Effects of Pixel Resolution, Mapping Window Size, and Spectral Species Classification on Remote Sensing of Plant Beta Diversity Using biodivMapR and Hyperspectral Imagery

Kevin M. Robertson[1], Eli Simonson[1], Natali Ramirez-Bullon[1], Benjamin Poulter[2], and Richard Carter[3]

[1]Tall Timbers Research Station, Tallahassee, FL, USA

[2]NASA Goddard Space Flight Center, Biospheric Sciences Lab., Greenbelt, MD, USA

[3]Valdosta State University, Valdosta, GA, USA

Corresponding author: Kevin Robertson (krobertson@talltimbers.org)

Key Points:

- Ability to map plant beta diversity on landscapes using biodivMapR and similar algorithms depends on having a sufficient number of pixels per mapping window.
- Increasing mapping window size to accommodate sufficient numbers of pixels per window decreases spatial resolution of beta diversity maps.
- Assigning the appropriate number of spectral species is important for generating a dissimilarity matrix that appropriately reflects actual plant beta diversity.

Abstract

Using imaging spectroscopy (hyperspectral imaging), we sought to assess the effects of image pixel resolution, size of mapping windows composed of pixels, and number of spectral species assigned to pixels on the capacity to map plant beta diversity using the biodivMapR algorithm, in support of the planned NASA Surface Biology and Geology (SBG) satellite remote sensing mission. BiodivMapR classifies pixels as spectral species, then calculates beta diversity as dissimilarity of spectral species among mapping windows each composed of multiple pixels. We used NEON airborne 1 m resolution hyperspectral images collected at three sites representing native longleaf pine ecosystems in the southeastern U.S. and aggregated pixels to sizes ranging from 1-90 m for comparative analyses. Plant community composition was groundtruthed. Results show that the capacity to detect plant beta diversity decreases with fewer pixels per mapping window, such that pixel resolution limits the size of mapping windows effective for representing beta diversity. Mapping window size in turn limits the spatial resolution of beta diversity maps composed of mapping windows. Assigning too few pixels per window, as well as assigning too many spectral species per image, results in overestimation of dissimilarity among locations that have plant species in common. This overestimation undermines the capacity to contrast mapping window dissimilarity within versus among community types and reduces the information content of beta diversity maps. These results demonstrate the advantage of maximizing spatial resolution of hyperspectral imaging instruments on the anticipated NASA SBG satellite mission and similar remote sensing projects.

Plain Language Summary

Mapping beta diversity, or the differences in species composition among different parts of a landscape, is an important goal of satellite remote sensing. Different remote sensing products have different sizes of pixels that make up the image, which can effect how much information the image displays. NASA is interested in knowing the effects of image pixel size for designing future satellite missions, including the upcoming Surface Biology and Geology (SBG) mission. We used remote sensing data taken from aircraft at NEON research sites in the southeastern U.S. to test how different pixel sizes, ranging from 1-90 m, affect how well beta diversity can be mapped. We used the program biodivMapR, which creates maps from square "mapping windows", which each contain multiple pixels that are classified according to their reflectance data, and which are used to tell how different one mapping window is from the other. We found that the larger the pixel size, the larger the mapping window has to be to have enough pixels to map biodiversity well. The tradeoff is that larger mapping windows result in coarser biodiversity maps. Using 30-45 m pixels, which has been recommended for the SBG mission, relatively large areas covered by different natural community types can be distinguished, but smaller features like isolated ponds and narrow streams may not be detected. The study shows the importance of having image pixel sizes that are as small as possible, in addition to having high quality information per pixel.

1 Introduction

The impact of human activities on the function and and sustainability of earth's biological and physical systems places high priority on tracking global patterns of biodiversity and ecosystem change. Given the high cost and limited geographic distribution of field plots for ecosystem monitoring, remote sensing will play an increasingly important role in systematically monitoring trends in biodiversity and ecosystem health. Remote sensing approaches depend on demonstrated links between field-observed data and remotely sensed reflectance data in order to develop automated systems for ecological interpretation of imagery covering wide geographic areas (Pereira et al., 2017). The increasing availability of imaging spectroscopy, hereafter hyperspectral imagery, in which hundreds of reflectance wavelengths are measured for each pixel, promises a paradigm shift in the capacity to remotely monitor biodiversity (Schimel et al., 2020). However, current application of hyperspectral technology remains limited by spatial and temporal coverage (Cawse-Nicholson et al., 2021).

In light of these needs, the 2017-2027 Decadal Survey organized by the United States National Academy of Sciences, Engineering, and Medicine (NASEM, 2018) established surface biology and geology (SBG) as a designated observable using satellite remote sensing. The SBG project is anticipated to use hyperspectral visible to shortwave infrared (VSWIR; 380-2500 nm) imagery in a mission to be lead by the National Air and Space Administration (NASA) (Cawse-Nicholson et al., 2021). To date, the Decadal Survey recommends a VSWIR instrument with 30-45 m pixel resolution, as well as ≤16 day global revisit time and 10 nm spectral resolution in the 380-2500 nm range. During the current formulation phase, it is important to consider tradeoffs among choices of parameter specifications. One important parameter under consideration is spatial resolution of imagery, which may have a significant influence on the capability to

remotely sense spatial distributions of biodiversity (Gamon et al., 2020).

The current study focuses on implications of spatial resolution of imagery on effectiveness of remotely sensing plant community beta diversity using the algorithm biodivMapR (Féret and de Boissieu, 2020). Beta diversity refers to spatial variation, or turnover, of species composition among plant communities at the landscape scale. BiodivMapR also estimates alpha plant diversity (local diversity) as diversity of spectral species (classified pixels) within larger mapping windows, which depends on the Spectral Variation Hypothesis (Palmer et al., 2002; Rocchini et al., 2004), or the assumption that local variation in spectral signatures among pixels corresponds to plant functional diversity and species diversity (Gamon et al., 2020). However, as the size of pixels increasingly exceeds the size of individual plant species, this assumption becomes much less certain (Féret and Asner, 2014; Féret and de Boissieu, 2020; Rocchini et al., 2018). For example, in grasslands where there multiple species per m^2, this relationship appears to break down at pixel sizes larger than about 5 m (Gholidezah et al. 2019; 2021; Gamon et al., 2020). Within the range of spatial resolutions suggested for the SBG mission (30-45 m), alpha diversity procedures will likely not be effective in herb-dominated plant communities such as those in the current study, which can have > 20 species per m^2 (Glitzenstein et al., 2003). Remote sensing of beta diversity is less limited by spatial resolution, as pixels classified as spectral species can represent local community composition instead of individual species (Rocchini et al., 2018). Thus, the degree of similarity in biodiversity between two areas of interest can be estimated by the similarity in spectral species composition. BiodivMapR uses the spectral species concept (Rocchini et al., 2010) to calculate dissimilarity in spectral species among larger mapping windows containing multiple pixels. In this light, biodivMapR might be effectively applied for detection of plant beta diversity using ranges of image spatial resolution

recommended for the SBG mission and currently available from existing hyperspectral satellite missions (e.g., PRISMA, DESIS).

The capacity to remotely sense beta diversity at spatial resolutions larger than individual plants depends on the plant community concept, specifically that plant species belonging to particular assemblages are adapted to certain environmental conditions (Lortie et al., 2004) and collectively have definable reflectance characteristics (Cavender-Bares et al., 2020). Ability to map beta diversity also depends on the pixel resolution of available imagery relative to the spatial scale of plant community turnover (Gamon et al. 2020). Landscapes within the study region, the southeastern U.S. Coastal Plain, provide a useful scenario for assessing methods of detecting beta diversity, given their complex spatial arrangements at small spatial scales (Carr et al. 2010), geologically active karst topography that influences community distribution (Lane and D'Amico, 2010), strong responses of vegetation to slight elevation changes (Drewa et al. 2002), and varying coverage and effects of frequent prescribed fire (Robertson et al. 2019), which contribute to the region being recognized as a global biodiversity hotspot (Noss et al., 2015). Thus, the ability to differentiate communities for mapping beta diversity will depend on both image pixel size and dimensions of the mapping windows composed of pixels classified as spectral species, among which dissimilarity of spectral species is calculated for mapping beta diversity (Asner and de Boissieu, 2020). Larger mapping windows contain more pixels and spectral species and thus have more refined capacity to estimate dissimilarity, but at the cost of spatial resolution of the beta diversity map. Effective mapping of beta diversity may also be influenced by the specified number of spectral species into which pixels are classified (Féret and de Boissieu, 2020).

In this study, we use biodivMapR to compare estimates of beta diversity among levels of

image resolution ranging in pixel size from 1 m to 90 m. Our approach was to use 1 m resolution imagery from airborne sensors at three National Ecological Observatory Network (NEON; neonscience.org) sites representing different natural landscapes representative North American Coastal Plain pine communities. We aggregated pixels to simulate coarser resolution imagery and ran biodivMapR algorithms to assess its capacity to detect beta diversity using different pixel resolutions using nearly constant sized mapping windows. We also explored the effects of mapping window size and the number of assigned spectral species on the ability to distinguish natural communities. We used field-collected data to confirm the similarity within and dissimilarity among plant communities with regard to actual plant species composition and to provide points of reference to estimates by biodivMapR. We use the results to discuss the implications of spatial resolution requirements of imagery for measuring biodiversity on natural landscapes using space-based hyperspectral imagery and indicate potential applications and limitations of such imagery and to provide guidance for development of the anticipated SBG mission.

2 Materials and Methods

2.1 Study Sites

We used remote sensing and field-collected plant presence and percent cover data provided by NEON (2022) from sites at three properties representing different longleaf pine (*Pinus palustris*) savanna ecosystems within the southeastern U.S. Coastal Plain. The properties were the Disney Wilderness Preserve (DSNY), Jones Ecological Research Center (JERC), and Ordway-Swisher Biological Station (OSBS) (Fig. 1). The sites were selected for the availability of NEON airborne hyperspectral imagery with 1 m resolution acquired annually within a 10 km x 10 km area centered on the property and the availability of surface vegetation data. Also, the

sites represent the three historically dominant upland ecosystems in the Coastal Plain region and their associated community types, described below for each site.

2.1.1 Disney Wilderness Preserve (DSNY)

DSNY (28.1287°N, -81.4303°W) is a 4,600 ha property owned and managed by The Nature Conservancy. Elevation of the property is approximately 15 m above sea level (asl), and low to high average monthly temperatures are 15.6°C in January and 27.8°C in July. Native pine savannas dominating the DSNY landscape are more specifically flatwoods pine savannas (FNAI, 2010), historically the most widespread natural community in Florida. Soils are mostly in the Spodosol order (Soil Survey Staff 2022), which are sandy with organic matter and saturated part of the year. The flatwoods pine community has sparse tree canopy dominated by longleaf pine with an understory of evergreen shrubs, especially saw palmetto (*Serenoa repens*) and gallberry (*Ilex glabra*), and a diverse herbaceous plant community dominated by wiregrass (*Aristida beyrichiana*) in most areas. It is dependent on frequent prescribed fire (Glitzenstein et al., 2003), applied at three-year or shorter intervals at DSNY. Cypress forests (syn. dome swamp, basin swamp; FNAI, 2010) occur at slightly lower elevations with long hydroperiods and a nearly closed canopy of pond cypress

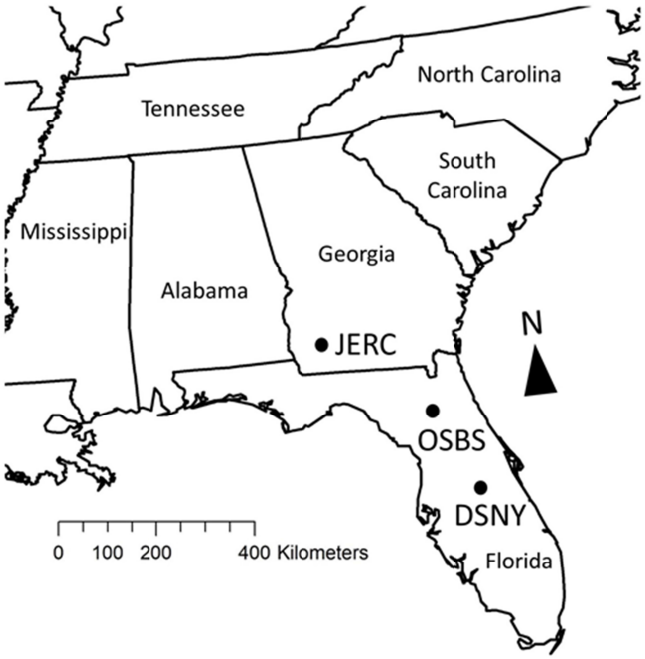

Figure 1. Locations of the three NEON study sites within the southeastern U.S.: Disney Wilderness Preserve (DSNY), Jones Ecological Research Station (JERC), and Ordway-Swisher Biological Station (OSBS).

(*Taxodium ascendens*) with evergreen shrubs and shade-tolerant herbs in the understory. Evergreen wetland forests (syn. baygall; FNAI, 2010) occur in similar though often shallower physical locations but are dominated by evergreen broadleaf trees, primarily titi (*Cyrilla racemiflora*) and sweetbay (*Magnolia virginiana*). Grass marshes (syn. depression marsh, basin marsh; FNAI, 2010) are herbaceous wetlands associated with shallow karst depressions with fluctuating water levels. Scrub is composed of shrub-like trees with sparse herbaceous surface vegetation in slightly raised areas of nearly pure sand, and they typically burn with crown fires less frequently than flatwoods (FNAI, 2010). Native pastures are former flatwoods communities where trees and most woody vegetation were removed but otherwise have similar herbaceous vegetation and are frequently burned.

2.1.2 Jones Ecological Research Center (JERC)

JERC (31.2205°N, -84.4793°W) is a 12,000 ha private research center. Elevation ranges from approximately 30-50 m asl. Low to high average monthly temperatures are 10.2°C in January and 28.4°C in July. The dominant soils are in the Ultisol order (Soil Survey Staff, 2022), consisting mostly of sand with a clayey subhorizon. Native pine savannas dominating JERC are specifically upland pine communities (FNAI, 2010). This community type has an open canopy of mostly longleaf pine and fire-tolerant broadleaf trees (mostly genera *Quercus* and *Carya*), surface vegetation of resprouting broadleaf tree and shrub species, and a diverse herbaceous community dominated by wiregrass in most areas (Carr et al., 2010). It is dependent on frequent fire and is typically burned at two-year intervals at JERC. Bottomland forest communities occur in occasionally flooded areas and have a closed-canopy dominated by mesic broadleaf deciduous trees with shrubs and sparse shade-tolerant herbaceous plants in the understory (FNAI, 2010). Old-field pine communities are former row crop sites that have been planted with longleaf pine

and managed with frequent fire similar to the upland pine savannas, such that it contains a subset of native savanna plant species (Kirkman et al., 2004; Dixon et al., 2021). Cultivated crop sites are annually tilled and planted mostly with cotton, peanuts, corn, or soybeans and harbor a variety of agricultural weeds.

2.1.3 Ordway-Swisher Biological Station (OSBS)

OSBS (29.6903°N, -82.0176°W) is a 3,800 ha property owned and managed by the University of Florida. Elevation of the property ranges from 30-55 m asl, and low and high average monthly temperatures are 12.6°C in January and 27.6°C in August. The dominant soils are in the Entisol order (Soil Survey Staff, 2022) consisting mostly of sand. Its native pine community is specifically sandhill pine (FNAI, 2010) burned at three-year intervals. This community has an open canopy of mostly longleaf pine and fire-tolerant broadleaf trees and relatively xeric surface vegetation, though also dominated by wiregrass like the other sites. Open wetland forests (syn. basin swamps; FNAI, 2010), with an open canopy of mostly pond cypress and black gum (*Nyssa biflora*), occur in locations with long hydroperiods and have understory vegetation consisting of wetland shrubs, ferns, and floating and emergent wetland herbaceous vegetation. Upland mixed forests (syn. upland hardwood forest; FNAI, 2010) occupy areas that were previously sandhill pine communities but fire-excluded for decades, resulting in a closed canopy of oaks and residual pines and sparse understory vegetation. Disturbed areas are sandhill pine communities with a history of intensive soil disturbance and are characterized by few pine and broadleaf trees and sparse ruderal forbs and grasses. Bottomland forests and grass marshes at OSBS are similar to those described for DSNY and JERC, respectively.

2.2 Field Data

We used field data to confirm that locations within areas classified as a particular

community type have similar plant species composition relative to other community types. We the plant presence and percent cover datasets provided for each site through the NEON portal (NEON, 2022), as well as plots that we established and censused on each of the properties, with the goal of providing multiple plots per common community type. Plots used in the study were distributed such that each represented an individual natural community feature or management unit. Both the NEON plots and our additional plots were 20 m x 20 m (400 m^2) and were censused for presence of all vascular plant species during the growing season. The NEON plots had all been censused within the previous three years. We censused additional plots at DSNY in late March, OSBS in April, and JERC in May of 2022. We also field-validated our interpretation of community types at virtual plot locations chosen from aerial photography for selecting points representing those communities for the biodivMapR analyses, described below. For the dominant pine community types, total numbers of plots ranged from 20-27 among the three study sites, and for the other community types numbers of plots ranged from 2-12 (average = 6) (Table 1).

Table 1. Plant community types, numbers of field-measured plots (NEON plant presence and percent cover plots and our plots combined), and numbers of virtual plot locations remotely chosen for analysis for the three study sites. DSNY = Disney Wilderness Preserve, JERC = Jones Ecological Research Center, OSBS = Ordway-Swisher Biological Station.

Site	Community	Field plots	Virtual plots
DSNY	Cypress forest	2	5
DSNY	Evergreen wetland forest	2	6
DSNY	Flatwoods pine savanna	25	8
DSNY	Grass marsh	4	5
DSNY	Pasture	12	4
DSNY	Scrub	4	6
JERC	Bottomland forest	3	5
JERC	Cultivated crops	5	5

JERC	Old-field pine savanna	3	4
JERC	Upland pine savanna	26	6
OSBS	Bottomland forest	7	4
OSBS	Disturbed	5	6
OSBS	Grass marsh	10	6
OSBS	Open wetland forest	5	5
OSBS	Upland mixed forest	7	6
OSBS	Sandhill pine savanna	13	11

For each of the three sites, we ran nonmetric multidimensional scaling (NMS) ordinations using PC-ORD v. 7 (McCune and Mefford, 2018) to confirm that communities are relatively definable in terms of plant species composition. We also ran multi-response permutation procedures (MRPP) based on the Bray-Curtis dissimilarity matrices, which averages the within-community dissimilarity and total dissimilarity to calculate within-community agreement (A) (McCune and Grace, 2002) for comparison to biodivMapR results, described below.

2.3 Airborne Hyperspectral Reflectance Data

Surface reflectance data were acquired by the NEON Imaging Spectrometer (NIS) on the NEON Airborne Observation Platform (AOP) and accessed through the NEON Data Portal (data.neonscience.org). The NIS design is based on Next Generation Airborne Visible/Infrared Imaging Spectrometers (AVIRISng), which was developed under the Next-Generation Imaging Spectrometer (NGIS) program at NASA's Jet Propulsion Laboratory (JPL). The raw data include 426 bands collected at 1--m spatial resolution in the visible-to-shortwave infrared (VSWIR) range between 0.38 and 2.5 microns and a spectral sampling of 5 nm (Karpowicz and Kampe, 2015). Images were collected by NEON as flightlines approximately 500 m in width which were mosaiced for the 10 km x 10 km area and subsequently separated into 1 km^2 tiles available for

download (Karpowicz and Kampe, 2015). We downloaded and mosaiced tiles to cover our area of interest, with numbers of tiles ranging from 25 to 50 tiles among the three study sites.

Reflectance data used in this study were collected in September, 2021 for all three sites. Although time since the previous prescribed fire no doubt had some influence on the reflectance properties of fire-dependent communities, the time between burning in the spring and imaging in September is sufficient for pine savannas to have considerable recovery by resprouting perennial vegetation characteristic of these communities (Picotte and Robertson, 2011).

NEON reflectance data were initially converted from at-sensor radiances to surface reflectance using the ATCOR atmospheric correction (Karpowicz and Kampe, 2015) and then provided to the community as georectified images in ENVI format using the neonhs R package (https://www.earthdatascience.org/neonhs/). No additional corrections were performed. For each site, we mosaiced multiple 1 km x 1 km flightline mosaics provided by NEON to create a seamless product that included most of the NEON vegetation plots. Finally, we used a python script from the Space-based Imaging Spectroscopy and Thermal pathfindER (SISTER) resample repository (https://github.com/EnSpec/sister-resample) to aggregate the NEON mosaics from their native resolution of 1 m to 5 m, 15 m, 30 m, 40 m, 60 m, and 90 m. These final mosaics, including the original 1 m resolution mosaic, served as the inputs into the biodivMapR package (https://jbferet.github.io/biodivMapR/index.html).

R scripts provided through the GitHub page were used to guide the workflow. The first step masked irrelevant pixels (e.g., non-vegetated, cloudy, shadow) based on a spectral thresholding of NDVI and the Blue/NIR domains. We used the default thresholds for Blue and NIR but lowered the Normalized Difference Vegetation Index (NDVI) threshold from the default of 0.5 to 0.1 to include lightly vegetated areas characteristic of some frequently burned areas. A

series of processing steps were then applied to the remaining data, including band removal, continuum removal, and dimensionality reduction using principal component analysis (PCA). The wavelengths removed from the analysis corresponded to atmospheric water absorption or otherwise had a high signal to noise ratio (Sousa et al., 2022), specifically 0-400 nm, 895-1005 nm, 1320-1480 nm, 1780-2040 nm, and 2400-3000 nm.

After the data were normalized and transformed, we performed a selection of principal components that were most relevant to the mapping of biodiversity in our study areas. While some components highlighted differences in vegetation properties, others showed information related to sensor characteristics or very high noise level. Therefore, it was important to visualize each component and follow the published recommendations for component selection (Féret and de Boissieu, 2020). The next step partitioned the selected components into a predefined number of clusters (spectral species) by using k-means clustering and assigned a cluster ID to each pixel. We used either 50 spectral species (default) and then 20 spectral species for comparison, described below. BiodivMapR then calculates the Bray-Curtis dissimilarity index for each pair of mapping windows based on abundance of each spectral species, and then uses an ordination technique to assign three numbers to each pixel as the basis for visualizing maps of beta diversity. BiodivMapR can also provide a BC matrix including only dissimilarities among points of interest, which we used for analyses described below.

In our first analysis, the goal was to test for effect of pixel size on the ability of biodivMapR to distinguish natural community types within each site in terms of dissimilarity in spectral species in pairwise comparisons among locations within and among communities. For this analysis, we used similarly size mapping windows (270-300 m) for comparison among pixel resolutions, such that the number of pixels per mapping window varied by several orders of

magnitude (Table 2). We used reference real color imagery from NEON used to create virtual plots by placing points within homogeneous areas of a given community type large enough to contain one mapping window. We took this approach instead of using the locations of the 20 m x 20 m plots with field data because of the spatial mismatch between the plots and the much larger mapping windows, and so we could choose a more balanced representation of community types than provided by the field plots (Table 2). However, we confirmed in the field that the virtual plot locations accurately represented the remotely interpreted community type. The function 'biodiversity from plots' was then used to extract the Bray-Curtis dissimilarity matrices comparing spectral species composition among mapping windows centered on the virtual plot point locations.

Table 2. Pixel resolution, square mapping window width, and number of pixels per mapping window used to calculate diversity metrics from NEON imagery.

Pixel size (m)	Window size (m)	Pixels per window
1	270	72,900
5	270	2,916
15	270	324
30	270	81
40	280	49
60	300	25
90	270	9

Using the Bray-Curtis dissimilarity matrices, we ran MRPP analyses using the vegan package and the function mrpp in R (Oksanen et al., 2022) to provide the average within-community dissimilarity and total dissimilarity for each study site and pixel resolution (Table 2) and using 50 versus 20 spectral species. From these values we calculated within community

agreement (A) as A = 1 - (average within variance / average total variance) (McCune and Grace, 2002). We also calculated the percentage of total dissimilarities equal to 1 (no spectral species in common) to assess the method's ability to identify relative dissimilarity as opposed to absolute dissimilarity. We charted trends in each of these metrics with increasing pixel size to visualize the effects of image resolution on capacity to discriminate natural communities as reflected in the A statistic.

In a second analysis, we assessed the effects of changing the sizes of the both pixel size and mapping windows on ability to discriminate among community types. For this analysis, the sizes of mapping windows were adjusted according to pixel size to maintain numbers of pixels per window within the range of 49-81 (Table 3), which is within the 50-400 range recommended by Féret and de Boissieu (2020). In this analysis we used 20 spectral species, as the first analysis revealed that this number provides higher resolution among community types. We used only pixel sizes 1 m, 5 m, 15 m, 30 m, and 40 m, as larger pixel resolutions would require mapping windows with 480 m or greater dimensions, which is larger than the area of any natural community feature in the study. Similar to the first analysis, we ran MRPP analyses to derive within community dissimilarity, total dissimilarity, the A statistic, and percentage of

Table 3. Pixel resolution, square mapping window width, and number of pixels per mapping window used to calculate diversity metrics from NEON imagery.

Pixel size (m)	Window size (m)	Pixels per window
1	8	64
5	40	64
15	120	64
30	270	81
40	280	49

dissimilarities equal to one, and values were charted to visualize trends among pixel sizes.

3. Results

NMS analyses of field collected data generally confirmed that community types were well defined by their plant species composition, represented by presence or absence of species, as visualized using NMS ordination (Figure 2).

The analysis using varying pixel resolutions (1-90 m) with similar sized mapping windows (270-300 m) showed a fairly strong decrease in capacity to identify beta diversity with increasing pixel size (Figure 3). The average dissimilarity among plots within community types increased with coarser pixel resolution (Figure 3a). Average total dissimilarity among all plots also

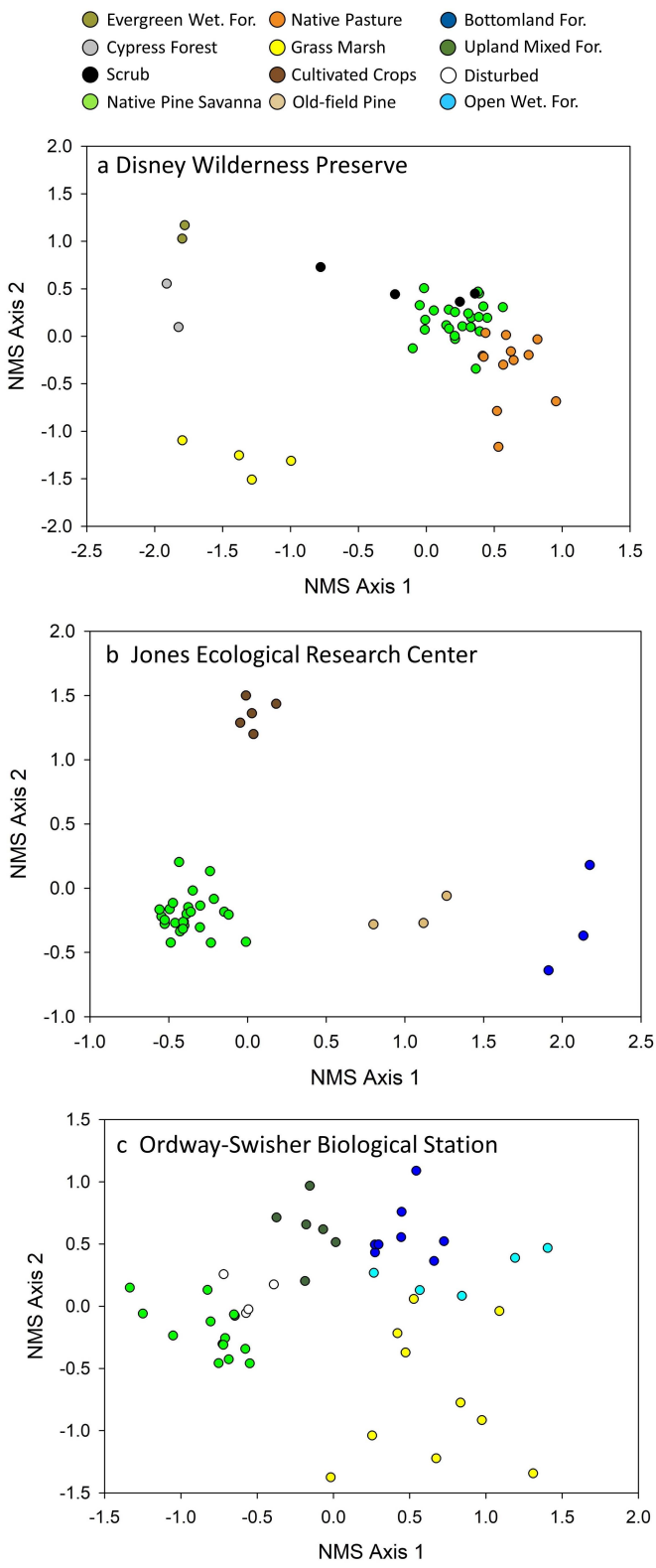

Figure 2. Results from NMS ordination on field collected plant species presence-absence data for each community type and study site. Symbols represent data from 20 m x 20 m field plots.

increased, but more gradually (Figure 3b). These patterns resulted in a general decrease in within community agreement (A = 1 - (within dissimilarity / total dissimilarity)) with coarser pixel resolution (Figure 3c). The percentage of pairwise comparisons with dissimilarity = 1 (no spectral species in common) increased with coarser pixel resolution (Figure 3d). These patterns can also be visualized through results of NMS analyses reflecting the Bray-Curtis dissimilarity matrices comparing spectral species among communities (Figure 4a-i). Plant communities generally can be distinguished in the ordination plots using 1 m pixels (Figure 4a-c) and 30 m pixels (Figure 4d-f), but the capacity to distinguish community types has largely broken down at 90 m (Figure 4g-i).

The trends were similar between analyses using 50 spectral species versus 20 spectral species. However, the 20

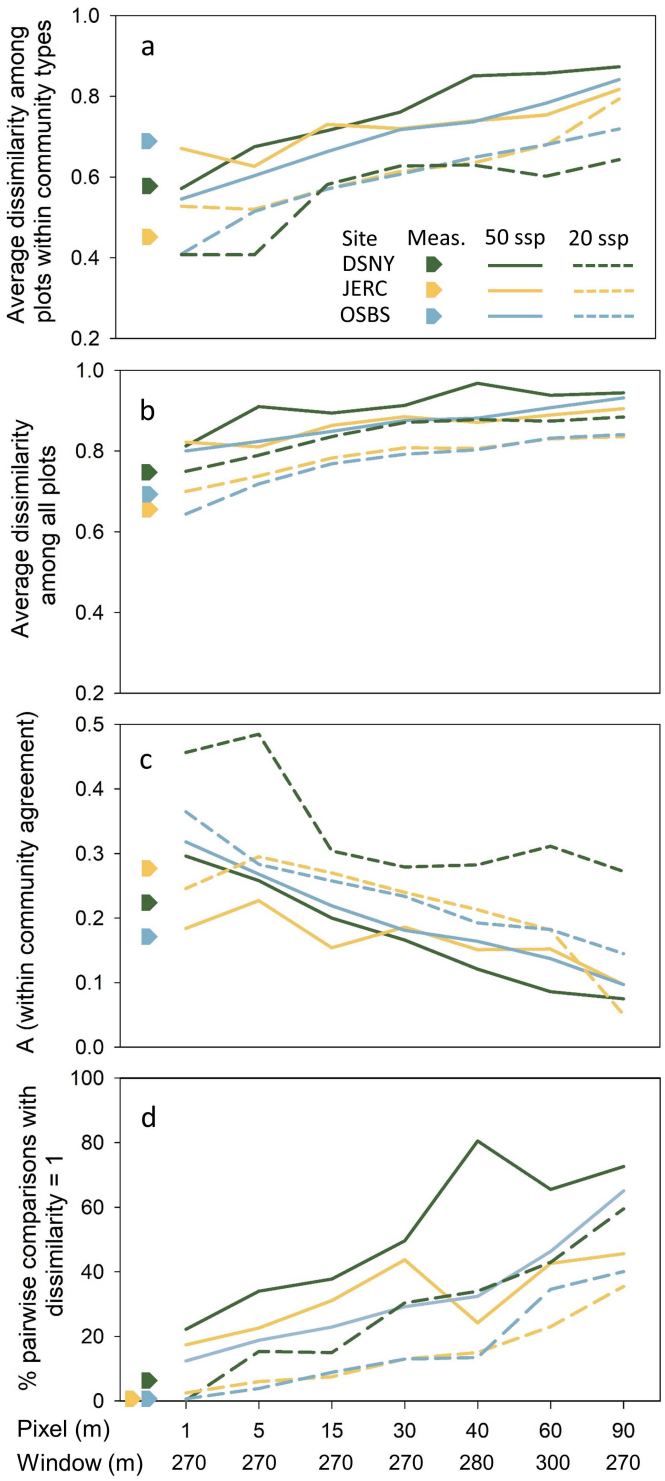

Figure 3. Spectral species dissimilarity within community types, dissimilarity among all plot locations, within community agreement (A), and percentage of comparisons with absolute disagreement for each pixel resolution, study site, and 50 versus 20 spectral species using 270-300 m mapping windows. Symbols represent values for field-measured plots using actual plant species.

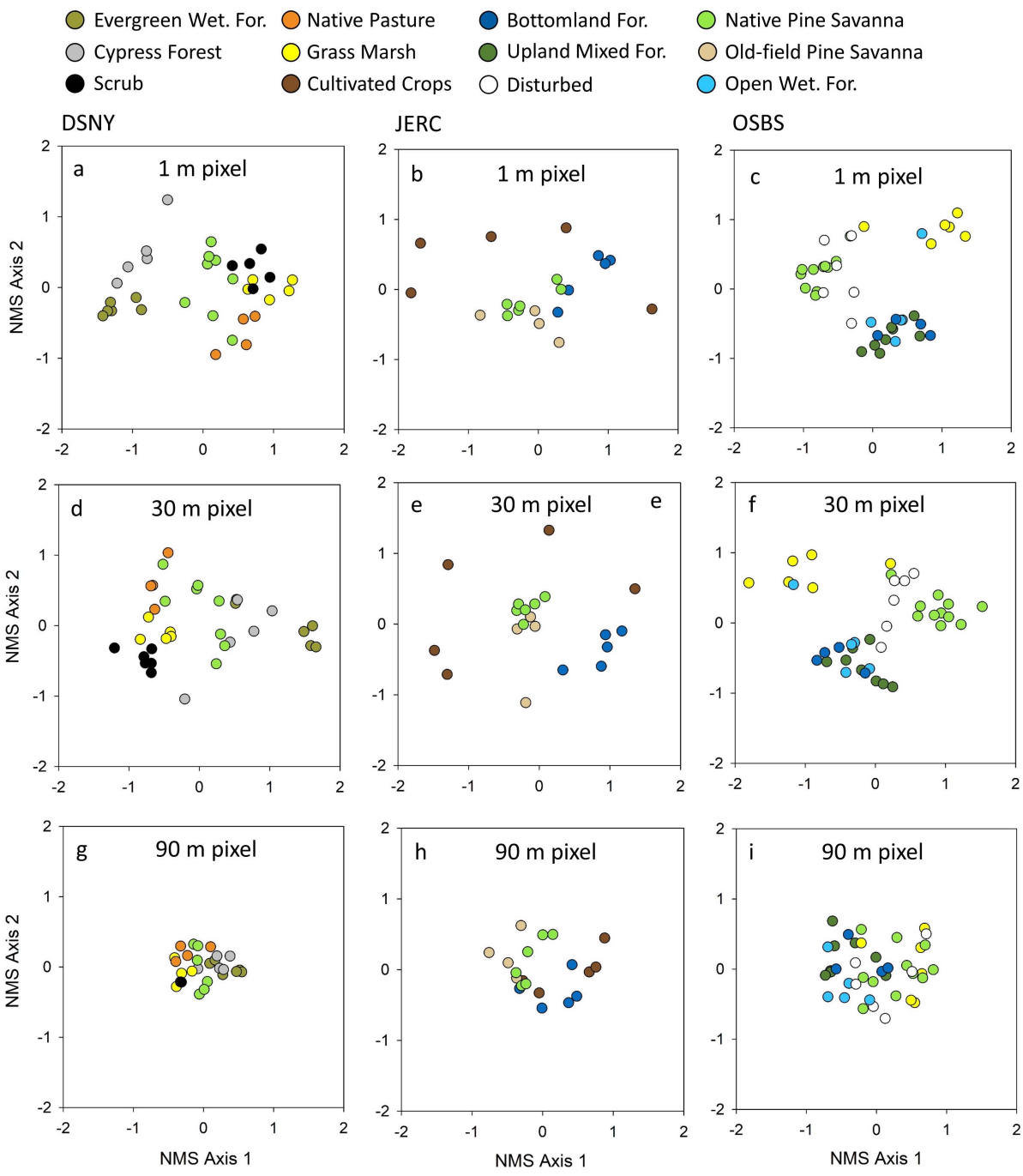

Figure 4. Results from NMS analyses from Bray-Curtis dissimilarity matrices generated by biodivMapR at using three pixel resolutions (1m, 30 m, 90 m) and 270 m mapping windows using 50 spectral species at each of the three study sites (DSNY = Disney Wilderness Preserve, JERC = Jones Ecological Research Center, OSBS = Ordway-Swisher Biological Research Station). Symbol colors correspond to community types.

384

spectral species invariably resulted in lower dissimilarity among plots, higher within community agreement, and fewer pairwise comparisons with dissimilarity = 1, and thus overall higher resolution in distinguishing plots among community types (Figure 3a-d).

Field measured values for dissimilarity metrics were generally most similar to remote sensing estimates that used the finest pixel resolution (Figure 3a-d). Field-measured plots had very few pairwise comparisons with zero plant species in common, in sharp contrast to virtual plot dissimilarities based on spectral species (Figure 3d).

For our second analysis, which compared varying pixel resolutions and mapping window sizes with similar numbers of pixels per window (49-81), there were no strong trends evident (Figure 5a-d). The values were generally similar to those for intermediate pixel

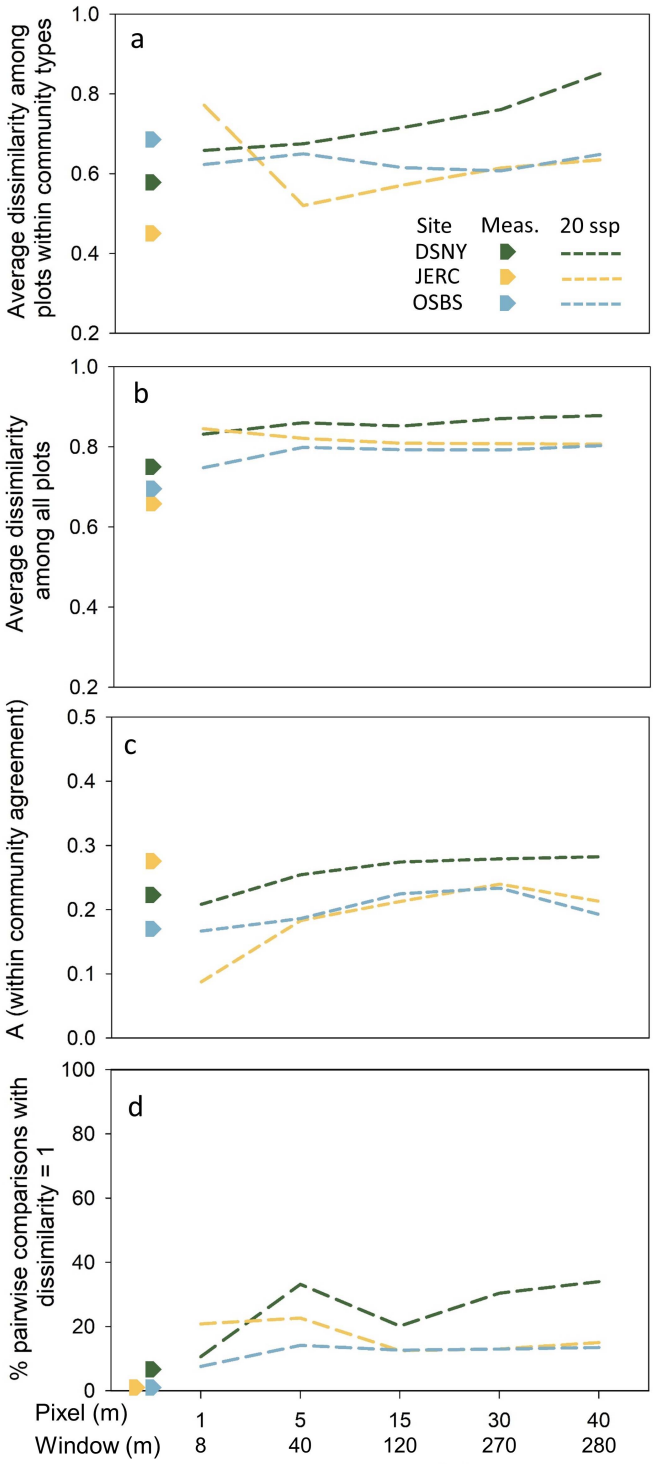

Figure 5. Spectral species dissimilarity within community types, dissimilarity among all plot locations, within community agreement (A), and percentage of comparisons with absolute disagreement for each pixel resolution and study site using 20 spectral species and varying sized mapping windows with 49-81 spectral species per window. Symbols represent values for field-measured plots using actual plant species.

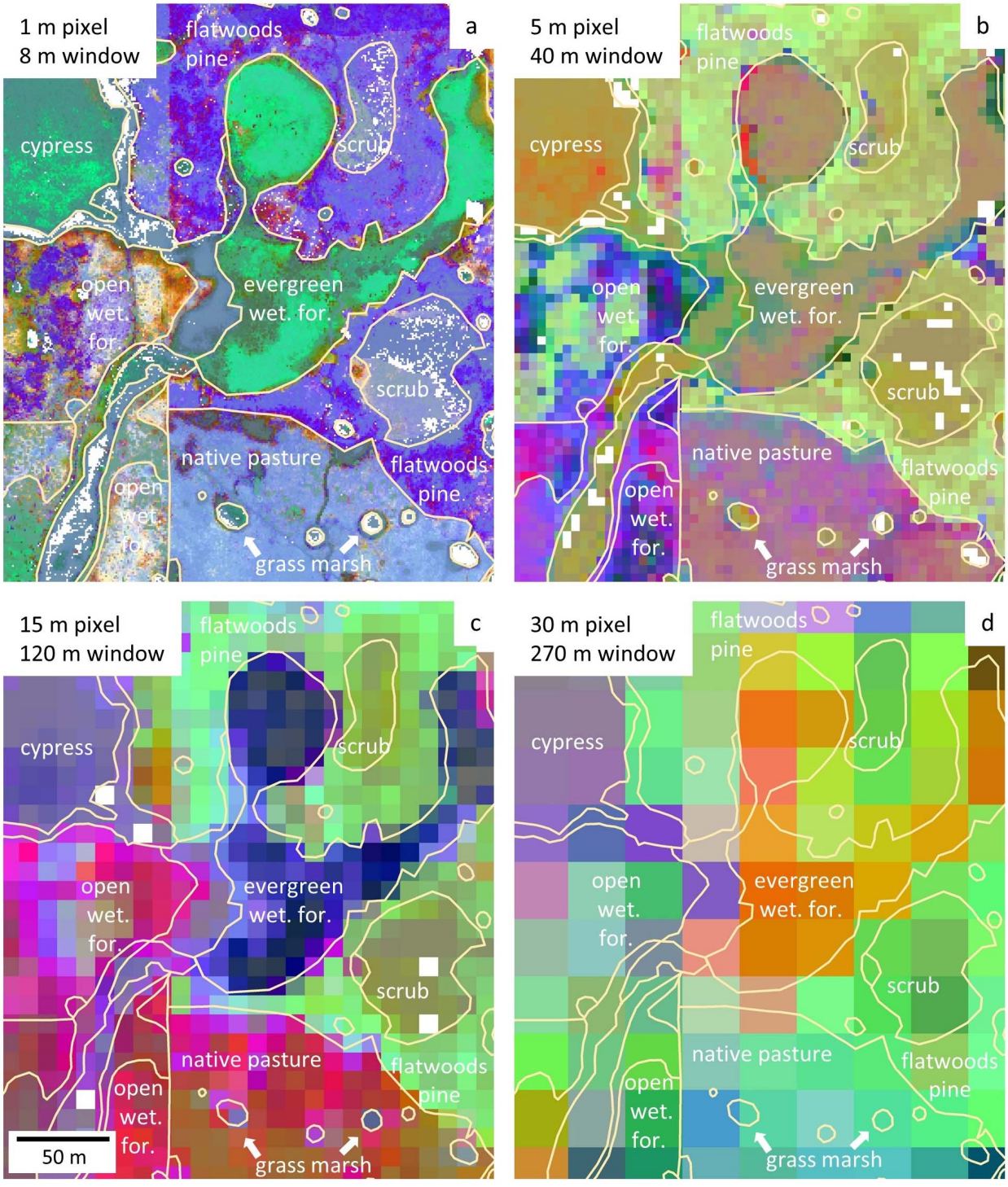

Figure 6. Beta diversity maps generated by bioDivMapR using different pixel resolutions and mapping window sizes to represent part of the Disney Wilderness Preserve (DSNY). Beige lines indicate the extent of natural community types interpreted from aerial photographs, and some natural community types are labeled.

resolutions (30-40 m) using the 270-300 m mapping windows (Figure 3a-d).

Taken together, these results show that the number of pixels per window has the strongest influence on ability to discern community types using biodivMapR. However, beta diversity maps generated from varying window sizes with similar numbers of pixels per window show that increasing window size to incorporate more pixels decreases the spatial resolution of beta diversity maps, which at some point decreases the capacity to spatially distinguish community types (Figure 6a-d).

4 Discussion

4.1 Assessment of biodivMapR outputs

Results of our analysis indicate that spatial resolution of imagery has a strong effect on the capacity to identify beta diversity using the algorithm biodivMapR. The key variable influencing the ability to detect relative levels of dissimilarity among locations is the number of pixels per mapping window, where more pixels provide greater resolution. Increasingly coarse pixel resolution can be compensated by increasing the mapping window size, though at the cost of decreasing spatial resolution of beta diversity maps built from the windows.

The decreasing capacity to identify beta diversity with fewer pixels per mapping unit appears to result primarily from overestimation of dissimilarity among mapping windows within community types, such that there is a loss of distinction between the within-community dissimilarity and total dissimilarity. This overestimation results from pairwise comparisons between windows showing increasingly few or zero spectral species in common as pixels per window decreases, even if comparisons are within the same community type. Where there were fewer than 50 pixels per mapping window (60-90 m pixels within 270-300 m mapping windows), more than half of the total dissimilarities were equal to one (Figure 3d). Such absolute dissimilarities provide limited information, even if between different community types, as they

suggest that communities within the same landscape have no more species in common than communities on different continents. In fact, however, all pairwise comparisons among community types using field data at JERC and OSBS showed at least some generalist species in common, and only about 6% of comparisons at DSNY had no species in common. We suggest that using the appropriately sized mapping window relative to image pixel size, particularly with the goal of minimizing the number of dissimilarities equal to one, is essential for producing the most meaningful beta diversity maps. The recommendation by Féret and de Boissieu (2020) that there be a minimum of 50 pixels per mapping window seems appropriate, although, as they point out, assigning an appropriate number of spectral species is also important.

The strong effect of number of spectral species on detection of beta diversity also relates to overestimation of dissimilarity in pairwise comparisons among mapping windows. In this study, reducing the number of spectral species from 50 to 20 considerably improved the capacity to distinguish community types. This effect may seem counterintuitive, but reducing the number resulted in there being more spectral species in common among windows within community types, which more strongly contrasted average dissimilarity among all community types. The advantage of having fewer spectral species may be a special case where pixels represent local plant community composition rather than individual plant species. Where two pixels are assigned different spectral species when in fact they have some species in common, their dissimilarity is overestimated as 1 (Rocchini et al., 2022). For the studied community types, it is easy to imagine that variations in the local abundance of potentially high-cover species, such as wiregrass, longleaf pine, and saw palmetto, might cause different spectral species classifications despite the overall plant community composition is quite similar (Ostertag and Robertson, 2007). Of course at some point reduction of spectral species will cause pixels classifications to be overly

homogenized and will not effectively represent degree of dissimilarity among communities. Currently the appropriate number of spectral species must be determined by trial and error with validation data, but eventually better guidelines might be determined based on the spatial resolution of imagery relative to the spatial scale of plant community complexity.

4.2 Implications for the SBG mission

Our results indicate the importance of maximizing the spatial resolution of imagery for the most effective mapping and monitoring of biodiversity. Using biodivMapR or similar algorithms, the spatial resolution of imagery will determine the minimum size of mapping windows with enough pixels to produce meaningful maps, which in turn limits the spatial resolution of those maps. The implications of mapping window size depend on the spatial distribution of natural community types representing plant beta diversity on a given landscape. In the southeastern U.S. Coastal Plain, pixel resolutions ≥ 30 m, such as those recommended for the SBG mission, corresponding to mapping windows > 270 m in order to contain > 50 pixels, would be insufficient for identifying the contribution to beta diversity by certain community types, such as isolated ephemeral ponds and narrow riparian features. However, coarser resolution data might still identify beta diversity among larger ecological features, such areas dominated by native pine communities, scrub, cypress forests, and native pastures (Figure 6).

Our study provides some context for assessing tradeoffs in investment among image spatial resolution, spectral resolution, signal-to-noise ratio, and flyover return interval. Although our focus was on spatial resolution, the ability for biodivMapR to distinguish natural community types as well as it did even with much coarser spatial resolution than those generally recommended for remote sensing of biodiversity (Gamon et al., 2020) presumably benefited from the immense spectral resolution afforded by hyperspectral imagery (Thorpe et al., 2013).

Flights for the NEON project were chosen on clear days, whereas utility of satellite remote sensing data is limited by cloudiness, underscoring the importance of sufficiently frequent returns to collect cloud-free data in regions where clear days are limited. Data acquisition at a frequency sufficient to account for seasonal effects and plant phenological changes and to monitor changes in land use and ecological status over time is also critical. For example, analyses in this study were simplified by having full coverages of the areas of interest within a few days and at the height of the growing season. However, within these limits, our analysis underscores the need to maximize the spatial resolution of imagery for effective mapping of plant beta diversity.

5 Conclusions

Using the algorithm biodivMapR with hyperspectral remote sensing imagery, we show that the capacity to detect plant beta diversity as represented by plant community types decreases with number of pixels per mapping window. It follows that pixel resolution places a lower limit on size of mapping windows that are effective for distinguishing community types, which in turn limits the spatial resolution beta diversity maps composed of mapping windows. When image pixel size is much larger than individual plants, the effect of having too few pixels per window, as well as assignment of too many spectral species per image, has the effect of overestimating dissimilarity among locations that in fact may have many plant species in common. This overestimation undermines the capacity to contrast mapping window dissimilarity within versus among community types and thus reduces the information content of beta diversity maps. These results demonstrate the advantage of maximizing spatial resolution of hyperspectral imaging instruments on the anticipated NASA Surface Biology and Geology satellite mission and similar

remote sensing projects.

Acknowledgments

Funding for the project was provided by NASA grant 80NSSC21K1956 and Tall Timbers Research Station. The National Science Foundation National Ecological Observatory Network (NEON) provided remote sensing imagery and plot field data on which the study was based. We thank Beatriz Pace-Aldana of The Nature Conservancy Disney Wilderness Preserve, Jeffery Cannon of the Jones Ecological Research Center, and Andrew Rappe of Ordway-Swisher Biological Station for providing access to sites and support for field work. We thank Joe Noble and Morgan Varner for administrative and technical support.

Open Research

Software used in analyses were:

PC-ORD, citation McCune, B., & Mefford, M. J. (2018) *PC-ORD. Multivariate Analysis of Ecological Data v. 7.08.* Oregon: MjM Software Design, available for download at https://www.wildblueberrymedia.net/store/pc-ord-7-single-user-license-regular-new-user.

R, downloadable at https://www.r-project.org/. We specifically used package version 2.6-2. https://doi.org/10.1002/env.516 https://CRAN.R-project.org/package=vegan&gt

BiodivMapR was downloaded from https://github.com/jbferet/biodivMapR and installed with the command line devtools::install _ git('https ://github.com/jbfer et/biodi vMapR.git') as provided in the manuscript by Asner and de Boissieu (2020). Data command were those described in the manuscript text.

Data sources are described as follows:

NEON hyperspectral remote sensing data as described in the manuscript was accessed by logging into the NEON data portal (https://data.neonscience.org/home), where the user can browse to or search for the section called Spectrometer orthorectified surface directional reflectance - mosaic, where available data are listed by date for each NEON site.

NEON plant presence and percent cover as described in the manuscript was accessed by logging into the NEON data portal (https://data.neonscience.org/home), where the user can browse to or search for the section called Plant presence and percent cover, where data for each site are displayed for download.

The Bray-Curtis dissimilarity matrices for all analyses presented, and the shapefiles of field plot and virtual plot locations to which the matrices pertain, have been submitted for open access storage in Pangaea (pangaea.de). The data submission is pending review.

References

Carr, S. C., Robertson, K. M., & Peet, R. K. (2010) A vegetation classification of fire-dependent pinelands of Florida. *Castanea, 75*(2), 153-189. https://doi.org/10.2179/09-016.1

Cavender-Bares, J., Gamon, J. A., & Townsend, P. A. (2020) The use of remote sensing to enhance biodiversity monitoring and detection: A critical challenge for the twenty-first century. In J. Cavender-Bares, J. A. Gamon, P. A. Townsend (Eds.), *Remote sensing of plant biodiversity* (pp. 1-12). Cham, Switzerland: Springer.

Cawse-Nicholson, K., Townsend, P. A., Schimel, D., Assiri, A. M., Blake, P. L., Buongiorno, M. F., Campbell, P., Carmon, N., Casey, K. A., Correa-Pabón, R. E. & Dahlin, K.M. (2021) NASA's surface biology and geology designated observable: A perspective on surface

548 imaging algorithms. *Remote Sensing of Environment*, *257*, 112349.
549 https://doi.org/10.1016/j.rse.2021.112349
550 Dixon, C. M., Robertson, K. M., Ulyshen, M. D., & Sikes, B. A. (2021) Pine savanna restoration
551 on agricultural landscapes: The path back to native savanna ecosystem services. *Science of*
552 *the Total Environment, 818,* 151715. https://doi.org/10.1016/j.scitotenv.2021.151715
553 Drewa, P. B., Platt, W. J., & Moser, E. B. (2002) Community structure along elevation gradients
554 in headwater regions of longleaf pine savannas. *Plant Ecology*, *160*(1), 61-78.
555 https://doi.org/10.1023/A:1015875828742
556 Féret, J.-B. & Asner, G. P. (2014) Mapping tropical forest canopy diversity using high-fidelity
557 imaging spectroscopy. *Ecological Applications, 24*(6), 1289-1296.
558 https://doi.org/10.1890/13-1824.1
559 Féret, J. B., & de Boissieu, F. (2020) biodivMapR: An R package for α-and β-diversity mapping
560 using remotely sensed images. *Methods in Ecology and Evolution, 11*(1), 64-70.
561 https://doi.org/10.1111/2041-210X.13310
562 Florida Natural Areas Inventory. (2010) *Guide to the natural communities of Florida: 2010*
563 *edition.* Tallahassee, Florida: Florida Natural Areas Inventory.
564 Gamon, J. A., Wang, R., Gholizadeh, H., Zutta, B., Townsend, P. A., & Cavender-Bares, J..
565 (2020) Chapter 16: Consideration of scale in remote sensing of biodiversity. In J. Cavender-
566 Bares, J. A. Gamon, P. A. Townsend (Eds.), *Remote sensing of plant biodiversity* (pp. 425-
567 447). Cham: Springer.
568 Gholizadeh, H., Gamon, J. A., Townsend, P. A., Zygielbaum, A. I., Helzer, C. J., Hmimina, G.
569 Y., Yu, R., Moore, R. M., Schweiger, A. K., & Cavender-Bares, J. (2019) Detecting prairie

biodiversity with airborne remote sensing. *Remote Sensing of Environment, 221*, 38-49. https://doi.org/10.1016/j.rse.2018.10.037

Gholizadeh, H., Gamon, J. A., Helzer, C. J., & Cavender-Bares, J. (2020) Multi-temporal assessment of grassland α-and β-diversity using hyperspectral imaging. *Ecological Applications*, *30*(7), e02145. https://doi.org/10.1002/eap.2145

Glitzenstein, J. S., Streng, D. R., & Wade, D. D. (2003) Fire frequency effects on longleaf pine (*Pinus palustris* P. Miller) vegetation in South Carolina and Northeast Florida, USA. *Natural Areas Journal, 23*(1), 22-37.

Karpowicz, B., & Kampe, T. (2015) The NEON imaging spectrometer radiance to reflectance algorithm theoretical basis document. NEON document #NEON.DOC.001288. Boulder, Colorado: National Ecological Observation Network, National Science Foundation. https://data.neonscience.org/data-products/DP3.30006.001#collectionAndProcessing

Kirkman, L. K., Coffey, K. L., Michell, R. J., & Moser, E. B. (2004) Ground cover recovery patterns and life-history traits: implications for restoration obstacles and opportunities in a species-rich savanna. *Journal of Ecology, 92*(3), 409-421. https://doi.org/10.1111/j.0022-0477.2004.00883.

Lane, C. R., & D'Amico, E. (2010) Calculating the ecosystem service of water storage in isolated wetlands using LiDAR in north central Florida, USA. *Wetlands, 30*, 967–977. doi:10.1007/s13157-010-0085-z

Lortie, C. J., Brooker, R. W., Choler, P., Kikvidze, Z., Michalet, R., Pugnaire, F. I., & Callaway, R. M. (2004) Rethinking plant community theory. *Oikos, 107*(2), 433-438. https://doi.org/10.1111/j.0030-1299.2004.13250.x

McCune, B., & Grace, J. B. (2002) *Analysis of ecological communities.* Gleneden Beach, Oregon: MjM Software Design.

McCune, B., & Mefford, M. J. (2018) *PC-ORD. Multivariate Analysis of Ecological Data v. 7.08.* Oregon: MjM Software Design.

National Academies of Sciences, Engineering, and Medicine (2018) *Thriving on our changing planet: A decadal strategy for earth observation from space.* Washington, DC: The National Academies Press. https://doi.org/10.17226/24938

NEON (National Ecological Observatory Network). *Plant presence and percent cover* (DP1.10058.001), RELEASE-2022. https://doi.org/10.48443/pr5e-1q60. Dataset accessed from https://data.neonscience.org on June 28, 2022

Noss, R. F., Platt, W. J., Sorrie, B. A. Sorrie, Weakley, A. S., Means, D. B., Costanza, J., & Peet, R. K. (2015) How Global Biodiversity Hotspots May Go Unrecognized: Lessons from the North American Coastal Plain. *Diversity and Distributions, 21*(2), 236–244. https://doi.org/10.1111/ddi.12278

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H., Fitz, J. R, Friendly, M., Furneaux, B., Hannigan, G., Hill, M., Lahti, L., McGlinn, D., Ouellette, M., Ribeiro, Cunha, E., Smith, T., Stier, A., Ter Braak, C., Weedon, J. (2022) _vegan: Community Ecology Package_. R package version 2.6-2. https://doi.org/10.1002/env.516 https://CRAN.R-project.org/package=vegan&gt

613 Ostertag, T. E., & Robertson, K. M. (2007) A comparison of native versus old-field vegetation in
614 upland pinelands managed with frequent fire, South Georgia, USA. *Tall Timbers Fire*
615 *Ecology Conference Proceedings, 23*. 109-120.
616 Palmer, M. W., Earls, P. G., Hoagland, B. W., White, P. S., & Wohlgemuth, T. (2002).
617 Quantitative tools for perfecting species lists. *Environmetrics*, *13*, 121–137.
618 https://doi.org/10.1002/env.516
619 Pereira, H. M., Belnap, J., Böhm, M., Brummitt, N., Garcia-Moreno, J., Gregory, R., Martin, L.,
620 Peng, C., Proença, V., Schmeller, D. and Swaay, C. V. (2017) Monitoring essential
621 biodiversity variables at the species level. In *The GEO handbook on biodiversity observation*
622 *networks* (pp. 79-105). Cham, Switzerland: Springer.
623 Picotte, J. J., & Robertson, K. M. (2011) Timing constraints on remote sensing of wildland fire
624 burned area in the southeastern US. *Remote Sensing, 3*(8), 1680-1690.
625 https://doi.org/10.3390/rs3081680
626 Robertson, K. M., Platt, W. J., & Faires, C. E. (2019) Patchy fires promote regeneration of
627 longleaf pine (*Pinus palustris* Mill.) in pine savannas. *Forests 10*(5), 367.
628 https://doi.org/10.3390/f10050367
629 Rocchini, D., Chiarucci, A., & Loiselle, S. A. (2004) Testing the spectral variation hypothesis by
630 using satellite multispectral images. *Acta Oecologica, 26*(2), 117-120.
631 https://doi.org/10.1016/j.actao.2004.03.008
632 Rocchini, D., Balkenhol, N., Carter, G.A., Foody, G.M., Gillespie, T.W., He, K.S., Kark, S.,
633 Levin, N., Lucas, K., Luoto, M., & Nagendra, H. (2010) Remotely sensed spectral
634 heterogeneity as a proxy of species diversity: recent advances and open

challenges. *Ecological Informatics*, *5*(5), 318-329. https://doi.org/10.1016/j.ecoinf.2010.06.001

Rocchini, D., Luque, S., Pettorelli, N., Bastin, L., Doktor, D., Faedi, N., Feilhauer, H., Feret, J.B., Foody, G.M., Gavish, Y., & Godinho, S. (2018) Measuring β-diversity by remote sensing: a challenge for biodiversity monitoring. *Methods in Ecology and Evolution, 9*(8), 1787–1798. https://doi.org/10.1111/2041-210X.12941

Schimel, D., Townsend, P. A., & Pavlick, R. (2020) Prospects and pitfalls for spectroscopic remote sensing of biodiversity at the global scale. In J. Cavender-Bares, J. A. Gamon, P. A. Townsend (Eds.), *Remote sensing of plant biodiversity* (pp. 503-518). Cham, Switzerland: Springer.

Soil Survey Staff, National Resources Conservation Service, USDA. 2022 *Web soil survey.* https://websoilsurvey.nrcs (accessed 30 September 2022)

Sousa, D., Brodrick, P., Cawse-Nicholson, K., Fisher, J. B., Pavlick, R., Small, C., & Thompson, D.R. (2022) The Spectral Mixture Residual: A source of low-variance information to enhance the explainability and accuracy of surface biology and geology Retrievals. *Journal of Geophysical Research: Biogeosciences*, *127*(2), p.e2021JG006672. https://doi.org/10.1029/2021JG006672

Thorp, K. R., French, A. N., & Rango, A. (2013). Effect of image spatial and spectral characteristics on mapping semi-arid rangeland vegetation using multiple endmember spectral mixture analysis (MESMA). *Remote Sensing of Environment*, *132*, 120–130. https://doi.org/10.1016/j.rse.2013.01.008