Three decades of variability in California's giant kelp forests from the Landsat satellites

Tom W. Bell\textsuperscript{a,b,*}, James G. Allen\textsuperscript{a}, Kyle C. Cavanaugh\textsuperscript{b}, David A. Siegel\textsuperscript{b,c}

\textsuperscript{a} Earth Research Institute, University of California, Santa Barbara, CA 93016, United States of America
\textsuperscript{b} Dept. of Geography, University of California, Los Angeles, CA 90095, United States of America
\textsuperscript{c} Dept. of Geography, University of California, Santa Barbara, CA 93016, United States of America

\section{Introduction}

Long-term, large area monitoring of ecosystems can help explain complex questions in ecology by comparing the dynamics of an organism or suite of organisms to variable environmental conditions, discrete disturbances, or known associations with other organisms (Nouvillon et al., 2001; Wimberly and Reilly, 2007; Meigs et al., 2011). Spaceborne sensors are a powerful tool that can generate these spatiotemporal data and provide repeat global measurements on decadal scales (Wulder et al., 2008; Loveland and Dwyer, 2012). However, since many organisms remain cryptic to observation from space, most studies have focused on large swaths of primary producers (e.g. forest canopy, grasslands, phytoplankton) to address ecosystem processes driven or affected by these species (e.g. beetle outbreaks, wildebeest migration, carbon export to the deep sea; White et al., 2005; Boone et al., 2006; Siegel et al., 2014).

Giant kelp (Macrocystis pyrifera) is a globally distributed macroalga that serves as the foundation to an incredibly diverse and productive ecosystem on subtidal rocky reefs (Dayton, 1985). The kelp is fastened onto the reef using a holdfast and buoyant fronds to the surface, where it forms a thick floating canopy that can be contiguous for many kilometers in extent. The resulting three-dimensional structure supports economically important fish and invertebrate species, many of which are not found on the reef without the presence of giant kelp (Graham, 2004; Miller et al., 2018). The floating biomass of the surface canopy can be quantified with high spatial resolution (≤30 m) multispectral...
sensors such as the Landsat satellites (Cavanaugh et al., 2011). This system is highly dynamic, as entire kelp forests can be dislodged from the reef during a single large storm, while the recruitment of new individuals to form a new forest canopy may occur within a few months (Dayton et al., 1992; Reed et al., 2006). The 16-day repeat cycle of the Landsat sensors should provide at least one seasonal observation of most giant kelp forests worldwide, allowing researchers to study the seasonal to multi-decadal dynamics of this foundation species over broad spatial scales (Bell et al., 2015a).

Giant kelp forests are subject to a variety of environmental processes that contribute to its growth, recruitment, and localized extinction over diverse time and space scales. These include seasonal wave disturbance and coastal upwelling, sea urchin grazing dynamics, and spore production and transport (Harrold and Reed, 1985; Lafferty and Behrens, 2005; Gaylord et al., 2006; Reed et al., 2008; Castorani et al., 2015, 2017). Over multiannual time scales, fluctuations in giant kelp have been related to large-scale, long-period cycles in ocean climate which can alter seawater temperature and nutrient patterns and modify the incidence of wave disturbance from storms (Dayton and Tegner, 1984; Edwards, 2004; Parnell et al., 2016; Bell et al., 2015b; Pfister et al., 2017). These climatic oscillations, such as the El Niño-Southern Oscillation (ENSO), Pacific Decadal Oscillation (PDO), and North Pacific Gyre Oscillation (NPGO), can have massive implications for marine ecosystems (Di Lorenzo et al., 2008). Bomb and burst cycles of zooplankton, predatory and forage fish, marine mammals, and seabirds have all been linked to these large-scale climatic changes (Trillmich and Limberger, 1985; Mantua et al., 1997; Beaugrand et al., 2002; Chavez et al., 2003).

Since these low-frequency climate cycles display decadal-scale periodicity and longer scale coupling patterns, several decades of time series data of the focal species are needed to fully understand their effects or to disentangle the potential effect of anthropogenic forcing due to climate change (Henson et al., 2016; Joh and Di Lorenzo, 2017). Shorter time series (5–10 years) may show strong trends associated with periodic swings in ocean climate conditions while longer time series (> 30 years) may balance this short-scale variability to reveal weaker trends but may provide more valuable evidence for long-term change (Henson et al., 2010). The spatial scale of observation may provide clues to the scales over which these drivers act. Populations with trends over small scales (tens to hundreds of meters) may be driven by a different set of environmental forcings than the same populations over larger scales (tens to hundreds of kilometers; Borcard et al., 2004; Lamy et al., 2018). Therefore, it is useful to observe population dynamics over a range spatiotemporal scales to better understand the effect that ocean climate drivers may have.

The Landsat sensors have provided imagery suitable for the estimation of giant kelp biomass since the launch of the Thematic Mapper sensor (TM: Landsat 4/5) in 1982 and continuing uninterrupted to the current Enhanced Thematic Mapper Plus (ETM+; Landsat 7) and Operational Land Imager (OLI; Landsat 8). However, merging these sensors to create a consistent kelp canopy biomass time series presents several challenges. The spectral response functions have changed between sensors, and the signal-to-noise ratio and radiometric quantization have improved through time (Barsi et al., 2014; Roy et al., 2014). Sensor degradation, instrument noise, and response function biases can lead to subtle but important trends when comparing multiple sensors (Okin and Gu, 2015; Sulla-Menashe et al., 2016). Mechanical failures, such as the Landsat 7 ETM+ scan line corrector failure, led to missing data which must be corrected for to produce a continuous time series (Scaramuzza et al., 2004). Additionally, the consequences of working in a marine system such as the effect of tides and currents increase complexity. To address these challenges and to produce a consistent, multi-decadal time series of giant kelp canopy biomass, we address the following objectives and questions:

1. Develop a consistent kelp canopy biomass time series from the TM, ETM+, and OLI sensors by correcting for spectral differences between sensors, developing a relationship with ground-truth data, and accounting for biases due to environmental variables.
2. Develop a robust procedure for gap-filling missing data due to the ETM+ scan line corrector failure.
3. How does time series length and spatial breadth relate to long-term trends and are changes in kelp biomass evident over longer time scales?

2. Methods

2.1. Study area

The study area encompasses the coast of California, USA (including the offshore Channel Islands) where the giant kelp (Macrocystis pyrifera) is the dominant canopy-forming macroalga, from Año Nuevo Island in the north to the USA/Mexico border. This area is comprised of eight Landsat tiles with a long, uninterrupted time series of cloud-free imagery. This ~1500 km segment of coastline can be primarily divided into three biogeographic regions: the Ensenadian (from the Mexico border to Los Angeles), the Southern Californian (Los Angeles west to Point Conception), and the Montereyan (the central coast of California north of Pt. Conception; Blanchette et al., 2008). Giant kelp in the Montereyan region displays a regular season pattern, where coastal upwelling drives robust growth in the spring and summer and creates a dense surface canopy, which is then removed by large swells originating from North Pacific storms in the fall and winter (Fig. 1a; Bell et al., 2015a). Much of the Southern Californian and Ensenadian region is protected from the large North Pacific swells by Pt. Conception and the offshore Channel Islands, and nutrient availability follows low-frequency fluctuations in oceanographic oscillations such as the strength of the North Pacific Gyre Oscillation, El Niño – Southern Oscillation, and Pacific Decadal Oscillation (Di Lorenzo et al., 2008; Parnell et al., 2016; Cavanaugh et al., 2011; Bell et al., 2015b). These differences in the disturbance and nutrient regimes lead to an intermittent seasonal cycle overlain on an interannual cycle producing high variability in the timing of maximum canopy biomass in southern California (Reed et al., 2011; Bell et al., 2015a).

2.2. Estimating giant kelp canopy from the Landsat sensors

Giant kelp canopy fraction was estimated from multispectral TM (1984–2011), ETM+ (1999–2017), and OLI (2013–2017) satellite imagery using a fully automated processing scheme. Landsat Collection 1 Level-2 reflectance data was downloaded from the United States Geological Survey Earth Explorer website for the areas of interest (earthexplorer.usgs.gov). The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model was used to mask out all pixels above 0 meters elevation (asterweb.jpl.nasa.gov/gdem.asp; Fig. 2.1). In order to mask out beaches and intertidal areas, a 120 m buffer was applied to this mask. A binary classification decision tree was then used to classify each pixel as one of four categories: seawater, cloud, land, and kelp canopy (Matlab function fitctree; Fig. 2.2). The decision tree classifier was trained by clustering pixels from a stacked, masked Landsat image (bands 1–5, 7 for TM/ETM+ and bands 2–7 for OLI) containing variable cloud and kelp canopy conditions using a k-means clustering algorithm (15 clusters; Matlab function kmeans). Each cluster was then manually binned into the four classes described above and used to train the decision tree classifier. In order to account for differences in spectral band widths, separate classifiers were trained for TM/ETM+ and OLI images using respective sensor training images. Once each image was classified, an additional cloud mask was then applied using the quality assessment band included with each Level-2 image (see Supplement A; Figs. S1, S2). After all the images were classified, we filtered errors of commission (free floating kelp paddles, spectral image errors) by removing any pixels
classified as 'kelp canopy' in < 1% of the time series images (see Supplement A; Fig. S3). An additional filter was used to remove spurious intertidal pixels, which may be covered by photosynthetic material such as intertidal algae or surfgrass. Known spatial synchrony patterns of giant kelp (Cavanaugh et al., 2013) were used to identify and eliminate these potentially spurious pixels. All pixels within 500 m of each composite kelp pixel were found and Pearson correlation coefficients were calculated between the focal pixel and the sum of all nearby pixels. Kelp pixels are known to have strong temporal correlations with nearby kelp pixels, while intertidal pixels will only be classified as 'kelp canopy' as a function of tidal exposure. Focal pixels with a mean correlation coefficient of < 0.1 with nearby kelp pixels were excluded from the composite map.

The fraction of kelp canopy covering each Landsat pixel classified as 'kelp canopy' was determined using Multiple Endmember Spectral Mixture Analysis (MESMA), by modeling each pixel as linear combinations of a single kelp endmember, which was constant between images, and one of 30 seawater endmembers, which were unique to each image to account for differences in seawater conditions (e.g. glint, sediment, phytoplankton blooms; Roberts et al., 1998; Cavanaugh et al., 2011). The MESMA process allows endmembers to vary on a per pixel basis by selecting from multiple endmembers for one or more cover types enabling spectral variability of cover types to vary in space and time. The 30 seawater endmembers were selected from consistently non-kelp covered areas within each Landsat scene. The locations where these seawater endmembers were collected did not change among image dates. The pixels of each image were modeled as a two-endmember mixture of kelp canopy and each of the 30 seawater endmembers which were free of cloud contamination. The final model (out of 30) chosen for each pixel was the model that minimized the root mean squared error (RMSE) when fit to the reflectance spectrum of that pixel (maximum RMSE = 0.25). The result of this process was a measure of the relative fraction of each pixel covered by kelp canopy (Fig. 2.3). The MESMA procedure was completed using bands 1–4 for TM/ETM+ and bands 2–5 for OLI.

2.3. Landsat TM, ETM+, and OLI spectral response function effects

The spectral response functions for each band changed slightly between the TM and ETM+ sensors and were significantly narrowed for OLI, especially the near infrared band. Since the kelp fraction estimation methods rely on these bands to model each pixel as linear combinations of kelp canopy and seawater using MESMA, it is essential to understand how these changes affect kelp fraction retrievals. To investigate these changes, we utilized hyperspectral imagery of giant kelp forests acquired as part of the HyspIRI Preparatory Mission using the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS; 18 m spatial resolution; 224 continuous 10 nm bands between 400 and 2500 nm). The AVIRIS imagery was atmospherically corrected using the Atmosphere Removal Algorithm (ATREM) with an additional empirical correction using invariant targets measured in situ (Gao et al., 1993; HyspIRI Preparatory Level 2 Surface Reflectance ftp://popo.jpl.nasa.gov/2013_HyspIRI_Prep_Data). We created simulated multispectral images of kelp forests in the Santa Barbara Channel using images acquired on April 11, 2013. We simulated TM, ETM+, and OLI images by spectrally resampling the hyperspectral imagery to known spectral
response functions for each sensor. The MESMA method was used to estimate the kelp fraction of the pixels for each simulated image. The same kelp endmember was used for these simulated images and 30 seawater endmembers were selected from the same sites for each image. Kelp fractions estimated by MESMA were then compared between these images using linear regressions.

2.4. Comparison with diver determinations

Landsat estimated kelp canopy fraction was compared to several diver estimated kelp variables at two sites in the Santa Barbara Channel as part of the Santa Barbara Coastal Long Term Ecological Research project (SBC LTER) at Mohawk (34° 23.660′ N, 119° 43.800′ W) and Arroyo Quemado (34° 28.127′ N, 120° 07.285′ W) reefs between 2003 and 2017. Canopy biomass, frond density, and plant density were estimated by divers in two permanent 40 × 40 m plots, one at each reef, using five transects (40 × 2 m) per plot. Briefly, divers counted the number and length of subsurface (> 1 m in length) and canopy fronds and used empirical allometric relationships to estimate biomass (Rassweiler et al., 2008). Each plot was overlapped by four Landsat pixels, so the kelp fraction in each of the four pixels was adjusted to its respective proportion of the 40 × 40 m plot. A reduced major axis linear regression was used to compare the mean satellite estimated kelp fractional cover to the diver estimated variables because both estimates contain error. Kelp fraction estimates were compared to diver-based estimates if the survey date was within 5 days of the image acquisition.

Tides and currents are known to affect the proportion of canopy reaching the surface (Britton-Simmons et al., 2008). A higher tide level will pull a portion of the fronds under water as the holdfast is static on the rocky reef below and will decrease the amount of frond that lies horizontal on the surface. Differences in tides and currents between the timing of in-field sampling and satellite overpasses were compared to the residuals of the kelp fraction to canopy biomass relationship. Predicted tides for the region were determined using the t_tide package in Matlab (Pawlowicz et al., 2002) and local currents were determined at half meter intervals from 1.5 m below the surface to a depth of 9 m at each site using acoustic doppler current profilers (RDI Instruments: Work Horse, 600 kHz).

2.5. Effects of tides on patch scale biomass estimates

In order to assess whether tidal height has an effect on patch scale (~100 m and larger) biomass dynamics, Landsat kelp pixel fractions were summed into patches delineated using a modularity analysis (detailed methods in Cavanaugh et al., 2014), across three Landsat tiles encompassing the Santa Barbara Channel and Northern Channel Islands, Southern Channel Islands and the Los Angeles coastline, and the Orange County and San Diego coastlines in California, USA (scene IDs: 042036, 041037, 040037).

We compared TM and ETM+ canopy biomass estimates from imagery collected 8 days apart between July 1999 and May 2003, the period when the ETM+ scan line corrector was functional. While giant kelp is the fastest growing marine autotroph and is subject to removal from discrete disturbance events (Clendenning, 1971; Reed et al., 2008), comparing images 8 days apart should provide for reliable comparisons to assess the importance of tidal biases. The mean patch biomass for all dates was then compared using a reduced major axis linear regression. Patches that were fully or partially obscured by clouds were excluded from analysis for that image date.

2.6. Landsat ETM+ biomass gap-filling algorithm

Enhanced Thematic Mapper Plus images acquired after May 2003 were subject to the scan line corrector failure, which resulted in about 20% of the data in each image being lost as missing data lines. These missing data manifest as alongscan black lines which increase in width towards the edge of the image to a maximum width of 450 m. Since many kelp patches were affected by these lines, we developed a gap-
filling algorithm to fill in missing pixels and estimate patch scale biomass dynamics, which was especially important during the time period when neither TM nor OLI sensor data were available (December 2011 – April 2013). Giant kelp canopy biomass is known to display high, but exponentially decreasing, spatial synchrony over the first several hundred meters in distance (Cavanaugh et al., 2013; Morton et al., 2016). We leveraged this phenomenon to predict canopy biomass in a missing pixel using a combination of the biomass state of nearby pixels and their relationship to the missing pixel through time. All pixels within a 300-meter radius from the missing pixel were identified and the linear relationship of the biomass time series between the missing pixel and each nearby pixel was found using a reduced major axis linear regression. For each significant relationship with a correlation coefficient > 0.8, a biomass estimate for the missing pixel was generated using the regression slope and offset. The mean of these estimates was used as the missing pixel fill value and the standard error was retained as a measure of uncertainty. Missing pixels where > 70% of nearby pixels show zero detected canopy biomass were filled with a value of zero. Missing pixels with no nearby kelp pixels with correlation coefficients > 0.8 were filled using a piecewise cubic interpolation of the missing pixel through time.

To validate this synchrony-based gap filling algorithm we selected six TM images across the study period and masked out pixels using a scan line gap mask from an ETM+ image. The dates used for validation were: November 16, 2000; October 5, 2002; July 14, 2004; November 22, 2005, August 8, 2008; and January 17, 2009. We then filled these missing pixels and compared the predicted biomass to the actual canopy biomass measured by the TM sensor using linear regression analysis. In order to produce a consistent, seasonal time series of canopy biomass, we calculated the mean biomass of each pixel across all Landsat imagery inside each 3-month time period.

2.7. Effect of scale on detecting long-term trends in biomass

In order to examine the effect of temporal and spatial scale on kelp biomass temporal trends, we analyzed the 34-year dataset as sets of progressively sized spatiotemporal series. This exercise represents a hypothetical experiment in which a researcher had started a monitoring site of varying size for any length of time, at any point in the time series. The native Landsat pixel size of 30 m was resampled to six alongshore segment scales (1 km, 2 km, 10 km, 25 km, 50 km, 100 km) in order to simulate a research sampling design of increasing scale. Kelp biomass values were summed through time inside each resampled site. The effect of temporal scale was estimated by subsetting the entire dataset into time series of progressively longer length at each site. Each hypothetical time series was started at every point in the dataset which could accommodate the number of samples in that series length, between 10 and 34 years. Significant temporal trends through time were calculated with generalized least squares regressions in order to account for temporal autocorrelation between the residuals of the model (R package nlme; Pinheiro et al., 2016). Each time series was standardized as a proportion of the maximum biomass value observed and log transformed to meet the assumptions of the model. The mean slope of each significant (p < 0.05) trend was calculated for all simulated time
series of a particular length, at each spatial scale. The same procedure, without standardization and transformation, was completed for low frequency climate oscillations known to have an effect on kelp biomass such as the NPGO (http://www.o3d.org/npgo/), PDO (http://research.jisao.washington.edu/pdo/), and Multivariate ENSO (MEI) indices (https://www.esrl.noaa.gov/psd/enso/mei/). The mean significant trends of kelp canopy biomass and of the ocean climate indices were then compared to each other using simple linear correlations.

3. Results

3.1. Landsat TM, ETM+, and OLI spectral response function effects

Kelp fraction retrievals from Landsat TM and ETM+ data simulated from the hyperspectral image showed a strong, positive significant linear relationship and were nearly equivalent \((y = 1.003x + 0.001; r^2 = 0.999; p < 0.0001; \text{Fig. 3a})\). The relationships between Landsat OLI and TM/ETM+ showed a greater amount of scatter, with OLI underestimating the fractional cover of kelp canopy \((y = 0.806x - 0.013; r^2 = 0.994; p < 0.0001; \text{Fig. 3b, c})\). The relationships between OLI and TM/ETM+ were best approximated by the quadratic function \(y = -0.229x^2 + 1.449x - 0.018 (r^2 = 0.996; p < 0.0001)\). When the OLI fractional kelp retrievals were corrected with this function, the corrected OLI retrievals (OLI_c) showed strong linear relationships to TM and ETM+ estimates \((y = 0.998x + 0.001; r^2 = 0.996; p < 0.0001; \text{Fig. 3d–f})\).

3.2. Comparison with diver determinations

Reasonably strong linear relationships between Landsat estimated kelp fraction and diver estimated canopy biomass density were found for all three sensors across the two SBC LTER sites (Figs. 1b, 4). Significant linear relationships between kelp fraction and frond/plant density were also found (Table 1). Equations fitted for each sensor between kelp fraction and diver estimated canopy biomass density displayed similar linear trends where the slopes were all within the 95% confidence interval of each other (Table 1). Since the slopes were not significantly different from each other we found the relationship between all kelp fraction estimates and diver canopy biomass estimates across all three sensors (Eq. (1)),

\[y = 6.53x + 0.30\]

where \(x\) equals the MESMA estimated kelp fraction and \(y\) equals the giant kelp canopy biomass density in fresh kg m\(^{-2}\). Kelp fraction residuals were examined for any relationship in tide and current speed differences between in-field diver estimates and Landsat observation times. There was no significant relationship between the residuals and current speed at any depth, however there was a significant relationship found between the residuals and the difference in tidal height \((r = -0.21, p < 0.001)\), although this weak relationship does not explain much of the scatter in the relationship.

3.3. Effects of tide on patch scale biomass estimates

Upon examination of the regional patch-scale kelp biomass data, inconsistencies in biomass estimates were apparent between the TM and ETM+ sensors even though the sensors should produce equivalent biomass estimates on a pixel scale. These inconsistencies were not stable across the domain with some Landsat tiles showing consistently higher biomass estimates for one sensor, while an adjacent Landsat tile would show the opposite. These discrepancies were attributed to the 8-day repeat difference between the satellites synching with tidal cycles across the time range of measurement (Fig. 5a). Each Landsat scene showed a biomass difference consistent in direction and magnitude with the tidal effect (Fig. 5b). The role of tide in these biomass differences was confirmed by examining the differences in tidal magnitude between successive Landsat TM and ETM+ passes. Because the effect of tide is also a function of the amount of canopy biomass present at the time of the overpasses, the relative proportional biomass difference of each patch between the two sensors was compared to the tidal state at the time of the Landsat pass relative to the region’s mean tidal height. The mean relationship across all patches across all three Landsat tiles was then used to correct for the differences in tide, adjusting the proportional biomass based on local tidal patterns (Fig. 5c). When we accounted for these differences in tidal state, there is a more consistent relationship between kelp patch biomass estimated from TM and ETM + sensors, which falls close to the 1:1 line, our expectation based on our analysis in Section 3.1. While this tidal adjustment method works well at the patch scale, entire Landsat pixels may lose kelp canopy during high tide events around the edge of the patch. Additionally, this tidal correction increased the error in low biomass patches (Fig. 5c). In order to maximize dataset versatility for spatial analyses across scales, we corrected for tidal effects on a 30 m pixel scale by averaging Landsat biomass estimates across a seasonal scale (3 months).

3.4. Landsat ETM+ biomass gap-filling algorithm

Across the six dates, the gap-filling algorithm was used to fill 80,286 missing kelp pixels. Of that total, 59% were filled using the spatial synchrony method and 41% were filled using temporal interpolation. Pixels filled using spatial synchrony had a coefficient of determination of 0.83 \((p < 0.0001; y = 0.94x + 74)\). Pixels filled using temporal interpolation had a coefficient of determination of 0.53 (Fig. S4a; \(p < 0.0001; y = 0.74x - 6)\). The overall relationship had a coefficient of determination of 0.77 \((p < 0.0001; y = 0.92x + 16)\). The results showed that the pixels filled using the spatial synchrony method were closer to actual canopy biomass than those filled using interpolation between two close dates. Overall, the total gap filling algorithm performed well in estimating the canopy biomass on a pixel scale, leading to general confidence in the algorithm to fill the scan line missing data gaps (Fig. 6).

Since most studies using the Landsat kelp biomass dataset have combined pixels together into coastline segments or patches, we...
RMSE is also reported for each sensor and the overall relationship. All relationships are significant at the p < 0.05 level.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>TM</th>
<th>ETM+</th>
<th>OLI</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canopy biomass</td>
<td>0.62</td>
<td>0.63</td>
<td>0.55</td>
<td>0.64</td>
</tr>
<tr>
<td>Plant density</td>
<td>0.19</td>
<td>0.13</td>
<td>0.21</td>
<td>0.17</td>
</tr>
<tr>
<td>Frond density</td>
<td>0.32</td>
<td>0.44</td>
<td>0.31</td>
<td>0.41</td>
</tr>
<tr>
<td>Equation canopy biomass</td>
<td>$y = 6.21x + 0.48 (0.56, 0.17)$</td>
<td>$y = 6.52x + 0.28 (0.44, 0.13)$</td>
<td>$y = 7.25x + 0.18 (0.87, 0.11)$</td>
<td>$y = 6.53x + 0.30 (0.31, 0.09)$</td>
</tr>
<tr>
<td>Biomass RMSE (kg m$^{-2}$)</td>
<td>0.973</td>
<td>1.063</td>
<td>0.613</td>
<td>0.968</td>
</tr>
</tbody>
</table>

compared the total canopy biomass of coastline segments affected by scan line missing data gaps. We aggregated pixels into 500 m coastline segments by assigning each pixel to its closest coastline point along a 500 m grid. We then determined if part of that segment contained any missing data lines and excluded those which did not from the analysis (Fig. 5b). The overall relationship had a coefficient of determination of 0.96 (p < 0.0001; $y = 0.98x + 16,000$). This analysis shows that the gap filling algorithm provides excellent data for studies that aggregate data into coastline segments or patches.

#### 3.5. Effect of scale on detecting long-term trends in biomass

The length of the kelp biomass time series had a considerable effect on the magnitude and direction of significant biomass trends. There was a cyclical pattern in the strength of the trend, as measured by the mean slope of biomass vs. time, across all sites and was dependent on the length of the time series studied. Shorter time series between 10 and 13 years in length displayed a relatively low mean slope, mostly between 0.004 and 0.007, while longer time series of between 15 and 25 years in length increased to between 0.005 and 0.0085, before decreasing again for time series longer than 20 years (Fig. 7a). Low frequency climate oscillations (NPGO, MEI, PDO) displayed varied patterns across the time period of study, but with a general trend of absolute maximum slopes between 12 and 22 years and generally weakening mean slopes for time series > 20 years (Figs. 7b, S5). The cyclical patterns in mean slope across time series length for both kelp biomass and the NPGO/MEI were significantly related across all spatial scales with a lag of one year (Table 2). Shorter period time series (e.g. 10 years) of the NPGO showed a weak mean slope, due to a mixture of positive and negative trends (Fig. 7c), while longer time series (e.g. ~14–22 years) showed more consistent trends through time (Fig. 7d).

The effect of temporal scale was variable across the study area and spatial scale of study. Positive mean slopes were more prominent in the southern portion of the study area; however positive and negative trends were apparent throughout (Fig. 8). As the dataset was resampled to larger spatial scale time series the cyclical pattern of mean slopes across time series lengths was conserved (Fig. 7a). However, the proportion of sites with significant trends across longer (34 years) time series was reduced and more consistent biomass trends were observed (Fig. 9).

### 4. Discussion

#### 4.1. Accounting for spectral differences and environmental biases among sensors

The three Landsat sensors examined in this study differ in several important respects, including radiometric quantization and signal to noise ratio. The greatest concern was associated with changes in the spectral band response functions between the TM/ETM+ and OLI sensors, which had the potential to bias kelp canopy biomass estimates. The spectral bands of the OLI sensor, although broadly comparable to the TM and ETM+ sensors, were narrowed and in particular the near infrared band was optimized to avoid the water absorption feature at 0.825 μm (Roy et al., 2014). While there are several methods for correcting for these spectral differences, including the use of a different endmember for the MESMA analysis depending on the sensor, we decided to use existing hyperspectral imagery to simulate Landsat imagery for each sensor. We identified a nonlinear relationship between the TM/ETM+ and OLI sensors and were able to correct for this bias to standardize biomass estimates and retain the use of a consistent kelp canopy spectral endmember (Fig. 3).

The use of hyperspectral imagery also allowed for the examination of simulated imagery from each sensor without confounding variables.

*Fig. 5.* a. The difference in tidal height between all proximate (8 days difference) Landsat 5 TM and Landsat 7 ETM+ images during the overlap period without the scan line corrector error (1999–2003) for the Landsat tiles associated with the Santa Barbara + Northern Channel Islands, Los Angeles + Southern Channel Islands, and San Diego coastline. The dashed lines show the mean tidal difference across all dates for each Landsat tile. b. Relationships between the patch scale canopy biomass across the TM and ETM + Landsat scenes before tidal correction. c. Relationships between the patch scale canopy biomass across the TM and ETM + Landsat scenes after the tidal correction. The dashed line is the 1:1 line and the solid lines are the reduced major axis linear regression fit line for each Landsat tile.
like canopy biomass accumulation and loss or differences in environmental variables like tides and currents. Since giant kelp is one of the fastest growing organisms on Earth (frond elongation can be ≥15 cm d⁻¹) and entire stands can be dislodged by large wave events, a level of caution should be used when comparing pixel scale biomass estimates produced even a few weeks apart. However, there are differences in the amount of visible canopy structure which occurs over the space of a few hours due to changes in tide. The influence of tide has been well studied in other coastal ecosystems like salt marshes, where increasing tidal inundation decreased the near infrared reflectance and shifted the location of the red-edge (Kearney et al., 2009; Turpie, 2013).

We found differences in patch scale canopy biomass that mirrored the differences in tidal state between the TM and ETM+ sensors (Fig. 5). If tidal state occurred randomly, this would simply lead to a greater amount of scatter when comparing the sensors. However, the 8-day difference in sensor overpass time linked with the tidal cycle during our period of comparison, leading to a systematic bias that varied by Landsat tile. Tidal differences present an additional layer of complexity to coastal remote sensing that is akin to ‘hitting a moving target’ as portions of the landscape are periodically submerged. We were able to account successfully for tidal state on a patch scale (Fig. 5c), however entire pixels may be lost if canopy biomass is submerged or drops below the detection limit.

Kelp fraction estimates were significantly related to diver estimated kelp variables with canopy biomass displaying the strongest relationship (Table 1). All three sensors had nearly equivalent linear relationships with the lowest coefficient of determination attributed to the OLI sensor (Table 1; Fig. 4). This is probably due to the limited variability of kelp observed at the monitoring sites since the launch of Landsat 8 in 2013. The years 2014–2016 were characterized as low biomass years due to large scale ocean warming events that prevented the formation of a thick, high biomass canopy (Reed et al., 2016). Work should continue to verify that the OLI relationship does not deviate from those of TM and ETM+ as higher canopy densities are observed. Kelp fractions were also related to frond and plant density, but the relationships with these variables were not as strong as with canopy biomass (Table 1). Since the sensor is only observing floating canopy there is bound to be a decrease in explanatory potential as one tries to infer structures that are partially or completely submerged. Furthermore, giant kelp is morphologically plastic, with traits that are strongly influenced by environmental conditions (reviewed in Graham et al., 2007). While the subsurface architecture may change drastically, the physics of canopy reflectance are likely conserved across phenotypes. These are important considerations for researchers aiming to answer population-based questions across regions while using canopy biomass as a proxy.

It should be noted that the in situ diver data is imperfect and so some of the uncertainty in the relationship between diver and Landsat estimates was likely due to inaccuracies and imprecision in the diver estimation.
estimates. Additionally, since diver estimates are limited to a few sites, it would be beneficial to expand the spatial coverage of ground truth areas and confirm that the frond length to weight relationship is conserved across regions. New technologies such as unmanned aerial vehicles can be used to rapidly obtain more spatially comprehensive data of kelp canopy across many areas. This method can be more easily timed with Landsat passes to avoid tidal effects and provide fine resolution estimates of extensive canopy area, which can be linked to Landsat scale density estimates across many canopy-forming kelp species.

4.2. Towards a continuous kelp canopy time series

A continuous, long-term satellite record is paramount to understanding the world’s most pressing environmental challenges (Birdsey et al., 2009). Giant kelp is incredibly dynamic through time and across space and several studies have required continuous time series of these dynamics in order to estimate patch modularity, investigate persistence, and track extinction and recolonization dynamics (Cavanaugh et al., 2014; Young et al., 2016; Castorani et al., 2015, 2017). Furthermore, a complete record of canopy biomass was useful for comparing recent canopy declines associated with a large-scale anomalous ocean warming event to previous kelp declines on regional scales (Fig. 1b; Reed et al., 2016). The ETM+ scan line corrector failure compromised the continuity of the canopy biomass time series and necessitated the creation of a procedure for filling missing data values. This was especially true during the period from December 2011 to April 2013 when ETM+ was the only operational Landsat sensor. We leveraged the high degree of synchrony of giant kelp biomass over the first several hundred meters to develop an improved method versus other gap-filling techniques such as temporal interpolation. The synchrony method sufficiently produced pixel scale biomass patterns observed in the validation imagery and displayed even better performance when larger scale spatial sums where compared (Fig. S4). This process of using known spatial associations through time has the potential to be generalized for filling reflectance values across ETM+ bands to produce a gap-filled reflectance product for a variety of long-term environmental studies.
which require continuous data.

Through the merging of the three Landsat sensors, the accounting of environmental biases, and the estimation of missing data through the gap filling method, we have created a continuous, seasonal time series of giant kelp canopy biomass along the coast of California (Fig. 1b). These biomass data, along with biomass uncertainties, are publicly available through the Santa Barbara Coastal Long Term Ecological Research project website for use by researchers (Bell et al., 2017). Work is ongoing extending the automated Landsat algorithm to areas in Alaska, Oregon, Washington, and Baja California towards a time series of kelp canopy dynamics across the Northeast Pacific.

4.3. Effect of time and space scale on long-term trends

The examination of temporal scale on the detection and direction of long-term trends in giant kelp biomass revealed a cyclical pattern associated with low frequency marine climate oscillations. The relationship was related most closely to the NPGO, echoing results shown in previous studies, while relationships with the PDO were weak, at least over the timescales observed in this study (Table 2; Cavanaugh et al., 2011; Bell et al., 2015b). Shorter scale (~10 years) time series are subject to periodic increasing and decreasing trends associated with the periodicity of these decadal scale climate oscillations leading to a high instance of trend detectability but weaker mean slopes when all 10-year time series are considered, especially for the NPGO (Fig. 7a, c). However, if each trend is viewed on its own and outside of the context of these low frequency oscillations it has the potential to be misinterpreted as a long-term trend in kelp biomass state for a specific site. Additionally, these marine climate oscillations are not entirely stable through time and may be subject to natural trends over longer time
continuous or short time series may lead to spurious conclusions of examined (Krumhansl et al., 2016). Furthermore, the use of dis-

tervals of signiﬁcant biomass trends at the 34 year time scale and the error bars represent the standard deviation of those signiﬁcant slopes relative to the maximum standard deviation of 0.0036 at a 1 km spatial scale. The red line is the best ﬁt line. (For interpretation of the references to color in this ﬁgure legend, the reader is referred to the web version of this article.)

periods (Mcphaden et al., 2011). For instance, we found that during the time period of the kelp biomass dataset (1984–2017) the NPGO has shown mean increasing trends over longer scale (~20 years) time series, while the MEI and PDO have shown mean decreasing trends (Figs. 7b, d; SS). Taken together these index trends represent a trend towards cooler, more nutrient rich waters in the California Current across this longer time scale, leading to a favorable environment for giant kelp biomass accumulation (Bell et al., 2018). Not surprisingly, we see mean increasing biomass trends for giant kelp at these longer scale time series, with a peak between 15 and 25 years (Fig. 7a). Recent large scale warming events such as the 2014 anomalous North Paciﬁc warming event and the 2015–2016 El Niño have begun to reduce the mean trends for ocean climate indices and kelp biomass time series longer than 25 years (Fig. 7a, b; Bond et al., 2015). However, it remains to be seen whether these longer trends (> 25 years) are indicative of climate change effects or an even longer-term or coupled natural cycle leading to the return of the unfavorable environmental conditions observed throughout the 1980’s and 90’s (Joh and Di Lorenzo, 2017).

An inverse relationship was found between the detection of long-term kelp biomass trends (34 years) and the spatial scale of study (Fig. 9). While there was a mix of positive and negative long-term trends seen through space at the 1 km scale, there was a reduction in the proportion of signiﬁcant trends detected as spatial scale increased (Fig. 8). This could mean that many of the smaller spatial scale long-term trends seen in the dataset are indicative of local scale e

Acknowledgements

We would like to acknowledge the support of the US National Science Foundation which provided funding for the Santa Barbara Coastal LTER (grant numbers OCE 0620276 & 1232779) and Coastal SEES grant (grant number OCE 1600230), and also support by the National Aeronautics and Space Administration grant (grant number NNX14AR62A) as part of the Biodiversity and Ecological Forecasting program. The National Aeronautics and Space Administration’s support of T.W.B. through the Earth and Space Sciences Fellowship program made this work possible. We would also like to extend supreme gratitude to the many ﬁeld technicians and student volunteers who conduct ﬁeld and laboratory work for the Santa Barbara Coastal LTER. Special thanks go to guest editors Dr. Tom Loveland, Dr. Curtis Woodcock, and Dr. Martin Herold and to three anonymous reviewers whose comments improved this manuscript.

References


