Australian Government

Australian Institute of Marine Science

Field photogrammetry in 4D: *Digitisation and 2D metric extraction*

Reef Restoration and Adaption Program (EcoRRAP)

Standard Operational Procedure Number 17 (No. 3 of series)

Maren Toor, Marine Lechene, Madison Becker, Tiny Remmers-Barry, Sophie Gordon, John Stratford, Renata Ferrari

AIMS: Australia's tropical marine research agency.

www.aims.gov.au

Australian Institute of Marine Science

PMB No 3
Townsville MC Qld 4810

PO Box 41775
Casuarina NT 0811

Indian Ocean Marine Research Centre
University of Western Australia, M096
Crawley WA 6009

This report should be cited as:
Toor, M., Lechene, M.A., Becker, M.L., Remmers-Barry, T., Gordon, S., Stratford, J., Ferrari, R. (2025) Field Photogrammetry in 4D: Digitisation and 2D metric extraction. Reef Restoration and Adaptation Program (EcoRRAP). Standard Operating Procedure 17 (No. 3 of series). Australian Institute of Marine Science, Townsville. (77pp).

Project Leader shall ensure that documents have been fully checked and approved prior to submittal to client				
Revision History:		*Name*	*Date*	*Comments*
1	Prepared by:	Maren Toor	10/4/2025	
	Contributed by:	Marine Lechene Madison Becker Tiny Remmers-Barry Sophie Gordon John Stratford	10/4/2025	
	Reviewed by:	Renata Ferrari	15/4/2025	
	Reviewed by:	Manuel Gonzalez Rivero	20/5/2025	
	Approved by:	David Wachenfeld	22/5/2025	

Cover photo: An orthomosaic with segmented coral colonies, colours represent different coral taxa. Image credit: M. Toor.

Acknowledgement

This work was undertaken by the Ecological Intelligence for Reef Restoration and Adaptation subprogram (EcoRRAP) of the Reef Restoration and Adaptation Program (RRAP). Funded by the partnership between the Australian Governments Reef Trust and the Great Barrier Reef Foundation, partners include: the Australian Institute of Marine Science, CSIRO, the Great Barrier Reef Foundation, Southern Cross University, the University of Queensland, Queensland University of Technology and James Cook University.

The RRAP partners acknowledge Aboriginal and Torres Strait Islander Peoples as the first marine scientists and carers of Country. We acknowledge the Traditional Owners of the places where RRAP works, both on land and in sea Country. We pay our respects to elders; past, present, and future; and their continuing culture, knowledge, beliefs, and spiritual connections to land and sea Country.

We specifically acknowledge and thank the following Traditional Owners of sea Country that this report relates to:

Location	**Traditional Owner Group**
Torres Strait	Masigalgal, Porumalgal, Warraberalgal
Northern Great Barrier Reef	Gunggandji, Ngurruumungu, Gingaal
Central Great Barrier Reef	Manbarra, Bindal
Southern Great Barrier Reef	Woppaburra, Bailai, Gurang, Gooreng Gooreng, Taribelang Bunda

CONTENTS

LIST OF FIGURES

LIST OF TABLES

SUMMARY

The Ecological Intelligence for Reef Restoration and Adaptation Program (EcoRRAP) team at the Australian Institute of Marine Science (AIMS) quantifies natural rates of ecological and genetic reef recovery and adaptation in response to acute and chronic disturbances, as well as key environmental variables related to different coral reef communities. This information is used to inform the Reef Restoration and Adaptation Program (RRAP) restoration interventions. The RRAP is a collaboration across many research institutes and experts, managed by AIMS.

This document is Standard Operational Procedure (SOP) Volume 3, produced by the EcoRRAP sub-program at AIMS. It details the standard procedures for generating and extracting demographic data of benthic sessile organisms from two-dimensional (2D) orthomosaics produced using photogrammetry, at sub-mm resolution over extents of 75 - 1500 square meters with high precision in temporal co-registration. This SOP specifically details the workflow for generating, extracting, and managing data using the software programs TagLab (Pavoni et al. 2022), ArcGIS Pro (ESRI, 2023), and R (R Core Team, 2023).

An introduction to the aims, theoretical background, and sampling design of EcoRRAP is provided in the first SOP of this series (SOP 1, Table **1**). Details for other surveying methods used by EcoRRAP, including field-based image collection, post-fieldtrip image processing and 3D model processing can be found in a series of Standard Operational Procedures described in Table **1** and are published online at: AIMS's SOP page (Reef monitoring sampling methods | AIMS), EcoRRAP Metadata records (EcoRRAP Metadata), and the EcoRRAP Website (EcoRRAP | gbrrestoration.org).

Table 1. EcoRRAP 3D photogrammetry tasks and associated standard operating procedures (SOPs).

Task	Associated SOP
Overview and in-field workflow	Field photogrammetry in 4D: No. 1 of series
Model processing	Field photogrammetry in 4D: No. 2 of series
Digitisation and 2D metric extraction	Field photogrammetry in 4D: No. 3 of series (current doc.)

Information regarding data generated by the EcoRRAP program can be accessed through the Australian Institute of Marine Science's metadata records (EcoRRAP Metadata). Additional links to project outputs can be found throughout this document. The EcoRRAP Database (internal document), data management files, and folder templates can be found here: EcoRRAP Data Management Templates, while processing scripts can be accessed through the AIMS GitHub page.

1 INTRODUCTION

1.1 Overview

***Note**: A detailed overview of the EcoRRAP subprogram, sampling design, and image collection techniques used are provided in the first SOP of this series "SOP 1: Overview and in-field workflow" (Gordon et al. 2023).*

The Reef Restoration and Adaptation Program (RRAP) brings together leading experts from Australia and around the world to help protect the future of the Great Barrier Reef, other Australian reefs, and coral reefs globally. The 'EcoRRAP' subprogram aims to maximise the success of restoration interventions by advising on the 'what', 'where', and 'when' of interventions, and by filling crucial gaps in ecological knowledge of the Great Barrier Reef (GBR)(EcoRRAP).

EcoRRAP uses close-range photogrammetry to quantify structural complexity, benthic communities, and demographic rates of corals on reefs across spatial and temporal scales. Two key outputs are created from the images collected by EcoRRAP: (1) 3D digital surface models (DSMs), used to quantify landscape metrics of habitat structural complexity, and: (2) 2D orthomosaics, used to quantify benthic community composition and demographic rates from corals of several taxa and morphologies. All photogrammetry outputs are generated using Structure from Motion (SfM) algorithms (Ferrari et al. 2016; Aston et al. 2022; Lechene et al. 2024; Gordon et al. 2023), which locate and track correspondence between images and use these trajectories to reconstruct their location in 3D space and thereby create representations of reef topography. The incorporation of model co-registration techniques further enables changes in 2D and 3D outputs to be precisely examined to describe changes in landscape metrics and community compositions, and to quantify demographic rates of benthic taxa with mean precision of 1.37mm (Lechene et al. 2024).

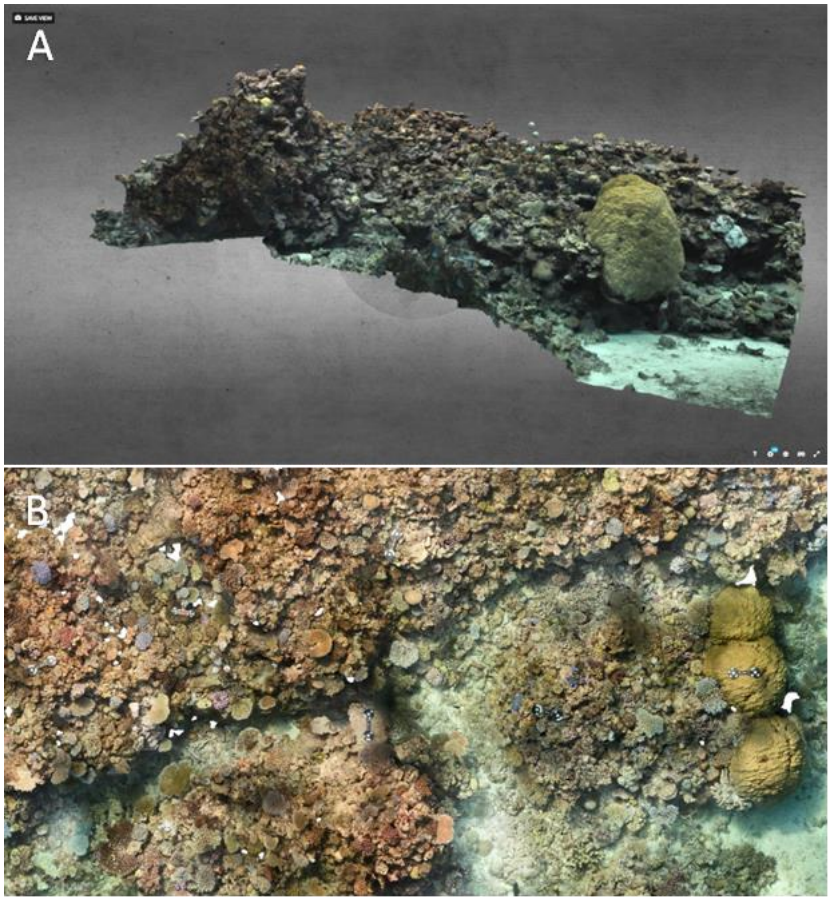

Figure 1. Three- and two-dimensional representations of the reef topography at an EcoRRAP plot produced using Structure from Motion (SfM) techniques. A) A 3D digital surface model (DSM), and B) a 2D orthomosaic. Image: S. Gordon, reproduced from SOP No. 1.

1.2 How to use this Standard Operational Procedure (SOP)

To quantify benthic community composition and demographic rates of coral morphologies, data is generated and extracted from the 2D orthomosaic products that are produced through the photogrammetry workflow. In this workflow, this is achieved through digitisation of orthomosaics, here defined as a two-step process of segmentation: the generation of a polygon around an object; and annotation: the classification of the object into a predetermined class. Select coral morpho-taxa have been digitised within EcoRRAP orthomosaics that provide key spatial and temporal demographics for reef communities. Such metrics include colony size (measured in 2D planar area), species density, colony survival, colony growth (change in 2D planar area across time) and more.

This workflow uses the AI-powered segmentation software, TagLab (v2022.02.11-v2025.2.15, Pavoni et al. 2022), to digitise coral colonies and extract demographic data. TagLab is specifically designed to support the analysis of large orthomosaic images created through a photogrammetric pipeline and provides semi- and fully-automatic segmentation modality to accelerate and optimise the data generation and extraction processes, speeding the process by at least two times what it would be using an alternative manual approach (e.g. Photoshop) (Pavoni et al. 2022).

This SOP explains how to use the software TagLab, including AI-assisted and fully automatic segmentation, to digitise sessile benthic organisms in an orthomosaic. It also covers the metric extraction and temporal dataset generation pipeline required to produce an analysis-ready dataset for coral demographics.

The EcoRRAP photogrammetry workflow consists of three key stages, here presented in three SOPs (Figure **2**, Table **1**):

1. Field-based data collection and model building
2. Office-based model building using HPC
3. Metric extraction

The final stage (metric extraction) can be divided into 3 data outputs (Figure **2**, purple):

1. Coral demography dataset (this SOP)
2. Benthic community composition (SOP in prep.)
3. Structural complexity metrics (SOP in prep.)

The workflow described in this SOP covers data output 1 of the metric extraction phase of the EcoRRAP photogrammetric workflow: the generation and extraction of demographic data from 2D orthomosaics. All prior steps are presented in SOP 1: Overview and In-field Workflow, and SOP 2: 3D Model Processing, and should be completed before proceeding with the steps presented in this SOP.

As with SOP 1 & 2, some aspects of this manual are specific to the equipment and aims of the EcoRRAP program, however this SOP is also designed to provide information for general use. This document is intended to be used as a guide to assist users with the process of digitising objects from 2D images (here focusing on orthomosaics, however single images can also be used).

For the purpose of this SOP, 'digitisation' is defined as a two-step process that consists of: 1) segmenting the object of interest, in this case coral colonies, and; 2) annotating the object by assigning it a class. Each step is defined as follows:

- Segmentation: the generation of a polygon around an object that aligns with the object borders
- Annotation: the classification of the object within the polygon (from a class in a predetermined dictionary)

Polygons retain information about the planar area of the object (strictly 2D) and can be used to monitor changes in two dimensional size through time.

At the time of writing, the TagLab version in use is v2025.2.15, and as such, instructions are reflective of this, except where stated otherwise. TagLab is updated regularly, and future versions may require additional or alternative steps than presented here and alternative workflows may be made available by the developers. A bi-annual review of this SOP will be conducted during the life of the sub-program (see disclaimer page).

Procedures that are regularly updated, and/or are AIMS specific, are described in AIMS 3D Modelling OneNote (internal link: AIMS 3D Modelling OneNote) rather than in the current document to ensure the current SOP remains relevant and useful. Some steps will also require an AIMS user account with the right approvals and access to the AIMS server "PEARL." If you are using a non-AIMS issued computer and want to access the AIMS computing network you will need to establish a remote connection (described in the AIMS 3D Modelling OneNote). If you are using your own network, you will need to ensure correct set up is done before following the steps in this SOP.

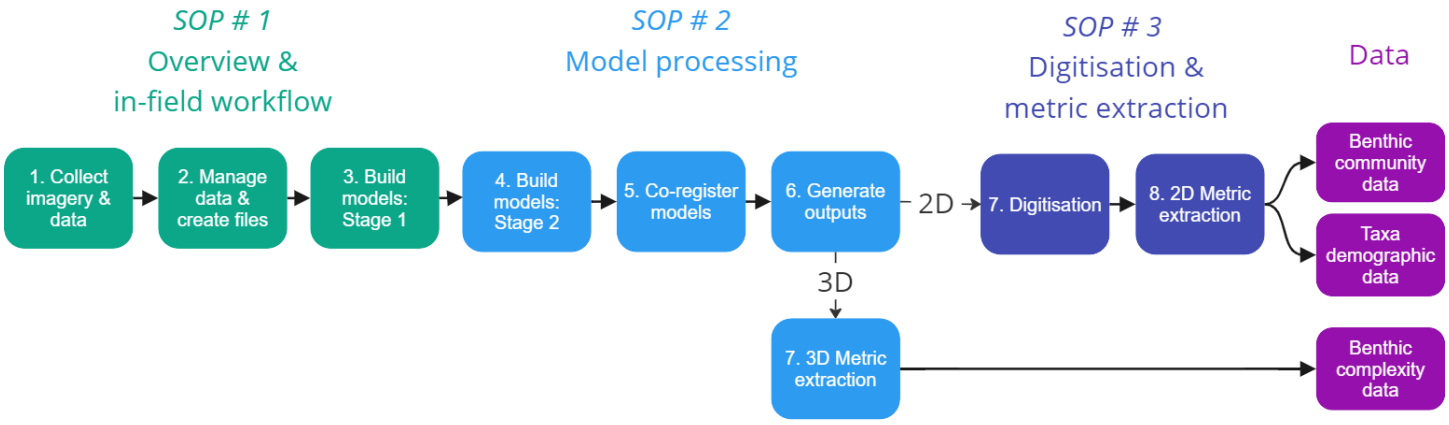

Figure 2. EcoRRP photogrammetry workflow, related SOPs, and three key data outputs. Image: S. Gordon, reproduced from SOP No. 1.

2 HARDWEAR AND SOFTWARE

2.1 Hardware

All steps and settings recommended are the result of rigorous testing and are strongly influenced by available computing capacity. All digitisation steps were designed to be completed locally on consumer-grade desktop computers or high-powered laptops. Computers with lower specifications than Table **2** may also be able to follow the described workflow, however, the time required to complete certain steps may increase significantly.

Equipment required to complete the digitisation workflow presented in this SOP is listed in Table **2**. The specific models listed are suggestions only, alternative providers and models can be substituted as required.

Table 2. List of equipment and specifications required to complete workflow presented in this SOP. Equipment with an asterisk (*) are not required to complete the workflow but are recommended to improve accuracy.

Equipment	Use in workflow	Recommended specifications
Computer	Desktop or laptop required to access software	Windows (for ease of use with TagLab) 64GB memory (low RAM, e.g. 16GB, can lead to issues with insufficient memory) NVIDIA RTX 2000 Ada, 8GB video card
Mouse	Improves segmentation accuracy and ease of software use	Ergonomic mouse
Touch screen display with pen*	Drawing and editing polygons Improves segmentation accuracy	Wacom Cintiq 22 display (or similar) Wacom Pro Pen 2 (or similar)
Pen tablet*	Drawing and editing polygons Improves segmentation accuracy Can be used on flat surface	Wacom One tablet (or similar) Wacom One pen (or similar)
Hard drive/storage	Storage for orthomosaics/images, TagLab projects, TagLab data outputs	Dependent on number of orthomosaics and data As a reference, EcoRRAP orthomosaics are ~2GB, TagLab projects are ~500-2000KB, and TagLab outputs (datatables and shapefiles) are between 15-2000KB each A minimum of 2.1GB is required for one EcoRRAP plot

2.2 Software and applications

TagLab has multiple software dependencies required to run and/or perform all available functions. All software required for this workflow are listed in Table **3**. Full instructions on how to download, install, and operate each software program are listed in section 3.2.1.

Table 3. List of software required to conduct EcoRRAP digitisation workflow. Software with an asterisk (*) require a license.

Provider	Software	Use in workflow
ISTI-CNR, Visual Computing Lab	TagLab	Software platform used to digitise 2D coral colonies
Anaconda	Anaconda Navigator	Used to create and manage python-based environments. Here used to create TagLab-specific environment with packages required to run TagLab (64bit python will be downloaded in Navigator environment)
	Anaconda Prompt	Terminal used to load required python environment and launch TagLab.
Nvidia	CUDA Toolkit	Used for GPU-accelerated computations Recommended but not required – where an Nvidia graphics card is not available CPU will be used
Microsoft	Visuals C++ Build Tools	Required in TagLab installations due to a dependency from the package pycocotools
	Visual Studio Redistributable	Required if PyQt package is not working properly (see TagLab's installation wiki on GitHub)
ESRI	ArcGIS Pro*	Used to link temporal data by conducting a spatial join on multiple shapefiles. Recommended for large datasets but not required, see section 3.6.2 for more information.

2.3 Prerequisites

This is the third SOP in the Field Photogrammetry in 4D SOP series. Prerequisites include the outputs from both previous SOPs in this series, including the field-based image collection in SOP 1 (Gordon et al. 2023) and the data processing outputs (3D models and 2D orthomosaics) produced in SOP 2 (Gordon et al. 2024). However, this SOP may still be relevant to users with alternative methods to produce 2D orthomosaics or to segment objects from single images.

In addition to the above outputs, if this workflow is to be used to digitise benthic reef communities, an intermediate to expert level of identification of reef benthos is required, depending on the level of identification desired.

3 WORKFLOW

3.1 Summary

The generation of demographic data for coral colonies (or any defined benthic sessile organism) from orthomosaics using TagLab requires the following steps (covered in detail in the following sections):

1. Installation of TagLab and software dependencies
2. Project setup in TagLab
3. Digitisation of coral colonies in an orthomosaic at time 0 (temporal reference)
4. Transfer of colonies through time on subsequent surveys after reference (T_1, T_2, T_3, etc.)
5. Data exportation
6. Data management and quality assessment/quality control (QAQC)
    a. For a single timepoint (e.g. no additional surveys through time)
    b. For multiple time points (e.g. the combination/joining of data across surveys through time)

This section covers all steps required to produce an analysis-ready demographic dataset, starting from the installation process of required software programs to the final QAQC steps taken by EcoRRAP to ensure data quality. Subsections include general information on how to use required software programs, such as TagLab, to generate colony data, and ArcGIS Pro, to link colony polygons through time. Each subsection also includes information specific to the EcoRRAP workflow, such as criteria used to determine whether a colony can be digitised and what notes are used to track colony status. Steps for non-EcoRRAP users should be amended to suit the user's specific requirements.

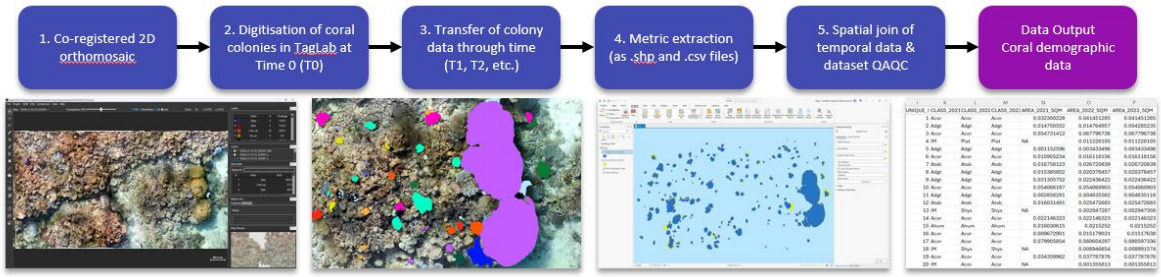

Figure 3. EcoRRAP coral demographic data generation workflow and key dataset output.

3.2 Installation and setup

The current section covers the software installation and setup required to conduct the full EcoRRAP digitisation workflow using the program 'TagLab' TagLab (cnr.it). For more information regarding each software program and tools used, please visit their respective webpages and links provided throughout this document.

Note that code to be written in command lines during the installation and launch processes is written in **bold**. This syntax can be copied exactly, except where specified.

3.2.1 Installations

The following software are presented in the recommended order of installation. Make sure all steps are completed prior to commencing the next stage to ensure software is fully operational.

Note that if using an AIMS-issued computer, users will need administrator privileges on the computer to install any software. Use the Elevate program to temporarily enable administrator mode (enabled for 5 minutes) and install software while enabled. Users will need to enter their personal AIMS log in credentials in order to install.

See the AIMS 3D Modelling OneNote – AIMS System Configuration section (internal link: AIMS 3D Modelling OneNote) for steps on how to enable the Elevate program.

3.2.1.1 Anaconda

Table 4. Steps to download, install, and set up Anaconda Environment.

Step	Key task	Notes
1	Download and install Anaconda Individual Edition	• Download Now | Anaconda (Figure 4a) • Download Distribution Installers • Leave Advanced Installations Options as default - "Create shortcuts" and "Register Anaconda3 as my default Python" should be ticked, all other options unticked
2	Open Anaconda Navigator from start menu	• Figure 4b
3	Create a new environment	• Click 'Environments' tab (left-hand side, Figure 4c • Click 'Create' (bottom of window) • Select python 3.11 (n.b. check TagLab github Wiki to see if the required python version changes with future updates, and select python version accordingly) - Tip: name the environment specific to yourself if using a shared device, e.g. taglab_env_your initials

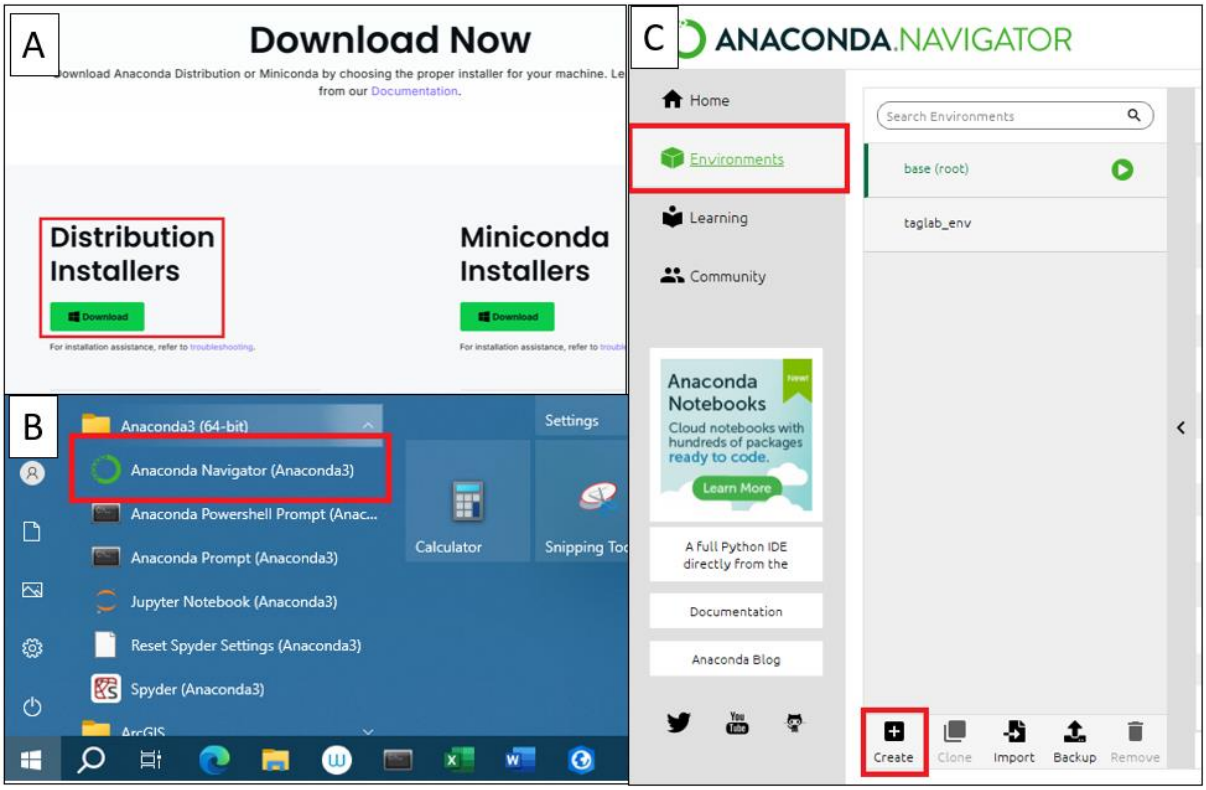

Figure 4. A) Anaconda website with download button outlined in red, B) location of Anaconda Navigator in start menu, and C) location of 'Environments' tab and 'Create' button within Anaconda Navigator used to create a new environment.

3.2.1.2 Microsoft Visual Studio

Table 5. Steps to install Microsoft dependencies for TagLab.

Step	Key task	Notes
1	Download Microsoft Visual Studio and C++ Build Tools	• Microsoft C++ Build Tools - Visual Studio • Click on "Download Build Tools" at the top of the page • N.b. the version at time of writing is Visual Studio 2022

2	Install Build Tools	• If using AIMS computer, enable Elevate (see section 3.2.1).
		• Open the folder where it was downloaded and double click the .exe file
		  - Default folder is the Downloads folder
		• Follow default installation prompts until installation window with package download options appears (Figure 5)
		• Tick box for "Desktop Development with C++"
		  - Leave all other options unticked
		• Click Install (bottom right of window)
		  - If initially installed without Desktop Development with C++ package included, open VS Studio Installer from Start menu and click "Modify". This will open installation options and the package can be installed

- Note that the Desktop Development with C++ package installs the Microsoft Visual C++ Redistributable mentioned in the TagLab website installation wiki (Install TagLab · cnr-isti-vclab/TagLab Wiki · GitHub)

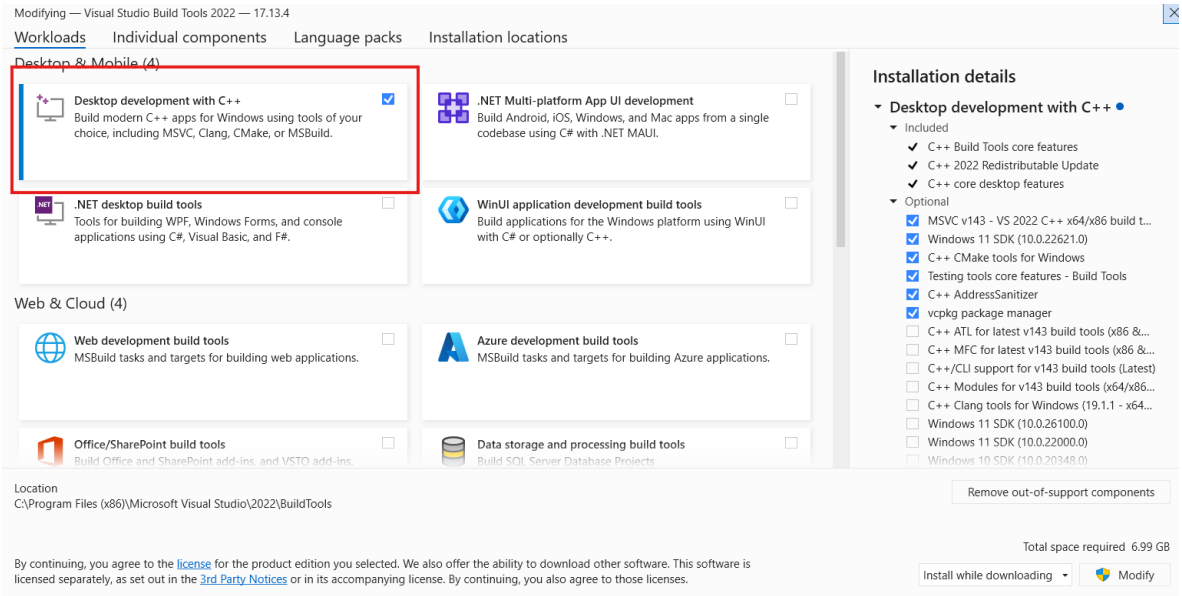

Figure 5. Microsoft Visual Studio Build Tools Install window. Select 'Desktop development with C++' to install.

3.2.1.3 Nvidia CUDA Toolkit

Download Microsoft Visual Studio (Table **5**) prior to installing the Nvidia CUDA Toolkit, otherwise the installation prompt will display the following warning: "No supported version of Visual Studio was found. Some components of the CUDA Toolkit will not work properly. Please install Visual Studio first to get the full functionality."

Table 6. Steps to download and install the Nvidia CUDA Toolkit. Text in **bold** is exact code to be written in the command line of Anaconda Prompt. This syntax can be directly copied except where specified.

Step	Key task	Notes
1	Download Nvidia CUDA Toolkit	• CUDA Toolkit 12.8 Update 1 Downloads | NVIDIA Developer (Figure 6a) • Note: some compatibility issues may arise with older versions of the CUDA Toolkit and Microsoft Visual Studio - It's recommended to install CUDA Toolkit 12.8 for use with Visual Studio 2022 (any prior versions of Visual Studio require a subscription)
2	Install CUDA Toolkit	• If using AIMS computer, enable Elevate (see section 3.2.1) • Open the folder where CUDA Toolkit was downloaded and double click the exe file (e.g. cuda_12.8.0_511.23_windows) - Default folder is the Downloads folder • Follow installation prompts

		• Select “Express” under Installation Options, all other options remain as default
3	Confirm successful installation and correct version	• Open Anaconda Prompt from start menu (Figure 8a) • Type **nvcc --version** (Figure 6b first red box) and hit Enter, the Anaconda Prompt will display the version of CUDA installed - It should say “release 12.8” (Figure 6b, second red box)

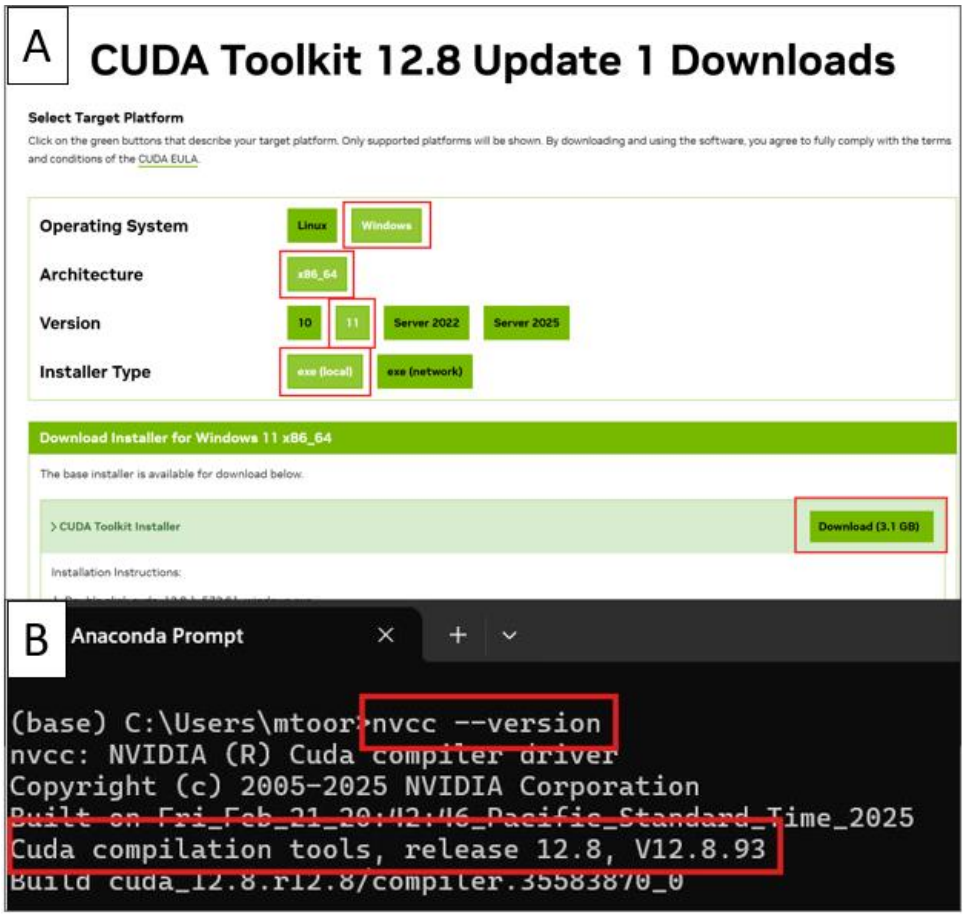

Figure 6. A) Nvidia Developer website with settings and download link outlined in red, and B) Anaconda Prompt with command to check the successful installation of CUDA Toolkit (top) and version installed (bottom) in red.

3.2.1.4 TagLab

Program Download

The download and installation steps have been divided into Table **7** and Table **8**. Table **7** contains the steps to download the TagLab code while Table **8** contains the steps to install the program.

Table 7. Steps to download TagLab.

Step	Key task	Notes
1	Open TagLab GitHub repository	• https://github.com/cnr-isti-vclab/TagLab
2	Download TagLab code	• Click on green Code button (top right) • Click 'Download ZIP' (Figure 7)
3	Extract 'TagLab-main' folder	• Extract ZIP folder and save on desktop (or easily accessible location)

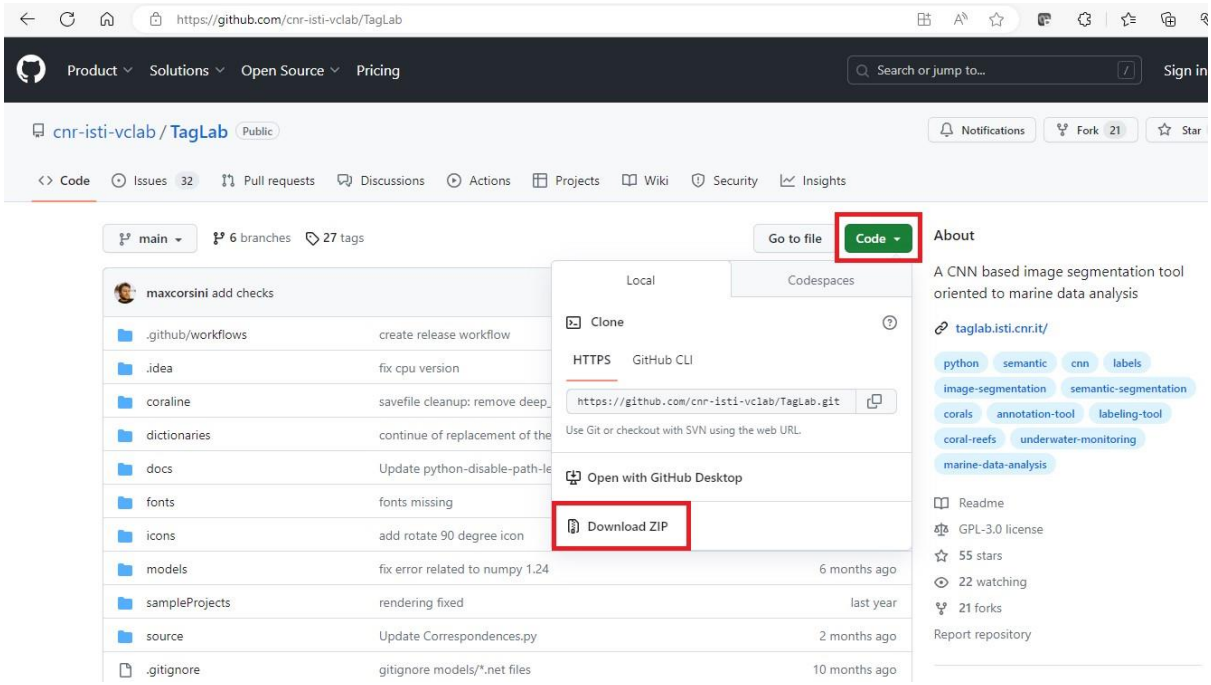

Figure 7. TagLab GitHub repository with green code button and downloadable ZIP file outlined in red.

Program Installation

The steps in Table **8** are the general steps to install TagLab, however, as new versions are released, changes to this process may occur. Alternatively, versioning changes in python packages may interfere with the installation or functionality of tools within TagLab. If issues arise during the installation process, see appendix 5.1.1 in this document, the AIMS 3D Modelling OneNote (internal link: AIMS 3D Modelling OneNote), or review the Issues page of the TagLab GitHub repository to see if other users have experienced similar issues.

Table 8. Steps to install TagLab. Text in bold is exact code to be written in the command line of Anaconda Prompt. This syntax can be directly copied except where specified.

Step	Key task	Notes
1	Open Anaconda Prompt	• Figure 8a • If the wrong version of python was selected in the Anaconda environment setup (Table **4**, step 3), a different version can be installed now

		- Type **conda install python==x.x** (replace x.x with desired version, e.g. 3.11) • Type 'y' when prompted
2	Change the directory to the TagLab-main folder	• Type **cd** with a space after it in Anaconda Prompt • Drag and drop the TagLab-main folder into prompt (Figure 8b)
3	Activate the TagLab environment	• Type **activate [name of environment]** (replace with your environment name, Figure 8c) • Use the environment created during Anaconda installation (Table **4**)
4	Install TagLab	• Type **python install.py** (Figure 8c) • If Nvidia CUDA Toolkit is not installed, install TagLab using the command **python install.py cpu** - Note that automatic classifiers cannot be used with this cpu version • If issues arise, see Appendix section 5.1.1

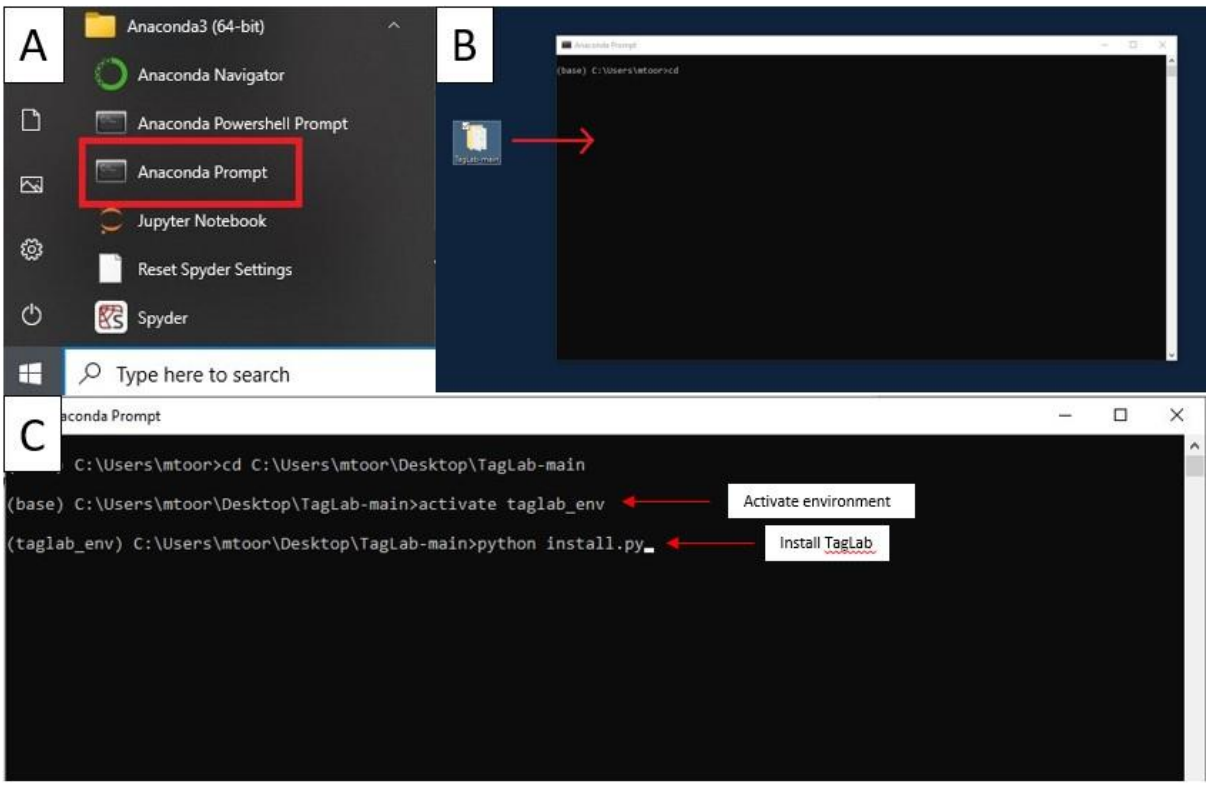

Figure 8. A) Location of Anaconda Prompt in start menu, B) set the directory by typing "cd " and dragging TagLab-main folder from location into Anaconda Prompt window, and C) command lines to activate the desired Anaconda environment and install TagLab.

3.3 Project setup and how to use TagLab

3.3.1 Launch TagLab

Follow the steps in Table **9** to launch the TagLab interface (Figure 9). This process is required to open TagLab each time, the program cannot be opened by clicking on a pre-existing project.

Table 9. Steps to launch and update TagLab. Text in **bold** is exact code to be written in the command line of Anaconda Prompt. This syntax can be directly copied except where specified.

Step	Key task	Notes
1	Open Anaconda Prompt	• Same as Table **8**, step 1 • Figure 8a
2	Change the directory to the TagLab-main folder	• Same as Table **8**, step 2 • Type **cd** in Anaconda Prompt ('cd' + space) • Drag and drop the TagLab-main folder into prompt (Figure 8b)
3	Activate the TagLab environment	• Same as Table **8**, step 3 • Type **activate [name of environment]** (replace with your environment name, Figure 8c) • Use the environment created during Anaconda installation (Table **4**)
4	Launch TagLab	• Type **python taglab.py** • Interface will launch (might take a minute) • Tip: after having done steps 2-4 once, copy each line that you entered and save these lines somewhere accessible (such as a .txt file) so they can be copied and pasted to launch TagLab in the future
5	Update TagLab	• This is only required when the TagLab developers update the software • A message will appear instructing to update when attempting to launch TagLab • Type **python update.py**

3.3.2 Interface and tools

When TagLab opens, it will show a blank interface with tools along the left-hand side of the window and various panels along the right-hand side (Figure 9). See the TagLab website for details on all tools, the ones relevant to this SOP are discussed in the following sections. Note that to use many of the tools listed, an orthomosaic will need to be loaded first (see section 3.3.3 for how to set up a project with an orthomosaic).

Table 10. TagLab navigation tools.

Step	Tool	Notes
1	Zoom	• Zoom in and out of an orthomosaic with the mouse wheel • Screen will zoom in on location of mouse
2	Pan/Move	• Left click and drag orthomosaic to move location • If using a different tool, hold control and left click to move orthomosaic without switching tools

3	Select polygon	• Double click inside a polygon • To select multiple polygons, select one then hold shift and double click additional polygon(s) • Note: polygons cannot be selected while using certain tools (e.g. 4-click segmentation)
4	Deselect polygon	• Double click on background orthomosaic or press the escape key • Note: if using split screen, pressing the escape key will only deselect polygon on the orthomosaic that is activated
5	Polygon visibility	• Check/uncheck boxes at top of TagLab window to view or hide segment fill, border, ID labels, or an overlaid grid • Use the sliding bar to adjust transparency of polygons

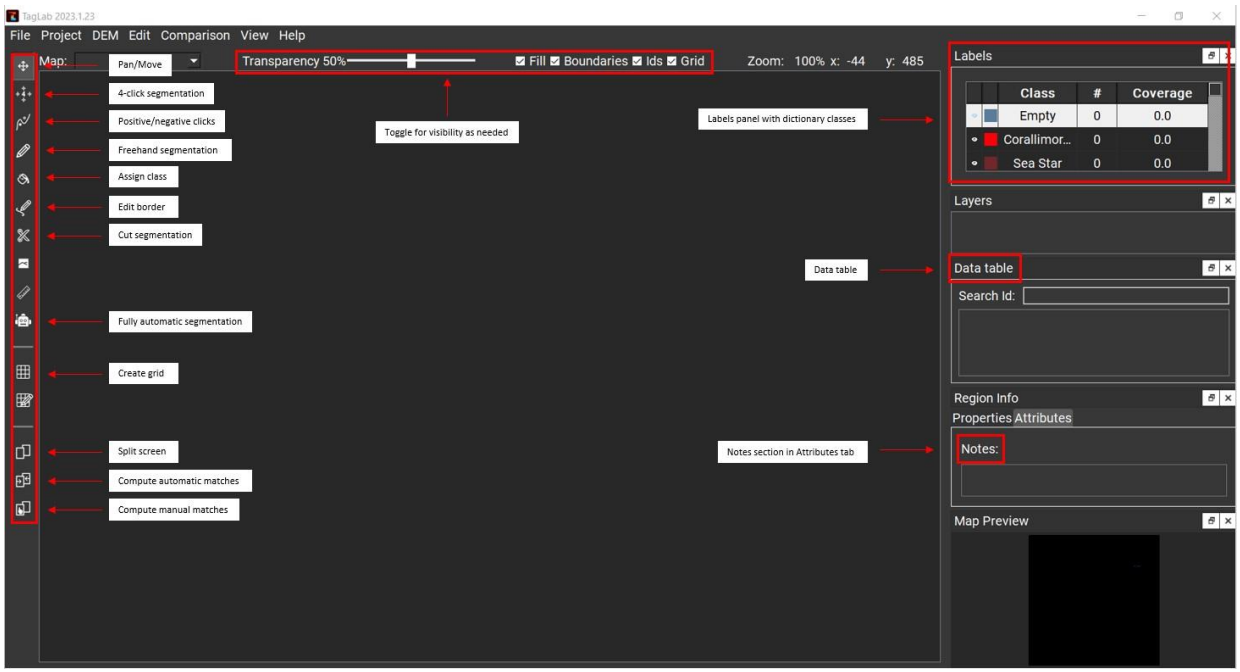

Figure 9. Blank TagLab interface with tools and relevant panels identified and outline in red.

3.3.3 Set up a new project

Once TagLab is open, an orthomosaic or image file can be loaded. Within TagLab, an orthomosaic or image file is referred to as a 'map', however, this SOP will use the term orthomosaic (they are considered interchangeable for this purpose). Multiple orthomosaics can be loaded into one project and will appear in the 'Layers' panel on the right-hand side of the interface (Figure 10, outlined in red). In this SOP, a project is defined as a TagLab file that contains one or more orthomosaics.

The time required to load an orthomosaic can vary depending on computing power, the size of the orthomosaic file, where the orthomosaic file is stored (e.g. locally or on a network), and the internet connection (if stored in a network location). Loading time can range anywhere from 30 seconds to one hour. The orthomosaic will appear in the center window of the interface once it loads (Figure 10).

See section 3.3.5 for EcoRRAP-specific project information.

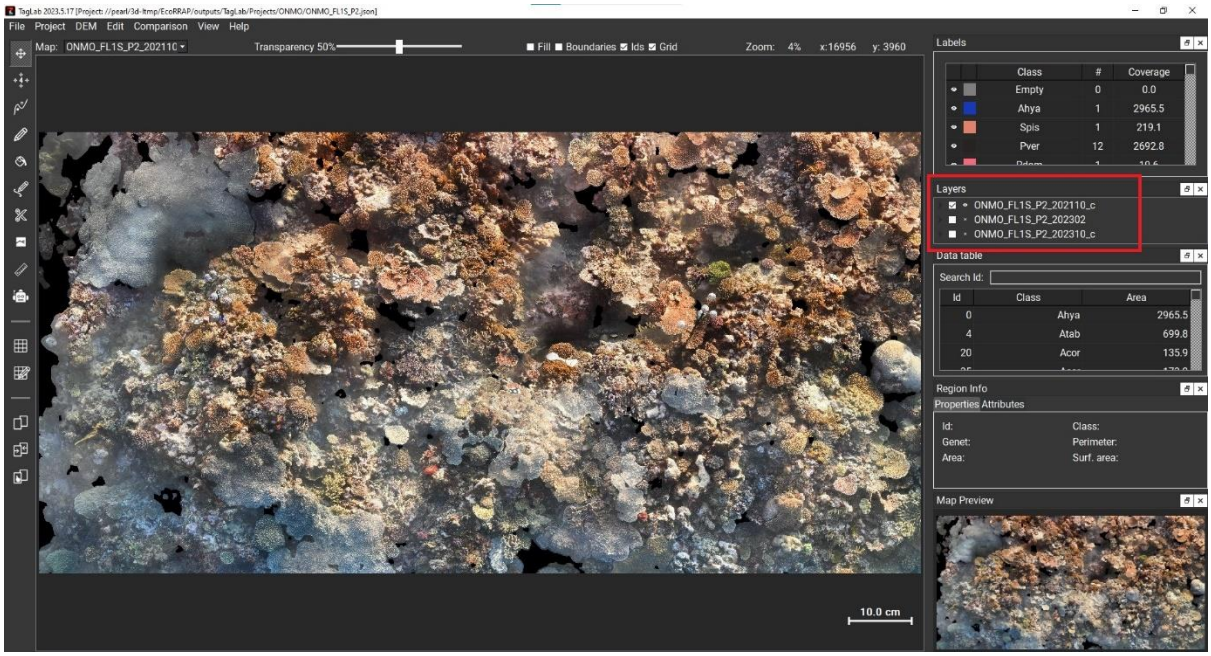

Figure 10. TagLab interface with loaded orthomosaic. The 'Layers' panel, outlined in red, shows all orthomosaics within the project.

Table **11** provides the steps to create a new project, including loading and editing an orthomosaic, loading or creating a dictionary for polygon annotation, and saving the project.

Table 11. Steps to set up a new project in TagLab. See section 3.3.5 for EcoRRAP-specific project information.

Step	Key task	Notes
1	Launch TagLab	• See Table **9** for steps, if required
2	Add a map (orthomosaic)	• Project > Add a new map • Orthomosaic must not exceed TagLab's max size of 32767x32767 pixels
3	Fill in map settings	• Figure 11a • Map Name: EcoRRAP standard is to match the filename of the orthomosaic (.tif) file • RGB Image: provide the file path to the orthomosaic, click '…' and navigate to file location • Depth Image: leave blank if not using DEM. If using DEM, import here to calculate colony surface area (2.5D) but note that EcoRRAP does not use this and no other information on this is provided in this SOP • Acquisition Date: fill in the date when the site was imaged in the format YYYY-MM-DD. This is required to compute multi-temporal comparisons. - For EcoRRAP purposes enter the year and the first day of the month of imaging (e.g. 2021-05-01) • Pixel size: comes from the orthomosaic (.tif) after it's exported from Metashape (or other software).

		- Check orthomosaic properties in ArcGIS Pro or similar (see appendix section 5.1.2 for steps). • Click Apply to load orthomosaic (may take some time) • This step only needs to be completed once per orthomosaic
4	Edit or delete an orthomosaic	• Project > Maps Editor • Click 'Edit' in the top right-hand corner of Maps Editor window to change the orthomosaic name, file pathway, acquisition date, or pixel size • Click 'Delete' to remove the orthomosaic from the project
5	Load a dictionary	• 'Dictionary' refers to the list of classes that polygons can be assigned to within a project ('Labels' panel, top right, Figure 9) • TagLab comes loaded with the default Scripps dictionary • To load a different dictionary (.json file): - Make sure the new dictionary is located in the 'dictionaries' folder of the TagLab-main folder - Project > Labels Dictionary Editor (Figure 11b) - Click 'Load' and navigate to dictionary file - Select 'Replace existing dictionary' when prompted - Class list will change in the Labels window on the top right-hand side of the screen - Review the colours of classes to ensure that no duplicated RGB values exist within the dictionary (see appendix 5.1.3 for more information on this) - See Table **13** for information about the EcoRRAP dictionary • To create a new dictionary: - Project > Labels Dictionary Editor - Click 'New' - Add a label name for a desired class along with a *unique* RGB value (ensure each RGB value is only used once per dictionary) - Click add - Continue this process for as many classes as desired - Click 'Save' to save the dictionary file externally (.json), which can then be added to other projects
6	Add additional orthomosaics	• Follow steps 2-3 to add additional orthomosaics to a single project, for example with multiple timepoints of the same site • All orthomosaics within a project will appear in the 'Layers' panel on the right-hand side of the TagLab window (Figure 10)
7	Save project	• File > Save Project • Save to desired location • Save frequently to avoid data loss

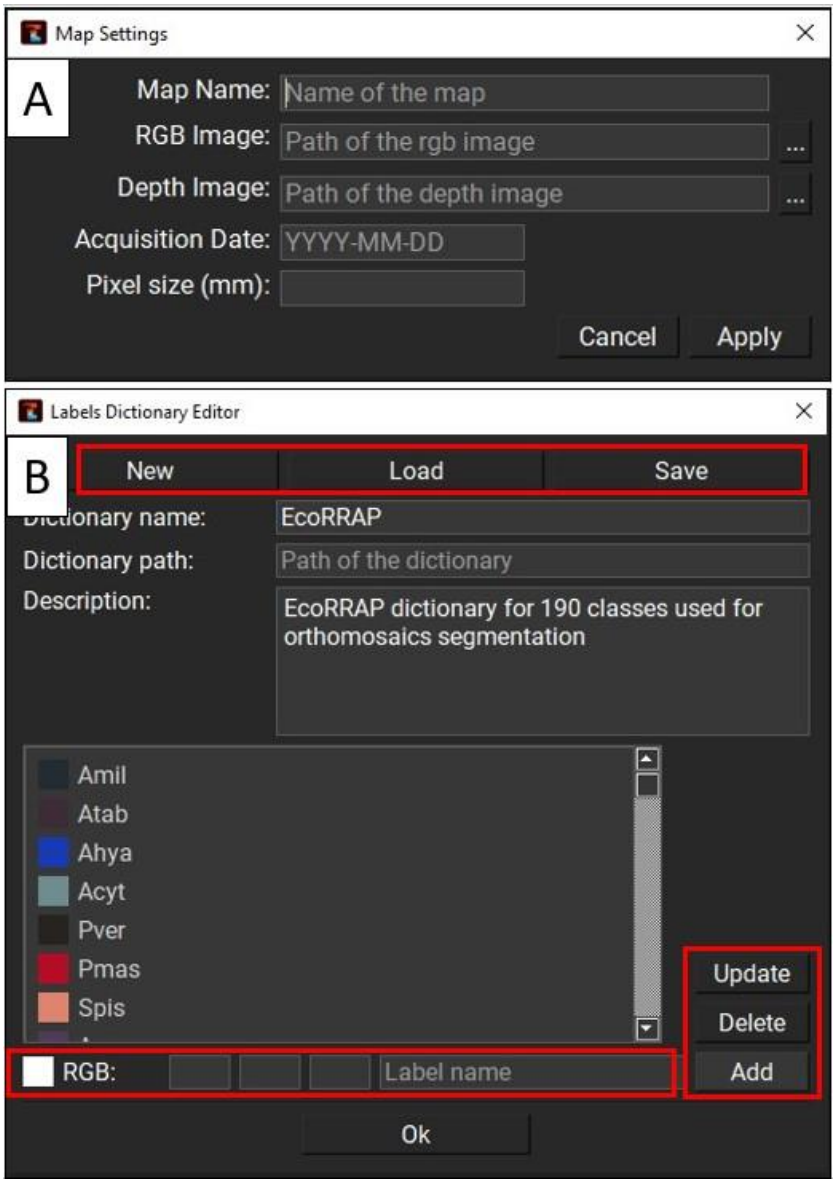

Figure 11. A) Map settings window with details to fill when loading an orthomosaic, and B) Labels dictionary editor window to create a new dictionary, load a pre-existing dictionary, or edit the currently loaded dictionary.

3.3.4 Edit an existing project

Existing projects must be opened from the TagLab interface, they cannot be opened by clicking on the project file (.json). Follow steps in Table **12** to open an existing project and load multiple orthomosaics within a single project, if applicable. See section 3.3.5 for EcoRRAP specific project information.

Table 12. Steps to open an existing project in TagLab.

Step	Key task	Notes
1	Launch TagLab	• See Table **9** for steps, if required
2	Add dictionary to TagLab-main folder (first time only)	• If the project to open has a different dictionary from the pre-loaded Scripps dictionary, then the correct dictionary will need

		to be added into the 'dictionaries' folder within the TagLab-main folder
3	Open project	• File > Open Project • Navigate to project location and select desired file • Click 'Open' (may take time to load)
4	Load multiple orthomosaics within a project	• Once open, the project will show the orthomosaic from the first timepoint (based on 'Acquisition Date') • Each additional orthomosaic will be listed in the Layers window. To toggle between orthomosaics: - Click the box to the left of the orthomosaic name - Alternatively, select the desired orthomosaic from the 'Map:' drop down list in the upper left corner of the TagLab screen - Note that each time a project is opened and each time an orthomosaic is accessed within a project, it will take time to load
5	Remap file pathway	• If the filename or file location of a previously loaded orthomosaic changes, a pop-up window will appear when the project is opened asking to remap the file pathway (file name appears at the top of the pop up window) • Navigate to the correct orthomosaic file and click 'Okay'

3.3.5 EcoRRAP-specific project information

Table **13** contains information about the EcoRRAP TagLab projects, including information on file locations, required inputs, and information on file naming. Note that projects already exist for all EcoRRAP plots.

Table 13. EcoRRAP-specific information for TagLab projects, including file locations and notes.

Option	File type	EcoRRAP-specific information
1	Orthomosaics	• File location (internal link): \\pearl\3d-ltmp\EcoRRAP\outputs\orthomosaics - Navigate to desired year folder - 'c' in folder names refers to 'clipped'. These orthomosaics have been clipped to the same extent as the orthomosaic from the reference year (usually 2022) for alignment purposes • Map name in TagLab (same as orthomosaic name): - 'ClusterReef_SiteZone_Plot_YearMonth_c' • E.g. CBHE_BA1D_P1_202105_c
2	EcoRRAP dictionary	• File location (internal link): \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Dictionaries - Copy dictionary to 'dictionaries' subfolder within TagLab-main folder • Follow step 5 in Table **11** to load dictionary • Note: all existing EcoRRAP projects already have this dictionary as the default dictionary but it still needs to be transferred to each individual's TagLab-main folder • This dictionary contains 190 classes • See section 5.2.1 for more information on EcoRRAP taxa

3	TagLab projects	• Existing EcoRRAP projects are saved within reef subfolders. Any new EcoRRAP projects should be saved within the corresponding reef folder here (internal link): • \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Projects

3.4 Digitisation of coral colonies in an orthomosaic at time zero

There are multiple AI-assisted tools that can be used to segment and annotate objects of interest in TagLab, including the 4-click tool, positive/negative click segmentation, and freehand segmentation. Most tools can be found in the left-hand column of the TagLab window (Figure 9) and those relevant to this workflow are explained in Table **14** and Table **15**. Also see the tools section of the TagLab website for further details (new tools are sometimes added by the developers as part of TagLab updates so checking the website regularly is recommended).

3.4.1 Digitising new colonies at time zero

It is recommended to establish a standardised (repeatable) search procedure of an orthomosaic to locate all colonies of interest. The steps in Table **14** outline one method of searching by using a grid overlay to divide the orthomosaic into smaller sections. Each section can be thoroughly searched and, once all colonies have been digitised, can be marked as complete to progressively work through the entire orthomosaic.

Table 14. Steps to digitise new colonies in an orthomosaic in TagLab using AI-assisted tools.

Step	Key task	Notes
1	Overlay a grid on the orthomosaic	• Click 'create grid' tool (left-hand side) • Edit number of rows and columns as desired, 5x5 recommend
2	Search each grid cell to locate colonies of interest	• Use the 'move' tool to navigate around each grid cell to search for colonies (see Table **10**, step 2) • Once a grid cell has been searched, click the 'Active/deactive grid operations' button in the left-hand tools panel - Right click on completed cell and select 'Mark cell as complete' (or other relevant option)
3	Segment substrate (4-click segmentation)	• Select 4-click segmentation tool • Hold shift and left click the four extremes of the colony (red 'x' will appear where clicked, Figure 12a) • Tool will automatically create a polygon when the fourth extreme is clicked (Figure 12b) • The polygon class will appear as 'Empty' and will need to be annotated with the correct class (step 5)
4	Edit polygon border	• Use the 'edit border' tool to manually adjust border if the AI-generated polygon doesn't align with colony edges well enough • Use the 'Refine' function (R button on keyboard) to adjust colony border by a few pixels • See Table **15** for more details about editing polygons
5	Annotate polygon	• Double click on a polygon (border will turn white)

		• Double click on the desired class in the Labels window (top right), polygon will fill with colour of the selected class (Figure 12c) • Alternatively, use the 'assign class' tool: - With polygon selected, click on the tool and then click once on the desired class in the Labels window
6	Add any relevant notes	• Click 'Attributes' in the Region Info window (right-hand side) • Type any desired notes - **Ensure that the borders of the polygon are finalised before adding notes. If a note is made and then an edit made (specifically if the 'Refine' tool is used) the note will be deleted.** • See Table **17** for the list of notes used in the EcoRRAP dataset

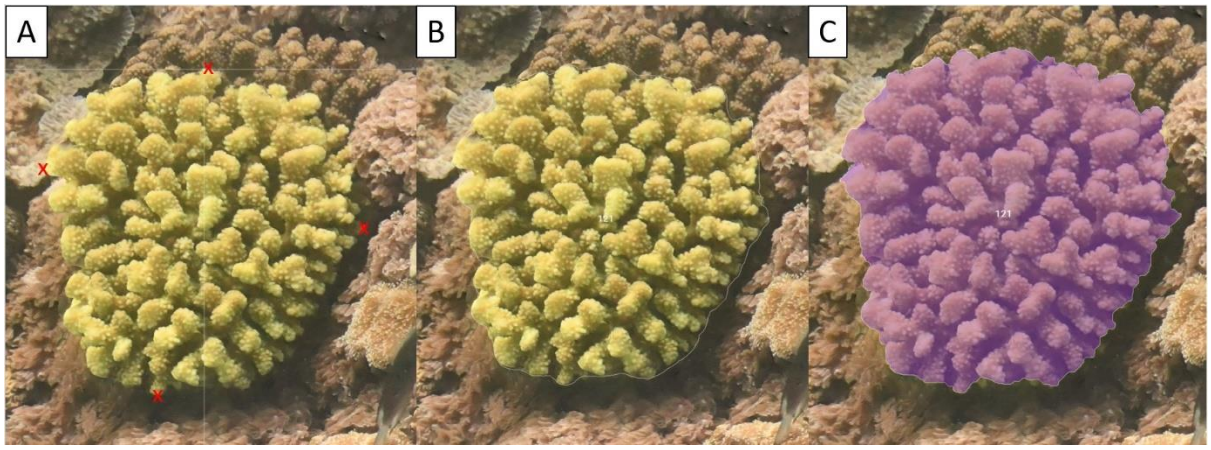

Figure 12. Example of 4-click segmentation, including A) marking the four extremes of the colony, B) the automatic polygon generation by the tool, and C) the annotated polygon after a class is assigned.

3.4.2 Editing existing colonies

It may be necessary to edit an automatically generated polygon (e.g. from the 4-click segmentation tool) if it doesn't align well enough with the colony border. There are multiple tools that can be used to edit polygons depending on the type of edit required. Tools specific to this workflow are presented in Table **15**. These editing options are especially relevant for transferring colonies in a timeseries (see section 3.6 for more information on timeseries data).

It's recommended to use a touch screen display and stylus pen (Table **2**) to improve accuracy of segmentation as well as to minimise impact on wrists. If a screen and stylus are not available, an ergonomic mouse can also be used to minimise impact from repetitive movements.

Table 15. Processes that can be used to edit existing polygons in TagLab.

Option	Key task	Notes
1	Edit border	• Add or remove sections of the polygon with the 'edit border' tool: - Select tool - Select polygon (border turns white)

		- Draw a line along the edge of the colony, make sure to intersect the existing polygon border in two places - Press the space bar to incorporate edit • Use the 'Refine' tool ('R' key) to better align the border automatically - Note: effectiveness of the 'Refine' tool is colony dependent
2	Split polygons	• Select the Cut tool to divide one polygon into two • Draw a line across a polygon where it should be split, making sure to intersect the existing border in two places • Press the space bar to split polygon – generates a new polygon with the same class
3	Merge polygons (Figure 13A-D)	• Polygons to be merged must overlap (use edit border to overlap if needed, Figure 13B) • Select one polygon • Press shift and select second polygon (Figure 13C) • Press M button on keyboard or right click and select "Merge Overlapped Labels" (Figure 13C)
4	Create adjacent, non-overlapping polygons (Figure 13 E-I)	• Use when polygons should share the same border (e.g. adjacent coral colonies) • Segment the substrate whose edge will be used as the border (polygon 1) • Segment the second substrate (polygon 2) whose edge will align to polygon 1 • Use the edit border tool to overlap polygon 2 with the shared edge of polygon 1 (Figure 13F) • Select polygon 1 then hold shift and select polygon 2 (note that the order of selection matters, Figure 13G-H) • Press D on the keyboard or right-click and select "Divide Labels" (Figure 13H) • If more than two colonies share borders, continue this process step-wise with each pair of polygons • This can also be used when a polygon falls fully inside another polygon (e.g. one colony growing on another) • Tip: refine non-shared edges of both polygons prior to the divide step to avoid unintentional edits to the shared edge that would require repeating the divide step
5	Positive/negative clicks (Figure 13 J-L)	• Align TagLab window so that the entirety of the polygon to edit is visible and fills majority of screen • Select polygon • Select tool • Press shift and left-click (checkerboard pattern will appear, Figure 13J) • Press shift and left-click on areas of the substrate to be added into the polygon (green dot appears, Figure 13K) • Press shift and right-click to remove areas of the substrate from the polygon (red dot appears, Figure 13K) • Continue adding/subtracting until polygon border aligns with substrate as desired • Press space bar to accept changes (Figure 13L)

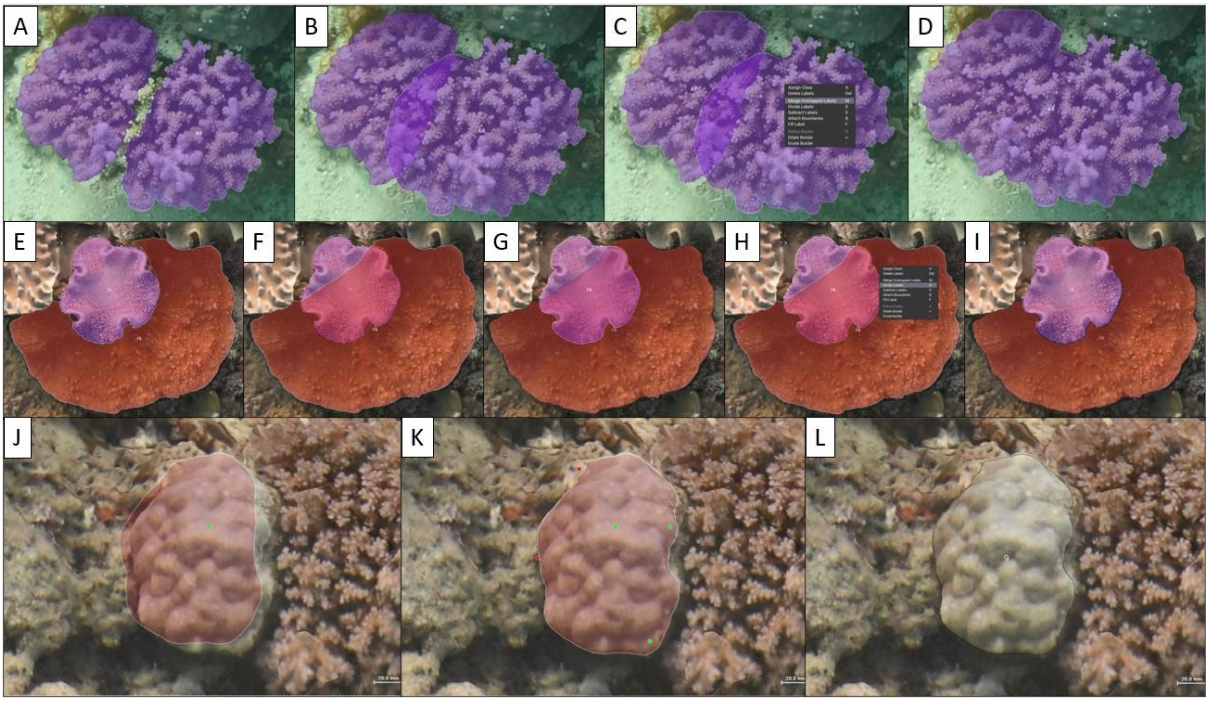

Figure 13. Steps to edit polygons, including merging polygons (A-D), creating adjacent non-overlapping polygons (E-I), and adjusting a polygon with the positive/negative clicks tool (J-L).

3.4.3 EcoRRAP digitisation criteria and colony notes

The following sections contain information about the EcoRRAP digitisation process, including criteria used to determine whether coral colonies should be segmented and what notes were used to track colony information.

3.4.3.1 EcoRRAP digitisation criteria

To standardise the EcoRRAP digitisation process as much as possible, a set of criteria were established to determine whether a coral colony could be digitised within the EcoRRAP orthomosaics. Criteria in Table **16** were followed to digitise colonies in the first timepoint as well as any new colonies added in later timepoints. If a digitised colony did not meet these criteria in any subsequent timepoints, it was given a note to indicate which criteria wasn't met and whether it could be used for certain analyses (e.g. growth or survival). If a colony was deemed unusable for analyses, it was identified as such during the dataset quality control process based on the note attributed (see section 3.6.2.3 for more information about the QAQC process). See Table **17** for further information on notes used in the EcoRRAP dataset.

Table 16. Criteria for digitising coral colonies in EcoRRAP projects.

Category	Criteria
Colony identification	• Colony taxon could be identified with at least 80% certainty • Identification resources used include: - Indo Pacific Coral Finder, 2022 Edition (Kelley et al. 2022)

	- Corals of the World (Veron et al. 2016) - Discussion and confirmation between benthic ecologists
Colony reconstruction	• Colony was reconstructed well in the orthomosaics so that it was not blurry or warped • The entirety of the colony was within orthomosaic bounds and did not fall within a hole in the orthomosaic
Colony size	• Any colony that was identifiable based on criteria above was digitised regardless of size
Colony visibility	• The majority of the colony (at least 75%) was clearly visible from above (e.g. not cryptic or hidden due to its orientation) • Colonies that were partially obscured by another substrate were given a note to indicate such (see Table **17** for more information on notes)
Colony status	• The colony appeared healthy with minimal or no partial mortality in the first time point - If they presented partial mortality in subsequent years, they were still digitised in T_0 • Only areas of live tissue were digitised (see 'Partial Mortality' in Table **17** for more detail on areas of dead tissue) • Colonies that displayed partial mortality were given a note to indicate such • Separate, remnant sections of a colony that appeared to previously belong to one larger colony were digitised as separate polygons
Staghorn Acropora	• A colony was clearly separated from others around it so that it had distinct edges (Figure 14a) • Segment edges aligned with external branches and internal sections of empty space between branches were not removed from polygon (Figure 14b) • Large thickets that could not be identified as individual colonies were given a note to identify it as a thicket and should not be used for size metrics (see Table **17** for note information) - If separate colonies merged into a thicket in future timepoints the colony was noted as such and no longer used for size metrics

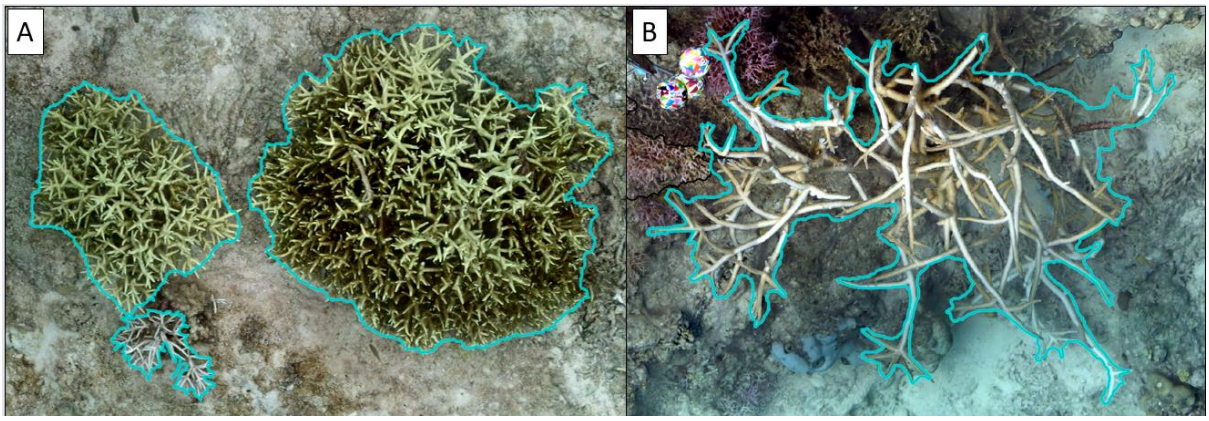

Figure 14. Examples of digitised *Acropora* staghorn colonies, including A) colonies with distinguishable edges indicating an individual colony (polygon edges in blue), and B) polygons roughly aligned with external branches, internal space between branches included within polygon (polygon edges in blue).

3.4.3.2 EcoRRAP colony notes

Notes are used during the digitisation process, including transferring colonies in a timeseries, to track aspects of colony status (e.g. whether it is distorted in the orthomosaic), as well as ecological processes (e.g. partial mortality or colonies merging). Notes used in the EcoRRAP dataset are listed in Table **17** along with an explanation and any additional instructions for digitisation. These notes are used in the QAQC process (section 3.6.2.3) to determine whether a colony can be used for certain analyses, such as growth or survival analyses, thus are critical to include. Note that EcoRRAP users should be using the exact syntax stated in the table. If two or more notes are to be added, separate each note with a comma. Visual examples of notes can be seen in Figure 15.

Table 17. EcoRRAP colony notes attributed during the digitisation process. See Figure 15 for visual examples.

Option	Key task	Notes

1	Partial mortality ('pm')	• If a colony shows any observable partial mortality (up to 99%)
		• If morality occurs along the border of a colony:
		  - Edit the polygon so that only areas of live tissue are included (polygon edge should be on the border between live tissue and dead skeleton)
		• If partial mortality occurs centrally on the colony:
		  - Leave it included in the polygon (i.e. do not cut out)
		• Type "pm" in the notes section (under Attributes) for any partial mortality (central or along the border of colony)
		• Type "pm base" for staghorn Acropora where the base section has died but branch tips remain alive
		• See option 10 if partial mortality has split a colony into separate remnant sections
		• See Figure 15 H and I

2	Obscured colonies ('o')	• A colony is considered obscured if it is overtopped or partially blocked by another substrate or piece of equipment (e.g. sphere tree, dumbbell) but is still visibly alive • Type an "o" in the notes section • See Figure 15E
3	Unobscured colonies ('unobscured')	• A colony is considered 'unobscured' if it was previously overtopped or obscured by another substrate, which is no longer there so that the entire colony can now be seen. • Type "unobscured" in the notes section
4	Bleached colonies ('bleached')	• Type "bleached" in the notes section if a colony is displaying any level of bleaching, including partial or full bleaching • Confirm with EcoRRAP team member if unsure
5	Dead colonies ('d')	• A colony is considered dead when there is 100% mortality and the structure is still visible • Type a "d" in the notes section • See Figure 15A
6	Missing colonies ('m')	• A colony is considered missing when it cannot be visually located within the orthomosaic but the location where the colony previously was can be identified (i.e. not obscured by other substrate). • Type an "m" in the notes section • Note: if an orthomosaic has shifted even slightly between time points, colonies may not align perfectly. Search within the nearby area for the colony, using the previous time point as a reference • See Figure 15B

7	Detached (‘detached’)	• A colony is considered detached when it can be visually located within the orthomosaic and is visibly alive but is no longer attached to the reef in its previous location - Observer must be 100% sure that it is the same colony • If the colony is both detached and dead, both notes can be entered • Type “detached” in the notes section • See Figure 15F

8	Bad colonies (‘b’)	• A colony is considered bad if it meets any of the following criteria:
		  - Falls partially outside of the orthomosaic bounds and the visible area is still alive
		  - Falls partially within a hole in the orthomosaic and the visible area is still alive
		  - The colony appears blurry or distorted
		  - The quality of the orthomosaic is too poor to determine the colony border but it is still visibly alive
		• Type a “b” in the notes section
		  - Note that “Bad” has also been used previously and is the equivalent of “b”
		• See Figure 15 C and D

9	Unknown colonies (‘unknown’)	• A colony is considered unknown if it:
		  - Falls fully outside of the orthomosaic bounds
		  - Falls fully within a hole in the orthomosaic
		  - Is fully obscured by another substrate
		  - Falls partially outside of orthomosaic bounds or partially within a hole in the orthomosaic and the visible area is dead
		  - Orthomosaic quality is too poor to determine if the colony is alive or dead
		• Type “unknown” in the notes section
		• See Figure 15G

10	Merged colonies (‘merged’)	• When two or more digitised colonies merge so that a distinct edge cannot be detected (e.g. staghorn Acropora) • Combine polygons using the merge tool or remove one polygon and edit the border of the other to encompass the entire merged area • When a digitised colony merges with another non-digitised colony, refine the polygon to include the newly merged section • Type “merged” in the notes section of the polygon
11	Remnant/split colonies (‘remnant’)	• If mortality splits a colony in two or more separate sections:

		- Use the cut tool to split the original polygon into multiple polygons and digitise the living sections of the colony separately - Alternatively, refine the border of the original polygon around the largest remaining section and generate new polygons around remnant sections of the colony (if a new polygon is generated be sure to annotate it with the same class) • Type “remnant” in the notes section for all polygons of the colony as well as “pm” to indicate partial mortality • NB: this step was only applied from the 2022-2023 transfer onwards - In the 2021-2022 transfer, only the largest living section of the colony was retained in the polygon, smaller remnant sections were not digitised
12	New colonies (‘new’)	• Any new colony added after T_0 (2021 for EcoRRAP) should be noted as such to differentiate it during the joining process • Type “new” in the notes section • Note: Due to improvements in EcoRRAP orthomosaic quality between the T_0 and T_1 (2022) timepoints, new colonies were added to the T_1 orthomosaics - No new colonies have been added to future timepoints after T_1 (current as of December 2024)
13	Genetics colonies (‘ng####’)	• See section 5.2.3 for details on genetics colonies within EcoRRAP plots • Any genetically sampled colonies in the EcoRRAP plots contain a note with “ng” for ‘new genetics’ and a unique identification number from the genomics database, e.g. ng1234 • Note: the “ng” note will appear in the previous timepoint from when it was collected in the field. For example, a colony nubbin collected in 2022 will have the note in the 2021 orthomosaic because the 2022 orthomosaic will not have been generated at the time of collection.
14	ID validated colonies (‘v’, ‘validated’)	• A colony whose species identification was validated *in-situ* by a coral taxonomist • See section 5.2.2 for more information on the ID validation process
15	Acropora staghorn thicket (‘thicket’)	• A staghorn Acropora cluster where individual colonies are indistinguishable • Type “thicket” in notes section • This note has only been used from 2024 onwards due to the addition of staghorn Acropora thickets for bleaching purposes. Any staghorn Acropora in previous years will have been digitised as an individual colony.

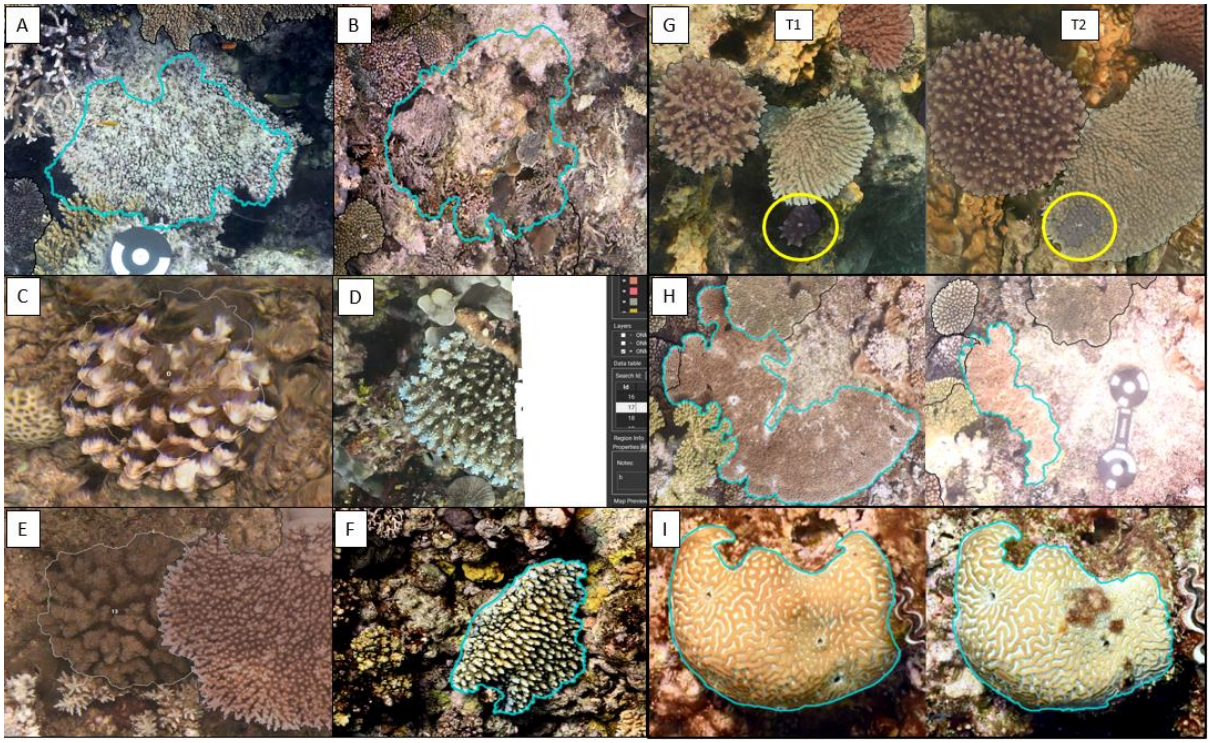

Figure 15. Examples of colonies that meet EcoRRAP note criteria, including A) dead colony, B) missing colony, C) 'bad' colony due to blurriness/distortion, D) 'bad' colony due to falling partially outside orthomosaic boundary, E) obscured colony, F) detached colony, G) unknown colony status due to complete overgrowth, circled in yellow H) partial mortality on edge of polygon, I) partial mortality internal to polygon. Panels G, H, and I show an example of what the colony looked like in the previous timepoint (T1) and in the timepoint where the note was attributed (T2). See Table **17** for full explanation of each note.

3.5 Exporting data from TagLab

When all colonies of interest in an orthomosaic have been digitised, the data can be exported from TagLab in various formats. This document will focus on three data formats: shapefiles (.shp), data tables (.csv), and image files (.png). See the TagLab website for other export options. See Table **19** for EcoRRAP specific file locations and file naming.

When data is exported from TagLab, regardless of the data format, data from each orthomosaic within a project must be exported separately and the data file will only contain information from that individual orthomosaic. If more than one orthomosaic exists within a project, data from each one must be exported separately and can be merged externally to TagLab to create a combined dataset. For a single time point, exported .csv files can be merged using the statistical software R (R Core Team, 2023). For timeseries data, see section 3.6.2 for how to link colony polygons across years.

Only classes that are visible (i.e. that have an eye icon next to their name in the 'Labels' panel) at the time of export will be included in the exported data, for all data formats.

Table 18. File formats and steps to export data products from TagLab.

Option	Key task	Notes
1	Shapefiles (.shp)	• Contain polygon shapes and attributes (area, notes, etc.) • File>Export>Export Annotations as Shapefile (Figure 16) • Tip: save shapefile with same name as orthomosaic for clarity
2	Data table (.csv)	• Contain polygon attributes (area, notes, etc.) but not polygon shapes • File>Export>Export Annotations as Data Table (Figure 16) • Tip: save file with the same name as the orthomosaic (as above)
3	Labelled Image (.png)	• Contain polygons as an image without attributes • Used here for transferring polygons across years (see section 3.6.1) • File>Export>Export Annotations as Labelled Image (Figure 16) • Tip: save file with the same name as the orthomosaic (as above)
4	Export a subset of classes	• Data from a subset of classes can be exported as any of the above data formats by hiding the visibility of any non-desired classes - Click the 'eye' icon to the left of the class name so that it appears with a strikethrough and polygons in that class are not visible on the orthomosaic • This can be done for any data format • It is not possible to export a subset of polygons within the same class

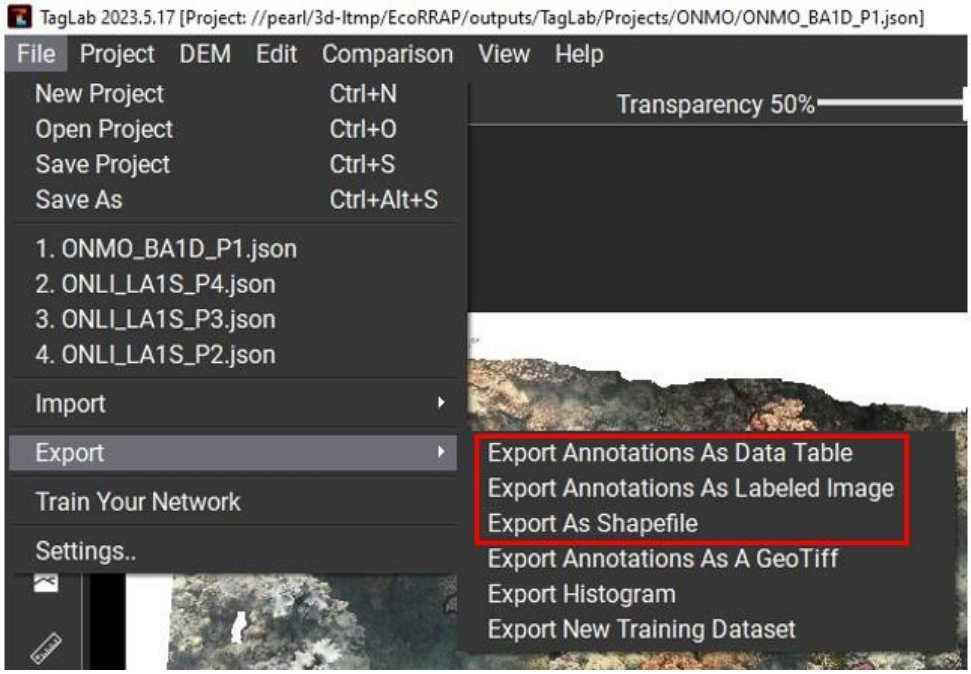

Figure 16. Location of export functions. Type of data products used in workflow are outlined in red.

3.5.1 EcoRRAP data export locations

Table **19** contains information on where files exported from TagLab should be saved within the EcoRRAP pearl network as well as file name specifications.

Table 19. File locations and file naming system for EcoRRAP TagLab exports.

Option	Key task	Notes
1	Shapefiles (.shp)	• Save raw shapefiles here: - \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Shapefiles\Demography\Raw • Save filenames as: 'ClusterReef_SiteZone_Plot_YearMonth' • E.g. CBHE_BA1D_P1_202105
2	Data table (.csv)	• Save raw datatables here: - \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Datatables\Prelim • Filename will automatically match that of the orthomosaic layer in TagLab
3	Labelled Image (.png)	• Save labelled images here: - \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Shapefiles\Labelled_images • Save filenames as: 'ClusterReef_SiteZone_Plot_YearMonth' • E.g. CBHE_BA1D_P1_202105

3.6 Timeseries datasets

Timeseries data provides information about a colony's survival and growth through time, as opposed to a single timepoint that only provides information on colony size and species density. If orthomosaics exist for multiple timepoints for a site, polygons from one timepoint can be overlaid on the orthomosaic from the following timepoint and edited to align with the colony borders in the new timepoint, hereafter referred to as polygon 'refining'. In TagLab, survival data is generated by assessing whether the colony remains alive in the new timepoint (recorded within the polygon notes), while growth data is calculated as change in 2D planar area of the colony between each timepoint.

3.6.1 Transferring colonies across timepoints

To transfer colonies across timepoints, polygons are exported from one time point as a labelled image (.png) and imported onto the orthomosaic from the next time point. This ensures that the same colonies are digitised and tracked through time. To successfully transfer colonies using this approach, steps from SOP 2 in this series must be completed beforehand to ensure that 3D models are co-registered and 2D orthomosaics are cropped appropriately so that orthomosaics align with high precision between timepoints. See SOP 2 in this series for more information about the co-registration process.

Table **20** contains steps for transferring and refining polygons across timepoints in TagLab.

Table 20. Steps to transfer polygons across timepoints.

Option	Key task	Notes
1	Export polygons as labelled image (.png) from previous time point	• See Table **18** for steps on how to export, if required • See Table **19** for location of EcoRRAP labelled images • Note: this step in the transfer process should be possible with shapefiles as well but TagLab may crash when attempting to import the shapefile.
2	Load orthomosaic for following time point	• See Table **11** for steps on how to add a new orthomosaic, if required
3	Import polygons as labelled image	• File>Import>Labelled Image • Navigate to desired labelled image in Input Label Map File window • Click Open • Polygons should appear on orthomosaic (may take a minute) and polygon data will appear in the Data Table • Note: the colours in the dictionary must match those in the labelled image. If a labelled image is exported and then a colour in the dictionary is changed, the polygons for that class will not be recognised and will not be imported • Note: if two classes have the same colour, all polygons will be imported as the class that is listed first in the dictionary - See appendix 5.1.3 for more information on this
4	Edit polygons	• See Table **15** for how to edit existing polygons • Use the data table to move through all polygons - Zoom in to desired level anywhere on the orthomosaic and click the first row in the data table. TagLab will automatically move the screen to the location of the polygon. Once that polygon is edited, click the next row in the data table to move to the next polygon – this avoids searching manually and reduces likelihood of missing polygons - Note that the order of polygons in the table can change if swapping between orthomosaics in a project or if a project is closed and reopened. Unedited polygons remain at the top of the table. Use the table headers to reorder table if desired • Note: the colony ID number is not important so, if required, a polygon can be deleted and remade as per Table **14**, but it must be annotated with the correct class
5	Add notes	• Add any desired notes to polygons, for example to record if a colony is dead, missing, etc. - Click ‘Attributes’ under Region Info to access ‘Notes’ section • If editing EcoRRAP projects, only add notes from existing categories. - See Table **17** for EcoRRAP note guide
6	Review polygons	• Use the data table to go back through and review each polygon to ensure none were missed and all edits are correct • Check that required notes have been added, retained, and are spelled correctly - If using EcoRRAP notes, ensure each note matches the exact syntax in Table **17**.

7	Add new colonies	• Additional colonies can be digitised following steps in Table **14** • This may be done to add new juvenile colonies or to include colonies that could not be digitised in previous time points due to orthomosaic quality • Add "new" in the notes section to distinguish it as a newly added colony in that timepoint • For EcoRRAP projects, consult Maren or Rena before adding any new colonies (this is to maintain a standard in taxa ID)
8	Export data	• See Table **18** for steps on how to export data • See Table **19** for where to save EcoRRAP files

3.6.2 Managing timeseries data

To generate a dataset that contains temporal data (survival, growth, etc.), polygons from each timepoint need to be matched. This can be done within TagLab or externally with software that manages spatial data, such as ArcGIS. Here, steps for both options are explained but note that the EcoRRAP workflow follows methods described in section 3.6.2.2 using ArcGIS Pro (ESRI, 2023). Also note that this is only one option applicable to this workflow and that methods and settings should be adjusted depending on desired data output.

EcoRRAP rationale for using external software to match polygons include:

- Volume of data
    - The matching function in TagLab needs to be run manually in each individual project. Therefore, if there are a large number of projects, it becomes time consuming to run this in each one.
    - Additionally, the matching function is not always 100% accurate and some polygons require manual matching. If there are a large number of polygons per project, the manual matching process also becomes time consuming.

- Projects must be finalized before conducting the matching
    - Matching within TagLab can only be conducted once all timepoints have been added, i.e. no more data is to be collected and transfers of all timepoints have been completed.

- Notes are not exported with matched data (as of version 2023.05.17)
    - Polygon notes are not included in the data table that is generated for matched polygons and are therefore not included in the final data export.
    - If notes contain crucial information about colonies, this information will be lost.

If any of these reasons are applicable, it is recommended to conduct polygon matching externally to TagLab. **Note that the polygon IDs created within TagLab are not consistent across orthomosaics within a project and thus cannot be used to link polygons via a data table.**

3.6.2.1 Linking colonies across timepoint in TagLab

TagLab has the functionality to automatically match polygons between multiple orthomosaics within the same project. This is done using the Compute Automatic Matches function and links polygons based on location within the orthomosaic. This function can only be used after all timepoints in a timeseries have been added and all polygons transferred and refined in each orthomosaic. It is not

possible to add another timepoint to a project after running the match function. **Note that this function does not automatically transfer polygons across years, it only links polygons that already exist in two or more timepoints and thus is only useful where both T1 and T2 have already been fully annotated (i.e. after colonies have been transferred across timepoints and refined)**.

The process documented in this section is based on TagLab v2023.5.17.

Table 21. Steps to match polygons in TagLab using the Compute Automatic Matches function.

Option	Key task	Notes
1	Segment and transfer colonies across timepoints	• See Table **14** for how to segment colonies • See Table **20** for how to transfer polygons across time points
2	Open project and load all orthomosaics within project	• See Table **11** for how to load orthomosaics
3	Open split screen	• Press the split screen button (bottom left-hand side, hover cursor over buttons to see names, or Compare > Enable Split Screen, or Alt + S) • Select the orthomosaic from the first time point in the left panel and the orthomosaic from the second timepoint on the right (use drop down menu on top of window)
4	Compute automatic matches	• Select Compare > Compute Automatic Matches • A new data table will appear on the right-hand side of the screen with matched polygons
5	Review matched polygons	• Zoom in on the orthomosaics and click through the data table to ensure that polygons are correctly matched • Delete incorrect matches by selecting the row in data table and pressing the delete button on keyboard
6	Conduct manual matching	• Some polygons may not have matched and will appear as only a single ID in the data table - Polygons identified in the second timepoint that are not matched will appear as “born” in the action column of data table • To manually match polygons: - Select the correct polygons in both time points (shift + double click) - Press the space bar
7	Match more than two timepoints	• Complete steps 3-6 for the first two orthomosaics in a time series • Change orthomosaics in split screen to display second orthomosaic and the following timepoint • Conduct steps 3-6 again • Repeat process for all orthomosaics within a project • This process should create a single, unique genet ID for all matched colonies
8	Export data	• Matched data can be exported as a data table (.csv) or as shapes (.svg) • Export data table: Comparison > Export Genet Data as CSV

		- This contains polygon ID number, class, and area from each matched timepoint in a project (i.e. data from all timepoints within one file) - These files do not contain 'Action' information (grow, shrink, born, same) • Separate data tables for each step in the timeseries can also be produced: Comparison > Export Matches - This will export separate files comparing data from each transition separately (e.g. T_0 to T_1, T_0 to T_2, T_1 to T_2) • Export shapes: Comparison > Export Genet Data as Shapes
9	Considerations	• Notes are not included with polygon data when matched, therefore are not included in the matched data table - If any crucial information is recorded in the Notes section, it's recommended to follow the joining process in section 3.6.2.2

3.6.2.2 Matching colonies across timepoints in ArcGIS Pro

Matching polygons in ArcGIS Pro is a three-step process that includes a data standardisation step, a joining step, and a quality control (QAQC) step. Each step contains its own series of sub-steps, which are explained in detail in the following sections. The process uses the shapefiles exported from each orthomosaic within a TagLab project (see section 3.5 for steps on how to export data).

1. Standardisation: standardises shapefiles exported from TagLab in ArcGIS Pro to ensure there's no discrepancies in data (projection, planar area, attribute table, etc.) between orthomosaics.
2. Join: links colonies through time by matching shapefiles from all timepoints in ArcGIS Pro and performs a spatial join on the colony polygons. Once colonies are linked, a .csv file is produced where each row represents a single colony that contains data for each timepoint in the columns
3. QAQC: removes errors from join process, clarifies colony notes, and generates clean, user-friendly dataset

All steps in the matching process have been scripted in python and are contained in a single .ipynb file (also available as an html file). This script is referred to as the "spatial join script" throughout the following sections and can either be found on the AIMS GitHub page or by contacting a member of the EcoRRAP team. Each step is explained throughout the script file as well as in sections below.

Table **22** and Table 23 explain the manual steps that can be done in ArcGIS Pro for a single pair of shapefiles for the standardisation and join processes, respectively. Table 24 explains the steps to run the spatial join script on multiple sets of shapefiles within a folder.

Standardisation

The standardisation step is a quality control measure aimed at removing any potential discrepancies in the size data of polygons across projects, for example if an orthomosaic did not have the pixel size input correctly. To control for this, size metrics are standardised in ArcGIS by defining the same coordinate system projection for all shapefiles and subsequently calculating the planar area of each polygon.

The EcoRRAP standardisation step carries out the following for each shapefile:

- Defines projection as WGS 1984 Web Mercator (auxiliary sphere)
- Calculates planar area in m^2
- Adds EcoRRAP Plot_ID and Date fields
- Fills empty cells with "NA"
- Removes extraneous TagLab fields: TL_Date, TL_Genet, TL_Cx, TL_Cy, TL_Area, TL_SurfA, TL_Perim (see Table **22** for further details on these fields)
- Exports one csv table per standardised shapefile

Table **22** explains how to conduct the above steps manually in ArcGIS Pro for a single shapefile (see Table 24 for steps to run standardization on multiple files within a single folder with the spatial join script). Note that this process modifies the input shapefiles directly, it is therefore recommended to run this process on duplicated files in order to retain the original, unmodified raw files. Certain steps, noted with an asterisk (*) in Table **22**, are specific to the EcoRRAP dataset. See the ArcGIS website for more information about the specific tools mentioned in Table **22**.

Table 22. Manual steps for standardisation of the projection, planar area, and attribute table in a single shapefile within ArcGIS Pro. Steps with an asterisk (*) are specific to the EcoRRAP dataset.

Step	Job name	Job description	Settings/manual steps

1	Load shapefile in ArcGIS Pro	Open ArcGIS Pro and load selected shapefile(s) to standardise	• Click 'Start without template'
			• Click New Map (top let corner in Insert tab)
			• Add folder connection to shapefile location
			  - Right click on "Folders" in Catalog pane (right hand side of screen)
			  - Click "Add Folder Connection"
			  - Navigate to folder with the shapefile and click "OK"
			• The selected folder connection will now appear as a subfolder in "Folders"
			  - Click the drop-down arrow to the left of the folder and locate selected shapefile
			  - Drag and drop shapefile in map screen

2	Set projection	Use the Define Projection tool to set the projection to "WGS 1984 Web Mercator (auxiliary sphere)"	• Click the Analysis tab at top of screen
			• Click "Tools" (Geoprocessing pane should open on right hand side)
			• Type "define projection" in search bar of geoprocessing pane and select tool
			• Input dataset: select loaded shapefile
			• Coordinate system:
			  - Click globe symbol to the right of "Coordinate System" menu
			  - Click arrow to the left of "Projected Coordinate System"
			  - Click arrow to the left of "World"
			  - Scroll down to find and select "WGS 1984 Web Mercator (auxiliary sphere)"
			• Press Run

3	Calculate geometry attributes	Uses Calculate Geometry Attributes tool to calculate the planar area of each polygon in square meters **Note that this tool modifies the input shapefile directly. To retain the raw file, run this process on a duplicate file.**	• Return to Geoprocessing pane • Search for "Calculate Geometry Attributes" and select tool • Input feature: select loaded shapefile • Geometry attributes: - Field: enter name of calculation (will appear as a column header), type "Area" - Property: Area (geodesic) - Area unit: Square meters - Coordinate system: select the shapefile layer • Press Run
4	Add relevant fields*	Uses the Calculate Field tool to: • Add column for EcoRRAP 'Plot_ID' based on .shp filename • Add column for date of EcoRRAP plot imaging (year and month) based on .shp filename	• Return to Geoprocessing pane • Search for "Calculate Field" and select tool • Input table: select loaded shapefile • Field name: name of field to add (Plot_ID or Date) • Field type: Text • Expression type: Python • Expression: - Fields: leave empty. If adding a new column, the Field Name will populate a new column - Enter text to populate field in box above 'Code Block' with quotation marks - E.g. for plot ID, should appear as Plot_ID = "CBHE_BA1D_P1" • Leave "Enforce Domains" unticked • Press Run • Repeat for each field to add
5	Remove extraneous fields*	Uses the Delete Field tool to remove fields populated by TagLab that are unnecessary for the EcoRRAP workflow: TL_Date: date entered in TagLab when loading orthomosaic, unnecessary for EcoRRAP purposes	• Return to Geoprocessing pane • Search for "Delete Field" and select tool • Input table: select loaded shapefile • Method: Delete fields • Fields: Check box(es) of fields to delete

		TL_Genet: unique ID attributed to each polygon that correspond to a specific timepoint, unnecessary for EcoRRAP purposes TL_Cx: polygon centroid coordinate, unnecessary for EcoRRAP purposes TL_Cy: polygon centroid coordinate, unnecessary for EcoRRAP purposes TL_Area: planar area calculated in TagLab, unnecessary for EcoRRAP purpose since area has now been calculated in ArcGIS Pro TL_SurfA: polygon surface area, requires DEM, unnecessary for EcoRRAP purposes TL_Perim: polygon perimeter, unnecessary for EcoRRAP purposes	• For EcoRRAP, the following fields are deleted: TL_Date, TL_Genet, TL_Cx, TL_Cy, TL_Area, TL_SurfA, TL_Perim
6	Save standardised shapefile output	Updated file is automatically saved as this process overwrites the original file To save an additional, separate file, follow steps to the right	• Right click on the shapefile name in the Contents pane (left side of screen) • Click Data > Export features • Save file in desired location
7	Export data table for each standardised shapefile	Exports a .csv with the new attribute table generated within the standardised shapefile	• Right click on the shapefile name in the contents pane • Click Data > Export table • Save in desired location

Spatial join to match coral colony polygons across timepoints:

The join step matches polygons in the standardised shapefiles from the same site across timepoints. Polygons are matched sequentially across timepoints, using the shapefile from the later timepoint as the reference. For example, first T_0 is matched to T_1 with T_1 as the reference file, then the $T_{0\text{-}1}$ file is matched to T_2 with T_2 as the reference file, and so on. This is so that any additional polygons added in later timepoints are retained within the final joined shapefile output.

As with the standardisation step, this process can be run manually in ArcGIS following the steps in Table 23 or can be automated with a python script. Table 24 explains how to run the EcoRRAP spatial join script in ArcGIS Pro (see section 3.6.2.2 for how to access script). The script performs the same steps as listed in the Standardisation (Table **22**) and Join (Table 23) processes but is automated for all shapefiles within designated folders. The script itself also contains step-by-step information for each code chunk.

All folders and filenames within the script are set to EcoRRAP requirements. Other users will need to change file paths and file naming to fit their data. Additionally, certain steps, such as matching shapefiles across timepoints, are based on the EcoRRAP file naming system, which contains the plot ID and year within the filename. Certain sections of the script will need to be edited to fit the user's needs.

Table 23. Manual steps to complete spatial join of shapefiles from two timepoints in ArcGIS Pro. Settings that require an input are underlined. Steps with an asterisk (*) are specific to the EcoRRAP dataset.

Step	Job name	Job description	Notes and settings
1	Load shapefiles to join	Open ArcGIS Pro and load shapefiles to be matched	• See Table **22** for steps on how to load shapefiles to ArcGIS • Drag and drop desired shapefiles into the map (i.e. 2 shapefiles from the same site from two time points). • Polygons from both shapefiles should align spatially. If not, the orthomosaics are likely not coregistered correctly – see SOP 2 for more information

2	Perform spatial join	Uses Spatial Join tool to match polygons across the two shapefiles. This step creates a new, joined shapefile with data from matched polygons (e.g. polygons in the new file contain data from both time points – note that polygons will retain the shape of the reference file).
		Settings provided here are specific to EcoRRAP requirements and should be customized as needed.
		*EcoRRAP uses a 5cm search radius because polygons (especially small ones) may have shifted slightly due to orthomosaic alignment and may not directly overlap. After testing multiple distances (5cm, 10cm, and 15cm) it was decided that 5cm allowed for many of these polygons to match without introducing incorrect matches between other adjacent colonies. Other users may want to adjust or remove this to fit their data.
			• Click the Analysis tab at top of screen
			• Click "Tools" (Geoprocessing pane should open on right hand side)
			• Search for "Spatial Join" in Geoprocessing pane and select tool
			• Target feature: select reference shapefile, recommended to use most recent of the two timepoints (e.g. T_1 if joining T_0 and T_1)
			• Join feature: select non-reference shapefile (timepoint prior to target feature)
			• Output feature class: filename of new shapefile to be generated
			  - For EcoRRAP, use following naming:
			  - ClusterReef_SiteZone_Plot_j (e.g. CBHE_BA1D_P1_j)
			• Join operation: Join one to one
			• Keep all target features: ensure this is ticked (retains new polygons in reference shapefile that do not match a polygon in shapefile from previous timepoint)
			• Match option: Largest overlap
			• *Search radius: 5cm
			• Leave everything else as default

			Press Run (new shapefile layer will appear in contents pane when finished)
3	Export shapefile	Saves new, joined shapefiles in selected folder	• Right click on the newly created feature layer in the contents pane (left side of screen) • Click Data > Export features • Save in desired location
4	Export data table	Exports attribute table from joined shapefiles as .csv files to selected folder	• Right click on the newly created feature layer in the contents pane • Click Data > Export table • Save in desired location

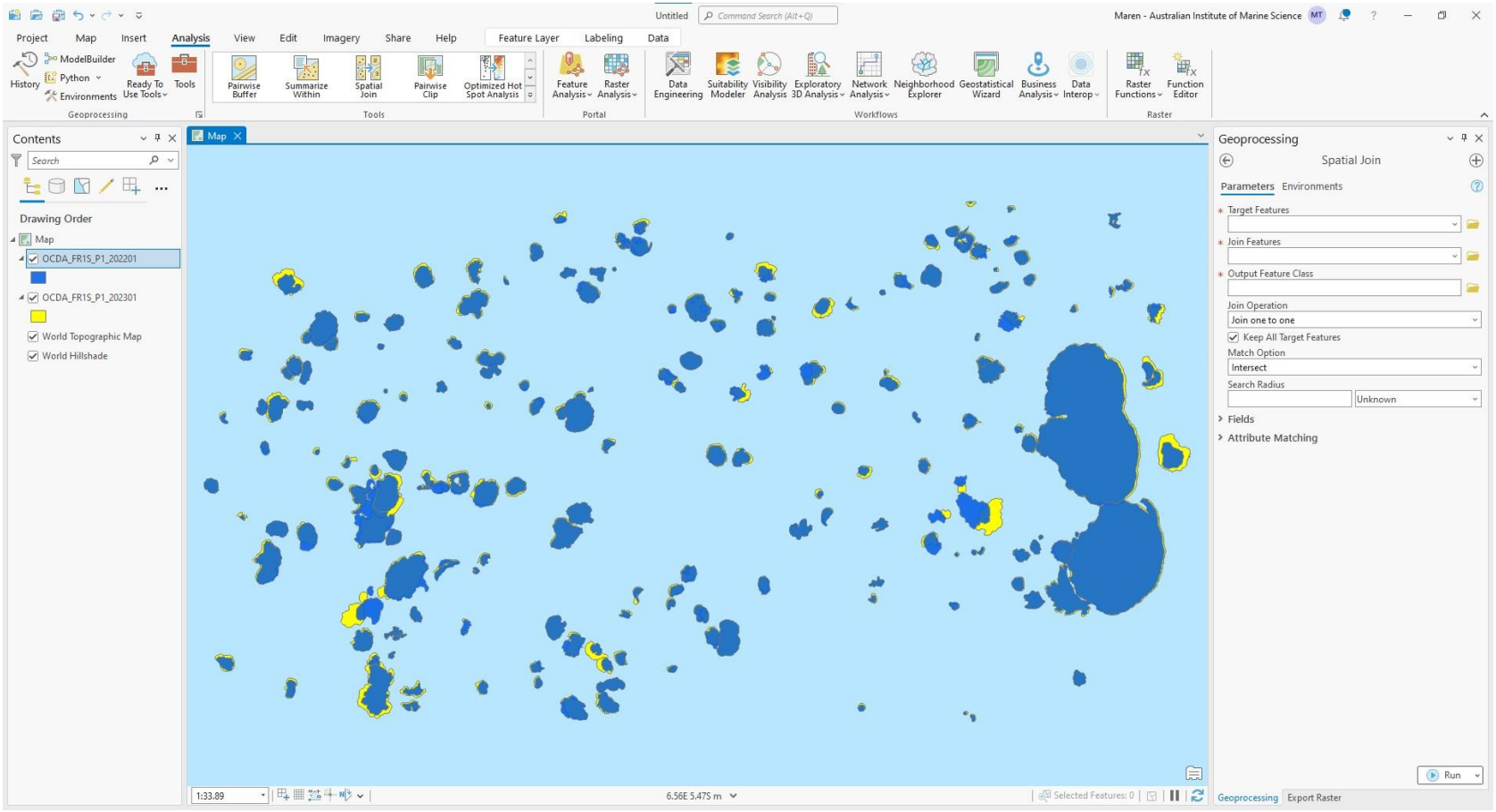

Figure 17. Shapefiles from two timepoints directly overlaid in ArcGIS Pro in preparation for the spatial join process Target features (reference shapefile, e.g. T_1) appears in blue, while the join features (non-reference shapefile, e.g. T_0) appears underneath in yellow. Settings for spatial join to be entered in Geoprocessing pane on right-hand side.

Table 24. Steps to run EcoRRAP spatial join script in ArcGIS Pro. Non-EcoRRAP users will need to edit filepaths, file naming, and any steps that rely on EcoRRAP file names to run, noted with an asterisk (*) in the table.

Step	Job name	Job description	Notes and settings
1	Create folder for Join files	Users can either copy the EcoRRAP file structure or create a file structure they prefer Note that file paths will need to be updated in the script as folder structure/names change	• EcoRRAP folder structure template: \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Shapefiles\Demography\Join_EmptyFolderStructure • Copy template folder and change folder name to YYYY_Join (replace YYYY with year of most recent timepoint to join) • Change subfolder titled "YYYY_Join" to the most recent year that is being joined, e.g. 2024_Join
2	Open spatial join script	This script is available as an .ipynb file and an html file, code can be copied from either file.	• Script location (internal): \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Shapefiles\Demography\Spatial_Joins\Scripts • Script location (GitHub): AIMS GitHub Contact Maren Toor if unable to access
3	Open ArcGIS, locate 'python window'	Open ArcGIS Pro and locate python window to run ArcPy scripts	• Click Analysis tab at top of screen • Click the 'Python' drop-down menu (top left corner) • Select 'Python Window' • A pop-up window should appear that says "Python" and has a line to enter text that says "Enter python code here"
4	Check all shapefiles are present in raw folder	This step is to double check that all required files have been exported and exist within the folder	• Create a .txt file with a list of all filenames that are expected to be in the folder - For EcoRRAP users, see appendix 5.2.4 for how to create list • Save file within the same folder as shapefiles • Copy and paste code chunk from "Preliminary Steps" into the python window in ArcGIS, change any filepaths and file names as needed • Press control and enter to run code • Output will list any files that are missing (based on .txt file) or state that all files are present • Export and add any missing files to the folder

5	Copy raw shapefiles	The standardization step in this process modifies and overwrites the shapefiles so best practice is to duplicate files to a separate folder (created in step 1)	• EcoRRAP raw shapefile folder: \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Shapefiles\Demography\Raw • Destination folder: "Standardised_Step1_Area" folder within YYYY_Join folder
6	Shapefile standardisation	*This step uses filenames to generate EcoRRAP plot IDs and dates. Non-EcoRRAP users will need to edit this code chunk or have the same naming system for this step to work. This script also produces a second set of shapefiles where each polygon has been edited to generate the convex hull and diameter values (saved to Standardised_Step2_Diameter folder). Depending on the number of files, this may take some time. For example, for EcoRRAP to run this on a folder with ~1000 .shp files it takes approximately 6 hours.	• Copy the code chunk in Step 1 of script and paste into the Python window of ArcGIS • Double check filepaths are correct for input and output folders - Input folder should be that of duplicated files (Standardised_Step1_Area) - Diameter files will be output to a separate folder (Standardised_Step2_Diameter) - Exported .csv files will be output to a separate folder (Standardised_csv) • Press control and enter to run
7	Separate standardised files by year	For the join part of this script to run, shapefiles need to be copied to separate folders based on year *This step uses filenames to allocate files to different folders. Non-EcoRRAP users will need to edit this code chunk or have the same naming system for this step to work.	• Copy code chunk from Step 2 in script and paste into ArcGIS python window • Change filepaths as required • Press control and enter to run • If additional years are being joined to EcoRRAP data the code will need to be edited to reflect this.
8	Run first spatial join	Files from T_0 and T_1 will be joined	• Copy code chunk from Step 3 in script and paste into ArcGIS python window • Change filepaths as required

		For EcoRRAP this is the 2021-2022 join. *This step uses filenames to match shapefiles from two time points. Non-EcoRRAP users will need to edit this code chunk or have the same naming system in order for this step to work. If only one transition exists (e.g. two timepoints), this is the final step in the join process and users can proceed to the QAQC process. Only proceed to further steps in the join process if multiple transitions need to be joined (e.g. three or more timepoints).	• Press control and enter to run
9	Incorporate any non-sampled plots from T1	If any plots were not sampled in T_1 (not imaged in the field or not processed in Metashape), the T_0 files will need to be joined to empty shapefiles so that all joined files maintain the same structure and it's identified that no data exists for that year.	• This step is divided into 4 sub-steps in the script due to complexity - Generate empty shapefiles for non-sampled sites - Copy empty files to Join folder - Run first spatial join with sampled year (T_0) as the reference file - Populate empty columns for non-sampled year with NS to indicate "not sampled" • See script for more detail on each step • Copy and paste each code chunk into ArcGIS separately and run sequentially • Change filepaths as required
10	Run second spatial join	Joined files from T_0-T_1 will be joined with T_2 For EcoRRAP this is the 2022-2023 join. *This step uses filenames to match shapefiles from two time points. Non-EcoRRAP users will need to edit this code	• Copy code chunk from Step 5 in script and paste into python window in ArcGIS • Change filepaths as required • Press control and enter to run

		chunk or have the same naming system in order for this step to work.	
11	Incorporate any non-sampled plots from T2	If any plots were not sampled in T_2 (not imaged in the field or not processed in Metashape), the joined $T_{0\text{-}1}$ files will need to be joined to empty shapefiles so that files maintain the same structure and it's identified that no data exists for that year.	• This step is divided into 4 sub-steps in the script due to complexity - Generate empty shapes for non-sampled sites - Copy empty files to Join folder - Run second spatial join with sampled year ($T_{0\text{-}1}$ join) as the reference file - Populate empty columns for non-sampled year with NS to indicate "not sampled" • See script for more detail on each step • Copy and paste each code chunk into ArcGIS separately and run sequentially • Change filepaths as required • Press control and enter to run
12	Repeat steps 10-11	If additional time points need to be joined, repeat Join and Incorporation steps for any missing plots as required	• Will require following substeps for each timepoint, if any plots are not sampled - Generate empty shapefiles for non-sampled sites - Copy empty files to Join folder - Run spatial join with sampled year as the reference file - Populate empty columns for non-sampled year with NS to indicate "not sampled"
13	Finalisation	All joined shapefiles should now be present within a single folder All exported csv files should be present within a separate sub-folder To merge exported csv files into final combined dataset, follow steps in the Data QAQC process below	• Joined shapefiles should be located in YYYY_Join subfolder • Exported csv files should be located in 'csv' folder within YYYY_Join subfolder

3.6.2.3 Dataset QAQC (cleaning)

The QAQC process is required to ensure data quality. The steps listed in this section explain the process used for the EcoRRAP dataset, however, QAQC steps will vary for other users and should be adapted as required. All steps here are conducted in R studio (v2023.09.1) with a script written specifically for this process. This script can either be found on the AIMS GitHub page or by contacting a member of the EcoRRAP team (internal link: Scripts).

There is a margin of error in the joining process that can lead to incorrectly matched polygons. Mis-matched polygons can occur due to natural processes on the reef if colonies shift or overgrow adjacent colonies. They can also occur if orthomosaics do not align perfectly between timepoints, either due to minor issues with the co-registration or distortion within the orthomosaic. Even slight shifts in orthomosaic orientation can lead to improper colony alignment, especially with small colonies. This may cause two different colonies to incorrectly join or cause a colony to have no match at all during the joining process (the latter problem is partially accounted for during the join process with the 5cm search radius when joining polygons). The steps taken here serve to remove those errors where possible, however, there is potential for mis-matched colonies to remain in the dataset, which cannot be detected without manual review of every colony.

The QAQC process used for the EcoRRAP dataset removes incorrectly matched colonies and adds columns to clarify the EcoRRAP naming codes. Each step in the cleaning process is explained in Table **25**. In addition to these steps, the cleaning process incorporates and clarifies notes attributed in TagLab for ease of use (for example changes the note 'd' to 'Dead', 'm' to "Missing', etc.). Clarified notes are then used to create additional columns that explain whether a colony can be used for survival or growth analyses in the EcoRRAP dataset – for example, a distorted colony (note: 'b' for Bad) should not be used for growth metrics but can be used for survival metrics. See Table **17** for notes used in the EcoRRAP dataset.

For more information about filtering columns and the EcoRRAP Demography dataset, contact Maren Toor (m.toor@aims.gov.au) for access to the associated readme file.

Table **25** explains the steps taken in the EcoRRAP Demography Dataset QAQC script. The script itself also contains information about each code chunk.

Table 25. Steps in the EcoRRAP data cleaning process. Steps with an asterisk (*) are specific to the EcoRRAP dataset.

Option	Notes
Merge csv files	• Merges individual .csv files (one per plot) produced during the join process into a single .csv file • Creates a file for the T_0-T_1 (2021-2022) join and a file for the T_0-T_1-T_2 (2021-2022-2023) join
*Change EcoRRAP plot ID names	• Changes the EcoRRAP Plot_ID (e.g. CBHE_BA1D_P1) to full names and separates the ID into the following columns: - Cluster (e.g. Offshore southern) - Reef (e.g. Heron Island) - Site (e.g. Back1) - Zone (e.g. Deep) - Plot (e.g. Plot1)

Remove data for incorrect matches	• Finds and replaces data in years where a colony polygon has incorrectly matched during the join process • Where the colony Class is the same in two years but different in the other year, data in the different year is replaced with "IM" to identify it as an "incorrect match" (e.g. Class_2021 is Acor but Class_2022 and Class_2023 are Pmas, the data in 2021 is replaced with IM). Data is replaced with "IM" in the following columns: - Class - Note - Date - Area: data in this column is changed to NA (instead of IM) to keep the column numeric - ID_YYYY: data in this column is changed to NA (instead of IM) to keep the column numeric • Note that if a dataset only includes one transition, then any colonies where class does not match in both timepoints should be removed
Remove 0s in non-sampled plots	• For plots that were not sampled in a timepoint, Area and ID columns were populated with 0 values during the join step • This changes the 0 values to NA to avoid incorrect numeric values
Remove colonies with data from only one timepoint	• Can occur from: - Shapefile misalignment that results in no polygon overlap so no colony joins - New colony added in T_1 at a site that was not imaged at T_2 • Removes rows where the Class is only present in one year - Note that this specific dataset is for the purpose of temporal analyses so data from only a single year does not contribute to the dataset.
Remove incorrect "new" colonies	• Removes colonies that are identified as New in T_1 but colony data is present in T_0 columns • Occurs due to mis-matched polygons that are adjacent • Note that this occurs after the removal of incorrect matches and many of these colonies will already have had the T_0 data replaced with "IM" (if the class from T_1 and T_2 match) and data for the following timepoints will remain - This step does not remove colonies where the 2021 data is "NA" or "IM"
Remove non-coral polygons	• Data is unreliable across years and irrelevant to dataset output • Includes: - Calcareous algae (Calg) - Rubble (Rubl) - Reef substrate (Rsub) - Sand (Sand)
Remove colonies with different class in each timepoint	• For example, multiple incorrect matches across all timepoints • In these cases, it is impossible to know which class is correct unless reviewing the individual colonies
Clarify notes	• Translates all notes made in TagLab into clear, user-friendly terminology

	- All possible notes are translated into the 14 notes listed in Table **17** - Interprets any misspellings or typos in notes that may have occurred during the digitisation process in Taglab • Creates new columns with translated notes for each year (Clarified_Note_YYYY)
Create columns for dataset filtering	• To make the dataset user-friendly, new columns are created for ease of filtering out certain colonies that should not be used for analyses • For more information on this, see EcoRRAP Demographic Dataset readme file (internal link: 2023 demographic dataset readme)
Remove incorrect "missing" colonies	• Removes colonies that have the note "Missing" in T_1 but data is present in T_2 • Can occur due to the following: - Colonies are mismatched during the join process but due to having the same class it is impossible to know which colony is correct unless reviewing the individual colonies - Observer error during the transfer process where either the colony is correctly missing in T_1 and misidentified in T_2 or the colony is incorrectly identified as missing in T_1 and relocated in T_2
Rename column headers	• For clarity and to include additional information • Adds units (square meters) to Area columns (Area_YYYY -> Area_YYYY_sqm) • Identifies TagLab ID columns (ID_YYYY -> TagLabID_YYYY)
Add unique colony ID	• Adds a unique numeric ID to each colony in the dataset

Even with the QAQC steps taken in Table **25**, it is possible for errors to persist within a dataset produced in this manner. The steps above are provided as a guide but users are advised to implement their own QAQC steps as required to ensure the integrity of their data.

3.6.2.4 Additional considerations and limitations

Table **26** lists additional considerations for data collected and processed in the manner explained throughout this SOP. Certain considerations apply to all photogrammetric methodologies while others are specific to the EcoRRAP dataset.

Table 26. Additional considerations and limitations for the EcoRRAP demographic dataset.

Option	Notes
Potential overestimation of mortality	• Colonies can only be viewed from one two-dimensional angle, therefore, when there is no live tissue within the colony polygon it gets labelled as dead. However, it is possible that small sections of live tissue may still be present on the sides or undersides of colonies, which are not visible from the 2D angle. This is a limitation of this method and such colonies cannot be identified within this dataset.
Use of staghorn Acropora (Asta)	• Colonies are digitised based on branch extent without internal, interstitial space cut out from polygons.

	• It is recommended not to use these colonies for growth analyses due to rough colony margins but they can be used for survival metrics. • Colonies often merge into thickets through time and are noted as such.
Incorrect colony matches	• The majority of mis-matched colonies identified during the join process are removed during the QAQC process, however, it is still possible that a small number of colonies with the same class have incorrectly matched across multiple timepoints. • This can happen if colonies of the same taxa are located adjacent to each other. While uncommon, users should be aware of this as a potential source of error in the dataset.
Coral taxonomy	• The EcoRRAP dataset uses species names that have since been taxonomically revised • See annexure for EcoRRAP species list, including naming system within the dataset as well as the corresponding updated taxonomic names

3.7 Automatic digitisation in TagLab

TagLab has the capability to train automatic classifiers by applying the user's data to optimize the digitisation process. Automatic digitisation in TagLab includes three main steps:

1. Preparing the training data
2. Training the classifier
3. Running the classifier

Each step is covered in further detail below but it is important to note that classifier quality is dependent on the amount and quality of training data as well as on the training process. Selection of training data and training settings is also highly dependent on the user's dataset. The following instructions are intended to be used as a general workflow guide with settings altered as needed to suit the user's requirements. See the Learning Pipeline section of the TagLab website for further details.

It should also be noted that this process requires the Nvidia CUDA Toolkit installation explained in section 3.2.1.3 and TagLab should not be installed with the "cpu" command.

EcoRRAP has trained nine automatic classifiers that work to varying degrees. Once trained, these classifiers were used to locate colonies within orthomosaics, which was followed by manual review of all polygons to refine and add notes where required. See Table **34** for information on which taxa EcoRRAP has built automatic classifier for using TagLab.

3.7.1 Training automatic classifiers in TagLab overview

The process here explains how to build a classifier for a single class. Training tests conducted on EcoRRAP data determined that building more than one class in a classifier resulted in a class imbalance with the poor identification of classes with fewer and smaller segments in the training data (see Pavoni et al. 2020 for more information). However, it is possible to train multiple classes in

a single classifier and various options should be tested for each individual dataset to ensure classifiers are optimised on a case-by-case basis.

Building an automatic classifier in TagLab requires the generation of training data and the subsequent model training with the generated data. Training data constitutes all of the information that the classifier is provided about your data. A classifier can only be as good as its training data.

The amount of training data required is subjective, but the general rule is that the more training data input, the better the classifier output will be, up to a certain threshold. It should be noted, however, that a larger quantity of training data will increase the amount of time required for both data extraction and training so the optimal quantity of training data may need to be determined through testing. Additionally, if the classifier is to be used on reefs from different locations or habitats, it's recommended to extract training data from different areas in order for the model to better generalise across locations (see Pavoni et al. 2020 for more information).

To create training data from orthomosaics, TagLab splits the 'Area to export' (i.e. the region that you instruct TagLab to extract training data from) into many small, square tiles (see Learning Pipeline on TagLab website for example). TagLab uses 75% of tiles in the 'Training' stage, and 15% in the 'Validation' and 'Testing' stages. Note that the unit used throughout the training process is 'tiles', not 'polygons' (i.e. 75% of *tiles* are used in the Training stage, not necessarily 75% of polygons).

3.7.1.1 Preparing training data

General recommendations for the selection of training data in TagLab are listed in Table **27**. Steps to export training data are listed in Table **28**.

Table 27. General recommendations for the selection of classifier training data.

Option	Notes
Number of orthomosaics	Use training data from fewer orthomosaics with a higher quantity of polygons as opposed to data from more orthomosaics with fewer polygons in each. • May be driven by abundance of target class • As an example, EcoRRAP has built the following classifiers, which have produced relatively successful results with the amount of data used: - Tabulate *Acropora*: 513 colonies across 20 orthomosaics - *Pocillopora damicornis*: 318 colonies across 18 orthomosaics
Variation in orthomosaic imagery	• If imagery is variable between orthomosaics, it is recommended to use orthomosaics from each type of environment so that the classifier recognises the class in each type of imagery (e.g. different habitats, lighting, coloration, etc.) • See Figure 18 for examples of different EcoRRAP orthomosaics
Area to export	• Ensure that all colonies of your target taxa located within the 'Area to export' are digitised. - Any missed colonies in this area will impair classifier training and reduce classifier quality (n.b. not digitising a colony included in the Area to export is equivalent to labelling it as 'Background' and actively saying that it is *not* what you want the classifier to locate).

- Ideally ensure that colonies of your target taxa are spread approximately evenly across the Area to export.
    - This helps to provide the desired split of polygons (not just tiles) across Training, Validation and Testing data.
    - If target taxa are not distributed equally around the orthomosaic, consider adjusting the Area to export to only encompass the area in which the target taxa are found

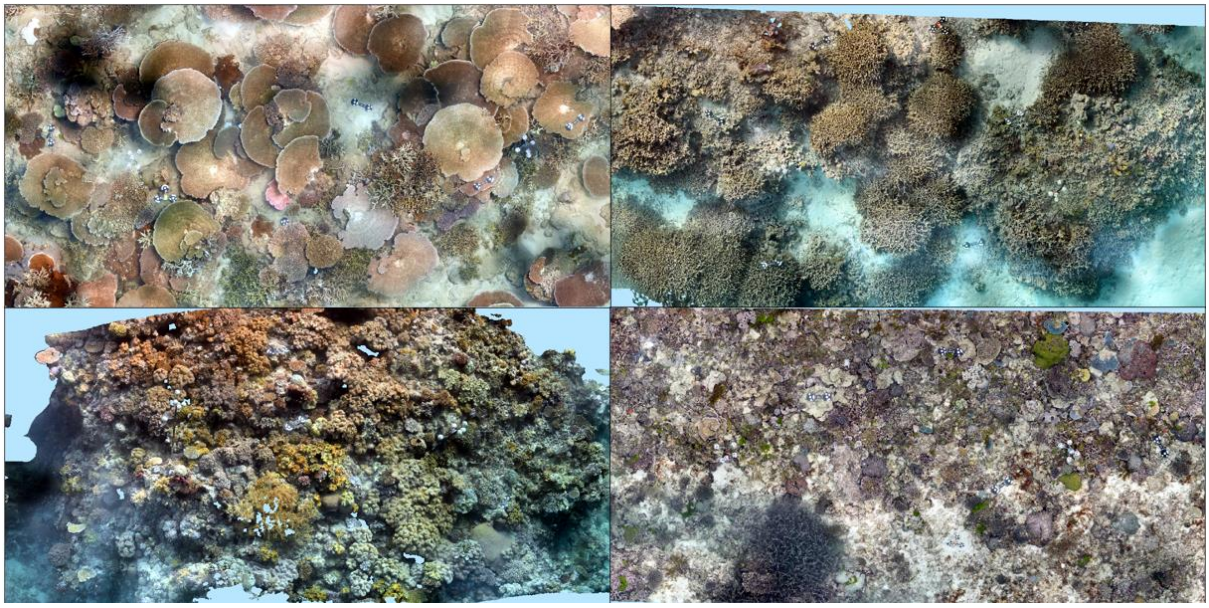

Figure 18. Examples of four EcoRRAP orthomosaics with variation in lighting, coloration, and community composition.

The steps in Table **28** describe the process involved in exporting training data from TagLab to build automatic classifiers.

Table 28. Steps to export training data to generate automatic classifiers in TagLab. Settings listed are recommendations, users should adjust as required for their dataset.

Option	Key task	Notes
1	Manually digitise sufficient training data	• Ensure all colonies of your target taxa located within the 'Area to export' are digitised • Tip: it's beneficial to have more polygons in a single orthomosaic for training than the same number of polygons spread across multiple orthomosaics • EcoRRAP classifiers are built with anywhere between 100-600 polygons
2	Create data folder	• Create a folder for TagLab to deposit training data into • Tip: name the folder '[*classifier name*] training data' • For EcoRRAP, create folder here: \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Classifiers\Training_files
3	Hide non-target classes	• Turn off visibility of any classes that should not be included in the classifier (anything hidden will not be exported) • Click the 'eye' icon to the left of non-desired class name(s) in Labels panel (eye will turn into an 'x' and polygons in this class will not be visible on orthomosaic, Figure 19a)

		• Tip: Press control and left-click on icon of desired class and TagLab will hide all other classes simultaneously - To reset visibility: press shift and left-click on the class to turn visibility on for all classes except the one clicked
4	Export training data	• File > Export > Export New Training Dataset • Fill export settings (Figure 19b): - Dataset folder: click '…', navigate to data folder created in step 2 - Area to export: leave as default for entire orthomosaic, or click the square icon to the right and draw the perimeter around the desired area - Dataset split: Uniform (vertical) ▪ Up to user but note that TagLab has crashed when attempting to use 'biologically-inspired' split - Target scale: should match pixel size of orthomosaics ▪ Note that changes to the target scale will impact the time required to export data and train the classifier - Check 'Show exported tiles' (useful to check the distribution of polygons across the tiles; see step 5) - Leave 'Export in coco panoptic format' unchecked • Once TagLab starts the export, three sub-folders and a text file will appear within the data folder: - Test - Training - Validation - Text file contains target pixel size set during export • Export may be slow (e.g. up to a couple of hours), depending on computing power and the size of the Area to export • See TagLab website or Pavoni et al. 2022 for detailed explanations
5	Check training data	• If 'Show exported tiles' was ticked, TagLab will create a summary image of the training data showing the distribution of tiles and polygons within tiles - See the Learning Pipeline section of the TagLab website for example images • To check summary image: - Open TagLab-main folder and navigate to 'tiles_cutted' folder (n.b. this file will be written-over the next time training data is exported) - Check that polygons are well distributed across training, validation, and testing tiles
6	Export data from multiple projects/orthomosaics	• Combine training data from multiple TagLab projects or orthomosaics by selecting the same dataset folder in the Export New Training Dataset window during each export • Target scale needs to be the same for every set of exported data - If the pixel size differs between projects, use an average size

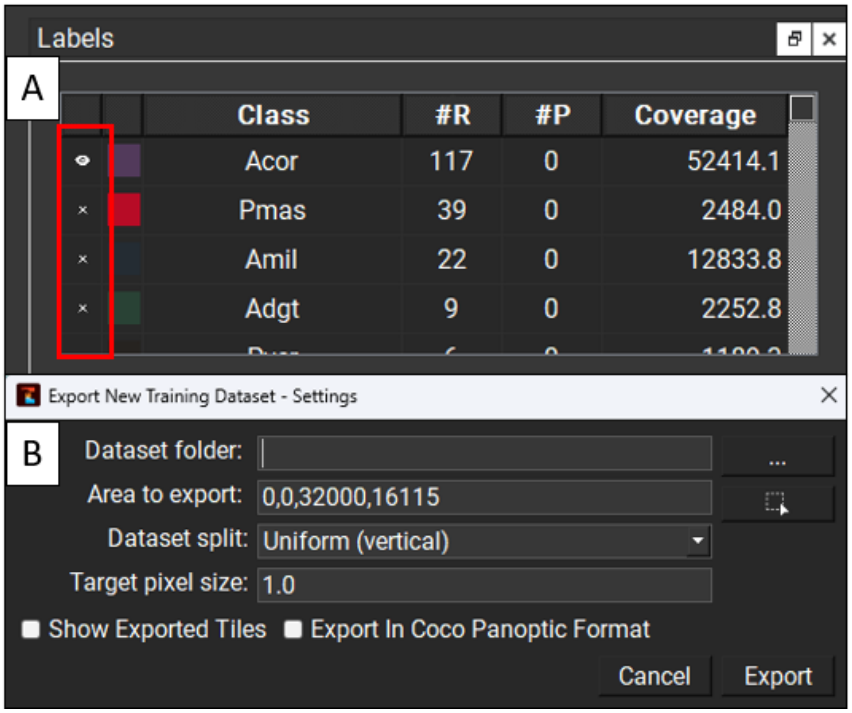

Figure 19. A) Labels panel in TagLab with class visibility outlined in red, turn off visibility of non-target classes to only export data from target class, and B) Export Training Dataset settings window.

3.7.1.2 Classifier training

Once training data has been exported, the model can be trained within TagLab. Training time is highly variable depending on computing power, the amount of training data and the settings used for training. It can take anywhere between a few hours to a few days. Follow the steps in Table **29** to train a classifier in TagLab.

Table 29. Steps to train an automatic classifier in TagLab. Settings listed are recommendations, users should adjust as required for their dataset.

Option	Key task	Notes
1	Launch TagLab	• See Table **9** for how to open TagLab • Ensure dictionary contains the class of the classifier being trained
2	Open Network Training window	• File > Train Your Network
3	Fill in training information	• Network name: Enter desired name for the classifier - For EcoRRAP, use: Class_v# (e.g. Atab_v3) • Dataset folder: Click '…' and navigate to the training data folder that was created in Table **28**, step 2 • Classes to recognise: will automatically fill once it's read training data, will display class exported and 'background' (can take a few minutes to load) • Number of epochs: recommended to use 50-60 - Determines how many times the whole training dataset is passed through the network during training. A higher number of epochs will increase the accuracy of the

		classifier but will also increase the amount of time it takes to train • Learning rate: leave as default • L2 regularization: leave as default • Batch size: recommended to use between 8-16 - Determines the number of tiles processed simultaneously before the model is updated during each epoch. A higher batch size will likely produce a better classifier but may be limited by computing capability • See TagLab website for more details on settings
4	Train classifier	• Press 'Train' • Depending on the amount of computing power, training data and settings used, training may take anywhere between a few hours to a few days
5	Test classifier	• A results window will appear when training is finished - Will display results (accuracy and mIoU) and training graph • Test results by pressing 'Select' in the Predictions box and choosing a training tile - This will show the ground-truthed data and the predicted classifier data side by side • Review multiple tiles to see results • Note that if the class is sparsely located in orthomosaics, it can be difficult to tell how well the classifier will work on new data from these results - See TagLab website for details on training results
6	Save classifier	• If results are successful, press 'Confirm' to save the classifier • See Table **30** for how to run classifier on full orthomosaic

3.7.2 Running automatic classifiers in TagLab

Once a classifier has been trained and saved, it will appear within the Fully Auto Segmentation window in TagLab and can be run on an orthomosaic. Note that any classifier trained by a user will only appear in the TagLab window that is running out of the same TagLab-main folder where it was trained. To use a classifier on a different computer, that specific TagLab-main folder will need to be transferred across. Steps in Table **30** explain how to run automatic classifiers in TagLab.

Table 30. Steps to run automatic classifiers in TagLab.

Option	Key task	Notes
1	Open or create a project	• See Table **11** for steps on how to create a project, if required • Make sure the pixel size is input when loading the orthomosaic (classifier will take longer to run otherwise) • Make sure that the dictionary loaded in the project contains the class label that the classifier was trained for
2	Select classifier	• Click the 'Fully auto semantic segmentation' button in the left-hand panel (see Figure 9 for location) • Select the classifier to run from the '*Classifier:*' drop-down menu in the Select Classifier window (Figure 20)

		• Details about classifier will appear below drop-down menu • All classes listed in the 'Classes Recognized' row must also be listed in the dictionary loaded (except 'Background')
3	Check classifier prediction (optional)	• Select the dashed-square icon under 'Check Classifier Prediction' section of Select Classifier window (Figure 20) • Draw a rectangle around a section of the orthomosaic with the target substrate – the section will appear in left-hand panel of window • Click Apply in bottom right-hand corner of window • Classifier output for that section will appear in the right-hand panel of the window
4	Run classifier	• If 'Check classifier prediction' was run in previous step, close and re-open the Select Classifier window (step 2) • Select desired classifier from '*Classifier:*' drop down menu • Click Apply • Depending on computing power, classification may take between 5-45 minutes to run • A progress bar will appear in the top left-hand corner of the TagLab screen
5	Save project	• Once finished, a pop-up window will appear to save the project (required) • Once saved, the project will automatically close and will need to be reopened to review • Tip: if the project is already saved, it's possible to save over the previous file to reduce the duplication of projects
6	Running multiple classifiers	• Repeat steps 1-5 for as many classifiers as desired • Note: You can have multiple TagLab windows open as long as they are operating out of separate Anaconda Prompt windows, but do not run classifiers in multiple TagLab windows at the same time because they interfere with each other
7	Review polygons	• Review each polygon that the classifier generates to: - Confirm that it has digitised the correct substrate and delete any false positives - Edit polygon border, if required - Add any notes required • Review the full orthomosaic to manually digitise any colonies that the classifier missed (false negatives).

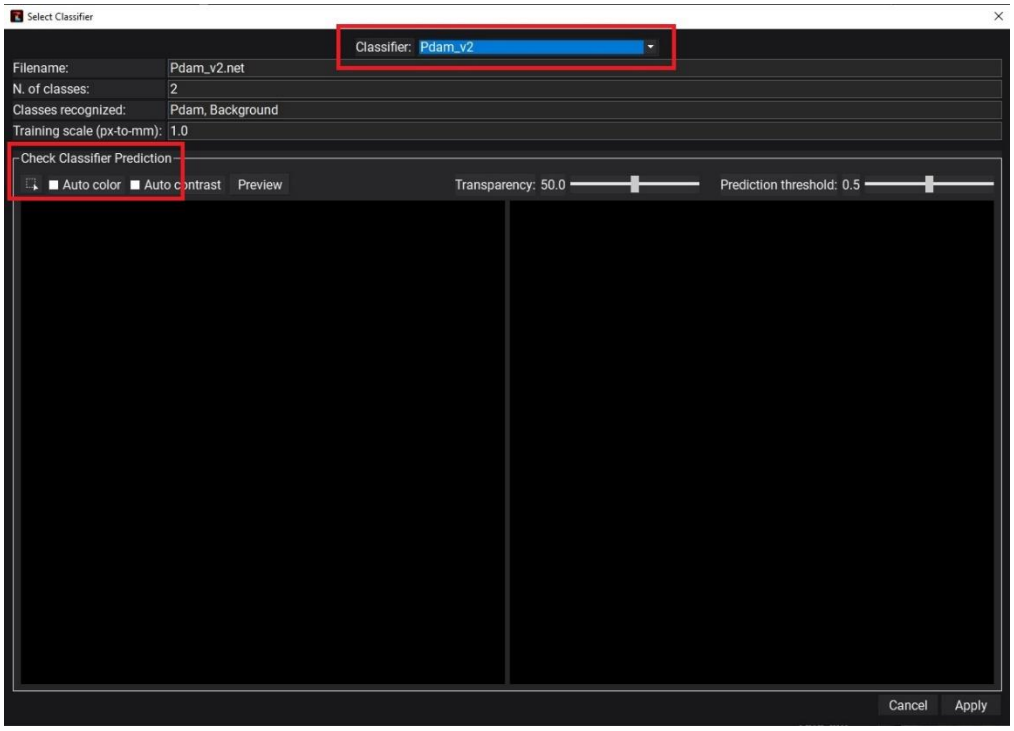

Figure 20. Select Classifier window with drop-down menu and Check Classifier Prediction options outlined in red.

Depending on training data, training settings, and the orthomosaics that the classifier is being run on, colony detection and accuracy of segmentation can vary. See Figure 21 for examples of EcoRRAP classifier outputs.

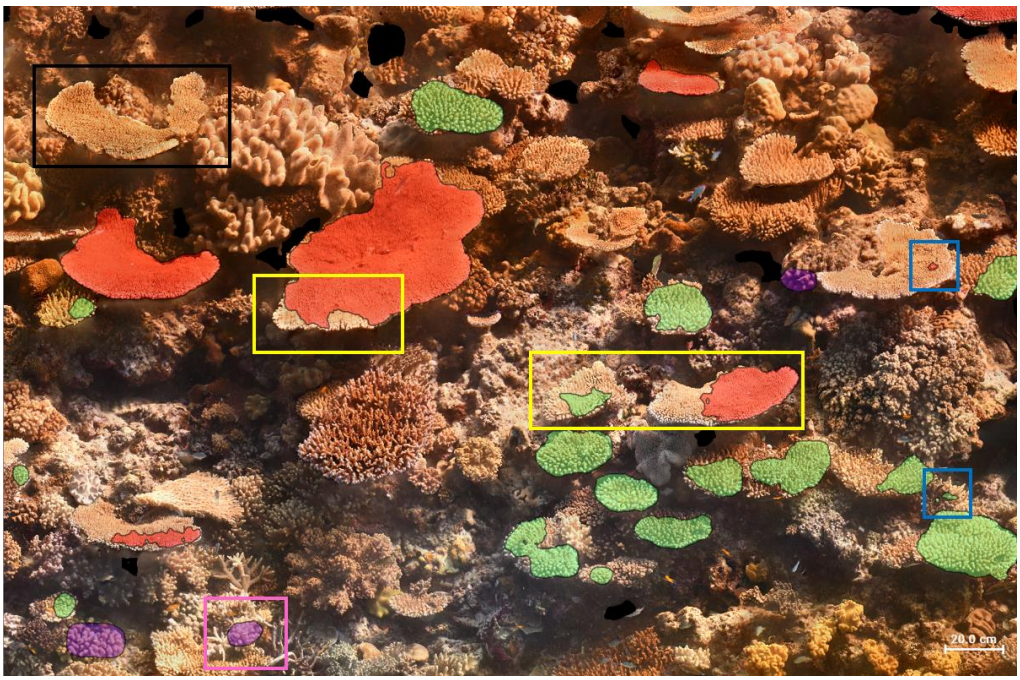

Figure 21. Examples of EcoRRAP classifier outputs for tabular Acropora (red colonies), corymbose Acropora (green colonies) and Pocillopora (purple colonies). Boxes highlight errors in output: black box indicates false negatives, pink indicates false positives, yellow indicates incomplete colony delineation, and blue indicates small segments that only encompass a fraction of the intended colony.

4 ADDITIONAL RESOURCES

4.1 Additional resources/quick links

- EcoRRAP links
  - Website: EcoRRAP (ecological intelligence for reef restoration)(gbrrestoration.org)
  - SOPs: Reef monitoring sampling methods | AIMS
  - Metadata: EcoRRAP Metadata
  - Data management templates: EcoRRAP Photogrammetry Data Management Templates
  - EcoRRAP Equipment Schematics: EcoRRAP Photogrammetry Equipment Schematics
  - GitHub: Australian Institute of Marine Science

- TagLab
  - Website: TagLab
  - GitHub: GitHub - cnr-isti-vclab/TagLab: A CNN based image segmentation tool oriented to marine data analysis

- Anaconda
  - Download software: Download Now | Anaconda
  - Documentation: Anaconda Documentation — Anaconda documentation

- Nvidia
  - Download software: CUDA Toolkit 12.8 Update 1 Downloads | NVIDIA Developer
    Documentation: CUDA Toolkit Documentation 12.8

- Microsoft Build Tools
  - Download: Microsoft C++ Build Tools - Visual Studio

4.2 EcoRRAP file links

Table **31** contains the internal file locations for key EcoRRAP demographic data files. Only internal users with access to the AIMS network and to the specified network folder will be able to access files through these links. Any other users should speak with Renata Ferrari or Maren Toor about accessing these files.

Table 31. Key EcoRRAP files relevant for the demographic workflow with file pathways (internal only).

Option	Notes
Orthomosaics	• \\pearl\3d-ltmp\EcoRRAP\outputs\orthomosaics
TagLab master projects	• \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Projects
EcoRRAP TagLab dictionary	• \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Dictionaries
EcoRRAP annotation log	• EcoRRAP_Annotation_Log
EcoRRAP classifier files	• \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Classifiers
TagLab datatable exports (.csv)	• \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Datatables\Prelim
TagLab shapefile exports (.shp)	• \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Shapefiles\Demography\Raw
TagLab labelled image exports (.png)	• \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Shapefiles\Labelled_images
Temporally-joined data (all years)	• \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Shapefiles\Demography\Spatial_Joins
Joined datatable outputs (QAQC processed)	• \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Shapefiles\Demography\2023_Join\Data_Outputs
Join and QAQC scripts	• \\pearl\3d-ltmp\EcoRRAP\outputs\TagLab\Shapefiles\Demography\Spatial_Joins\Scripts
ID validation files	• \\pearl\3d-ltmp\EcoRRAP\outputs\ID_Validation

5 APPENDICES

5.1 Troubleshooting

5.1.1 TagLab installation

It is not uncommon for errors to arise during the TagLab installation process. Issues can arise with version compatibility if the software is updated by the developers or python packages are updated. If issues arise, check the TagLab Github repository under the Discussion and Issues pages to see if any other users have encountered the same problem.

Common problems encountered are listed in Table **32**, however, the package versions or python syntax may change in the future and therefore any code listed here may not resolve the issue. If issues persist, it's recommended to google the error stated in the Anaconda Prompt.

Table 32. Potential errors that can arise during the TagLab installation process with steps for how to troubleshoot each. Text in **bold** is exact code to be written in the command line of Anaconda Prompt. This syntax can be directly copied except where specified.

	Error	Troubleshooting option
1	OsGeo error	• Due to bug in installation script (since updated in any 2025 versions) • Requires direct download of gdal and rasterio packages (should normally happen automatically) - Download packages here: taglab.isti.cnr.it/wheels/ - Version to download will correspond with version of python (e.g. "cp38" for python 3.8) - Move downloads to TagLab-main folder • Install gdal and rasterio packages in Anaconda Prompt - Type **conda install** in Anaconda Prompt, then copy and paste file pathway of gdal package (or drag and drop file in prompt) - Repeat for rasterio package - Use **conda forge** or **pip install** commands if **conda install** doesn't work • Rerun TagLab installation command (**python install.py**)
2	Non-functioning 'Refine' tool	• Likely due to versioning issue of Scikit package • Type **Pip install –iv scikit-image==0.19.3** in Anaconda Prompt - This will downgrade the Scikit package so that it is compatible with TagLab - If above doesn't work, try **Pip install scikit-image==0.19.3**
3	'Specified module could not be found'	• Identify name of module/package that caused error • Most errors related to packages or modules can be fixed by running **conda install -c conda-forge <package name>** command • Alternatively, specify a version that is known to work **conda install -c conda-forge *<insert package name>*==*<insert version>*** • Rerun TagLab installation or launch command

5.1.2 Map (orthomosaic) pixel size

Depending on the settings used to export orthomosaics from 3D model processing software, the pixel size may change from what's shown in processing software, for example, if the orthomosaic size is manually set. This can lead to errors with scaling and polygon size in TagLab, therefore, it's recommended to record the pixel size of the orthomosaic after it's exported. Steps to locate the pixel size of an orthomosaic in ArcGIS Pro are listed in Table **33**.

Table 33. Steps to identify the pixel size of an orthomosaic using ArcGIS Pro.

Step	Key task	Notes
1	Open ArcGIS Pro	• Click 'Start without a template'
2	Add folder connection	• Right click on the Folders tab in the Catalog pane (right-hand side of window) • Click Add folder connection • Navigate to folder with orthomosaics - Click OK
3	Locate orthomosaic in Catalog	• Use arrows to the left of folder icons to expand folders • Navigate within folders to the desired orthomosaic (Figure 22a)
4	Check orthomosaic properties	• Right-click on desired orthomosaic • Click Properties – this will open the Raster Dataset Properties window (Figure 22b) • Locate Raster Information (expand section if required) • Locate pixel size (called Cell Size in ArcGIS, red box in Figure 22b) • Note: the value shown there is in meters and will need to be converted to millimetres for TagLab (x1000)

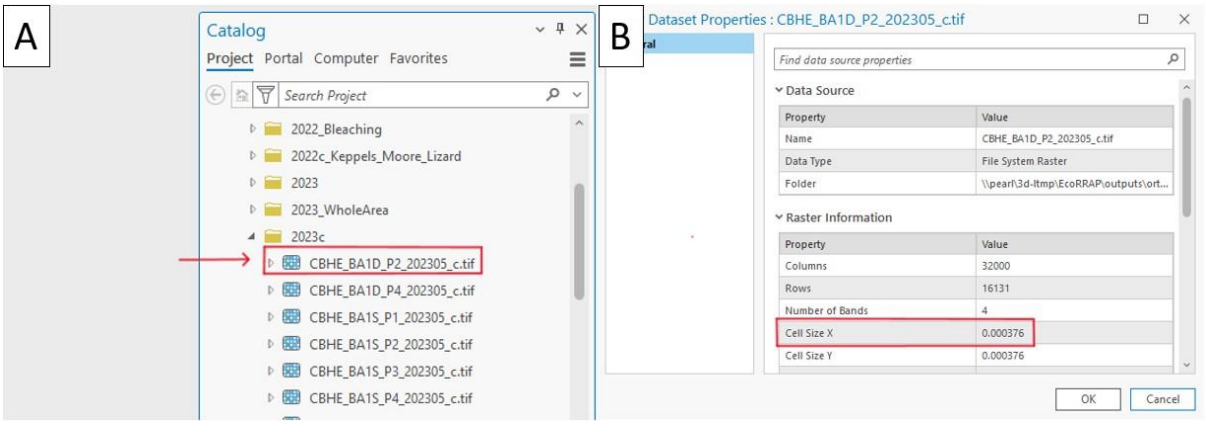

Figure 22. ArcGIS Pro window showing A) the location of raster files in the Catalog pane, after a folder location has been set, and B) the location of the pixel size (cell size) in the Dataset Properties window underneath the Raster Information tab.

5.1.3 TagLab dictionary troubleshooting

The TagLab script allows for multiple classes to have the same RGB colour. This is problematic when transferring polygons between years because all classes with the same RGB value get changed to a single class (the one listed first in the dictionary). As of version 2023.05.17 this could not be fixed by

simply changing the dictionary colour in TagLab, when the colour of one class was changed it changed all classes with that colour. Additionally, it was not fixed by creating a new dictionary and uploading that into an existing project, any taxa with duplicated RGB values in the previous dictionary were still duplicated in the new one.

This occurred with the EcoRRAP dictionary and in September 2023 a new dictionary was created with all new and unique RGB values attributed to the 190 classes in the EcoRRAP dictionary. To allow for the new dictionary to be uploaded to all projects without introducing errors, a minor section of the TagLab.py script that allowed for this duplication to happen was hashed out. A temporary script called TagLab_fix.py was used to correctly update the dictionary.

Be sure to check the dictionary when opening a project to verify that classes do not have the same values. Also note that if any changes are made to the dictionary, labelled images with classes that contain previous RGB values will not be imported into projects because TagLab will not recognise them. In this case, re-export the labelled image before importing.

5.2 EcoRRAP specific information

5.2.1 EcoRRAP species list

The EcoRRAP master species list can be found here (internal link). This list contains 190 taxa, all of which are included in the EcoRRAP dictionary. From this list, approximately twenty taxa, divided into three priority levels, were selected for digitisation to quantify demographic rates within EcoRRAP plots (Table **34**). Taxa selection was based on overall abundance, the representation of different life history strategies, the accuracy of identification from orthomosaics (determined by an *in-situ* validation process conducted by a coral taxonomist, see section 5.2.2), and the relevance of species in restoration research. Taxon identification followed a hierarchical structure so that if species-level identification could not be confirmed with at least 80% confidence, colonies were identified to a broader taxonomic level (genus or morpho-taxa). It is important to note that because the EcoRRAP dataset is ~70% offshore (clear water) reefs and ~30% inshore reefs, the priority taxa selected are more representative of mid-shelf and offshore reefs. Morpho-taxa descriptions follow that of the CATAMI classification scheme (Althaus, Hill, Edwards, Ferrari, 2013) and the Indo Pacific Coral Finder (Kelley 2022).

As of January 2025, only taxa in Priority Level 1 have been digitised uniformly across all EcoRRAP plots. Priority 2 & 3 taxa, as well as non-priority taxa, appear in the dataset but have not been consistently digitised across all sites – sporadic colonies have been added because they were collected for genetic sampling in the field or had their ID field-validated by a coral taxonomist.

Four non-coral classes were also digitised on the T_0 orthomosaics with the intention of generating training data for AI classifier development. These polygons were not refined across time points.

Table 34. EcoRRAP species list – shortened to taxa currently present in EcoRRAP dataset. Priority 1 taxa have been digitised uniformly across all sites. Priority 2 & 3 taxa, as well as non-priority taxa, appear in the dataset but have not been uniformly digitised across all sites; sporadic colonies have been added because they were either collected for genetic sampling or had their ID field-validated by a coral taxonomist. Taxa in **bold** within Priority 2, 3, and non-priority levels have been digitised consistently in all sites of the EcoRRAP Central Cluster from 2022 onwards. Taxa with an asterisk (*)

have automatic classifiers built for them (to varying degrees of success, contact EcoRRAP team for details or access to classifiers). See CATAMI classification scheme (Althaus, Hill, Edwards, Ferrari, et al. 2013) for morpho-taxa descriptions. Taxonomic updates follow those on the World Register of Marine Species (WoRMS Editorial Board 2025).

EcoRRAP Priority Level	EcoRRAP TagLab Code	Species/Morpho-taxa name (revised taxonomy in parentheses)	**Notes**
1	Atab*	*Acropora* table	Includes all tabular species
	Ahya	*Acropora hyacinthus*	Labelled as *Acropora* table when unidentifiable to species
	Acyt	*Acropora cytherea*	Labelled as *Acropora* table when unidentifiable to species
	Acor*	*Acropora* corymbose	Includes all corymbose species
	Amil	*Acropora millepora*	Labelled as *Acropora* corymbose when unidentifiable to species
	Aten	*Acropora tenuis* (*Acropora kenti*)	What has been identified in this dataset as *A. tenuis* has been taxonomically revised to *A. kenti* (Bridge et al. 2024) Labelled as *Acropora* corymbose when unidentifiable to species
	Adgt*	*Acropora* digitate	Includes all digitate species
	Ahum	*Acropora humilis*	Labelled as *Acropora* digitate when unidentifiable to species
	Agem	*Acropora gemmifera*	Labelled as *Acropora* digitate when unidentifiable to species
	Pver*	*Pocillopora verrucosa*	
	Pdam*	*Pocillopora damicornis*	
	Spis*	*Stylophora pistillata*	
	Pmas*	*Porites* massive	Species complex that includes *P. lobata, P. lutea, P. australiensis, P. solida, P. murrayensis, P. mayeri, P. myrmidonensis*
	Plat*	*Platygyra*	Includes all *Platygyra* species
	Pdae	*Platygyra daedalea*	Labelled as *Platygyra* when unidentifiable to species
	Gspp*	*Goniastrea sp.*	Includes all *Goniastrea* species
	Gedw	*Goniastrea edwardsi*	Labelled as *Goniastrea sp.* when unidentifiable to species
	Gret	*Goniastrea retiformis*	Labelled as *Goniastrea sp.* when unidentifiable to species
2	**Asta**	***Acropora* staghorn**	Includes all staghorn species Digitised uniformly in EcoRRAP Central cluster
	Dhel	*Diploastrea heliopora*	
	Mtab	*Montipora* table/flat tiers/laminae	Includes all laminate/foliose *Montipora* species
	Pcyl	*Porites cylindrica*	
	Shys	***Seriatopora hystrix***	Digitised uniformly in EcoRRAP Central cluster
3	Adiv	*Acropora divaricata*	
	Alat	*Acropora latistella*	
	Amyr	*Astreopora myriophthalma*	
	Lphr	*Leptoria phrygia*	
	Lhem	*Lobophyllia hemprichii*	Labelled as *Lobophyllia* when unidentifiable to species
	Lobo_sp	***Lobophyllia***	Includes all *Lobophyllia* species

			Digitised uniformly in EcoRRAP Central cluster
	Pspe	*Pachyseris speciosa*	
Non-coral	Calg	Calcareous algae	Crustose coralline red algae
	Rsub	Reefal substrate (settleable substrate)	Substrate suitable for coral settlement (e.g. hard substrate with sparse turf algae)
	Rubl	Rubble	Loose fragments of dead coral, any size (Althaus, Hill, Edwards, Ferrari, et al. 2013)
	Sand	Sand	Coarse sand. Grainy look with fragments of material such as shells (Althaus, Hill, Edwards, Ferrari, et al. 2013)
Non-priority	Aabr	*Acropora abrotanoides*	Non-priority species have not been uniformly digitised across all sites
	Acul	*Acropora aculeus*	
	Aacum	*Acropora acuminata*	
	Aant	*Acropora anthocercis*	
	Aaus	*Acropora austera*	
	Abra	*Acropora branching*	
	Acer	*Acropora cerealis*	
	Acla	*Acropora clathrata*	
	Adig	*Acropora digitifera*	
	Aflo	*Acropora florida*	
	Afor	*Acropora formosa / muricata*	
	Agla	*Acropora glauca*	
	Agran	*Acropora granulosa*	
	Aint	*Acropora intermedia*	
	Alor	*Acropora loripes*	
	Amic	*Acropora microphthalma*	
	Anan	*Acropora nana*	
	Anas	*Acropora nasuta*	
	Asam	*Acropora samoensis*	
	Asar	*Acropora sarmentosa*	
	Asec	*Acropora secale*	
	Asel	*Acropora selago*	
	Asol	*Acropora solitaryensis*	
	Astri	*Acropora striata*	
	Asub	*Acropora subulata*	
	Aval	*Acropora valida*	
	Ayon	*Acropora yongei*	
	Amyr	*Astreopora myriophthalma*	
	Cyp	*Cyphastrea*	
	Favi	*Dipsastraea*	What has been identified in this dataset as *Favia* is considered *Dipsastrea* (Hoeksema and Cairns, 2024b)
	Egem	*Echinopora gemmacea*	
	Elam	*Echinopora lamellosa*	
	Fliz	*Favia lizardensis*	
	Fmat	*Favia matthai*	
	Fpal	*Favia pallida*	

	Fste	*Favia stelligera*	
	Favt	*Favites*	
	Fabd	*Favites abdita*	
	Frot	*Favites rotundata*	
	Gast	*Galaxea astreata*	
	Gfas	*Galaxea fascicularis*	
	Gaus	*Goniastrea australensis*	
	Iso_sp	*Isopora*	
	Lept	*Leptastrea*	
	Lpru	*Leptastrea pruinosa*	
	Lcor	*Lobophyllia corymbosa*	
	Mcur	*Montastrea curta*	What has been identified in this dataset as *Montastrea curta* is considered *Astrea curta* (Huang et al. 2014)
	Menc	***Montipora encrusting***	Digitised uniformly in EcoRRAP Central cluster
	Mver	*Montipora verrucosa*	
	Mele	*Mycedium elephantotus*	
	Oulo	*Oulophyllia*	
	Oxy	*Oxypora*	
	Prug	*Pachyseris rugosa*	
	Plam	*Platygyra lamellina*	
	Psin	*Platygyra sinensis*	
	Pacu	*Pocillopora acuta*	
	Peyd	*Pocillopora eydouxi*	What has been identified in this dataset as *P. eydouxi* has been taxonomically revised to *P. grandis* (Hoeksema and Cairns, 2024c)
	Pmea	*Pocillopora meandrina*	
	Plic	*Porites lichen*	
	Prus	*Porites rus*	
	Psma	*Psammocora*	
	Srad	*Symphyllia radians*	What has been identified in this dataset as *Symphyllia* has been taxonomically revised to *Lobophyllia* (Hoeksema and Cairns, 2024b)
	Srec	*Symphyllia recta*	
	Turb	*Turbinaria*	

5.2.2 EcoRRAP species ID validation workflow

An ID validation process was undertaken at the start of the EcoRRAP program to determine the accuracy of image-based identification of coral taxa. This process, explained below, was one contributing factor for the selection of the EcoRRAP priority taxa.

Orthomosaics generated from the first year of fieldwork were digitised by an experienced benthic ecologist (fully computer-based ID), who initially spent up to 2 hours per plot digitising all common taxa. Digitisation was done by looking at both the orthomosaics and the underlying high resolution (40 MP) images, to identify colonies to the highest taxonomic level possible (i.e. species > morpho-genus > morphology), and labels were only assigned when the benthic ecologist was 80% certain of the class. After all the central inshore and offshore plots were digitised, a list of priority taxa based on abundance was produced, and the benthic ecologist spent a maximum of thirty minutes per plot

to digitise as many coral colonies as possible across all EcoRRAP sites with up to eight individuals per single taxon. The benthic ecologist focused on digitising the most common taxa as per the priority list but also added any new taxa that were abundant. If a taxon was added because it was abundant in the new dataset, all previous plots were reviewed to ensure equal effort per taxa per plot was used across all plots.

Digitised maps (orthomosaic with labeled colonies) were then printed on waterproof paper and taken into the field during the second year of fieldwork along with their associated TagLab files. An experienced coral taxonomist then reviewed the maps underwater, spending approximately 15-20 minutes per plot to locate and either validate or correct the taxon ID attributed during the computer-based digitisation for as many colonies as possible. Any colonies that could not be located because they had disappeared, died, or could not be located due to time constraints were noted as such as such so that all digitised colonies were attributed with a note.

Following the ID validation dive, colonies were reviewed in TagLab and given one of the following notes:

- Validated: taxon ID attributed by benthic ecologist during computer digitisation was correct
- Changed: the taxon ID was changed from what the benthic ecologist had labeled (with original ID retained in separate file)
- Dead
- NA: Not found/missing
- No time: Not reviewed due to time constraints, correct ID unconfirmed

Field ID's validated by the taxonomist were compared with the ID's assigned by the benthic ecologist to determine the level of accuracy for each taxon that had been originally digitised. Along with various other contributing factors (see section 5.2.1 for priority taxa selection criteria) taxa with a higher level of computer-based identification accuracy were selected as EcoRRAP priority species.

Due to various factors, orthomosaics from the first year of fieldwork were re-digitised in new TagLab files to focus on priority taxa. ID validation information (along with genetic sample information, see section 5.2.3) is contained in separate TagLab files as well as in exported shapefiles. ID validation files were reviewed (either in TagLab or in ArcGIS Pro) during the digitising of the new, master files to include a note for any colony that had been field-validated. These colonies are identifiable by a 'v' or 'validated' in the notes column of the dataset.

5.2.3 EcoRRAP genetic sampling workflow

During the 2021-2023 EcoRRAP fieldwork trips, the EcoRRAP genetic adaptation team collected small samples (i.e. nubbins) from coral colonies within or nearby to the EcoRRAP photogrammetry plots. For samples collected from colonies within plots, printed orthomosaics were taken on collection dives to identify the sampled colonies in the orthomosaic. Colonies were then digitised within TagLab projects post-dive with a note ('ng') to identify it as a colony collected for genetic sampling along with its sample ID. These colonies can be identified within the dataset by this note (see Genetics SOP for more details, internal link only. This SOP covers field and office procedures for genetic sampling during a bleaching event but the process is largely the same as for non-bleaching collection).

5.2.4 Check files based on list script (.txt file creation)

This section explains the steps to create a .txt file for EcoRRAP plots that is used in the spatial join process to ensure that all shapefiles have been exported from TagLab and are located within the correct folder. The .txt file can be created using Notepad and must list the names of all files that should be present in the folder (including file extension, e.g. ".shp").

To check EcoRRAP shapefiles:

- Copy and paste the plot list from the annotation log (internal link) into a separate excel file
    - Any year is fine, each one lists all 352 plots
    - Copy the "Metashape File" column (column F)
        - This matches the shapefile name structure
- Change any existing file extensions to ".shp" (highlight column and use Find and Replace function)
    - For example, if using the Metashape File name from the annotation log type ".psx" into the Find section and type ".shp" in the replace section
- Copy file list into a Notepad document
    - Save file in the same folder as shapefiles (for EcoRRAP users this is the Raw shapefile folder)
- Run script (see Table 24 for how to run and script file)
- The names of any missing files will be listed in the python window, otherwise it will state that all files are present

6 REFERENCES

Althaus F., Hill N., Edwards L., Ferrari R., et al. (2013). CATAMI Classification Scheme for scoring marine biota and substrata in underwater imagery – A pictorial guide to the Collaborative and Annotation Tools for Analysis of Marine Imagery and Video (CATAMI) classification scheme. Version 1. http://catami.org/classification; Accessed: December 10, 2024.

Aston, E. A., Duce, S., Hoey, A. S., & Ferrari, R. (2022). A Protocol for Extracting Structural Metrics From 3D Reconstructions of Corals. *Frontiers in Marine Science*, *9*, 854395. https://doi.org/10.3389/fmars.2022.854395

Aston, E.A. (2023). Quantifying and assessing coral reef structural complexity and its role in ecological processes and dyanmics using photogrammetry (PhD Thesis). James Cook University, Townville, Australia. https://doi.org/10.25903/1hw1-a964

Bridge, T. C. L., Cowman, P. F., Quattrini, A. M., Bonito, V. E., Sinniger, F., Harii, S., Head, C. E. I., Hung, J. Y., Halafihi, T., Rongo, T., & Baird, A. H. (2024). A *tenuis* relationship: Traditional taxonomy obscures systematics and biogeography of the ' *Acropora tenuis* ' (Scleractinia: Acroporidae) species complex. *Zoological Journal of the Linnean Society*, *202*(1), zlad062. https://doi.org/10.1093/zoolinnean/zlad062

Esri. (2023). ArcGIS Pro (Version 3.2.1) [Software]. Esri.

Gordon S., Aston E., Lechene M., Harianto J., Bray P., Figueira W., Gonzalez Rivero M., Ferrari Legorreta R. (2023). Field photogrammetry in 4D. Reef Restoration and Adaption Program (EcoRRAP). Standard Operational Procedure Number 14 (1 of series): Overview and in-field workflow. Australian Institute of Marine Science, Townsville. (62 pp) https://doi.org/10.25845/SE7T-PS86

Gordon S., Aston E., Lechene M., Remmers T., Piccaluga A., Brinker E., Cowasji Z., Harianto J., Bray P., Miller G., Watkin B., Figueira W., Gonzalez Rivero M., Ferrari Legorreta R. (2024). Field photogrammetry in 4D: Model Processing. Reef Restoration and Adaption Program (EcoRRAP). Standard Operational Procedure 16 (2 of series): Model processing. Australian Institute of Marine Science, Townsville. (57 pp) http://doi.org/10.25845/mwkv-e70

Hoeksema, B. W.; Cairns, S. (2024a). World List of Scleractinia. *Symphillia* Milne Edwards & Haime, 1848. Accessed through: World Register of Marine Species at: https://www.marinespecies.org/aphia.php?p=taxdetails&id=1597510 on 2024-12-09

Hoeksema, B. W.; Cairns, S. (2024b). World List of Scleractinia. *Dipsastraea* Blainville, 1830. Accessed through: World Register of Marine Species at: https://www.marinespecies.org/aphia.php?p=taxdetails&id=718746 on 2024-12-09

Hoeksema, B. W.; Cairns, S. (2024c). World List of Scleractinia. *Pocillopora eydouxi* Milne Edwards, 1860. Accessed through: World Register of Marine Species at: https://www.marinespecies.org/aphia.php?p=taxdetails&id=206958 on 2024-12-09

Huang, D., Benzoni, F., Fukami, H., Knowlton, N., Smith, N. D., & Budd, A. F. (2014). Taxonomic classification of the reef coral families Merulinidae, Montastraeidae, and Diploastraeidae (Cnidaria:

Anthozoa: Scleractinia). *Zoological Journal of the Linnean Society*, *171*(2), 277–355. https://doi.org/10.1111/zoj.12140

Kelley R. (2022). Indo Pacific Coral Finder, 2022 edition, BYOGUIDE.

Lechene, M. A. A., Figueira, W. F., Murray, N. J., Aston, E. A., Gordon, S. E., & Ferrari, R. (2024). Evaluating error sources to improve precision in the co-registration of underwater 3D models. *Ecological Informatics*, *81*, 102632. https://doi.org/10.1016/j.ecoinf.2024.102632

R Core Team (2023). R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Pavoni, G., Corsini, M., Callieri, M., Fiameni, G., Edwards, C., & Cignoni, P. (2020). On Improving the Training of Models for the Semantic Segmentation of Benthic Communities from Orthographic Imagery. *Remote Sensing*, *12*(18). https://doi.org/10.3390/rs12183106

Pavoni, G., Corsini, M., Ponchio, F., Muntoni, A., & Cignoni, P. (2021). *TagLab: A human-centric AI system for interactive semantic segmentation* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2112.12702

Pavoni, G., Corsini, M., Ponchio, F., Muntoni, A., Edwards, C., Pedersen, N., Sandin, S., & Cignoni, P. (2022). TagLab: AI-assisted annotation for the fast and accurate semantic segmentation of coral reef orthoimages. *Journal of Field Robotics*, *39*(3), 246–262. https://doi.org/10.1002/rob.22049

Veron J.E.N., Stafford-Smith M.G., Turak E. and DeVantier L.M. (2016). Corals of the World. Accessed 01 Dec 2024. https://www.coralsoftheworld.org/species_factsheets/

WoRMS Editorial Board (2025). World Register of Marine Species. Available from https://www.marinespecies.org at VLIZ. Accessed 2025-04-07. doi:10.14284/170.