esa

PEOPLE

Ecosystem Restoration

D6: MEKONG DELTA DEMONSTRATION VALIDATION REPORT

January 2024

esa Hatfield THE UNIVERSITY OF BRITISH COLUMBIA VTT ICAS

Prepared for:

IUCN Vietnam

Hatfield Consultants LLP

#200 - 850 Harbourside Drive
North Vancouver, British Columbia, Canada V7P 0A3
Tel: 1.604.926.3261 • Fax: 1.604.926.5389
www.hatfieldgroup.com

PIONEER EARTH OBSERVATION APPLICATIONS FOR THE ENVIRONMENT – ECOSYSTEM RESTORATION (PEOPLE-ER)

D6: MEKONG DELTA DEMONSTRATION VALIDATION REPORT

Prepared for:

IUCN VIETNAM

Prepared by:

HATFIELD CONSULTANTS LLP

#200 - 850 HARBOURSIDE DRIVE

NORTH VANCOUVER, BC

CANADA V7P 0A3

TEL: 1.604.926.3261 • WWW.HATFIELDGROUP.COM

JANUARY 2024

ESA11339

VERSION 1

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DISTRIBUTION LIST

The following individuals/firms have received this document:

Name	Firm	Email	SharePoint	Project Website
Frank Martin Seifert	ESA	✓	-	-

AMENDMENT RECORD

This report has been issued and amended as follows:

Issue	Description	Date	Approved by	
1	D6 – Mekong Delta Demonstration Validation Report	20240104	(insert signature)	(insert signature)
			Andy Dean Project Manager	Marcos Kavlin Assistant Project Manager

1.0 INTRODUCTION

Ecosystem Restoration (ER) is important to reverse biodiversity loss and is a critical element of nature-based solutions (NBS) for climate change mitigation and adaptation, food security, and disaster risk reduction. ER is needed on a large scale to achieve the United Nations (UN) sustainable development agenda and as part of the UN Decade on Ecosystem Restoration (2021−2030). At the Convention on Biological Diversity (CBD) COP 15 in Montreal in December 2022, nations adopted a target to “Ensure that by 2030 at least 30 per cent of areas of degraded terrestrial, inland water, and coastal and marine ecosystems are under effective restoration, in order to enhance biodiversity and ecosystem functions and services, ecological integrity and connectivity[1].

Effective planning, monitoring, and assessment of ER is required to evaluate ecosystem functions and to determine whether ER is having the desired impact. ER investments must be data-driven, requiring historical information on ecosystem disturbance and degradation, to enable planning of interventions, which are then monitored for their impact. There is a huge opportunity for satellite Earth Observation (EO) applications for ER, to meet the needs for regular, repeat measures of ER processes over long time periods covering large, often remote, areas.

To support ER investments, innovative methods are required to deliver high-quality EO-based products and indicators targeting high-priority forest, wetland, and biodiversity variables.

The Pioneer Earth Observation apPlications for the Environment (PEOPLE) ER project financed by the European Space Agency (ESA) is a trailblazer project to develop innovative high-quality EO-based application products, indicators, and methods, targeting ER research and development (R&D) priorities.

PEOPLE-ER is led by Hatfield Consultants – a science-driven service-oriented company that builds solutions to complex environmental challenges, with a depth of experience in ER projects in Canada and around the world. Hatfield is a trusted partner for the development of cutting-edge and practical EO technologies. The PEOPLE-ER consortium includes:

- VTT – the remote sensing team at VTT Technical Research Centre of Finland produces EO data processing chains for domestic and international users. The team is internationally known, particularly for its forest monitoring applications and the Forestry TEP cloud processing platform. VTT is ranked among the leading European Research and Technology Organisations (RTO).
- University of British Columbia, Faculty of Forestry – Dr. Nicholas Coops leads the Integrated Remote Sensing Studio (IRSS) and is a leading international research scientist in the application of EO technologies for forest ecosystem assessment and monitoring, including ER and the prioritization of methods and products for remote sensing essential biodiversity variables (RS-EBVs).

The Early Adopters are:

- **National Institute for Research and Development in Forestry (INCDS)** (Romania) – formally a member of the consortium, INCDS is the main organisation of research and development in forestry from Romania. INCDS is in charge for the forest resources assessment and monitoring

[1] HYPERLINK "https://www.cbd.int/article/cop15-cbd-press-release-final-19dec2022"https://www.cbd.int/article/cop15-cbd-press-release-final-19dec2022

in Romania through National Forest Inventory. INCDS has also secured the support of two Romanian NGOs as documented in letters of support: Forestry Society Association and Fundatia Grupul Verde Oradea.

- **IUCN** (Vietnam) – established in 1948, IUCN is an international authority working on a wide range of themes related to nature conservation, forests, ecosystem management, protected areas, global policy and governance and rights.
- **African Parks** – a leading non-profit conservation organisation that takes on the complete responsibility for the rehabilitation and long-term management of protected areas across Africa in partnership with governments and local communities.
- **Society for Ecosystem Restoration in northern British Columbia (SERNbc) (Canada)** – a key enabler for ER in forested ecosystems affected by cumulative disturbances from forest operations, energy exploration, wildfires, and forest pests/diseases.
- **Natural Resources Institute (Luke) (Finland)** – as one of the biggest clusters of bioeconomy expertise in Europe, Luke develops knowledge-based solution models and services for renewable natural resources management and supports decision-making in society.

The following PEOPLE-ER Tools are defined:

1. **Vegetation Spectral Recovery** – The PEOPLE-ER Vegetation Spectral Recovery tool provides a flexible, powerful set of EO data analytics solutions to support forest landscape ER assessment. The tool provides a method for high-resolution satellite EO data time series analysis to enable monitoring of vegetation recovery in forested ecosystems from boreal to tropical biomes.
2. **K Nearest Neighbour Tool** – The PEOPLE-ER k-NN tool enables wall-to-wall prediction of target variables of interest using field reference data and selected EO datasets.
3. **Wetland Function Trends** – The PEOPLE-ER Wetland Function Trends tool provides a flexible, powerful set of EO data analytics tools to support wetland ER assessment. The tool provides methods for high-resolution satellite EO data time series analysis to enable monitoring of inundation dynamics and trends in natural to heavily modified wetland ecosystems.

These tools were identified following assessment of the current State of the Art (Deliverable 2a), Policy and Stakeholder Analysis (D2b), and Early Adopter Value Proposition (D3). The Tools are fully defined in the Algorithm Theoretical Baseline Documents (D7).

1.1 SCOPE OF DELIVERABLE D6 VERSION 2

Deliverable 6 (D6) version 1 defined the validation methodology for the PEOPLE-ER methods and algorithms to be developed and the specific demonstrations with Early Adopters. This updated version of D6 is the Validation Final report, presenting the tools/algorithm applied in each demonstration and the full set of results and associated accuracy assessment.

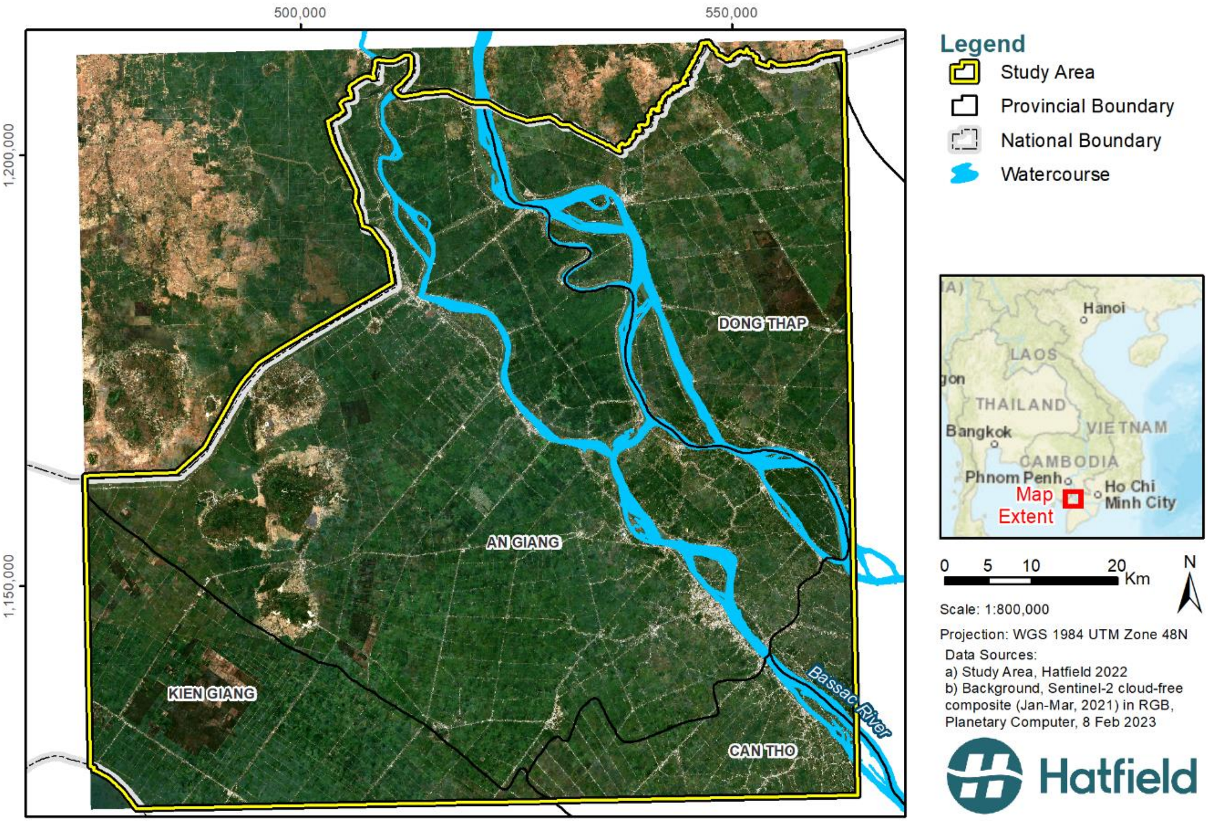
2.0 IUCN VIETNAM – MEKONG DELTA FLOODPLAIN RE-NATURALIZATION

2.1 DEMONSTRATION STUDY AREA AND OBJECTIVE

IUCN has worked closely with the World Bank, the Embassy of the Netherlands, and GIZ to develop an alternative economic, social, and environmental strategy for land use in the Mekong Delta, balancing food production with environmental values. This culminated in Resolution 120 in 2017, which provided the political basis for a transition to flood-based agriculture such as lotus and the traditional long-stem flood-adapted rice varieties. The shift toward a more natural hydrology allows for restoration of the Mekong floodplain, which has major biodiversity and climate change adaptation benefits (IUCN 2023b). While the area of natural wetlands in protected areas is very small, large areas of restored semi-natural wetlands could be considered "other effective area-based conservation measures" (OECMs) (IUCN 2023a). The initiative is to transition away from a third annual rice crop into nature-based solutions (NBS) of flood-based agriculture which will help mitigate extreme floods and droughts. This policy change involves ending investment in high dikes, which prevent flooding from the Mekong River, and enabling floodplain reconnection and restoring wetland functions. Farming in the wet season would transition to flood-based agriculture.

The objective of the PEOPLE-ER demonstration was to use the wetland function assessment tools/workflow to assess the impact of management interventions and initiative to re-naturalize a substantial area within Vietnam's upper Mekong Delta – the Plain of Reeds (PoR) and Greater Long Xuyen Quadrangle (G-LXQ). The study area and location of IUCN field survey is shown in Figure 1. The analysis needed to demonstrate the change in flood patterns between 2018-2022 to compare restoration areas to reference areas still protected by high dikes, or the restoration area before/after high dike modifications. Challenges included: (1) persistent cloud cover during the rainy season; (2) timing of the flood pulse can change be several weeks from year-to-year; (3) lack of available landscape parcel/unit geospatial data and complexity of land use within land parcels.

Figure 1 Study area and IUCN field validation survey in the upper Mekong Delta, Vietnam.



PEOPLE-ER

K:\Data\Project\ESA11339\A_MXD\ESA11339_MekongDelta_20220218_v0_3_AnS_Victor_2.mxd

2.2 DATA AND METHODS

2.2.1 EO Data

All Sentinel-1 Synthetic Aperture Radar (SAR) images were collected over the period spanning from 2018 to 2022 within the An Giang Province, situated in the Mekong Delta of Vietnam. The collected Sentinel-1 images (1146 images) were systematically divided into annual subsets, and images in each subset were aggregated by 10-day period using a median compositing method. Consequently, this procedure generated a gap-free time series of Sentinel-1 composites at regular intervals, facilitating subsequent time series analysis and classification.

Sentinel-2 Level-2A images were acquired during two distinct dry periods (January to March and April to June) in both 2020 and 2021, resulting in the creation of four cloud-free composites covering identical areas as the corresponding Sentinel-1 images. These dry-season composites are particularly adept at revealing precise landscape boundaries unaffected by flood inundation. Comprising solely four bands—blue, green, red, and near-infrared—the composites boast a high spatial resolution of 10 meters, which significantly benefits landscape segmentation tasks. The multi-date composites when stacked together serve as inputs for landscape segmentation to provide a structure for Sentinel-1 time series analysis of floodplain function changes.

2.2.2 Reference Data

Reference data were required for interpreting the distinct temporal patterns in the Sentinel-1 time series related to specific floodplain functions. Moreover, these labeled data are used for validating the classification of floodplain functions. IUCN conducted extensive fieldwork to investigate the land-use type of a selection of landscape units. In total, 639 landscape units were labelled into ten categories (see Table 1), with a significant portion identified as triple and double-rice cropping method areas that are strong indicators of floodplain connectivity in this region. Example field photos are shown in Figure 2.

Table 1 Summary of IUCN-labelled Landscape Units.

Landuse 2022	Floodplain connection	Count
Annual Crops	No (but variable)	26
Aquaculture	No (but variable)	4
Built-Up Land	No	3
Mono Rice Crop	No (but variable)	14
Perennial Trees	No	61
Rice - Shrimp	Yes (but variable)	2
Double Rice	Yes	223
Double Rice-Vegetable	Yes (but variable)	1
Triple Rice	No	303
Wetland	Yes	2

Figure 2 IUCN field validation survey (October 2023): selected photos.

Double Rice	
	
Triple Rice	
	
Perennial Trees	
	

2.2.3 Floodplain Function Cluster Analysis and Classification Methods

Technical details regarding the methods for floodplain function change analysis be found in the PEOPLE-ER Algorithm Theoretical Baseline Document (ATBD) (Hatfield Consultants 2023) and the notebooks and documentation on the PEOPLE-ER GitHub page[2].

Identification of floodplain functions were implemented on a landscape unit basis. The landscape units were delineated based on boundary information derived from cloud-free Sentinel-2 dry-season composites. To amplify boundary information, multiple-date composites were stacked together, serving as inputs for the CEWS workflow (Watkins and Van Niekerk 2019a; 2019b) specifically designed for landscape segmentation. The landscape segmentation produced a landscape unit dataset that supported our field-based analysis of floodplain function.

Utilizing the Sentinel-1 image time series, annual temporal profiles in the VH polarization band were obtained for each landscape unit, serving as signatures to discern wetland function types. t-distributed Stochastic Neighbor Embedding (t-SNE), a state-of-the-art embedding technique to visualize the similarity of temporal patterns (Maaten and Hinton 2008), was used to project the time series of into a 2-Dimensional map.

The clustering approach employed to identify predominant temporal patterns in the dataset was the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN; McInnes et al, 2017), a density-based clustering algorithm widely used for identifying clusters or groups within data, especially when dealing with datasets of varying densities and shapes.

The average time series from selected clusters served as reference temporal patterns characterizing corresponding floodplain functions. Dynamic Time Warping (DTW) (Salvador and Chan 2007) was used to measure the similarity between landscape unit time series and these reference patterns. Consequently, landscape units were assigned to floodplain function classes sharing the most similar temporal patterns.

2.2.4 Impact Assessment Method

The time series classification was implemented for the years of 2018 and 2022, covering the period before and following the resolution of a transition to flood-based agriculture, enabling the comparison of classification maps. This comparison highlighted landscape units exhibiting changes in floodplain functions over time, facilitating a summary of alterations in flood-connected areas.

2.2.5 Validation Assessment Method

The validation assessment was based on the reference samples for double-rice (floodplain connected) and triple-rice (floodplain isolated) samples due to their prevalence in the IUCN-labelled dataset and predominance in the study area. It leveraged the abundance of double and triple-rice samples to ensure a robust and in-depth examination of their temporal characteristics and patterns.

---

[2] https://github.com/PEOPLE-ER/Wetland-Function-Assessment/tree/gh-pages

Field samples exhibiting low temporal variability (see Eq 1) were not the priority for this analysis due to lack of temporal characteristics of floodplain functions. These samples were identified and not included in the classification or validation of floodplain connectivity.

$$\text{Temporal Variability} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad \text{[eq 1]}$$

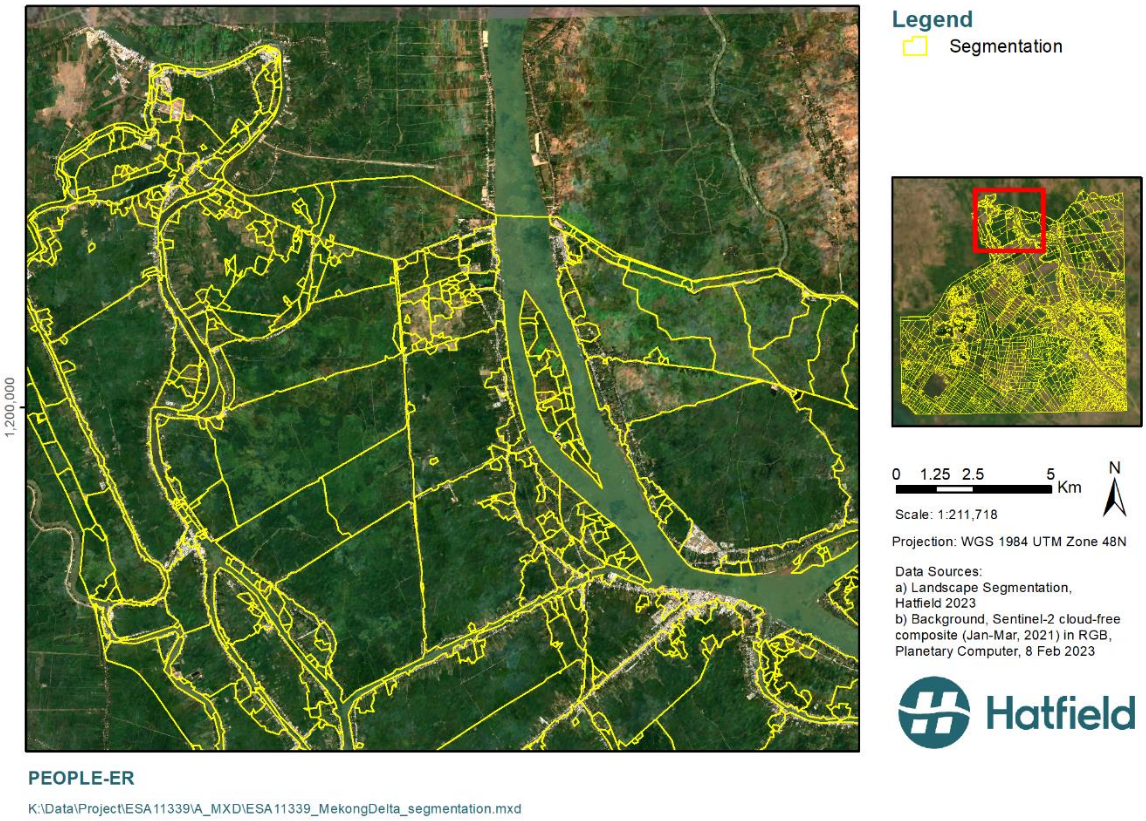
where x_i denote the one of the *n* elements in time series, and $\bar{x}$ is the average of elements.

2.3 RESULTS, VALIDATION AND DISCUSSION

2.3.1 Landscape Segmentation

The landscape segmentation of Sentinel-2 provided a suitable basis for the further implementation of the methodology using Sentinel-1 time series (Figure 3). A minimum polygon size was implemented (0.25 ha), but some limitations in the shape of polygons and their boundaries were apparent given the complex multi-scale structure of the landscape. We noted that if available, a user may want to use very high-resolution input satellite imagery or digitized field boundaries / cadastral data to provide a precise landscape structure for the analysis.

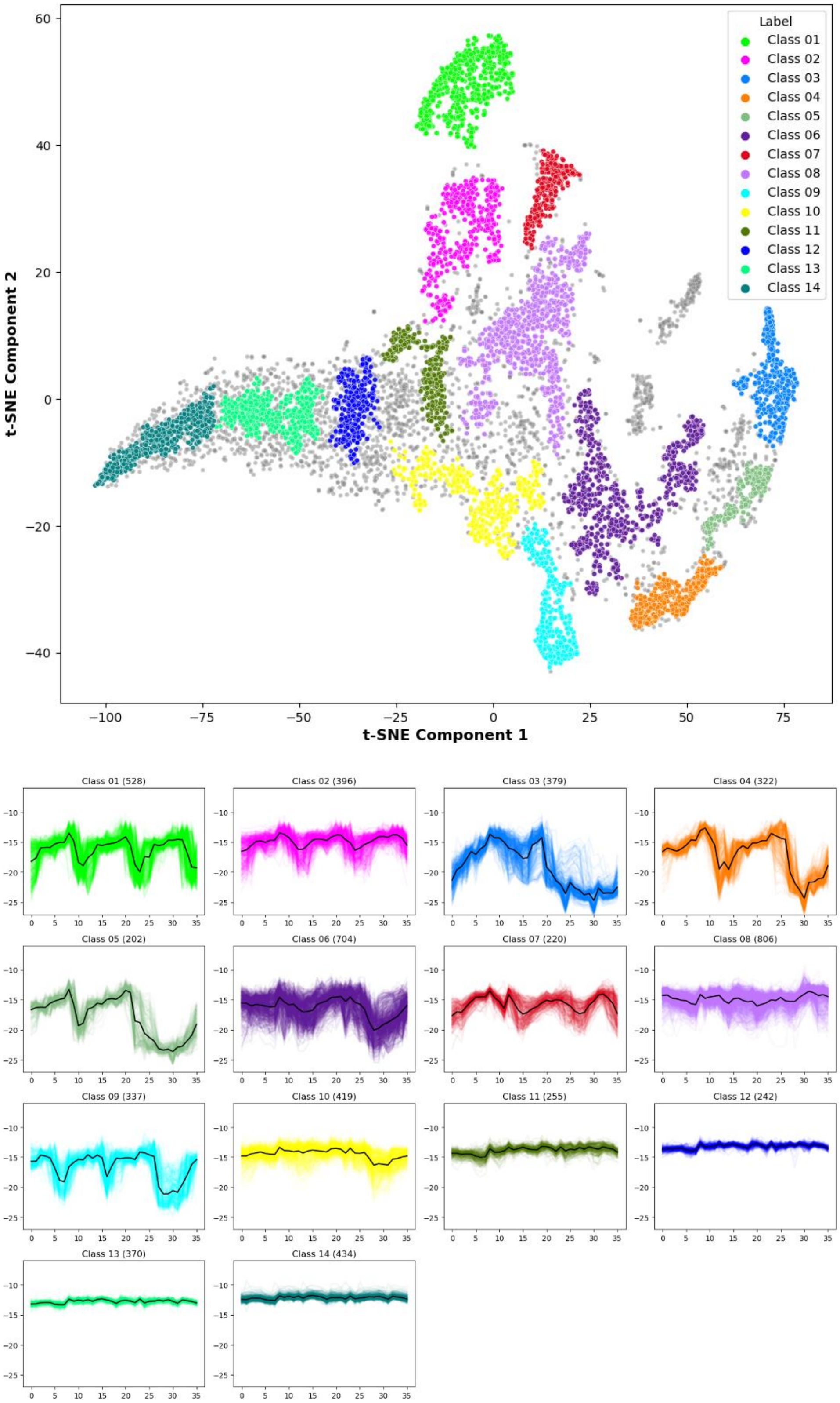
Figure 3 Landscape segmentation using Sentinel-2 in the Mekong Delta.



2.3.2 Cluster Analysis

A diverse range of distinct temporal patterns were discovered from the time series dataset through the application of t-SNE and HDBSCAN cluster analysis method (Figure 4).

Figure 4 tSNE and HDBSCAN cluster analysis results.



Upon visualization and comparison of temporal patterns derived from cluster analysis, it is evident that IUCN-labelled double and triple rice contain several subclasses, as their samples were distributed across various clusters identified by the HDBSCAN algorithm. Most of the triple-rice samples are found to belong to Cluster 1, 2, and 9, as well as majority of the double-rice samples come from Cluster 4, 5, and 6 (Figure 5). The average time series of Cluster 1, 2, and 9 is shown in Figure 6 and the average time series of Cluster 4, 5, and 6 is shown in Figure 7.

Figure 5 Cross matrix for double- and triple- rice.

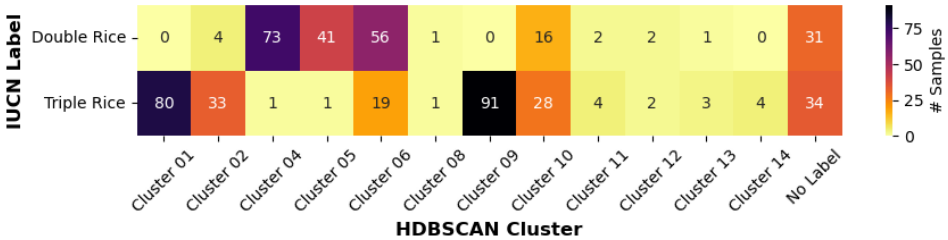


Figure 6 Time series of IUCN-labelled triple rice. Black lines represent the reference time series of the clusters.

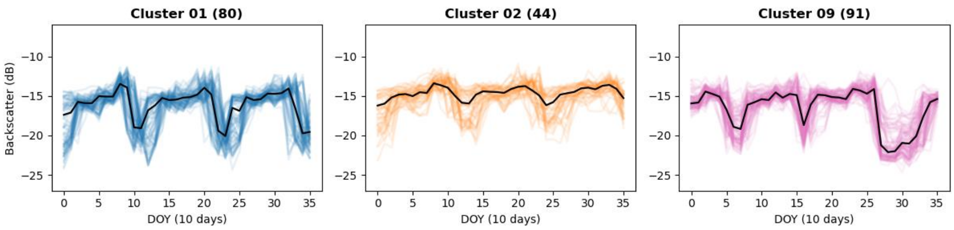
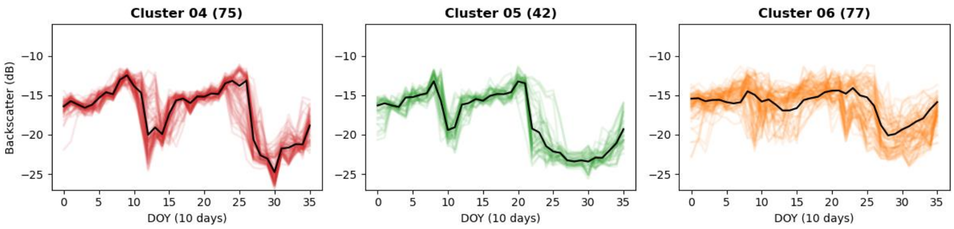


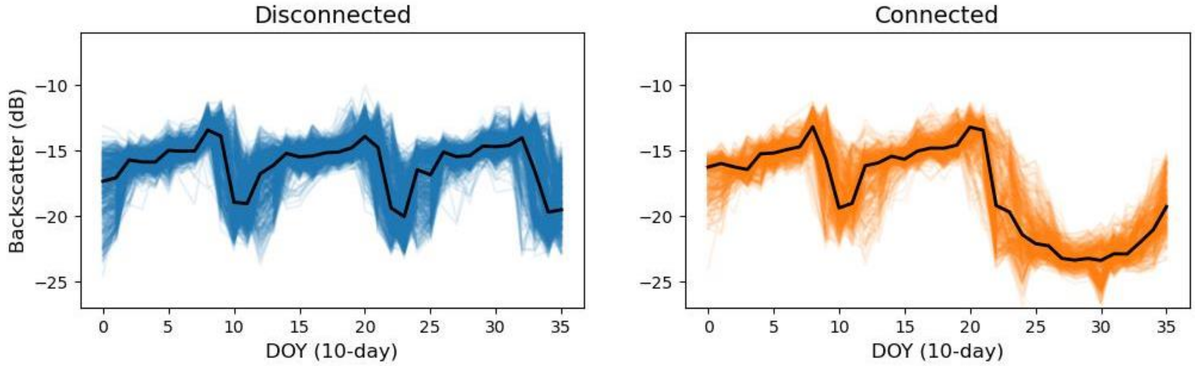
Figure 7 Time series of IUCN-labelled double rice. Black lines represent the reference time series of the clusters.



According to IUCN's labelled samples, landscape units of perennial tree (e.g., forest and plantation) and buildup areas, are usually less affected by water inundation during flood season, and often present stable backscatter reflectance across the years. Additionally, mixed land-use practices, which were observed within some large landscape units, also exhibit low temporal variation. Landscape units with high internal variability likely weaken temporal characteristics and therefore lead to “flat” median time series that represent the temporal profile of the landscape units. In future work, conducting landscape segmentation at a finer level could reduce the number of mixed landscape units and preserve more temporal characteristics within the study area.

To capture representative temporal profiles for specific floodplain functions, the average time series within each cluster was chosen as the reference. By aligning with IUCN's labels, two reference time series (depicted by black lines in Figure 8) were identified to characterize the temporal patterns of connected and disconnected floodplain areas, specifically denoting double- and triple-rice fields.

Figure 8 Reference time series (in black) for disconnected and connected floodplain function types and the analogous time series of landscape units based on DTW scores.



2.3.3 Classification Using DTW

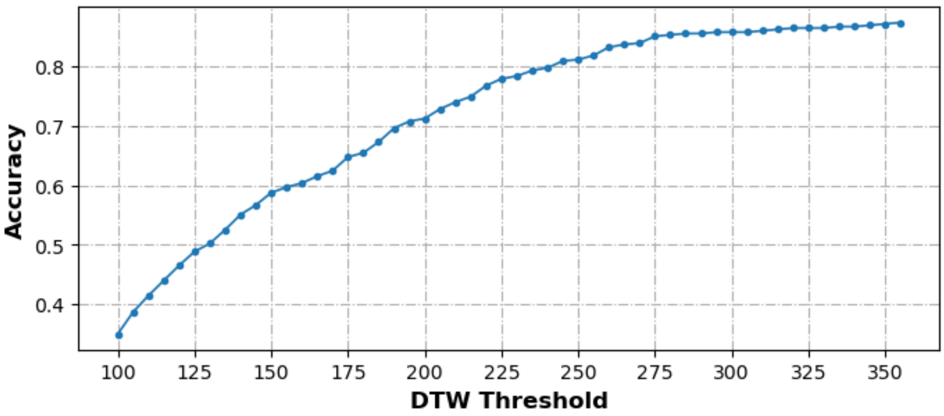
The DTW scores for the time series of selected IUCN-labelled floodplain-connected (double rice) and floodplain-isolated (triple rice) samples were calculated against the reference time series of the six clusters (i.e., Cluster 1, 2, 4, 5, 6, and 9). Subsequently, each sample was classified based on the most similar temporal pattern (indicated by the lowest DTW score) to its respective reference time series. Time series were then reclassified - those assigned to Clusters 1, 2, and 9 were relabeled as triple rice, while those assigned to Clusters 4, 5, and 6 were identified as double rice. Time series exceeding a DTW threshold against any of the selected reference time series, were classified as “others”.

The agreement between these reclassified labels and the IUCN labels served as validation for the DTW-based classification method's performance. Notably, the chosen time series, based on the DTW threshold, exhibited consistent temporal variations akin to the reference ones. This manifests the effectiveness of the DTW-based approach in identifying landscape units that share analogous temporal profiles with the established references for floodplain functions.

IUCN’s samples of low temporal variation were omitted from the validation assessment. Evaluation of the temporal variability values of the 64 samples from the two categories that are less impacted by flood inundation (i.e., perennial tree and buildup areas) revealed that 99% of those samples exhibited a temporal variability lower than 2.0. This finding led to the adoption of a threshold of 2.0 to eliminate low temporal variation samples from further analysis. In total, 434 IUCN-labelled samples from double- and triple-rice categories were selected and applied to validate the DTW-based time series classification.

The selection of the DTW threshold significantly impacted the classification accuracy. Testing thresholds from 100 to 360 revealed that an optimal threshold, yielding high agreement between IUCN's labels and classified labels, was identified at 275. Accuracy notably increased from 100 to 275; however, beyond 275, the rate of improvement slowed, stabilizing in a consistent zone (Figure 9). This suggests that 275 represents a suitable DTW threshold for achieving high agreement with IUCN field investigation. The overall accuracy of classification employing a DTW threshold of 275 is 86.93%.

Figure 9 DTW threshold and classification accuracy.



It is important to acknowledge that the reference time series were established using a 2022 dataset. However, given that the temporal patterns of floodplain functions are significantly influenced by the annual variations in precipitation and streamflow, inter-annual variability is expected among time series across different years even for identical landscape units. To accommodate inter-annual variability, it is necessary to enhance the flexibility of the DTW alignment process when applying this DTW-base classification scheme to a different year. In this use case, when conducting the classification for the 2019 dataset, the DTW window size was expanded from 5 to 7, which allowed for a more flexible alignment to reference time series.

The classification maps of triple and double rice were compared between 2019 and 2022 to identify the transition of floodplain functions within the study area. A 22% decrease in the total area of triple rice was detected between 2019 and 2022 within the entire study area (Table 2 and Table 3), with some hot spots of transition of floodplain functions observed and highlighted on the maps (see yellow boxes in Figure 10).

Figure 10 DTW-based classification for landscape units in Vietnam Mekong for 2019 and 2022.

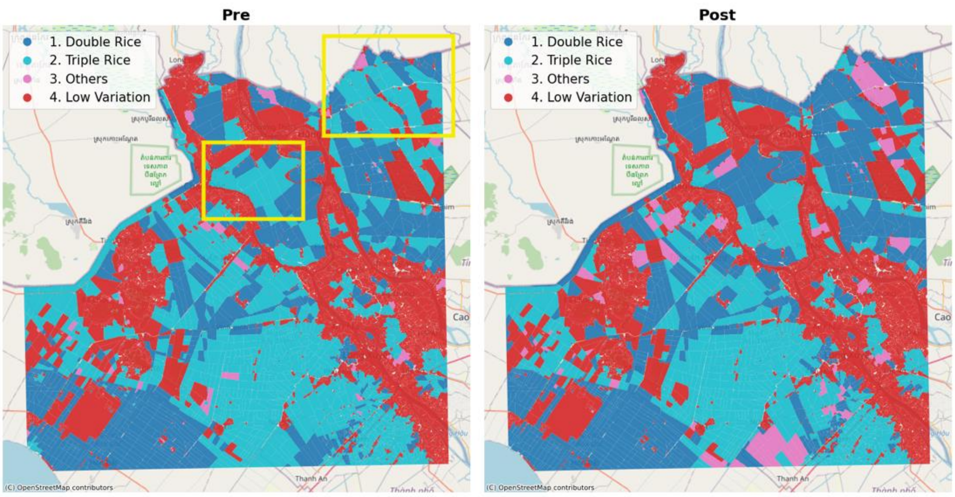


Table 2 Cross matrix of number of fields for different floodplain functions between 2019 and 2022.

		2022			2019 Total
		Double Rice	Triple Rice	Others	
2019	Double Rice	**696**	85	42	**823**
	Triple Rice	211	**835**	82	**1128**
	Others	26	37	**36**	**99**
	2022 Total	**933**	**957**	**160**	**2050**

Table 3 Cross matrix of area (km²) for different floodplain functions between 2019 and 2022.

Unit: km^2		2022			2019 Total
		Double Rice	Triple Rice	Others	
2019	Double Rice	**1318.88**	228.36	71.96	**1619.21**
	Triple Rice	605.20	**1519.46**	194.92	**2319.58**
	Others	22.31	44.01	**14.31**	**80.63**
	2022 Total	**1946.38**	**1791.83**	**281.19**	**4019.41**

2.3.4 Main Conclusions on the Vietnam Demonstration Area

The main conclusions in respect to the Early Adopter needs and the demonstration objectives can be summarized as follows:

- The PEOPLE-ER wetland function assessment tool can be effectively used to monitor the effectiveness of policies aimed at restoring natural floodplain cycles on a landscape. The IUCN-Vietnam use case demonstrated how this workflow can be used to monitor the impact and progress made by of these policies and restoration efforts. The key benefits are:
    - The ability to classify a landscape based on diverse flood cycles and wetland typologies.
    - Synoptic nature of EO data, with the potential to complete the analysis over the entire An Giang province.
    - The flexibility built into the tool enables the user to adapt the tool's workflow to best suite their needs.
- The PEOPLE-ER wetland function assessment tool can be used in a variety of computational platforms such as the Planetary Computer Hub or local desktop environments. As such, it can be used by the user in whichever environment they feel most comfortable working in.
- The PEOPLE-ER wetland function assessment tool was shared as a series of Jupyter notebooks so as to facilitate the uptake and learning that might required in order to carry out the workflow.
- To select and validate a DTW threshold, field reference data are required each year moving forwards.

3.0 REFERENCES

Hatfield Consultants. 2023. “PEOPLE-ER Algorithm Theoretical Baseline Document.” www.people-er.info.

IUCN. 2023a. “IUCN WCPA Other Effective Area-Based Conservation Measures Specialist Group.” 2023. https://www.iucn.org/our-union/commissions/group/iucn-wcpa-other-effective-area-based-conservation-measures-specialist.

IUCN. 2023b. “Transboundary Dialogue on Wetlands and Agriculture in the Mekong Delta.” 2023. https://www.iucn.org/story/202302/transboundary-dialogue-wetlands-and-agriculture-mekong-delta.

Maaten, Laurens van der, and Geoffrey Hinton. 2008. “Visualizing Data Using T-SNE.” *Journal of Machine Learning Research* 9 (86): 2579–2605.

Salvador, Stan, and Philip Chan. 2007. “Toward Accurate Dynamic Time Warping in Linear Time and Space.” *Intelligent Data Analysis* 11 (5): 561–80. https://doi.org/10.3233/IDA-2007-11508.

Watkins, Barry, and Adriaan Van Niekerk. 2019a. “Automating Field Boundary Delineation with Multi-Temporal Sentinel-2 Imagery.” *Computers and Electronics in Agriculture* 167: 105078.

———. 2019b. “A Comparison of Object-Based Image Analysis Approaches for Field Boundary Delineation Using Multi-Temporal Sentinel-2 Imagery.” *Computers and Electronics in Agriculture* 158 (March): 294–302. https://doi.org/10.1016/j.compag.2019.02.009.