Spatially-referenced estimates of seagrass depth of colonization

Marcus W. Beck[1], **James D. Hagy III**[2]

[1]*ORISE Research Participation Program*
USEPA National Health and Environmental Effects Research Laboratory
Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL 32561
Phone: 850-934-2480, Fax: 850-934-2401, Email: beck.marcus@epa.gov

[2]*USEPA National Health and Environmental Effects Research Laboratory*
Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL 32561
Phone: 850-934-2455, Fax: 850-934-2401, Email: hagy.jim@epa.gov

1 Introduction

Issues related to excessive nutrient pollution have motivated a substantial body of research to understand and address impacts on coastal waters. Eutrophication, defined as an increase in the rate of supply of organic matter to an ecosystem (Nixon 1995), is primarily caused by anthropegenic inputs of limiting nutrients that exceed background concentrations of receiving waters. Adverse impacts on aquatic resources are well-documented and have included increased occurrence in the frequency and severity of harmfal algal blooms (Cloern 1996), reduction of dissolved oxygen necessary to support heterotrophic organisms (Justíc et al. 1987, Diaz and Rosenberg 2008), and loss of ecosystem functioning through food web simplification (Tewfik et al. 2007). Although management activities have been successful in mitigating or reversing eutrophication impacts (e.g., Greening and Janicki 2006), the evaluation of critical response endpoints remains an important topic given that ecosystem changes in response to nutrient reduction are not fully understood nor anticipated (Duarte et al. 2009). The most appropriate indicators of ecosystem response are those that exhibit clear biological linkages with water quality changes, such that the potential effects of management actions can be unambiguously characterized through known cause and effect pathways. Critical management decisions may be forced by tentative assessments, political motivations, or qualitative criteria in the absence of empirical methods to identify adequate indicators of ecosytem response (Duarte et al. 2009).

Seagrasses have been used for decades as eutrophication indicators such that their presence, coverage, and diversity are often used as biological endpoints for assessment. Seagrasses are ‘ecosystem engineers’ (Jones et al. 1994, Koch 2001) that serve a structural and functional role in altering aquatic habitat often through complex feedback mechanisms with other ecosystem components. For example, seagrass beds create habitat for juvenile fish and crabs by reducing wave action and stabilizing sediment (). Seagrasses also respond to changes in water clarity through direct physiological linkages with light availability. In short, light mechanism.. Consequently, healthy seagrass communities

Seagrass related to habitat quality and strongly affected by water clarity

restoration targets, load limits

Extensive datasets describing historical and current seagrass growth patterns and distribution in Florida estuaries

No consistent approach for estimating depth of colonization (DoC) to establish restoration targets {acro:doc}

WBID has been considered appropriate management unit although considerable spatial heterogeneity in seagrass growth Reproducible and empirical approaches can be developed that leverage multiple types of information to provide more consistent estimates for restoration targets or nutrient criteria

Objectives

Use information-rich datasets to estimate seagrass DoC by incorporating spatially referenced information

Provide a basis for using these estimates to inform nutrient criteria development using empirical relationships with water clarity

2 Methods

Development of a spatially-referenced approach to estimate seagrass DoC relied extensively on data and partially on methods described in Hagy, In review. The following is a summary of locations and data sources, methods in Hagy, In review, methods and rationale developed to incorporate spatial information in seagrass DoC, and evaluation of the approach including relationships with water clarity.

2.1 Locations and data sources

Four unique locations were chosen for the analysis: Choctowatchee Bay (Panhandle), Big Bend region (northeast Gulf of Mexico), Tampa Bay (central Gulf Coast of Florida), and Indian River Lagoon (east coast) (). These locations were chosen to represent the different geographic regions in the state, in addition to data availability and observed gradients in water clarity that likely contributed to hetereogeneity in seagrass growth patterns. For example, the Big Bend region was chosen to an outflow of the Steinhatchee River where higher concentrations of dissolved organic matter are observed. Seagrasses near the outflow were observed to grow at shallower depths as compared to locations far from the river source. Coastal regions and estuaries in Florida are divided into individual spatial units based on a segmentation scheme developed by US Environmental Protection Agency (EPA) for the development of numeric nutrient criteria. {acro:EPA
One segment from each geographic location was used for the analysis to evaluate estimates of seagrass DoC. The segments included numbers 0303 (Choctowatchee Bay), 0820 (Big Bend region), 0902 (Tampa Bay), and 1502 (Indian River Lagoon), where the first two digits indicate the estuary and the last two digits indicate the segment within the estuary.

Data used to estimate seagrass DoC included a suite of publically available Geographic {acro:GIS
Information System (GIS) products. At the most generic level, spatially-referenced information describing seagrass aerial coverage combined with co-located bathymetric depth information were used to estimate DoC. These data products are available in coastal regions of Florida

through the US Geological Survey, Florida Department of Environmental Protection, and watershed management districts. Data are generally more available in larger estuaries that are of specific management concern, e.g., Tampa Bay, Indian River Lagoon. For example, seagrass coverage data are available from 1950 (Tampa Bay) to present day (multiple estuaries), with more recent products available at annual or biennial intervals. Seagrass coverage maps are less frequent in areas with lower population densities (e.g., Big Bend region) or where seagrass is naturally absent (northeast Florida). Seagrass maps were produced using photo-interpretations of aerial images to categorize coverage as absent, discontinuous (patchy), or continuous. For this analysis, we considered seagrass coverage as being only present (continuous and patchy) or absent since the former did not represent unequivocal categories between regions.

Seagrass coverage maps were combined with bathymetric depth layers to characterize location and depth of growth in each location. Bathymetric depth layers for each location were obtained from the National Oceanic and Atmospheric Administration's (NOAA) National Geophysical Data Center as either Digital Elevation Models (DEMs) or raw sounding data from {acro:DEM
hydroacoustic surveys. Tampa Bay data provided by the Tampa Bay National Estuary Program are described in Tyler et al. (2007). Bathymetic data for the Indian River Lagoon were obtained from the St. John's Water Management District (Coastal Planning and Engineering 1997). NOAA products were referenced to mean lower low water, whereas Tampa Bay data were referenced to the North American Vertical Datum of 1988 and the Indian River Lagoon data were referenced to mean sea level. Depth layers were combined with seagrass coverage layers using standard union techniques of raster and vector layers in ArcMap 10.1 (Environmental Systems Research Institute 2012). To reduce computation time, depth layers were first masked using a 1 km buffer of the seagrass coverage layer. The final layer used for analysis was a point layer with attributes

describing location (latitude, longitude, segment), depth (m), and seagrass (present, absent). Additional details describing the data are available in Hagy, In review.

2.2 Segment-based estimates of seagrass depth of colonization

Methods in Hagy, In review describe an approach for estimating seagrass DoC at individual coastal segments. Specifically, the combined seagrass depth data described above are used to estimate maximum (Z_{cMax}) and median ($Z_{c50\%}$) seagrass DoC, where the maximum depth is defined as the deepest depth at which a "significant" coverage of seagrasses occured in a segment and the median depht is defined as the median depth occurring at the deep water edge. The seagrass depth points are grouped into bins and the proportion of points within each depth bin that contain seagrass are quantified. Both seagrass DoC estimates are obtained from the plot of proportion of points occupied at each depth bin. In general, the plot is characterized by a decreasing trend such that the proportion of occupied points by depth bin decreases and eventually flattens with increasing depth. A regression is fit on this descending portion of the curve such that the intercept point on the x-axis is considered the maximum depth of colonization. The median portion of this curve is considered the median depth of the deepwater edge of seagrass.

Considerable spatial heterogeneity in the observed seagrass growth patterns suggests that a segment-wide estimate of seagrass DoC may be inappropriate, particularly for the examples in the current analysis. Fig. 1 illustrates spatial variation in seagrass distribution for a location in the Big Bend region of Florida. Using methods in Hagy, In review, the estimate for median seagrass DoC for the segment is over- and under-estimated for different areas of the segment. In particular, DoC is greatly over-estimated at the outflow of the Steinhatchee where high concentrations of dissolved organic matter naturally limit seagrass growth. This example suggests that estimates of

DoC may be needed at finer spatial scales to provide a more robust determination of restoration targets and nutrient criteria.

2.3 Estimating seagrass depth of colonization using spatial information

The approach used to estimate seagrass DoC with spatial information has a similar theoretical foundation as the original, although several key differences should be noted. In general, seagrass DoC estimates are based on empirical measures of the frequency occurrence of seagrass by increasing depth. The first difference is that the maximum DoC is estimated from a logistic growth curve fit through the data, as compared to a simple linear regression in the previous example. Second, a third measure describing the depth at which seagrass were most commonly located, as compared to maximum depth of growth, was defined using these methods. The third and most important difference is that the estimates are specific to discrete locations, using either a grid of points or as a single location of interest. Methods and implications of these differences are described below.

The spatially-referenced approach for estimating DoC begins by creating a grid of evenly-spaced points within the segment. The same process for estimating DoC is used for each point. Alternatively, a single location of interest can be chosen rather than a grid-based design with multiple estimates. Seagrass depth data that occur within a set radius from each grid point are selected (Fig. 2). An estimate of seagrass DoC for each point in the grid is obtained using the sampled seagrass depth points. The seagrass DoC estimate for each grid location is quantified from a plot of the proportion of bathymetric soundings that contain seagrass at each depth bin (Fig. 3a). A radius around a grid point for sampling seagrass depth points that is sufficient to quantify depth of colonization typically has a plot similar to Fig. 3a. The proportion of points that are occupied by seagrass should decrease continuously with increasing depth.

A decreasing logistic growth curve is fit to the sampled seagrass depth points for the grid location to create a monotonic and asymptotic function. This curve is fit using non-linear regression to characterize the reduction in points occupied by seagrass as a function of depth. The logistic growth curve is fit by minimizing the residual sums-of-squares with the Gauss-Newton algorithm (Bates and Chambers 1992) and user-supplied starting parameters that are an approximate estimate of the curve characteristics. The model has the following form:

$$Proportion = \frac{\alpha}{1 + \mathrm{e}^{(\beta - Depth)/\gamma}} \tag{1}$$

where the proportion of points occupied by seagrass at each depth is defined by a logistic curve with an asymptote α, a midpoint inflection β, and a scale parameter γ. Starting values α, β, and γ were estimated empirically from the observed data.

Finally, a simple linear curve is fit through the inflection point (β) of the logistic curve to estimate depth of colonization (Fig. 3c). The inflection point is the depth at which seagrass are decreasing at a maximum rate and is used as the slope of the linear curve. Three measures are obtained from the linear curve. The maximum depth of seagrass colonization, DOC_{max}, is the x-axis intercept of the linear curve. The depth of maximum seagrass occupancy, SG_{max} is the location where the linear curve intercepts the asymptote of the logistic growth curve. The median depth of seagrass colonization, DOC_{med}, is the depth halfway between SG_{max} and DOC_{max}. DOC_{med} was typically but not always the inflection point of the logistic growth curve. The estimation process is repated for each point in the grid.

Estimates for each of the three DoC measures are obtained only if specific criteria are met.

These criteria were implemented as a safety measure that ensures a sufficient amount and appropriate quality of data are used. First, estimates are provided only if a sufficient number of seagrass depth points are present within the radius of the grid point to estimate a logistic growth curve. This criteria applies to the sample size as well as the number of points with seagrass in the sample. That is, the curve cannot be estimated for small samples or if an insufficient number of points contain seagrass regardless of sample size. Second, estimates are provided only if an inflection point is present on the logistic curve within the range of the sampled depth data. This criteria may apply under two scenarios where the curve is estimated but a trend is not adequately described by the sampled data. That is, a curve may be estimated that describes only the initial decrease in points occupied as a function of depth but the observed points do not occur at depths deeper than the predicted inflection point. The opposite scenario may occur where a curve is estimated but only the deeper locations beyond the inflection point are present in the sample. Finally, the estimate for SG_{max} is set to zero if the linear curve through the inflection point intercepts the asympote at x-axis values less than zero. The estimate for DOC_{med} is also shifted to the depth value halfway between SG_{max} and DOC_{max}.

All estimates were obtained using custom-made functions in program R that were based on the `nls` and `SSlogis` functions to fit a nonlinear least squares using a self-starting logistic growth model (Bates and Chambers 1992, R Development Core Team 2014). All seagrass depth shapefiles were imported and processed in R using functions in the `rgeos` and `sp` packages (Bivand et al. 2008, Bivand and Rundel 2014).

2.4 Comparison with segment-based approach and sensitivity analysis

Spatially-referenced estimates for seagrass DoC were obtained for each of the four segments described above. Segment-wide estimates obtained using methods in Hagy, In review were used as a basis of comparison such that departures from these values were evidence of spatial heterogeneity in seagrass growth patterns within each segment. A sampling grid of locations for estimating each of the three depth values in Fig. 3 was created for each segment. This grid is a set of evenly spaced points with a random starting location for the first point. The grid is masked by the segment boundaries to remove locations that did not occur on the water. Initial spacing between sample points was chosen arbitrarily as 0.02 decimal degrees, which is approximately 2 km at 30 degrees N latitude. Similarly, the sampling radius around each sampling location in the grid was chosen as 0.06 decimal degrees, or approximately 6 km.

Three factors influence the estimates at each sampling point, as well as the ability compare values between points. First, the starting location of the first point of the sampling grid is chosen arbitrarily such that a unique grid is obtained comparisons of within-segment estimates may vary slightly given the starting location. Second, the spacing between sampling points affects the degree of collinearity between estimates that are near each other. For a set sampling radius around each point, estimates will be less correlated at larger spacing between sampling points, whereas the converse is true for smaller spacing. Third and most important, the radius around each sampling point determines the number of seagrass depth points that are included in the estimate. The chosen radius is considered an explicit area within which the estimate applies. As before, increasing the radius around each sample point will increase the collinearity between estimates at adjacent points for a set grid spacing. Collinearity between sample points based on the sampling scheme is not inherently problematic provided the results are interpreted in the context of the question of interest. For example, small spacing and large sampling radii will create very similar estimates between points. This approach does not necessarily invalidate the estimate at each point, although comparisons between points become less valid as the estimates are not related to a unique sampling area for each location. Similarly, a grid with large spacing and small radii facilitates comparison between points as each location represents a unique collection of samples, although each estimate is relevant for a small location with undescribed and potentially important variation in seagrass growth patterns between points.

A systematic approach was used to evaluate validity of comparisons between sampling points given parameters that influence collinearity or spatial autocorrelation. For the analysis, 'validity' is considered relative uniqueness of estimates at each point in the context of grid spacing and sampling radius. The effect of the random starting location of each grid was considered negligible for this analysis and set constant between analyses for comparison. For each segment, unique combinations of grid spacing and sampling radii were used to estimate maximum seagrass DoC. Spatial autocorrelation between all estimates was measured by semivariance

2.5 Developing a spatially coherent relationship of water clarity with depth of colonization

3 Results

4 Discussion

References

Bates DM, Chambers JM. 1992. Nonlinear models. In: Chambers JM, Hastie TJ, editors, Statistical Models in S, pages 421–454. Wadsworth and Brooks/Cole, Pacific Grove, California.

Bivand R, Rundel C. 2014. rgeos: Interface to Geometry Engine - Open Source (GEOS). R package version 0.3-8.

Bivand RS, Pebesma EJ, Gómez-Rubio V. 2008. Applied Spatial Data Analysis with R. Springer, New York, New York.

Cloern JE. 1996. Phytoplankton bloom dynamics in coastal ecosystems: A review with some general lessons from sustained investigation of San Francisco Bay, California. Review of Geophysics, 34(2):127–168.

Coastal Planning and Engineering. 1997. Indian River Lagoon bathymetric survey. A final report to St. John's River Water Management District. Technical Report Contract 95W142, Coastal Planning and Engineering, Palatka, Florida.

Diaz RJ, Rosenberg R. 2008. Spreading dead zones and consequences for marine ecosystems. Science, 321:926–929.

Duarte CM, Conley DJ, Carstensen J, Sánchez-Camacho M. 2009. Return to *Neverland*: Shifting baseline affect eutrophication restoration targets. Estuaries and Coasts, 32(1):29–36.

Environmental Systems Research Institute. 2012. ArcGIS v10.1. ESRI, Redlands, California.

Greening H, Janicki A. 2006. Toward reversal of eutrophic conditions in a subtrophical estuary: Water quality and seagrass response to nitrogen loading reductions in Tampa Bay, Florida, USA. Environmental Management, 38(2):163–178.

Hagy JD. In review. Seagrass depth of colonization in Florida estuaries.

Jones CG, Lawton JH, Shachak M. 1994. Organisms as ecosystem engineers. OIKOS, 69(3):373–386.

Justíc D, Legovíc T, Rottini-Sandrini L. 1987. Trends in oxygen content 1911–1984 and occurrence of benthic mortality in the northern Adriatic Sea. Estuarine, Coastal and Shelf Science, 25(4):435–445.

Koch EW. 2001. Beyond light: Physical, geological, and geochemical parameters as possible submersed aquatic vegetation habitat requirements. Estuaries, 24(1):1–17.

Nixon SW. 1995. Coastal marine eutrophication: A definition, social causes, and future concerns. Ophelia, 41:199–219.

R Development Core Team. 2014. R: A language and environment for statistical computing, v3.1.2. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Tewfik A, Rasmussen JB, McCann KS. 2007. Simplification of seagrass food webs across a gradient of nutrient enrichment. Canadian Journal of Fisheries and Aquatic Sciences, 64(7):956–967.

Tyler D, Zawada DG, Nayegandhi A, Brock JC, Crane MP, Yates KK, Smith KEL. 2007. Topobathymetric data for Tampa Bay, Florida. Technical Report Open-File Report 2007-1051 (revised), US Geological Survey, US Department of the Interior, St. Petersburg, Florida.

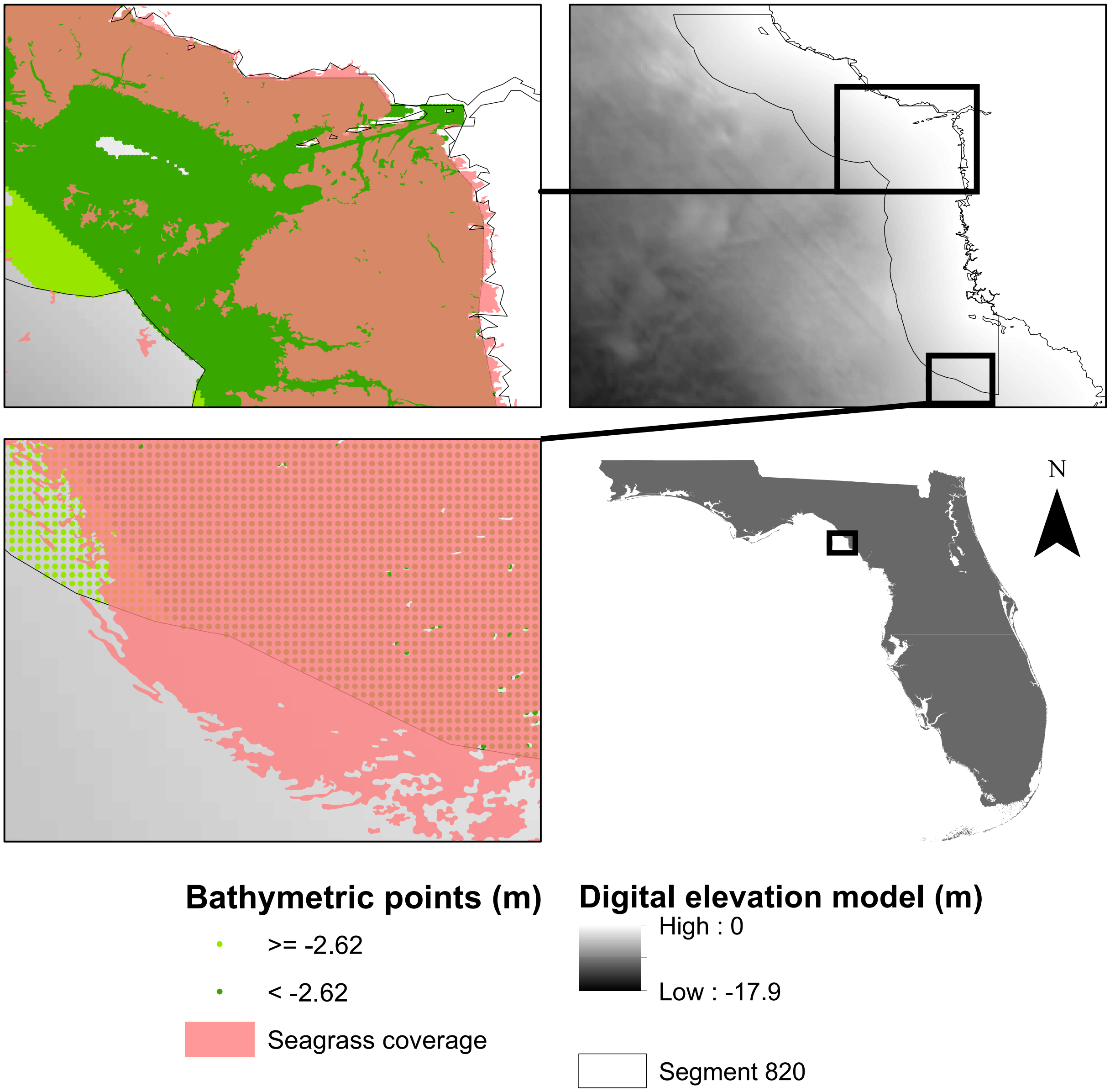


Fig. 1: Example of over- and under-estimates for seagrass depth of colonization for a segment in the Big Bend region, Florida. The top-left figure indicates over-estimation and the bottom-left indicates under-estimation. Bathymetric points are color-coded by the median depth of colonization estimate for continuous seagrass in the segment.

{fig:wbid_

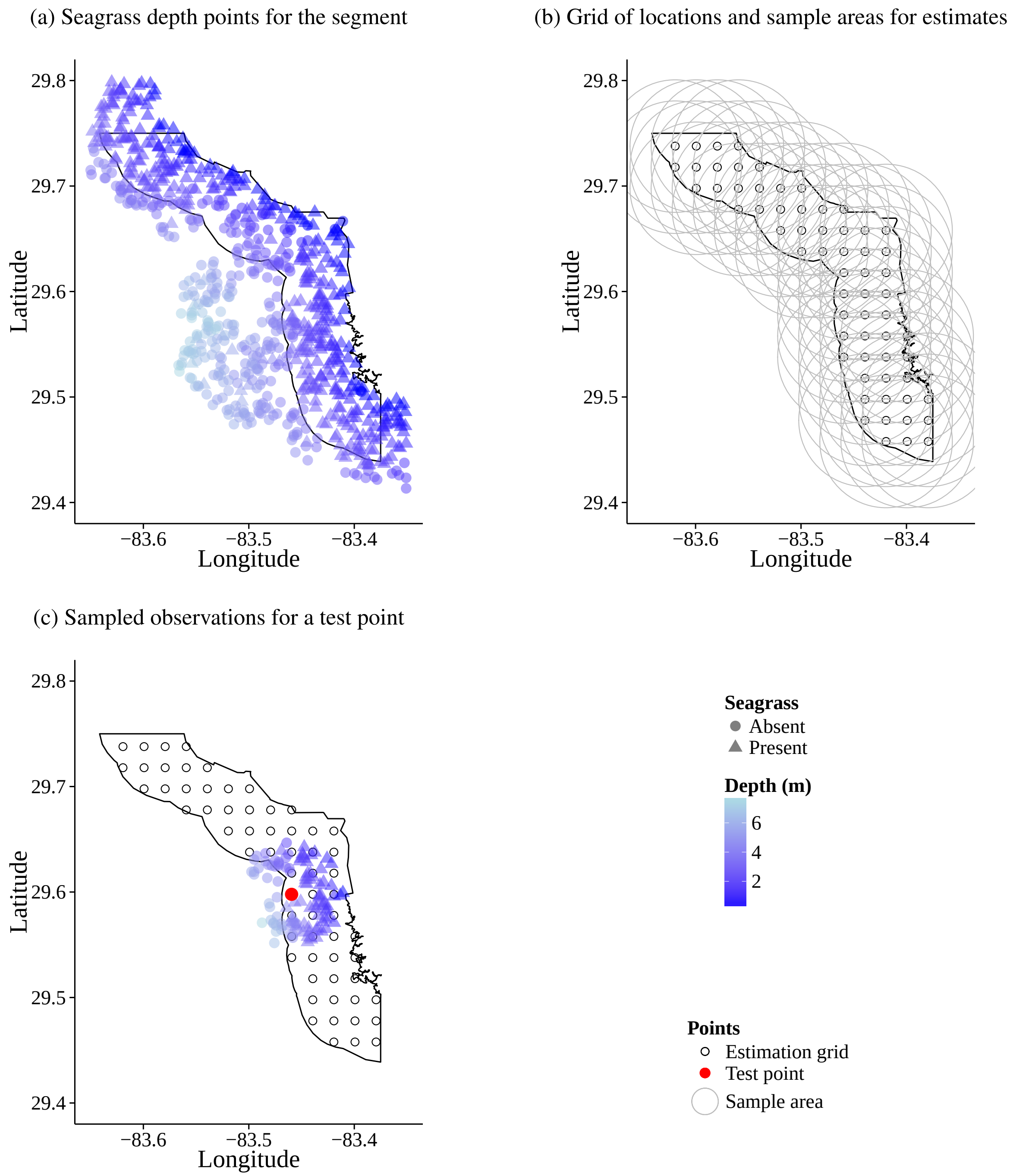
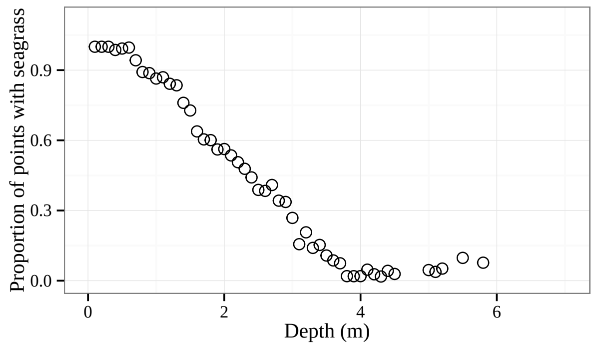
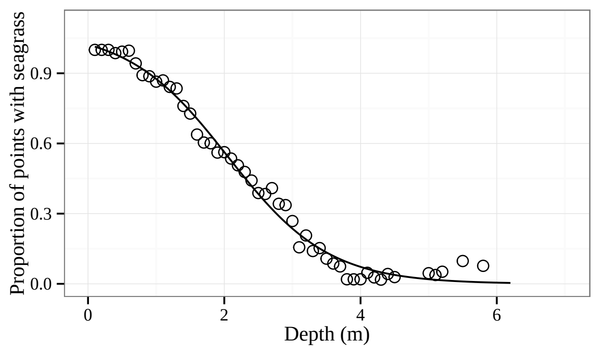


Fig. 2: Examples of data and grid locations for estimating seagrass depth of colonization for a region of the Big Bend, Florida. Fig. 2a shows the seagrass depth points that are used for sampling, Fig. 2b shows a grid of locations and sampling radii for estimating seagrass DoC, and Fig. 2c shows an example of sampled seagrass depth points for a location. Estimates in Fig. 3 were obtained from the sampled location in Fig. 2c. {fig:buff_

(a) Proportion of points with seagrass by depth



(b) Logistic growth curve fit through points



(c) Depth estimates

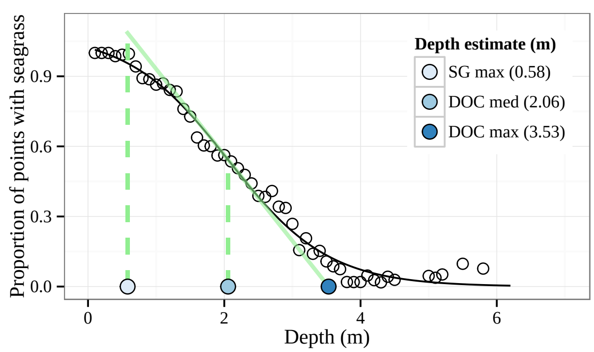


Fig. 3: Methods for estimating seagrass depth of colonization using sampled seagrass depth points around a single location. Fig. 3a is the proportion of points with seagrass by depth using depth points within the buffer of the test point in Fig. 2. Fig. 3b adds a decreasing logistic growth curve fit through the points. Fig. 3c shows three depth estimates based on a linear curve fit through the inflection point of logistic growth curve. {fig:est_e