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$), highlighting the improved clarity and definition afforded by the CNMF algorithm.

classes: bare soil, grass, industrial roofs, trees, residential roofs/ asphalt, and water bodies. The selection and refinement of spectral signatures across the six LULC ensured that only the most representative training pixels were employed. The automatic pixel purification algorithm utilizing spectral similarity angle (SAM) and spectral distance (ED) in n-dimensional space, considering implied confidence interval (95%), effectively identified the mistakenly selected pixels as outliers and removed them from the analysis, thereby enhancing the quality of our training set. This innovative and automatic approach, detailed in our methodology, has yielded a set of refined spectral signatures that have been cleaned of outliers and mix-pixels, as demonstrated in Figure 6 (a-f).

The transformation from pre- to post-purification spectral signatures is particularly stark. For instance, the spectral signatures of water bodies prior to purification (Figure 6a, representing for three LULC classes) display substantial variance, with certain reflectance values peaking indicative of the presence of mixed pixels and outliers. Post-purification, Figure 6b shows a notable reduction in this variance, resulting in a smoother, more consistent reflectance curve that closely mirrors the 'endmember' profile for water bodies, as evidenced by the yellow points on the dark line. Figure 6c presents the spectral signatures for trees before purification, where the spread of data points suggests a mix of spectral responses, potentially from the heterogeneous nature of vegetation or surrounding material contamination. After purification, Figure 6d demonstrates a clear smoothing of spectral responses, a narrowing of the spread, and an alignment of data points around a central trend, highlighted by a yellow-dotted line. This refined clarity accentuates the distinct spectral characteristics of trees, thereby improving their identifiability in the classification process. Similarly, Figure 6e illustrates the spectral signatures for residential roofs/asphalt prior to purification, with significant spread indicating potential contamination. The post-purification spectral signatures shown in Figure 6f reveal a marked reduction in variability, with the average spectral line depicting a clear and distinct profile, essential for accurately classifying urban materials, key to assessing impervious surfaces in urban analysis. Overall, the pixel purification process has systematically enhanced the training data quality for each LULC class, minimizing the impact of mixed pixels and outliers and, consequently, reinforcing the accuracy classification process. This detailed and careful approach is anticipated to result in more precise LULC change detection, critically important for monitoring urban development and understanding its ecological ramifications.

Furthermore, the Principal Component Analysis (PCA) analysis was performed to visually appraise the effectiveness of automated pixel purification algorithm in improving the differentiation between land cover types. PCA excels at maximizing spectral variance, thereby aiding in the discrimination of land cover types within a multidimensional

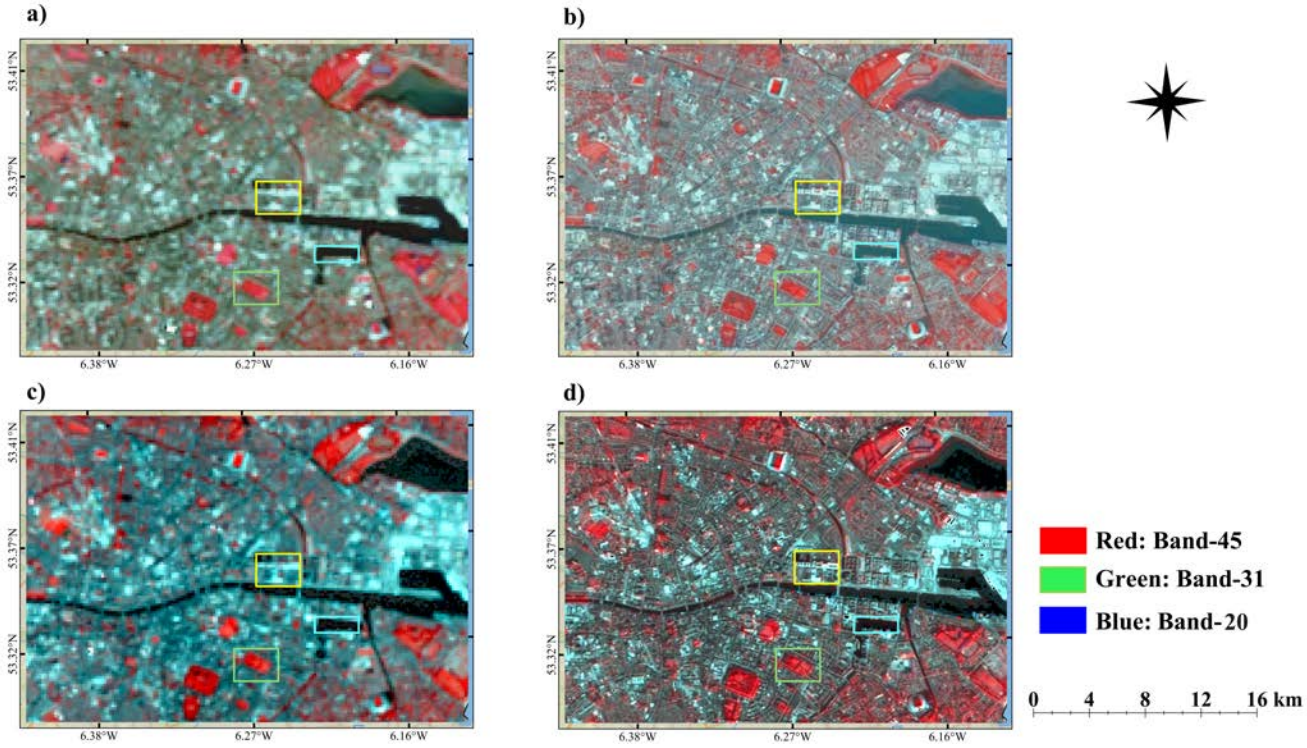

FIGURE 4. Detailed zoomed-in views of the CNMF image fusion process applied to PRISMA hyperspectral imagery of Dublin. Panels a and c provide a closer look at the original imagery before fusion, while panels b and d display the enhanced resolution post-fusion. The yellow box emphasizes urban areas, the blue box delineates water bodies, and the green box focuses on vegetated regions, underlining the spatial resolution improvements across different land covers.

component space. By converting the original spectral data into orthogonal principal components, PCA enabled the visualization of class separability, which may be obscured in the raw spectral data. The results of the PCA analysis (PC1, PC2, and PC3), post-pixel purification of $PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$, are displayed in Figure 7 (a,b) and Figure 8 (a,b), respectively, offering a clear perspective on the enhancement of class differentiation due to the purification process. Figure 7a illustrates the PCA scatter plot for the $PRISMA_{Fused,2021}$ image prior to pixel purification, indicated a certain overlap among classes and spectral variability that could potentially hinder accurate classification. There was a noticeable dispersion of classes, with boundaries between industrial roofs, residential roofs/asphalt, and bare soil blending. Contrastingly, Figure 7b, which represents the post-purification PCA scatter plot for the same image, demonstrates a substantial improvement in class clustering. The spectral points for water bodies, for example, became more tightly packed, suggesting a reduced risk of misclassification. Trees and grasslands displayed clear separations, reflecting a more uniform spectral signature within each class, post-purification.

The $PRISMA_{Fused,2022}$ image PCA plots further validate our pixel purification algorithm's effectiveness. Figure 8a presents the pre-purification PCA plot with noticeable overlaps among LULC classes, mirroring earlier observations. Notably, the spectral overlaps among industrial roofs, residential roofs, residential roofs/asphalt, and bare soil classes present classification challenges due to their spectral similarities. Figure 8b highlights the algorithm's robust performance, where the post-purification PCA scatter plot reveals significantly enhanced class separability. This is highlighted in the water bodies class, where a noticeably closer clustering illustrates a successful purification, reducing the risk of misclassification and echoing the benefits seen in the previous year's dataset. Such consistent improvements across different datasets underscore the algorithm's reliability and potential for broad application in urban LULC monitoring. Overall, the PCA results affirm the pixel purification algorithm's ability to improve spectral separability for LULC classification. This enhanced delineation in the PCA space lays the groundwork for more accurate and dependable classifications, a crucial aspect in monitoring urban land cover dynamics and their environmental impact.

Employing our automatic pixel purification algorithm to enhance the spectral signatures of the training pixels, we proceeded to LULC classification of the purified pixels from the $PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$, images. This refined classification, essential for accurately detecting shifts between impervious and pervious surfaces, leveraged the RF algorithm enhanced by high-quality training and test data. The resultant LULC maps, drawn from the purified clas-

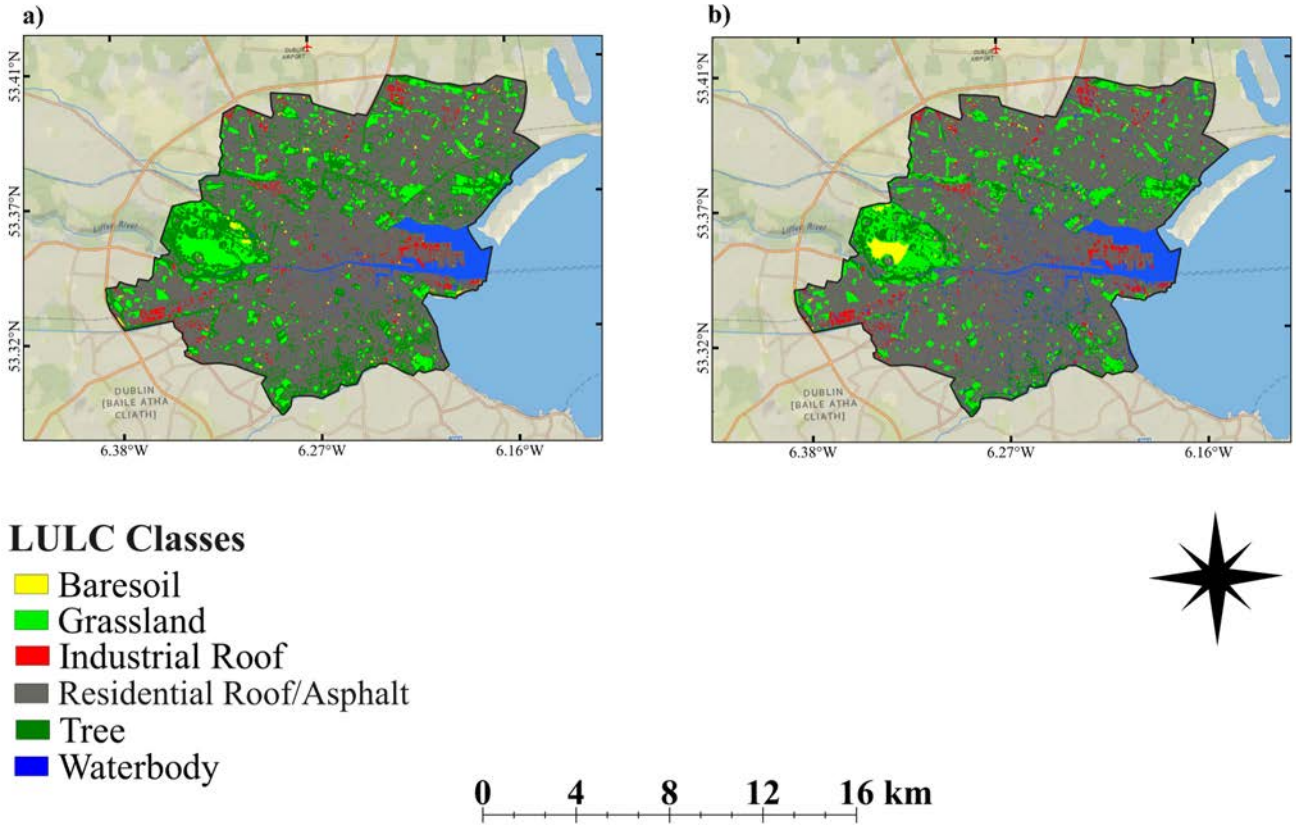

FIGURE 5. LULC maps derived from RF algorithm: a) LULC map generated pre-purification for 2021 (C_{2021}) and b) and LULC map generated pre-purification for 2022 (C_{2022}).

sification (Figure 9), offered a reliable foundation for the assessment of urban permeability changes. The classification outcomes elucidate the urban landscape's transformation, which will be explored in detail with interpretations of the classification maps for 2021 and 2022. Figure 9 showcases the post-purification RF classification maps; panel a illustrates the LULC map where the delineation of grassland and trees is now apparent, water bodies are well-defined, and the distinction among land covers has improved. Panel b reflects the 2022 LULC map with a discernible expansion in bare soil areas, suggesting land use changes, construction activities, over the one-year span. The enhanced classification delineates the land covers more precisely, showing notably clearer water bodies and a better-resolved differentiation between industrial roofs and residential roof/asphalt areas,

In a closer examination, Figure 10 (a-d) compares the LULC maps for a zoomed-in area of Dublin, illustrating the pixel purification effects for 2021 and 2022. Figures 10a and b show the LULC map before and after purification, where the former displays spectral mixing and the latter visually demonstrates modified, refined and improved classification accuracy with more homogeneous land cover representations. The same for Figures 10c and d for the 2022 LULC maps.,. These maps visually verify the pixel purification's effectiveness on both images 2021 and 2022.

D. ACCURACY ASSESSMENT AND PIXEL-PURIFICATION ALGORTIHM PERFORMANCE EVALUATION

The robustness of our pixel purification process was rigorously evaluated through a detailed accuracy assessment using a comprehensive validation dataset (refer to purification section 2.1.7). Utilizing the confusion matrix, we computed key metrics such as OA, κ, and UA to quantitatively benchmark the classification performances before and after purification [93], [95]. Figures 11, and Figure 12, graphically and quantitatively represent these accuracy enhancements, where the post-purification classification demonstrates notable improvements. For instance, the pre-purification OA stood at an impressive 94.04%, with a κ of 92.60%, yet post-purification, these figures rose to 96.69% and 95.91%, respectively. The trend continued in 2022, with initial OA and κ values of 92.80% and 91.17%, which increased to 95.05% and 94.04% following purification (Please refer to Table S1).

The UA percentages depicted in Figure 12 for various land cover classes further substantiate the pixel purification's efficacy. Remarkably, for the soil class in 2021, the UA achieved a perfect score of 100% post-purification, rising from a pre-purification UA of 96.49% (Please refer to Table S2). The residential roofs/ asphalt class witnessed a significant UA improvement, ascending from 92.07% to 95.50% in 2021 and from 89.36% to 93.93% in 2022. Similarly, the

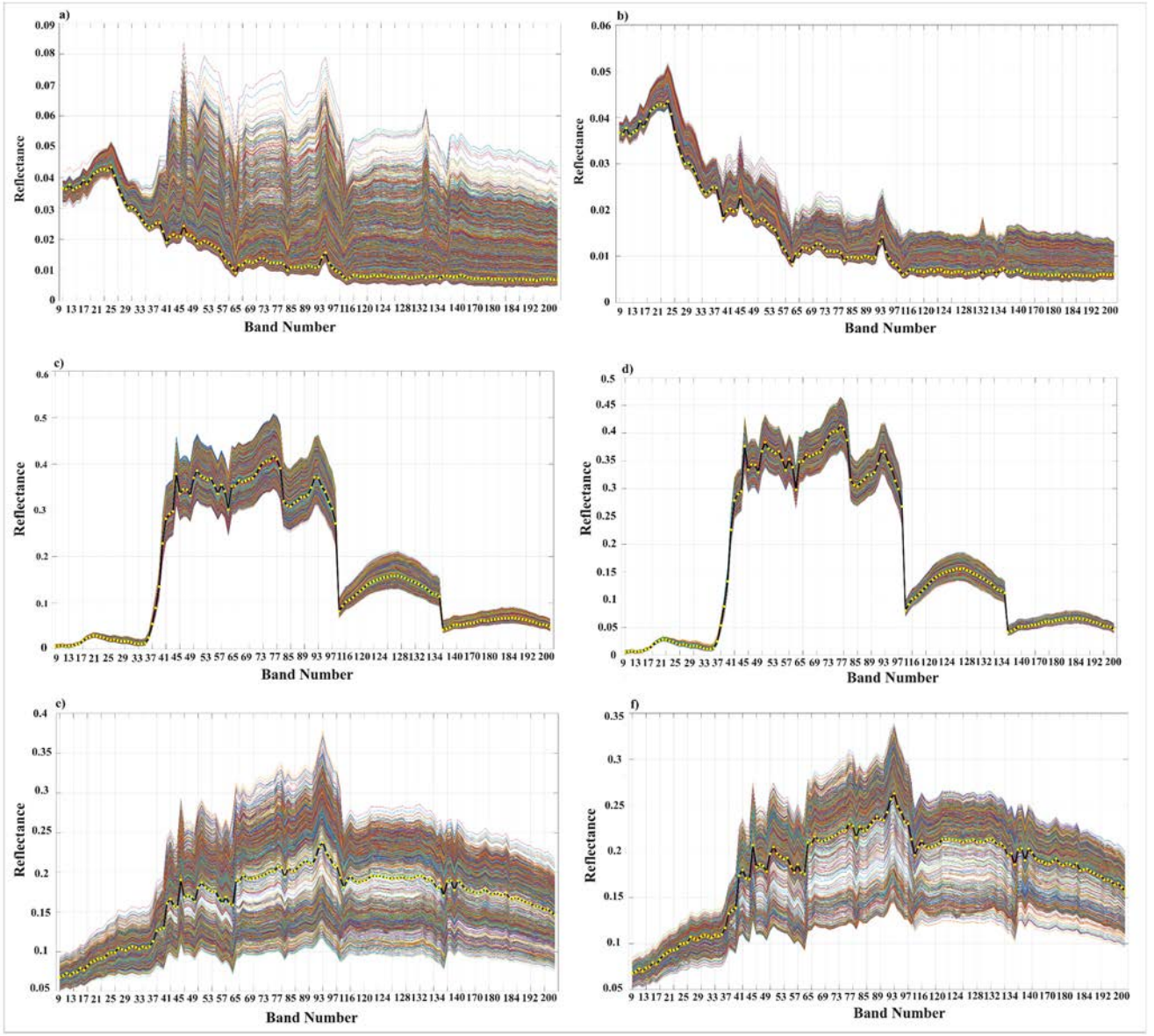

FIGURE 6. Pixel purification algorithm's impact on spectral signatures for water bodies (a,b), trees (c,d), and residential/asphalt (e,f) in Dublin. Panels a, c, and e present the original spectral profiles before purification, while panels b, d, and f display the refined signatures after purification. Yellow-dotted lines indicating the mean spectral signature.

tree class displayed increased UA scores post-purification, indicative of the purification process's capability to refine vegetative cover differentiation. The grassland areas, too, experienced an uplift in UA, albeit modest, from 90.47% to 93.07% post-purification in 2021, with a stable performance in 2022. Most strikingly, water bodies' UA improved significantly, from 94.67% to 99.74% in 2021 and from 93.82% to 98.66% in 2022, showcasing the algorithm's refined precision in classifying hydrological features post-purification (Please refer to Table S2).

These substantial improvements across nearly all land cover classes affirm the pixel purification's positive impact. Such enhancements are particularly crucial for classes that initially exhibited lower accuracy, highlighting the process's significance in the remote sensing classification workflow.

E. IMPERVIOUS-PERVIOUS SURFACE MAPPING AND CHANGE DTECTION

Following the enhanced LULC classification facilitated by pixel purification, we advanced to delineate impervious and pervious surfaces from the classified PRISMA fused images of 2021 and 2022. We reclassified the original six land cover classes, distinguishing between impervious surfaces (industrial Roofs and residential roofs/asphalts) and pervious surfaces (bare soil, grasslands, and trees), with water bodies

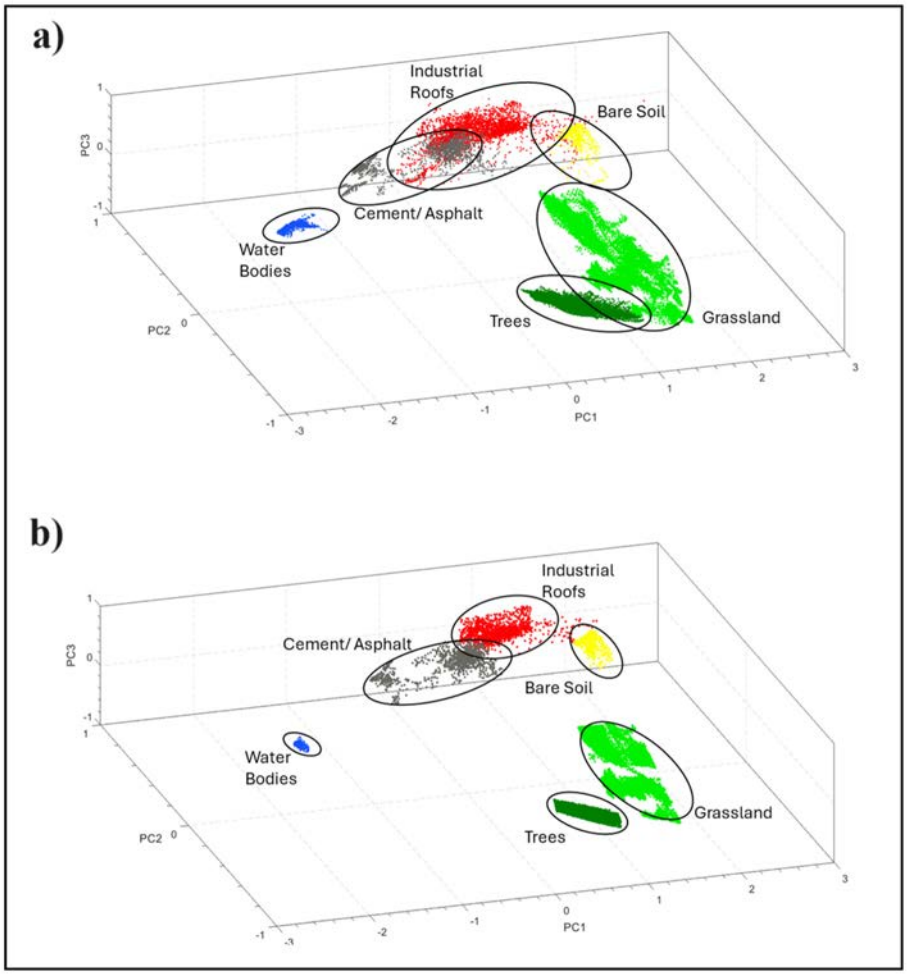

FIGURE 7. PCA analysis pre- and post-pixel purification for $PRISMA_{Fused,2021}$; Panel a): PCA scatter plot before pixel purification, Panel b): scatter plot after purification.

set apart. This categorization was integral to the assessment of urban permeability changes. We then employed a differential analysis, comparing reclassified images to detect and measure the conversion from pervious to impervious surfaces, indicative of urban expansion as previously discussed in section 2.1.8 (Equation 2-5). The outcome of this analysis is depicted in Figure 13, which illustrates the degree and direction of land cover changes in Dublin, revealing the areas that have transitioned from pervious to impervious between 2021 and 2022. Table S3 provides a quantitative summary of these land cover changes, indicating the extent of urban encroachment onto previously permeable surfaces. To enhance the robustness of our findings, we employed high-resolution Google Earth imagery for a detailed visual inspection, affirming the significant changes detected by our analysis (red boxes in Figure 13) and presented in Figure 14 (a-d). Figure 13 visually captures the evolution of Dublin's urban landscape, contrasting the spread of impervious and pervious surfaces over one year. The maps show a discernible increase in impervious areas, indicative of construction and development, resulting in a decreased pervious surface area, and permeability of surfaces and increased potential urban flooding in the region. Crucially, the red boxes emphasize regions that have undergone considerable change, where the proliferation of impervious surfaces indicates new urban development or the intensification of existing built-up areas. These changes suggest an ongoing trend of urban sprawl, with potential implications for the city's hydrological balance and ecological sustainability.

In Figure 14, the Google Earth images provide tangible evidence of the land cover changes between 2021 and 2022 for two selected regions (red boxes in Figure 13). The transformation from green to red outlines, representing the conversion of permeable to impermeable surfaces, validates the shifts we've identified using our algorithm.. The validation process facilitated by these Google Earth snapshots is pivotal, as it not only confirms the algorithm's efficacy in discerning the nuanced changes from pervious to impervious surfaces but also visually substantiates the urbanization trends identified within the study.

Figure 15 illustrates the overall pervious to impervious transformation, where the extent of green areas representing

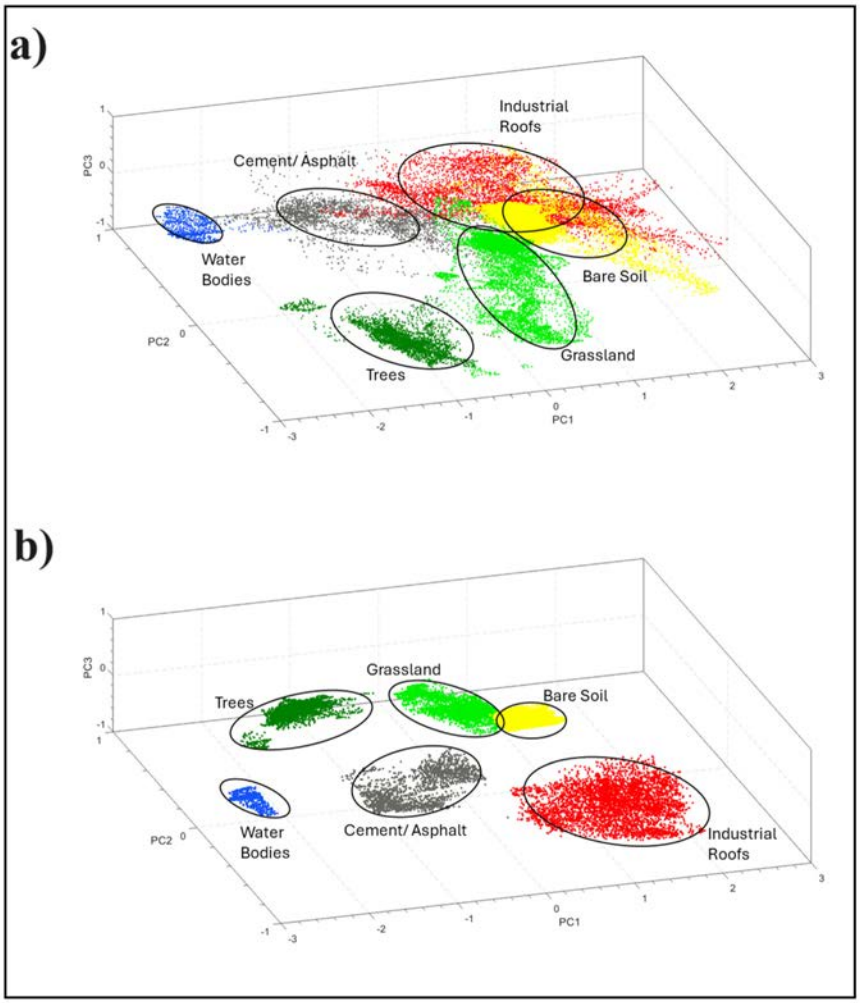

FIGURE 8. PCA analysis pre- and post-pixel purification for $PRISMA_{Fused,2022}$. Panel a): PCA plot before pixel purification. Panel b): PCA plot post-purification.

pervious surfaces like bare soil, grassland, and trees has diminished notably from 29.4% in 2021(Figure 15a) to 25.3% in 2022 (Figure 15b).

The results indicate that the pervious surface cover in Dublin decreased from 33.29 km^2 (133,1649 pixels) to 28.08 km^2 (1,123,443 pixels), a reduction of 4.08%. Conversely, impervious surface area increased from 79.96 km^2 (3,198,454 pixels) to 82.92 km^2 (3,316,992 pixels), an increase of 4.09% (please refer to Table S3). Overall, approximately 1.07 km^2 (1.35%) of Dublin has been converted to impervious cover in one year; assuming annual precipitation of 900 mm, this landcover change displaces an additional 1,000,000 m3 of rainwater. The spatial detail provided by this approach is needed to manage the impacts of landcover change, much of which occurs at scales that are too small to capture using conventional approaches. Moreover, the method developed here can provide the temporal resolution needed to capture rapid change and identify where it is concentrated.

IV. DICUSSION

A. IMPORTANCE OF IMAGE FUSION IN ENHANCING SPATIAL RESOLUTION

The incorporation of image fusion techniques in enhancing the spatial resolution of hyperspectral images is essential for achieving refined LULC classifications. The implementation of CNMF algorithm in our study has been paramount in enhancing the spatial detail of PRISMA hyperspectral imagery to a resolution of 5m, facilitating an intricate analysis of land cover transitions. Image fusion's role in boosting LULC classification is well-established [103], [104], [105], and our application of CNMF, a specialized unmixing algorithm, advances this domain by maintaining a delicate balance between spatial and spectral integrity [68], [106]. At a broader scale, the enhanced clarity in urban delineation underscores the fusion's contribution to improved classification accuracy, as indicated by.Ghassemian [23]. The CNMF algorithm's proficiency in safeguarding spectral data while enhancing spatial resolution was also evidenced in prior

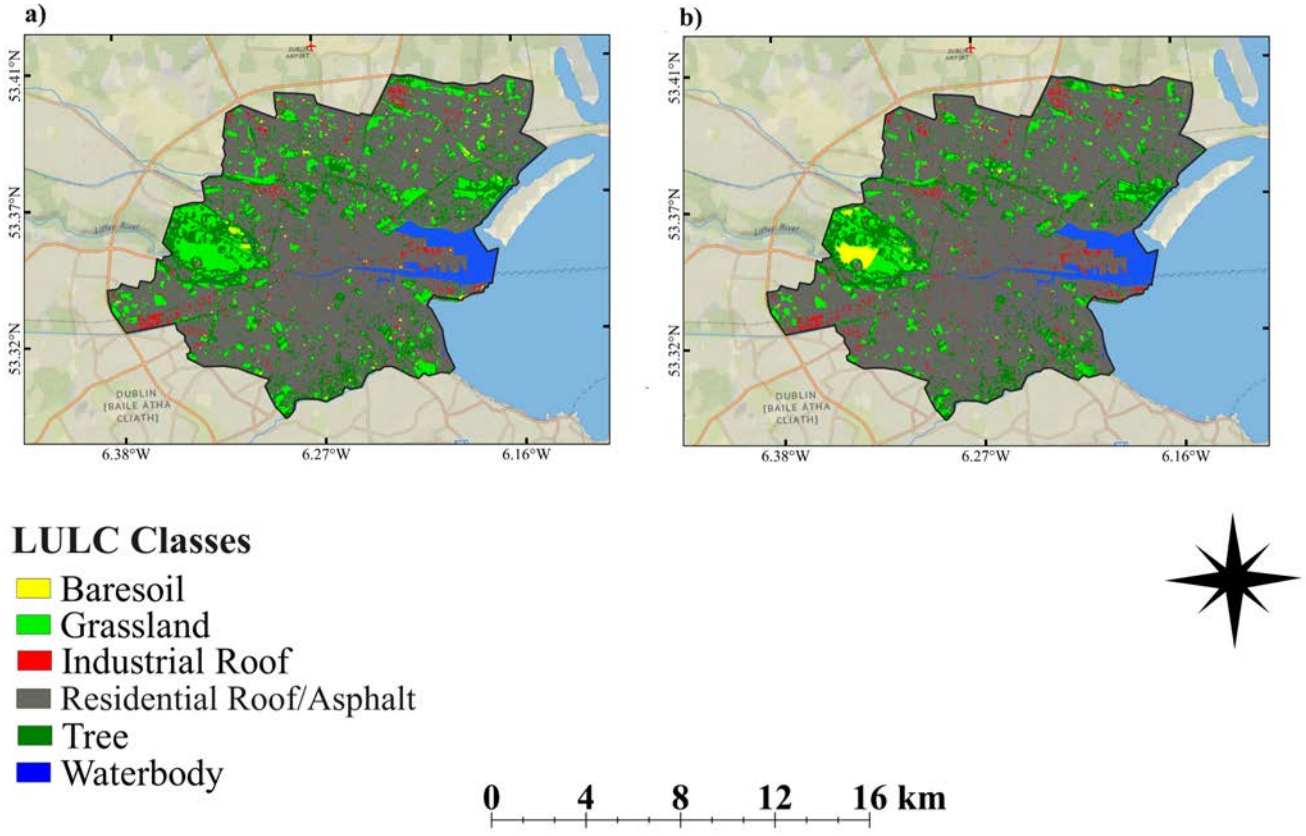

FIGURE 9. LULC classification post-pixel purification for 2021 (a) and 2022 (b).

researches [22], [68]. The fused images, unmarred by spectral distortions, provide unprecedented insights into the differentiation between urban and natural landscapes, a task that traditional hyperspectral imaging methods struggle with due to inherent spectral-spatial trade-offs and data dimensionality challenges. The detailed urban views shed light on the algorithm's finesse in portraying complex cityscapes where distinct building footprints and vegetative cover are now easily distinguishable, corroborating with Myint, et al. [107], on the importance of high-resolution data for detailed urban studies. Moreover, the detailed representation of water bodies and green spaces in Figure 6 parallels with the findings of Tewkesbury et al. (2015), highlighting the improved monitoring of various surface objects in urban environments. This level of detail is essential for environmental management and urban sustainability.

B. AUTOMATIC PIXEL PURIFICATION AND ITS IMPORTANCE IN LULC CLASSIFICATION

The CNMF algorithm showed a significant improvement in the spatial resolution of PRISMA imagery for 2021 and 2022. However, LULC classification is often evaluated by the accuracy of pixel representation within spectral datasets. Urban environments, with their inherent spectral confusion caused by mixed pixels and outliers, pose considerable classification challenges, often exacerbated in high-resolution datasets [108]. While fusion algorithms like CNMF enhance spatial resolution, they fall short in resolving the ambiguities of mixed pixel values, leading to spectral signatures that may have inconsistencies, as evidenced in the initial analysis of water bodies, trees, and residential roof/asphalt classes, undermining the accuracy and reliability that CNMF fusion.

Our study introduces a novel automatic pixel-purification algorithm that gaps this bridge by employing a spectral similarity angle (SAM) and spectral distance (ED) in n-dimensional feature space, specific confidence interval [86], [87], to accurately identify and eliminate mixed pixels, outliers, and erroneously selected pixels during the algorithm training phase to enhance the LULC accuracy of hyperspectral imageries like PRISMA. This purification is essential for the RF supervised classification algorithm, enhancing spectral purity for accurate change detection and LULC classification. The robustness of this process is demonstrated by the refined spectral signatures across land covers, which align more closely with true 'endmember' profiles, thus enhancing the accuracy of LULC classification [109]. The pixel purification process's effectiveness is underscored by its impact on PCA results, which clearly show enhanced spectral separability and reduced class overlap post-purification. This transformation is critical in urban settings, where accurate

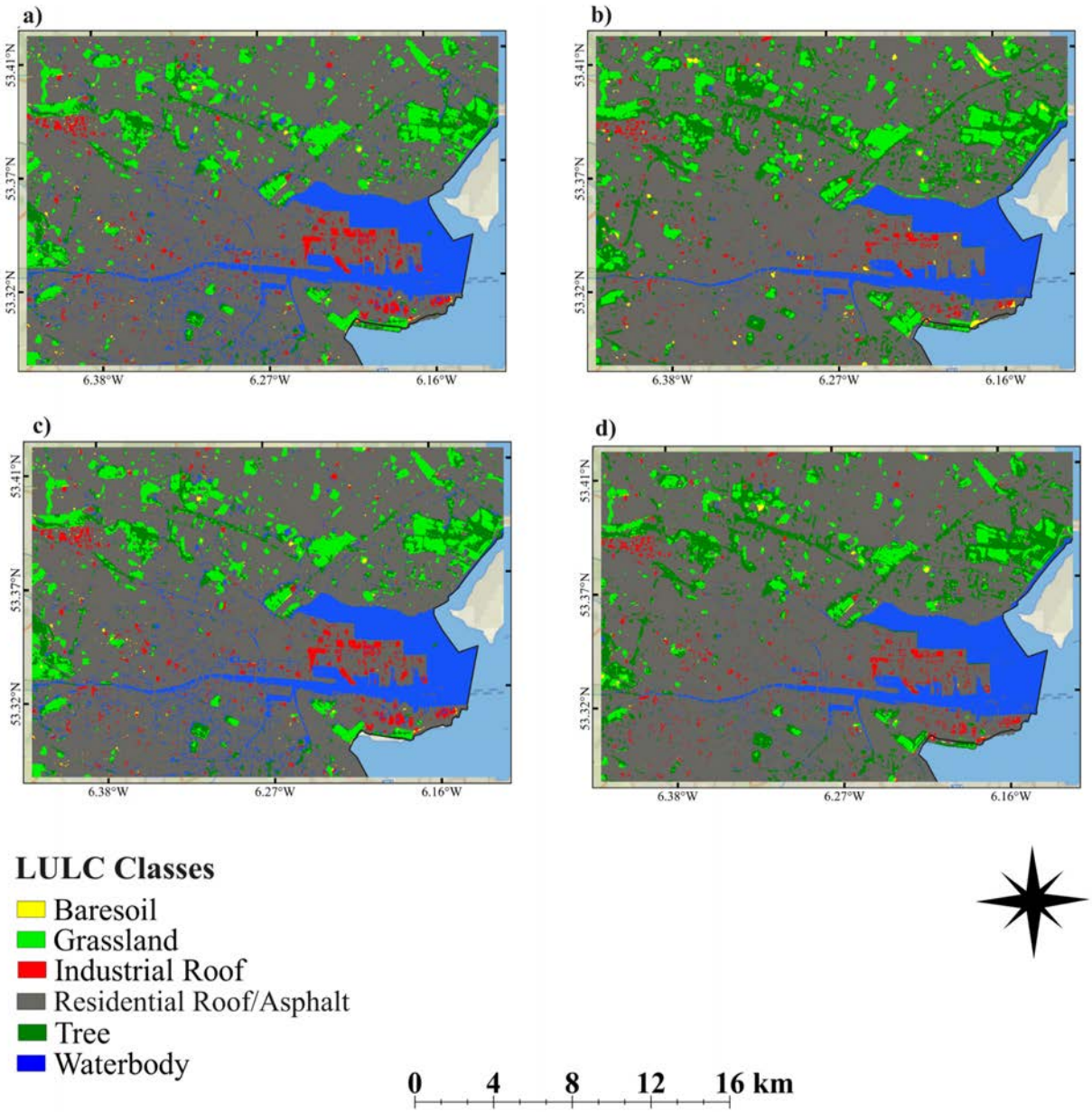

FIGURE 10. Comparison of LULC classification pre- and post-pixel purification for a zoom-in area, Panels a) and c) present the LULC classification before pixel purification for 2021 and 2022, respectively. Panels b) and d) show the classification post-purification.

classification is essential for effective urban planning and ecological monitoring. The clarified spectral data ensures that classes such as water bodies, grasslands, and urban materials are distinctly separated, reducing the risk of misclassification.

In the complex urban mosaic, where differentiation between land covers can be subtle, the clarity achieved through pixel purification is paramount. The reduced overlap and increased separability in post-purification PCA plots are critical for precise classification [110], confirming the indispensable role of pixel purification in achieving accurate LULC classification. Furthermore, the comparative analysis of LULC mapping from RF algorithm on both purified and non-purified datasets using performance metrics (OA, K, and UA) highlights the significant improvements in classification accuracy brought about by our approach. Initial challenges observed in the RF algorithm's performance due to mixed pixels and spectral similarities are effectively mitigated post-purification,, affirming that improvements and enhanced spectral purity in training data quality can lead to enhanced classification accuracy and minimize the class ambiguities [79], [111], [112]. The nuanced analysis of LULC maps, on a tested area (Dublin), reveals the transformative impact of pixel purification on classification outcomes, facilitating a more distinct demarcation of land covers, especially in water bodies and vegetation areas. Such improvement in RF classifier distinction highlights the value of accurate spectral information in achieving LULC classification precision, aligning with previous findings on the importance of accurate

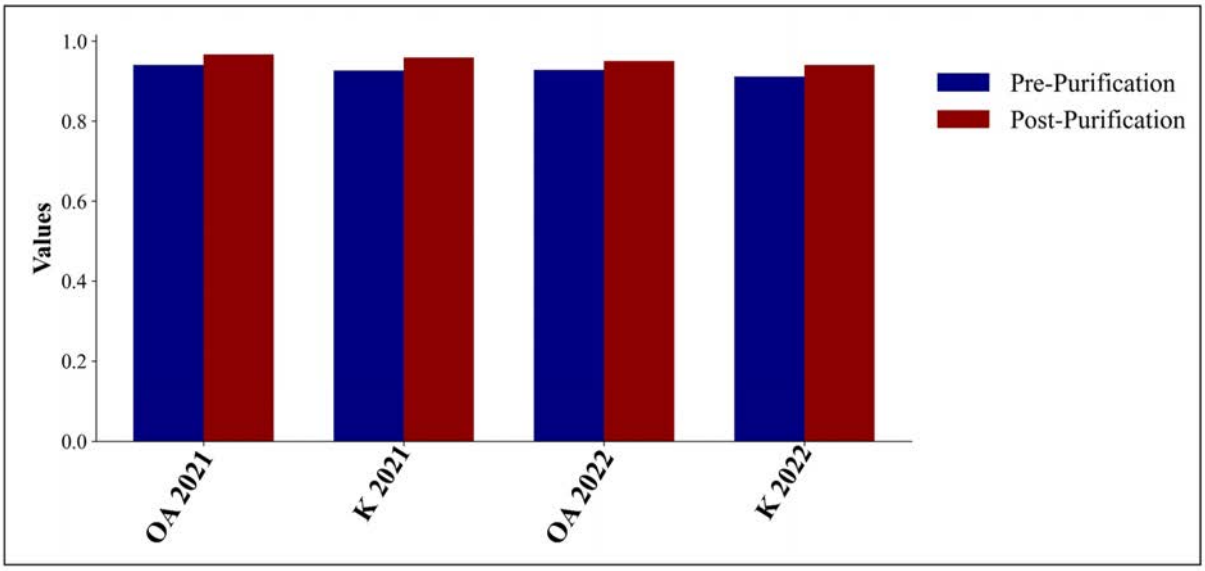

FIGURE 11. Accuracy assessment metrics (Overall Accuracy, OA, and Kappa Coefficient, K) for LULC classification pre- and post-pixel purification.

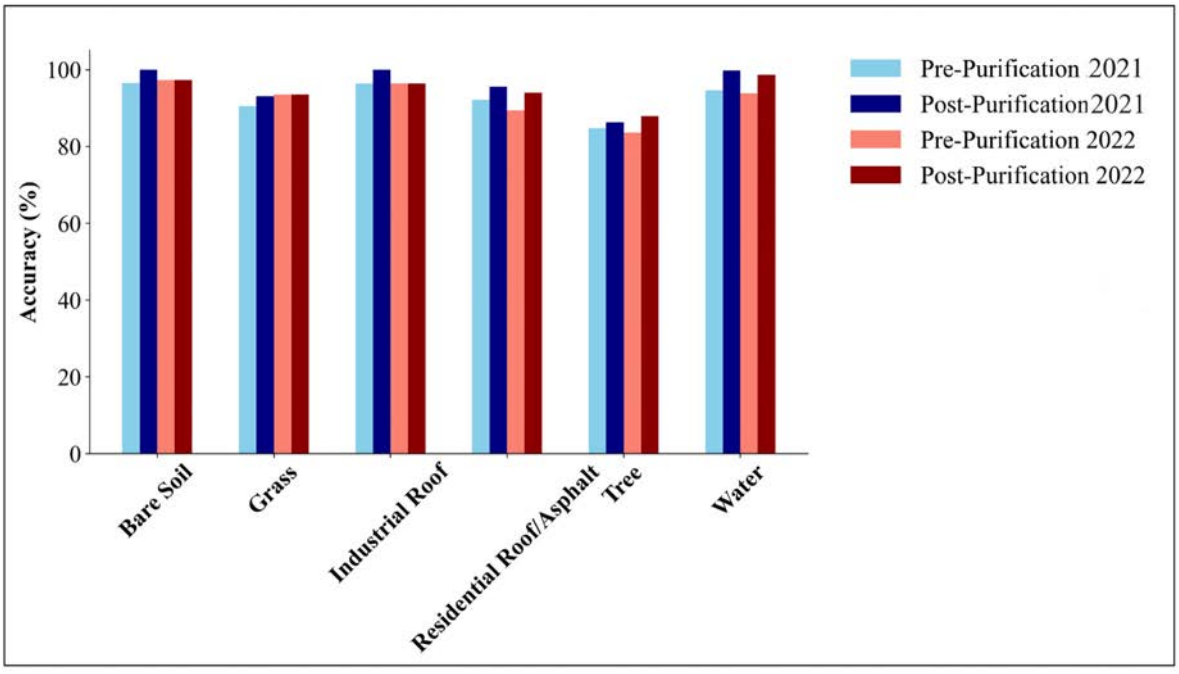

FIGURE 12. User Accuracy (UA) comparisons for the six LULC classes pre- and post-pixel purification.

training data for enhanced urban material classification [33], [95], [113]. Furthermore, previous research has highlighted the RF algorithm's inclination towards high accuracy, contingent upon the quality of training data [39], [114].

C. IMPERVIOUS-PERVIOUS SURFACE CHANGE DETECTION

In this study, we employed a pixel-by-pixel differential approach to quantify the shift from pervious to impervious surfaces in a highly urbanized region of Dublin using $PRISMA_{Fused,2021}$ and $PRISMA_{Fused,2022}$ Imageries. This method enabled accurate differential analysis of (im)pervious surfaces, revealing a 4.08% decrease in pervious surfaces (from 33.29 km^2 to 28.08 km^2, particularly affecting green space) and a 4.09% increase in impervious surfaces (from 79.96 km^2 to 82.92 km^2) over the year. These figures underscore the significant impact of urbanization in reduction in green spaces and expansion of urban structures. The combination of CNMF fusion and automatic pixel purification provides the spectral clarity needed for high-resolution mapping and change detection in urban areas, tackling the issues of mixed pixels and spectral ambiguities [79], [113]. Observations from Dublin's rapid urbanization and pervious surface reduction underscore the need for urgent

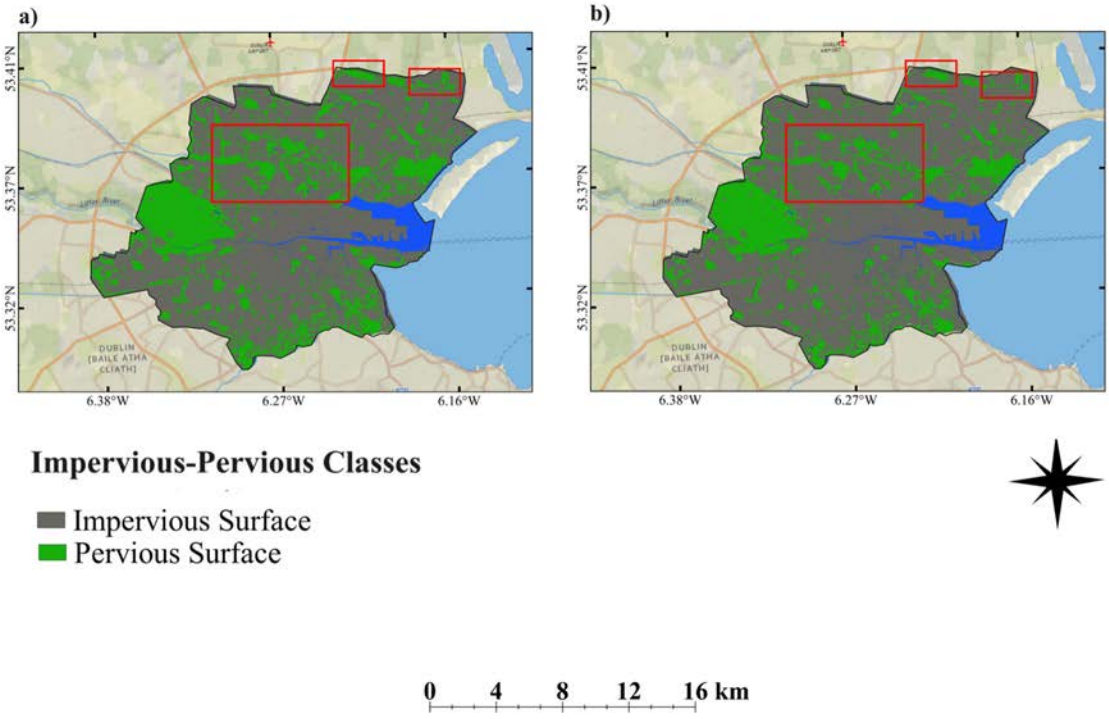

FIGURE 13. Map of impervious (gray) and pervious (green) surfaces in Dublin in 2021 (a) and 2022 (b) using pixel-purification algorithm. Red boxes represent the areas with significant surface changes (cfr. Figure 14).

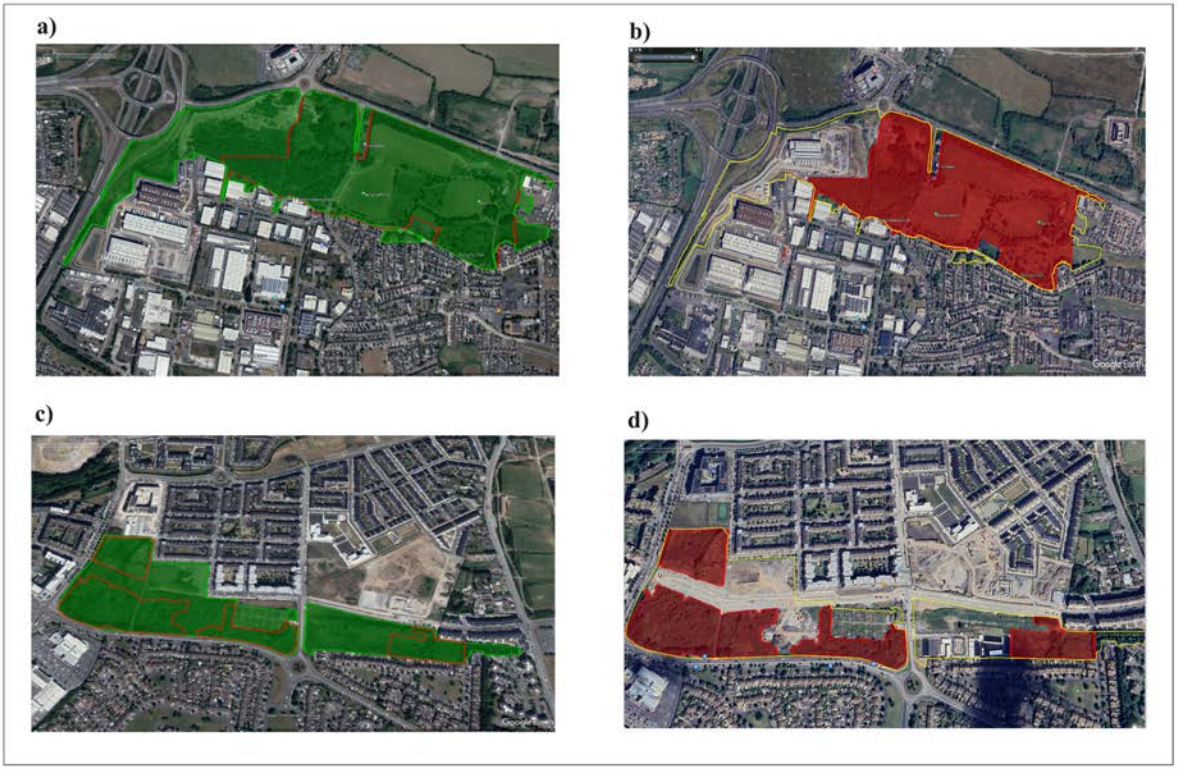

FIGURE 14. Visual validation of impervious-pervious using Google Earth Imagery in selected Dublin areas. Panels a) and b) show the transition from pervious (green in a) to impervious surfaces (red in b) from 2021 to 2022. The same for another area within Dublin is shown in panels c) and d).

sustainable urban planning practices. By providing comprehensive insights into the proliferation of impervious surfaces, our research supports policymakers and urban planners in promoting ecological sustainability amid rapid urban development. In summary, the integration of our novel automatic pixel purification with fusion algorithms' capability (CNMF)

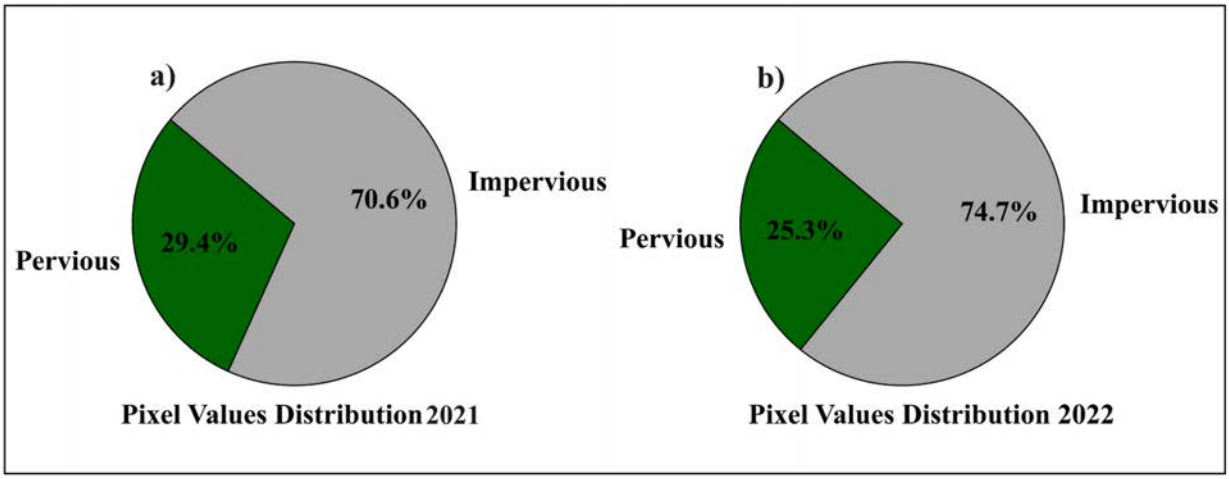

FIGURE 15. Pie charts illustrating the distribution of impervious (grey) and pervious (green) surfaces within Dublin, comparing the years a) 2021 and b) 2022.

in spatial resolution enhancement offers a valuable and highly efficient tool for detecting and monitoring changes between impervious and pervious surfaces. This methodology not only facilitates precise urban environmental analysis but also encourages its broader application in diverse urban studies, aiming to enhance classification accuracy and support sustainable urban development strategies.

V. CONCLUSION

A methodological innovation in detecting LULC changes in highly urbanized areas (tested on Dublin), focusing on the impact of urbanization on surface permeability, is assessed. We employed the CNMF algorithm to enhance the resolution of PRISMA hyperspectral (30m) imagery by fusing it with high-resolution PRISMA-PAN (5m) data. Initial applications of the RF algorithm on fused images for LULC classification highlighted challenges such as susceptibility to misclassification and mixed pixel selection errors. To address these issues, we developed an automated pixel purification algorithm that leverages spectral similarity angles (SAM) and spectral distances (ED) in n-dimensional feature space to accurately remove mixed pixels, outliers, and erroneously selected pixels. This purification significantly enhanced the integrity of spectral signatures, particularly in challenging classes like trees and residential roofs/asphalts, and significantly improved the RF algorithm's accuracy.

These advancements underscore the critical role of our pixel purification method in enhancing classification reliability and providing detailed insights into the transition from previous to impervious surfaces within the Dublin area. The refined classification process indicated a significant transformation in landcovers with a 4.09% increase in impervious surfaces and a 4.08% decrease in pervious surfaces from the year 2021 to 2022. The core innovation of our study is the automatic pixel purification process, designed to address challenges from high-resolution datasets by eliminating mixed pixels and outliers. The effectiveness of this process depends on the quality of the input data and may vary with different spectral inconsistencies across environments. In addition, the range of confidence intervals and implied thresholding may also influence the algorithm's performance and need to be experimentally determined.

The spatial detail provided by our novel and automatic approach is needed to manage the impacts of landcover change and address the intensification of landscape transformations leading to the loss of green and blue spaces. This intensification is occurring as cities become denser in both build and population, leading to more efficient use of space but also significant local environmental impacts, particularly affecting urban hydrology, such precise and accurate LULC map is necessary at fine scales, to capture using conventional approaches, to effectively assess urban change and its implications for urban resilience. Moreover, the method developed here can provide the temporal resolution needed to capture rapid change and identify where it is concentrated.

A future step of this approach will be the evaluation of this pixel-purification algorithm with other spectral data sources to enhance its versatility and effectiveness across diverse landscapes. We also aim to test the scalability of this approach for broader geographic applications and assess its utility with emerging remote sensing technologies to further support sustainable urban development.

REFERENCES

[1] C. Kamusoko and C. Kamusoko, *Image Classification*. Berlin, Germany: Springer, 2019.

[2] Y. Wang, Y. Sun, X. Cao, Y. Wang, W. Zhang, and X. Cheng, "A review of regional and global scale land use/land cover (LULC) mapping products generated from satellite remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 206, pp. 311–334, Dec. 2023, doi: 10.1016/j.isprsjprs.2023.11.014.

[3] B. J. Thompson, Z.-U. Rahman, and S. K. Park, "Multiscale retinex for improved performance in multispectral image classification," *Proc. SPIE*, vol. 4041, pp. 34–44, Jun. 2000.

[4] Y. Liu, S. Lu, X. Lu, Z. Wang, C. Chen, and H. He, "Classification of urban hyperspectral remote sensing imagery based on optimized spectral angle mapping," *J. Indian Soc. Remote Sens.*, vol. 47, no. 2, pp. 289–294, Feb. 2019, doi: 10.1007/s12524-018-0929-1.

[5] S. Gakhar and K. Chandra Tiwari, "Spectral–spatial urban target detection for hyperspectral remote sensing data using artificial neural network," *Egyptian J. Remote Sens. Space Sci.*, vol. 24, no. 2, pp. 173–180, Aug. 2021, doi: 10.1016/j.ejrs.2021.01.002.
[6] Q. Man, P. Dong, and H. Guo, "Pixel- and feature-level fusion of hyperspectral and LiDAR data for urban land-use classification," *Int. J. Remote Sens.*, vol. 36, no. 6, pp. 1618–1644, Mar. 2015.
[7] M. Khodadadzadeh, J. Li, S. Prasad, and A. Plaza, "Fusion of hyperspectral and LiDAR remote sensing data using multiple feature learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2971–2983, Jun. 2015.
[8] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.
[9] B. Yang, S. Wang, S. Li, B. Zhou, F. Zhao, F. Ali, and H. He, "Research and application of UAV-based hyperspectral remote sensing for smart city construction," *Cognit. Robot.*, vol. 2, pp. 255–266, Jan. 2022, doi: 10.1016/j.cogr.2022.12.002.
[10] D. Peijun, X. Junshi, C. Wen, and W. Xiaoling, "Extraction of urban impervious surface from hyperspectral remote sensing image," in *Proc. 2nd Workshop Hyperspectral Image Signal Process., Evol. Remote Sens.*, Jun. 2010, pp. 1–5, doi: 10.1109/WHISPERS.2010.5594892.
[11] Q. Weng, "Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends," *Remote Sens. Environ.*, vol. 117, pp. 34–49, Feb. 2012, doi: 10.1016/j.rse.2011.02.030.
[12] Y. Ji, F. Tan, S. Zhao, A. Feng, C. Zeng, H. Liu, and C. Wang, "Spatial–spectral resolution tunable snapshot imaging spectrometer: Analytical design and implementation," *Appl. Opt.*, vol. 62, no. 17, pp. 4456–4464, 2023.
[13] Y. Sun, A. Zia, and J. Zhou, "High spectral spatial resolution synthetic HyperSpectral dataset form multi-source fusion," 2023, *arXiv:2309.00005*.
[14] J. M. Ramirez and H. Arguello, "Multiresolution compressive feature fusion for spectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9900–9911, Dec. 2019, doi: 10.1109/TGRS.2019.2930093.
[15] R. U. Shaik, S. Periasamy, and W. Zeng, "Potential assessment of PRISMA hyperspectral imagery for remote sensing applications," *Remote Sens.*, vol. 15, no. 5, p. 1378, Feb. 2023.
[16] P. Ghamisi, N. Yokoya, J. Li, W. Liao, S. Liu, J. Plaza, B. Rasti, and A. Plaza, "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 37–78, Dec. 2017.
[17] M. Ehlers, "17 new developments and trends for urban remote sensing," in *Urban Remote Sensing*, Q. Weng and D. A. Quattrochi, Eds. Boca Raton, FL, USA: CRC Press, 2006, pp. 357–375.
[18] H. Huang, G. Shi, H. He, Y. Duan, and F. Luo, "Dimensionality reduction of hyperspectral imagery based on spatial–spectral manifold learning," *IEEE Trans. Cybern.*, vol. 50, no. 6, pp. 2604–2616, Jun. 2020.
[19] W. Gardner, D. A. Winkler, R. Maliki, S. M. Cutts, S. Ellis, R. L. Anderson, B. W. Muir, and P. J. Pigram, "Fusing ToF-SIMS images for spatial–spectral resolution enhancement using a convolutional neural network," *Adv. Mater. Interfaces*, vol. 9, no. 34, Dec. 2022, Art. no. 2201464.
[20] Q. Weng, "What is special about global urban remote sensing," in *Global Urban Monitoring and Assessment Through Earth Observation*. Boca Raton, FL, USA: CRC Press, 2014, pp. 1–12.
[21] S. Kahraman, G. Yesilyurt, A. Ertürk, and S. Ertürk, "Block-based and segmentation-based approaches for component substitution based hyperspectral pansharpening," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2018, pp. 8054–8057.
[22] L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, J.-Y. Tourneret, M. A. Veganzones, G. Vivone, Q. Wei, and N. Yokoya, "Hyperspectral pansharpening: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.
[23] H. Ghassemian, "A review of remote sensing image fusion methods," *Inf. Fusion*, vol. 32, pp. 75–89, Nov. 2016.
[24] G. Vivone, R. Restaino, G. Licciardi, M. D. Mura, and J. Chanussot, "Multiresolution analysis and component substitution techniques for hyperspectral pansharpening," in *Proc. IEEE Geosci. Remote Sens. Symp.*, Jul. 2014, pp. 2649–2652.
[25] W. Dong, Y. Yang, J. Qu, S. Xiao, and Q. Du, "Hyperspectral pansharpening via local intensity component and local injection gain estimation," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
[26] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and pan imagery," *Photogrammetric Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.
[27] S. Bakken, D. Langer, J. L. Garrett, S. Y. Le Moan, and T. A. Johansen, "Spatial resolution enhancement of hyperspectral images from miniaturized sensors," *Proc. SPIE*, vol. 12733, pp. 24–38, Oct. 2023.
[28] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, Jan. 2000.
[29] A. M. S. Ang and N. Gillis, "Algorithms and comparisons of nonnegative matrix factorizations with volume regularization for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 4843–4853, Dec. 2019.
[30] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 2, pp. 29–56, Jun. 2017.
[31] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.
[32] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 67, pp. 93–104, Jan. 2012, doi: 10.1016/j.isprsjprs.2011.11.002.
[33] M. Belgiu and L. Drăgut, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogramm. Remote Sens.*, vol. 114, pp. 24–31, Apr. 2016.
[34] G. M. Foody, "Status of land cover classification accuracy assessment," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 185–201, Apr. 2002.
[35] P. Gong and P. J. Howarth, "The use of structural information for improving land-cover classification accuracies at the rural-urban fringe," *Photogrammetric Eng. Remote Sens.* vol. 56, no. 1, pp. 67–73, 2021. [Online]. Available: http http://hub.hku.hk/bitstream/10722/296491/1/re01.htm
[36] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 3, pp. 247–259, May 2011, doi: 10.1016/j.isprsjprs.2010.11.001.
[37] C. D. Lippitt, J. Rogan, Z. Li, J. R. Eastman, and T. G. Jones, "Mapping selective logging in mixed deciduous forest," *Photogrammetric Eng. Remote Sens.*, vol. 74, no. 10, pp. 1201–1211, Oct. 2008.
[38] M. Pal and P. M. Mather, "An assessment of the effectiveness of decision tree methods for land cover classification," *Remote Sens. Environ.*, vol. 86, no. 4, pp. 554–565, Aug. 2003.
[39] M. Pal, "Random forest classifier for remote sensing classification," *Int. J. Remote Sens.*, vol. 26, no. 1, pp. 217–222, Jan. 2005.
[40] A. Mellor, S. Boukir, A. Haywood, and S. Jones, "Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 155–168, Jul. 2015, doi: 10.1016/j.isprsjprs.2015.03.014.
[41] K. Millard and M. Richardson, "On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping," *Remote Sens.*, vol. 7, no. 7, pp. 8489–8515, Jul. 2015. [Online]. Available: https://www.mdpi.com/2072-4292/7/7/8489
[42] P. Du, A. Samat, B. Waske, S. Liu, and Z. Li, "Random forest and rotation forest for fully polarized SAR image classification using polarimetric and spatial features," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 38–53, Jul. 2015, doi: 10.1016/j.isprsjprs.2015.03.002.
[43] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, and M. Marconcini, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. 110–122, Sep. 2009.
[44] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
[45] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," *Remote Sens.*, vol. 10, no. 3, p. 482, Mar. 2018.
[46] L. Ma, M. Li, X. Ma, L. Cheng, P. Du, and Y. Liu, "A review of supervised object-based land-cover image classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 130, pp. 277–293, Aug. 2017, doi: 10.1016/j.isprsjprs.2017.06.001.
[47] L. Ma, L. Cheng, M. Li, Y. Liu, and X. Ma, "Training set size, scale, and features in geographic object-based image analysis of very high resolution unmanned aerial vehicle imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 102, pp. 14–27, Apr. 2015.

[48] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[49] L. Ma, M. M. Crawford, and J. Tian, "Local manifold learning-based k-nearest-neighbor for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4099–4109, Nov. 2010.

[50] D. Spiller, L. Ansalone, S. Amici, A. Piscini, and P. P. Mathieu, "Analysis and detection of wildfires by using prisma hyperspectral imagery," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 43, pp. 215–222, Jun. 2021, doi: 10.5194/isprs-archives-xliii-b3-2021-215-2021.

[51] E. Bedini and J. Chen, "Prospection for economic mineralization using PRISMA satellite hyperspectral remote sensing imagery: An example from central east Greenland," *J. Hyperspectral Remote Sens.*, vol. 12, no. 3, pp. 124–130, May 2022.

[52] N. Bohn, B. Di Mauro, R. Colombo, D. R. Thompson, J. Susiluoto, N. Carmon, M. J. Turmon, and L. Guanter, "Glacier ice surface properties in South-West Greenland Ice Sheet: First estimates from PRISMA imaging spectroscopy data," *J. Geophys. Res., Biogeosciences*, vol. 127, no. 3, Mar. 2022, Art. no. e2021JG006718.

[53] J. Verrelst, J. P. Rivera-Caicedo, P. Reyes-Muñoz, M. Morata, E. Amin, G. Tagliabue, C. Panigada, T. Hank, and K. Berger, "Mapping landscape canopy nitrogen content from space using PRISMA data," *ISPRS J. Photogramm. Remote Sens.*, vol. 178, pp. 382–395, Aug. 2021.

[54] E. Alevizos, T. Le Bas, and D. D. Alexakis, "Assessment of PRISMA level-2 hyperspectral imagery for large scale satellite-derived bathymetry retrieval," *Mar. Geodesy*, vol. 45, no. 3, pp. 251–273, May 2022.

[55] R. Loizzo, M. Daraio, R. Guarini, F. Longo, R. Lorusso, L. Dini, and E. Lopinto, "Prisma mission status and perspective," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 4503–4506.

[56] D. O. McInerney and C. Walsh, "Spatial distribution of urban land-use change in the Dublin city-region: 1990–2006," *Irish Geography*, vol. 42, no. 2, pp. 207–223, Apr. 2014.

[57] L. Busetto and L. Ranghetti. (2020). *Prismaread: A Tool for Facilitating Access and Analysis of PRISMA L1/L2 Hyperspectral Imagery V1. 0.0*. [Online]. Available: https://lbusett.github.io/prismaread/

[58] M. Ranghetti, M. Boschetti, L. Ranghetti, G. Tagliabue, C. Panigada, M. Gianinetto, J. Verrelst, and G. Candiani, "Assessment of maize nitrogen uptake from PRISMA hyperspectral data through hybrid modelling," *Eur. J. Remote Sens.*, vol. 56, no. 1, Dec. 2023, Art. no. 2117650.

[59] D. Upreti, W. Huang, W. Kong, S. Pascucci, S. Pignatti, X. Zhou, H. Ye, and R. Casa, "A comparison of hybrid machine learning algorithms for the retrieval of wheat biophysical variables from Sentinel-2," *Remote Sens.*, vol. 11, no. 5, p. 481, Feb. 2019.

[60] P. Sajadi, M. Gholamnia, S. Bonafoni, and F. Pilla, "Spectral-based regression model for destriping of abnormal pixel values in PRISMA hyperspectral image," *Eur. J. Remote Sens.*, vol. 55, no. 1, pp. 622–643, Dec. 2022.

[61] Y. Behroozi, M. Yazdi, and A. Z. Asli, "Hyperspectral image denoising based on superpixel segmentation low-rank matrix approximation and total variation," *Circuits, Syst., Signal Process.*, vol. 41, no. 6, pp. 3372–3396, Jun. 2022.

[62] C. Li, C. Zhou, L. Ma, L. Tang, and X. Wang, "A stripe noise removal method of interference hyperspectral imagery based on interferogram correction," *Proc. SPIE*, vol. 8537, pp. 71–77, 2012.

[63] M. J. Mashala, T. Dube, B. T. Mudereri, K. K. Ayisi, and M. R. Ramudzuli, "A systematic review on advancements in remote sensing for assessing and monitoring land use and land cover changes impacts on surface water resources in semi-arid tropical environments," *Remote Sens.*, vol. 15, no. 16, p. 3926, Aug. 2023. [Online]. Available: https://www.mdpi.com/2072-4292/15/16/3926

[64] L. Kühnlein and S. Keller, "Evaluation of transformers and convolutional neural networks for high-dimensional hyperspectral soil texture classification," in *Proc. 12th Workshop Hyperspectral Imag. Signal Process., Evol. Remote Sens. (WHISPERS)*, Sep. 2022, pp. 1–5, doi: 10.1109/WHISPERS56178.2022.9955087.

[65] T. Zhang, W. Wang, J. Wang, Y. Cai, Z. Yang, and J. Li, "Hyper-LGNet: Coupling local and global features for hyperspectral image classification," *Remote Sens.*, vol. 14, no. 20, p. 5251, Oct. 2022. [Online]. Available: https://www.mdpi.com/2072-4292/14/20/5251

[66] Y. Ji, S. M. Park, S. Kwon, J. W. Leem, V. V. Nair, Y. Tong, and Y. L. Kim, "MHealth hyperspectral learning for instantaneous spatiospectral imaging of hemodynamics," *PNAS Nexus*, vol. 2, no. 4, Apr. 2023, Art. no. pgad111, doi: 10.1093/pnasnexus/pgad111.

[67] K. R. Shahi, P. Ghamisi, B. Rasti, R. Gloaguen, and P. Scheunders, "MS^{2}A-net: Multiscale spectral–spatial association network for hyperspectral image clustering," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 6518–6530, 2022, doi: 10.1109/JSTARS.2022.3198137.

[68] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled non-negative matrix factorization (CNMF) for hyperspectral and multispectral data fusion: Application to pasture classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2011, pp. 1779–1782.

[69] K. Priya and K. K. Rajkumar, "Minimum simplex nonlinear nonnegative matrix factorization for hyperspectral unmixing," in *Smart Intelligent Computing and Applications*, vol. 1, V. Bhateja, S. C. Satapathy, C. M. Travieso-Gonzalez, and T. Adilakshmi, Eds. Singapore: Springer, 2022, pp. 161–175.

[70] C.-H. Lin, F. Ma, C.-Y. Chi, and C.-H. Hsieh, "A convex optimization-based coupled nonnegative matrix factorization algorithm for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1652–1667, Mar. 2018, doi: 10.1109/TGRS.2017.2766080.

[71] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[72] X. Tian, W. Zhang, Y. Chen, Z. Wang, and J. Ma, "HyperFusion: A computational approach for hyperspectral, multispectral, and panchromatic image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2022, Art. no. 5518216, doi: 10.1109/TGRS.2021.3128279.

[73] K. N. Loukika, V. R. Keesara, and V. Sridhar, "Analysis of land use and land cover using machine learning algorithms on Google Earth engine for Munneru River Basin, India," *Sustainability*, vol. 13, no. 24, p. 13758, Dec. 2021. [Online]. Available: https://www.mdpi.com/2071-1050/13/24/13758

[74] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, Oct. 2001.

[75] S. Oliveira, F. Oehler, J. San-Miguel-Ayanz, A. Camia, and J. M. C. Pereira, "Modeling spatial patterns of fire occurrence in Mediterranean Europe using multiple regression and random forest," *Forest Ecology Manag.*, vol. 275, pp. 117–129, Jul. 2012.

[76] U. Grömping, "Variable importance assessment in regression: Linear regression versus random forest," *Amer. Statistician*, vol. 63, no. 4, pp. 308–319, Nov. 2009.

[77] C. Yang, G. Wu, K. Ding, T. Shi, Q. Li, and J. Wang, "Improving land use/land cover classification by integrating pixel unmixing and decision tree methods," *Remote Sens.*, vol. 9, no. 12, p. 1222, Nov. 2017. [Online]. Available: https://www.mdpi.com/2072-4292/9/12/1222

[78] M. Y. Lilay and G. D. Taye, "Semantic segmentation model for land cover classification from satellite images in gambella national park, Ethiopia," *Social Netw. Appl. Sci.*, vol. 5, no. 3, p. 76, Mar. 2023, doi: 10.1007/s42452-023-05280-4.

[79] V. F. Rodriguez-Galiano, M. Chica-Olmo, F. Abarca-Hernandez, P. M. Atkinson, and C. Jeganathan, "Random forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture," *Remote Sens. Environ.*, vol. 121, pp. 93–107, Jun. 2012, doi: 10.1016/j.rse.2011.12.003.

[80] R. Pu, *Hyperspectral Remote Sensing: Fundamentals and Practices*. Boca Raton, FL, USA: CRC Press, 2017.

[81] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. Berlin, Germany: Springer, 2003.

[82] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz, "The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data," *Remote Sens. Environ.*, vol. 44, nos. 2–3, pp. 145–163, May 1993.

[83] F. van der Meer, "The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 8, no. 1, pp. 3–17, Jan. 2006, doi: 10.1016/j.jag.2005.06.001.

[84] N. Khetan, L. Kejriwal, and S. Indu, "Enhancement of degraded manuscript images using adaptive Gaussian thresholding," *Int. J. Future Gener. Commun. Netw.*, vol. 10, no. 1, pp. 47–60, Jan. 2017.

[85] A. Khmag, "Additive Gaussian noise removal based on generative adversarial network model and semi-soft thresholding approach," *Multimedia Tools Appl.*, vol. 82, no. 5, pp. 7757–7777, Feb. 2023.

[86] J. L. Folks and R. S. Chhikara, "The inverse Gaussian distribution and its statistical application—A review," *J. Roy. Stat. Soc. Ser. B, Stat. Methodol.*, vol. 40, no. 3, pp. 263–275, Jul. 1978.

[87] Z. Xu, C. Su, and X. Zhang, "A semantic segmentation method with category boundary for land use and land cover (LULC) mapping of very-high resolution (VHR) remote sensing image," *Int. J. Remote Sens.*, vol. 42, no. 8, pp. 3146–3165, Apr. 2021.
[88] M. T. Eismann, "Resolution enhancement of hyperspectral imagery using maximum a posteriori estimation with a stochastic mixing model," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Dayton, Dayton, OH, USA, May 2004.
[89] P. Javadi, "Use satellite images and improve the accuracy of hyperspectral image with the classification," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 40, pp. 343–349, Dec. 2015.
[90] O. Rajadell, P. García-Sevilla, and F. Pla, "Spectral–spatial pixel characterization using Gabor filters for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 860–864, Jul. 2013.
[91] J. Aryal, C. Sitaula, and A. C. Frery, "Land use and land cover (LULC) performance modeling using machine learning algorithms: A case study of the city of Melbourne, Australia," *Sci. Rep.*, vol. 13, no. 1, p. 13510, Aug. 2023, doi: 10.1038/s41598-023-40564-0.
[92] T. Kavzoglu and F. Bilucan, "Effects of auxiliary and ancillary data on LULC classification in a heterogeneous environment using optimized random forest algorithm," *Earth Sci. Informat.*, vol. 16, no. 1, pp. 415–435, Mar. 2023, doi: 10.1007/s12145-022-00874-9.
[93] T. D. H. Le, L. H. Pham, Q. T. Dinh, N. T. T. Hang, and T. A. T. Tran, "Rapid method for yearly LULC classification using random forest and incorporating time-series NDVI and topography: A case study of thanh hoa province, Vietnam," *Geocarto Int.*, vol. 37, no. 27, pp. 17200–17215, Dec. 2022.
[94] M. Wang, D. Mao, Y. Wang, K. Song, H. Yan, M. Jia, and Z. Wang, "Annual wetland mapping in metropolis by temporal sample migration and random forest classification with time series Landsat data and Google Earth engine," *Remote Sens.*, vol. 14, no. 13, p. 3191, Jul. 2022. [Online]. Available: https://www.mdpi.com/2072-4292/14/13/3191
[95] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.*, vol. 37, no. 1, pp. 35–46, Jul. 1991.
[96] J. Wickham, S. V. Stehman, D. G. Sorenson, L. Gass, and J. A. Dewitz, "Thematic accuracy assessment of the NLCD 2019 land cover for the conterminous United States," *GIScience Remote Sens.*, vol. 60, no. 1, Dec. 2023, Art. no. 2181143.
[97] F. A. Islami, S. D. Tarigan, E. D. Wahjunie, and B. D. Dasanto, "Accuracy assessment of land use change analysis using Google Earth in Sadar Watershed Mojokerto regency," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 950, Jan. 2022, Art. no. 012091.
[98] D. Lu, E. Moran, and S. Hetrick, "Detection of impervious surface change with multitemporal Landsat images in an urban–rural frontier," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 3, pp. 298–306, May 2011, doi: 10.1016/j.isprsjprs.2010.10.010.
[99] D. Lu, P. Mausel, E. Brondízio, and E. Moran, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2401, 2004.
[100] R. E. Kennedy, P. A. Townsend, J. E. Gross, W. B. Cohen, P. Bolstad, Y. Q. Wang, and P. Adams, "Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects," *Remote Sens. Environ.*, vol. 113, no. 7, pp. 1382–1396, Jul. 2009.
[101] C. L. Arnold and C. J. Gibbons, "Impervious surface coverage: The emergence of a key environmental indicator," *J. Amer. Planning Assoc.*, vol. 62, no. 2, pp. 243–258, Jun. 1996.
[102] G. Xian, M. Crane, and J. Su, "An analysis of urban development and its environmental impact on the Tampa bay watershed," *J. Environ. Manag.*, vol. 85, no. 4, pp. 965–976, Dec. 2007.
[103] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.
[104] C. Pohl and J. Van Genderen, *Remote Sensing Image Fusion: A Practical Guide*. Boca Raton, FL, USA: CRC Press, 2016.
[105] R. Dian, S. Li, B. Sun, and A. Guo, "Recent advances and new guidelines on hyperspectral and multispectral image fusion," *Inf. Fusion*, vol. 69, pp. 40–51, May 2021.
[106] D. Gerçek, D. çesmeci, M. K. Güllü, A. Ertürk, and S. Ertürk, "Unmixing-based fusion of hyperspectral images with high spatial resolution images," in *Proc. 21st Signal Process. Commun. Appl. Conf. (SIU)*, Apr. 2013, pp. 1–4.
[107] S. W. Myint, P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng, "Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery," *Remote Sens. Environ.*, vol. 115, no. 5, pp. 1145–1161, May 2011.
[108] B. Johnson, "Scale issues related to the accuracy assessment of land use/land cover maps produced using multi-resolution data: Comments on 'the improvement of land cover classification by thermal remote sensing,'" *Remote Sens.*, vol. 7, no. 10, pp. 13436–13439, Oct. 2015. [Online]. Available: https://www.mdpi.com/2072-4292/7/10/13436
[109] V. N. Mishra, R. Prasad, P. Kumar, D. K. Gupta, P. K. S. Dikshit, S. B. Dwivedi, and A. Ohri, "Evaluating the effects of spatial resolution on land use and land cover classification accuracy," in *Proc. Int. Conf. Microw., Opt. Commun. Eng. (ICMOCE)*, Dec. 2015, pp. 208–211, doi: 10.1109/ICMOCE.2015.7489727.
[110] L. Qu, Z. Chen, M. Li, J. Zhi, and H. Wang, "Accuracy improvements to pixel-based and object-based LULC classification with auxiliary datasets from Google Earth engine," *Remote Sens.*, vol. 13, no. 3, p. 453, Jan. 2021. [Online]. Available: https://www.mdpi.com/2072-4292/13/3/453
[111] A. E. Maxwell, T. A. Warner, and F. Fang, "Implementation of machine-learning classification in remote sensing: An applied review," *Int. J. Remote Sens.*, vol. 39, no. 9, pp. 2784–2817, May 2018, doi: 10.1080/01431161.2018.1433343.
[112] A. E. Maxwell, M. P. Strager, T. A. Warner, C. A. Ramezan, A. N. Morgan, and C. E. Pauley, "Large-area, high spatial resolution land cover mapping using random forests, GEOBIA, and NAIP orthophotography: Findings and recommendations," *Remote Sens.*, vol. 11, no. 12, p. 1409, Jun. 2019, doi: 10.3390/rs11121409.
[113] D. C. Duro, S. E. Franklin, and M. G. Dubé, "Multi-scale object-based image analysis and feature selection of multi-sensor Earth observation imagery using random forests," *Int. J. Remote Sens.*, vol. 33, no. 14, pp. 4502–4526, Jul. 2012.
[114] D. R. Cutler, T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler, "Random forests for classification in ecology," *Ecology*, vol. 88, no. 11, pp. 2783–2792, Nov. 2007.
[115] G. Vivone, R. Restaino, and J. Chanussot, "Full scale regression-based injection coefficients for panchromatic sharpening," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3418–3431, Jul. 2018.
[116] G. Vivone, M. Simões, M. Dalla Mura, R. Restaino, J. M. Bioucas-Dias, G. A. Licciardi, and J. Chanussot, "Pansharpening based on semiblind deconvolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1997–2010, Apr. 2015.
[117] A. A. J. Deijns, A. R. Bevington, F. van Zadelhoff, S. M. de Jong, M. Geertsema, and S. McDougall, "Semi-automated detection of landslide timing using harmonic modelling of satellite imagery, Buckinghorse River, Canada," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 84, Feb. 2020, Art. no. 101943.

PAYAM SAJADI received the Ph.D. degree in remote sensing and GIS from Jawaharlal Nehru University, in 2018. He is currently a Principal Investigator and a Senior Postdoctoral Researcher with the Spatial Dynamics Laboratory, University College Dublin (UCD). He is also a leading Researcher specializing in the convergence of remote sensing and big geospatial data science. His work leverages advanced machine learning and deep learning methodologies to tackle critical environmental and sustainability issues, demonstrating significant contributions to the field through his innovative approaches in data analysis and model development. He distinguished himself with groundbreaking research at Jawaharlal Nehru University, that has since been cited in numerous peer-reviewed articles. His expertise is further recognized through his role with the Spatial Dynamics Laboratory, UCD, where he mentors emerging scientists and engineers, fostering a new generation of experts in geospatial data science.

MEHDI GHOLAMNIA received the Ph.D. degree in remote sensing and GIS from the University of Tehran, in 2017. He developed a strong foundation in both theoretical and applied aspects of geospatial sciences at the University of Tehran. He is currently a Postdoctoral Fellow with the Spatial Dynamics Laboratory, University College Dublin (UCD), where he engages in cutting-edge research on next-generation energy systems (NexSys) and the CAMEO project, which focuses on creating architectures for manipulating Earth observation data. His research interests remote sensing data analysis, pattern recognition, machine learning, geospatial processing, and statistical and mathematical modeling. He is also skilled in time series and signal processing techniques, applying these methods to various complex environmental, and geospatial datasets to extract meaningful patterns and insights.

STEFANIA BONAFONI (Senior Member, IEEE) received the Laurea (cum laude) and Ph.D. degrees in electronic engineering from the University of Perugia, Perugia, Italy, in 1997 and 2000, respectively. Since 2005, she has been with the Department of Engineering, University of Perugia, where she teaches a class on remote sensing. Her research activity concerns the remote sensing of the Earth's atmosphere and surface: land surface temperature and albedo from satellite-based sensors, downscaling techniques for satellite images, microwave radiometry from satellite and ground-based sensors for the estimation of atmospheric parameters, GPS for propagation delay and water vapor analyses, and GPS radio occultation for atmospheric profiling.

GERALD MILLS received the B.A. and M.A. degrees from University College Dublin (UCD), Ireland, with a focus on the synoptic climatology of precipitation, and the Ph.D. degree from The Ohio State University, with a focus on numerical cartography and physical climatology. After seven years in academia in the USA, mainly at UCLA, he returned to UCD, in 1997. He is currently a leading Physical Geographer with the School of Geography, UCD, specializing in the climatology of urban areas. He has coauthored influential books, such as *Urban Climates* (Cambridge University Press, 2017) and *The Urban Heat Island: A Guidebook* (Elsevier, 2021). His research interests include urban climates, climate change, and urban greening. He has held prestigious positions, including the President of the International Association for Urban Climates and the Irish Meteorological Society and has contributed to the WMO's Integrated Urban Services Initiative.

YAN-FANG SANG is currently a Distinguished Professor, specializing in hydrometeorology, natural disasters, climate change, and the sustainable management of urban water resources. He is also affiliated with the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, and the Key Laboratory of Compound and Chained Natural Hazards Dynamics under the Ministry of Emergency Management of China. He has authored over 80 publications in top-tier international journals, including *Nature Reviews Earth and Environment*, *One Earth*, *Water Resources Research*, and *Journal of Climate*. His research interests include urban and mountain flooding disasters, urban rainwater harvesting systems, and the dynamics of natural disasters in high mountain areas. He currently serves as an Editor for journals, such as *Natural Hazards Research* and *Hydrology*.

ZETING LI received the M.Sc. degree in geospatial data analysis, in 2021. He is currently pursuing the Ph.D. degree with the School of Geography, University College Dublin (UCD), specializing in remote sensing and big geospatial data modeling. His current research interest includes assessing risk vulnerability across major global cities. His work aims to develop methodologies that enhance urban resilience by integrating advanced spatial analysis techniques.

SALMAN KHAN received the Ph.D. degree in hydrology and water resources, in 2020, specializing in large-scale hydrological modeling using remote sensing techniques. He is currently a Postdoctoral Researcher with the Spatial Dynamics Laboratory, University College Dublin (UCD). His research journey includes a previous postdoctoral appointment at CIRAD, France, in 2022, where he focused on plant-available water estimation and soil mapping in the tropical Amazon region. His expertise spans watershed modeling, flood forecasting, and ecosystem preservation, with proficiency in GIS/RS, machine learning, and data analysis using the R programming language.

JIAZHENG HAN received the B.S. degree in survey and mapping engineering from China University of Mining and Technology, Xuzhou, China, in 2017, where he is currently pursuing the Ph.D. degree in geodesy and surveying engineering with the School of Environment Science and Spatial Informatics. His current research interests include remote sensing observations and ecological restoration of urban and coal mining areas.

FRANCESCO PILLA is currently a Full Professor and the Chair of smart cities with University College Dublin (UCD), Ireland. His work is positioned at the nexus of urban development and technology, aiming to enhance urban life through technological innovation, and citizen participation. His specialization in smart cities encompasses geospatial analysis and modeling of urban dynamics. This includes the development of GIS-based models and decision-support tools to anticipate the effects of human-environment interactions. He employs a variety of pervasive and community-sensing applications to calibrate and validate these tools. Further integrating these GIS models with advanced machine learning algorithms, he provides deeper insights into the spatial dynamics within cities, aiming to foster more sustainable, and efficient urban environments.