2020

Performance Across Worldview-2 and RapidEye for Reproducible Seagrass Mapping

Megan M. Coffer

Blake A. Schaeffer

Richard C. Zimmerman
Old Dominion University, rzimmerm@odu.edu

Victoria Hill
Old Dominion University, vhill@odu.edu

Jiang Li
Old Dominion University, jli@odu.edu

See next page for additional authors

Original Publication Citation

Coffer, M. M., Schaeffer, B. A., Zimmerman, R. C., Hill, V., Li, J., Islam, K. A., & Whitman, P. J. (2020). Performance across WorldView-2 and RapidEye for reproducible seagrass mapping. *Remote Sensing of Environment*, *250*, 1-14, Article 112036. https://doi.org/10.1016/j.rse.2020.112036

Remote Sensing of Environment 250 (2020) 112036

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Performance across WorldView-2 and RapidEye for reproducible seagrass mapping

Megan M. Coffer[a,b,*], Blake A. Schaeffer[c], Richard C. Zimmerman[d], Victoria Hill[d], Jiang Li[e], Kazi A. Islam[e], Peter J. Whitman[a]

[a] ORISE fellow, U.S. Environmental Protection Agency, Office of Research and Development, Durham, NC, USA
[b] Center for Geospatial Analytics, North Carolina State University, Raleigh, NC, USA
[c] U.S. Environmental Protection Agency, Office of Research and Development, Durham, NC, USA
[d] Department of Ocean, Earth & Atmospheric Sciences, Old Dominion University, Norfolk, VA, USA
[e] Department of Electrical & Computer Engineering, Old Dominion University, Norfolk, VA, USA

ARTICLE INFO

Keywords:
Remote sensing
Atmospheric correction
WorldView-2
RapidEye
Seagrass
Image classification

ABSTRACT

Satellite remote sensing offers an effective remedy to challenges in ground-based and aerial mapping that have previously impeded quantitative assessments of global seagrass extent. Commercial satellite platforms offer fine spatial resolution, an important consideration in patchy seagrass ecosystems. Currently, no consistent protocol exists for image processing of commercial data, limiting reproducibility and comparison across space and time. Additionally, the radiometric performance of commercial satellite sensors has not been assessed against the dark and variable targets characteristic of coastal waters. This study compared data products derived from two commercial satellites: DigitalGlobe's WorldView-2 and Planet's RapidEye. A single scene from each platform was obtained at St. Joseph Bay in Florida, USA, corresponding to a November 2010 field campaign. A reproducible processing regime was developed to transform imagery from basic products, as delivered from each company, into analysis-ready data usable for various scientific applications. Satellite-derived surface reflectances were compared against field measurements. WorldView-2 imagery exhibited high disagreement in the coastal blue and blue spectral bands, chronically overpredicting. RapidEye exhibited better agreement than WorldView-2, but overpredicted slightly across all spectral bands. A deep convolutional neural network was used to classify imagery into deep water, land, submerged sand, seagrass, and intertidal classes. Classification results were compared to seagrass maps derived from photointerpreted aerial imagery. This study offers the first radiometric assessment of WorldView-2 and RapidEye over a coastal system, revealing inherent calibration issues in shorter wavelengths of WorldView-2. Both platforms demonstrated as much as 97% agreement with aerial estimates, despite differing resolutions. Thus, calibration issues in WorldView-2 did not appear to interfere with classification accuracy, but could be problematic if estimating biomass. The image processing routine developed here offers a reproducible workflow for WorldView-2 and RapidEye imagery, which was tested in two additional coastal systems. This approach may become platform independent as more sensors become available.

1. Introduction

Seagrass meadows occupy less than 0.2% of the ocean floor, but store approximately 10% of the ocean's carbon per year (Duarte et al., 2013; Fourqurean et al., 2012). Despite their prominent role in the global carbon cycle, carbon storage in these ecosystems is poorly quantified, with estimates ranging from 4.2 to 19.9 Pg of total organic carbon currently stored in seagrass ecosystems (Fourqurean et al., 2012). This uncertainty results, to a large extent, from a poor understanding of global seagrass coverage, ranging between 150,000 and 4,320,000 km^2 (Duarte, 2017). The 2016 Commission for Environmental Cooperation report concluded that seagrasses represent the greatest data gap in Blue Carbon habitat mapping, where Blue Carbon represents carbon captured by ocean and coastal sediments (CEC, 2016).

Seagrass extent has been poorly quantified largely due to logistical and labor challenges involved in aerial and ground-based efforts, including the high cost of comprehensive mapping (Björk et al., 2008; Dekker et al., 2006). In contrast, satellite remote sensing can offer an effective approach for mapping and monitoring seagrass in coastal

* Corresponding author at: U.S. Environmental Protection Agency, Office of Research and Development, Durham, NC, USA.
E-mail address: mmamanat@ncsu.edu (M.M. Coffer).

https://doi.org/10.1016/j.rse.2020.112036
Received 4 May 2020; Received in revised form 31 July 2020; Accepted 5 August 2020
Available online 18 August 2020

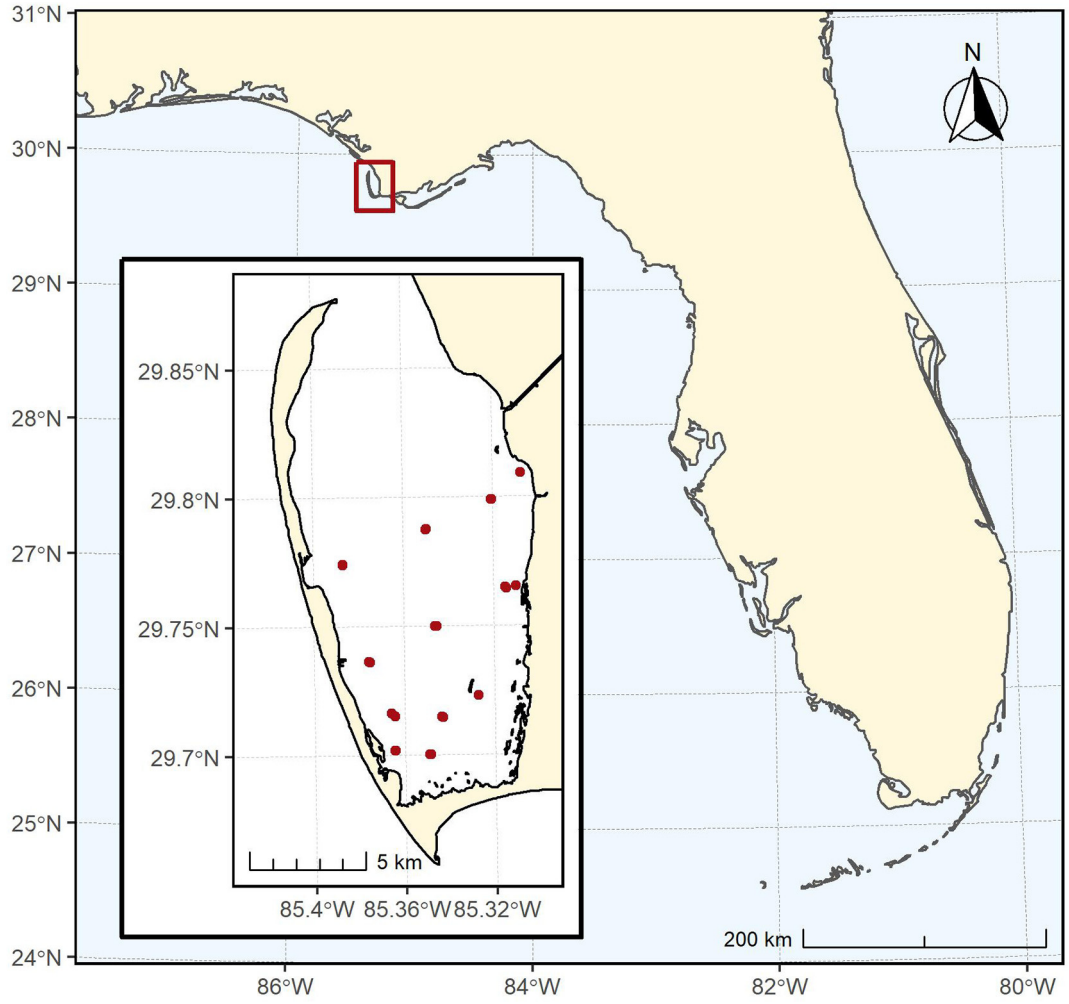

Fig. 1. St. Joseph Bay is located along the Florida panhandle in the northern Gulf of Mexico (29.797°N, 85.353°W). Hyperspectral field measurements were taken at 24 locations throughout the bay (red points) on 2, 6, and 9 Nov 2010. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ecosystems (Baumstark et al., 2013). Several classification techniques have been used for seagrass detection, including maximum likelihood, object based image analysis (OBIA), support vector machines, and random forests. These algorithms have been applied to a range of satellite platforms (Hill et al., 2014), with accuracies ranging from 73% to 85% (Traganos and Reinartz, 2017; Pasqualini et al., 2005; Pu et al., 2014; Pu and Bell, 2013; Meyer and Pu, 2012; Traganos and Reinartz, 2018).

The Landsat series, which provides imagery at 30 m spatial resolution in five visible to near infrared spectral bands (Dekker et al., 2005; Mumby et al., 1997; Ward et al., 2003; Hossain et al., 2015), and other moderate-scale sensors such as Sentinel-2A at 10 m spatial resolution (Fauzan et al., 2017; Thalib et al., 2018), have been used to monitor seagrass ecosystems. Commercial platforms such as RapidEye at 5 m (Traganos and Reinartz, 2018), PlanetScope at 3 m (Wicaksono and Lazuardi, 2018), Quickbird-2 at 2.4 m (Phinn et al., 2008a; Lyons et al., 2011), and WorldView-2 at 2 m (Baumstark et al., 2016; Roelfsema et al., 2014) offer finer spatial resolution, but have been included in fewer studies due to costs associated with the imagery. High spatial resolution (< 10 m) allows for improved seagrass biomass retrieval, productivity estimates, and patch size statistics (Hill et al., 2014).

Although commercial platforms have been successfully employed for seagrass delineation, previous studies have not presented a consistent nor transparent processing regime, limiting reproducibility. Atmospheric correction, for example, is an essential step for creating separability in spectra (Huang et al., 2016), required for both classification (Knudby and Nordlund, 2011) and comparison across space and time (Coppin et al., 2004). However, optimization of atmospheric correction parameters specific to commercial satellite platforms remains largely an ad hoc process, limited at least in part by the low signal to noise ratio of commercial platforms over dark pixels.

Moreover, radiometric performance of commercial imagery in coastal environments has not been quantitatively assessed. Analyzing agreement of radiometric measurements is important for increasing user confidence, determining appropriate applications, and to allow informed use of the imagery (Mélin et al., 2016). An understanding of radiometric performance is necessary for tasks such as merging different datasets (Pottier et al., 2006) and performing trend analyses (Vantrepotte and Mélin, 2011).

Machwitz et al. (2014) found RapidEye reflectance values to be higher than field observations in land-based applications. Less quantitative validation exists for WorldView-2 data. Looking into the literature of the radiometric performance of WorldView-2 in terrestrial systems, Latif et al. (2012) quantified disagreement between WorldView-2 reflectance measurements and field measurements of a forest habitat. The two datasets varied by about 5%, and reflectance measurements in shorter shorter wavelengths were higher in WorldView-2 imagery as compared to field observations and lower in longer wavelengths.

We used DigitalGlobe's WorldView-2 satellite platform, which

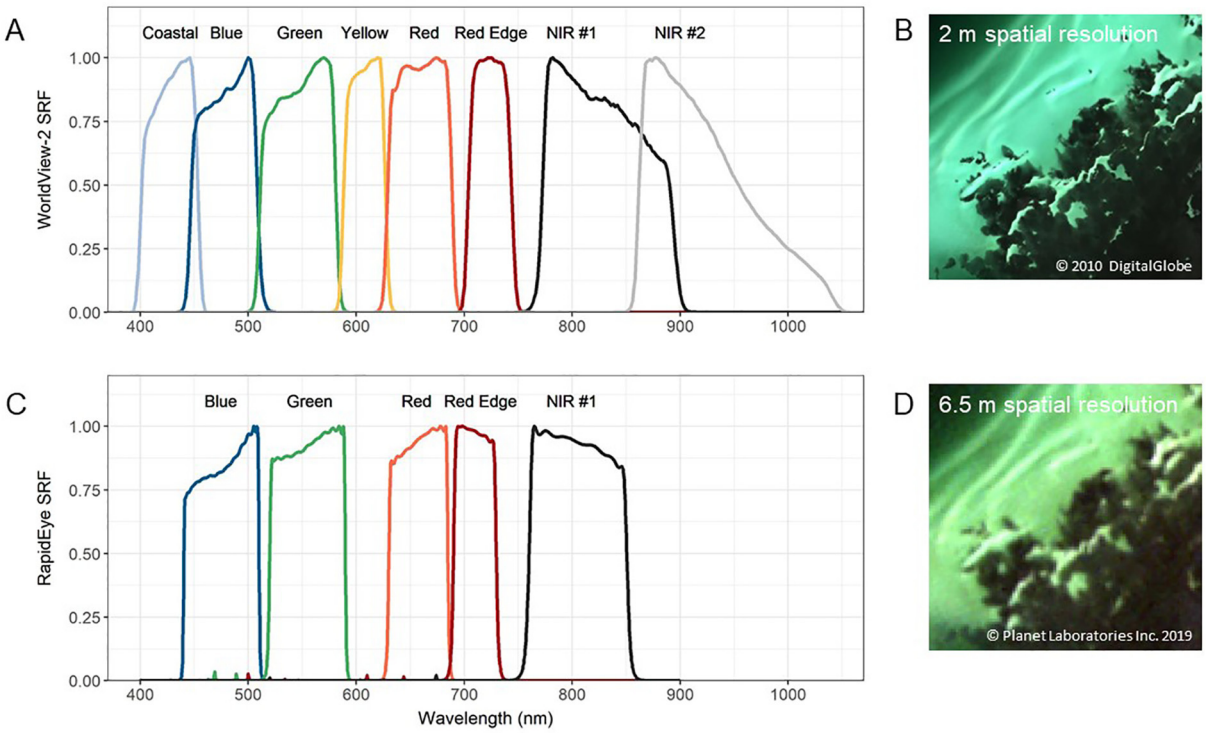

Fig. 2. (A) Spectral response function (SRF) of visible and near infrared (NIR) channels for DigitalGlobe's WorldView-2 satellite and (B) a true color image of 2 m WorldView-2 data. (C) SRF of Planet's RapidEye satellite constellation and (D) a true color image of 6.5 m RapidEye data (copyright Planet Laboratories Inc. 2019 all rights reserved).

collects imagery in eight multispectral bands at 2 m resolution, and Planet's RapidEye satellite constellation, which collects imagery in five multispectral bands at 6.5 m resolution. Here we report our efforts to:

1. Develop a reproducible workflow for processing commercial satellite imagery from DigitalGlobe's WorldView-2 satellite platform and Planet's RapidEye satellite constellation.
2. Evaluate and compare radiometric performance of each sensor by analyzing agreement between satellite-derived reflectance measurements and field-derived reflectance measurements.
3. 3. Apply a deep learning approach presented in Islam et al. (2020) to estimate seagrass coverage and assess classification agreement with an aerial photointerpretation.

2. Data and methods

2.1. Study area

St. Joseph Bay is located along the Florida panhandle in the northern Gulf of Mexico (29.797°N, 85.353°W; Fig. 1). Selection of St. Joseph Bay was motivated by the presence of existing field data and extensive familiarity with the site (i.e., Hill et al., 2014). St. Joseph Bay is approximately 24 km long and 10 km wide, opening in the north to the Gulf of Mexico. The Gulf County Canal flows into the northeast portion of the bay at the town of Port St. Joe. Water depth ranges up to 12 m at the center of the bay. At the southern end, an extensive shallow area less than 3 m deep occupies about a third of the entire bay (Hill et al., 2014).

Waters in St. Joseph Bay are characterized by moderately high concentrations of phytoplankton, colored dissolved organic matter (CDOM), and suspended nonalgal particulate (detritus and sediment), creating a potentially challenging optical environment (Conmy et al., 2017). The seagrass within the bay is dominated by turtlegrass (*Thalassia testudinum*) which forms dense meadows at depths less than 2 m (Hill et al., 2014). Manateegrass (*Syringodium filiforme*), shoalgrass (*Halodule wrightii*), widgeongrass (*Ruppia maritima*), and stargrass (*Halophila engelmannii*) are also present in lower densities. Manateegrass can reach depths of up to 2 m while the other species tend to dominate at shallower depths.

2.2. Satellite data

Satellite data were obtained from two commercial platforms, WorldView-2 and RapidEye (Supplementary Table S1). DigitalGlobe launched WorldView-2 in October 2009, offering multispectral data at a 1.84 m ground sample distance (GSD) at nadir. Basic level 1B data were obtained from DigitalGlobe's EnhancedView Web Hosting Service (evwhs.digitalglobe.com) at a spatial resolution of 2 m for 14 Nov 2010 (Maxar, 2019), which was the closest cloud-free overpass to field data collection.

RapidEye was a constellation of five satellite sensors launched in August 2008 and decomissioned in December 2019 (Planet Team, 2017). While new imagery will not be collected from the RapidEye constellation, over a decade of archived imagery exists and Planet has announced a plan to launch Super-Dove platforms with similar resolutions as RapidEye. Each Planet's RapidEye satellite offered multispectral data at a 6.5 m GSD at nadir. Basic level 1B data were obtained for the RapidEye-2 satellite through Planet Explorer (https://www.planet.com/explorer/) for 11 Nov 2010 (Planet Labs Inc, 2019), which, again, was the closest cloud-free overpass to field data collection. Off-nadir images were intentionally selected to avoid challenges presented with sunglint. View angles for the selected images ranged from 10° to 16°.

RapidEye offered five multispectral bands with four in the visible and one in the near infrared (NIR). WorldView-2 offers six bands in the visible, two in the NIR, and an order-of-magnitude finer spatial resolution than RapidEye (Fig. 2, Table 1). The dynamic range, which describes the amount of information that can be measured within a

Table 1
Spectral characteristics of each sensor.

Band	WorldView-2			RapidEye		
	FWHM	Center	SNR	Range	Center	SNR
Coastal	400–450	427.3	22.22			
Blue	450–510	477.9	7.40	440–510	475	3.10
Green	510–580	546.2	2.65	520–590	555	1.71
Yellow	585–625	607.8	1.64			
Red	630–690	658.8	1.41	630–685	657.5	1.17
Red edge	705–745	723.7	0.58	690–730	710	1.24
NIR #1	770–895	832.5	0.45	760–850	805	1.18
NIR #2	860–1040	908	0.57			

FWHM represents the full width half maximum for each band. SNR indicates the signal to noise ratio computed across the scene following Dadon et al. (2011) as the ratio between the standard deviation and the mean of the satellite signal for each band, and excluded land pixels. The exclusion of land pixels is explained in Section 2.3.

pixel, was higher in RapidEye (12-bit) than WorldView-2 (11-bit). Signal to noise ratio (SNR) was computed following Dadon et al. (2011). WorldView-2 offered improved SNR compared to RapidEye in shorter wavelengths; RapidEye offered improved SNR compared to WorldView-2 in longer wavelengths (Table 1).

2.3. Data processing

WorldView-2 and RapidEye scenes were processed to produce orthorectified, radiometrically calibrated, and atmospherically corrected images from the Level 1B data products. The order of some steps is important (e.g., atmospheric correction must be applied after radiometric calibration), but the sequence of other steps is interchangeable. Each scene was delivered as multiple tiles. The WorldView-2 scene was divided into three tiles, and the RapidEye scene into two. Data processing was performed for each tile before being mosaicked into a single scene. Satellite data processing was performed in ENVI 8.7.0 and IDL 8.7.2 (Exelis Visual Information Solutions, Boulder, Colorado).

2.3.1. Radiometric calibration

The radiometric calibration values provided with WorldView-2 imagery were adjusted using updated vicarious calibration factors provided by Kuester (2017). Radiometric calibration required the adjusted and raw calibration coefficients for WorldView-2 and RapidEye, respectively. Next, all pixel values were divided by π to convert dimensionless reflectance values to remote sensing reflectance (R_{rs}) in units of inverse steradians (sr^{-1}).

2.3.2. Atmospheric correction

A dark object subtraction (DOS) approach was used to remove atmospheric contamination from the R_{rs} signal (Chavez, 1988). Three parameters can be optimized when performing DOS: the starting spectral band, the threshold at which pixels are considered to represent dark objects, and the value of the exponent used to correct shorter wavelengths. This study employed the red edge band as the starting spectral band, a threshold of the median of lowest 5% of the distribution of red edge pixel values, and a Rayleigh exponent of 4.75. NIR bands are commonly used to establish DOS adjustment values (Green et al., 2000); however, the signal to noise ratio of WorldView-2 and RapidEye were insufficient such that radiometrically calibrated values over the darkest water pixels in the image frequently yielded negative radiances in the NIR bands before atmospheric correction, preventing their use for DOS (Table 1).

Optimization of the reference contamination value and the exponent were explored by iterating through a range of values and noting the effect on the resulting spectra. Red edge distribution percentage medians ranging from 5% through 50% in increments of 5% were considered, while maintaining a constant Rayleigh exponent of 4.75. Rayleigh exponents ranging from 3.5 to 6.5 in increments of 0.25 were tested, while maintaining a constant red edge distribution percentage of 5%. Resulting spectra were visually compared against field observations.

A reproducible workflow was developed in an effort to reduce the arbitrary nature of selecting dark pixels used for DOS (Fig. 3). A shapefile indicating the extent of estuaries along the coastline was used as an initial separation between land and water (Schaeffer and Myer, 2020), ensuring red edge anchor values were not selected from inland water pixels such as lakes and rivers. After spatial subsetting, some pixels along the shoreline remained. Therefore, a spectral threshold was applied. The normalized difference water index (NDWI) was computed (McFeeters, 1996), and pixel values flagged as water (i.e., NDWI value above zero) were retained for dark pixel consideration.

The median of red edge R_{rs} values was computed from those representing the darkest 5% of the distribution. We assumed half the red edge radiance represented atmospheric contamination, while the remainder represented true water-leaving radiances. This assumption was motivated by Vanhellemont and Ruddick (2014) in which a linear relationship between the two surrounding bands, red and NIR, was used to characterize atmospheric contamination in Landsat 8 imagery and was further established through a sensitivity analysis in which approximately half of the red edge signal generated retrievals that did not result in negative R_{rs} values, where negative R_{rs} indicate an atmospheric overcorrection. Determination of a red edge anchor value was repeated for all tiles within the scene, and the lowest of these values was used to compute the scattering factor for the entire scene. Red edge anchor values across tiles did not differ drastically, ranging from 0.000655 to 0.000657 for WorldView-2 and 0.00269 to 0.00279 for RapidEye. The scattering factor was computed as:

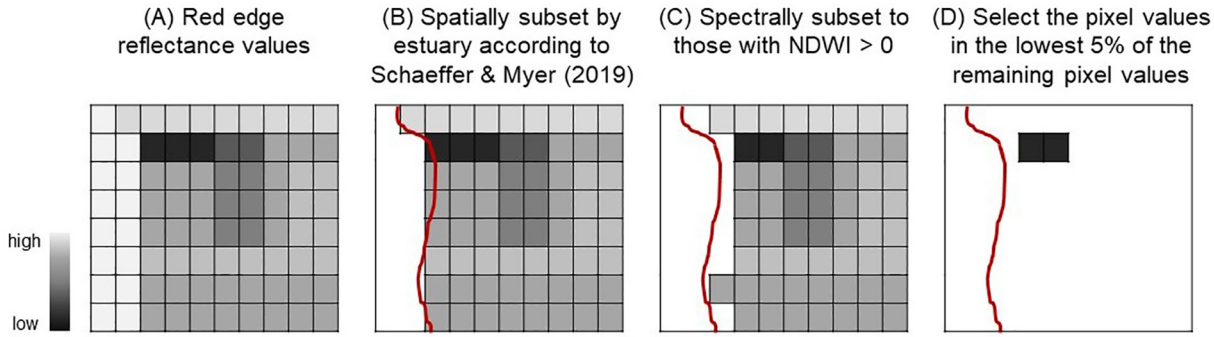

Fig. 3. (A) Image showing the distribution of red edge R_{rs} values scaled from "low" to "high" across a hypothetical image. (B) Pixels retained after applying the estuary boundary of Schaeffer and Myer (2020). (C) Additional pixels removed with NDWI > 0. (D) Remaining pixels representing the lowest 5% of the R_{rs} values in the red edge band. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$\sigma = \lambda_{red\ edge}^{4.75} \times \alpha_{red\ edge} \quad (1)$$

where σ is the scattering factor and α is the anchor value. The scattering factor was then applied to the wavelength of each band to compute the band-specific value subtracted from each pixel in each band, as:

$$\eta_{band} = \frac{\sigma}{\lambda_{band}^{4.75}} \quad (2)$$

where η is the subtraction value. The resulting band-specific values of η were then subtracted from R_{rs} values in each band for all pixels in the scene.

2.3.3. RPC Orthorectification

Rational polynomial coefficient (RPC) orthorectification was applied to geometrically correct the image and remove distortions from image capture. The Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) dataset is provided with ENVI at a spatial resolution of 30 arc-seconds, and was used in the processing protocol presented here to ensure large-scale reproducibility and since there were no terrain-specific issues over water (Danielson and Gesch, 2011).

2.3.4. Mosaicking and resampling

Orthorectified tiles were then mosaicked into a single scene. In addition to the native resolution WorldView-2 and RapidEye scenes, a third raster was generated by resampling the 2 m WorldView-2 imagery via bilinear interpolation to match the 6.5 m spatial resolution of RapidEye. Comparing native resolution to downsampled WorldView-2 imagery revealed differences stemming from spatial resolution, given spectral resolution was the same. Comparing downsampled WorldView-2 imagery to native resolution RapidEye imagery revealed differences stemming from spectral resolutions given most other image parameters were held constant.

2.4. Field measurements of reflectance

Satellite estimates of atmospherically corrected R_{rs} were compared against hyperspectral reflectance measurements taken at 24 locations throughout the bay (Fig. 1). Statistical analyses were conducted in R Version 3.4.0 (R Core Team, 2017). Data were collected as described in Hill et al. (2014). Hyperspectral field measurements of R_{rs} were resampled to match the eight WorldView-2 and five RapidEye spectral bands by taking the mean of the hyperspectral reflectance values within the FWHM for each band. For each of the 24 observations, field measured R_{rs} was compared to satellite derived R_{rs} using the mean absolute deviation (MAD) and associated bias.

2.5. Classification of seagrass extent

A deep convolutional neural network (DCNN) developed by Islam et al. (2020) was used to classify image pixels into five classes: deep water, land, submerged sand, seagrass, and intertidal. Intertidal represents a region of filamentous red algae and microbial films, but not seagrass (Hill et al., 2014), along the land-water interface that would be submerged during high tide and exposed during low tide. Characterization of an intertidal region can be important for time series analyses in which different tidal stages are represented.

A DCNN was chosen as it achieves high accuracies with less computational complexity compared to other approaches (Islam et al., 2018, 2020). The DCNN model was developed with the Keras package in Python 3.5 (Python Core Team, 2015; Chollet, 2015). A DCNN model requires known input classes for training, which are provided through spectral information contained in regions of interest (ROIs). ROIs were generated across the image based on a combination of local knowledge, expected spectral response, and visual confirmation, with three polygons representing each of the five classes (Fig. 4). The same ROIs were used for both WorldView-2 and RapidEye, ensuring differences in classification results between the two sensors were not due to differences in ROIs.

We used an input patch size of 3 × 3 to extract the training samples from the selected ROIs. A sequential model, or linear stack, of six hidden layers was applied. The first layer was a convolutional layer consisting of 32 filters with a kernal size of 1 × 1 followed by a rectified linear activation function. Next a dropout layer randomly set 1% of the outputs from the first layer to zero. The second convolutional layer consisted of 16 filters with a kernal size of 3 × 3, again, followed by a rectified linear activation function. Then, another dropout layer with a dropout rate of 1% was added before flattening the model. Finally, a dense layer with a SoftMax activation function was used to compute pixel-based probabilities for each class (Bishop, 2006). This DCNN model was trained for 500 epochs, where an epoch represents the point at which all training data have been processed one time, using a batch size of 256. More information about the model can be found in Appendix A.

2.6. Aerial estimates of seagrass extent

Seagrass pixels identified in the satellite imagery were compared to a shapefile obtained from the Florida Fish & Wildlife Conservation Commission (FL FWC) representing the distribution of seagrass beds in St. Joseph Bay derived from aerial photography. Aerial photographic images were interpreted by Quantum Spatial, (2010) formerly Photo Science, Inc.) based on October 2010 natural color aerial photography acquired by the Florida Department of Transportation (Statistics Canada, 2008; Great Britain, 2009). Images were taken at high tide and only visible features were mapped.

FL FWC classified features into tidal flats, bays & estuaries, continuous seagrass, patchy seagrass, and unclassified from the aerial imagery. Descriptions of these classes were not available; thus, the ROIs chosen here were based on those presented in Islam et al. (2020) rather than attempting to match the classes presented in the FL FWC dataset. Additionally, the distinction between patchy and continuous seagrass is described as a perceived texture difference rather than percent cover; therefore, continuous and patchy seagrass were combined into a single seagrass class. FL FWC warns that the accuracy of this dataset was not verified. Typically, state programs strive to achieve a 10% mark for ground verification. Thus, this analysis assesses agreement among three remote sensing datasets, none of which have been verified through on the ground measurements: two datasets using a deep learning classifier on satellite imagery, the third using human photointerpretation of aerial imagery.

The aerial shapefile was subset to include just St. Joseph Bay and rasterized to match the spatial resolution of each satellite sensor. Before computing statistics, pixels classified in the satellite imagery as intertidal were discarded in both satellite and aerial estimates to account for differences in tidal heights between each dataset. An agreement matrix was developed for seagrass presence and seagrass absence, and agreement between aerial imagery and each satellite platform was assessed according to Congalton (1991) via the *carot* package in R (Kuhn et al., 2020).

Agreement matrices compare, on a class-by-class basis, the relationship between reference data (in this case, aerial imagery) and the corresponding classification results from analysis of the satellite images. Errors of ommission occur when a feature is incorrectly left out of the category being evaluated, i.e., a false negative. Errors of commission occur when a feature is incorrectly included in the category being evaluated, i.e., a false positive. Overall accuracy was computed as the number of pixels labeled as the same class in aerial and satellite imagery normalized to the total number of pixels in the scene.

The Kappa coefficient indicates how well the resulting classification performs compared to a random classification (Cohen, 1960; Goodman and Kruskal, 1954). It is represented as a ratio between −1 and 1 with higher Kappa coefficients indicating higher agreement between

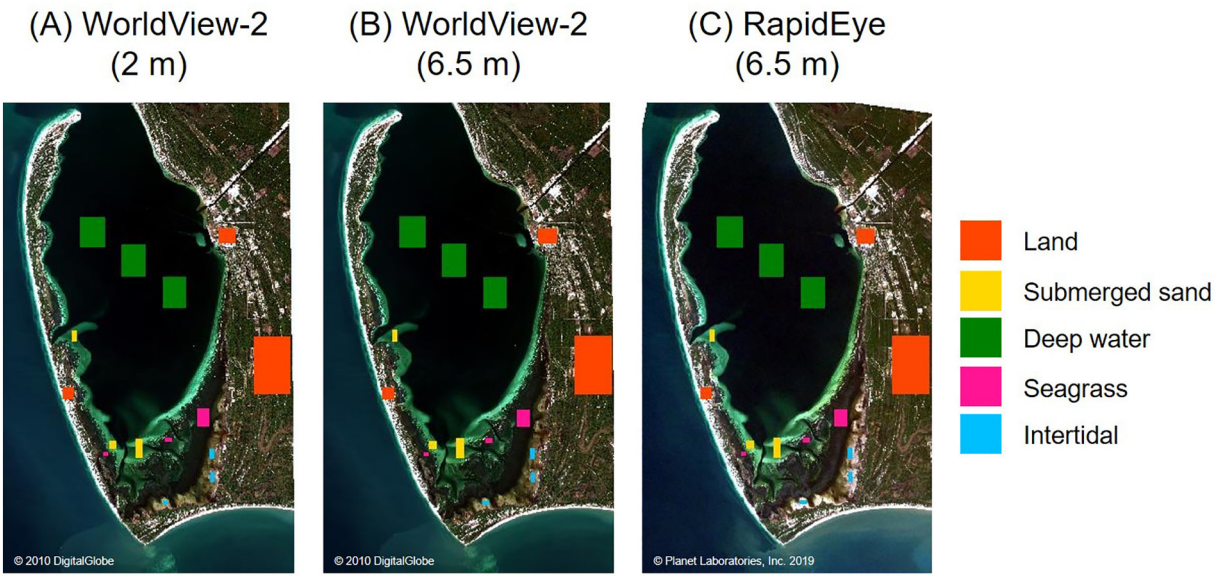

Fig. 4. Regions of interest (ROIs) for each of the five classes overlaid on (A) native resolution WorldView-2 imagery, (B) downsampled WorldView-2 imagery, and (C) native resolution RapidEye imagery. ROIs were generated based on a combination of local knowledge provided by coauthors Zimmerman and Hill, expected spectral response, and visual confirmation.

classification approaches. Difference maps were generated using the ArcGIS Image Analysis window to visualize areas in which satellite and aerial classifications differed (ESRI, 2016).

3. Results and discussion

3.1. Atmospheric correction

3.1.1. Optimizing DOS parameters

Adjustment of the red edge distribution percentage had a smaller effect on reflectance values than adjustment of the Rayleigh exponent, as shown for a single field location characterized by submerged sand (Fig. 5). In WorldView-2 imagery, adjusting the red edge distribution percentage did little to change the resulting reflectance values, meaning distribution percentages ranging between 5% and 50% produced similar results. This same range of red edge distribution values applied to RapidEye data yielded a larger range of reflectance values, but the spectral shapes of corrected imagery still mimicked that of field observations.

When iterating the Rayleigh exponent, similarity in spectral shape between corrected satellite-derived reflectance and field reflectance was sought. A Rayleigh exponent of 4.75 provided necessary correction in shorter wavelengths while preserving non-negative values (i.e., avoiding over-correction) across the entire spectrum. In WorldView-2 imagery, a Rayleigh exponent of 6.5 was still incapable of producing reasonable reflectance values in the coastal blue and blue spectral bands. In RapidEye imagery, higher Rayleigh exponents resulted in negative reflectance values in shorter wavelengths.

Even with more stringent atmospheric correction through an increased red edge distribution and Rayleigh exponent, comparatively high reflectance values in the coastal blue and blue spectral bands of WorldView-2 could not be lowered sufficiently to match the in situ measurements of R_{rs}. Although the present study only examined one WorldView-2 scene, manual inspection of multiple WorldView-2 scenes indicates this offset is common across this platform, suggesting that radiometrically corrected values in the coastal blue and blue spectral bands are much higher than they should be over these relatively dark water targets.

A primary component of standardizing data processing for commercial imagery was the development of an atmospheric correction approach that did not require user-supplied values on a scene-by-scene basis, which is a known limitation of DOS. In addition to creating spectral separability necessary for classification, atmospheric correction is considered a requirement for change detection applications (Coppin et al., 2004), such as assessing temporal changes in seagrass extent. This effort begins to assess the applicability of a standardized processing regime, and it is encouraging that both images settled on the same parameters.

St. Joseph Bay is characterized by a large central basin of optically deep water, which explains similar reflectance values while iterating through various red edge distributions. In coastal areas characterized by waters with a more diverse distribution of optical properties, the percentage of the red edge reflectance values used to characterize dark water will certainly have a larger influence on the resulting spectra. In such environments, using a higher percentage of the red edge reflectance values will likely result in over-correction, yielding negative reflectance values in shorter wavelengths. Therefore, it is likely that a conservative threshold of 5% will allow proper atmospheric characterization even in scenes with a small area of optically deep water, but future efforts should focus on spatial and temporal applicability.

Preserving the minimum red edge anchor value for all tiles within the scene further increases confidence that the red edge anchor is representative of deep water pixels, increasing the likelihood that the atmospheric correction is an accurate representation of atmospheric contamination. If a minimum of one tile within the scene is characterized by optically deep water in at least 5% of the surface area, this DOS approach should yield a sufficient atmospheric correction. A division factor of two was used to capture atmospheric contamination while maintaining true water-leaving radiance in the red edge spectral band.

Relatively high Rayleigh exponents have been found to be more appropriate in clear conditions, whereas relatively low Rayleigh exponents are more appropriate in hazy conditions (Curcio, 1961; Slater et al., 1983; Chavez, 1988). This is because in clear conditions, shorter wavelengths are scattered more relative to longer wavelengths while in hazy conditions, scattering across all wavelengths is more equal. Thus, the focus of DOS in clearer conditions is to decrease reflectance in shorter wavelengths, thereby adjusting the spectral shape. In hazier conditions, the focus of DOS is to decrease reflectance across the entire spectrum while, more or less, preserving the spectral shape. Both scenes

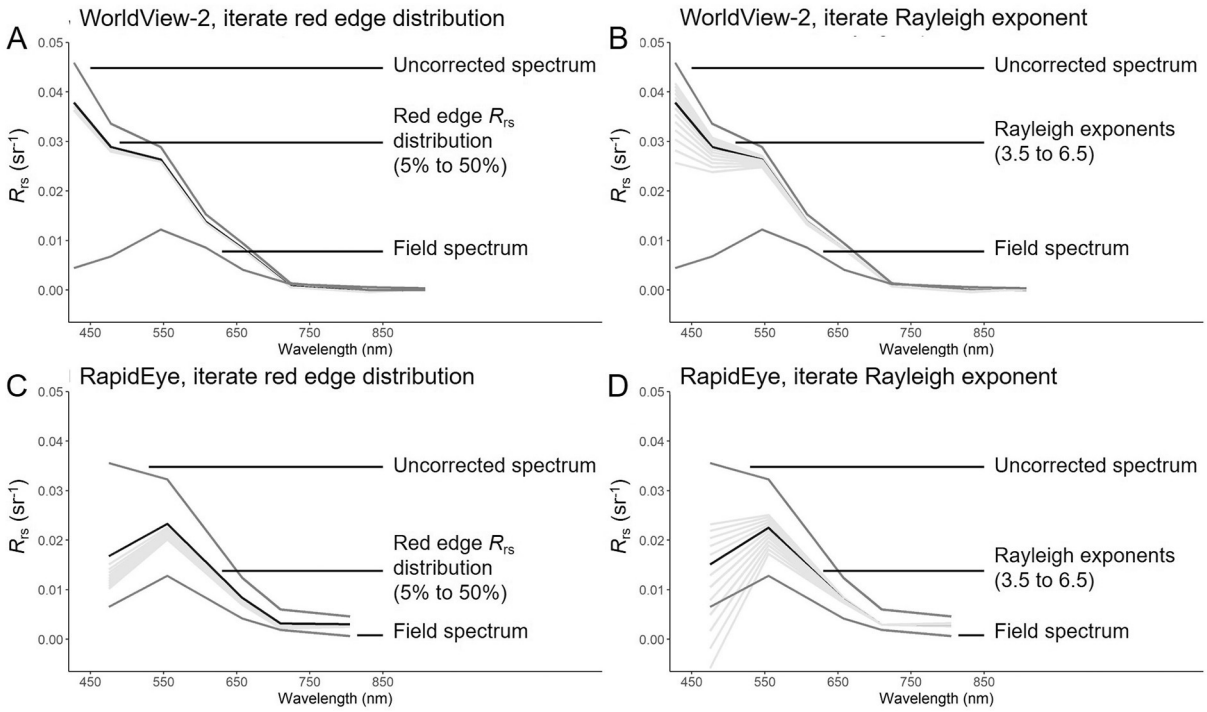

Fig. 5. Example of the effect of red edge distribution percentage and Rayleigh exponent on the atmospherically corrected R_{rs} spectrum for a single pixel representing optically shallow sand. (A) WorldView-2 uncorrected spectra and field spectra compared against a DOS regime considering red edge distribution percentages between 5% and 50% in increments of 5%, while maintaining a constant Rayleigh exponent of 4.75. (B) WorldView-2 uncorrected spectra and field spectra compared against a DOS regime considering Rayleigh exponents between 3.5 and 6.5 in increments of 0.25, while maintaining a constant red edge distribution percentage of 5%. (C) Same as (A), but for RapidEye. (D) Same as (B), but for RapidEye. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

considered in this study represent visibly clear conditions, therefore a high Rayleigh exponent – as opposed to the idealized value of 4 – to represent the relative scattering model for this DOS approach is reasonable. Additionally, given that clear, cloud-free scenes are desired when extracting information at the surface, a Rayleigh exponent of 4.75 will likely be appropriate across a range of satellite scenes.

3.1.2. Atmospheric correction in the literature

Other automated atmospheric correction regimes have been applied to satellite remote sensing imagery, including ENVI's Fast Line-of-Sight Atmospheric Analysis of Hypercubes (FLAASH) model (Kovacs et al., 2018; Tamondong et al., 2013; Roelfsema et al., 2014; Traganos and Reinartz, 2018; Pu et al., 2014). However, FLAASH has been found to provide poorer results than those derived with DOS (Collin and Hench, 2012; Wicaksono and Hafizt, 2018).

DOS has successfully been applied to satellite imagery from a variety of platforms for seagrass detection (Hossain et al., 2015; Fauzan et al., 2017; Thalib et al., 2018; Wicaksono and Lazuardi, 2018). Despite its employment in commercial satellite processing, an investigation into its unbiased optimization has not been performed until now. Many studies apply atmospheric correction to a single scene, but the approach presented here to identify the darkest pixels in a scene can reduce bias requiring little, if any, manual intervention to the current workflow. The presented workflow could also be modified for sensors or scenes containing NIR values with better SNR. Employing NIR channels may result in an improved atmospheric correction given no assumptions need to be made regarding the actual versus perceived atmospheric contributions to the NIR reflectance values.

3.2. Deviation and bias in satellite-derived reflectance

RapidEye overpredicted reflectance values compared to field measurements for all spectral bands, but only slightly, while WorldView-2 overpredicted in shorter wavelengths (coastal blue through red) and underpredicted in longer wavelengths (Fig. 6, Supplementary Table S2). MAD and bias associated with the coastal blue and blue spectral bands of WorldView-2 was nearly 0.03 and 0.025, respectively, while results for the remaining bands fell below 0.015.

Higher reflectance values in RapidEye imagery compared to field observations is consistent with previous studies in terrestrial ecosystems (Machwitz et al., 2014; Chander et al., 2013), and offset between WorldView-2 data and field observations is also supported (Latif et al., 2012). However, previous studies have not assessed radiometric performance over lower-reflectance aquatic environments. Difficulties in the radiometric correction are likely the reason for high reflectance values in shorter wavelengths of WorldView-2 imagery. Updated calibration factors presented in Kuester (2017) likely do not capture sensor drift, meaning updated values do not represent accurate correction factors for the 2010 image presented here. Additionally, these values should be revisited annually, but more recent correction values were not available.

While MAD and bias provide a comparison of sensor performance, a primary objective of atmospheric correction is to create realistic and separable spectral shapes for each class. Therefore, while the magnitude of MAD is a valid measure of sensor performance, a consistent MAD is more relevant for classification. Deviation associated with RapidEye was more consistent over its spectral bands than WorldView-2, suggesting that, although there was an offset, the spectral shapes of RapidEye data were more similar to those of field observations.

The temporal offset between field measurements and satellite overpass ranged from 5 to 12 days for WorldView-2 and 2 to 9 days for RapidEye. Deviation and bias for observations collected on 2 Nov 2010 ($n = 2$) were nearly double those of observations collected on 6 Nov 2010 ($n = 11$) or 9 Nov 2010 ($n = 11$; Supplementary Table S3). Field observations taken on 9 Nov 2010 and early observations taken on 6 Nov 2010 coincided with the tidal stage at which satellite overpass

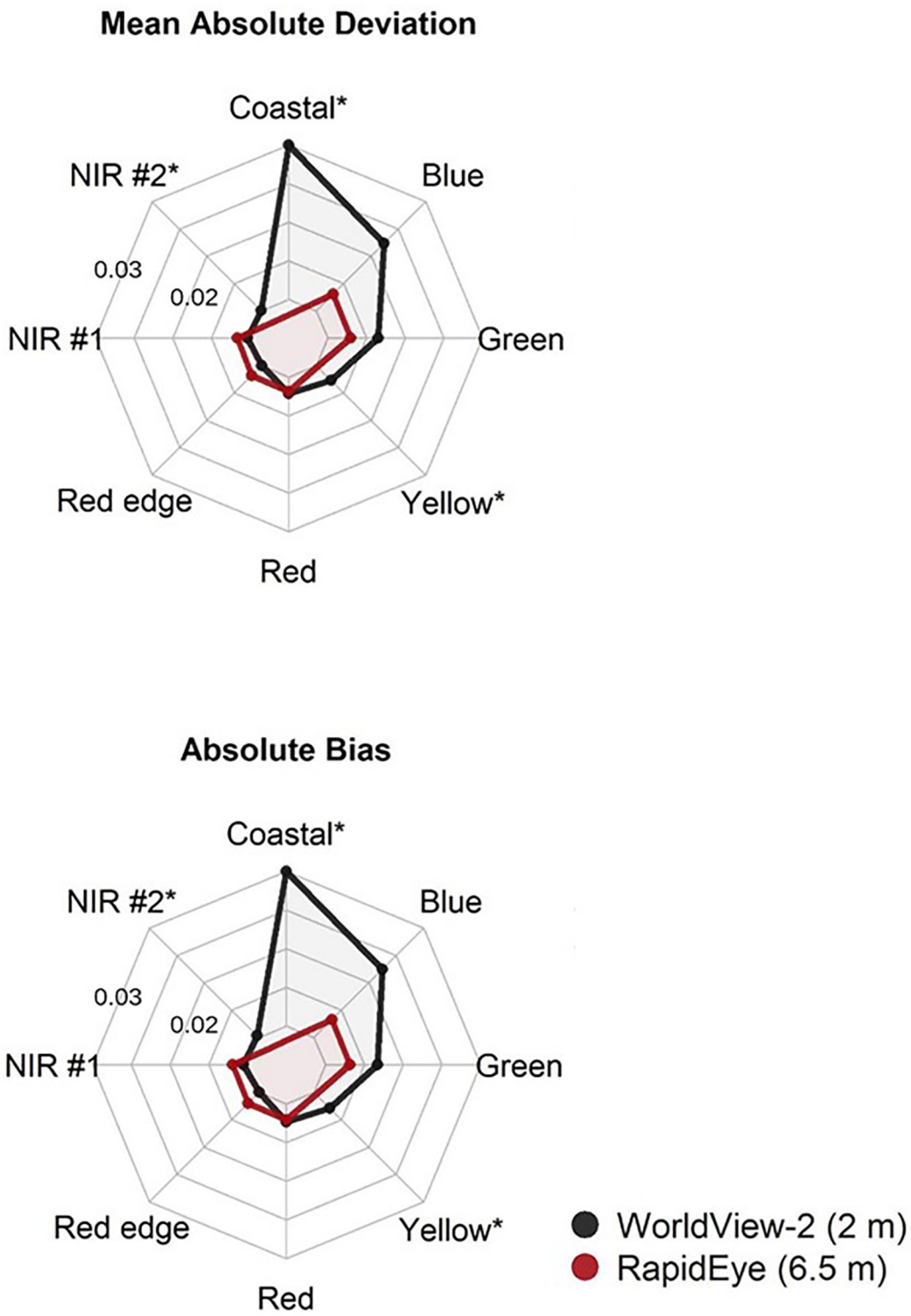

Fig. 6. Radar plots indicating the average mean absolute deviation and mean absolute bias across 24 field sites throughout St. Joseph Bay. Asterisks indicate the wavelengths at which RapidEye does not collect data; therefore, RapidEye statistics were not computed for these bands.

occurred for both WorldView-2 and RapidEye (Fig. 7). Tide predictions were obtained from the NOAA Tides & Currents database (https://tidesandcurrents.noaa.gov/). The tidal heights during field data collected on 2 Nov 2010 and later observations taken on 6 Nov 2010, however, was much lower.

3.3. Agreement between satellite and aerial classifications

3.3.1. Regions of interest

Reflectance within ROIs varied between WorldView-2 and RapidEye (Fig. 8). The intertidal class, for example, was characterized by much higher spectral signatures for RapidEye compared to WorldView-2, nearly resembling a land signature. Because of the tidal stage (Fig. 7), deeper water at the time of image acquisition produced lower values in

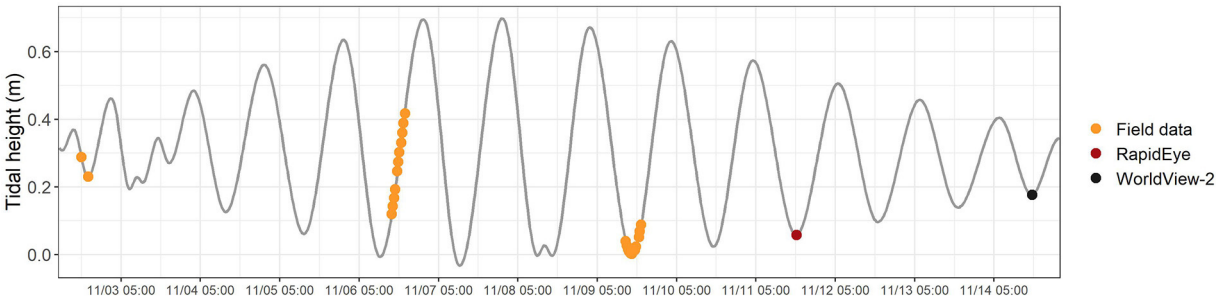

Fig. 7. Predicted tidal height at Port St. Joe, FL obtained from the NOAA Tides & Currents database (https://tidesandcurrents.noaa.gov/). Acquisition times of field data are shown in orange, overpass time of RapidEye is shown in red, and overpass time of WorldView-2 is shown in black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the WorldView-2 image relative to RapidEye.

Reflectance values within RapidEye ROIs also demonstrated increased variability compared to WorldView-2, as shown through a larger spread between the 5*th* and 95*th* percentiles across nearly all classes. Additionally, chronically high reflectance values associated with coastal blue and blue spectral bands of WorldView-2 imagery was evident, particularly in the submerged sand spectra; RapidEye and WorldView-2 had high agreement in the green through NIR bands, but diverged in the shorter wavelengths as RapidEye followed a more expected spectral shape, decreasing in reflectance values, and WorldView-2 increased in reflectance values.

3.3.2. Classification results

Classification results were visually similar across all three datasets throughout much of St. Joseph Bay. However there were clear differences for each sensor compared to aerial photointerpretation (Fig. 9, Supplementary Table S4). Notably, the DCNN classified large areas of intertidal for both satellite platforms that were manually classified as seagrass from the aerial imagery. Although tidal height differed between aerial imagery acquisition and satellite overpass, Hill et al. (2014) reported much of the intertidal region to be colonized by filamentous red algae and microbial films, not seagrass. DCNN results also indicated higher coverage area of submerged sand, much of which was designated as deep water in the photointerpreted aerial imagery. Deeper water at the time of the aerial image acquisition likely explains the under-reporting of optically shallow water consisting of unvegetated sand.

Quantitative assessment indicated that both satellite platforms were well-suited for seagrass detection, with overall accuracy above 96% (Tables 2, Supplementary Table S5). The Kappa coefficient (around 0.87) was also similar among the three images. Kappa coefficients between 0.80 and 0.90 are considered strong with 64–81% of data being reliable (McHugh, 2012). Thus, the higher disagreement between WorldView-2 and field-measured reflectance values did not appear to affect the performance of the image classification, although poor agreement could be problematic if using the imagery for pure radiometric values or atmospherically corrected reflectance values. For example, derived products such as density estimation require low disagreement (Hill et al., 2014).

Visibly identifiable seagrass pixels were misclassified by both WorldView-2 and RapidEye as sand along the fringes of seagrass before transitioning to submerged sand (Fig. 10). RapidEye captured these areas with less accuracy than WorldView-2 as shown by large areas of orange along the the transition zone in the western portion of St. Joseph Bay. Poor capture in this area is likely due to the patchy nature of these regions, which can be problematic for classification in both satellite imagery (Knudby and Nordlund, 2011; Phinn et al., 2008b; Green et al., 1996; Baumstark et al., 2016; Pu et al., 2014; Pu and Bell, 2013) and photointerpretation (McKenzie et al., 2001; Meehan et al., 2005).

Differences between satellite and aerial classifications along the transition from submerged sand to deep water were apparent in native resolution WorldView-2 imagery. A faint magenta line is visible along this transition zone (Fig. 10), likely indicative of misclassification by the DCNN algorithm. Coarser spatial resolution could explain the absence of this artifact in both downsampled WorldView-2 imagery and RapidEye imagery.

WorldView-2 appears to overclassify seagrass in the southern portion of the bay compared to both the aerial photointerpretation and the RapidEye classification. Many of the inlets FL FWC labeled as either submerged sand or deep water are classified as seagrass in both the native resolution and resampled WorldView-2 imagery. Additionally, across all platforms, it appears that many smaller patches of submerged sand embedded in the seagrass beds were misclassified as seagrass. These misclassifications are likely because the spectra appear dark green to the satellite sensor since they were in relatively deep water. A spectral linear unmixing approach could offer insight into subpixel composition, allowing for improved habitat characterization in patchy environments.

The area surrounding the Gulf County Canal creates an artifact in the DCNN classification. All three scenes misidentify seagrass in this area of high CDOM discharge from the canal due to similar spectral shapes. In a spectrally similar environment, image classification approaches can fail to discern different cover types accurately (Knudby and Nordlund, 2011). The inclusion of a CDOM class could improve the classification in the future.

Agreement matrices could not be generated to compare satellite classifications to aerial photointerpretation across all classes because class labels differed between the two datasets. Instead, an agreement matrix was generated to compare classification results for deep water, land, submerged sand, seagrass, and intertidal between resampled WorldView-2 imagery with native resolution RapidEye imagery (Supplementary Table S6).

Previous studies have found classification improvements through further image processing, particularly the application of a water column correction (Pu et al., 2014). However, in our system this correction was unnecessary to achieve sufficient spectral separability between seagrass and other classes. Additionally, none of the images considered were affected by sunglint which could present difficulties in both the proposed processing regime and the classification algorithm. Unmixing-based denoising has also improved accuracy assessments of remotely sensed seagrass (Traganos et al., 2017). Results presented here report merely on the presence or absence of seagrass. Satellite imagery has been used to report on percent cover of seagrass using a five class classification scheme of submerged aquatic vegetation (Pu and Bell, 2013; Pu et al., 2012; Roelfsema et al., 2009). Such an approach could improve classification along the transition from seagrass to another class, but ground measurements are needed to properly assess agreement.

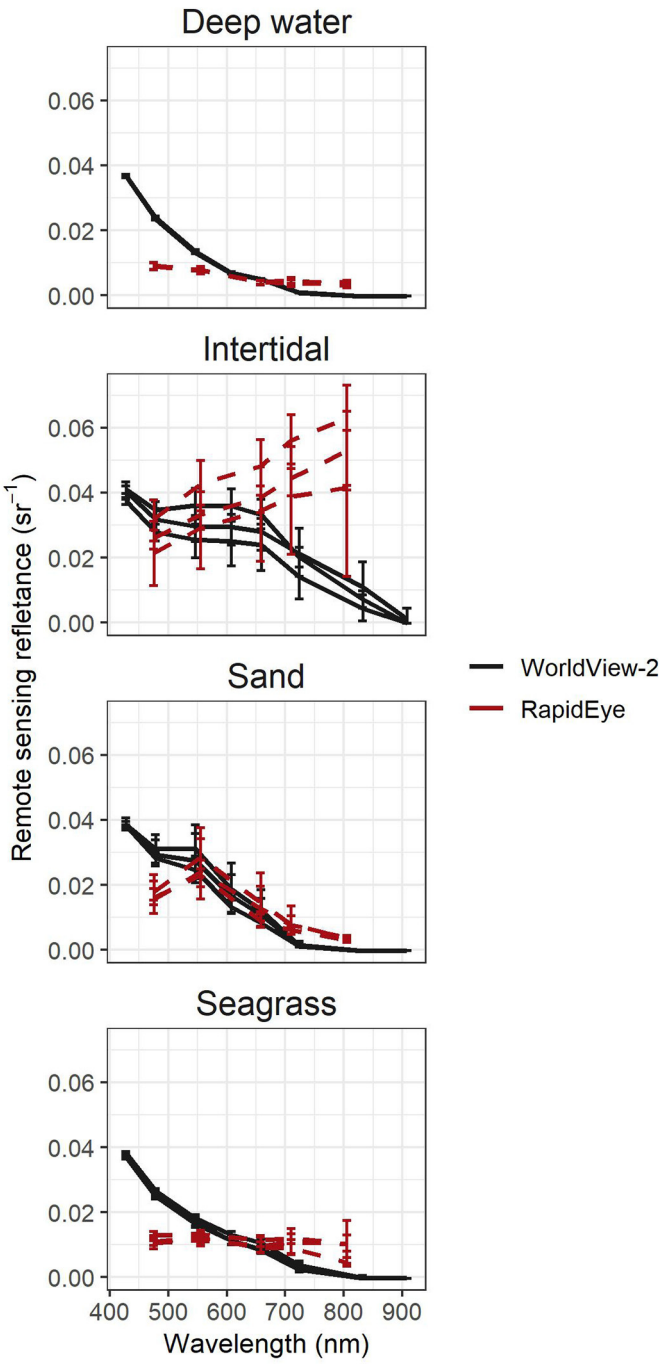

Fig. 8. The 5*th*, 50*th*, and 95*th* percentile of the R_{rs} distribution for each region of interest (ROI) and sensor. Three ROIs were generated for land as well, but their spectra were excluded from this figure.

3.3.3. Applicability across other systems

The processing protocol presented here and classification algorithm introduced in Islam et al. (2020) were used in three additional scenarios to test applicability across other systems (Fig. 11). Because of the decomissioning of RapidEye in December 2019, only WorldView scenes were considered here, including imagery from both WorldView-2 and WorldView-3. Scenes at St. Joseph Bay, Tampa Bay, and St. George Sound were considered for the years 2010, 2016, and 2014, respectively. Tampa Bay and St. George Sound both represent similar systems to St. Joseph Bay. Tampa Bay is located along Florida's Gulf coast and St. George Sound is located along the Florida panhandle just east of St. Joseph Bay.

First, the WorldView-2 scene presented earlier in the manuscript was considered, but using only the blue, green, red, and NIR #1 bands to mirror the spectral characteristics of other commercial platforms. Results were compared to photointerpreted aerial imagery and continued to demonstrate strong agreement (Supplementary Table S7). Next, a WorldView-3 scene collected over Tampa Bay on 1 Mar 2016 was processed and classified using the methods presented here. Visual inspection of the results are promising, with the DCNN algorithm able to identify the patchy nature of seagrass in this system. A WorldView-2 scene collected over St. George Sound on 27 Apr 2014 also demonstrated agreement through visual inspection.

Results at Tampa Bay and St. George Sound were not quantitatively assessed for agreement with seagrass coverage maps for two reasons. First, the DCNN algorithm has already been demonstrated across sites (Islam et al., 2020). The focus of this study was instead to develop a reproducible processing regime for commercial satellite imagery, which can be demonstrated through successful creation of separable spectra required for an accurate classification. Second, the quality of photointerpretation results can vary drastically across regions. Visual comparison of true-color imagery against seagrass shapefiles at Tampa Bay, for example, revealed an important limitation of human drawn maps: areas of patchy seagrass are often aggregated into a single polygon, whereas a pixel-based approach such as a DCNN can provide more detailed estimates of seagrass coverage. Thus, using more generalized seagrass shapefiles can artificially reduce the agreement of the satellite classification as neither dataset conveys a validated representation of true seagrass coverage.

Although archives of WorldView-2 and RapidEye imagery exist, Schaeffer et al. (2013) found that mission continuity is critical for end-users. Several current satellite missions, including Airbus Defense and Space's SPOT-6 and -7 constellation and DigitalGlobe's GeoEye-1 satellite, are potentially well-suited for coastal seagrass detection, but each of these platforms offer only 4 multispectral bands: blue, green, red, and NIR. Results presented here demonstrate promise that the proposed methods are reproducible and potentially could move toward sensor independence, but additional testing across a larger range of images collected across space and time is needed. Additionally, more detailed field data would be required to quantitatively validate the performance of the DCNN at other locations.

3.3.4. The use of photointerpretation as reference data

Many studies assess the accuracy of remote sensing classifications by comparing results to photointerpretation, assuming photointerpretation results are correct. However, this assumption rarely holds and can result in a biased evaluation of remote sensing products (Congalton, 1991). Error associated with reference data is seldom known and, thus, impossible to consider when assessing classification accuracy. Moreover, while they did not offer an alternative, Rutchey and Vilchek (1999) cautioned against quantitative comparisons between aerial photointerpretation and satellite classification, noting the two approaches rely on very different mechanisms for classifying ground cover data. This is a gap that must be closed in order to transition from human to automated interpretation of remotely sensed imagery.

McKenzie et al. (2001) noted a particular area of error along the boundary of seagrass beds, which was an area of disagreement between the satellite and aerial imagery. The authors suggested these regions be assigned a quality flag to reflect lower confidence associated with human error. Moreover, findings from Meehan et al. (2005) argue that photointerpretation methods can overestimate seagrass area by amalgamating disjunct seagrass patches into continuous meadows. Edwards and Lowell (1996) and Thierry and Lowell (2001) proposed a 'fuzzy boundary' concept, where transition zones from one class to another were labeled or interpolated to indicate decreased confidence. These suggestions were not adopted in the FL FWC data, although such an approach could help quantify uncertainty in aerial estimates.

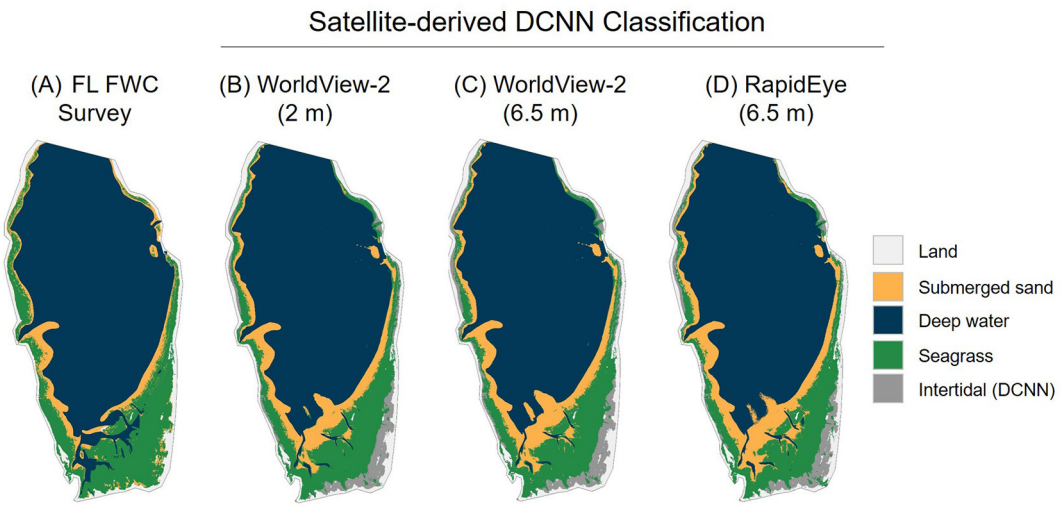

Fig. 9. Classification results from Florida Fish & Wildlife Conservation Commission (FL FWC) and satellite data from native resolution WorldView-2 imagery, downsampled WorldView-2 imagery, and native resolution RapidEye imagery. FL FWC data was based on photointerpreted aerial imagery collected at high tide in October 2010. WorldView-2 and RapidEye imagery were obtained during low tide in November 2010 (WorldView-2 tidal height 0.2 m; RapidEye tidal height 0.06 m).

To our knowledge, no comprehensive study exists to assess the human error rates associated with image classification in aquatic systems. One study assessed the overall accuracy of a single photointerpretation in a wetland ecosystem, finding a 95% agreement with 34 field survey sites (Rutchey and Vilchek, 1999). While not explicitly quantified, Edwards and Lowell (1996) found texture differences and the presence of straight versus curved lines to influence photointerpretation accuracy.

In addition to error inherent in a single photointerpretation event, inconsistencies also exist from one individual to another (Edwards and Lowell, 1996), and for the same individual from one occasion to another (Nantel, 1993). Interpretor error can significantly affect results when comparing maps for temporal trends (Ward et al., 1997). Artificial intelligence approaches for image classification such as a DCNN are likely to be more consistent over time versus traditional photointerpretation methods.

3.4. Future work

This study aimed to retrieve seagrass extent in coastal waters. However, satellite imagery has also demonstrated utility for estimation of leaf area index (LAI) and biomass (Dierssen et al., 2003; Hill et al., 2014), which could aid in addressing uncertainty regarding global estimates of seagrass carbon storage. Moreover, the image processing and classification regime presented here could be used to monitor additional water quality parameters, such as chlorophyll. Future work testing this approach for other applications would require extensive field data for validation.

Ongoing work is focused on testing whether ROIs defined for each class for a single area of interest can be reused in other scenes from that same area of interest, or if ROIs defined for each class for a single area of interest can be reused in other areas of interest. Islam et al. (2020) suggests promise in this pursuit by successfully defining a generalizable DCNN model for seagrass detection from one location to another. Excellent performance was achieved by utilizing labeled samples from new environments to adapt a previously trained classifier. Additionally, extensive evaluation of which machine learning algorithms are best for seagrass detection is planned for future efforts.

Applicability of the proposed processing regime to 4-band sensors was shown here, but the incorporation of additional sensors could improve temporal resolution as compared to commercial platforms. Future work should focus on expanding the presented methods to openly available imagery from sensors such as the Landsat series and Sentinel-2. The Landsat series offers a longer imagery archive than either of the platforms presented here, and Sentinel-2 offers consistent temporal coverage with a revisit frequency of 2 to 3 days.

4. Conclusions

A reproducible workflow was developed to process commercial satellite imagery from WorldView-2 and RapidEye. This workflow transforms basic, level 1B imagery into an atmospherically corrected product. Corrected imagery were compared to field observations of R_{rs}. ENVI/IDL source code and associated documentation accompanying this workflow can be found at doi:10.23719/1518572. WorldView-2 imagery indicated higher disagreement and a positive bias in shorter wavelengths, particularly the coastal blue and blue spectral bands. Longer wavelengths had lower disagreement and a slight negative bias. RapidEye imagery showed lower overall disagreement, and a slight positive bias compared to field data. Despite spectral and spatial

Table 2
Agreement between machine learning classification of satellite imagery and photointerpretation of aerial imagery for seagrass presence and absence.

	Errors of ommission		Errors of commission		Overall accuracy	Kappa
	Present	Absent	Present	Absent		
WorldView-2 (2 m)	6%	3%	14%	1%	97%	0.88
WorldView-2 (6.5 m)	9%	2%	12%	2%	97%	0.88
RapidEye (6.5 m)	14%	2%	9%	3%	96%	0.86

Agreement was quantified through errors of ommission, errors of comission, overall accuracy, and the Kappa coefficient. Areas classified as intertidal by the DCNN were removed before computing agreement statistics to correct for differences in tidal height at time of data collection.

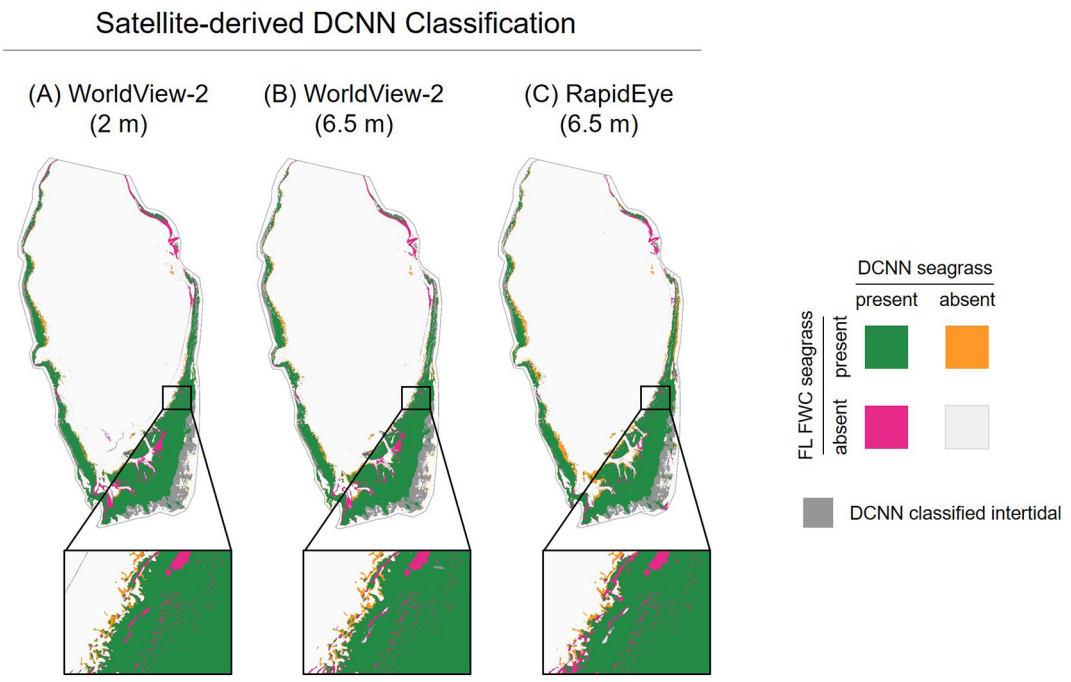

Fig. 10. Difference maps illustrating areas classified through satellite imagery as seagrass present but classified through aerial imagery as seagrass absent (shaded in magenta) and areas classified through satellite imagery as seagrass absent but classified through aerial imagery as seagrass present (shaded in orange). Insets (black box) illustrate a portion of the classification difference to demonstrate performance differences along the transition from seagrass to another class. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

differences, image classification via a DCNN demonstrated nearly 97% agreement for both satellite platforms compared to aerial photo-interpretation, although neither the satellite classification nor the aerial photointerpretation were validated through ground measurements. Methods were tested at two additional Florida sites, St. George Sound and Tampa Bay, and results demonstrated visual agreement with the imagery, suggesting the presented methods can be transferable to other locations.

Declaration of Competing Interest

The authors declare that they have no known competing financial

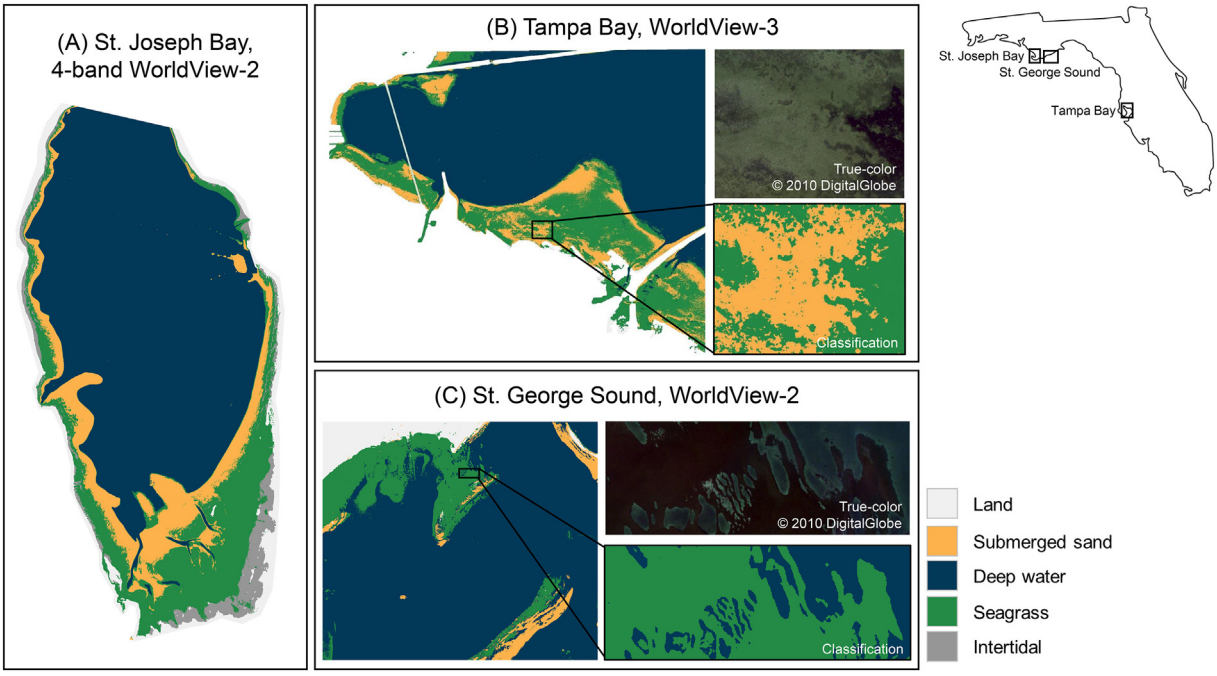

Fig. 11. Classification results for (A) the same WorldView-2 scene used above, but using only the blue, green, red, and NIR #1 spectral bands, (B) a WorldView-3 scene from Tampa Bay, FL, USA from 1 Mar 2016 including a subset of the classification compared to the true-color imagery, and (C) a WorldView-2 scene from St. George Sound, FL, USA from 27 Apr 2014 including a subset of the classification compared to the true-color imagery. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank M. C. Lebrasse, D. Graybill, W. Salls, J. Darling, and P. Colarusso for their contributions in both method development and manuscript review. This work was supported by the NASA Ocean Biology and Biogeochemistry Program NRA #NNH16ZDA001N and augmented by the NASA Commercial Data Buy. Work was also supported by the U.S. EPA, and Oak Ridge Institute for Science and Technology (ORISE). Digital Globe imagery was made available through the U.S. Government NextView license. Planet's RapidEye imagery was made available through the NASA Commercial Data Buy. This article has been reviewed by the Center for Environmental Measurement and Modeling and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use by the U.S. Government. The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA. Lastly, we thank three anonymous reviewers for their assistance in improving the structure and communication of the study.

A. An overview of deep convolutional neural networks (DCNN)

Deep convolutional neural networks (DCNN) have been successfully employed for seagrass detection in multispectral imagery (Islam et al., 2020). A DCNN is a deep learning approach which achieved state-of-the-art classification performances in many image recognition tasks (Krizhevsky et al., 2012). A DCNN typically consists of two types of layers: a convolutional layer and a fully connected layer. Convolutional layers use small kernals to filter the input image and extract useful features for classification. A fully connected layer is also referred to as a dense layer. Dense layers mimic the functionality of the traditional neural network for classification.

A DCNN model can have one or more convolutional layers depending on the complexity of the tasks. These layers can be thought of as feature identifiers are are adept at extracting hierarchical features for classification. Lower layers extract low-level characteristics of the image such as edges, curves, and colors. Higher layers compute more complex features as different combinations of low-level features to achieve robust image classification. The final step before the classification layer is to flatten the model. Flattening transforms the features into a vector that can be fed into the dense layer for classification. The dense layer takes the vector and computes probabilities for the input image. Additionally, dropout layers are usually dispersed throughout the network, helping control overfitting. During training, if the output probabilities are incorrect, the errors are then backpropagated back to the model to adjust the model parameters. This process allows the model to slowly approach the optimum point where the model performs best.

In our experiments, instead of training the DCNN model on the full image, we used small subsets of the image to acheive pixel-wise classification. These subsets of the image are known as patches. We assigned a class label to each patch according to the center pixel of the patch. We extracted fifty thousand patches per class from the selected ROIs to train the DCNN model. Once trained, the model scanned across the entire image patch-by-patch to produce a pixel-wise classification of the entire image. We set the size of the patch to 3×3 and set the step size of the scanning process to one so that adjacent patches were overlapped during testing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2020.112036.

References

Baumstark, R., Dixon, B., Carlson, P., Palandro, D., Kolasa, K., 2013. Alternative spatially enhanced integrative techniques for mapping seagrass in Florida's marine ecosystem. Int. J. Remote Sens. 34, 1248–1264. https://doi.org/10.1080/01431161.2012.721941.

Baumstark, R., Duffey, R., Pu, R., 2016. Mapping seagrass and colonized hard bottom in springs coast, Florida using WorldView-2 satellite imagery. Estuar. Coast. Shelf Sci. 181, 83–92. https://doi.org/10.1016/j.ecss.2016.08.019.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer, New York, NY. http://cds.cern.ch/record/998831.

Björk, M., Short, F., Mcleod, E., Beer, S., 2008. Managing Seagrasses for Resilience to Climate Change. Technical Report. IUCN, Gland, Switzerland.

CEC, 2016. North America's Blue Carbon: Assessing Seagrass, Salt Marsh and Mangrove Distribution and Carbon Sinks. Technical Report. Commission for Environmental Cooperation, Montreal, Canada.

Chander, G., Haque, M.O., Sampath, A., Brunn, A., Trosset, G., Hoffmann, D., Roloff, S., Thiele, M., Anderson, C., 2013. Radiometric and geometric assessment of data from the RapidEye constellation of satellites. Int. J. Remote Sens. 34, 5905–5925. https://doi.org/10.1080/01431161.2013.798877.

Chavez, P.S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sens. Environ. 24, 459–479. https://doi.org/10.1016/0034-4257(88)90019-3.

Chollet, F., 2015. Keras. https://github.com/fchollet/keras.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20, 37–46. https://doi.org/10.1177/001316446002000104.

Collin, A., Hench, J.L., 2012. Towards deeper measurements of tropical Reefscape structure using the WorldView-2 Spaceborne sensor. Remote Sens. 4, 1425–1447. https://doi.org/10.3390/rs4051425.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ. 37, 35–46. https://doi.org/10.1016/0034-4257(91)90048-B.

Conmy, R.N., Schaeffer, B.A., Schubauer-Berigan, J., Aukamp, J., Duffy, A., Lehrter, J.C., Greene, R.M., 2017. Characterizing light attenuation within Northwest Florida Estuaries: Implications for RESTORE Act water quality monitoring. Marine Pollution Bulletin 114, 995–1006. http://www.sciencedirect.com/science/article/pii/S0025326X16309493 https://doi.org/10.1016/j.marpolbul.2016.11.030.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., Lambin, E., 2004. Digital change detection methods in ecosystem monitoring: a review. Int. J. Remote Sens. 36, 1565–1596. https://doi.org/10.1080/0143116031000101675.

Curcio, J.A., 1961. Evaluation of atmospheric aerosol particle size distribution from scattering measurements in the visible and infrared. J. Opt. Soc. Am. 51, 548–551. https://doi.org/10.1364/JOSA.51.000548.

Dadon, A., Ben-Dor, E., Beyth, M., Karnieli, A., 2011. Examination of spaceborne imaging spectroscopy data utility for stratigraphic and lithologic mapping. J. Appl. Remote. Sens. 5, 1–15. https://doi.org/10.1117/1.3553234.

Danielson, J., Gesch, D., 2011. Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010). Technical Report. U.S. Geological Survey Open-File Report 2011–1073. http://pubs.usgs.gov/of/2011/10pdf/of2011-1073.pdf.

Dekker, A.G., Brando, V.E., Anstee, J.M., 2005. Retrospective seagrass change detection in a shallow coastal tidal Australian lake. Remote Sens. Environ. 97, 415–433. https://doi.org/10.1016/j.rse.2005.02.017.

Dekker, A., Brando, V., Anstee, J., Fyfe, S., Malthus, T., Karpouzli, E., 2006. Remote sensing of Seagrass ecosystems: Use of Spaceborne and airborne sensors. In: Larkum, A., Orth, R., Duarte, C. (Eds.), Seagrasses: Biology. Ecology and Conservation. Springer Netherlands, Dordrecht, pp. 347–359.

Dierssen, H.M., Zimmerman, R.C., Leathers, R.A., Downes, T.V., Davis, C.O., 2003. Ocean color remote sensing of seagrass and bathymetry in the Bahamas banks by high-resolution airborne imagery. Limnol. Oceanogr. 48, 444–455. https://doi.org/10.4319/lo.2003.48.1_part_2.0444.

Duarte, C.M., 2017. Reviews and syntheses: hidden forests, the role of vegetated coastal habitats in the ocean carbon budget. Biogeosciences 14, 301–310. https://doi.org/10.5194/bg-14-301-2017.

Duarte, C.M., Losada, I.J., Hendriks, I.E., Mazarrasa, I., Marbà, N., 2013. The role of coastal plant communities for climate change mitigation and adaptation. Nat. Clim. Chang. 3, 961–968. https://doi.org/10.1038/nclimate1970.

Edwards, G., Lowell, K., 1996. Modeling uncertainty in photointerpreted boundaries. Photogramm. Eng. Remote. Sens. 62, 377–390.

ESRI, 2016. ArcGIS Desktop: Release 10.5.

Fauzan, M.A., Kumara, I.S.W., Yogyantoro, R., Suwardana, S., Fadhilah, N., Nurmalasari, I., Apriyani, S., Wicaksono, P., 2017. Assessing the capability of sentinel-2A data for mapping Seagrass percent cover in Jerowaru, East Lombok. Indones. J. Geogr. 49, 195–203. https://doi.org/10.22146/ijg.28407.

Fourqurean, J.W., Duarte, C.M., Kennedy, H., Marbà, N., Holmer, M., Mateo, M.A., Apostolaki, E.T., Kendrick, G.A., Krause-Jensen, D., McGlathery, K.J., Serrano, O., 2012. Seagrass ecosystems as a globally significant carbon stock. Nat. Geosci. 5, 505–509. https://doi.org/10.1038/ngeo1477.

Goodman, L.A., Kruskal, W.H., 1954. Measures of association for cross classifications. J. Am. Stat. Assoc. 49, 732–764. https://doi.org/10.1080/01621459.1954.10501231.

Great Britain, 2009. Ordnance Survey. EDINA Digimap Ordnance Survey Service. OS Code- Point with Polygons [Shapefile geospatial data] Aberdeen District. Updated May 2008. http://edina.ac.uk/digimap.

Green, E.P., Mumby, P.J., Edwards, A.J., Clark, C.D., 1996. A review of remote sensing for the assessment and management of tropical coastal resources. Coast. Manag. 24, 1–40. https://doi.org/10.1080/08920759609362279.

Green, E.P., Mumby, P.J., Edwards, A.J., Clark, C.D., 2000. Remote Sensing Handbook for Tropical Coastal Management.

Hill, V.J., Zimmerman, R.C., Bissett, W.P., Dierssen, H., Kohler, D.D.R., 2014. Evaluating light availability, Seagrass biomass, and productivity using hyperspectral airborne

remote sensing in Saint Joseph's bay, Florida. Estuar. Coasts 37, 1467–1489. https://doi.org/10.1007/s12237-013-9764-3.
Hossain, M.S., Bujang, J.S., Zakaria, M.H., Hashim, M., 2015. Application of Landsat images to seagrass areal cover change analysis for Lawas, Terengganu and Kelantan of Malaysia. Cont. Shelf Res. 110, 124–148. https://doi.org/10.1016/j.csr.2015.10.009.
Huang, H., Roy, D., Boschetti, L., Zhang, H., Yan, L., Kumar, S., Gomez-Dans, J., Li, J., 2016. Separability analysis of sentinel-2A multi-spectral instrument (MSI) data for burned area discrimination. Remote Sens. 8, 873. https://doi.org/10.3390/rs8100873.
Islam, K.A., Pérez, D., Hill, V., Schaeffer, B., Zimmerman, R., Li, J., 2018. Seagrass detection in coastal water through deep capsule networks. In: Lai, J.H., Liu, C.L., Chen, X., Zhou, J., Tan, T., Zheng, N., Zha, H. (Eds.), Pattern Recognition and Computer Vision. Springer International Publishing, Cham, pp. 320–331.
Islam, K.A., Hill, V., Schaeffer, B., Zimmerman, R., Li, J., 2020. Semi-supervised adversarial domain adaptation for seagrass detection using multispectral images in coastal areas. Data Sci. Eng. 5 (2), 111–125. https://doi.org/10.1007/s41019-020-00126-0.
Knudby, A., Nordlund, L., 2011. Remote sensing of seagrasses in a patchy multi-species environment. Int. J. Remote Sens. 32, 2227–2244. https://doi.org/10.1080/01431161003692057.
Kovacs, E., Roelfsema, C., Lyons, M., Zhao, S., Phinn, S., 2018. Seagrass habitat mapping: how do Landsat 8 OLI, Sentinel-2, ZY-3A, and Worldview-3 perform? Remote Sens. Lett. 9, 686–695. https://doi.org/10.1080/2150704X.2018.1468101.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. In: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc, pp. 1097–1105.
Kuester, M., 2017. Absolute Radiometric Calibration: 2016v0. Technical Report. DigitalGlobe.
Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., Team, R. Core, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., Hunt, T., 2020. Caret: Classification and Regression Training. https://github.com/topepo/caret/.
Latif, Z.A., Zamri, I., Omar, H., 2012. Determination of tree species using Worldview-2 data. In: 2012 IEEE 8th International Colloquium on Signal Processing and its Applications, pp. 383–387. https://doi.org/10.1109/CSPA.2012.6194754.
Lyons, M., Phinn, S., Roelfsema, C., 2011. Integrating Quickbird multi-spectral satellite and field data: mapping bathymetry, Seagrass cover, Seagrass species and change in Moreton Bay, Australia in 2004 and 2007. Remote Sens. 3, 42–64. https://doi.org/10.3390/rs3010042.
Machwitz, M., Giustarini, L., Bossung, C., Frantz, D., Schlerf, M., Lilienthal, H., Wandera, L., Matgen, P., Hoffmann, L., Udelhoven, T., 2014. Enhanced biomass prediction by assimilating satellite data into a crop growth model. Environ. Model. Softw. 62. https://doi.org/10.1016/j.envsoft.2014.08.010.
Maxar, 2019. System-Ready Imagery. https://www.digitalglobe.com/resources.
McFeeters, S.K., 1996. The use of the normalized difference water index (NDWI) in the delineation of open water features. Int. J. Remote Sens. 17, 1425–1432. https://doi.org/10.1080/01431169608948714.
McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochemia Med. 22, 276–282. https://www.ncbi.nlm.nih.gov/pubmed/23092060 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/.
McKenzie, L.J., Finkbeiner, M.A., Kirkman, H., 2001. Methods for mapping seagrass distribution. In: Short, F.T., Coles, R.G. (Eds.), Global Seagrass Research Methods. Elsevier Science, Amsterdam, pp. 101–121. https://doi.org/10.1016/B978-044450891-1/50006-2. chapter 5.
Meehan, A.J., Williams, R.J., Watford, F.A., 2005. Detecting trends in seagrass abundance using aerial photograph interpretation: problems arising with the evolution of mapping methods. Estuaries 28, 462–472. https://doi.org/10.1007/BF02693927.
Mélin, F., Sclep, G., Jackson, T., Sathyendranath, S., 2016. Uncertainty estimates of remote sensing reflectance derived from comparison of ocean color satellite data sets. Remote Sensing of Environment 177, 107–124. http://www.sciencedirect.com/science/article/pii/S0034425716300426 https://doi.org/10.1016/j.rse.2016.02.014.
Meyer, C.A., Pu, R., 2012. Seagrass resource assessment using remote sensing methods in St. Joseph sound and Clearwater Harbor, Florida, USA. Environ. Monit. Assess. 184, 1131–1143. https://doi.org/10.1007/s10661-011-2028-4.
Mumby, P., Green, E., Edwards, A., Clark, C., 1997. Measurement of seagrass standing crop using satellite and digital airborne remote sensing. Mar. Ecol. Prog. Ser. 159, 51–60.
Nantel, J., 1993. A new and improved digitizing method based on the Thiessen (Voronoi) algorithm. In: Proceedings Sixth Annual Genasys International Users Conference, Fort Collins, Colorado, pp. 12–25.
Pasqualini, V., Pergent-Martini, C., Pergent, G., Agreil, M., Skoufas, G., Sourbes, L., Tsirika, A., 2005. Use of SPOT 5 for mapping seagrasses: an application to Posidonia oceanica. Remote Sens. Environ. 94, 39–45. https://doi.org/10.1016/j.rse.2004.09.010.
Phinn, S., Roelfsema, C., Dekker, A., Brando, V., Anstee, J., 2008a. a. Mapping seagrass species, cover and biomass in shallow waters: an assessment of satellite multi-spectral and airborne hyper-spectral imaging systems in Moreton Bay (Australia). Remote Sens. Environ. 112, 3413–3425. https://doi.org/10.1016/j.rse.2007.09.017.
Phinn, S., Roelfsema, C., Dekker, A., Brando, V., Anstee, J., 2008b. b. Mapping seagrass species, cover and biomass in shallow waters: an assessment of satellite multi-spectral and airborne hyper-spectral imaging systems in Moreton Bay (Australia). Remote Sens. Environ. 112, 3413–3425. https://doi.org/10.1016/j.rse.2007.09.017.
Planet Labs Inc, 2019. Planet Imagery Product Specifications. Technical Report. https://assets.planet.com/docs/Planet_Combined_Imagery_Product_Specs_letter_screen.pdf.
Planet Team, 2017. Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. https://api.planet.com.
Pottier, C., Garçon, V., Larnicol, G., Sudre, J., Schaeffer, P., Le Traon, P.Y., 2006. Merging SeaWiFS and MODIS/Aqua Ocean color data in north and equatorial Atlantic using weighted averaging and objective analysis. IEEE Trans. Geosci. Remote Sens. 44, 3436–3451.
Pu, R., Bell, S., 2013. A protocol for improving mapping and assessing of seagrass abundance along the west central coast of Florida using Landsat TM and EO-1 ALI/Hyperion images. ISPRS J. Photogramm. Remote Sens. 83, 116–129. https://doi.org/10.1016/j.isprsjprs.2013.06.008.
Pu, R., Bell, S., Meyer, C., Baggett, L., Zhao, Y., 2012. Mapping and assessing seagrass along the western coast of Florida using Landsat TM and EO-1 ALI/Hyperion imagery. Estuar. Coast. Shelf Sci. 115, 234–245. https://doi.org/10.1016/j.ecss.2012.09.006.
Pu, R., Bell, S., Meyer, C., 2014. Mapping and assessing seagrass bed changes in Central Florida's west coast using multitemporal Landsat TM imagery. Estuar. Coast. Shelf Sci. 149, 68–79. https://doi.org/10.1016/j.ecss.2014.07.014.
Python Core Team, 2015. Python: A Dynamic, Open Source Programming Language. https://www.python.org/.
Quantum Spatial, 2010. Panhandle_Seagrass_2010. Vector Digital Data. Quantum Spatial (URL: http://quantumspatial.com/).
R Core Team, 2017. R: A Language and Environment for Statistical Computing. http://www.R-project.org/.
Roelfsema, C.M., Phinn, S.R., Udy, N., Maxwell, P., 2009. An integrated field and remote sensing approach for mapping Seagrass cover, Moreton Bay, Australia. J. Spat. Sci. 54, 45–62. https://doi.org/10.1080/14498596.2009.9635166.
Roelfsema, C.M., Lyons, M., Kovacs, E.M., Maxwell, P., Saunders, M.I., Samper-Villarreal, J., Phinn, S.R., 2014. Multi-temporal mapping of seagrass cover, species and biomass: a semi-automated object based image analysis approach. Remote Sens. Environ. 150, 172–187. https://doi.org/10.1016/j.rse.2014.05.001.
Rutchey, K., Vilchek, L., 1999. Air photointerpretation and satellite imagery analysis techniques for mapping cattail coverage in a Northern Everglades impoundment. Am. Soc. Photogramm. Remote Sens. 65, 185–191.
Schaeffer, Blake, Myer, Mark, 2020. Resolvable estuaries for satellite derived water quality within the continental United States. Remote Sens. Lett. 11 (6), 535–544. https://doi.org/10.1080/2150704X.2020.1717013.
Schaeffer, B.A., Schaeffer, K.G., Keith, D., Lunetta, R.S., Conmy, R., Gould, R.W., 2013. Barriers to adopting satellite remote sensing for water quality management. Int. J. Remote Sens. 34, 7534–7544. https://doi.org/10.1080/01431161.2013.823524.
Slater, P.N., Doyle, F.J., Fritz, N.L., Welch, R., 1983. Photographic systems for remote sensing. In: Colwell, R.N. (Ed.), Manual of Remote Sensing, 2 ed. American Society of Photogrammetry, Falls Church, VA, pp. 231–291 Chapter 6.
Statistics Canada, 2008. 2006 Census. Census subdivisions (cartographic boundary file gcsd000b06a_e). In: Indian Reserves in Canada. Created by McMaster University Library Lloyd Reeds Map Collection, using ArcView 3.2, as a subset of the original dataset with csdtype = IRI.
Tamondong, A.M., Blanco, A.C., Fortes, M.D., Nadaoka, K., 2013. Mapping of seagrass and other benthic habitats in Bolinao, Pangasinan using Worldview-2 satellite image. In: 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS, pp. 1579–1582. https://doi.org/10.1109/IGARSS.2013.6723091.
Thalib, M.S., Nurdin, N., Aris, A., 2018. The ability of Lyzenga's algorithm for Seagrass mapping using sentinel-2A imagery on Small Island, Spermonde archipelago, Indonesia. In: IOP Conference Series: Earth and Environmental Science. 165, 012028. https://doi.org/10.1088/1755-1315/165/1/012028.
Thierry, B., Lowell, K., 2001. An uncertainty-based method of photointerpretation. Photogramm. Eng. Remote. Sens. 67, 65–72.
Traganos, D., Reinartz, P., 2017. Mapping Mediterranean seagrasses with Sentinel-2 imagery. Mar. Pollut. Bull. 134, 197–209. https://doi.org/10.1016/j.marpolbul.2017.06.075.
Traganos, D., Reinartz, P., 2018. Interannual change detection of Mediterranean Seagrasses using RapidEye image time series. Front. Plant Sci. 9, 96. https://www.frontiersin.org/article/10.3389/fpls.2018.00096.
Traganos, D., Cerra, D., Reinartz, P., 2017. CubeSat-derived detection of seagrasses using planet imagery following unmixing-based denoising: is small the next big? In: ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-1/W1, pp. 283–287. https://doi.org/10.5194/isprs-archives-XLII-1-W1-283-2017.
Vanhellemont, Q., Ruddick, K., 2014. Turbid wakes associated with offshore wind turbines observed with Landsat 8. Remote Sens. Environ. 145, 105–115.
Vantrepotte, V., Mélin, F., 2011. Inter-annual variations in the SeaWiFS global chlorophyll a concentration (1997–2007). Deep-Sea Res. I Oceanogr. Res. Pap. 58, 429–441.
Ward, D.H., Markon, C.J., Douglas, D.C., 1997. Distribution and stability of eelgrass beds at Izembek lagoon, Alaska. Aquat. Bot. 58, 229–240. https://doi.org/10.1016/S0304-3770(97)00037-5.
Ward, D.H., Morton, A., Tibbitts, T.L., Douglas, D.C., Carrera-González, E., 2003. Long-term change in eelgrass distribution at Bahía san Quintín, Baja California, Mexico, using satellite imagery. Estuaries 26, 1529. https://doi.org/10.1007/BF02803661.
Wicaksono, P., Hafizt, M., 2018. Dark target effectiveness for dark-object subtraction atmospheric correction method on mangrove above-ground carbon stock mapping. IET Image Process. 12, 582–587. https://doi.org/10.1049/iet-ipr.2017.0295.
Wicaksono, P., Lazuardi, W., 2018. Assessment of PlanetScope images for benthic habitat and seagrass species mapping in a complex optically shallow water environment. Int. J. Remote Sens. 39, 5739–5765. https://doi.org/10.1080/01431161.2018.1506951.