Environmental Modelling and Software 124 (2020) 104580

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: http://www.elsevier.com/locate/envsoft

Percentile-Range Indexed Mapping and Evaluation (PRIME): A new tool for long-term data discovery and application

Shimelis B. Dessu[a,b,*], René M. Price[a,b], John S. Kominoski[b,c], Stephen E. Davis[d], Adam S. Wymore[e], William H. McDowell[e], Evelyn E. Gaiser[b,c]

[a] Department of Earth and Environment, Florida International University, Miami, FL, 33199, USA
[b] Southeast Environmental Research Center (SERC), Florida International University, Miami, FL, 33199, USA
[c] Department of Biological Sciences, Florida International University, Miami, FL, 33199, USA
[d] The Everglades Foundation, Palmetto Bay, FL, 33157, USA
[e] Department of Natural Resources and the Environment, University of New Hampshire, Durham, NH, 03824, USA

ARTICLE INFO

Keywords:
Percentile
Ecological mapping
PRIME
Florida Coastal Everglades
Long Term Ecological Research
Datascape

ABSTRACT

Percentile-Range Indexed Mapping and Evaluation (PRIME) is a new tool to visualize and quantifying spatio-temporal dynamics of long-term datasets. PRIME is based on categorical partitioning of magnitude based on user defined indices assigned to ranges of percentile and mapping subsets of data at selected percentiles of long-term data. Indices can reflect attributes such as water management decisions, tolerable range of water quality to a species, ecological risk, response to and recovery from disturbance, and values of ecosystem services. PRIME provides visual and robust datascapes and flexibility to evaluate variability in space and time for long-term environmental assessment. Here, we demonstrate the utility of PRIME using 16 years of hydrologic and salinity data from 14 sites representing three unique hydrological systems in the Florida Coastal Everglades (FCE). The resulting PRIME datascapes reveal interaction between water management and sea-level rise to drive salinity levels in the FCE.

Software/data availability

Name of Software Percentile-Range Indexed Mapping and Evaluation (PRIME)
Developer Shimelis B. Dessu
Contact Tel. 305-401-5898, Email: sbehailu@gmail.com
Availability http://go.fiu.edu/prime
Required hardware PC
Required Software Windows 7 or latest and Text editor
Programming Language MatLab 2018b

1. Introduction

Long-term environmental observatory programs have enabled collection of a large amount of hydrological, physio-chemical and biological data to reveal the complex drivers of ecological change, and to inform decision making. National and international observatory networks are becoming more common in order to address the spatio-temporal context of long-term environmental change including the U. S. Long Term Ecological Research (LTER) (Callahan, 1984; Hobbie et al., 2003), National Critical Zone Observatory (CZO), the Global Lakes Ecological Observatory Network (GLEON) (Hanson et al., 2018), National Ecological Observatory Network (NEON) (Dalton, 2000), and hundreds of International LTER sites (Kim, 2006) generating massive quantities of high-resolution environmental data. Likewise, there are many government and regulatory agencies with long-term water quality and hydrological monitoring programs across the globe (Hobbie et al., 2003; Read et al., 2017). Many of these platforms now include high-frequency sensor-based data streams which further magnify the volume of data available to researchers. The challenge of environmental monitoring and assessment is now moving from lack or scarcity of data to data mining, visualization and big data synthesis (Carey et al., 2015; Hamilton et al., 2015; Read et al., 2017). The best use of these data requires novel synthesis methods and statistical tools to perform cross-cutting spatio-temporal analyses at global, regional and local scales to develop and test hypotheses. Simple yet comprehensive

* Corresponding author. Department of Earth and Environment, Florida International University. 11200 SW 8th Street, AHC5-364, Miami, FL, 33199, USA.
E-mail addresses: sdessu@fiu.edu, sbehailu@gmail.com (S.B. Dessu).

https://doi.org/10.1016/j.envsoft.2019.104580
Received 13 May 2019; Received in revised form 2 August 2019; Accepted 4 November 2019
Available online 8 November 2019
1364-8152/Published by Elsevier Ltd.

analyses and visualizations are essential to communicate results and findings of long-term studies among the scientific community and the public (Farley et al., 2018).

Here, we introduce the development and application of Percentile-Range Indexed Mapping and Evaluation (PRIME), a new tool designed to assess long-term hydrological and ecological processes, develop and test novel hypotheses, and facilitate informed decision making. As an end-product, PRIME creates datascapes which are multi-layered, semi-quantitative, highly visual and easy to interpret matrices to compare variability at multiple scales from days to years, pre- and post-events and overall long-term patterns. The overarching objective of this paper is to present the conceptual framework, algorithm, and practical application of PRIME for long-term hydrological and ecological data. To meet this objective, we describe the mathematical benefits and flexibility of percentile curves (PCs) compared to commonly used statistical approaches to establish mapping indices for the assessment of long-term data. We demonstrate how a complex eco-hydrological system can be mapped to datascapes using PRIME to explore and understand processes using multiple variables, while incorporating the spatiotemporal variability of critical environmental drivers. To accomplish this, we use 16 years of data from the Florida Coastal Everglades (FCE) collected as part of the FCE LTER program along two freshwater-to-marine gradients in Everglades National Park (ENP) since 2000.

2. Comparison of commonly used statistical approaches to percentile curves

Environmental data are often collected and presented as time series of values observed in calendar time step. Consider two datasets of variables plotted as time series showing their long-term relationship with each other (Fig. 1a). Such data are often used to test hypotheses using descriptive statistics (Fig. 1b), analysis of variance (ANOVA, Fig. 1c) and regression analysis (Fig. 1d), to name a few. Descriptive statistics summarize the central tendencies (e.g., mean, median, maximum and minimum) and variation (e.g., standard deviation and variance) of data integrated over time (Fig. 1b). Analysis of trends in mean values, however, can be misleading because natural processes are seldom normally distributed. More meaningful insight into long-term variation and trends could be generated by quantifying the percent of time that the mean is likely to be observed. The 1-factor ANOVA provides a better insight into the variability of data and results are often presented as box-plots (Fig. 1c). Extreme values are often treated as 'outliers' depending on selection and setting of underlying probability distribution used in the 1-factor ANOVA. In long-term studies, the 'outliers' are likely to represent extreme events which can have long-lasting effects on the dynamics of the system. As such, box plots are often used and offer a simple representation of the range of data (whiskers) along with median, first and third quartiles for comparison.

Trend analysis is yet another common method used to visualize (e.g. scatter plots) and summarize relationship among multiple data sets with a common attribute such as time (Fig. 1d). Trends between a response variable are regressed with a set of drivers to determine cause-and-effect relationships. Even though linear regressions are quite common, many hydrological and ecological relationships display non-linearities and analyses may warrant non-linear or step-wise multi-variate regressions. Trend and regression analyses results however, are also prone to the large influence by extreme values and outliers. The underlying bias due to extreme values is attributed to the mathematical dependence of these methods on the relative distance (deviation) of observations. Since rare and extreme events tend to initiate a lasting response reverberating through a system, frequency-based approaches have been used to circumvent the limitations of deviation-based analyses and capture their contribution in synthesized results. Percentile curves (PC) are such a frequency-based approach which utilize the linear ranking of values in ascending order to transform data into useable visualizations that are easier to differentiate and interpret (Fig. 1e).

Flow duration curves (FDCs) are a widely used version of PCs in water resources applications such as water management, flood forecasting (Vogel and Fennessey, 1995) and water control structures such as reservoirs. FDCs are PCs ranked in descending order and illustrate the signature characteristics of the flow regime (Yilmaz et al., 2008) with frequency and duration of occurrence of a specific observation over a sampling list or period (Qian, 2015). Exceedance probabilities (Gunderson, 1994; Todd et al., 2010) and exceedance curves (Dessu et al., 2018, 2019) are variants of PCs frequently used to assess environmental compliance (e.g. concentration duration frequency curves, U.S. EPA, 2005) in order to integrate the frequency component (in percentile) with magnitude and duration.

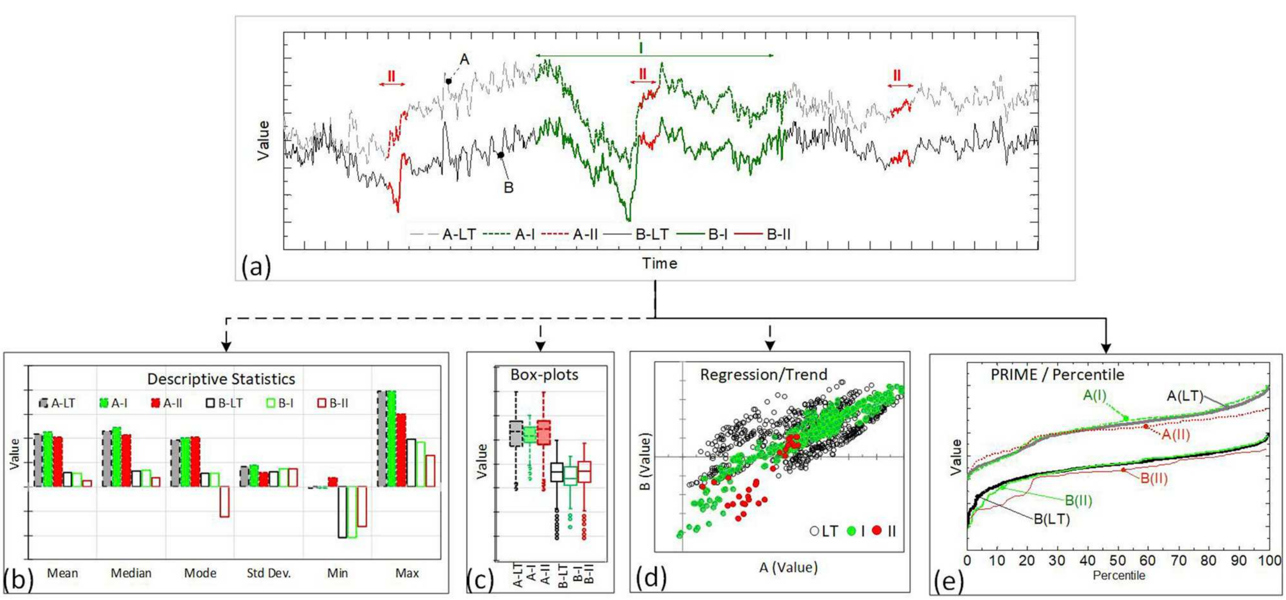


Fig. 1. Comparison of common approaches for visualizing long-term time series data (A and B), their sub-sets representing a continuous segment in the middle (A–I and B–I) and combination of three smaller segments (A–II and B–II). Statistical summaries are presented to show relative variability A against B and relative to their corresponding sub-dataset. (a) Time-series plot of a variable measured at two separate sites. (b) Descriptive statistical summaries, (c) summary of analysis of variance, (d) scatter plot showing trend of B relative to A, and (e) percentile curve representation as used in PRIME, of the long-term and sub data sets.

Quantification via percentile transform the calendar time steps in to rank based on the percentage of the total number of observations. Since values and percentiles are monotonically increasing/decreasing in PCs, the curve can be segmented into mutually exclusive ranges of percentile or value. PCs spread values over a percentile range of zero to hundred, irrespective of the number and type of data enabling comparison of sub-data sets against their long-term trend or other relevant data-sets. These ranges can be assigned descriptive labels based on their value and percentile as extreme high/low, high/low and normal ranges. For example, the values regarded as outliers (observations falling beyond the whiskers in Fig. 1c) can be grouped as extreme low or high on the percentile curve (Fig. 1e). Hence, the percentile curve of the long-term data represents the signature characteristics of the system against which all sub-datasets can be compared with to assess the long-term dynamics.

3. Percentile range index mapping and evaluation (PRIME)

3.1. General PRIME framework

PRIME utilizes percentile curves (PCs) to synthesize long-term time-series datasets by categorically comparing magnitude and frequency of sub-datasets against long-term observations and trends. Mapping and evaluation procedures are based on comparison of subsets of the time-series at selected percentiles against the intervals of magnitudes from percentile ranges of the long-term data. PRIME consists of five steps: (1) input and initialization where the user provides data in suitable format and sets mapping parameters (Fig. 2a); (2) process the long-term percentile curve and extract baseline values for mapping (Fig. 2b); (3) partition the long-term data into groups, generate their respective percentile curve and extract the value corresponding to defined mapping percentile (Fig. 2c); (4) derive status level datascapes (Fig. 2d); and (5) evaluate relationships between two corresponding maps (Fig. 2e). In this section, we present the algorithm of PRIME (Fig. 2) along with step-by-step implementation using average daily sea level data collected from January 1, 2001 to December 31, 2016 at the Key West tide gage station (Holgate et al., 2012; PSMSL, 2016). The sea level is referenced to the North American Vertical Datum 1988 (NAVD88) to align with water level data discussed in section 4.

3.2. Input data

Consider a multi-variable time-series dataset containing time t, variable/site name, mapping ID (I, J) and value, in column format [time, Name, i, j, value] format (Fig. 2a). Mapping identifiers define rows and columns of the final datascape output. For a selected variable from the dataset, PRIME extracts the data in [t, name, i, j, var] format (Fig. 2a). Hence, the input data is grouped into i x j columns, representing individual mapping unit.

3.3. Percentile-ranges and indices

PRIME procedures are based on comparison of percentile curves of subsets of the data defined by the combinations of mapping identifiers against the long-term PC. Percentile curve is defined as.

$$PC(\ P(x_n), x_n\ ) = \left[\ \frac{n}{N+1} \times 100,\ x_n\ \right] \tag{1}$$

Where: PC is the percentile curve, $P(x_n)$ is the percentile of the nth ranked observation x_n of the variable in ascending order. N is total number of observations. Since percentiles range from zero to a hundred, the long-term PC can be partitioned into mutually exclusive percentile intervals (Fig. 2b) defined by user specified cut-off percentiles. Each of these intervals are assigned a unique index to reflect the status of observations falling within the range as:

$$PRI = \begin{cases} Index_1, & P \le \Omega_1 \\ Index_2, & \Omega_1 < P \le \Omega_2 \\ \vdots & \vdots \\ Index_n, & P > \Omega_{n-1} \end{cases} \tag{2}$$

where: PRI is the set of percentile-range indices; *Index* is user defined index; n is the number of intervals; and Ω is set of percentiles defining each interval. The boundary value of the intervals is extracted from the long-term PC. For each cut-off percentile, the boundary of the corresponding long-term percentile-range values *PRV*(Ω) is extracted.

The sea level time series data was extracted as [Date, Site, Year, Month, Value] and the PCs were generated for the long-term data (implementation of Fig. 2a and b). Five percentile intervals are defined by cut-off percentiles (Ω = 15%, 33%, 67%, and 85%). Percentile-range indices (PRI = LL, L, M, H, and HH) were established representing relative sea level status as lower low (LL: P ≤ 15%), low (L: 15% < P ≤ 33%), median (M: 33% < P ≤ 67%), high (H: 67% < P ≤ 85%) and higher high (HH: P > 85%) (Eq. (2), Fig. 2a) observations. Percentile range values (PRV(Ω) = −32.5 cm, −26.5 cm, −17.2 cm and −9.8 cm) corresponding to (Ω = 15th, 33rd, 67th and 85th) percentiles of the long-term sea level PC were extracted, (Fig. 2b).

3.4. Datascapes

Once the long-term percentile intervals and indices are established, PRIME partitions the long-term data by the mapping ID (I, J) (Fig. 2c) and generates percentile curves, *pc*(*i, j* (Eq. (1)). Let *ω*(*k*) be the set of *k* percentiles of interest for mapping and evaluation and *pv*(*i,j,k*), be the value extracted from *pc*(*i, j*) at *ω*(*k*) (Fig. 2c). By comparing the long-term PV and the mapping unit pv, the datascape is given by:

$$Dscape(i,j,k) = \begin{cases} index_1, & pv(i,j,k) < PRV(\Omega_1) \\ index_2, & PRV(\Omega_1) \le pv(i,j,k) < PRV(\Omega_2) \\ \vdots & \vdots \\ index_n, & pv(i,j,k) \ge PRV(\Omega_{n-1}) \end{cases} \tag{3}$$

where: *Dscape (i, j, k)* is the datascape with indices, $index_n$, at row *i*, column *j* of mapping percentile *k*. *PRV* is the long-term value corresponding to the cutoff percentile, *Ω*.; and *pv* is the percentile value extracted at mapping percentiles. The datascape represents the changes in the system at selected percentiles with respect of the long-term trend (Fig. 2d). PRIME assigns a unique color for a quick visual understanding the datascape.

Three sets of mapping categories (i, j) were used to map sea level status (Fig. 2c). The first map used [All Years, Month] mapping categories in which all values falling in the same month were combined to make the monthly PC. The second map used [All Months, Year] to map the relative annual status. The third mapping category was [Month, Year], in which PCs of individual month were generated. The long-term and individual monthly datascape was used to evaluate the seasonal pattern of variables across sites. The annual datascape was intended to capture annual variabilities across sites associated with major events such as drought, storms and hurricanes that have extended impact on the ecosystem. The 33rd, average, median (50th) and 67th percentiles (i.e., ω(1) = 33%, ω(2) = mean(x), ω(3) = median 50% and ω(4) = 67%) of the PCs from sub-datasets grouped by the mapping categories were extracted and compared with values of the five percentile ranges of the long-term data to produce datascape (Fig. 2d).

PCs for the long-term daily sea-level and overlays of selected sub-datasets extracted by Month (June), Year (2009), and Month-(February and October 2009) are shown in Fig. 3a. Percentile range values (PRV = −32.5 cm, −26.5 cm, −17.2 cm and −9.8 cm) corresponding to (PRI = 15th, 33rd, 67th and 85th) percentiles of the long-term sea level PC were extracted, respectively as shown in Fig. 3a. PRIME generates 12 PCs for the sea level sub-datasets grouped in Month category, 16 for the Year category and 192 PCs for the Month-Year

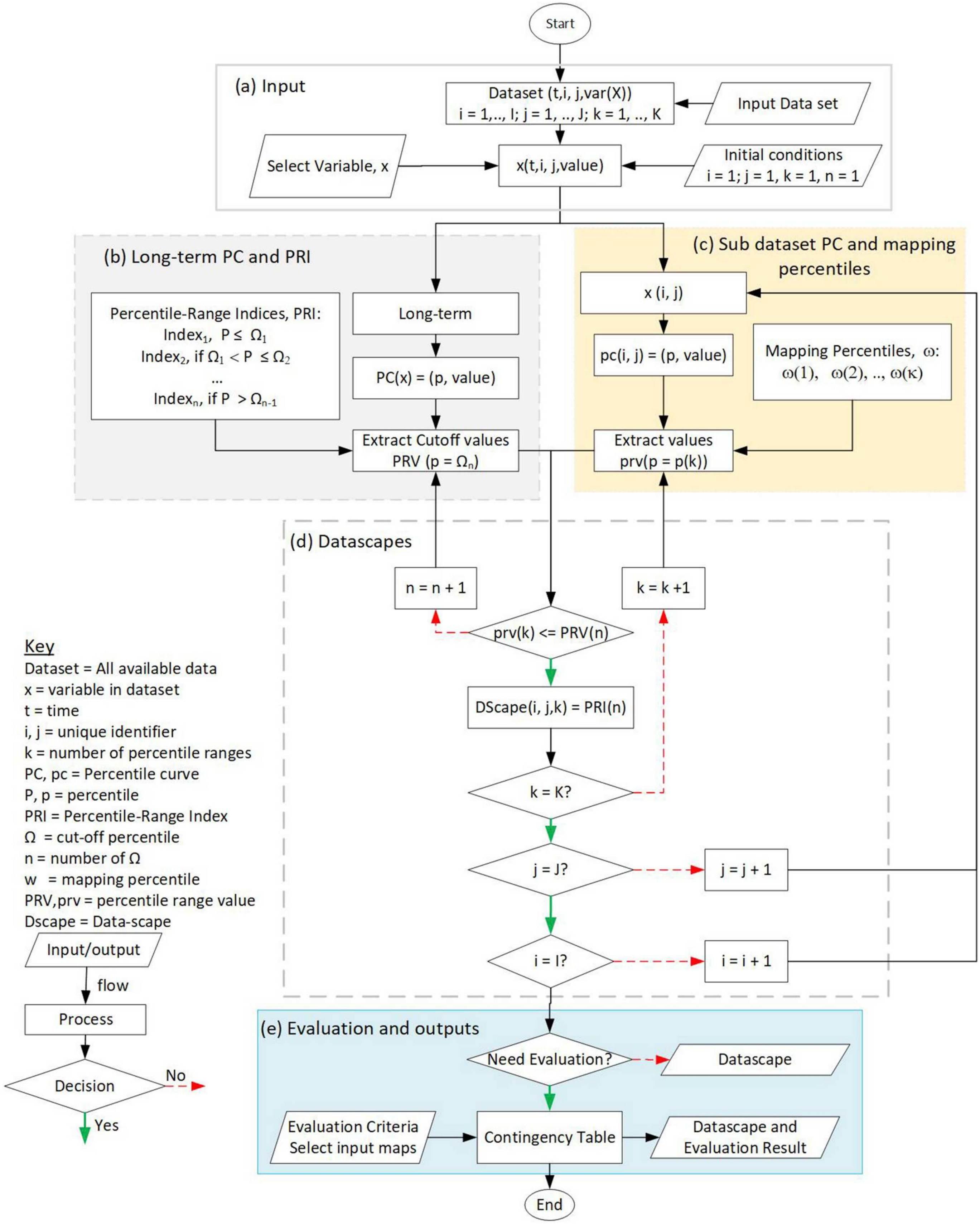
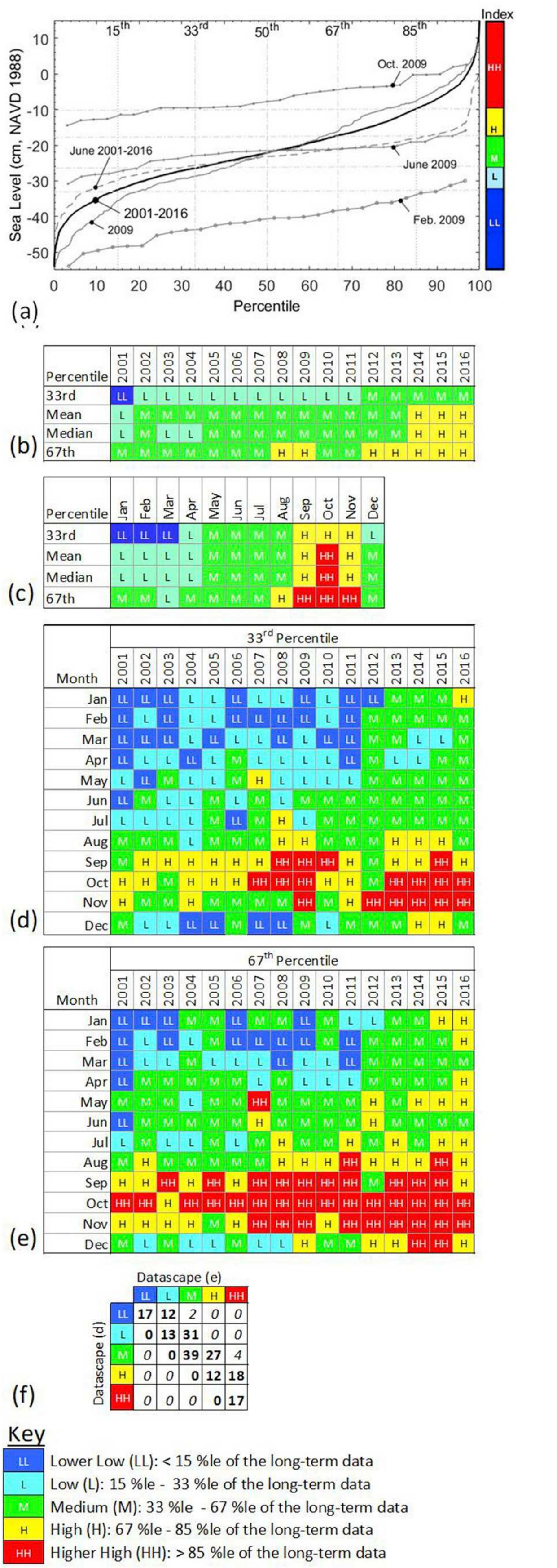


Fig. 2. General process algorithm of PRIME. (a) Input selection, (b) generate PC for long-term and extract values at selected percentiles, (c) generate PC for each mapping sub-dataset and extract values at mapping percentiles, (d) generate datascape, and (e) Evaluation and outputs.



(caption on next column)

Fig. 3. Illustration of the PRIME setup and input/output procedure using sixteen years of daily sea-level data (2001–2016) from Key West tide gage station to produce intervals of percentile curves (PC) at 15th, 33rd, 67th and 85th percentile and corresponding ranges were defined as LL, L, M, H and HH. (a) Long-term percentile curves of sea level data and selected subset data by month (June), by year (2009), and by year-month (February and October 2009), (b) the long-term monthly datascape of the 33rd percentile, mean, median and 67th percentile, (c) the long-term annual sea level datascape of the mean, median and 33rd and 67th percentiles, (d) month-year sea level datascape at 33rd percentile, (e) month-year sea level datascape at 67th percentile, and (f) contingency table showing comparative evaluation of datascapes in (d) and (e). The chi-square statistics of the contingency table is 307.9 with $P < 0.01$.

category.

Compared to the five status ranges of the long-term data, the 33rd and 67th percentiles of the 2009 PC (−28.8 cm, −15.9 cm) fall in L and H ranges of the long-term PRI, respectively (Fig. 3b). Similarly, when the June 2001–2016 PRVs at the 33rd and 67th mapping percentiles are extracted and compared with the long-term PRVs, both fall within the range assigned M status index of long-term data (Fig. 3c). The monthly 33rd and 67th percentiles of sea-level observations of 2009 were mapped to LL (Fig. 3d) in February and HH (Fig. 3e)in October, respectively.

3.5. Evaluation

PRIME datascapes are multi-dimensional maps where the relative location of indices is important in determining the degree of association and correspondence. Hence, common one-dimensional (rank-based) correlation such as Spearman's rank-order correlation coefficient (Spearman, 1904), Kendall's rank-correlation test (Kendall, 1955) may not fully capture the association among datascapes. Contingency table and chi-square test are widely used methods to assess relationship between maps (Everitt, 1992) and here we employ both contingency table and chi-square to measure the degree of correspondence and association between PRIME datascapes. Chi-square test whether there is a significant relationship between datascapes based on the associated *P*-value. After this significance test, the degree of correlation between datascapes is measured by contingency coefficient (CC) as:

$$CC = \sqrt{\frac{\chi^2}{N + \chi^2}} \tag{4}$$

where CC is the contingency coefficient. χ^2 is chi-square statistics, N is the total number of counts and k is the number of rows or columns of a contingency table. Contingency coefficient values vary between 0 and 1, close to 0 shows little relationship and close to 1 shows strong relationship.

PRIME also employs a new non-parametric diagonal correspondence coefficient (DCC) to measure the overall correspondence of datascapes as:

$$DCC = \begin{cases} \dfrac{\sum\limits_{i=j} CT(i,j) + 0.5 \times \sum\limits_{i=j\pm1} CT(i,j)}{\sum\limits_{all} CT(i,j)}, & \sum\limits_{i=j} CT(i,j) \geq \sum\limits_{i=j} CT_F(i,j) \\ -\dfrac{\sum\limits_{i=j} CT_F(i,j) + 0.5 \times \sum\limits_{i=j\pm1} CT_F(i,j)}{\sum\limits_{all} CT_F(i,j)}, & Otherwise \end{cases} \tag{5}$$

where *CT* (*i, j*) is the element of row *i* and column *j* of the contingency table. $CT_F(i, j)$ is the flipped contingency table determined as $CT_F(i, j) = CT(i, J - j + 1)$. DCC is the ratio of the sum of all elements along the diagonal ($i = j$) plus half of the elements above and below the diagonal

($i = 1 \pm j$) divided by the total sum of the contingency table. If the datascapes have an inverse relationship, DCC will be negative and is calculated after flipping the contingency table. DCC values vary from −1 to 1 and values close to 0 suggest little correspondence. Positive and negative DCC show direct and inverse correspondence, respectively, as per the order of indices defined by the percentile range. A strong correspondence may be suggested when $|DCC| \geq 1.5k/(3k\text{-}2)$ for contingency tables with four or more rows and columns, $k \geq 4$.

The distribution and pattern of counts in the contingency table can provide further insight about relationships between datascapes not captured with the methods included in PRIME. However, the user should exercise caution in the use and interpretation of contingency table and measures of association. Statistical inference should be supplemented with site and disciplinary knowledge. PRIME also includes statistical summaries of input data to supplement evaluation and interpretation of datascapes.

To illustrate the evaluation procedures of PRIME datascapes, consider the 33rd and 67th percentile sea level datascapes as two separate datascapes (Fig. 3d and e). The resulting contingency table (Fig. 3f) has a total count of N = 192 pair-wise indices, chi-square (×2 = 307.9), and *P*-value close to 0. Based on the *P*-value, the two contingency table statistically significant association between the two datascapes. The CC is 0.78 (Eq. (4)) suggests a strong association. The diagonal correspondence coefficient, DCC = (98 + 0.5 × (0 + 88))/192 = 0.74, also suggests strong correspondence between the two datascapes (Eq. (5)). As expected, the 67th percentile sea level datascape has equal or greater status levels compared to the 33rd percentile datascape and as such, the contingency table has zero counts below the diagonal.

4. Application of PRIME: mapping long-term changes in water levels and salinity in the Florida Everglades

4.1. Introduction

The FCE is a large subtropical wetland ecosystem in south Florida, USA and is the southern-most portion of the remaining wetlands of the Greater Everglades, which also include the Water Conservation Areas (WCA) (Fig. 4). Managed freshwater canals border the FCE to the north and east while the Florida Keys and the Gulf of Mexico form the southern and southwestern boundaries, respectively (Fig. 4). Florida Bay is an open water area between mainland Florida and the Florida Keys. The FCE receives both fresh and saline surface water and groundwater along the boundaries of ENP. FCE is susceptible to a projected SLR of 1–2 m by 2100 (Haigh et al., 2014), as much of the landscape is less than 1.5 m above mean sea-level (Titus and Richman, 2001). The combination of rising sea-level with reductions in fresh water flow (due to water management) has increased salt water intrusion into the FCE and its underlying aquifer allowing for salt-tolerant communities such as mangroves to overtake formerly freshwater species (Dessu et al., 2018; Karamperidou et al., 2013; Krauss et al., 2011; Ross et al., 2000). An acceleration in sea-level rise is expected to increase coastal erosion and soil loss, potentially replacing coastal wetlands with non-vegetated open water areas (Todd et al., 2012; Trenberth et al., 2014; White and Kaplan, 2017; Wilson et al., 2018).

The Comprehensive Everglades Restoration Plan (CERP) was authorized in 2000 to restore the remaining wetland portions of the Everglades to its pre-drainage conditions and ensure sustainability. The long-term impact of CERP projects in the FCE relies on maintaining suitable quantity and quality of water, nutrient flux, productivity and hydrologic connectivity (Sklar et al., 2005). Continual monitoring and assessment are essential to understand and evaluate restoration, identify challenges and opportunities to inform management decisions and mitigate undesirable or unintended consequences.

Here, we demonstrate an application of Percentile Range Indexed Mapping and Evaluation (PRIME) to facilitate data synthesis and the assessment of long-term environmental restoration in the Everglades. We apply PRIME along two main freshwater flow-ways of the FCE, the Shark River slough (SRS) and Taylor Slough (TS), and within Florida Bay (FB) following the FCE-LTER monitoring stations to map the status and dynamics of hydrologic drivers and changes in salinity from 2001 to 2016 (Fig. 4). The specific objectives are to: (1) create datascapes describing spatial and temporal changes in water level, sea level, fresh-to-marine head difference (FMHD) and salinity, (2) evaluate the impact of water level in freshwater marshes to counteract salinity, and (3) isolate signatures of disturbances and their spatial and temporal extent.

Output from our analysis includes long-term PCs for annual, seasonal and monthly datascapes of water level, sea level, FMHD and salinity. Relationships among these variables are then evaluated using contingency tables. Findings from this application of PRIME provide a holistic understanding of the hydrologic conditions of FCE in a context of how freshwater flow management can be improved through CERP to reduce the negative impacts of SLR.

4.2. Methods

4.2.1. Site description

Freshwater inputs to the FCE come from direct rainfall and surface water inflow through gated structures connecting storage in water conservation areas (WCA) with Shark River Slough (SRS) and Taylor Slough (TS). Freshwater inflow from WCA-3A to SRS is managed by the operation of water control structures located along Tamiami Trail, the northern boundary of ENP. Inflow to TS is controlled by a series of detention areas and structures along the eastern boundary of ENP (Fig. 4).

Sites SRS1 to SRS3 and TS/Ph1 to TS/Ph3 are considered freshwater marsh sites dominated by sawgrass (Fig. 4). Sites SRS4 to SRS6, and TS/Ph6-TS/Ph7 are brackish water sites dominated by mangroves. The region between SRS3 and SRS4, and TS/Ph3 and TS/Ph6 are ecotones, transitioning between freshwater and estuarine habitats. We use PRIME to map patterns of water level and sea level and evaluate the implications on salinity at the coastal sites of SRS (SRS3 to SRS6) and TS (TS/Ph3, TS/Ph6 and TS/Ph7) and the Florida Bay (TS/Ph9, TS/Ph10 to TS/Ph11) sites of the FCE-LTER (Fig. 4).

4.2.2. Data

Sixteen years of daily water level data from 2001 to 2016 for each of the FCE SRS and TS sites were obtained from nearby stations (NP201/SRS1, P36/SRS2, MO-215/SRS3, Shark River/SRS5, TE/SRS4, NTS1/TS/Ph1, TSB/TS/Ph2, E146/TS/Ph3), Upstream Taylor River (UTR, TS/Ph6) and Taylor River Mouth (TRM, TS/Ph7) operated by the United States Geological Survey (USGS) and the Everglades Depth Estimation Network (EDEN) project (USGS/EDEN, 2019) (Figs. 4 and 5 a& b). Daily sea-level data from the Key West station operated by the National Oceanic and Atmospheric Administration (NOAA) was used due to its comparatively longer and continuous daily sea-level data available since 1913 (Holgate et al., 2012; PSMSL, 2016) (Fig. 5a and b). Tidal gage sea-level data were converted to the 1988 North America Vertical Datum (NAVD88) for comparison with water-level data. FMHD is determined by subtracting sea level from water levels at the most downstream freshwater site, SRS3 for Shark River slough and TS/Ph3 for Taylor Slough, respectively (Fig. 5c and d). Surface water salinity data were obtained from FCE-LTER monitoring stations in SRS (SRS3 to SRS6) (Gaiser and Childers, 2017) (Fig. 5c), TS sites (TS/Ph3, TS/Ph6 and TS/Ph7) (Troxler, 2017) (Fig. 5d) and Florida Bay (TS/Ph9, TS/Ph10 and TS/Ph11) (Briceno, 2017; Fourqurean, 2017) (Fig. 5e). Surface water salinity data were three-day composites of water samples collected at each of those sites.

4.2.3. PRIME setup and output

Long-term PCs were generated for all variables based on Eq. (1) (Fig. 2a and b). Five percentile-range indices (PRI) were established representing the status level of the variable as lower low (LL: $P \leq 15\%$)

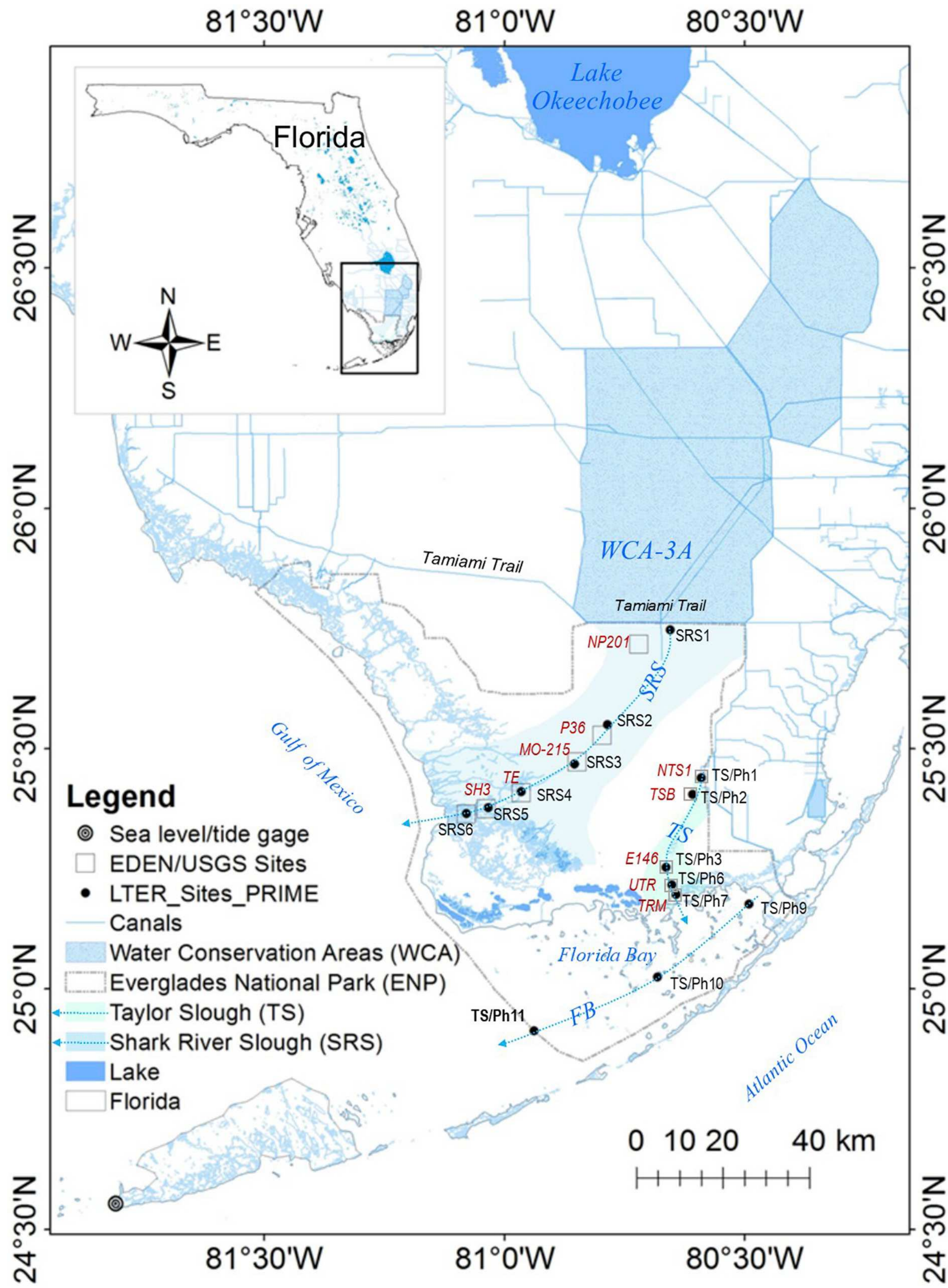


Fig. 4. Location map of the Florida Coastal Everglades Long Term Ecological Research (FCE-LTER) Program hydrologic and water quality monitoring stations in the Shark River Slough (SRS), Taylor Slough (TS) and Florida Bay (FB) transect within Everglades National Park (ENP), Florida.

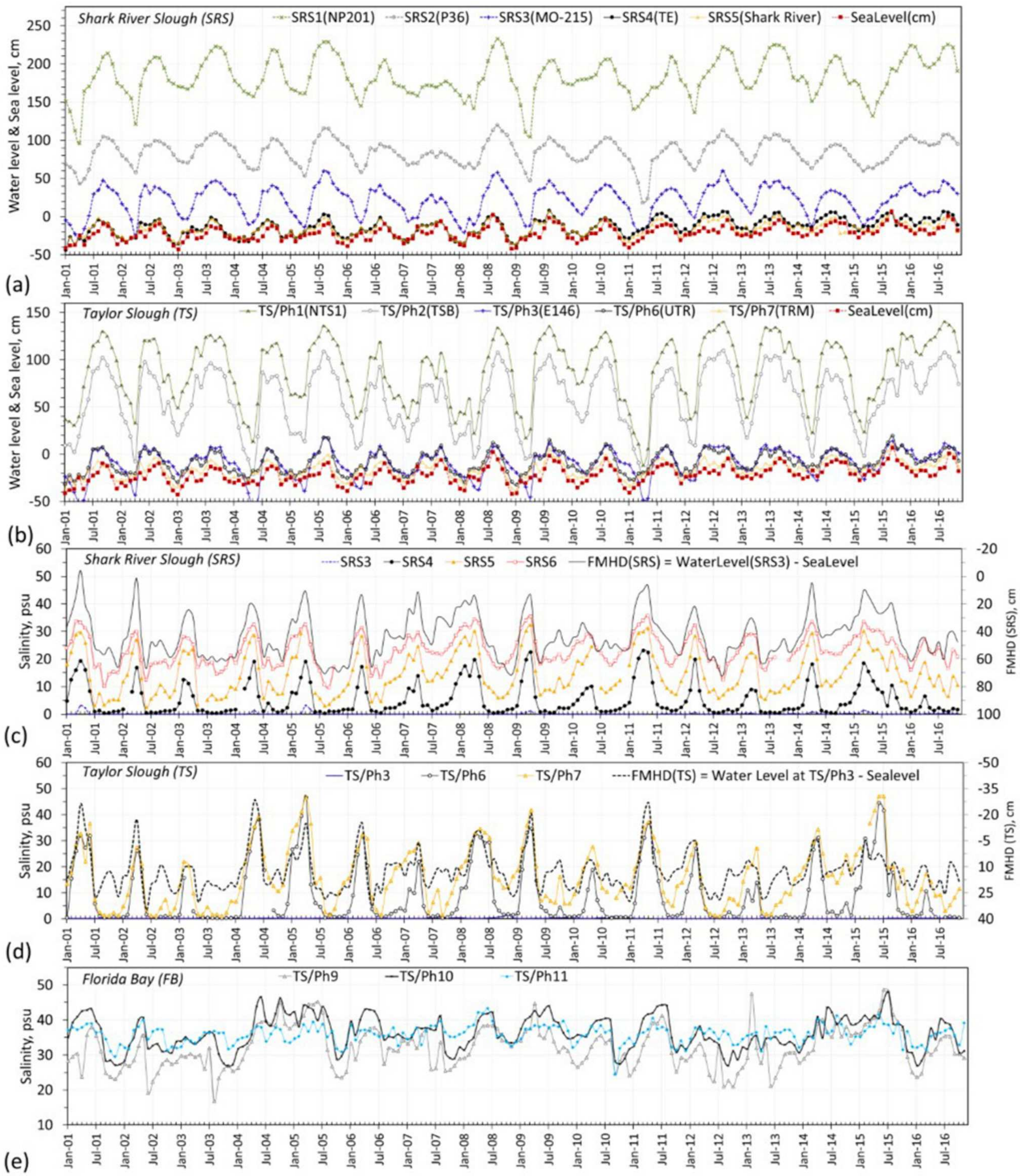


Fig. 5. Long-term mean monthly water level, sea level and salinity time series data from January 2001 to December 2016. Water level and sea level are referenced to NAVD 1988 datum. (a) Water level in Shark River Slough (SRS) and sea level, (b) water level in Taylor slough (TS) and sea level, (c) salinity in SRS, (d) monthly salinity in TS, and (e) Salinity in Florida Bay (FB).

segment observations, low observations (L: 15% < P ≤ 33%), observations within ±17% of the median (M: 33% < P ≤ 67%), high observations (H: 67% < P ≤ 85%) and observations exceeding the 85th percentile and termed higher high (HH: P > 85%) (implementation of Eq. (2), Fig. 2b). These PRI's were applied across all variables. The PRI selection criteria were: (1) statistical margins based on standard deviation of approximate normal distribution at (±0.5σ and ±1σ), (2) to provide sufficient flexibility to use the datascape for ecological assessment such as frequency of exposure to ranges of salinity; (3) to capture spectrum of seasonal variability, and (4) to extract mapping percentiles for data sets as small as 9 observations (e.g. water quality sampling interval for SRS4 to SRS6 is a 3-day composite that translates to a maximum of 10 observations per month).

As all the variables used in this application are in the same (daily) time step and assessed with similar user-defined criteria, the example procedures employed to map sea level in section 3.3 are used to setup input/output and evaluation using PRIME. Salinity responses of SRS4 and TS/Ph6 to changes in freshwater level, FMHD and sea level were evaluated using contingency tables based on the PRIME datascapes. South Florida experiences two seasons, a summer-time wet season (May to October) and a winter-time dry season (November to April). As wet season is characterized by low salinity and high-water level, sea level and FMHD (Fig. 5), the 33rd percentile salt-scapes are compared with 67th percentile hydro-scapes, and vice versa.

4.3. Results

4.3.1. Long-term spatial variability

The observed gap among water level PCs reflected the relative ground elevations of the individual stations along the freshwater flow direction (Fig. 6a) in SRS and TS. The water level PCs for sites SRS4 & 5 as well as TS/Ph3, 6 & 7 tended to cluster together following a similar pattern as sea level. Long-term water level PCs were above sea level, except at TS/Ph3. Water levels at SRS3 and TS/Ph3 were below their

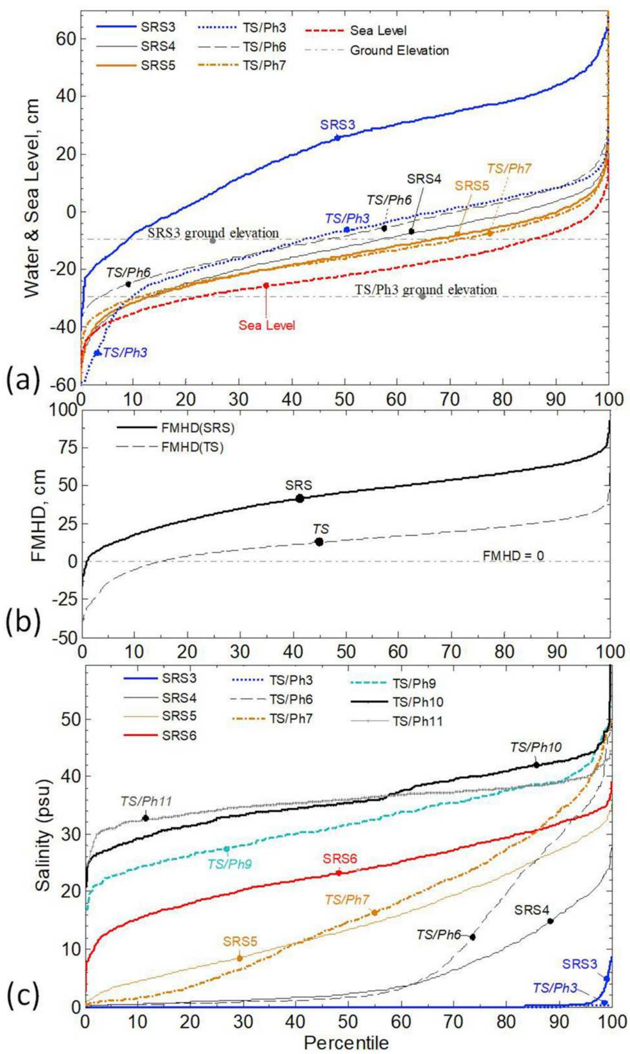


Fig. 6. Long-term hydrologic and salinity percentile curves (PCs) of FCE. (a) Water-levels, ground level and sea-level, (b) FMHD and (c) salinity in Shark River Slough (SRS3, 4, 5 and 6), Taylor Slough (TS/Ph3, 6 and 7) and Florida Bay (TS/Ph9, 10 and 11) from 2001 to 2016.

respective mean ground level for about 10% of the time (Fig. 6a). Water levels at TS/Ph3 were below sea level for 5% of the time. Sea levels exceeded mean ground levels of SRS3 and TS/Ph3 for 15% and 75% of the time (Fig. 6a). Water-level trends at SRS5 and TS/Ph7 are relatively similar but SRS4 was consistently lower than TS/Ph6 by ~5 cm. Compared to the sea-ward hydraulic gradient of water levels in SRS, the order of water level PCs between TS/Ph3 and TS/Ph6 interchanged at the lower (<~35 percentile) and higher (>90 percentile) ends (Fig. 6a). FMHD in SRS was higher compared to TS/Ph3 (Fig. 6b).

Salinity levels were below 2 psµ 95% of the time at SRS3 and below 10 psµ for about two thirds of the time at SRS4 (Fig. 6c). Conversely, salinity at FB site TS/Ph10 ranged between 25 and 50 psµ. The seasonal variability of salinity was higher at SRS4 and TS/Ph6 compared to SRS5 and TS/Ph7, their respective downstream sites. Compared to SRS, TS experiences extreme high and low salinity levels (Fig. 6c). The maximum salinity levels in SRS and TS are bounded by salinity levels at SRS6 and TS/Ph10, respectively. Despite their relatively close geographic proximity, the gap between salinity PCs at TS/Ph6 and TS/Ph7 increased steadily up to the 65th percentile after which the gap decreased rapidly while the salinity values at both sites approached those observed at FB. TS/Ph6 displayed the highest rate of rise in salinity after the ~65th percentile from 10 psµ to 50 psµ (Fig. 6c). The long-term averages and values at percentiles defining the range of indices are listed in Table 1.

4.3.2. Cross site annual and seasonal datascapes

4.3.2.1. Hydro-scapes (water level, sea level and FMHD). Annual water levels and sea level showed increasing trends from 2001 to 2016 (Fig. 7a). These temporal patterns is reflected in the hydro-scapes as change in status levels from L to M and M to H. Relatively dry years (e.g. 2001, 2005, 2007, 2011, and 2015) exhibited either LL/L status in freshwater level at the 33rd percentile and L status at the mean and median. High (H) water level status was observed at the 33rd percentile in 2016. H/HH water level status became dominant at most SRS and TS sites at the median and 67th percentile between 2012 and 2016 (Fig. 7a). H/HH water level status was most frequently observed in 2016. The FMHD status levels were L/M at the mean and 33rd percentile, and M/H at 67th percentile for both SRS and TS. Unusually low and high FMHD status levels were observed in SRS in 2015 and 2005, respectively (Fig. 7a).

The monthly hydro-scape across all sites showed strong seasonality with the L/LL and H/HH status generally aligned with the dry and wet months, respectively (Fig. 7b). Water levels and sea level status were H in September and October at the 33rd percentile and H/HH from August to November at 67th percentile (Fig. 7b). Water levels at the 33rd percentile indicated a three-month lag in LL status at the upstream freshwater sites compared to the salt-water sites in both SRS and TS (Fig. 7b). The FMHD for both SRS and TS tended to be lowest between March and June at the 33rd percentile and between April and May at the 67th percentile (Fig. 7b).

4.3.2.2. Salt-scape. Annual salinity generally shifted from L to M to H in SRS and TS at the 33rd percentile (Fig. 7c) between 2001 and 2015. The years 2008 and 2015 have either H or HH salinity status at the 67th percentile for all sites, except TS/Ph10 (Fig. 7c). The months of March through June corresponded with H and/or HH salinity conditions in both SRS and TS at both percentiles mapped (Fig. 7d).

Higher salinity status was observed in the FB sites between the months of March and August, with H status extending in to wet months at the 67th percentile (Fig. 7d).

4.3.3. Temporal (month-year) datascapes

Dry months from March to June were characterized by low FMHD and high salinity, whereas August to November was dominantly high

Table 1
List of variables, sites, and values corresponding to specific percentiles of the long-term data (2001 to 2016).

Variable	Site	Mean	Percentile						
			~0.01	15	33	50	67	85	~99.99
Water level (cm)	SRS1	182.8	88.4	162.2	171.3	178.3	196.3	213.1	237.7
	SRS2	84.5	7.9	67.4	76.5	86.6	94.5	101.5	122.2
	SRS3	21.3	-41.5	-2.7	14	26.2	32.9	40.2	67.7
	SRS4	-12.5	-68	-28	-18.6	-12.2	-6.1	1.5	60
	SRS5	-15.5	-65.5	-28.3	-20.7	-15.2	-9.8	-3	40.5
	TS/Ph1	88.3	-24.4	43	72.5	98.8	113.4	125.3	148.1
	TS/Ph2	58.2	-42.4	17.7	41.8	64	81.7	95.1	121.9
	TS/Ph3	-9.2	-64.6	-24.4	-15.2	-6.4	0	6.4	30.8
	TS/Ph6	-8.0	-42.1	-21.6	-14.6	-8.5	-1.8	5.5	69.2
	TS/Ph7	-15.7	-51.2	-27.7	-20.7	-16.2	-11.3	-4	79.2
Sea level (cm)	Key West	-21.5	-54.1	-32.5	-26.5	-22.1	-17.2	-9.8	18.4
FMHD (cm)	SRS	42.9	-16.2	22.8	37	45.6	52.4	60.6	93
	TS	12.2	-42.3	0.2	8.9	14.1	18.7	24.8	58.4
Salinity (psu)	SRS3	0.2	0	0	0	0	0	0.2	8.7
	SRS4	5.4	0	0.7	1.3	2.2	5.3	12.8	27.7
	SRS5	14.7	0.1	5.5	9.4	13.4	18.4	25.2	35.3
	SRS6	23.5	2.5	16.8	20.9	23.3	26.6	30.5	39.2
	TS/Ph3	0	0	0	0	0	0	0	0
	TS/Ph6	8.7	0	0.5	0.9	1.6	6.1	24.9	49.3
	TS/Ph7	16.1	0.1	2.4	8.0	14.7	21.3	30.0	49.7
	TS/Ph9	32.0	16.8	25.6	28.9	31.7	35.4	38.8	48.6
	TS/Ph10	36.1	26.3	31.5	34.3	35.5	39.8	42.4	47.9
	TS/Ph11	35.9	24.4	33.3	35.2	36.0	37.3	38.7	44.4

FMHD and low salinity both in the 33rd and 67th percentile (Fig. 8). In general, salinity levels were H/HH in the dry season from March to July at SRS4 and TS/Ph6 (Fig. 8b and c). Among the sixteen years, 2015 had the longest L/LL FMHD status (10 months) that extended into the wet season corresponding to seven months of H/HH salinity status. In general, the FMHD from July to November decreased over the sixteen years (Fig. 8).

4.3.4. Evaluation of datascapes

Relationships between the month-year saltscapes at SRS4 and TS/Ph6 with sea level, water level and FMHD were evaluated (see Table 2). Salinity at SRS4 and TS/Ph6 showed strong association and inverse correspondence with water levels (CC>0.5) and FMHD (DCC ≥ -0.58) (Table 2). Sea levels at Key West and salinity levels at SRS4 were weakly associated. There was no statistically significant association between salinity at TS/Ph6 and sea level (P value > 0.05) (Table 2). FMHD had a stronger association with salinity at TS/Ph6 compared to freshwater levels at TS/Ph3. Overall, FMHD had a consistent level of association with salinity at SRS4 and TS/Ph6. Water levels at TS/Ph3 showed a marginally better association with sea level compared to SRS3 (Table 2).

4.4. Discussion and interpretation

4.4.1. PRIME is a synthesis platform for long-term data-sets

With increasing volumes of data, exclusive use of descriptive statistical summaries even if they include estimates of variance over certain time intervals, may mask the impact of disturbances (e.g., hurricanes, drought, storm water surges and saltwater intrusion, etc.) or other sources of abrupt shifts. FCE-LTER has been collecting long-term hydrologic and ecological data along the SRS, TS and FB transects of the ENP since 2000 (Figs. 4 & 5, and Table 1). Results of PRIME extend previous FCE data synthesis efforts (e.g. Childers, 2006; Davis et al., 2018; Dessu et al., 2018; Kelble et al., 2007) and provide novel insight to the underlying hydrologic and salinity pattern at individual sites as well as potential connectivity across the three FCE transects for the five variables investigated. Long-term PCs captured signature pattern of the specific variables at each site. PRIME datascapes visualize patterns in finer resolutions to capture processes of interest, identify inherent relationships and correlations across variables or sites.

4.4.2. Long-term PCs show signature patterns of water level and salinity across FCE-LTER sites

Compared to the monthly average time series plots (Fig. 5), hydrologic and salinity PCs provided a clearer picture of underlying patterns (Fig. 6) at the site. For instance, water levels at SRS3 and TS/Ph3 experienced similar percentiles of wetness and dryness with water levels at or above their respective ground surface elevations for at least 90% of the daily observations from 2001 to 2016 (Fig. 6a). Vulnerability to sea level rise can be inferred from the long-term PCs when sea level exceeded the ground surface elevation at a site, which was 75% and 15% of the daily observations at TS/Ph3 and SRS3, respectively (Fig. 6a). Extended draw down in water levels either due to drought or a decrease in freshwater inputs could expose TS to increasing coastal erosion replacing coastal wetlands with non-vegetated open water areas (Ross et al., 2000; Todd et al., 2012; White and Kaplan, 2017). With the average sea level rise of 7.7 ± 4.3 mm/year from 2001 to 2016 at Key west (Dessu et al., 2018), the estuarine region of TS could be below sea level within ten to fifteen years.

As expected, salinity increased downstream and with distance from freshwater marshes of SRS and TS (Fig. 6b). Interestingly, the SRS estuaries experience a narrow range of salinity between SRS4 and SRS6 over a large spatial extent compared to the wide range of salinity over a much shorter distance between TS/Ph6 and TS/Ph7. The change in the salinity PC gradient, particularly at TS/Ph6 above a percentile of 60% may be attributed to a seasonal change in hydrologic connectivity. The combined effect of a negative FMHD and higher water levels at estuarine site TS/Ph6 compared to the upstream site TS/Ph3 may facilitate salt water intrusion and increased residence time at TS/PH6 resulting in not only the rapid increase in salinity but also the hypersaline conditions observed above a percentile of 95% (Fig. 6c). Discharge of brackish groundwater within the region of TS/Ph6 and TS/Ph7 also account for the higher salinity observed at those sites (Price et al., 2006). Small spikes in salinity observed at SRS3 are driven by the higher tidal influence in SRS (Smith and McCormick, 2001).

4.4.3. FMHD can inform freshwater delivery to coastal Everglades

FMHD along Shark River and Taylor Sloughs of ENP measures landscape-level vulnerability of downstream freshwater marshes through the oligohaline ecotone to reductions in freshwater flow (caused by drought and water management) and sea level rise. Recent

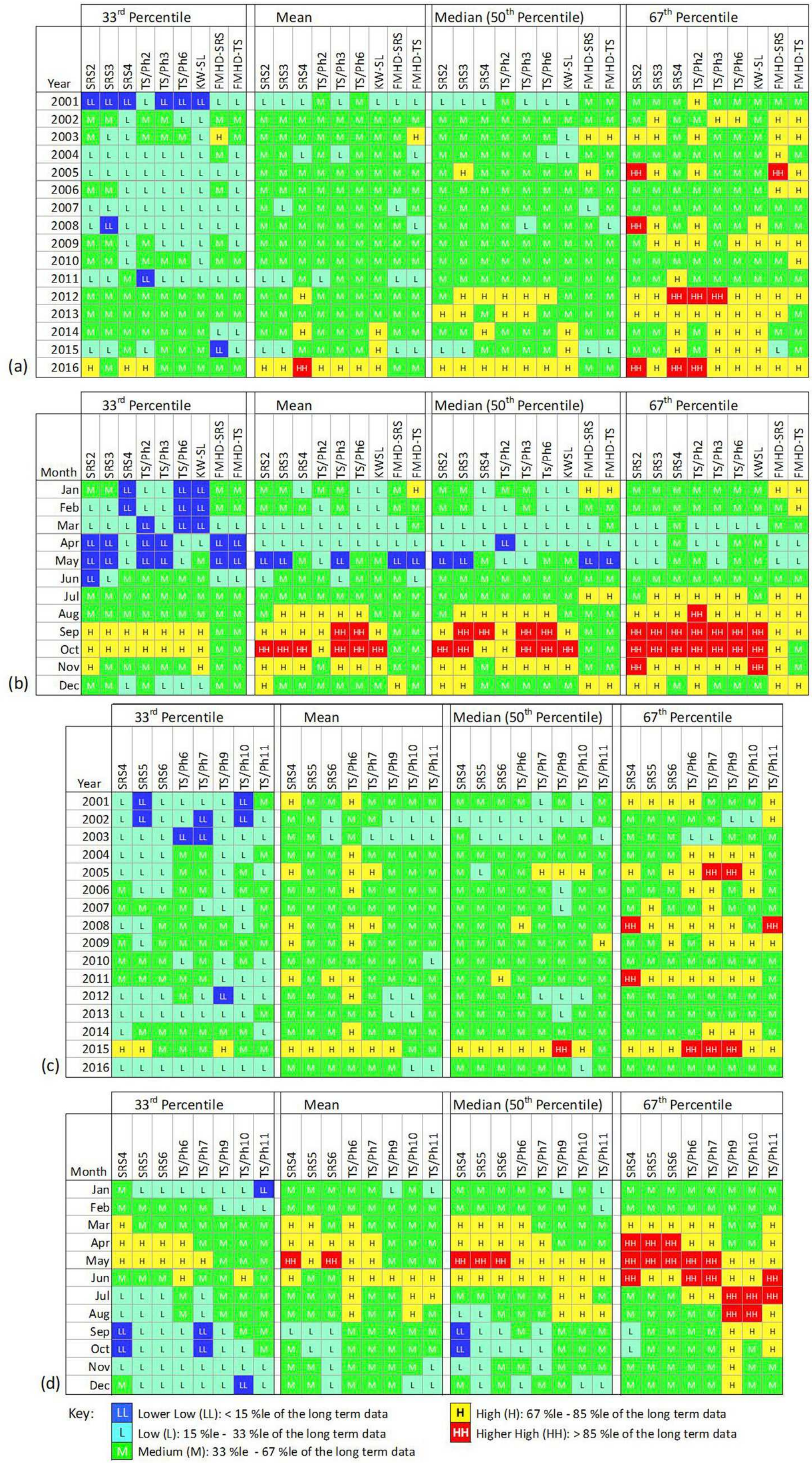


Fig. 7. Annual and seasonal (monthly) PRIME hydro-scape and salt-scapes of the FCE at 33rd and 67th percentile as compared to five percentile ranges of their respective long-term trend. (a) Annual water-level, sea level and FMHD datascapes, (b) Seasonal water-level, sea-level and FMHD (Fresh to Marine Head difference), (c) annual salinity, and (d) seasonal salinity.

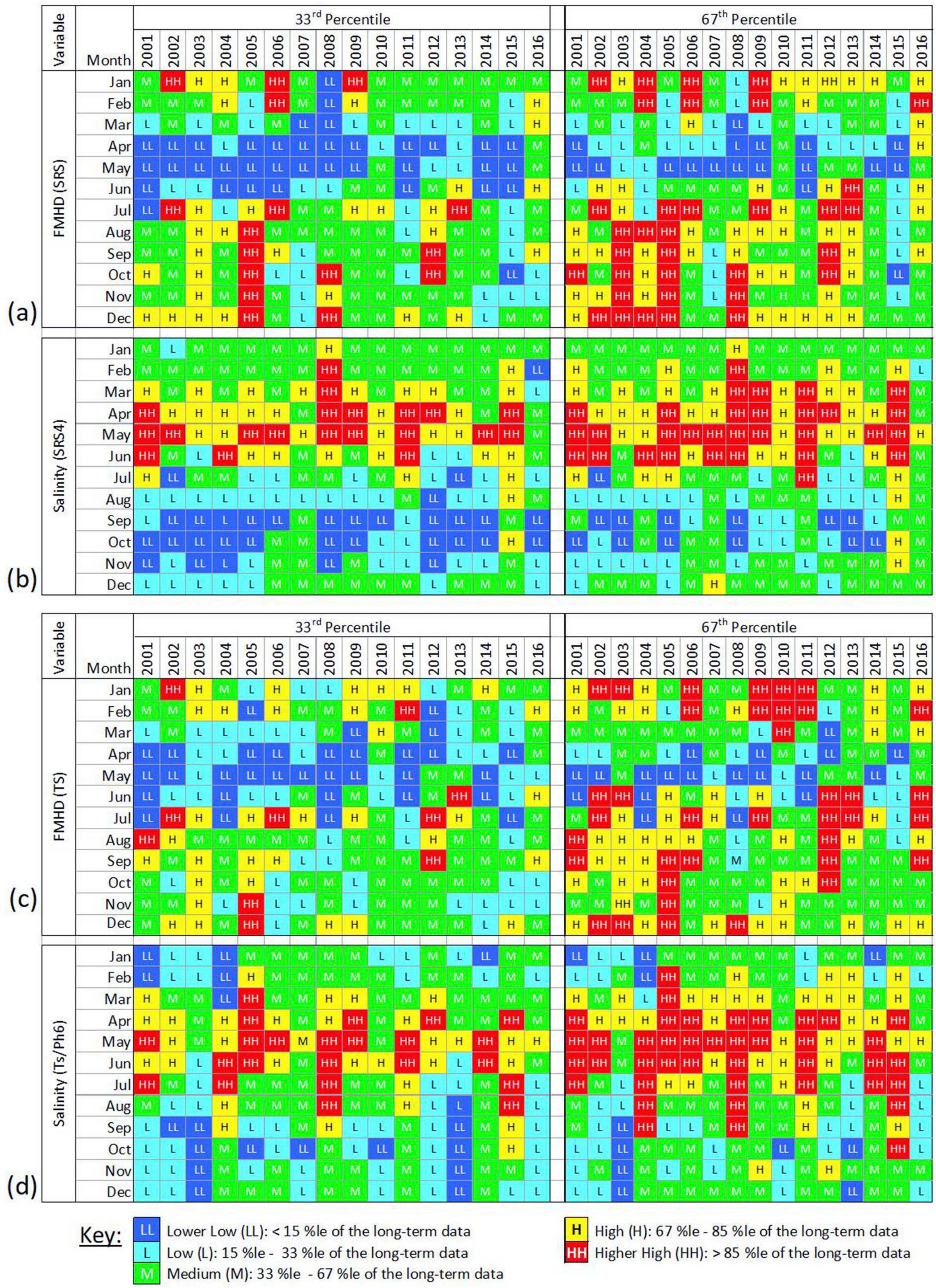


Fig. 8. Fresh-to-marine head difference (FMHD) and salt-scapes by month and year at 33rd and 67th percentile as compared to five percentile ranges of their respective long-term trend. (a) FMHD at Shark River Slough (SRS), (b) Salinity at SRS4, (c) FMHD at Taylor Slough (TS) and (d) Salinity at TS/Ph6. Refer Fig. 4 for map legend and description.

Table 2
Contingency table evaluation of month-year PRIME datascapes for selected pair of sites and variables. Chi-square (x^2) test and *P* value are used to determine statistical significance ($P < 0.05$) of the association represented by the contingency table. Contingency Coefficient (CC) and Diagonal Correspondence Coefficient (DCC) measure the level of association.

Datascape 1			Datascape 2			Summary Statistics		
Variable	Site	%le	Variable	Site	%le	x^2	CC[+]	DCC[++]
Salinity	SRS4	33	WL	SRS3	67	332.7	0.80	-0.84
		67			33	252.7	0.75	-0.80
		33	SL	KW	67	76.6	0.53	*-0.55*
		67			33	53.8	*0.47*	*-0.54*
		33	FMHD	SRS	67	219.4	0.73	-0.72
		67			33	169.4	0.68	-0.73
	TS/Ph6	33	WL	TS/Ph3	67	113.1	0.61	-0.60
		67			33	104.2	0.59	-0.65
		33	SL	KW	33	*25.3*	*Not significant*	
		67			67	*20.2*	*P value > 0.05*	
		33	FMHD	TS	67	147.9	0.66	-0.64
		67			33	144.2	0.65	-0.68
WL	SRS3	33	SL	KW	33	74.0	0.53	0.58
		67			67	69.3	0.52	0.58
	TS/Ph3	33	SL	KW	33	109.3	0.60	0.64
		67			67	102.2	0.59	0.61

Abbreviations: %le = percentile; WL = water level; KW = Key West; SL = sea level; FMHD = fresh-to-marine head difference.
[+] CC $\geq$ 0.5 suggests strong association between datascapes.
[++] |**DCC**| $\geq$ 0.58 suggests strong correspondence between datascapes for a 5 $\times$ 5 contingency table.

work demonstrated that a combination of salt exposure and dry-down makes FCE systems particularly vulnerable to peat collapse (Dessu et al., 2018; Wilson et al., 2018). FMHD PCs show the hydraulic gradient for freshwater marshes of SRS and TS relative to sea level (Fig. 6b). The higher positive FMHD in SRS may provide a higher hydraulic head to drive fresh surface and groundwater toward the coastal estuaries. Since FMHD is below zero in TS 15% of the time, all LL levels of TS FMHD datascapes represent sea level higher than freshwater levels at TS/Ph3. TS estuaries are at exposure to salt water intrusion at least 33% of the time in April and 50% of the time in May (Fig. 7b).

PRIME datascapes also highlighted temporal and spatial hot-spots for alternative mitigation and management options. The 2–3 months lag between the lowest freshwater levels relative to sea level corresponded with the lowest FMHD between April and June (Fig. 7b). This difference in status level is reflected in the FMHD-scapes as L/LL status. The dry season freshwater marsh water levels in SRS and TS lag by one status level from sea level in April and May resulting in extreme low FMHD and high salinity levels. Redistributing freshwater deliveries from August and September to April and May might increase the FMHD and reduce the extreme high salinities in those later dry season months. Compared to SRS, freshwater deliveries to TS may need to consider the prevailing salinity conditions and the change in salinity over a short distance to provide enough residence time to freshwater deliveries.

Water levels in the freshwater marsh sites of FCE depend on the amount of freshwater delivery (Dessu et al., 2018; Karamperidou et al., 2013). Hence, the relationship between freshwater levels and salinity can indicate the effects of water management on salinity levels. Monthly salt-scapes at SRS4 displayed a better correspondence with the water levels at SRS3 over the sixteen year period (Table 2). However, it is important to factor in the effect of sea level rise which will continue to reduce the effect of freshwater levels on salinity. The level of correspondence between freshwater marsh water levels (SRS3 and TS/Ph3) and sea level (DCC < 0.65, Table 2) indicates a possible mismatch in timing of water level and sea level patterns. Comparison of the long-term monthly and annual hydro-scapes (Fig. 6a and c) with salt-scapes (Fig. 6b and d) showed that FMHD is a consistent predictor of salinity status in SRS and TS compared to either freshwater levels or sea level.

4.4.4. Salinity in Florida Bay is influenced by freshwater levels in SRS and TS

PRIME datascapes also suggest a potential correlation between salinity levels in FB with FMHD of SRS and TS. Even though the effect of freshwater deliveries in TS and salinity levels in FB has been widely documented (Nuttle et al., 2000; Swart and Price, 2002), the link of water levels in SRS and salinity levels in FB is not yet clearly established (Marshall et al., 2011). The annual and seasonal salt-scapes of FB indicated a correspondence of high salinity levels at 67th percentile with low FMHD of either SRS or TS (Fig. 7a and c). The monthly salt-scapes of FB also showed a similar correspondence, but lagged by one or two months (Fig. 7b and d). Hence, the high salinity levels in FB may be reduced by maintaining suitable FMHD to drive fresh surface and groundwater flow from TS and SRS to FB.

4.4.5. PRIME datascapes facilitate detection and tracing of disturbance

Extreme disturbances and their impact can be detected as a single or successive hot-spot or nonconformity in PRIME datascapes. Droughts and hurricanes/tropical storms are major disturbances on the hydrology and water quality of the Everglades (Davis et al., 2018). The impact of Hurricanes and Tropical storms in 2003, 2005 and 2008 is observed in the PRIME results with higher water levels in those annual hydro-scapes and high levels of FMHD at SRS4 (Fig. 8a). Freshwater delivery to the SRS and TS also depends on the anticipation or occurrence of drought and hurricanes (Dessu et al., 2018; Light and Dineen, 1994). The anticipation of either a strong or weak hurricane complicates water management and freshwater delivery to the Everglades. Hurricanes or storm events that occur during a period of anticipated drought, or a passing hurricane changing course after water levels in reservoirs are brought to low levels may disturb the normal function of the Everglades. The excess freshwater from hurricanes helps to reduce salinity if stored and released in the following dry season.

Based on the monthly salinity status pattern, low and high salinity levels are expected from August to November and March to May, respectively. High (H/HH) salinity levels from August to February may suggest disturbance due to either lack of freshwater delivery or extended drought conditions (e.g. 2005, 2007/2008, 2011 and 2015) (Dessu et al., 2018). The dry condition in 2007 induced low water levels and FMHD that extended into 2008 until Tropical storm Fay (August 2008) resulting in a six month streak of high salinity status at SRS4 and TS/Ph6 between the months of January and September 2008 (Fig. 8). A period of M status months in the wet season at the 33rd percentile (2007 and 2015) suggests prolonged exposure of the system to above normally expected L/LL salinity levels. Prolonged salinity may induce a period of stress on the ecosystem that would have had relief from the seasonal flushing. Similarly, low salinity levels (LL/L/M) in the dry season (e.g. 2007 and 2016) may indicate high freshwater delivery due to excess freshwater stored in the upstream reservoirs from a hurricane or storm in the previous wet season, or expectation of a coming wetter year.

PRIME datascapes provided a visual representation to detect and quantify disturbances across sites (Fig. 7). The change in hydro-scapes from 2011 to 2012 stands out across SRS and TS as water levels and sea level each increased by at least one status level at the 33rd and 67th percentiles (Fig. 7a) with a corresponding decrease in salinity level by at least a level or two (Fig. 7c). The salinity levels of 2015 were unprecedented with high salinity levels observed in SRS and TS particularly at the 33rd percentile (Fig. 7c), whereas 2016 was characterized by the highest water levels and sea level, and low salinity levels (Fig. 7d).

4.4.6. PRIME can be used to evaluate potential benefits from Everglades restoration scenarios

The overall increasing trend of water levels in the freshwater marsh sites suggests progress towards restoring the pre-development flow conditions (McVoy et al., 2011; Sklar et al., 2005), as well as preserving the natural landscape of the Everglades (Nungesser et al., 2015). From a management and restoration perspective, understanding how increased freshwater flow through the Everglades can offset the adverse effects of sea level rise, maintaining suitable salinity conditions throughout the ecotone, and providing for the freshwater needs of downstream estuaries is critical for protecting and sustaining these resources for the future. While increasing freshwater levels can help to increase FMHD and combat sea water intrusion, the estuarine wetlands vulnerability to rising water levels needs to be factored in water delivery decision process. PRIME results have demonstrated the potential of using the FMHD variable to capture the benefits of increased freshwater delivery relative to sea level rise along SRS, TS and FB. Multiple CERP projects are being implemented towards restoration of the Everglades and improving freshwater delivery. PRIME datascapes from modeled restoration scenarios can help to compare benefits across plans, allowing for an understanding of increments of freshwater flow needed to offset recent rates of sea level rise. Hence, the potential benefits of CERP projects can be evaluated by mapping the FMHD datascapes from model simulation of CERP scenarios and comparing with the baseline historical relationship of FMHD and salinity datascapes.

5. Conclusions

The use of time-series analyses has become part of the statistical toolkit of many earth and environmental scientists. This is the direct result of the large volumes of data now available through hydrological and environmental observatory platforms including the LTER network, Critical Zone Observatories (CZOs), and the National Ecological Observatory Network (NEON). The need for effective, statistically robust, and easily interpretable time-series analyses is now required to compare patterns within and across these rich data sets and environmental gradients. As a direct result of this need throughout the earth and environmental science community, we developed PRIME as a tool based on the statistical advantages and flexibility of percentiles. PRIME results can directly assist environmental assessment and monitoring by comparing both magnitude and recurrence. PRIME produces visual datascapes summarizing a series of comprehensive analyses. Datascapes also provide a visual output that can be used to communicate results and findings to stakeholders and which demonstrate the importance of maintaining long-term observatories and data repositories. We demonstrate how a complex hydrological system such as the FCE can be mapped using PRIME to explore and understand processes across variables, time and space. We advocate the use of PRIME as a new statistical tool to support cross-cutting spatio-temporal analyses at global, regional and local scales and build comparative relationships of long-term ecological and hydrological studies.

Acknowledgment

This material is based upon work supported by the National Science Foundation (NSF) through the Florida Coastal Everglades Long-Term Ecological Research program under Grant No. DEB-9910514 [2000–2006] and Grant No. DBI-0620409 [2007–2012], and Grant No. DEB-1237517 [Dec. 2012–2018]. Support was also provided by the NSF under grant DEB-1545288, through the Long-Term Ecological Research Network Communications Office (LNCO) and the National Center for Ecological Analysis and Synthesis (NCEAS). Additional support for this work was also provided by The Everglades Foundation, the Southeast Environmental Research Center, and the Department of Earth and Environment of the School of the Environment, Arts and Society of the College of Arts, Sciences and Education of Florida International University. This is contribution number 932 from the Southeast Environmental Research Center in the Institute of Water and Environment at Florida International University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envsoft.2019.104580.

References

Briceno, H., 2017. Surface water quality monitoring data collected in south Florida coastal waters (FCE) from June 1989 to present. Long term ecological research network. Date Accessed 11/29/2017. https://doi.org/10.6073/pasta/ecc72ed682eb6d0e1481bdf0b88b7b57.

Callahan, J.T., 1984. Long-term ecological research. Bioscience 34 (6), 363–367.

Carey, C.C., Darner Gougis, R., Klug, J.L., O'Reilly, C.M., Richardson, D.C., 2015. A model for using environmental data-driven inquiry and exploration to teach limnology to undergraduates. Limnol. Oceanogr. Bull. 24 (2), 32–35.

Childers, D.L., 2006. A synthesis of long-term research by the Florida Coastal Everglades LTER Program. Hydrobiologia 569 (1), 531–544.

Dalton, R., 2000. NEON to shed light on environment research. Nature 404, 216.

Davis, S.E., Boucek, R., Castañeda-Moya, E., Dessu, S., Gaiser, E., Kominoski, J., Sah, J.P., Surratt, D., Troxler, T., 2018. Episodic disturbances drive nutrient dynamics along freshwater-to-estuary gradients in a subtropical wetland. Ecosphere 9 (6), e02296.

Dessu, S.B., Melesse, A.M., Bhat, M.G., Seid, A.H., Debebe, S.A., McClain, M.E., 2019. Development and application of a priority rated optimization model (PROM) for multi-sector water resource management systems. Environ. Model. Softw 113, 84–97.

Dessu, S.B., Price, R.M., Troxler, T.G., Kominoski, J.S., 2018. Effects of sea-level rise and freshwater management on long-term water levels and water quality in the Florida Coastal Everglades. J. Environ. Manag. 211, 164–176.

Everitt, B.S., 1992. The Analysis of Contingency Tables. Chapman and Hall/CRC.

Farley, S.S., Dawson, A., Goring, S.J., Williams, J.W., 2018. Situating ecology as a big-data science: current advances, challenges, and solutions. Bioscience 68 (8), 563–576.

Fourqurean, J., 2017. Florida Bay Physical Data, Everglades National Park (FCE), South Florida from September 2000 to Present. https://doi.org/10.6073/pasta/4f444ecc52b7f7eb83431d01a6d3429c. Environmental Data Initiative, Dataset accessed 06/15/2017, In: FCE-LTER (Ed.).

Gaiser, E., Childers, D., 2017. Water quality data (extensive) from the Shark River slough, Everglades national Park (FCE), from October 2000 to Present. Environmental Data Initiative. https://doi.org/10.6073/pasta/32826025b5f61d967f80aa502c712e1a. Dataset accessed 6/1/2017.

Gunderson, L.H., 1994. Vegetation of the Everglades: Determinants of Community Composition. Everglades: The ecosystem and its restoration, pp. 323–340.

Haigh, I.D., Wahl, T., Rohling, E.J., Price, R.M., Pattiaratchi, C.B., Calafat, F.M., Dangendorf, S., 2014. Timescales for detecting a significant acceleration in sea level rise. Nat. Commun. 5.

Hamilton, D.P., Carey, C.C., Arvola, L., Arzberger, P., Brewer, C., Cole, J.J., Gaiser, E., Hanson, P.C., Ibelings, B.W., Jennings, E., 2015. A Global Lake Ecological Observatory Network (GLEON) for synthesising high-frequency sensor data for validation of deterministic ecological models. Inland Waters 5 (1), 49–56.

Hanson, P.C., Weathers, K.C., Dugan, H.A., Gries, C., 2018. The Global Lake Ecological Observatory Network, Ecological Informatics. Springer, pp. 415–433.

Hobbie, J.E., Carpenter, S.R., Grimm, N.B., Gosz, J.R., Seastedt, T.R., 2003. The US long term ecological research program. AIBS (Am. Inst. Biol. Sci.) Bull. 53 (1), 21–32.

Holgate, S.J., Matthews, A., Woodworth, P.L., Rickards, L.J., Tamisiea, M.E., Bradshaw, E., Foden, P.R., Gordon, K.M., Jevrejeva, S., Pugh, J., 2012. New data systems and products at the permanent service for mean sea level. J. Coast. Res. 29 (3), 493–504.

Karamperidou, C., Engel, V., Lall, U., Stabenau, E., Smith III, T.J., 2013. Implications of multi-scale sea level and climate variability for coastal resources. Reg. Environ. Chang. 13 (1), 91–100.

Kelble, C.R., Johns, E.M., Nuttle, W.K., Lee, T.N., Smith, R.H., Ortner, P.B., 2007. Salinity patterns of Florida Bay. Estuar. Coast Shelf Sci. 71 (1), 318–334.
Kendall, M.G., 1955. Further contributions to the theory of paired comparisons. Biometrics 11 (1), 43–62.
Kim, E.-S., 2006. Development, potentials, and challenges of the international long-term ecological research (ILTER) network. Ecol. Res. 21 (6), 788–793.
Krauss, K.W., From, A.S., Doyle, T.W., Doyle, T.J., Barry, M.J., 2011. Sea-level rise and landscape change influence mangrove encroachment onto marsh in the Ten Thousand Islands region of Florida, USA. J. Coast. Conserv. 15 (4), 629–638.
Light, S.S., Dineen, J.W., 1994. Water control in the Everglades: a historical perspective. Everglades: The Ecosystem and its Restoration, 5, pp. 47–84.
Marshall, F., Smith, D., Nickerson, D., 2011. Empirical tools for simulating salinity in the estuaries in Everglades National Park, Florida. Estuar. Coast Shelf Sci. 95 (4), 377–387.
McVoy, C., Said, W.P., Obeysekera, J., VanArman, J.A., Dreschel, T.W., 2011. Landscapes and Hydrology of the Predrainage Everglades. University Press of Florida, Gainesville.
Nungesser, M., Saunders, C., Coronado-Molina, C., Obeysekera, J., Johnson, J., McVoy, C., Benscoter, B., 2015. Potential effects of climate change on Florida's Everglades. Environ. Manag. 55 (4), 824–835.
Nuttle, W.K., Fourqurean, J.W., Cosby, B.J., Zieman, J.C., Robblee, M.B., 2000. Influence of net freshwater supply on salinity in Florida Bay. Water Resour. Res. 36 (7), 1805–1822.
Price, R.M., Swart, P.K., Fourqurean, J.W., 2006. Coastal groundwater discharge–an additional source of phosphorus for the oligotrophic wetlands of the Everglades. Hydrobiologia 569 (1), 23–36.
PSMSL, P.S.f.M.S.L.P., 2016. Tide Gauge Data. Global Sea Level Observing System.
Qian, S.S., 2015. The frequency component of water quality criterion compliance assessment should be data driven. Environ. Manag. 56 (1), 24–33.
Read, E.K., Carr, L., De Cicco, L., Dugan, H.A., Hanson, P.C., Hart, J.A., Kreft, J., Read, J. S., Winslow, L.A., 2017. Water quality data for national-scale aquatic research: the Water Quality Portal. Water Resour. Res. 53 (2), 1735–1745.
Ross, M.S., Meeder, J.F., Sah, J.P., Ruiz, P.L., Telesnicki, G.J., 2000. The Southeast Saline Everglades revisited: 50 years of coastal vegetation change. J. Veg. Sci. 11 (1), 101–112.
Sklar, F.H., Chimney, M.J., Newman, S., McCormick, P., Gawlik, D., Miao, S., McVoy, C., Said, W., Newman, J., Coronado, C., 2005. The ecological–societal underpinnings of Everglades restoration. Front. Ecol. Environ. 3 (3), 161–169.
Smith, E.P., McCormick, P.V., 2001. Long-term relationship between phosphorus inputs and wetland phosphorus concentrations in a northern Everglades marsh. Environ. Monit. Assess. 68 (2), 153–176.
Spearman, C., 1904. The proof and measurement of association between two things. Am. J. Psychol. 15 (1), 72–101.
Swart, P., Price, R., 2002. Origin of salinity variations in Florida Bay. Limnol. Oceanogr. 47 (4), 1234–1241.
Titus, J.G., Richman, C., 2001. Maps of lands vulnerable to sea level rise: modeled elevations along the US Atlantic and Gulf coasts. Clim. Res. 18 (3), 205–228.
Todd, M.J., Muneepeerakul, R., Miralles-Wilhelm, F., Rinaldo, A., Rodriguez-Iturbe, I., 2012. Possible climate change impacts on the hydrological and vegetative character of Everglades National Park, Florida. Ecohydrology 5 (3), 326–336.
Todd, M.J., Muneepeerakul, R., Pumo, D., Azaele, S., Miralles-Wilhelm, F., Rinaldo, A., Rodriguez-Iturbe, I., 2010. Hydrological drivers of wetland vegetation community distribution within Everglades National Park, Florida. Adv. Water Resour. 33 (10), 1279–1289.
Trenberth, K.E., Dai, A., van der Schrier, G., Jones, P.D., Barichivich, J., Briffa, K.R., Sheffield, J., 2014. Global warming and changes in drought. Nat. Clim. Chang. 4 (1), 17–22.
Troxler, T.G., 2017. Water quality data (extensive) from the taylor slough, Everglades national Park (FCE), from April 1996 to present. Environmental data initiative. https://doi.org/10.6073/pasta/c3971619b22e1ded0f9a258e95f4e2a7. Dataset accessed 11/29/2017. TS/Ph6a,TS/Ph7a,TS/Ph8.
USGS/EDEN, 2019. Everglades Depth estimation network/USGS project. https://sofia.usgs.gov/eden.
Vogel, R.M., Fennessey, N.M., 1995. Flow duration curves II: a review of applications in water resources planning. JAWRA J. Am. Water Resour. Assoc. 31 (6), 1029–1039.
White, E., Kaplan, D., 2017. Restore or retreat? Saltwater intrusion and water management in coastal wetlands. Ecosyst. Health Sustain. 3 (1).
Wilson, B.J., Servais, S., Charles, S.P., Davis, S.E., Gaiser, E.E., Kominoski, J.S., Richards, J.H., Troxler, T.G., 2018. Declines in plant productivity drive carbon loss from brackish coastal wetland mesocosms exposed to saltwater intrusion. Estuar. Coasts 41 (8), 2147–2158.
Yilmaz, K.K., Gupta, H.V., Wagener, T., 2008. A process-based diagnostic approach to model evaluation: application to the NWS distributed hydrologic model. Water Resour. Res. 44 (9).