RESEARCH ARTICLE

Patterning emergent marsh vegetation assemblages in coastal Louisiana, USA, with unsupervised artificial neural networks

Gregg A. Snedden 🆔

U.S. Geological Survey, Wetland and Aquatic Research Center, Lafayette, Louisiana

Correspondence
Gregg A. Snedden, U.S. Geological Survey, Wetland and Aquatic Research Center, 700 Cajundome Blvd, Lafayette, LA 70506, USA.
Email: sneddeng@usgs.gov

Co-ordinating Editor: Hannes Feilhauer

Abstract

Questions: Are self-organizing maps (SOMs) useful for patterning coastal wetland vegetation communities? Do SOMs provide robust alternatives to traditional classification methods, particularly when underlying species response functions are unknown or difficult to approximate, or when a need exists to continuously classify new samples obtained under ongoing long-term ecosystem monitoring programs as they become available?

Location: Coastal Louisiana, USA.

Methods: A SOM was trained from in-situ observations of 559 vegetation species relative cover data from 2,526 samples collected over 8 years at 343 locations across coastal Louisiana. Hierarchical cluster analysis was applied to the SOM output to delineate vegetation community types, and indicator species analysis was conducted. Salinity and flood duration were compared across the delineated community types.

Results: The SOM patterned the 2,526 training samples into 260 output neurons, which were further clustered into eleven community types. Clear gradients in salinity and flood duration existed among the community types, and geographic zonation of the communities was evident across the landscape. At some locations assemblages were temporally stable; at other locations, they varied considerably. Samples not used in training the network were effectively projected onto the SOM and assigned to one of the delineated community types.

Conclusions: The SOM was effective in delineating plant communities in the region that were qualitatively similar to those obtained in previous investigations. Being robust to skewed distributions and the presence of outliers, SOMs provide an alternative to traditional distribution-based statistical approaches. Their ability to efficiently classify new data into existing community types makes their use an ideal approach to classifying samples obtained from ongoing, long-term ecological monitoring programs.

KEYWORDS
artificial neural network, classifying vegetation, northern Gulf of Mexico, self-organizing maps, wetlands

1 | INTRODUCTION

Under existing climate change scenarios, coastal wetland ecosystems are anticipated to experience profound impacts to coastal hydrology, which may drive shifts in marsh vegetation species composition (Callaway, Parker, Vasey, & Schile, 2007; Howard, Biagas, & Allain, 2016; Janousek & Folger, 2014; Snedden & Steyer, 2013). These impacts include altered salinity and inundation regimes associated with either sea-level rise (Bhuiyan & Dutta, 2012; Hilton, Najjar, Zhong, & Li, 2008) or changes in freshwater delivery brought about by altered precipitation regimes (Robins et al., 2016). The ability to objectively and efficiently classify coastal wetland vegetation communities can be essential to quantifying how zonation responds to changing environmental conditions, and is thus fundamental to successful coastal wetland ecosystem management and restoration in the face of climate change.

The wetlands in coastal Louisiana along the northern Gulf of Mexico account for roughly 40% of all the wetlands in the continental United States and have exhibited high rates of land loss over the last century, during which roughly 5,000 km^2 of land has transitioned to open water (Couvillion, Beck, Schoolmaster, & Fischer, 2017) largely due to a combination of submergence and diminished sediment supply (Penland et al., 2001). The region is now the subject of a globally unprecedented ecosystem restoration program estimated to cost $50–100 billion over the next 50 years (Peyronnin et al., 2017). Plant communities often reflect ecosystem processes and conditions at local and landscape scales more effectively than any other set of factors (Albert & Minc, 2004), and different communities can influence processes such as hydrodynamics (Leonard & Luther, 1995), vertical accretion (Pasternack & Brush, 2001), and carbon burial (Wang, Xu, & Rongrong, 2016) in distinct ways. As such, substantial efforts to monitor the zonation of marsh vegetation communities in coastal Louisiana and how it may respond to restoration activities and continued disturbance such as sea-level rise and tropical storm impacts have been in place for over 20 years (Folse et al., 2014; Steyer, Raynie, Steller, Fuller, & Swenson, 1995) and are expected to continue (Hijuelos & Hemmerling, 2016).

Previous efforts to visualize gradients in vegetation community composition in the region (Snedden & Steyer, 2013; Visser, Sasser, Chabreck, & Linscombe, 1998; Visser, Sasser, Linscombe, & Chabreck, 2000) have all relied on parametric, distribution-based multivariate statistical approaches. These approaches (two-way indicator species analysis, canonical correspondence analysis), along with other commonly used eigen-based analytical approaches to ordination such as principal components analysis (PCA) and redundancy analysis, all make assumptions regarding the shape of the species-abundance response along environmental gradients (e.g., linear vs. unimodal; Jongman, ter Braak, & van Tongeren, 1995), and the nature of these responses is often unknown or varies depending on the species in question. Additionally, these techniques perform best when species distributions are not strongly skewed (Legendre & Birks, 2012). Yet species-abundance datasets typically contain many zeros that arise from the occurrence of rare species or those with narrow realized environmental niches, and they may remain strongly skewed even after transformation by logarithmic or other functions (Lek et al., 1996; Brosse, Giraudel & Lek, 2001; Quinn & Keough, 2002).

Over the last two decades, a multitude of artificial neural network (ANN) approaches have gained appeal as an alternative to classical statistical methods for analyses of multidimensional data (Chon, 2011). Among ANNs, self-organizing maps (SOMs; Kohonen, 2001) have become increasingly popular for the analysis of large ecological datasets, including studies of diatoms (Tison et al., 2005), fish (Chen, Lek, Lek-Ang, & Zhongjie, 2012; Park, Grenouillet, Esperance, & Lek, 2006a), benthic macroinvertebrates (Bae, Chon, & Park, 2012; Lencioni, Marziali, & Rossaro, 2012; Li, Xiang, & Li, 2015), birds (Lee, Kwak, Lee, & Kim, 2007), forest vegetation (Adamczyk, Kurzak, Park, & Kruk, 2013), and water quality (Chea, Grenouillet, & Lek, 2016). Unlike parametric statistical approaches, SOMs make no assumptions related to distributions of variables or correlations between those variables (Giraudel & Lek, 2001). Essentially, SOMs are methods for non-linear ordination analysis in that they provide a non-linear projection of the training dataset onto fewer (usually two) dimensions that approximates its probability density function (Kohonen, 2001). This two-dimensional projection consists of a network of cells, or neurons, each described by a vector that corresponds to a discrete region of the training dataset's multidimensional space. After a SOM is trained, traditional clustering algorithms can be applied to the SOM results to delineate discrete community types. At this point, new samples can be projected onto the clustered SOM to classify them into the existing community types without altering the existing clustering scheme. In this fashion, SOMs are a particularly attractive approach for classifying samples obtained from ongoing, long-term ecological monitoring programs (Park, Chon, Kwak, & Lek, 2004), and they allow for the verification of predictive model outcomes as data become available to do so. In this study, a SOM was trained with marsh vegetation species abundance data obtained from a long-term regional coastal wetland monitoring program in Louisiana, USA, as an unsupervised machine-learning approach for classifying coastal wetland vegetation communities in the region.

2 | METHODS

2.1 | Study area and data collection

Vegetation cover data used in this study were obtained from 343 sites selected from the Coastwide Reference Monitoring System (CRMS, www.lacoast.gov/crms; Figure 1—see Supporting information Appendix S1 for a list of sites and years used for training), a statewide network of wetland monitoring stations distributed across coastal Louisiana that has been in place since 2006. Visual surveys of emergent marsh vegetation percent cover are conducted at ten 2 m × 2 m permanent vegetation plots situated along
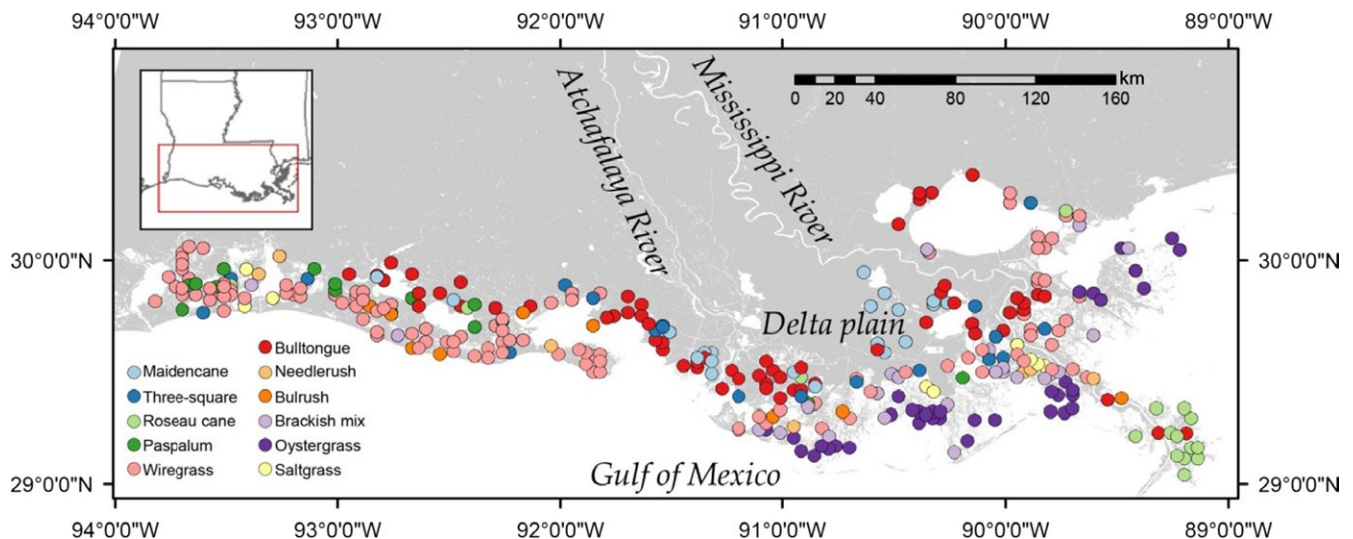
FIGURE 1 Location of 343 sites where vegetation cover, salinity, and water level were collected. Color indicates vegetation community type, as classified by the self-organizing map, based on taxa composition observed during the 2013 Coastwide Reference Monitoring System (CRMS) vegetation survey [Colour figure can be viewed at wileyonlinelibrary.com]

a 280-m transect at each site annually (July–September; Folse et al., 2014). Plant taxonomic nomenclature followed the USDA PLANTS Database (USDA, NRCS, 2019).

At each of the 343 sites, a logging instrument installed within 150 m of the vegetation transects recorded salinity and water surface elevation data at hourly intervals. For this study, it was assumed that salinity and water surface elevation did not vary over this distance and thus, values measured at each site were assumed to be identical to those in all vegetation plots at each site. Marsh surface elevation at each vegetation plot was surveyed in 2013 to the same vertical datum (North American Vertical Datum 1988) as the site's corresponding water level recorder. With these data, average annual salinity and percent time inundated were calculated for the time period October 2013–October 2014. Percent time inundated was calculated as the percent of time in which water surface elevation exceeded the marsh elevation of each vegetation plot.

2.2 | Modelling methods

Self-organizing maps are rooted in an unsupervised competitive learning process (Kohonen, 2001), and thus, they make no prior assumptions regarding the distribution of the data, nor do they require predefined target classes (community types in this case). Rather, training of the network and the resulting classes is entirely data-driven, whereby the network learns to classify by identifying patterns among the samples in the training data set. The SOM consists of two layers—an input layer presented to the network as a matrix **X** with rows corresponding to samples n and columns corresponding to variables p, and an output layer (or map) formed by nodes (or map units) arranged in a two-dimensional grid. For this study, each input node is a measure of the abundance of each taxon, and the output layer is composed of M output nodes arranged in a hexagonal lattice. A hexagonal, as opposed to a square, lattice is typically used because

it maximizes the number of neighbors for each node in the map. The number of map units along the vertical M_{vert} and horizontal M_{horz} axes of the map are typically set such that their ratio approximates that of the first two PCA eigenvalues obtained from the training data and their product $M_{vert} \times M_{horz} = M$ is approximately equal to $5\sqrt{N}$ (Vesanto & Alhoniemi, 2000), where N is the number of samples in the training set ($n = 1, 2, ..., N$). A few other parameters of the training algorithm must be predetermined (e.g., training iterations, neighborhood size, learning rate), but no further assumptions or decisions need to be made prior to training.

The relative cover values used to train the SOM were obtained from annual cover surveys conducted between 2006 and 2014 at the 343 selected sites (Figure 1) and calculated from species cover averaged across all plots within each site. This training data set consisted of 2,526 samples and 559 species, and the learning process was conducted using the SOM Toolbox (Alhoniemi, Himberg, Parhankangas, & Vesanto, 2000; http://www.cis.hut.fi/projects/somtoolbox/) in Matlab. Because SOMs make no distributional assumptions about the data (Kohonen, 2001), transformations are unnecessary and were not applied here. The training phase begins by first associating a weight vector w_m composed of P elements ($p = 1, 2, ..., P$; $m = 1, 2, ..., M$) where P is the number of variables (taxa) in the data set and M is the number of map units, to each map unit such that each map unit can be conceptualized as a virtual sample whose species' relative abundances are indicated by its weight vector (Figure 2). Here, initial values for each weight vector were assigned with the linear initialization algorithm (Kohonen, 2001), where PCA is first performed on the training data and the first two eigenvectors are interpolated in two dimensions across the map. Linear initialization ensures that, prior to training, the initial values of the weight vectors roughly approximate the probability density function of the training data, allowing for more rapid convergence during the training phase. The collection of the
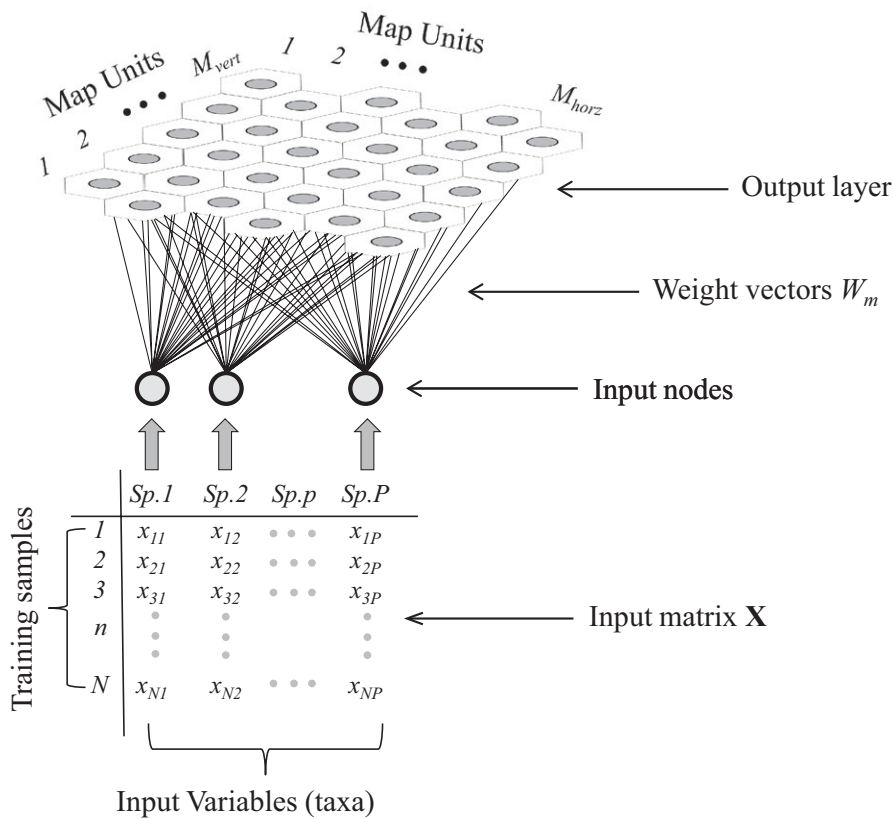
FIGURE 2 Schematic diagram of the structure of the self-organizing map, modified from Chon, Park, Moon, and Cha (1996)

pth element of all vectors in the map specifies the pth component plane of the map, which represents the distribution of the pth species' relative cover across the SOM. As the first step in the learning procedure, a sample x_n is drawn from the training dataset and compared to each map unit. The weight vector from the map unit c with the lowest Euclidean distance from the training sample is designated the best matching unit (BMU). The elements in the BMU's vector are altered to become more similar to the training sample's vector x_n, as are (to a lesser degree) those of neighboring map units according to the learning function:

$$w_{pm}(t+1) = w_{pm}(t) + h_{cm}(t)\left[x_{pn}(t) - w_{pm}(t)\right] \qquad (1)$$

Here h_{cm} is a Gaussian neighborhood function, or smoothing kernel

$$h_{cm} = \alpha(t)\exp\left(\frac{(\|r_m - r_c\|)^2}{2\sigma^2(t)}\right) \qquad (2)$$

where α is a learning rate factor and σ defines the width of the kernel, both of which decrease with increasing number of learning cycles t, and $\|r_m - r_c^2\|$ is the distance on the map between the BMU and the mth map unit. This procedure is repeated for all training samples, and then the entire learning process is repeated several times until the results converge.

The importance of each taxon in structuring the SOM was assessed with the global structuring index (GSI; Park, Gevrey, Lek, & Giraudel, 2005), computed as

$$\text{GSI}_p = \sum_{m=1}^{M}\sum_{j=1}^{m-1}\frac{\left|w_{pj} - w_{pm}\right|}{\|r_j - r_m\|} \qquad (3)$$

where $\|r_j - r_m\|$ is the distance on the map between the map units j and m. Thus, high GSI values indicate strong gradients in a given taxon's relative abundance across the map. Many taxa in species-abundance datasets are rare, and as such, they often add noise and provide little information to community analysis (Gauch, 1982; McCune & Grace, 2002). However, simply being rare does not necessarily preclude their informative value to ordination. Here the importance of taxa to ordination was assessed with GSI scores and used as the basis for inclusion in the SOM. First, GSI scores were calculated for a SOM trained with all taxa present in the training dataset, and values of all GSI scores were summed. This process was then repeated with fewer species, each time discarding the species with the lowest GSI score until only ten species remained in the model. Subsequently the differences of the summed GSI values between the original dataset (full model with all taxa) and reduced datasets were calculated, and a reduced SOM that struck a balance between minimizing this difference and reducing dimensionality (as indicated by the presence of a local minimum in a plot of number of species vs summed GSI scores of full and reduced species models) was selected (see Park et al., 2006b).

By identifying BMUs for each training sample, SOMs initially classify them into one of M groups projected onto discrete regions (map units) of a two-dimensional plane. Where M is large, additional compression into a smaller number of classes ($q = 1, ..., Q$) is often

desired. Here, a hierarchical cluster analysis was applied to the weight vectors of the SOM map units with Ward's linkage based on Euclidean distance. Rather than impose prior assumptions to the number of clusters obtained in the analysis, the samples were classified by varying the number of clusters from 2 to 20, and the classification scheme that produced the greatest mean silhouette statistic (Kaufman & Rousseeuw, 1990) was selected. At this point, dendrogram branches containing no significant indicator species (obtained from indicator species analysis; see below) were combined with neighboring branches into single groups (Abraão et al., 2010; Salovaara, Cárdenas, & Tuomisto, 2004). Between-cluster differences in environmental variables were evaluated with the Kruskal–Wallis test (non-parametric analysis of variance) followed by a post-hoc Tukey test for multiple comparisons (Zar, 1984).

Characteristic taxa for each community type were identified with indicator species analysis (Dufrêne & Legendre, 1997), where indicator values (IV) are calculated as

$$IV_{qp} = RA_{qp} \times RF_{qp}$$

where RA_{qp} is the relative abundance of species p in class q, computed as the mean relative cover of species p in class q, divided by the sum of mean relative cover values of species p across all classes. RF_{qp} is the relative frequency of species p in class q, defined as the proportion of samples in the qth class that contain species p. Thus, RA_{qp} is a measure of species specificity to a particular class, whereas RF_{qp} is a measure of fidelity. The maximum IV of 100 occurs when individuals of species p only occur in samples assigned to group q, and all samples classified as q contain species p. After computing IVs, a Monte Carlo significance test of observed maximum indicator values for a given species was applied with 999 permutations.

Samples collected under separate research or monitoring programs with compatible methods that were not used in model training were then projected onto the trained SOM for classification. For this exercise, high-density, lower-frequency helicopter-based vegetation surveys conducted at ~4,000 sites across coastal Louisiana every 5–10 years (see Chabreck, Linscombe, Hartley, Johnston, & Martucci, 2001; Linscombe & Hartley, 2011; Sasser, Visser, Mouton, Linscombe, & Hartley, 2008, 2014) were projected onto the SOM to facilitate delineation of the geographic zonation of the eleven community types at a landscape scale.

3 | RESULTS

A reduced SOM model trained with the 49 most important taxa (i.e., highest GSI scores; Table 1) was selected and used for all further analysis (see Supporting information Appendix S2). The SOM contained 260 map units arranged across a 20 × 13 grid, and BMUs for the 2,526 training samples were distributed broadly across the SOM (all but 12 map units were BMUs for at least one of the training samples; Figure 3). The dendrogram obtained from hierarchical cluster analysis was cut at 15 groups to maximize the mean silhouette width (0.37), after which branches containing no significant indicator

TABLE 1 The 49 taxa retained to train the reduced self-organizing map, along with their abbreviations and global structuring indices (GSIs)

Taxa	Abbreviation	GSI
Spartina patens	SPPA	1,093
Spartina alterniflora	SPAL	1,002
Distichlis spicata	DISP	572
Phragmites australis	PHAU	463
Juncus roemerianus	JURO	246
Bolboschoenus robustus	BORO	218
Sagittaria lancifolia	SALA	209
Schoenoplectus americanus	SCAM	179
Panicum hemitomon	PAHE	159
Paspalum vaginatum	PAVA	158
Polygonum punctatum	POPU	136
Alternanthera philoxeroides	ALPH	107
Schoenoplectus californicus	SCCA	99
Ipomoea sagittata	IPSA	82
Typha spp.	TYPHA	78
Eleocharis spp.	ELEOC	77
Leersia hexandria	LEHE	77
Typha latifolia	TYLA	77
Lythrum lineare	LYLI	68
Spartina cynosuroides	SPCY	67
Echinochloa walteri	ECWA	64
Iva frutescens	IVFR	58
Sacciolepis striata	SAST	52
Typha domingensis	TYDO	51
Avicennia germinans	AVGE	50
Cladium mariscus	CLMA	49
Zizaniopsis miliacea	ZIMI	48
Ludwigia grandiflora	LUGR	48
Colocasia esculenta	COES	43
Thelypteris palustris	THPA	41
Hydrocotyle umbellata	HYUM	37
Amaranthus australis	AMAU	36
Eleocharis macrostachya	ELMA	34
Bacopa monnieri	BAMO	33
Baccharis halimifolia	BAHA	32
Sagittaria latifolia	SALA2	31
Batis maritima	BAMA	31
Cicuta maculate	CIMA	30
Cyperus odoratus	CYOD	29
Symphyotrichum tenuifolium	SYTE	28
Eleocharis cellulosa	ELCE	28
Paspalum distichum	PADI6	27
Panicum dichotomiflorum	PADI	25

(Continues)

TABLE 1 (Continued)

Taxa	Abbreviation	GSI
Sagittaria platyphylla	SAPL	23
Leptochloa fusca	LEFU	20
Morella cerifera	MOCE	20
Bidens laevis	BILA	18
Schoenoplectus deltarum	SCDE	10
Nelumbo lutea	NELU	9

species were combined with neighboring branches into single groups to ultimately delineate eleven distinct community types (Figure 3) with a mean silhouette width of 0.42. The twelve map units that were not identified as BMUs for any of the training samples all fell along community type boundaries in the SOM, essentially corroborating results of the cluster analysis.

The component planes (Figure 4; Supporting information Appendix S3) depict the distribution of each component (vegetation taxa) in each SOM map unit and visually convey information regarding the specificity and fidelity of each species in the SOM to each of the delineated community types. Most of the taxa with high GSI scores (those most important in structuring the SOM) were clearly associated with single clusters. For example, *Spartina patens* was most prevalent in map units encompassed in the wiregrass cluster (Figure 3), even though it was present in smaller abundances in map units belonging to several other clusters. Other species, such as *Ipomoea sagittata*, span a wide region of the SOM and are thus present in a variety of community types, albeit possibly in small abundances. The component planes also confer qualitative indications regarding correlations among species. For example, some species tend to co-occur (e.g., *Sagittaria lancifolia* and *Alternanthera philoxeroides*) whereas others are rarely found together (e.g., *Spartina*

alterniflora and *Phragmites australis*). Some community types are dominated by different species, depending on which region of the SOM cluster a sample projects onto. For example, the upper region of the bulltongue cluster is dominated by *Sagittaria lancifolia*, whereas the lower region is dominated by *Polygonum punctatum* (with *Sagittaria lancifolia* present in smaller abundance).

3.1 | Dominant/indicator species and geographic trends of vegetation communities

The majority of the eleven community types delineated was characterized by multiple indicator species; however, five communities, wiregrass, needlerush, brackish mix, oystergrass, and saltgrass, had only one indicator species (Table 2). The species with the highest percent relative cover was an indicator species in all communities with the exception of the brackish mix community (Table 3). For the following community descriptions, geographic zonation is shown in Figure 1, and commonly occurring species were determined from relative frequency values in Table 2.

Maidencane communities are dominated by *Panicum hemitomon* (IV 86.4) and typically found in the inland extremities of the interdistributary basins of the Mississippi River delta plain, often near bulltongue sites (see below). Other species common to this community type include *Leersia hexandria*, *Thelypteris palustris*, *Alternanthera philoxeroides*, *Polygonum punctatum*, and *Sagittaria lancifolia*.

Three-square communities are dominated by *Schoenoplectus americanus* (IV 55.6) with *Spartina patens* occurring in these communities 91% of the time. Three-square sites are often located just inland from wiregrass communities. Other common species for this community include *Sagittaria lancifolia* and *Lythrum lineare*.

Roseau cane communities are dominated by *Phragmites australis* (IV 85), and often co-occur with *Spartina patens*, *Alternanthera*
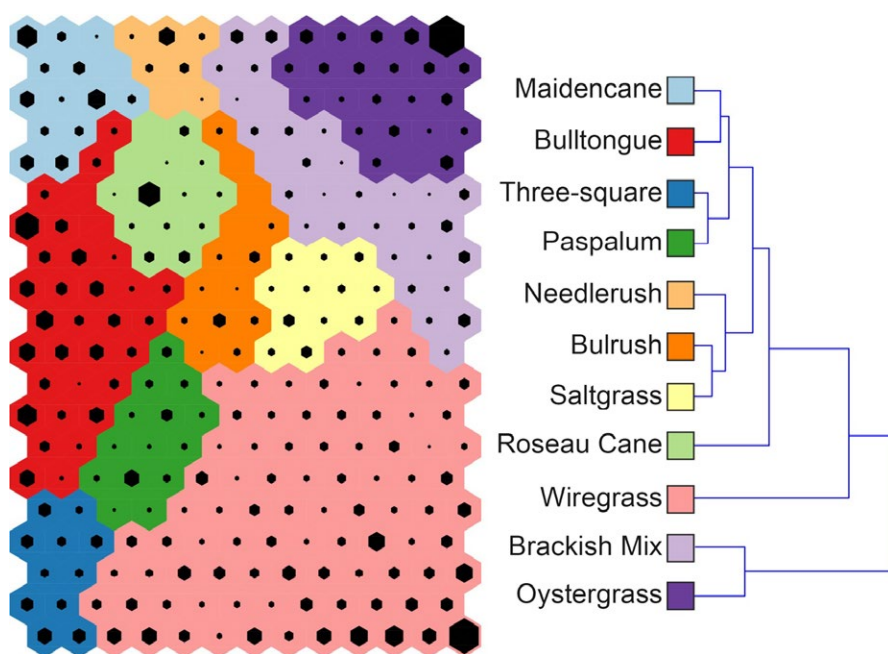

FIGURE 3 (left) Classification of the training samples according to the self-organizing map (SOM). The size of the black inset hexagon in each SOM unit is proportional to the number of samples assigned to that unit. The largest inset hexagon (upper right corner) represents 128 samples. Based on hierarchical cluster analysis with Ward's linkage performed on SOM output (right), SOM units were classified into 11 clusters (vegetation community types) [Colour figure can be viewed at wileyonlinelibrary.com]
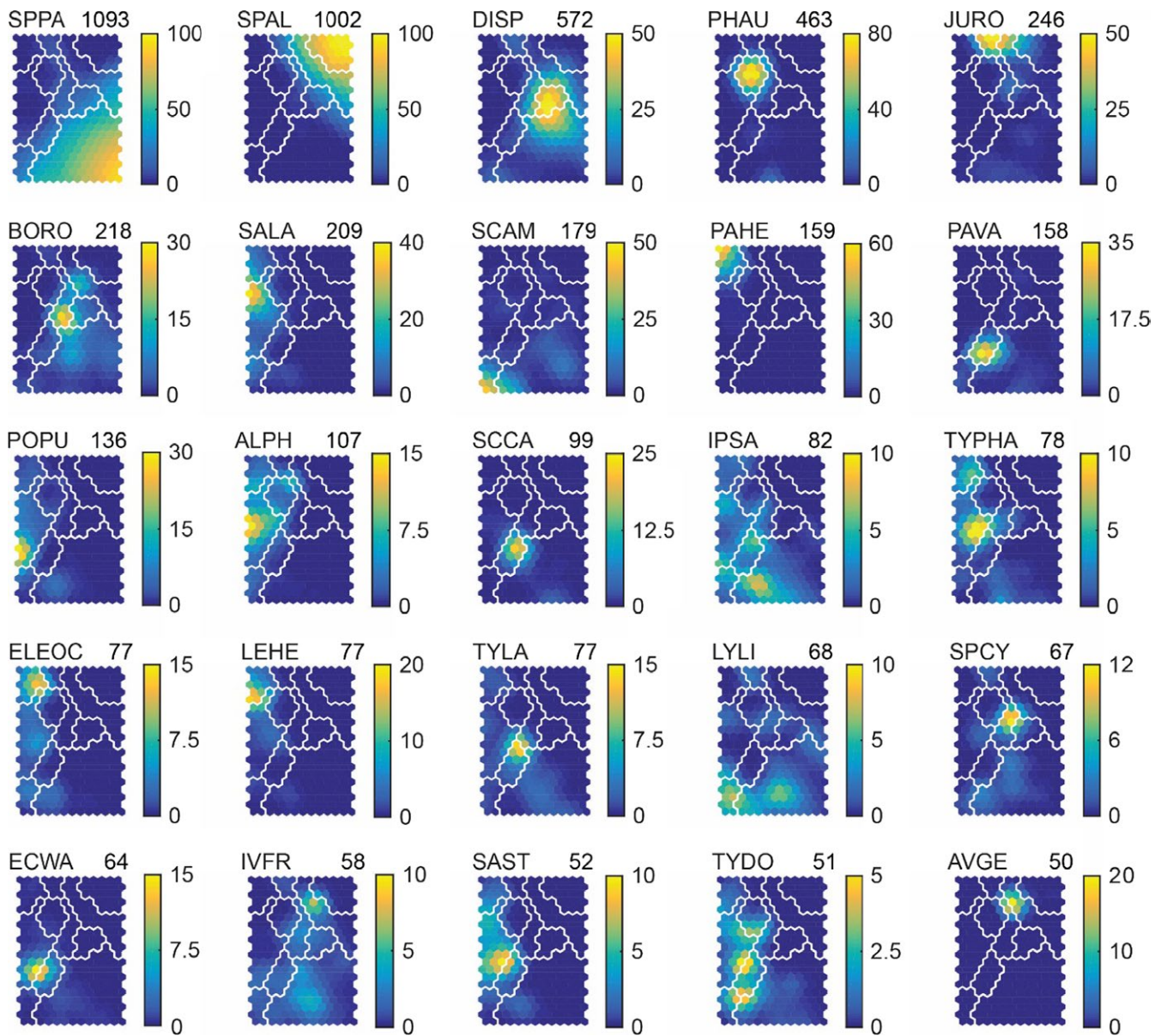
FIGURE 4 Component planes for the 25 taxa most important in structuring the self-organizing map (SOM). Each plane depicts the distribution of relative abundances for each particular taxon across the SOM. White lines delineate vegetation community types as depicted in Figure 3. Global structuring index (GSI) values are shown in the upper right-hand corner. Component planes for the remaining 24 taxa are presented in Appendix S3 [Colour figure can be viewed at wileyonlinelibrary.com]

philoxeroides, and *Polygonum punctatum*. Roseau cane communities are most prevalent in the marshes at the mouth of the Mississippi River.

Paspalum communities are dominated by *Paspalum vaginatum* (IV 52.4). Other species common to Paspalum communities include *Schoenoplectus californicus*, *Spartina patens*, *Typha latifolia*, *Ipomoea sagittata*, *Distichlis spicata*, and *Echinochloa walteri*. Paspalum sites are most prevalently located near large inland lakes along the western reaches of coastal Louisiana (Figure 1).

Wiregrass communities are dominated by *Spartina patens* (IV 39.0). Other common species for this community type include *Schoenoplectus americanus*, *Lythrum lineare*, *Ipomoea sagittata*, *Bolboschoenus robustus*, and *Distichlis spicata*. Wiregrass is perhaps the most widespread of all marsh community types across coastal

Louisiana. The relatively low IV for *Spartina patens* in wiregrass communities is a reflection of the low RA of this species, resulting from the fact that it is present at all of the eleven community types. However, the species was present at all sites classified as wiregrass.

Bulltongue communities are dominated by *Sagittaria lancifolia* (IV 32.4), and often contain *Polygonum punctatum*, *Ipomoea sagittata*, and *Alternanthera philoxeroides*. These sites are most prevalent along the inland reaches of the coastal zone, often intermixed with maidencane communities.

Needlerush communities are dominated by *Juncus roemerianus* (IV 70.1) and are generally co-dominated by varying combinations of *Distichlis spicata*, *Spartina alterniflora*, *Bolboschoenus robustus*, and *Spartina patens*. Needlerush sites are most prevalent toward the

TABLE 2 Indicator species for community types, indicator values (IV), relative abundance (RA) and relative frequency (RF) by community types for the 49 species included in the SOM analysis

Community	Taxon	IV	% RA/% RF by community type										
			MC	TS	RC	PA	WG	BT	NR	BR	BM	OG	SG
Maidencane	PAHE	**86.4**	**93/93**	1/15	0/1	0/5	0/1	5/28					
	LEHE	**53.1**	**73/73**	5/13	1/2	8/6	1/2	13/35					
	THPA	**41.7**	**75/55**	8/10	0/1		0/1	17/18					
	ELEOC	**33.7**	**59/57**	17/18	0/1	1/8	3/6	20/38		0/2			
	HYUM	**32.8**	**48/69**	14/27	0/1	3/11	0/2	35/41					
	TYPHA	**16.9**	**33/51**	6/19	0/6	4/8	4/8	31/36	0/6	16/11			5/1
	SALA2	**15.8**	**43/37**	3/5	14/7			40/21					
	MOCE	**15.4**	**50/31**	7/4	2/2	1/1	1/1	34/11	1/2	4/5			
	BILA	**12.9**	**48/27**	0/1	0/1	1/1	1/0	50/17					
Three-square	SCAM	**55.6**	1/12	**66/84**	3/11	4/18	12/48	4/30	0/6	3/23	2/10	0/1	4/19
	LYLI	**20.1**	0/4	**37/54**	5/7	1/10	16/40	9/20	19/15	8/11	3/11	0/1	2/11
	ELCE	**10.5**	12/16	**45/23**	1/2	6/6	2/5	28/27		4/2	2/2		
	ELMA	**8.7**	13/9	**49/18**	6/8	7/4	1/2	19/13	5/4	0/2	0/1		
	CLMA	**7.6**	0/2	**60/13**	1/1	4/6	4/5	24/16	0/2	7/5			
Roseau Cane	PHAU	**85.0**	0/6	1/13	**85/100**	3/19	2/13	1/8	3/15	5/23			1/5
	SAPL	**9.7**	0/1		**35/27**			60/5		5/4			
	SCDE	**2.0**			29/7			71/2					
Paspalum	PAVA	**52.4**	0/4	7/25	0/1	**83/63**	5/19	1/6		1/5	1/7		2/7
	SCCA	**43.5**	2/8	1/14	0/2	**89/49**	4/10	3/12		1/9			
	TYLA	**31.6**	9/23	6/13	1/2	**66/48**	6/13	8/16		5/2			
	LEFU	**22.4**		27/5		**65/35**	1/2	6/3					0/1
	ECWA	**19.3**	4/10	4/15	5/12	**42/46**	5/14	37/27		4/14	0/1		1/1
	CYOD	**16.7**	5/8	8/18	2/6	**47/36**	4/12	31/22		2/9	0/1		1/3
	AMAU	**12.9**	2/10	26/50	6/13	**27/48**	11/31	14/28	1/15	4/21	1/6		8/27
	IPSA	**10.4**	8/46	18/53	4/14	**24/44**	13/43	16/52	9/27	7/21	0/1		0/9
	BAMO	**8.7**	6/11	26/19	1/2	**28/31**	7/14	30/25		0/1	0/1		1/3
	TYDO	**4.2**	10/7	19/13	24/13	22/19	6/7	19/14		0/2			
Wiregrass	SPPA	**39.0**	0/13	12/91	3/32	7/73	**39/100**	1/30	5/85	8/91	12/70	1/17	13/96

(Continues)

TABLE 2 (Continued)

Community	Taxon	IV	% RA/% RF by community type										
			MC	TS	RC	PA	WG	BT	NR	BR	BM	OG	SG
Bulltongue	POPU	**42.9**	15/73	9/41	9/22	2/5	4/13	**56/77**	0/2	4/14	0/2		0/1
	SALA	**32.4**	27/75	16/45	1/15	3/20	1/11	**45/72**	3/8	3/20			0/1
	SAST	**26.4**	22/53	11/25	0/1	16/21	2/4	**48/56**		1/5			
	ALPH	**26.3**	21/61	10/40	20/41	5/24	1/7	**36/73**		8/14			
	LUGR	**20.2**	16/22	4/4		0/1		**79/25**					
	COES	**13.4**	26/15	0/2	19/12			**54/25**		1/2			
	CIMA	**10.2**	22/12	12/6	0/1		0/1	**66/15**					
	ZIMI	**9.1**	16/22	1/5	25/17	19/13	0/1	**38/24**		1/4			
	BAHA	8.2	6/10	12/17	5/4	8/5	8/12	**44/19**	1/6	11/11	0/1		4/4
	NELU	3.5	12/4			1/1		**88/4**					
Needlerush	JURO	**70.1**	0/1	1/17	0/7	0/12	2/20	0/4	**70/100**	6/29	13/35	5/42	1/13
Bulrush	BORO	**49.9**	0/2	3/31	2/18	5/45	6/50	1/8	5/54	**56/89**	7/21	0/4	15/55
	SPCY	**24.5**	1/2	3/7	1/4		6/8	3/6	16/35	**62/39**	0/2		7/11
	PADI6	**15.3**	0/1	1/3	0/1	5/4	2/1	6/2		**86/18**			1/1
	BAMA	**10.8**	0/1			3/7	2/5	0/1		**50/21**	17/8	7/5	20/16
	SYTE	6.7	0/4	17/38	1/8	15/32	9/29	13/20	13/48	19/36	5/20	1/4	7/27
	IVFR	5.4	0/1	6/17	8/8	10/30	12/23	6/10	10/25	20/27	16/9		13/35
	PADI	1.8	0/1	5/5	2/3		18/6	24/7	24/4	25/7			
Brackish Mix	AVGE	**7.7**									**96/8**	4/7	
Oystergrass	SPAL	**49.2**		0/10	2/22	0/7	1/18	0/4	8/85	4/63	28/100	**49/100**	8/71
Saltgrass	DISP	**50.8**	0/1	4/46	1/13	4/37	7/60	0/5	8/88	16/82	8/70	1/26	**51/100**

Notes. Values for %RA and %RF are rounded to the nearest whole percentage. Blank cells for %RA/%RF indicate values of 0 for %RA and %RF. IVs in bold font indicate species was a significant indicator of its community type for $\alpha = 0.01$. Community types are abbreviated as follows: BR: bulrush; BM: brackish mix; BT: bulltongue; MC: maidencane; NR: needlerush; OG: oystergrass; PA: paspalum; RC: roseau cane; SG: saltgrass; TS: three-square; WG: wiregrass.

TABLE 3 Mean % relative cover of the seven most abundant taxa for each community type

Maidencane		Three-square		Roseau Cane		Paspalum	
Panicum hemitomon	34	**Schoenoplectus americanus**	27	**Phragmites australis**	71	**Paspalum vaginatum**	24
Leersia hexandria	11	Spartina patens	19	Spartina patens	5	**Schoenoplectus californicus**	13
Sagittaria lancifolia	10	Sagittaria lancifolia	6	Alternanthera philoxeroides	4	Spartina patens	11
Eleocharis	7	**Lythrum lineare**	5	Spartina alterniflora	3	**Typha latifolia**	10
Thelypteris palustris	5	**Cladium mariscus**	4	Typha domingensis	2	**Ipomoea sagittata**	6
Alternanthera philoxeroides	4	**Eleocharis macrostachya**	4	Zizaniopsis miliacea	2	Distichlis spicata	3
Typha	4	Distichlis spicata	4	Polygonum punctatum	2	**Echinochloa walteri**	3
Wiregrass		**Bulltongue**		**Needlerush**		**Bulrush**	
Spartina patens	65	**Sagittaria lancifolia**	16	**Juncus roemerianus**	54	**Bolboschoenus robustus**	24
Distichlis spicata	7	**Polygonum punctatum**	11	Spartina alterniflora	15	Distichlis spicata	16
Schoenoplectus americanus	5	**Alternanthera philoxeroides**	7	Spartina patens	8	Spartina patens	13
Bolboschoenus robustus	3	**Ludwigia grandiflora**	4	Distichlis spicata	8	**Spartina cynosuroides**	8
Ipomoea sagittata	2	Typha	4	Lythrum lineare	2	Spartina alterniflora	7
Lythrum lineare	2	**Colocasia esculenta**	3	Phragmites australis	2	**Paspalum distichum**	5
Spartina alterniflora	2	**Sacciolepsis striata**	3	Bolboschoenus robustus	2	Juncus roemerianus	5
Brackish Mix		**Oystergrass**		**Saltgrass**			
Spartina alterniflora	53	**Spartina alterniflora**	93	**Distichlis spicata**	49		
Spartina patens	20	Juncus roemerianus	4	Spartina patens	21		
Juncus roemerianus	10	Spartina patens	1	Spartina alterniflora	15		
Distichlis spicata	7	Distichlis spicata	1	Bolboschoenus robustus	6		
Bolboschoenus robustus	3	Batis maritima	<1	Schoenoplectus americanus	2		
Avicennia germinans	3	Bolboschoenus robustus	<1	Iva frutescens	1		
Iva frutescens	2	Avicennia germinans	<1	Juncus roemerianus	1		

Note. Bold indicates species was a significant indicator for the group in question.

seaward extremities of the Mississippi River delta plain, often intermixed with brackish mix and oystergrass communities.

Bulrush communities are dominated by *Bolboschoenus robustus* (IV 49.9), often co-dominant with *Distichlis spicata*, *Spartina patens*, and *Spartina alterniflora*. Bulrush sites are mostly concentrated at the seaward reaches of southwest Louisiana (Figure 1).

Brackish mix communities are typically situated just inland from oystergrass communities (see below) along the seaward end of the Mississippi River delta plain (Figure 1), and are dominated by *Spartina alterniflora* (Tables 2, 3). Common co-occurring species of brackish mix communities include *Spartina patens*, *Juncus roemerianus*, and *Distichlis spicata*. Because the common co-occurring species in this community also strongly co-occur in other community types, brackish mix communities do not have a strong indicator species. The sole indicator species, *Avicennia germinans* (IV 7.7) was only present in eight percent of brackish mix sites, even though 96% of the occurrences of this species are found at sites classified as brackish mix communities (Table 2).

Oystergrass communities are dominated by *Spartina alterniflora* (IV 49.2; Tables 2, 3), and often co-occur with small coverages of *Juncus roemerianus*. Oystergrass sites are typically located at the seaward ends of the interdistributary basins of the Mississippi River delta plain (Figure 1).

Saltgrass communities are dominated by *Distichlis spicata* (IV 50.8), and typically are co-dominated by *Spartina patens*, *Spartina alterniflora*, and *Bolboschoenus robustus*. Saltgrass sites are sporadically intermixed with brackish mix and wiregrass communities.

3.2 | Variation in hydrologic variables

Salinity differed among the community types (χ^2 = 1494.32; *df* = 10; *p* < 0.001) and was highest for oystergrass communities and lowest for maidencane and bulltongue communities, which did not differ from each other (Figure 5). Wiregrass, bulrush, saltgrass, needlerush and brackish mix communities showed considerable overlap with respect to salinity, though wiregrass and brackish mix communities
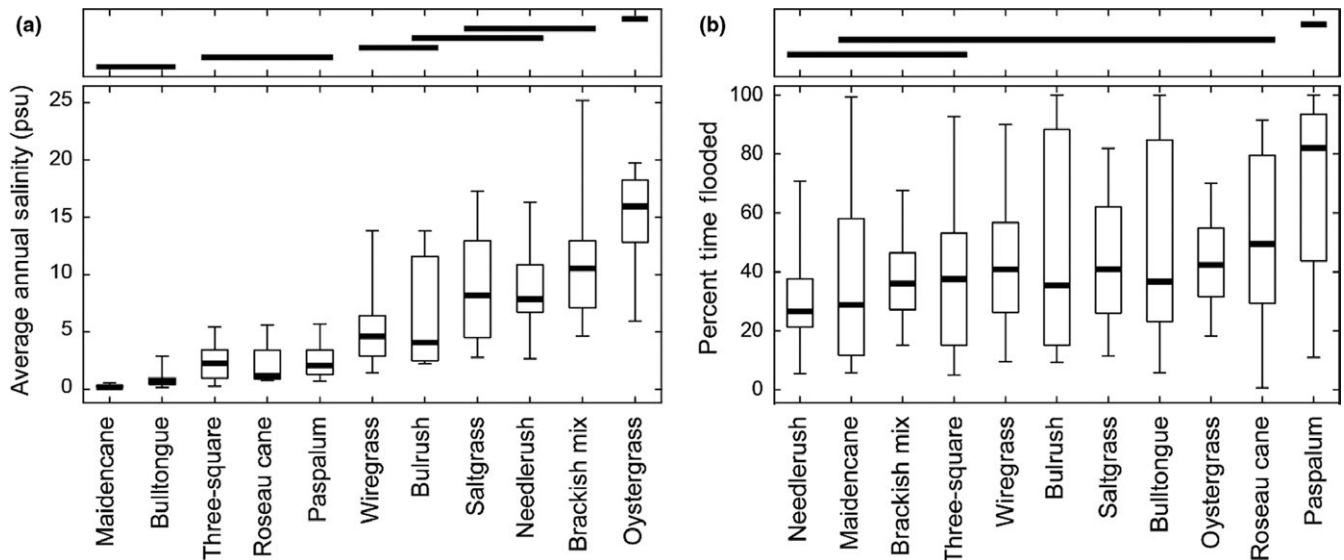
FIGURE 5 (lower) Boxplots indicating average annual salinity (a) and percent time flooded (b) for the 11 vegetation community types identified by clustering the self-organizing map (SOM). Community types are ordered from left to right by increasing rank sums. In each boxplot, thick horizontal line indicates median, box indicates interquartile range, and whiskers indicate maximum and minimum values. (upper) Community types sharing the same line did not differ statistically (α = 0.05) as determined by Tukey multiple comparisons
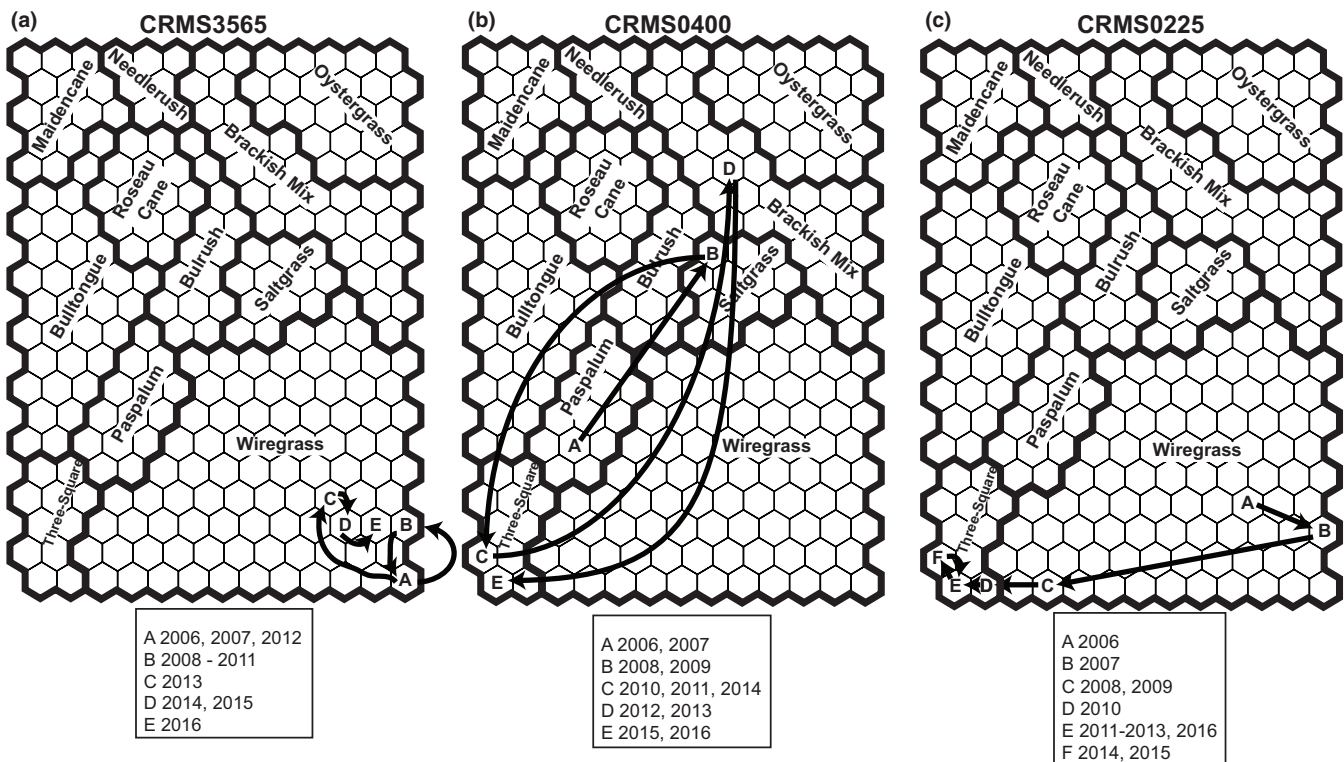
FIGURE 6 Examples of self-organizing map (SOM) temporal trajectories depicting annual variation in vegetation communities for CRMS3565 (a), CRMS0400 (b) and CRMS0225 (c). Samples obtained in 2015 and 2016 were not used in training the SOM but were projected onto the SOM after training, illustrating the ability of the trained SOM to classify new samples as they become available

showed clear separation from each other. Though percent time inundated varied among community types (χ^2 = 117.43; *df* = 10; *p* < 0.001), there was considerable overlap among the community types, with inundation time lowest in needlerush communities and highest in paspalum. All other community types grouped together with respect to inundation.
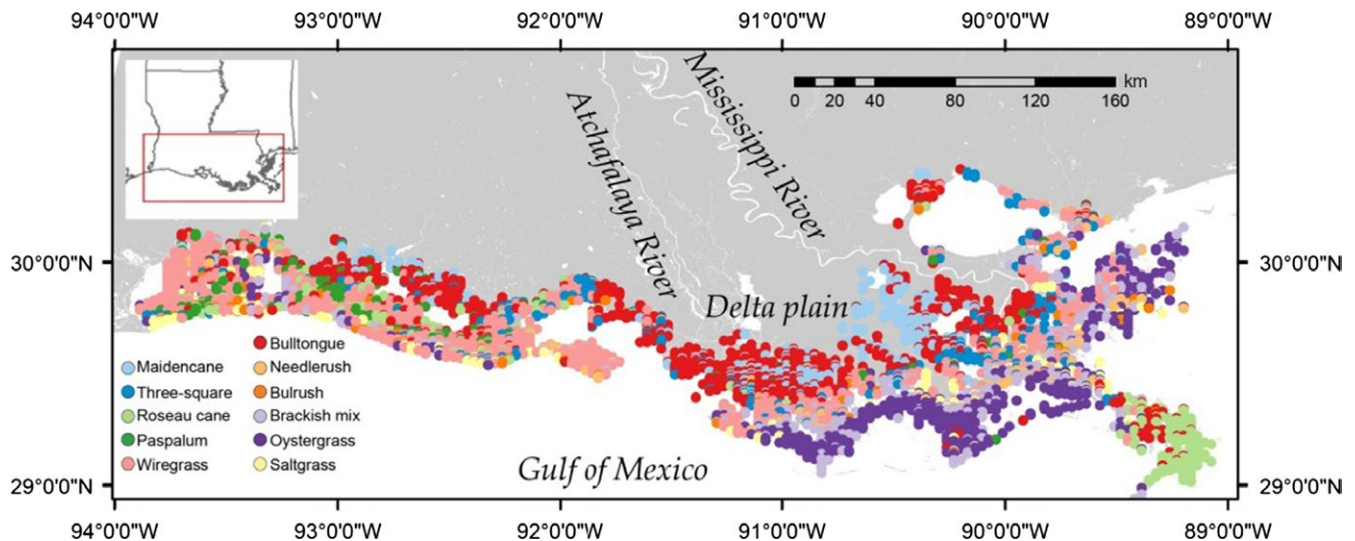
FIGURE 7 Self-organizing map classification of species cover data from 2013 helicopter-based vegetation surveys conducted at 4,215 locations across coastal Louisiana (Sasser et al., 2014) [Colour figure can be viewed at wileyonlinelibrary.com]

3.3 | Classifying new samples

Site data from 2015 and 2016, which were not used in the learning process, were classified as one of the established community types by projecting species vectors onto the SOM to determine BMUs. The BMU is determined simply by selecting the SOM map unit m for which the Euclidean distance between the weight vector w_m and the new sample in question is minimized. Figure 6 shows an example of assessing time-variation in community structure from three CRMS sites (CRMS3565, CRMS0400, CRMS0225; www.la-coast.gov/crms), from 2006 through 2016. Throughout this time period, CRMS3565 occupies the same cluster (wiregrass) and projects onto map units in a very limited region of the SOM, indicating a temporally stable community composition (Figure 6a). In contrast, CRMS0400 shows pronounced interannual shifts in community composition, spanning a wide range of community types (Figure 6b). CRMS0225 begins as a wiregrass community in 2006 and remains so through 2009, after which it transitions to a three-square community and remains so through 2016 (Figure 6c). The SOM enables visualization of these temporal trends and, in the case of CRMS0225 during 2008–2009, provides an indication that a community shift may be under way even though the species composition has not yet crossed the discrete boundary to a new community type. The SOM was also effective at classifying samples collected during the 2013 helicopter-based survey (Figure 7). Geographic zonation of the plant communities based on the helicopter survey is, qualitatively speaking, very similar to that obtained from the ground-based survey conducted in 2013 (Figure 1).

4 | DISCUSSION

Coastal emergent marsh vegetation assemblages of the northern Gulf of Mexico were patterned with a SOM according to distributional

similarities of species relative cover data from over 2,500 samples collected at nearly 350 sites spanning 8 years. Hierarchical cluster analysis, in conjunction with the silhouette statistic and indicator species analysis, delineated the SOM output into eleven distinct vegetation community types that, broadly speaking, resembled those put forth from previous efforts to delineate community types with conventional statistical approaches (Snedden & Steyer, 2013; Visser et al., 1998, 2000).

Salinity was clearly the most important hydrologic factor associated with the spatio-temporal variation of vegetation assemblages. The pronounced influence of salinity is manifest by the relatively small salinity interquartile ranges for each community type and the greater degree of separation in the Kruskal–Wallis multiple comparison groupings (Figure 5). The influence of salinity is also evident in the geographic zonation of the community types, whereby more salt-tolerant community types (e.g., oystergrass, brackish mix, saltgrass) are situated at the seaward end of estuarine basins far from mouths of major rivers; fresher community types (e.g., maidencane, bulltongue, roseau cane) tend to be located either further inland or in regions near strong fluvial influence (Figure 1). The overarching importance of salinity in structuring vegetation communities observed in this study is in agreement with numerous existing studies (Adams, 1963; Cooper, 1982; Ewing, 1983; Latham, Pearlstine, & Kitchens, 1994; Odum, 1988; Phleger, 1971). Qualitatively, the distribution of the eleven community types across the estuarine salinity gradient in the present study reflects findings of Snedden and Steyer (2013), though direct comparisons are challenging given the disparity in clustering approaches between the two studies.

The manner in which vegetation communities associated with inundation was less clear. Median percent time inundated for paspalum communities exceeded 80%, and was significantly greater than that of all other community types. Cantero, Cisneros, Zobel, and Cantero (1998) observed paspalum stands to be located in regions with high flood frequencies with durations exceeding 45 days.

Among the more saline community types in this study, there was a change in species composition from needlerush to brackish mix to oystergrass communities along a gradient of increased inundation, consistent with several previous investigations of vegetation zonation along elevation gradients in salt marshes (Eleuterius & Eleuterius, 1979; Latham et al., 1994; Niering & Warren, 1980; Pennings, Grant, & Bertness, 2005). Despite these subtle patterns of community zonation with respect to inundation, each of the eleven community types delineated in this study was, broadly speaking, present across a wide range of inundation regimes, as shown by the large interquartile ranges (Figure 5). Yet previous investigations have clearly shown the importance of inundation to primary production in emergent marsh macrophytes, particularly belowground, where increased inundation leads to diminished production for a variety of species that were dominants for the community types identified in this study, including *Spartina alterniflora* (oystergrass and brackish mix communities; Voss, Christian, & Morris, 2013; Snedden, Cretini, & Patton, 2015), *Juncus roemerianus* (needlerush communities; Voss et al., 2013), *Spartina patens* (wiregrass communities; Kirwan & Guntenspergen, 2015; Snedden et al., 2015), *Schoenoplectus americanus* (three-square communities; Schile, Callaway, Suding, & Kelly, 2017), and *Sagittaria lancifolia* (bulltongue communities; Visser & Sandy, 2009; Visser & Peterson, 2015). If decreased production is an ubiquitous response to increased inundation among the dominant species of several community types, decreased total biomass, rather than strong shifts in species composition, may be the primary impact of increased flooding. Thus, many assemblages observed here may often persist at low elevations corresponding to sub-optimal production, possibly due to the occurrence of competitive interactions at higher, less stressful elevations (Bertness, 1991a,b; Pennings et al., 2005).

The two-step classification procedure described here, whereby traditional cluster analysis is applied to SOM output is, in multiple ways, a robust alternative to classical parametric statistical approaches for examining gradients in vegetation species composition. The efficacy of this two-step approach can be compared to results obtained with hierarchical agglomerative clustering applied directly to the Bray–Curtis similarity matrix of the full (559 taxa) dataset by comparing mean silhouette widths between the two approaches. The mean silhouette width for the eleven-class SOM-based approach with the 49 selected taxa was 0.42, compared with 0.35 for the direct clustering approach on the full (559 taxa) species composition matrix with seven groups (seven groups produced a local maximum in mean silhouette width for this approach), and 0.29 when the dendrogram is cut at eleven groups (Supporting information Appendix S4), indicating comparable, possibly even slightly increased performance, for the SOM-based approach.

Species abundance datasets typically contain many rare taxa that are absent in the vast majority of samples. These absences, represented by zeros, can make their distributions strongly skewed and difficult to correct for with any transformation (98% of the values of the original training dataset in this study were zeros). One common approach to mitigating this problem has been to eliminate taxa that

are not present in some arbitrary minimum of percentage samples (e.g., 5%; Gauch, 1982). In addition to its arbitrary nature, this approach can be problematic in that common taxa occurring at nearly all sites may be selected for inclusion in the analysis (even though they minimally inform the ordination) whereas others that are abundant at relatively few sites may be excluded (even if they are strong indicators of a particular community type; Poos & Jackson, 2012). The method presented here is not hindered by these issues because the GSI, the basis for selecting taxa for inclusion in the reduced model, indicates species contribution to the overall organization of the SOM, by quantifying the gradient in species abundance across a given topological distance in the SOM. Had the 5% rule of thumb been used as the criterion for taxa selection in the present analysis, 20% of the taxa that were ultimately retained in the reduced model would have been excluded.

Another consequence of the multitude of rare taxa in ecological datasets is that they typically contain a great number of outliers, which present significant challenges to traditional clustering techniques (Johnson & Wichern, 1992). The problem of outliers is mitigated with SOMs because during the learning phase, map units topologically near each other (defined by the neighborhood function in Equation 3) activate each other to learn from the same training vector. This process has a smoothing effect on the neurons' weight vectors, and thus removes noise (Kohonen, 2001). Additionally, each outlier impacts only its BMU and neurons in its neighborhood, leaving other regions of the map are unaffected.

Incumbent upon nearly all ecosystem restoration and management programs is the need for predicting and monitoring ecosystem responses to natural variability, disturbance, and anthropogenic impacts, often as they relate to changes in species composition. In coastal wetlands of Louisiana, increased inundation is expected to accompany increased freshwater inflows (Snedden, Cable, & Wiseman, 2007) brought about by Mississippi River diversions designed to reintroduce sediments to the rapidly subsiding delta plain (Peyronnin et al., 2017). Excessive inundation has been shown to impede belowground marsh production of community dominants in regions where river diversions are being planned (Snedden et al., 2015), which can diminish organic soil accumulation, the primary contributor to marsh vertical accretion in the inactive regions of the Mississippi River delta plain (Cahoon, White, & Lynch, 2011; DeLaune, Kongchum, White, & Jugsujinda, 2013; DeLaune, Whitcomb, Patrick, Pardue, & Pezeski, 1989; Nyman, DeLaune, Roberts, & Patrick, 1993; Turner, Swenson, & Milan, 2002). As such, the success of these projects may largely hinge upon the ability of existing marsh vegetation communities to self-organize around conditions of increased inundation (Peyronnin et al., 2017) and thus there exists a need to continuously classify new samples of species composition data into these community types as new data become available. Traditional unsupervised clustering techniques are unstable to the addition of new data, as shifts in cluster membership of previously classified sites may occur when new samples are added to previously classified datasets. This instability is problematic in that each new classification exercise yields a new classification system that may not be comparable with

those previously obtained (Černá & Chytrý, 2005). With SOMs, new data can simply be projected onto the existing, trained network, allowing for objective, effective and efficient classification as they become available. Thus, differences between expected or predicted outcomes and actual observations can be readily detected, making SOMs an ideal tool for monitoring and adaptive management.

Though several aspects of SOMs make them an attractive technique for ordination and classification of floral and faunal assemblages it is important to bear in mind that, while offering an approach that is largely free of assumptions, SOMs still carry a large subjective component that requires analyst input (e.g., specifying network size, learning rates, clustering method, size of training dataset). Additionally, SOMs provide no significance level for the gradients they identify, and (unlike eigen-based techniques) they do not partition the variance explained by these gradients. Though SOMs are robust to issues such as skewness, outliers, rare taxa, and classifying new samples without altering the existing partition, these issues can also be addressed with other approaches such as data transformation, noise clustering (DeCáceres, Font, & Oliva, 2010), downweighting of rare species, fuzzy classification frameworks (DeCáceres et al., 2010), and supervised classification approaches (Černá & Chytrý, 2005; van Tongeren, Gremmen, & Hennekens, 2008). And like all unsupervised classification methods, SOMs only perform well on samples containing combinations of taxa that are similar to those presented to it during the learning phase. Thus, it is important to ensure that the set of training samples are representative of the set of new samples the analyst expects to classify. However, one key distinguishing feature of SOMs is their capacity to facilitate visualization of taxa on the SOM in a way that provides semi-quantitative information regarding their abundance in the space of their sampling locations through examination of component planes (Astel, Tsakovski, Barbieri, & Simeonov, 2007). In this fashion, SOMs offer an avenue for combining R-mode and Q-mode clustering (Li et al., 2015).

Future efforts should expand the utility of this assemblage classification approach by linking the community types delineated here to environmental variables with predictive models based on either statistical (e.g., multinomial logistic regression; see Snedden & Steyer, 2013) or neural network (e.g., multilayer perceptron; see Park et al., 2006a) predictive models. Additionally, sites classified with this approach into community types based on their in-situ species composition data can subsequently serve as training data for supervised classification of remotely sensed multispectral imagery to provide high-resolution (10 m) spatial data of marsh vegetation community zonation across coastal landscapes.

ACKNOWLEDGEMENTS

DATA ACCESSIBILITY

All vegetation cover, salinity, water level and marsh elevation used in this study data can be obtained at www.lacoast.gov/crms.

ORCID

Gregg A. Snedden https://orcid.org/0000-0001-7821-3709

REFERENCES

Abraão, M. B., Shepard, G. H., Nelson, B. W., Baniwa, J. C., Andrello, G., & Yu, D. W. (2010). Baniwa vegetation classification in the white-sand *Campinarana* habitat of the northwest Amazon, Brazil. In L. M. Johnson, & E. S. Hunn (Eds.), *Landscape ethnoecology: Concepts of biotic and physical space* (pp. 83–115). New York, NY: Berghahn Books.

Adamczyk, J. J., Kurzak, M., Park, Y.-S., & Kruk, A. (2013). Application of a Kohonen's self-organizing map for evaluation of long-term changes in forest vegetation. *Journal of Vegetation Science*, 24, 405–414. https://doi.org/10.1111/j.1654-1103.2012.01468.x

Adams, D. A. (1963). Factors influencing vascular plant zonation in North Carolina salt marshes. *Ecology*, 44, 445–455. https://doi.org/10.2307/1932523

Albert, D. A., & Minc, L. D. (2004). Plants as regional indicators of Great Lakes coastal wetland health. *Aquatic Ecosystem Health & Management*, 7, 233–247. https://doi.org/10.1080/14634980490461588

Alhoniemi, E., Himberg, J., Parhankangas, J., & Vesanto, J. (2000). SOM toolbox. http://www.cis.hut.fi/projects/somtoolbox

Astel, A., Tsakovski, S., Barbieri, P., & Simeonov, V. (2007). Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. *Water Research*, 41, 4566–4578. https://doi.org/10.1016/j.watres.2007.06.030

Bae, M.-J., Chon, T.-S., & Park, Y.-S. (2012). Characterizing differential responses of benthic macroinvertebrate communities to floods and droughts in three different stream types using a Self-Organizing Map. *Ecohydrology*, 7, 115–126. https://doi.org/10.1002/eco.1326

Bertness, M. D. (1991a). Interspecific interactions among high marsh perennials in a New England salt marsh. *Ecology*, 72, 125–137. https://doi.org/10.2307/1938908

Bertness, M. D. (1991b). Zonation of Spartina patens and Spartina alterniflora in New England salt marsh. *Ecology*, 72, 138–148. https://doi.org/10.2307/1938909

Bhuiyan, M. J. A. N., & Dutta, D. (2012). Assessing impacts of sea level rise on river salinity in the Gorai river network, Bangladesh. *Estuarine, Coastal and Shelf Science*, 96, 219–227. https://doi.org/10.1016/j.ecss.2011.11.005

Brosse, S., Giraudel, J. L., & Lek, S. (2001). Utilisation of non-supervised neural networks and principal component analysis to study fish assemblages. *Ecological Modelling*, 146, 159–166. https://doi.org/10.1016/S0304-3800(01)00303-9

Cahoon, D. R., White, D. A., & Lynch, J. C. (2011). Sediment infilling and wetland formation dynamics in an active crevasse splay of the Mississippi River delta. *Geomorphology*, 131, 57–68. https://doi.org/10.1016/j.geomorph.2010.12.002

Callaway, J. C., Parker, V. T., Vasey, M. C., & Schile, L. M. (2007). Emerging issues for restoration of tidal marsh ecosystems in the context of predicted climate change. *Madroño*, 54, 234–248. https://doi.org/10.3120/0024-9637(2007) 54[234:EIFTRO]2.0.CO;2

Cantero, J. J., Cisneros, J. M., Zobel, M., & Cantero, A. (1998). Environmental relationships of vegetation patterns in saltmarshes of central Argentina. *Folia Geobotanica*, 33, 133. https://doi.org/10.1007/BF02913341

Černá, L., & Chytrý, M. (2005). Supervised classification of plant communities with artificial neural networks. *Journal of Vegetation Science*, *16*, 407–414. https://doi.org/10.1111/j.1654-1103.2005.tb02380.x

Chabreck, R. H., Linscombe, G., Hartley, S., Johnston, J. B., & Martucci, A. (2001). *Coastal Louisiana: Marsh vegetation types*. Lafayette, LA: U.S.G.S. National Wetlands Research Center. Report to the Coastal Wetlands, Planning, Protection and Restoration Act, and the Louisiana Department of Wildlife and Fisheries [CD-ROM]

Chea, R., Grenouillet, G., & Lek, S. (2016). Evidence of water quality degradation in lower Mekong basin revealed by self-organizing map. *PLoS ONE*, *11*, e0145527. https://doi.org/10.1371/journal.pone.0145527

Chen, L., Lek, S., Lek-Ang, S., & Zhongjie, L. (2012). Predicting fish assemblages and diversity in shallow lakes in the Yangtze River basin. *Limnologica*, *42*, 127–136. https://doi.org/10.1016/j.limno.2011.09.007

Chon, T. S. (2011). Self-Organizing Maps applied to ecological sciences. *Ecological Informatics*, *6*, 50–61. https://doi.org/10.1016/j.ecoinf.2010.11.002

Chon, T.-S., Park, Y. S., Moon, K. H., & Cha, E. Y. (1996). Patternizing communities by using an artificial neural network. *Ecological Modelling*, *90*, 69–78. https://doi.org/10.1016/0304-3800(95)00148-4

Cooper, A. (1982). The effects of salinity and waterlogging on the growth and carbon uptake of salt marsh plants. *New Phytologist*, *90*, 263–275. https://doi.org/10.1111/j.1469-8137.1982.tb03258.x

Couvillion, B. R., Beck, H., Schoolmaster, D., & Fischer, M. (2017). Land area change in coastal Louisiana (1932–2016). *Scientific Investigations* Map 3381. https://doi.org/10.3133/sim3381

DeCáceres, M., Font, X., & Oliva, F. (2010). The management of vegetation classifications with fuzzy clustering. *Journal of Vegetation Science*, *21*, 1138–1151. https://doi.org/10.1111/j.1654-1103.2010.01211.x

DeLaune, R. D., Kongchum, M., White, J. R., & Jugsujinda, A. (2013). Freshwater diversions as an ecosystem management tool for maintaining soil organic matter accretion in coastal marshes. *Catena*, *107*, 139–144. https://doi.org/10.1016/j.catena.2013.02.012

DeLaune, R. D., Whitcomb, J. H., Patrick, W. H. Jr, Pardue, J. H., & Pezeski, S. R. (1989). Accretion and canal impacts in a rapidly subsiding wetland. I. ^{137}Cs and ^{210}Pb techniques. *Estuaries*, *12*, 247–259. https://doi.org/10.2307/1351904

Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, *67*, 345–366. https://doi.org/10.1890/0012-9615(1997)067[0345:SAAIST]2.0.CO;2

Eleuterius, L. N., & Eleuterius, C. K. (1979). Tide levels and salt marsh zonation. *Bulletin of Marine Science*, *29*, 394–400.

Ewing, K. (1983). Environmental controls in Pacific Northwest intertidal marsh plant communities. *Canadian Journal of Botany*, *61*, 1105–1116. https://doi.org/10.1139/b83-118

Folse, T. M., West, J. L., Hymel, M. K., Troutman, J. P., Sharp, L. A., Weiffenbach, D. K., … Wood, W. B. (2014). *A standard operating procedures manual for the coast-wide reference monitoring system-Wetlands: Methods for site establishment, data collection, and quality assurance/quality control*. Baton Rouge, LA: Louisiana Coastal Protection and Restoration Authority, Office of Coastal Protection and Restoration.

Gauch Jr, H. G. (1982). *Multivariate analysis in community ecology*. New York, NY: Cambridge University Press.

Giraudel, J. L., & Lek, S. (2001). A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. *Ecological Modelling*, *146*, 329–339. https://doi.org/10.1016/S0304-3800(01)00324-6

Hijuelos, A. C., & Hemmerling, S. A. (2016). *Coast wide and basin wide monitoring plans for Louisiana's System Wide Assessment and Monitoring Program (SWAMP), version III*. Prepared for and funded by the Coastal Protection and Restoration Authority (CPRA) under Task Order 6, Contract No. 2503-12-38. Baton Rouge, LA: The Water Institute of the Gulf.

Hilton, T. W., Najjar, R. G., Zhong, L., & Li, M. (2008). Is there a signal of sea-level rise in Chesapeake Bay salinity? *JGR Oceans*, *113*, C09002. https://doi.org/10.1029/2007JC004247

Howard, R. J., Biagas, J., & Allain, L. (2016). Growth of common brackish marsh macrophytes under altered hydrologic and salinity regimes. *Wetlands*, *36*, 11–20. https://doi.org/10.1007/s13157-015-0711-x

Janousek, C. N., & Folger, C. L. (2014). Variation in tidal wetland plant diversity and composition within and among coastal estuaries: assessing the relative importance of environmental gradients. *Journal of Vegetation Science*, *25*, 534–545. https://doi.org/10.1111/jvs12107

Johnson, R. A., & Wichern, D. W. (1992). *Applied multivariate statistical analysis*, 3rd ed.. Englewood Cliffs, NJ: Prentice Hall.

Jongman, R. H. G., ter Braak, C. J. F., & van Tongeren, O. F. R. (1995). *Data analysis in community and landscape ecology*. New York, NY: Cambridge University Press.

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data*. New York, NY: John Wiley & Sons.

Kirwan, M. L., & Guntenspergen, G. R. (2015). Response of plant productivity to experimental flooding in a stable and a submerging marsh. *Ecosystems*, *18*, 903–913. https://doi.org/10.1007/s10021-015-9870-0

Kohonen, T. (2001). *Self-organizing maps*, 2nd ed.. Berlin, Germany: Springer.

Latham, P. J., Pearlstine, L. G., & Kitchens, W. M. (1994). Species association changes across a gradient of freshwater, oligohaline, and mesohaline marshes along the lower Savannah River. *Wetlands*, *14*, 174–183. https://doi.org/10.1007/BF03160654

Lee, J., Kwak, I.-S., Lee, E., & Kim, K. A. (2007). Classification of breeding bird communities along an urbanization gradient using an unsupervised artificial neural network. *Ecological Modelling*, *203*, 62–71. https://doi.org/10.1016/j.ecolmodel.2006.04.033

Legendre, P., & Birks, H. J. B. (2012). From classical to canonical ordination. In H. J. B. Birks, A. Lotter, S. Juggins, & J. Smol (Eds.), *Tracking environmental change using lake sediments. Developments in paleoenvironmental research*, vol. 5. (pp. 201–248). Dordrecht, The Netherlands: Springer.

Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., & Aulagnier, S. (1996). Application of neural networks to modeling non-linear relationships in ecology. *Ecological Modelling*, *90*, 39–52. https://doi.org/10.1016/0304-3800(95)00142-5

Lencioni, V., Marziali, L., & Rossaro, B. (2012). Chironomids as indicators of environmental quality in mountain springs. *Freshwater Science*, *31*, 525–541. https://doi.org/10.1899/11-038.1

Leonard, L. A., & Luther, M. A. (1995). Flow hydrodynamics in tidal marsh canopies. *Limnology and Oceanography*, *40*, 1474–1484. https://doi.org/10.4319/lo.1995.40.8.1474

Li, T., Xiang, R., & Li, T. (2015). Application of a self-organizing map and canonical correspondence analysis in modern benthic foraminiferal communities: a case study from the Pearl River estuary, China. *Journal of Foraminiferal Research*, *45*, 305–318. https://doi.org/10.2113/gsjfr.45.3.305

Linscombe, R. G., & Hartley, S. B. (2011). *Analysis of change in marsh types of coastal Louisiana, 1978–2001*. Open-File Report 2010-1282. Denver, CO: US Geological Survey.

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

Niering, W. A., & Warren, R. S. (1980). Vegetation patterns and processes in New England salt marshes. *BioScience*, *30*, 301–307. https://doi.org/10.2307/1307853

Nyman, J. A., DeLaune, R. D., Roberts, H. H., & Patrick Jr., W. H. (1993). Relationship between vegetation and soil formation in a rapidly submerging coastal marsh. *Marine Ecology Progress Series*, *95*, 269–279. http://www.jstor.org/stable/24833555

Odum, W. E. (1988). Comparative ecology of tidal freshwater and salt marshes. *Annual Review of Ecology and Systematics*, *19*, 147–176. https://doi.org/10.1146/annurev.es.19.110188.001051

Park, Y.-S., Chon, T.-S., Kwak, I.-S., & Lek, S. (2004). Hierarchical community classification and assessment of aquatic ecosystems using artificial neural networks. *Science of the Total Environment*, *327*, 105–122. https://doi.org/10.1016/j.scitotenv.2004.01.014

Park, Y. S., Gevrey, M., Lek, S., & Giraudel, J. L. (2005). Evaluation of relevant species in communities: Development of structuring indices for the classification of communities using a self-organizing map. In S. Lek, M. Scardi, P. F. Verdonschot, J. P. Descy, & Y. S. Park (Eds.), *Modelling community structure in freshwater ecosystems* (pp. 369–380). Berlin, Germany: Springer.

Park, Y.-S., Grenouillet, G., Esperance, B., & Lek, S. (2006a). Stream fish assemblages and basin land cover in a river network. *Science of the Total Environment*, *365*, 140–153. https://doi.org/10.1016/j.scitotenv.2006.02.046

Park, Y.-S., Tison, J., Lek, S., Giraudel, J.-L., Coste, M., & Delmas, F. (2006b). Application of a self-organizing map to select representative species in multivariate analysis: A case study determining diatom distribution patterns across France. *Ecological Informatics*, *1*, 247–257. https://doi.org/10.1016/j.ecoinf.2006.03.005

Pasternack, G. B., & Brush, G. S. (2001). Seasonal variations in sedimentation and organic content in five plant associations on a Chesapeake Bay tidal freshwater delta. *Estuarine, Coastal and Shelf Science*, *53*, 93–106. https://doi.org/10.1006/ecss.2001.0791

Penland, S., Wayne, L., Britsch, L. D., Williams, S. J., Beall, A. D., & Butterworth, V. C. (2001). *Process classification of coastal land loss between 1932 and 1990 in the Mississippi River delta plain, southeastern Louisiana*.. U.S.GS. Open File Report 00-418. Woods Hole, MA: U.S. Geological Survey, Coastal and Marine Geology Program, Woods Hole Field Center.

Pennings, S. C., Grant, M. B., & Bertness, M. D. (2005). Plant zonation in low-latitude salt marshes: disentangling the roles of flooding, salinity and competition. *Journal of Ecology*, *93*, 159–167. https://doi.org/10.1111/j.1365-2745.2004.00959.x

Peyronnin, N. S., Caffey, R. H., Cowan, J. H., Justic, D., Kolker, A. S., Shirley, B. L., & Wilkins, J. G. (2017). Optimizing sediment diversion operations: working group recommendations for integrating complex ecological and social landscape interactions. *Water*, *9*, 368. https://doi.org/10.3390/w9060368

Phleger, C. F. (1971). Effect of salinity on growth of a salt marsh grass. *Ecology*, *52*, 908–911. https://doi.org/10.2307/1936042

Poos, M. S., & Jackson, D. A. (2012). Addressing the removal of rare species in multivariate bioassessments: The impact of methodological choices. *Ecological Indicators*, *18*, 82–90. https://doi.org/10.1016/j.ecolind.2011.10.008

Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists*. Cambridge, UK: Cambridge University Press.

Robins, P. E., Skov, M. W., Lewis, M. J., Giménez, L., Davies, A. G., Malham, S. K., & Jago, C. F. (2016). Impact of climate change on UK estuaries: A review of past trends and potential projections. *Estuarine, Coastal and Shelf Science*, *169*, 119–135. https://doi.org/10.1016/j.ecss.2015.12.016

Salovaara, K. J., Cárdenas, G. G., & Tuomisto, H. (2004). Forest classification in an Amazonian rainforest landscape using pteridophytes as indicator species. *Ecography*, *27*, 689–700. https://doi.org/10.1111/j.0906-7590.2004.03958.x

Sasser, C. E., Visser, J. M., Mouton, E., Linscombe, J., & Hartley, S. B. (2008). *Vegetation types in coastal Louisiana in 2007*. Open-File Report 2008–1224, 1 sheet, scale 1:550,000. Denver, CO: U.S. Geological Survey.

Sasser, C. E., Visser, J. M., Mouton, E., Linscombe, J., & Hartley, S. B. (2014). *Vegetation types in coastal Louisiana in 2013*. U.S. Geological Survey Scientific Investigations Map 3290, 1 sheet, scale 1:550,000. Denver, CO: U.S. Geological Survey.

Schile, L. M., Callaway, J. C., Suding, K. N., & Kelly, N. M. (2017). Can community structure track sea-level rise? Stress and competitive controls in tidal wetlands. *Ecology and Evolution*, *2017*(7), 1276–1285. https://doi.org/10.1002/ece3.2758

Snedden, G. A., Cable, J. E., & Wiseman, W. J. (2007). Subtidal sea level variability in a shallow Mississippi River deltaic estuary, Louisiana. *Estuaries and Coasts*, *30*, 802–812. https://doi.org/10.1007/BF02841335

Snedden, G. A., Cretini, K. F., & Patton, B. (2015). Inundation and salinity impacts to above- and belowground productivity in *Spartina patens* and *Spartina alterniflora* in the Mississippi River Deltaic Plain: Implications for using river diversions as restoration tools. *Ecological Engineering*, *81*, 113–139. https://doi.org/10.1016/j.ecss.2012.12.002

Snedden, G. A., & Steyer, G. D. (2013). Predictive occurrence models for coastal wetland plant communities: Delineating hydrologic response surfaces with multinomial logistic regression. *Estuarine, Coastal and Shelf Science*, *118*, 11–23. https://doi.org/10.1016/j.ecss.2012.12.002

Steyer, G. D., Raynie, R. C., Steller, D. L., Fuller, D., & Swenson, E. (1995). *Quality management plan for coastal wetlands planning, protection, and restoration act*. Open-file series 95-01. Baton Rouge, LA: Louisiana Department of Natural Resources, Coastal Restoration Division.

Tison, J., Park, Y.-S., Coste, M., Wasson, J. G., Ector, L., Rimet, F., & Delmas, F. (2005). Typology of diatom communities and the influence of hydro-ecoregions: A study on the French hydrosystem scale. *Water Research*, *39*, 3177–3188. https://doi.org/10.1016/j.watres.2005.05.029

van Tongeren, O., Gremmen, N., & Hennekens, S. (2008). Assignment of relevés to pre-defined classes by supervised clustering of plant communities using a new composite index. *Journal of Vegetation Science*, *19*, 525–536. https://doi.org/10.3170/2008-8-18402

Turner, R. E., Swenson, E. M., & Milan, C. S. (2002). Organic and inorganic contributions to vertical accretion in salt marsh sediments. In M. P. Weinstein, & D. A. Kreeger (Eds.), *Concepts and controversies in tidal marsh ecology* (pp. 583–596). New York, NY: Kluwer Academic Publishers.

USDA, NRCS (2019). *The PLANTS Database*. Greensboro, NC: National Plant Data Team. http://plants.usda.gov

Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, *11*, 586–600. https://doi.org/10.1109/72.846731

Visser, J. M., & Peterson, J. K. (2015). The effects of flood duration and salinity on three common upper estuary plants. *Wetlands*, *35*, 625–631. https://doi.org/10.1007/s13157-015-0644-4

Visser, J. M., & Sandy, E. R. (2009). The effects of flooding on four common Louisiana marsh plants. *Gulf of Mexico Science*, *27*, 21–29. https://doi.org/10.18785/goms.2701.03

Visser, J. M., Sasser, C. E., Chabreck, R. H., & Linscombe, R. G. (1998). Marsh vegetation types of the Mississippi river deltaic plain. *Estuaries*, *21*, 818–828. https://doi.org/10.2307/1353283

Visser, J. M., Sasser, C. E., Linscombe, R. G., & Chabreck, R. H. (2000). Marsh vegetation types of the Chenier Plain, Louisiana, USA. *Estuaries*, *23*, 318–327. https://doi.org/10.2307/1353324

Voss, C. M., Christian, R. R., & Morris, J. T. (2013). Marsh macrophyte responses to inundation anticipate impacts of sea-level rise and indicate ongoing drowning of North Carolina marshes. *Marine Biology*, *160*, 181–194. https://doi.org/10.1007/s00227-012- 2076-5

Wang, X., Xu, L., & Rongrong, W. (2016). Comparison on organic carbon within two typical wetland areas along the vegetation gradient of Poyang Lake, China. *Hydrology Research*, *47*(S1), 261–277. https://doi.org/10.2166/nh.2016.218

Zar, J. H. (1984). *Biostatistical analysis*. Upper Saddle River, NJ: Prentice Hall.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Coastwide Reference Monitoring System vegetation samples, by site and year, used to train the self-organizing map

Appendix S2. Differences in summed Global Structuring Indices between original (559 taxa) and reduced (49 taxa) datasets

Appendix S3. Component planes for the 24 taxa of the reduced self-organizing map not included in Figure 4

Appendix S4. Silhouette plots and dendrograms comparing various clustering approaches