Landscape Recovery in the Lower Chew- Mapping & Modelling Framework

Author: Murali Thoppil

Contents

[Introduction 2](#_Toc188882329)

[Project Background 2](#_Toc188882330)

[Objectives and Deliverables 2](#_Toc188882331)

[Data Sources 3](#_Toc188882332)

[Land Cover Classification Framework 3](#_Toc188882333)

[Ecosystem Services Assessment Framework 4](#_Toc188882334)

[Carbon Estimation 4](#_Toc188882335)

[Biodiversity Assessment 5](#_Toc188882336)

[Natural Flood Management 5](#_Toc188882337)

[Future Scenario Modelling 6](#_Toc188882338)

[Web Visualisation Framework 7](#_Toc188882339)

[Stakeholder Needs/Solution Matrix 7](#_Toc188882340)

# 

# Introduction

## Project Background

Avon Needs Trees (ANT) is a registered charity focused on creating permanent woodlands throughout the Bristol-Avon River catchment to address climate and nature emergencies. As part of DEFRA's Landscape Recovery scheme, ANT is developing the Lower Chew Valley Landscape Recovery Project, which aims to transform 422 acres into the South West's largest new woodland. This project requires robust methods to assess and monitor landscape changes and their impacts on ecosystem services.

The objective of this work is to develop a cost-effective and replicable remote sensing methodology that can assess current landscape conditions and model future scenarios in the Lower Chew Valley. This methodology will focus on creating detailed land cover maps using machine learning techniques and developing models to evaluate ecosystem services, particularly key carbon stocks, ecological status, and natural flood management potential. The approach will utilise open-source data and tools to ensure accessibility and replicability for future projects

## Objectives and Deliverables

This document outlines the methodological framework for developing a landscape assessment and monitoring system for the Lower Chew Valley Landscape Recovery Project. The framework encompasses three core components: (1) a machine learning approach for detailed land cover classification using open-source satellite data, (2) ecosystem service modelling methods for assessing carbon stocks, biodiversity, and flood management potential, and (3) a web-based visualisation system for stakeholder engagement.

The main aim of this framework is to document the intended approach, data requirement and outputs from the project. While this will be followed as close as possible, the actual implementation and techniques used during the project might change based on data access, model capability and software features. The realised method will be detailed in the guidance document as per deliverable 2.

# Data Sources

The following data sources will be utilised to demonstrate the web map and modelling. Open source and license free datasets for charities will be utilised where possible

* Copernicus Sentinel 2 imagery
* DEFRA LIDAR 1m Digital Surface Maps
* UKCEH Maps- Land Cover, Hedgerows, Crops, Flood
* UKCEH Gridded Estimated Area Rainfall- 1km resolution
* DEFRA Historic Flood Map
* British Geological Survey's (BGS) Permeability Index
* National Soil Resources Institute's Soilscapes dataset

When selecting satellite imagery for habitat classification, several options exist with varying spatial, temporal, and spectral resolutions. Landsat imagery, while offering a long historical record dating back to the 1970s, provides 30-meter spatial resolution which can be too coarse for detailed habitat mapping at the field scale. Commercial satellites like WorldView and Planet offer very high spatial resolution (sub-meter to 3-meter) but are costly and often lack the spectral bands crucial for vegetation analysis. MODIS provides daily coverage but at 250–500-meter resolution, making it suitable only for broad-scale studies. Aerial photography can offer extremely high spatial resolution but typically lacks the spectral information needed for automated classification and is expensive to acquire regularly.

Sentinel-2 emerges as the best choice for the Lower Chew Valley project. Its 10-meter spatial resolution in key bands is sufficient for field-level habitat mapping while providing coverage of large areas efficiently. The satellite's 13 spectral bands, including red-edge bands particularly sensitive to vegetation characteristics, enable detailed discrimination between habitat types. Importantly, Sentinel-2 data is freely available through the Copernicus program, making it cost-effective for long-term monitoring. These characteristics make Sentinel-2 particularly suitable for developing repeatable habitat classification methods that can be applied across the landscape recovery project area.

## Land Cover Classification Framework

Land classification over a large area can be costly and time-consuming requiring expert knowledge as well as map building technical skills. Over the past few years with the advance of high-resolution satellite imagery and increased computational power, machine learning methods have begun to be used to speed up the estimation and classification process.

Remote sensing for habitat classification can employ various machine learning approaches, each with distinct characteristics and applications. Support Vector Machines (SVM) excel at finding boundaries between habitat classes but can struggle with complex, overlapping habitat types. Neural Networks offer powerful pattern recognition capabilities but require extensive training data and computational resources, making them less practical for regional-scale habitat mapping. Maximum Likelihood Classification, while straightforward and widely used, assumes normal distribution of the data and can be less effective with the non-linear relationships common in natural landscapes.

Random Forest, a combination of averaging the outputs of multiple decisions trees, emerges as the optimal choice for the Lower Chew Valley habitat classification for several key reasons. It handles the non-linear relationships between environmental variables effectively, manages the inherent noise in remote sensing data, and performs well with limited training data - a common constraint in habitat mapping projects. The algorithm's ability to estimate classification confidence is particularly valuable for identifying areas that may need field verification. Furthermore, random forest provides insights into variable importance, helping understand which environmental factors most strongly influence habitat distribution. Its computational efficiency and robust performance across different environmental conditions make it particularly suitable for landscape-scale applications where habitats can be complex and heterogeneous.

The Semi-Automatic Classification Plugin in QGIS provides a user-friendly interface for implementing random forest classification. The workflow begins with data preparation, where Sentinel 2 imagery is collected and processed. Training areas for known habitat types are then defined, allowing the algorithm to learn the characteristics of different habitats. The classification process itself involves configuring the random forest parameters and executing the algorithm, followed by validation using independent test data.

After the initial classification, the results undergo rigorous accuracy assessment. This involves comparing the classified habitats with known ground truth data and calculating various accuracy metrics. If needed, the classification can be refined by adjusting parameters or adding more training data. Validation data will employ ANT ground surveys to ensure accurate classification and where necessary, Natural England land cover maps will be used in conjunction for areas where ground truth is not available.

The resulting habitat map will serve as the foundation for assessing habitat connectivity and analysing ecosystem services. Furthermore, it will provide essential baseline data for modelling future landscape scenarios and understanding potential changes in ecosystem service provision Training Data Collection

# Ecosystem Services Assessment Framework

## Carbon Estimation

The Intergovernmental Panel for Climate Change Good Guidance documentation will be used to estimate the carbon stocks of the following types:

* Aboveground biomass
* Belowground biomass
* Soil Organic Matter
* Dead Organic Matter

Default emission factors and activity data detailed in the Land Use and Land Use Change methodology will be implemented to calculate carbon stocks in the project area. Where possible, national emission factors from research journals will be investigated for further accuracy

## Biodiversity Assessment

When evaluating connectivity modelling approaches for landscape recovery projects, several key methods stand out, with Factorial Least Cost Paths emerging as the most suitable choice. While Circuit Theory, Graph Theory, and Spatially Explicit Population Modelling all offer valuable insights, each has limitations that make them less ideal for this landscape recovery project.

Factorial Least Cost Paths stands out as the optimal choice for several compelling reasons. This approach excels at incorporating multiple environmental variables such as land cover, slope, and flood risk, while maintaining computational efficiency across large landscapes. Its ability to weight different factors affecting connectivity makes it particularly valuable for landscape recovery planning, where multiple objectives need to be balanced. The method also integrates seamlessly with GIS systems, making it ideal for web mapping applications and visualisation.

While Circuit Theory offers sophisticated analysis of connectivity bottlenecks by modelling landscapes as electrical circuits, it becomes computationally intensive over large areas. Graph Theory, though excellent for analysing network properties, tends to oversimplify landscapes into discrete nodes and edges. Spatially Explicit Population Modelling, while powerful for species-specific analysis, requires extensive data and may be too complex for initial landscape recovery planning.

The practical advantages of Factorial Least Cost Paths make it particularly suitable for this project. It works efficiently with remote sensing data, scales well to the 800-hectare study area, and produces clear outputs that stakeholders can easily understand. Importantly, QGIS offers a dedicated Least Cost Path plugin that can be easily implemented, making the analysis more accessible and streamlined. The plugin provides a user-friendly interface for creating cost surfaces and calculating optimal paths, significantly reducing the technical complexity of implementation

## Natural Flood Management

This flood management layer will be developed by using a systematic approach to assess natural flood management potential across the Lower Chew Valley landscape using multiple environmental factors. The framework will integrate five key components that influence water retention and movement: underlying geology, soil characteristics, topographic slope, habitat type, and land management practices. Each component will be scored on a standardised 1-5 scale, where higher values indicate better flood management capability.

The geological assessment will utilise the British Geological Survey's (BGS) Permeability Index, which provides standardised classifications of how water moves through different rock types. This will be complemented by soil data from the National Soil Resources Institute's Soilscapes dataset, which characterises soil drainage properties. Topographic influence will be assessed using slope calculations derived from digital elevation models, with flatter areas receiving higher scores due to their greater water retention potential. Habitat types will be classified using remote sensing data, with scores assigned based on vegetation structure and known hydrological properties. Land management practices will be assessed using detailed farm-level data collected during the baseline surveys of the seven farms within the study area.

Each component will be weighted based on its relative influence on flood management (as consulted on by experts). These weights reflect both the immediate and long-term impacts of each factor on water movement through the landscape. The weighted scores will be combined to create a composite flood management potential index, visualised as a heat map across the study area. This approach will allow for both broad-scale assessment of flood management capability and identification of specific areas where interventions might be most effective.

The framework will be designed to be replicable and adaptable, utilising publicly available datasets where possible. While the implementation will focus on the Lower Chew Valley, the methodology could be applied to other Landscape Recovery projects across England. However, it's important to note that this strategic-scale assessment should be supported by field validation before making site-specific management decisions.

Alternative flood modelling approaches were considered, including the SCIMAP-Flood model and the open-source Fast Flood app. While these tools provide sophisticated hydrological modelling capabilities, they were deemed unsuitable for this project's specific needs. SCIMAP-Flood, while powerful, requires significant technical expertise and computational resources that may not be readily available to all Landscape Recovery projects. The Fast Flood app, though user-friendly, lacks the ability to model landscape interventions such as increased woodland cover or changes in land management practices - features that are central to the Landscape Recovery scheme's objectives. The proposed framework instead aims to strike a balance between scientific robustness and practical usability, making it more suitable for stakeholder engagement and adaptive management planning.

# Future Scenario Modelling

A forecasting model will be constructed for the woodland growth yields to determine the land classified to woodland, carbon stored over time as well as its impact on the ecological status of the project site. This scenario forecast will be constructed for the current layout of the project site as well as the two vision output maps. Prediction maps will be produced for 5, 10, 20, 30 year periods from the current date in order to assess how much carbon yield is generated over time and will primarily use the growth yield forecast data from the Woodland Carbon Code.

Modelling including the climate change scenario forecasts will be assessed, however due to the technical and time constraints of implementing the climate scenarios this will be carried out as a feasibility research and analysis module for possible further implementation during a separate work package.

# Web Visualisation Framework

## Frontend Mapping

The following options will be considered for the web map application: leaflet.js, QGIS Server and Felt. While all options have positives and negatives, emphasis will be placed on ensuring that the service is low cost and easy to use by a non-technical audience.

While leaflet.js offers an open-source fully customisable implementation of webmaps, the coding of the application itself requires knowledge in programming languages such as JavaScript, html and CSS. Any amendments or addition of features will require changing the codebase and might cause difficulties when working on other projects.

QGIS Server offers the possibility of hosting web maps and dashboards online and is readily tied into the analysis QGIS software itself. While this would be the automatic choice for the web map, ANT is subscribed to the FELT mapping application that offers a similar service. Using FELT would therefore allow the LVC web map to be based in the same location and offer ease of use to users who are already familiar with its toolbox and options. The main interactive feature of the webmap will allow the user to switch between the layers for the habitat classification, ecosystem services and forecasts. Where possible, graphs, charts and quantitative figures will be provided to ensure further information as to trends and seasonality patterns within the data.

## Output Formats

The web map will be hosted and accessible online in a graphical interactive format. PDF maps will also be produced of the different layers of the web map.

# Stakeholder Needs/Solution Matrix

The following highlights the stakeholder needs/requirements as per the project document and initial discussions. Each requirement is followed up by the solution provided and implementation.

## Baseline Mapping Needs

1. Need to develop replicable machine learning method for land cover mapping

*The solution will look to use a random forest machine learning method that classifies the LCV site according to the UKHab land classes.*

1. Need to use open-source data (e.g., Copernicus Sentinel)

*The solution will source data from Sentinel 2 satellite imagery (land classification), DEFRA 1m Digital Terrain Maps (flooding and hedgerow analysis) and where possible datasets that are open license to non-profit organisations. Data sources from ANT will also be utilised where necessary for training and validation.*

1. Need to use open-source software (e.g., QGIS)

*The primary software used in the analysis will be based in QGIS and FELT( as used by ANT)*

1. Need to follow UK habitat classification nomenclature

*The solution will classify land classes according to Level 3 UKHab classes*

## Ecosystem Service Assessment Needs

1. Need methods to assess above ground carbon stocks

*Carbon stocks will be analysed according to the IPCC LULUCF definitions and emission factors (tier 1). These will include: aboveground biomass, belowground biomass, soil organic matter(mineral soil) and dead organic matter where applicable.*

1. Need methods to assess habitat connectivity indicators

*The solution will analyse habitat connectivity according to the least cost paths method which quantifies the resistance level for woodland species to move across the LCV site.*

1. Need methods to assess natural flood management potential

*The solution will analyse the flood management potential using a rule based analysis method showing surface runoff potential. This model will analyse different land management scenarios and it’s effect on slowing flooding intensity for key infrastructure in the area.*

1. Need methods that can be replicated by others

*All methods mentioned above will be documented for replication in future project work including script and datasets.*

## Future Modelling Needs

1. Need to model scenarios at 5, 10, 20 and 30+ year intervals

*Future models for the business-as-usual and two vision maps for the project site will be carried out on woodland growth yields to determine the land transitioned to woodland and carbon stock changes*

1. Need to incorporate UNFCCC climate pathways

*UK climate forecasts will be researched to determine which model, parameters and datasets will be required for further implementation in future project work for the habitat and ecological services models.*

## Visualisation & Interface Needs

1. Need web-linked user interface for non-technical users

*A webmap will be produced with the ability to switch between the different mapping layers. This webmap will utilise a GUI to enable non-technical users to easily navigate the application.*

1. Need PDF versions of all maps

*PDFs of all mapping baseline and scenarios will be delivered along with the webmap.*

1. Need clear visualisation of baseline/future scenarios

*The baseline/future scenarios will be visualised using data techniques such as choropleth maps, showing carbon intensities, flood risks and biodiversity hotspots. Additional graphs or charts will be visualised where necessary.*

1. Need training materials for future users

*Documentation will be produced to guide users on the overview of building similar web-based apps and running model analyses.*