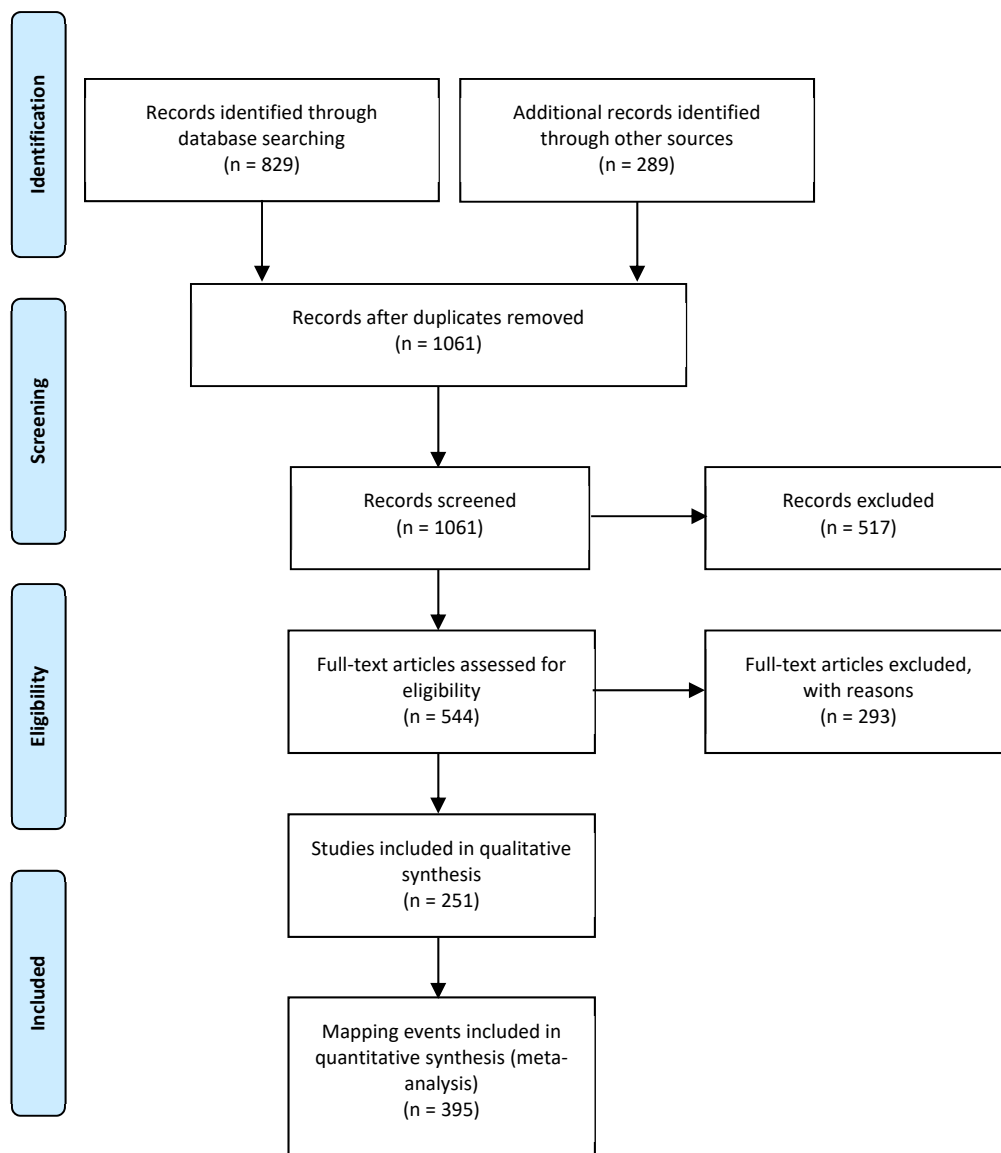# Title: Improving approaches to mapping seagrass within the Great Barrier Reef: from field to spaceborne earth observation.

Len J. McKenzie, Lucas A. Langlois, and Chris M. Roelfsema

Supplementary Material

## S1. Systematic review of seagrass mapping in the GBR

Figure S1. PRISMA [1] flow diagram of the literature selection process for the systematic review of seagrass mapping in the GBR. Note that some studies include multiple mapping events.

Table S2. Description of criteria used to evaluate the confidence in the mapping product for a mapping event. Modified [2]

| Criteria	Description
Maturity of methodology	Shows the confidence that the mapping method/s being used are tested and accepted broadly by the scientific community. Methods must be repeatable and well documented. Apart from methods used before the 1980s (before commercial availability of computer-based GIS tools), maturity of methodology is not a representation of the age of the method but the stage of development.
Validation (observing platforms)	Shows the degree of field validation that has been established and the types of observing platforms used to acquire the data. The reason for this criterion is to show the confidence that field observations/measures are close to a true value and to minimise error propagation.
Representativeness (AOI)	Shows the level of confidence that the field validation data actually represents the whole seagrass area being mapped (Area Of Interest) at the time the observations were made. This includes the level and temporal spread of field validation relative to the Area Of Interest. This criterion takes into consideration the natural spatial and temporal variability embedded in the data as well as the size of the seagrass area being mapped.
Directness (mapping approach)	This criterion looks at the relationship between the earth observing data and the approach being used to construct the final map product. An interpreted/modelled meadow boundary has lower confidence that an actual meadow boundary mapped directly in situ.
Measured mapping error	Incorporates known spatial errors, levels of statistical confidence and any other quantitative measures of accuracy (including correct classification) in the data used to make the map, including meadow boundaries. The confidence is also influenced by the scale and resolution of the map.

S2. Field validation
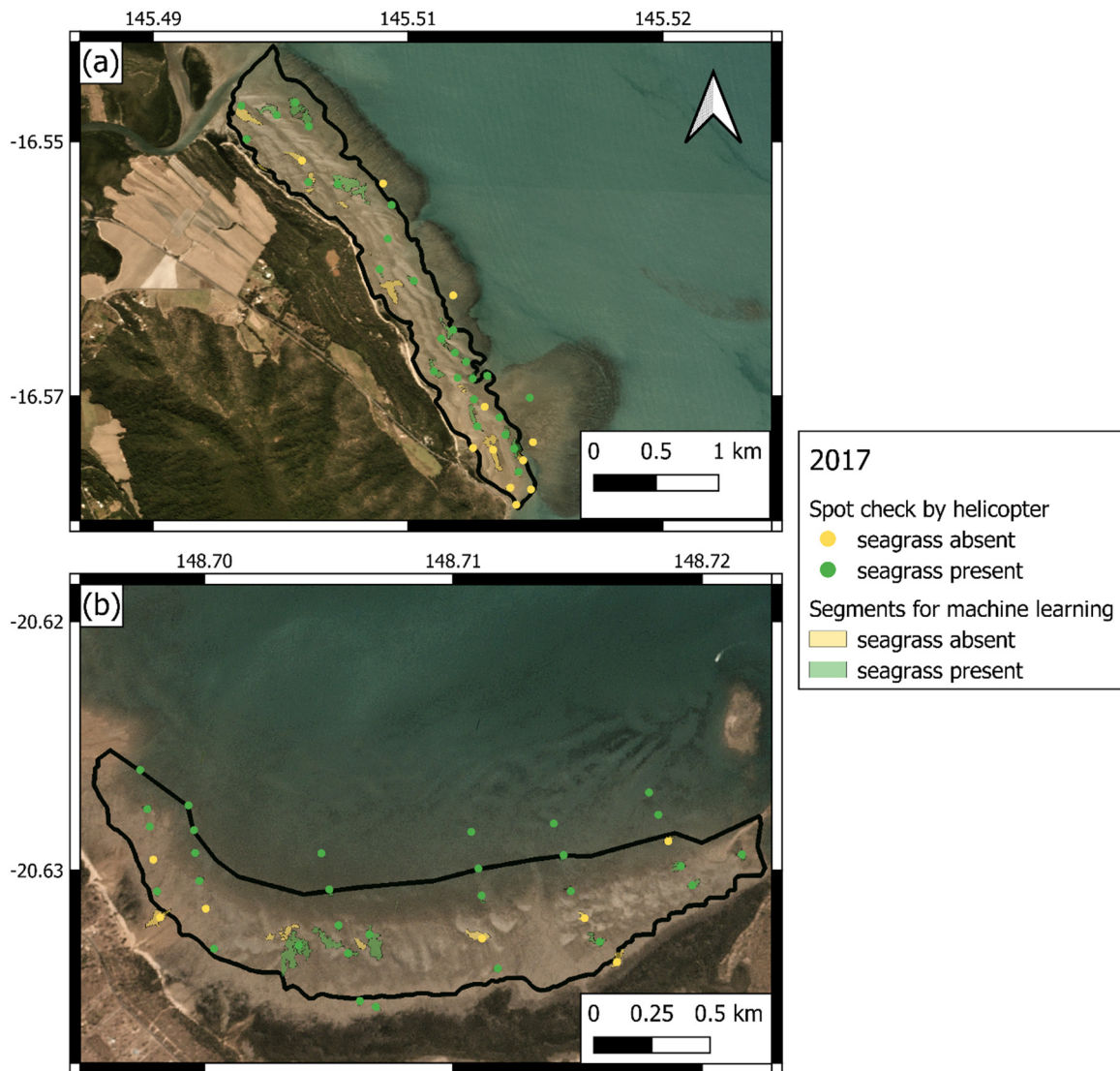
Figure S2. Positions of in situ field validation points (spot–checks) conducted using a helicopter, and derived segments within the coastal AOIs: (a) at clear water habitat (Yule Point) on 05 September 2017 (37 points); (b) and turbid water habitat (Midge Point) on 17 October 2017 (37 points).
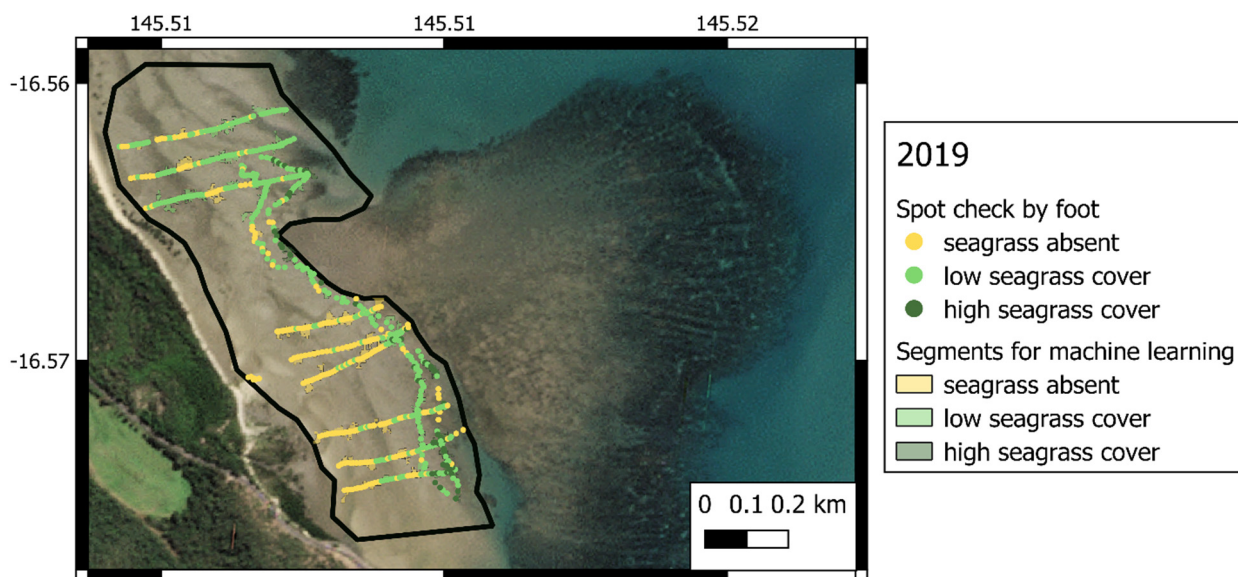
Figure S3. Coastal clear water habitats (Yule Point) AOI showing the positions of 690 in situ field validation points (spot–checks) assessed on foot 14-15 August 2019, and derived segments. Low seagrass cover defined as >0 ≤25% and high seagrass cover as >25%.
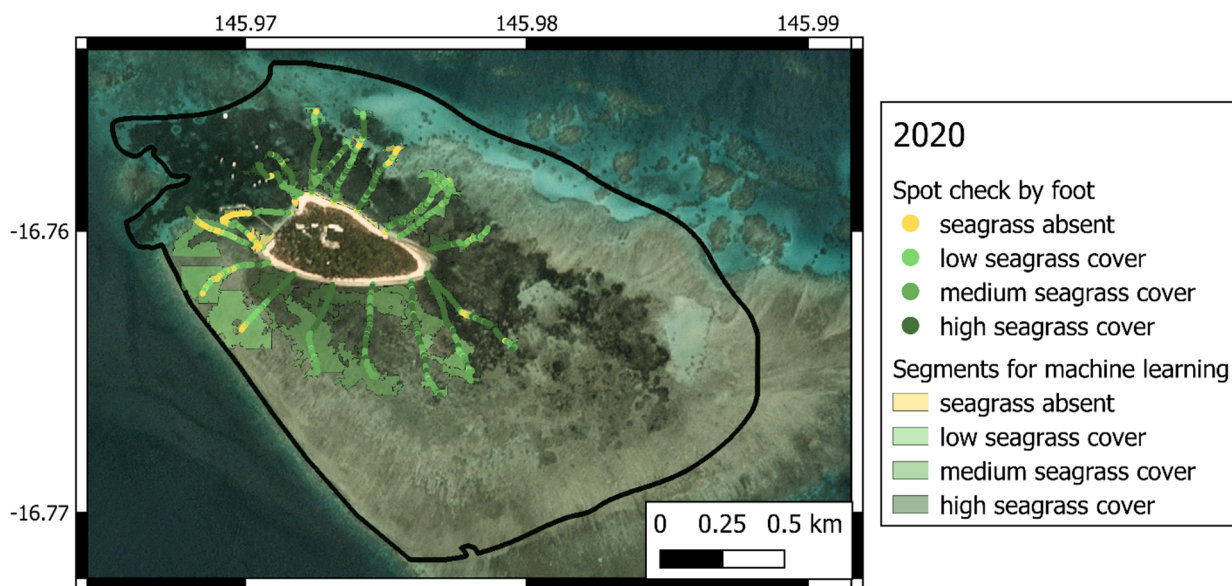
Figure S4. Reef clear water habitat (Green Island) AOI showing positions of 5,415 in situ field validation points (spot–checks) assessed on foot 25-27 November 2020, and derived segments. Low seagrass cover defined as >1 ≤15%, medium as >15 ≤50% and high seagrass cover as >50%.

S3. Machine- and Deep-learning additional method details and results

S3.1. Machine-learning

<u>Additional method details</u>

When examining seagrass abundance from the PlanetScope Dove imagery, we aimed for a balanced repartition across the different classes to cover each AOI as homogeneously as possible, but were constrained by the limited extent of the field validation data. For the coastal turbid water AOI we used only a binary classification of presence/absence of seagrass. For the coastal clear water AOI we used a binary classification for the 2017 map, but for the 2019 map we were able to identify two abundance classes by differentiating between low (<25%) and high (≥25%) seagrass cover, due to a more extensive field validation dataset to match the classes used for the UAV deep learning map (Figure S2). Similarly, at the reef clear water AOI, due the wider range of seagrass cover and much larger number of field validation points (Figure S3), we were able to identify four abundance classes: (1) absence of seagrass; (2) low seagrass cover (1-15%); (3) medium seagrass cover (>15-50%); (4) high seagrass cover (>50%).

Next, rather than using summary variables (e.g. mean of pixel values) to train the model and use them as predictors on the remaining unclassified the segments, we used all the raw pixel values within the classified reference segments. In addition of the red, blue, green and NIR band pixel value, a depth layer was added when appropriate from GBR30 Beaman [3]. NIR was not used in the model for Reef clear water AOI, as all seagrass meadows were submerged at the time of the image capture. This approach had multiple advantages: (1) eliminating trial-and-error to select the best set of summary variables, (2) creating a sufficiently large dataset to run a random forest classifier, while using a small number of unique segments that suitably accounted for heterogeneity, (3) produced a much higher resolution output that was not limited by the minimum size of the segments.

<u>Additional results</u>

Table S3. Summary of Machine-learning model outputs.

AOI Location and Year	Mean % Accuracy ± SE
Yule Point - 2017	98.78±0.0024
Midge Point - 2017	98.69±0.0018
Yule Point - 2019	93.67±0.0009
Green Island – 2020	94.70±0.0004

S3.2. Deep-learning

<u>Additional method details</u>

The orthomosaics from the Coastal clear water (Yule Point, 57332 × 57238 pixels, ~0.80 ha) and Reef clear water (Green Island, 10098 × 7444 pixels, ~4.51 ha) AOIs were decomposed into 256 × 256 pixel tiles to be used in deep-learning. A total of 398 (~1.56 ha, about 29% of the total area) and 510 (~0.013 ha, about 1.6% of the total area) of these tiles for the coastal and reef habitats, respectively, were selected for manually annotated (supervised) in Labelbox for training, testing and validation. Every pixel in each selected tile was assigned a seagrass abundance class. For the coastal habitat, which has a much lower seagrass density and composed of structurally smaller species, the classes used were: (1) Bare sediment, (2) Low cover seagrass (≤ 25%), (3) High cover seagrass (< 25%), (4) Macroalgae/Coral. For the reef habitat the classes used were: (1) absence of seagrass (0%), (2) Low cover seagrass (1-15%), (3) Medium cover seagrass (15-50%), (4) High cover

seagrass (> 50%). The resulting annotations were reasonably even across classes for the coastal habitat (29, 28, 25 and 17%, respectively), but were more unbalanced at the reef habitat (1, 17, 34 and 47%, respectively).

The Segmentation-Model library (https://github.com/qubvel/segmentation_models) in the Keras/TensorFlow frameworks was used in python. This library was a simple way to implement a popular Fully Convolutional Network Model called Unet [4] combined with a backbone model called Resnet34, pre-trained on the ImageNet datatset. The main advantage of the backbone model, used as an encoder of the Unet, is to optimize the feature extraction process for each tiles, making achieving high level accuracy much faster. The models were trained over 50 and 30 epochs (iterations) for the coastal and reef habitats, respectively. Increasing the number of epochs for the training did not result in further model improvement. The loss function used for the models was:

$$total\_loss = dice\_loss + (1 \times focal\_loss)$$

The dice loss was weighted for the reef habitat due to the class imbalance (1, 3, 2.5 and 1 weights for class 1 to 4, respectively).

The accuracy of the models was assessed using the mean Intersection-Over-Union (IOU) score with a threshold of 0.7 and the Dice Coefficient (F1 score). The IoU is defined as the area of overlap between the predicted segmentation and the ground truth divided by the area of union between the predicted segmentation and the ground truth. The F1 score is similar and defined as two times the area of overlap divided by the total number of pixels.

Once the deep-learning semantic segmentation models were trained, they were applied to make predictions over the full UAV orthomosaics and produce the final classified maps. In order to do so, the UAV orthomosaics were decomposed into overlapping tiles of 256 × 256 pixel on a moving window of 64 pixels. This was to smooth out the predictions by avoiding misclassification on the edges. Each pixel, in each tile, received a prediction probability for seagrass abundance by assigning the class with the highest probability. To compose the final prediction raster, we used the MosaicToRaster function in ArcGIS Pro, which applies a mean function over the predictions of overlapping pixels.

Additional results
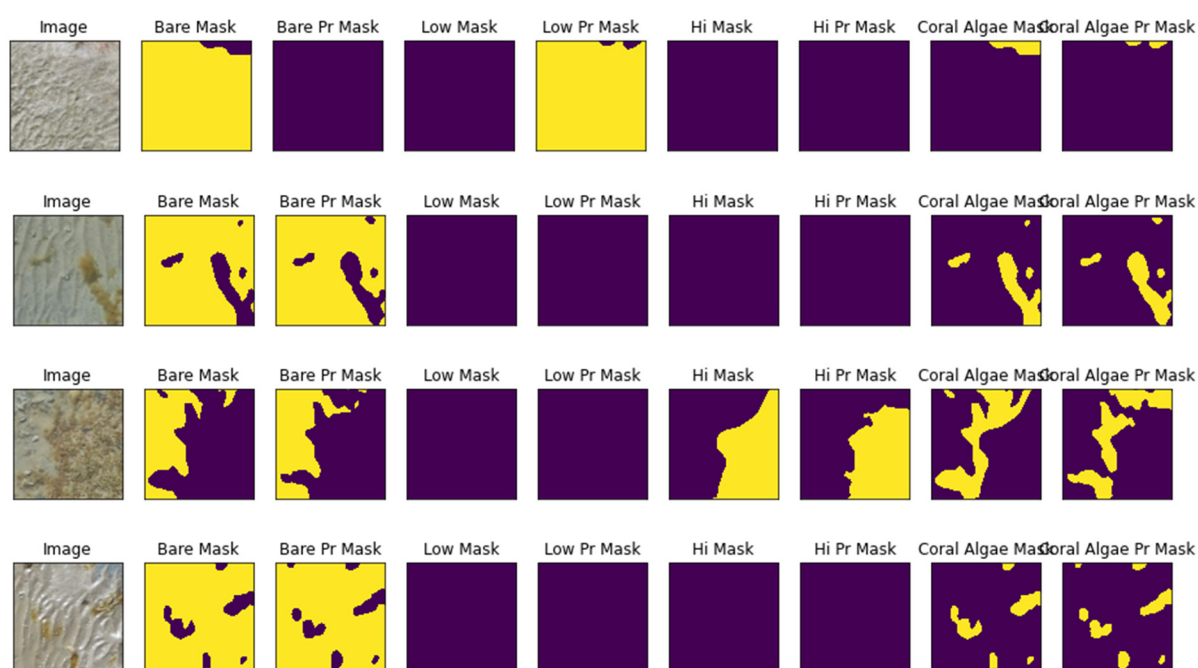
Figure S5. Example of deep-learning output from UAV captured imagery of coastal clear water habitat (Yule Point), 14–15 August 2019. Mask = manually digitized segmentation. PR mask = Segmentation from deep learning model. Masks are coded as 0(violet) – 1(yellow). Overall result: Loss = 0.739; mean IOU score = 0.78085; mean F1-score = 0.80752.
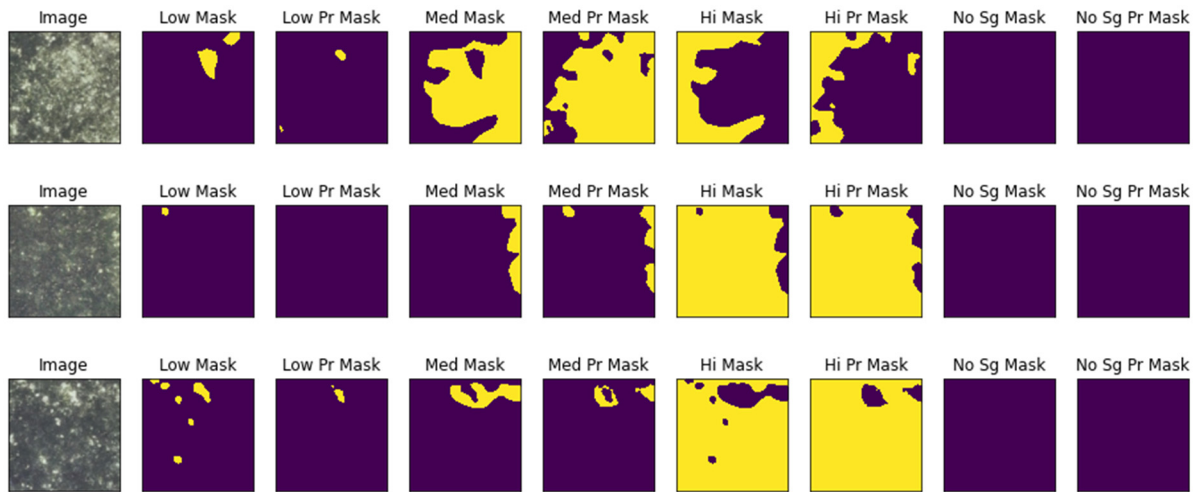
Figure S6. Example of deep-learning output from UAV captured imagery of a reef clear water habitat (Green Island) in 25–27 November 2020. Mask = manually digitized segmentation. PR mask = Segmentation from deep learning model. Masks are coded as 0(violet) – 1(yellow). Overall result: Loss = 0.57621; mean IOU score = 0.73876; mean F1-score = 0.7745.

S4. Applying enhanced mapping approach to historic data collected by traditional approach
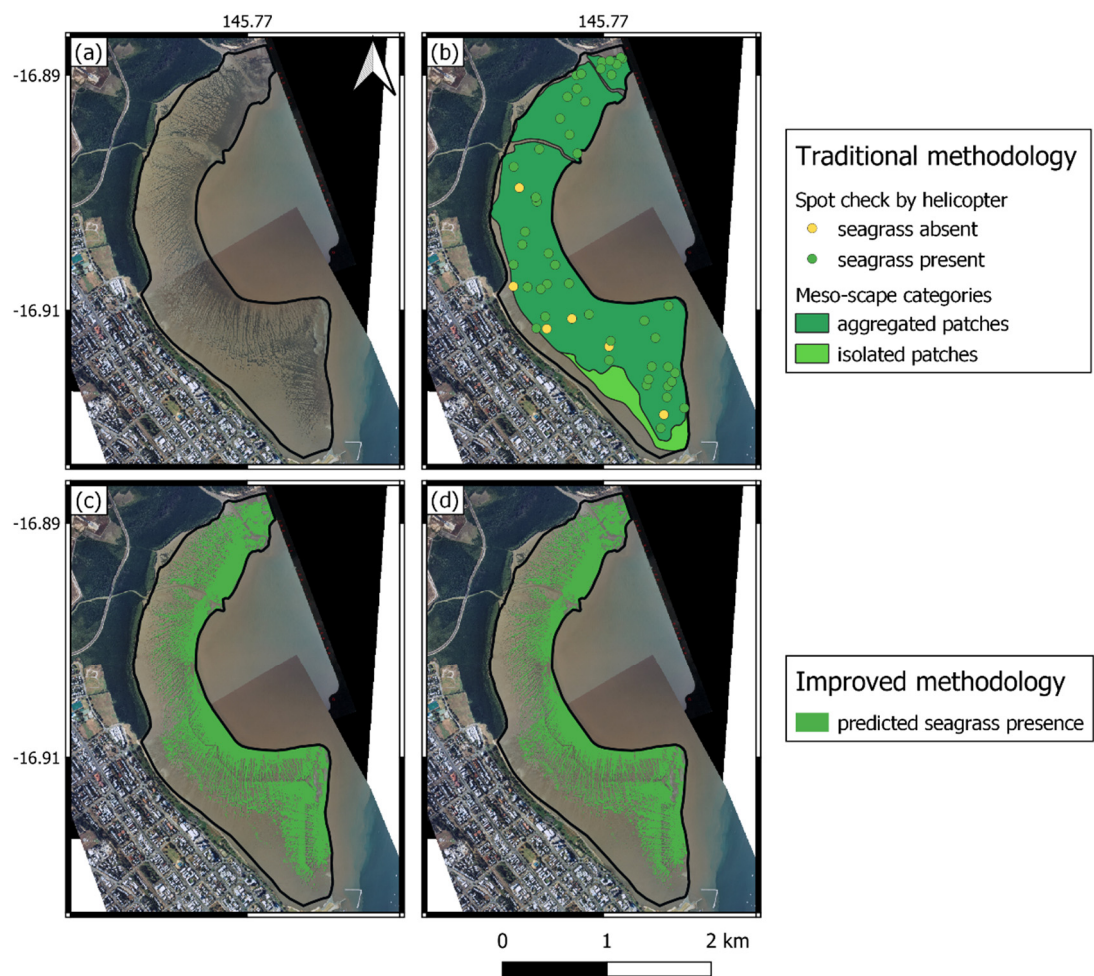
Figure S7. Comparison of mapping approaches at Cairns harbor: (a) the aerial photo of the AOI (black line) from 18 September 2001; (b) map produced from traditional methodology using a helicopter to collect spot–checks (collected 13–14 November 2001), and the meadow boundary drawn by hand on-screen, visually interpolated [5]; (c) map produced with the improved machine learning methodology showing prediction of 60% Bootstrap Probability; (d) and machine learning methodology showing prediction of 100% Bootstrap Probability. Data accessed from [6]

Table S4. Area of seagrass (hectares) mapped within Cairns Harbour AOI using traditional and enhanced approaches to create meso-scale spatially explicit seagrass maps. The range in seagrass area is presented as estimate of reliability for traditional approach, and Bootstrap Probability (BP) for enhanced approach.

Mapping approach	Seagrass area (ha)
Traditional method (±5 m)	279.71 (271.64 to 288.91)
Improved method 60% BP	142.59
Improved method 100% BP	132.92

S4. References

1. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G., Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* **2009**, *339*, b2535. https://doi.org/10.1136/bmj.b2535

2. Waterhouse, J.; Henry, N.; Mitchell, C.; Smith, R.; Thomson, B.; Carruthers, C.; Bennett, J.; Brodie, J.; McCosker, K.; Northey, A., *et al.*, *Paddock to Reef Integrated Monitoring, Modelling and Reporting Program, Program design 2018-2022.* Australian and Queensland, Government: Brisbane, 2018; p 137.

3. Beaman, R.J., High-resolution depth model for the Great Barrier Reef - 30 m. Geoscience Australia: Canberra, 2017. Retrieved from: http://dx.doi.org/10.4225/25/5a207b36022d2

4. Ronneberger, O.; Fischer, P.; Brox, T. In *U-Net: Convolutional networks for biomedical image segmentation*, International Conference on Medical image computing and computer-assisted intervention, 2015; Springer: pp 234-241 (https://arxiv.org/abs/1505.04597).

5. Campbell, S.; Rasheed, M.; Thomas, R., *Seagrass habitat of Cairns Harbour and Trinity Inlet: December 2001. DPI Information Series QI02059*. DPI: Cairns, 2002; p 25. https://bit.ly/3IwyqER

6. Carter, A.; McKenna, S.; Rasheed, M.; Collier, C.; McKenzie, L.; Pitcher, R.; Coles, R. *Seagrass mapping synthesis: A resource for coastal management in the Great Barrier Reef. NESP TWQ project 3.2.1 and 5.4*, ; Centre for Tropical Water & Aquatic Ecosystem Research(TropWATER), James Cook Univ. : Cairns, 2020. doi:10.25909/y1yk-9w85.