

Photoquadrat (PQ) endpoint information for data users

Background

Reef Life Survey (RLS) is a non-profit citizen science program in which trained SCUBA divers undertake standardised underwater visual census surveys of reef biodiversity around the world. The survey data collected includes census of fish and macroinvertebrates along 50m transects. At the same time, photographs of the seaweed and coral cover are taken every 2.5m along each transect. The photoquadrats (PQs) for each survey are compiled and managed by RLS and publicly available. For more information on RLS methods please visit <https://reeflifesurvey.com/methods/>

PQ scoring:

The method used for PQ annotation is a point intercept method; whereby the benthos directly below a point overlaid on the image is labelled with the corresponding category from a label scheme. Either the standard RLS label scheme is used, or a compatible alternative that may have nested categories of a higher taxonomic resolution. Where multiple layers of biology are found under a point, the **uppermost** one is to be labelled. Generally, 100 points are scored per survey and the percent cover per survey of each label is calculated on export of the data. Please review the [RLS instructions for PQ analysis in Squidle+](#) for detailed information on the current PQ scoring process.

Databases:

PQ endpoints contain data from 2 sources: the NRMN database and Squidle+. The NRMN database contains PQ data from a range of scoring platforms, whilst Squidle+ is the current platform used - allowing continued QA/QC and review of annotations. Please note the difference between the data sourced from each database in the endpoints:

- Due to the varying sources of data ingested into the NRMN database, this data may not have information in the fields “total_points” and “num_points” however there will always be percent_cover calculated.
- Data sourced from Squidle+ will have metadata used to track annotations in Squidle+ that is not available for the NRMN data. These fields are: sq_annotation_set_id, sq_media_collection_name, annotation_set_owner, sq_datashare_group, label_scheme_lineage.

Important information for data users:

- Dataset_id
 - An important field in the endpoints for data users. This denotes each unique set of PQ scores for an RLS survey – for which the percent cover of the analysis should sum to 100%. This field is necessary as the PQs for a survey may be analysed more than once using different label_schemes for different purposes.
 - Dataset_id is a concatenation of survey_id, the label scheme used, and the annotation_set_id (set to NULL for NRMN data). It is also appended with “_infilled” if the data are infilled (see point on infilled data)
- Survey_id
 - Survey_id matches to the RLS fish and invertebrate data collected on the same transect and all metadata available in the NRMN survey_list endpoints (available at <https://portal.aodn.org.au/search> by searching for NRMN)

- Partially scored surveys
 - The field “Partially_scored_survey” is set to TRUE if the PQs were only assessed for corals – leaving the points overlaid on non-corals blank or unfinished. These data should only be used for coral analyses and should be removed for any assessment of the full benthic community unless they are infilled (see following point on infilled data).
- No Corals Scored
 - This label is present where the survey is **partially scored** but no corals were present under the points scored. It is essentially a place holder so that the data user knows the survey has been reviewed for corals and the percent cover of coral should be set to 0%. There will be a zero present in the fields: num_points, total_points, and percent_cover for these labels.
 - This label is not present if the survey is fully scored but no corals are scored (the placeholder is not needed as the non-coral labels are present in the data), i.e. “No Corals Scored” is not an indication of all surveys with 0% cover coral, just the partially scored ones.
- Infilled data
 - Present in the “_infilled” endpoints only (see Table 1.)
 - If a survey is partially scored in Squidle+ but there are data available from another dataset, the non-coral scores have been ‘infilled’ – i.e., added to the coral scores from the partially scored survey **under the same dataset_id so that the percent cover still sums to 100%**. These can now be treated as fully scored datasets for assessment of the full benthic community (noting the caveats of separate analyses of the same PQs).
 - The dataset_id is appended with “_infilled” for these surveys.
 - Num_points is set to NA for the infill-data (the non-coral data)
 - The database, label scheme, label_scheme lineage, annotation_set_id and media_collection fields will still be true to where the data came from.
- RLS_category
 - Data from all label schemes are mapped to the standard RLS_category scheme for unified analysis. There are 75 of these RLS_category labels in the full data
 - RLS_category labels are appended with “_bleached” if the label is a coral and is bleached (TRUE is present in the “flag_bleached” field)
 - Converted to “Dead coral” if the biota is a coral and “flag_dead” is TRUE.
- RLE_category
 - Used to populate habitat data on the [Reef Life Explorer](#)
 - Includes the categories: coral, dead coral, kelp, substrate, and other sessile benthos, and “No Corals Scored”
 - Can be used to identify **hard** corals by filtering for the RLE_category “Coral”. This includes Scleractinians, Hydrocorals, Octocorals of the genera Heliopora and Tubipora that have a hard skeleton but no other “soft corals”.
- Label
 - For full resolution data endpoints this field will indicate the highest resolution scored. This may be species-level, genus level, morphological, or a combination of taxonomic and morphological (e.g., “Porites massive”) – depending on the label_scheme used. Data users should carefully consider the labels schemes they include in analysis to ensure compatibility.
- Label_scheme_lineage
 - Present for Squidle+ data, showing the lineage of hierarchical label_schemes. Does not represent a taxonomic lineage.

Table 1. PQ endpoint standard formats available

File name	Label resolution	Data included
PQ_FullRes_endpoint	Species-level where available	All data from NRMN and Squidle+ at the original resolution of the analysis (species-resolution in some cases). A field for the mapped RLS_category and RLE_category is also included which can be used to filter out groups of benthos or summarize data, for example. Partially scored surveys (where only corals were marked) are also included.
PQ_FullRes_infilled	Species-level where available	All data from NRMN and Squidle+ at the original resolution of the analysis (species-resolution in some cases). A field for the mapped RLS_category and RLE_category is also included which can be used to filter out groups of benthos or summarize data, for example. Partially scored surveys (where only corals were marked) are also included. Partially scored surveys have been infilled where data is available from another dataset. This increases the number of fully scored surveys that still have high taxonomic resolution coral labels.
PQ_RLS_endpoint	RLS label scheme	All data from NRMN and Squidle+ summarised to RLS categories. Partially scored surveys (where only corals were marked) are also included.
PQ_RLS_infilled	RLS label scheme	All data from NRMN and Squidle+ summarised to RLS categories. Partially scored surveys (where only corals were marked) are also included. Partially scored surveys have been infilled where data is available from another dataset, denoted by the field "infilled". This increases the number of fully scored surveys where the corals have been scored by the same labeller.
PQ_RLE_endpoint	RLE categories	All data from NRMN and Squidle+ summarised to RLE_categories (coral, dead coral, kelp, substrate and other). Partially scored surveys (where only corals were marked) are also included.