Giant Kelp and Bull Kelp Canopy Dynamics from the Landsat Satellite Sensors (TM, ETM+, OLI)
Santa Barbara Coastal LTER 2020

Figure 1. Timeline of the three Landsat sensors used for the kelp canopy time series. Landsat 5 Thematic Mapper (TM) acquired imagery from 1984 – 2011, while Landsat 7 Enhanced Thematic Mapper + (ETM+) and Landsat 8 Operational Land Imager (OLI) are both currently in operation. Landsat 7 ETM+ experienced a scan line corrector error in May of 2003, shown as the light-yellow section of the timeline (see more below).

Overview

The Landsat sensors have acquired 30 m spatial resolution multispectral imagery continuously from 1984 – present (Figure 1), with each sensor imaging the globe every 16 days. We use these images to estimate the canopy area of giant kelp (*Macrocystis pyrifera*) and bull kelp (*Nereocystis luetkeana)* and the canopy biomass of giant kelp along the coast of the northeast Pacific. Compared to seawater, emergent kelp canopy presents relatively high reflectance in the near infrared region of the electromagnetic spectrum and allows for accurate and consistent retrievals of the fraction of kelp canopy in each Landsat pixel across variable ocean conditions. By relating the estimated kelp fraction to higher resolution aerial imagery and long-term diver estimates of canopy biomass, we validated these fractional estimates and applied these relationships to kelp forests along the California, Oregon, Washington and Baja California coast.

Giant Kelp Canopy Biomass Methodology

A full explanation of this methodology can be found in Bell et al. (2020).

Giant kelp canopy biomass was estimated from Landsat TM, ETM+, and OLI 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 (Version 2) was used to mask out all pixels above 0 meters elevation (asterweb.jpl.nasa.gov/gdem.asp; Figure 2.1). In order to mask out beaches and intertidal areas, a 30 m buffer was applied to this mask. Additionally, an intertidal mask was implemented using the modified normalized difference water index (MNDWI) for

extreme low tide Landsat imagery. 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*; Figure 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. After all the images were classified, we filtered errors of commission (free floating kelp paddies, spectral image errors) by removing any pixels classified as 'kelp canopy' in < 1% of the time series images. Images collected at tidal height of < 1m were removed for areas where there was a significant influence of tide (mostly around the coastal areas of Washington).

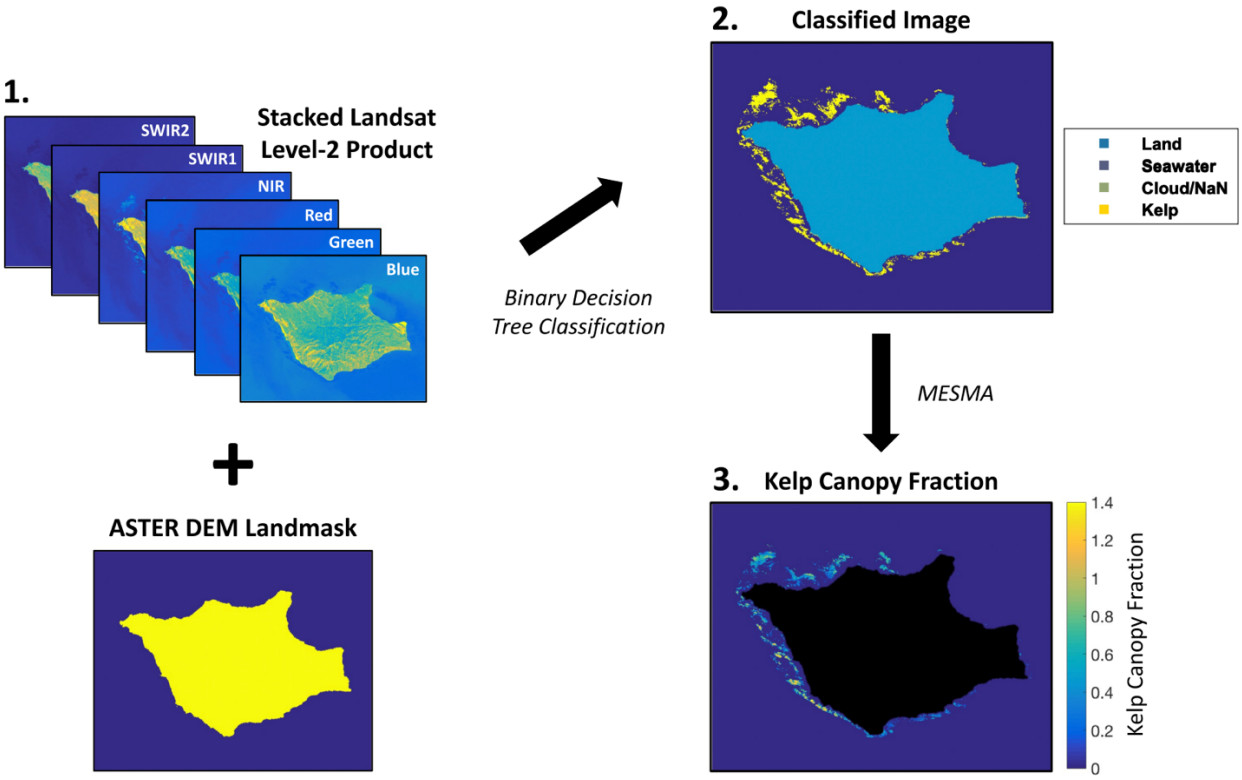

Figure 2. Conceptual model of the automated giant kelp canopy fraction processing scheme at Santa Rosa Island, California, USA (33.97 N, 120.11 W). **1.** USGS Level-2 Landsat Surface Reflectance images are stacked and land is masked using the ASTER DEM with a 30 m coastline buffer. **2.** Stacked images are classified into four categories using a binary classification decision tree trained using Landsat imagery. **3.** Landsat pixels classified as 'kelp' are modeled as combinations of kelp canopy and seawater using Multiple Endmember Spectral Mixture Analysis (MESMA) and fractional kelp canopy cover is estimated at a 30 m pixel scale.

We estimated relative kelp canopy density from the calibrated reflectance data using multiple endmember spectral mixing analysis (MESMA; Figure 2.3; Roberts et al. 1998). Spectral mixture analysis models the fractional cover of two or more endmembers within a pixel. Each endmember represents a pure cover type and endmembers are assumed to combine linearly (Adams et al. 1993). Standard spectral mixture analysis uses a uniform set of endmembers for the entire image. This approach was problematic for the estimation of giant kelp canopy because of varying water conditions across the image and through time. The reflectance of seawater in the near-shore marine environment is influenced by sun glint, breaking surface waves, phytoplankton blooms, dissolved organic matter, and suspected sediment. Due to the highly variable seawater reflectance, a single seawater endmember cannot be used.

The MESMA process allows endmembers to vary on a per pixel basis by selecting from multiple endmembers for one or more cover types. This technique can better capture the spectral variability of a cover type through space and time. We modeled pixel reflectance as the linear mixture of reflectance from two endmembers: giant kelp canopy and seawater. Thirty seawater endmembers were selected from consistently non-kelp covered areas within each Landsat scene. The spectral information collected from these seawater locations varied between images to account for changing seawater conditions. A single kelp endmember was selected by extracting kelp covered pixel spectra for a series of images and finding the single spectrum that fit the entire library of kelp spectra with the lowest root mean square error (RMSE; Dennison & Roberts 2003). The pixels of each image were modeled as a two-endmember mixture of kelp and each of the 30 water endmembers that were free of cloud contamination. The final model (out of 30) chosen for each pixel was the model that minimized the RMSE when fit to the spectrum of that pixel. The result of this process was a measure of the relative fraction of each pixel covered by kelp canopy. The MESMA process successfully estimated the relative canopy fraction of giant kelp under a variety of conditions including large amounts of sediment runoff and high levels of sun glint (Cavanaugh et al. 2011).

Each Landsat sensor differs in signal to noise ratio, radiometric calibration, and the number and width of their spectral bands. The most important difference for the Landsat kelp time series was the ‘optimization’ of the Landsat 8 OLI spectral bands relative to the Landsat 5 TM and Landsat ETM+ sensors. Due to the lack of temporal overlap between Landsat 5 and Landsat 8, the dynamic nature of the giant kelp canopy, and the fact that each image is acquired 8 days apart for the overlapping sensors, the best way to compare the kelp fraction retrievals from each sensor was by using simulated imagery. We took advantage of aerial hyperspectral imagery over the Santa Barbara Channel collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) mounted on an ER-2 aircraft flying at 65,000 feet to simulate hyperspectral satellite imagery. We resampled a hyperspectral image collected in April 2013 to all three Landsat sensors’ spectral bands using published spectral response functions. The MESMA process was then applied to each simulated image using 30 seawater endmembers from the same locations in each simulated image and the kelp endmember discussed above. Kelp fractions were then compared between each simulated sensor pair (Figure 3). Kelp fractions estimated from the Landsat 5 TM and Landsat 7 ETM+ were comparable however Landsat 8 OLI kelp fractions were consistently lower for the same amount of kelp canopy. To address this difference, we adjusted the Landsat 8 kelp fractions using the equation in Figure 3f to match those of the previous Landsat sensors.

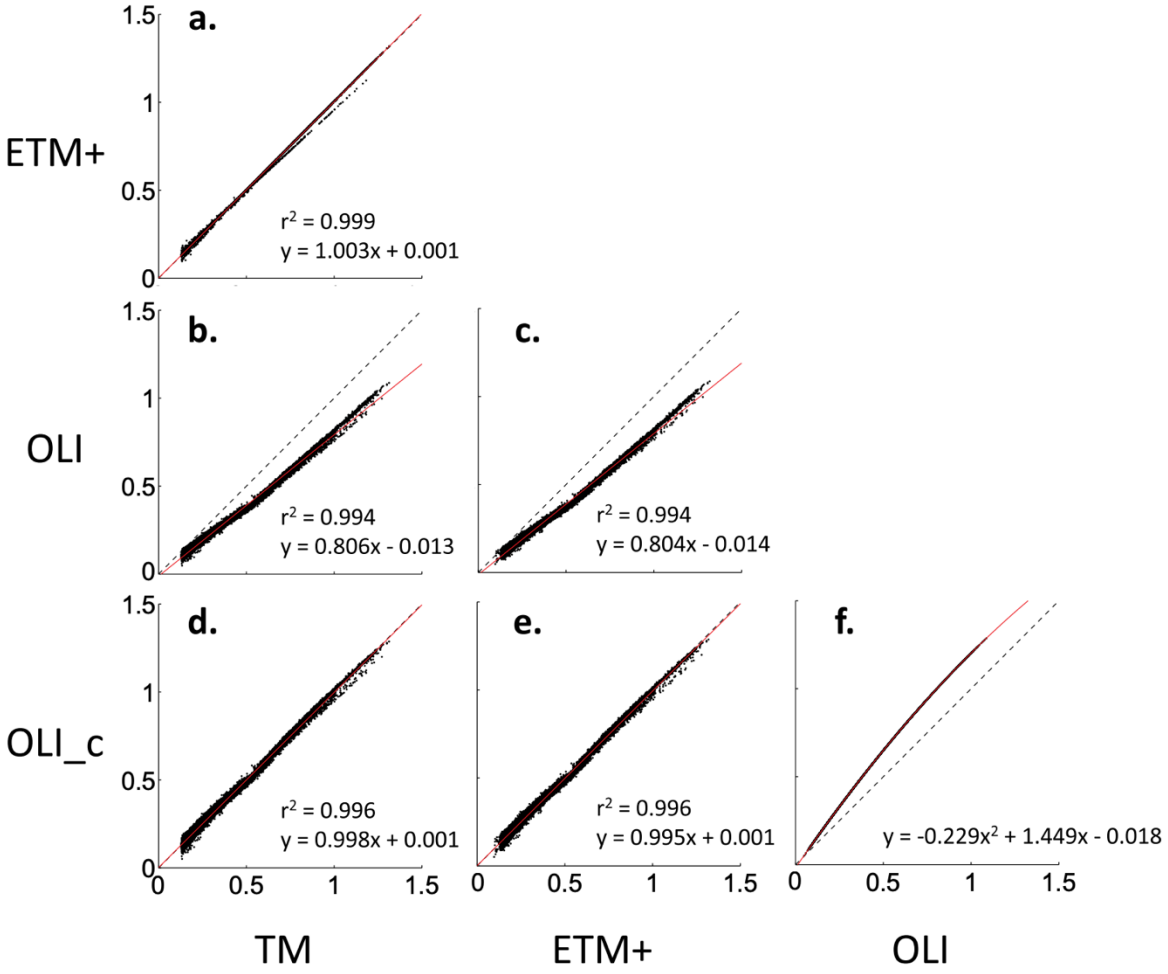

Figure 3. Scatterplot matrix of MESMA derived kelp fractions from the simulated images of each Landsat sensor (from AVIRIS hyperspectral imagery), compared against every other Landsat sensor used in the kelp canopy biomass time series. The dashed black lines show the 1:1 line while the red lines are the best fit lines for each individual scatterplot. TM represents the Thematic Mapper, ETM+ is the Enhanced Thematic Mapper+, OLI is the Operational Land Imager, and OLI_c it the Operational Land Imager kelp fraction corrected to match the TM and ETM+ kelp fraction estimates. All relationships are significant at the $p < 0.001$ level.

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. 2018). 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.

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 (Figure 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 (Equation 1),

$$y = 6.53x + 0.30 \qquad \text{Eq. 1}$$

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.

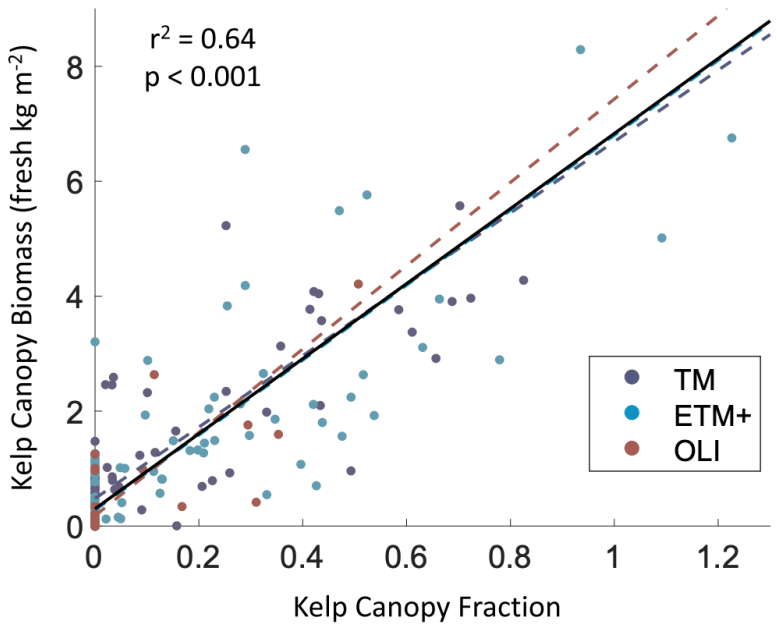

Figure 4. Validation of Landsat satellite kelp fraction estimates of the three sensors versus diver-estimated canopy biomass from the two Santa Barbara Coastal Long Term Ecological Research project study sites. Each dotted line represents the reduced major axis linear regression fit line for each Landsat sensor. The solid black line represents the reduced major axis linear regression fit line across all three sensors.

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

Table 1. The coefficient of determination (r^2) and reduced major axis linear regression line equations between Landsat estimated kelp fraction and diver-estimated kelp variables for each Landsat sensor and across all sensors (standard deviations for the slope and intercept are reported in parentheses). The root mean squared error (RMSE) is also reported for each sensor and the overall relationship. All relationships are significant at the $p < 0.05$ level.

Enhanced Thematic Mapper Plus images acquired after May 2003 were subject to the scan line corrector failure, which resulted in ~20% of the data in each image being lost as missing data lines. These missing data manifest as 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). Kelp canopy biomass is known to display high, but exponentially decreasing, spatial synchrony over the first few 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 500 m 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 > 0.7, 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. 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.

Pixels filled using spatial synchrony had a $r^2 = 0.83$ ($p < 0.0001$; $y = 0.94x + 74$). The 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 (Figure 5). Since most studies using the Landsat kelp biomass dataset have combined pixels together into coastline segments or patches, we 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. 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.

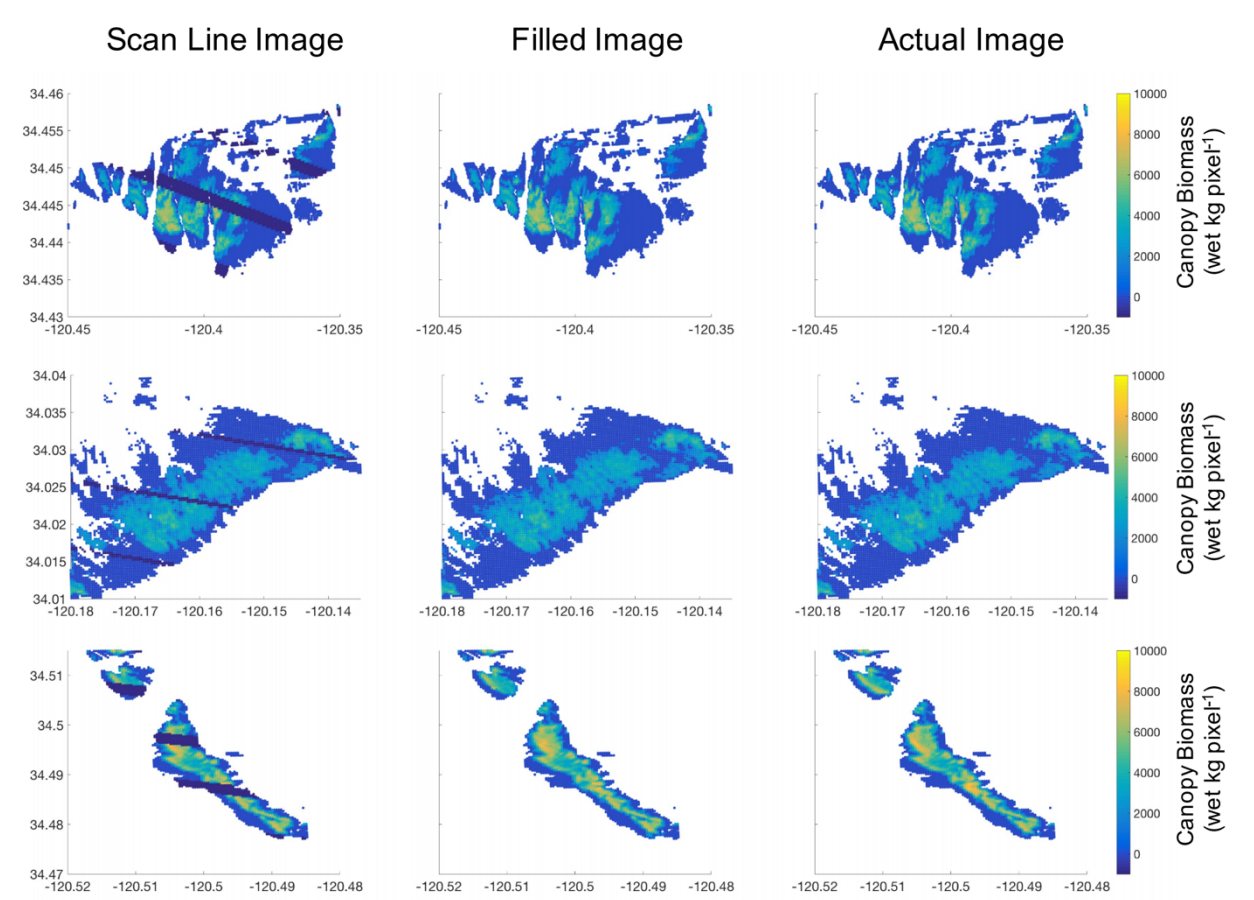

Figure 5. Results of the gap filling algorithm across three kelp forest canopies in the Landsat 5 TM validation images. In the first column, kelp forest pixels are masked with multiple dark blue simulated scan line error missing data lines. The second column shows the results of the gap filling algorithm. The third column shows the actual images before the scan line error missing data lines.

Kelp Canopy Area Methodology

A full explanation of this methodology can be found in Hamilton et al. (2020).

Kelp canopy area was estimated from Landsat imagery using the same automated classification protocol explained above. We compared Landsat-derived estimates of canopy fractional cover to high-resolution (~1 m) aerial color-infrared photographs taken by Bergman Photographic, Inc. as part of an effort by the Oregon Department of Fish and Wildlife to survey kelp populations statewide. These images of primarily bull kelp canopy, were taken for several years in the 1990s in late summer and early fall at neutral tides. For a more complete description of the aerial photography, see Fox et al. (1996). We used 11 of these photographs of Orford Reef, Cape Blanco Reef, and Rogue Reef that had a corresponding Landsat 5 TM or Landsat 7 ETM+ image of the same reef taken within 7 days of the aerial photograph. We georeferenced the aerial

photographs, manually removed rocks and islands, and identified kelp canopy in each aerial image pixel by calculating normalized difference vegetation index (NDVI; Equation 2):

$$NDVI = \frac{(NIR-Red)}{(NIR+Red)} \quad \text{Eq. 2}$$

where NIR stands for near infrared reflectance and the Red for red reflectance. Once kelp canopy was identified, we binned pixels in the aerial photos into 30 m cells corresponding to the 30 m grid of a Landsat image and summed the percentage of each 30 m cell that was covered in kelp. To quantify the percentage of each Landsat pixel covered in kelp, we used (1) a binary decision tree that determines whether a pixel contains kelp, seawater, cloud, or exposed rock/land based on blue, green, red, near infrared, and short-wave infrared reflectance; and (2) MESMA to estimate the kelp and seawater fractional cover within pixels that contain kelp. For more detailed information on the algorithm see above or Bell et al. (2020). We applied a two-dimensional Gaussian filter (σ = 0.5) to both images to account for potential changes in current direction, which can shift the canopy between image dates. We then compared per-pixel estimates of canopy cover derived from the aerial photos to the fractional canopy coverage derived from the Landsat images. We quantified accuracy using reduced major axis linear regression and by calculating the rate at which the algorithm misidentified kelp as water and vice versa. To minimize spatial autocorrelation in our algorithm validation statistics, we only included pixels that were > 150 m apart. Points with less than 1% canopy density and/or less than 1% kelp canopy fraction were removed from the analysis. A strong linear relationship was found across all dates (Figure 6a; r^2 = 0.779, $p < 0.001$, $y = 1.03x - 0.024$). Since we used high spatial resolution imagery to identify the proportion of kelp that covered each 30 m pixel, we calculated how often the algorithm misidentifies kelp as water and vice versa in Landsat imagery. We found that the Landsat sensors did an excellent job of identifying seawater and misidentified it as kelp less than 1% of the time. For kelp canopy, the algorithm identified 70% of the kelp-containing pixels from the aerial imagery and missed 30% of those pixels. We then investigated those times when kelp was missed. We found that canopy densities of > 20% were rarely missed by the algorithm in Landsat images and that the miss rate increased rapidly at canopy densities of 30 m away from exposed rocks or coastlines (Figure 6b). Across these three reefs, the total amount of kelp missed due to the miss rate plus coastline adjacency effects equals 10.2% when moving from the high-resolution aerial imagery to the Landsat method.

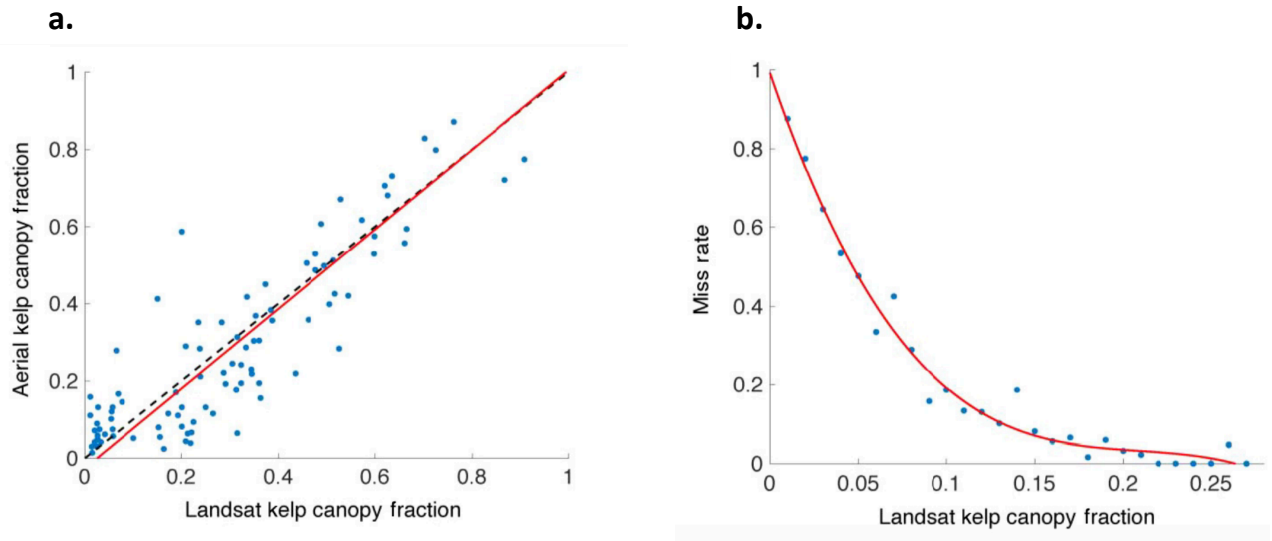

Figure 6. (a) Relationship between aerial photography derived canopy density and Landsat-derived canopy density. The red line is the best-fit linear relationship and the dashed black line a 1:1 line. Units are the fraction of each 30 m pixel that was covered by kelp canopy. (b) Miss rate as function of Landsat kelp canopy fraction. The red line shows the best-fit nonlinear relationship.

Dataset

Cloud-free imagery of the coastline is available about every 1 – 3 months during the time series. This frequency increases when multiple satellites are acquiring imagery. We take the mean of all area/biomass estimates within a season (3-months quarters). The current version of the dataset includes 25 Landsat scenes, which cover the entire coastline of California, Oregon, Washington, and Baja California, including offshore islands. The dataset is provided as a netCDF file that includes the mean kelp canopy area (bull kelp and giant kelp) and biomass (giant kelp) of every identified kelp containing pixel, for each season, from 1984 – 2020, along with relevant metadata.

References

Adams J, Smith M, Gillespie A (1993) Imaging spectroscopy: interpretation based on spectral mixture analysis. In: Pieters CM, Englert P (eds) Remote geochemical analysis: elemental and mineralogical composition. Cambridge University Press, New York, NY, p 145–166

Bell, T.W., Allen, J.G., Cavanaugh, K.C., Siegel, D.A., 2020. Three decades of variability in California's giant kelp forests from the Landsat satellites. Remote Sensing of Environment, Time Series Analysis with High Spatial Resolution Imagery 238, 110811. https://doi.org/10.1016/j.rse.2018.06.039

Cavanaugh, K.C., Kendall, B.E., Siegel, D.A., Reed, D.C., Alberto, F., Assis, J., 2013. Synchrony in dynamics of giant kelp forests is driven by both local recruitment and regional environmental controls. Ecology 94, 499–509. https://doi.org/10.1890/12-0268.1

Cavanaugh, K.C., Siegel, D.A., Reed, D.C., Dennison, P.E., 2011. Environmental controls of giant-kelp biomass in the Santa Barbara Channel, California. Marine Ecology Progress Series 429, 1–17. https://doi.org/10.3354/meps09141

Dennison, P.E., Roberts, D.A., 2003. Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. Remote Sensing of Environment 87, 123–135. https://doi.org/10.1016/S0034-4257(03)00135-4

Fox, D.S., Merems, A., Golden, J., Amend, M. 1996. 1996 Kelp/reef habitat assessment. Newport, OR: Oregon Department of Fish and Wildlife. 61pp.

Hamilton, S.L., Bell, T.W., Watson, J.R., Grorud-Colvert, K.A., Menge, B.A., 2020. Remote sensing: generation of long-term kelp bed data sets for evaluation of impacts of climatic variation. Ecology 101, e03031. https://doi.org/10.1002/ecy.3031

Morton, D.N., Bell, T.W., Anderson, T.W., 2016. Spatial synchrony of amphipods in giant kelp forests. Mar Biol 163, 32. https://doi.org/10.1007/s00227-015-2807-5

Rassweiler, A., Reed, D.C., Harrer, S.L., Nelson, J.C., 2018. Improved estimates of net primary production, growth, and standing crop of Macrocystis pyrifera in Southern California. Ecology 99, 2132–2132. https://doi.org/10.1002/ecy.2440

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., Green, R.O., 1998. Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models. Remote Sensing of Environment 65, 267–279. https://doi.org/10.1016/S0034-4257(98)00037-6