

Robust ocean color from drones: Viewing geometry, sky reflection removal, uncertainty analysis, and a survey of the Gulf Stream front

Patrick Clifton Gray¹*, Anna E. Windle,² Julian Dale,¹ Ivan B. Savelyev,³ Zackary I. Johnson,¹ Greg M. Silsbe,² Gregory D. Larsen,¹ David W. Johnston¹

¹Duke University Marine Laboratory, Nicholas School of the Environment, Duke University, Beaufort, North Carolina

²Horn Point Laboratory, University of Maryland Center for Environmental Science, Cambridge, Maryland

³U.S. Naval Research Laboratory, Washington, District of Columbia

Abstract

Accurate and robust retrieval of ocean color from remote sensing enables critical observations of aquatic natural systems, from open ocean biological oceanography, coastal biodiversity, and water quality for human health. In the last decade, studies have increasingly highlighted the important role of small-scale processes in coastal and marine ecology and biogeochemistry, but observation and modeling at these scales remains technologically limited. Unoccupied aircraft systems (UAS, aka drones) can rapidly sample large areas with high spatial and temporal resolution; but the challenge of accurately retrieving ocean color, particularly with common wide field-of-view multispectral imagers, has limited the adoption of this technology. As UAS endurance, autonomy, and sensor capabilities continue to increase, so does this technology's potential to observe the ocean at fine scales, but only if proper protocols are followed. The present study provides a guide for achieving (1) ideal viewing geometry of UAS-borne ocean color sensors, (2) techniques for the removal of sun glint and reflected skylight to derive water-leaving radiances, (3) characterization of uncertainty in these measurements, and (4) converting water-leaving radiances to remote-sensing reflectance for analytic end products such as chlorophyll *a* estimates. Documented open-source code facilitates replication of this emerging technique. Using this methodology, we briefly describing fine-scale variability of the Gulf Stream front off North Carolina alongside synoptic satellite data and in situ measurements for comparison. These results demonstrate how UAS-based ocean color measurements complement and enhance conventional ocean observations and modeling to resolve fine-scale variability and close the lacuna between satellite and in situ methods.

Fine-scale ocean color observations

Remote sensing of ocean color informs ocean biology by using light to describe spatial distributions, community compositions, types and gradients of organic matter, particle sizes, productivity levels, or simply chlorophyll *a* (Chl *a*) concentrations (International Ocean Color Coordinating Group [IOCCG] 2008; McClain 2009). Advancements in understanding ocean biogeochemistry have often accompanied

advancements in ocean color remote sensing: early experimental sensors in personal airplanes confirmed an association between decreasing blue light from the ocean with increasing Chl *a* (Clarke et al. 1970), and soon National Aeronautics and Space Administration's (NASA) upcoming Plankton, Aerosol, Cloud, ocean, Ecosystem mission will add hyperspectral ocean color measurements and polarimetry that will reveal new details of the carbon cycle and phytoplankton communities (Werdell et al. 2019). Even still, current and upcoming satellite remote-sensing approaches cannot describe submesoscale ($O[3-30 \text{ km}]$) and fine-scale ($O[< 3 \text{ km}]$) ocean features, cannot resolve coastal areas due to their scale and optical complexity, and are often hindered by cloud cover. Higher-resolution terrestrial satellites, such as Landsat and Sentinel-2 may aid in this work, but their application to marine systems is hindered by insufficient atmospheric correction, spectral band locations that are not tuned for marine applications and thus miss key absorption peaks, lower sensitivity, and long revisit times. In

*Correspondence: patrick.c.gray@duke.edu

Additional Supporting Information may be found in the online version of this article.

Author Contribution Statement: P.G.: conceptualization, data collection, data analysis, visualization, writing, funding. A.W.: conceptualization, data collection, synthesis, writing. J.D.: data collection, UAS design, UAS piloting, writing. I.S.: data collection, writing. Z.J.: data collection, writing. G.S.: data collection, writing. G.L.: data collection, writing. D.J.: conceptualization, synthesis, writing, funding.

addition, these terrestrial focused satellites typically do not downlink imagery of the open ocean where coastline is not in the image (Pahlevan et al. 2017; Wulder et al. 2019).

Submesoscale fluid dynamics can drive nutrient fluxes across the mixed layer, can subduct particles from the mixed layer below the thermocline, and can increase stratification by horizontally tilting density layers (Mahadevan 2016). Such processes occur at scales of time and space germane to the life cycles of phytoplankton, and accordingly they appear to strongly modulate primary productivity, ecosystem structure, and marine biodiversity (Lévy et al. 2018). Over the last two decades, this alignment of scale has inspired extensive study into the biological significance of submesoscale ocean dynamics (Lévy et al. 2001; D’Ovidio et al. 2010; Ruiz et al. 2019). This intensifying effort demands remote-sensing instruments that complement satellite observations at the scale of ocean basins (e.g., Aqua’s Moderate Resolution Imaging Spectroradiometer [MODIS] and Suomi NPP’s Visible Infrared Imaging Radiometer Suite [VIIRS] platforms) by resolving submesoscale dynamics.

Strong persistent mesoscale fronts, such as the Gulf Stream’s north wall, exhibit consistent submesoscale dynamics (Mcwilliams et al. 2019); they are linked to increased productivity in phytoplankton, zooplankton, and fish, and represent a hotspot of ocean predator diversity (Mann and Lazier 2005). The Gulf Stream front off Cape Hatteras, North Carolina is an ideal testbed for demonstrating fine-scale ocean observation techniques. Though the current is overall in geostrophic balance, western boundary currents like the Gulf Stream often manifest substantial ageostrophic circulation, moving both nutrients and organisms vertically at high speeds compared to most of the ocean. The Gulf Stream is often assumed to be an ecological barrier at the large scale (Bower et al. 1985), but instability in the front’s geostrophic balance not only injects nutrients into the euphotic zone, but drives the formation of submesoscale filaments, eddies, and streamers on both sides of the Gulf Stream (Gula et al. 2015). Submesoscale features may dominate biogeochemical exchange across the front (Klymak et al. 2016); however, their mechanics and ecological impacts remain poorly understood due to their ephemeral and fine-scale nature.

Unoccupied aircraft systems (UAS) address this submesoscale gap in observational capacity with the ability to capture fine spatial scales and repeat observations within a single day, as is necessary to adequately characterize processes of this scale (Gray et al. 2022). Diel variability, such as changing phytoplankton fluorescence (Behrenfeld et al. 2009), can characterize the physiological status of a phytoplankton population, and subkilometer oscillations in Chl *a* concentration can indicate underlying physical drivers; but both of these dynamics can easily be missed or aliased when remote-sensing observations are even slightly too coarse in space or time. At the other extreme, in situ oceanographic methods confront a “synopticity problem” (Martin 2003), whereby they cannot collect

data at a sufficient scale or resolution to characterize target variables before flow redistributes the system and its values. UAS methods readily achieve temporal and spatial coverage between the limits of satellite and vessel-based methods, complementing these conventional techniques to achieve the synoptic view necessary to describe submesoscale physical–biological interactions (Fig. 1).

Review of optical and logistical challenges

Accurate measurement of ocean color for higher level products such as Chl *a*, productivity estimates, and phytoplankton functional types requires methods to accurately derive the radiance of light leaving the water. The spectral composition of light changes as photons transit through the water column due to absorption and scattering. By sensing these specific photons that have transited through the water column, and not those reflected off the surface or scattered in the atmosphere, remote-sensing techniques can determine properties and constituents of the water column that altered that light.

Water-leaving radiance measurements (L_w) are typically normalized by downwelling irradiance measurements (E_d) to obtain the remote-sensing reflectance (R_{rs}); this R_{rs} value is the primary goal and first product in most workflows for processing ocean color measurements. However, it can be relatively challenging to derive R_{rs} from aerial sensors, where radiance measurements of the sea surface can be dominated by photons that never entered the water, such as specular sun glint (Fig. 2a), skylight that reflects off the water’s surface (Fig. 2b), and photons scattered by the atmosphere. Atmospheric scattering is minimal for sensors at low altitudes (100 s of meters) over water (Kirk 2010; Kim et al. 2013), so sun glint and reflected skylight dominate the noise in estimates of R_{rs} in these situations (IOCCG 2019). Maintaining certain viewing geometries can substantially reduce sun glint and reflected skylight (Mobley 1999) in above water radiometry, with a recommended sensor orientation of 40° off nadir (θ) and 135° azimuthally off the sun (ϕ ; Fig. 3). It is challenging to precisely maintain these angles on a ship or platform in the ocean due to wave movement and self-shading, but a multirotor UAS can achieve consistent, optimal geometry either by flying on a fixed heading or by holding the sensor’s orientation steady on a gimbal. In addition, the small shadow of an airborne UAS does not materially influence measured R_{rs} , in contrast to that of a ship or fixed platform.

A large and growing repertoire of techniques are available to address challenges of glare and reflection in aerial imagery of ocean color (Mobley 1999; Lee et al. 2010; Zhang et al. 2017; O’Shea et al. 2020; Windle and Silsbe 2021). One approach that is widely used and easily implemented uses three measurements to obtain R_{rs} : total sea radiance (L_t), sky radiance (L_{sky}), and downwelling irradiance (E_d) (Fig. 3; Ruddick et al. 2019). These measurements are used to remove the contribution of reflected sky radiance from the total

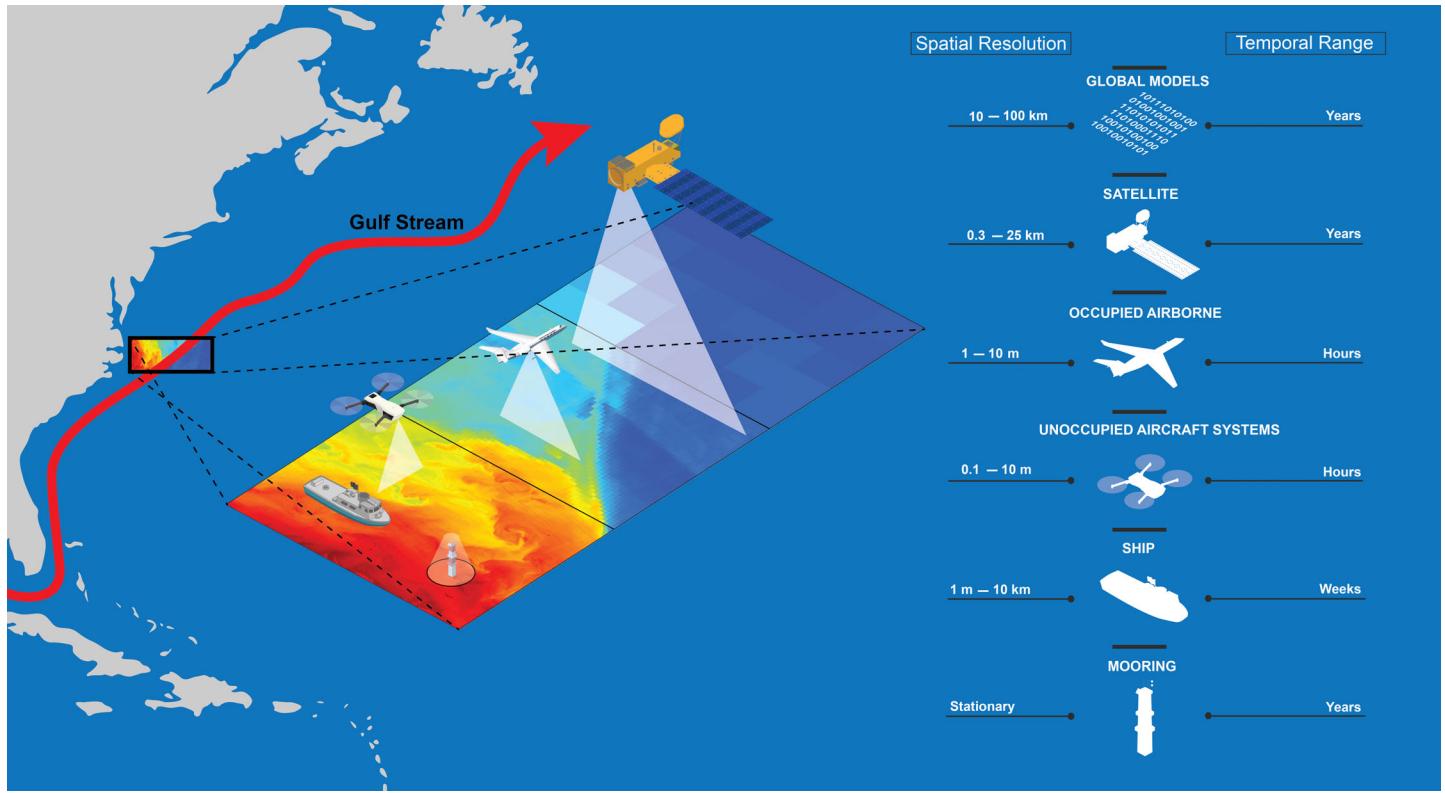


Fig. 1. Spatiotemporal scales of ocean observation platforms. Three scales of spatial resolutions are represented for a section of the Gulf Stream front off Cape Hatteras (inset) using resampled Chl α imagery from the Landsat 8 Aquatic Reflectance product: native 30-m resolution featuring submesoscale filaments and eddies; resampled 1-km resolution representing typical imagery from ocean observing satellites; resampled 10-km resolution representing typical “high-resolution” outputs of global- or basin-scale biogeochemical models. Figure inspiration from NASA ABoVE “Scaling Diagram” <https://above.nasa.gov/materials.html>.

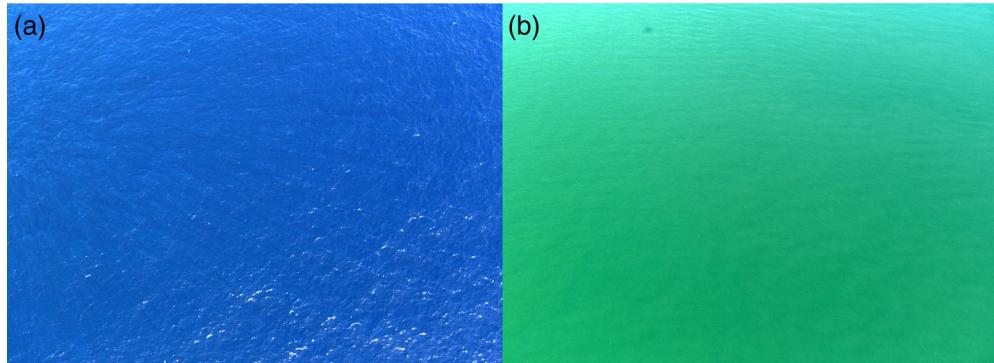


Fig. 2. Aerial measurements of ocean color are often challenged by light reflecting off the sea surface. UAS images can include specular sun glint (**a**, bottom right quadrant) that overwhelms the water leaving radiance and reflected skylight (**b**) that can shift the measured spectrum of ocean color toward that of the reflected skylight, as seen in the shift from green in the bottom of the image towards blue at the top.

radiance measured above the water’s surface, yielding only water-leaving radiance, using the following equation:

$$L_w(\theta, \varphi, \lambda) = L_t(\theta, \varphi, \lambda) - \rho(\theta, \varphi, \theta', \Omega_{\text{FOV}}, W) \times L_{\text{sky}}(\theta', \varphi, \lambda), \quad (1)$$

where θ and φ specify the polar and azimuthal viewing angles, respectively; λ is the wavelength; ρ is the effective sea-surface

reflectance of sky radiance (L_{sky})—a function of viewing geometry (θ, φ), sensor field-of-view (Ω_{FOV}); sky radiance distribution ($L_{\text{sky}}(\theta', \varphi, \lambda)$); and surface roughness, here represented by the common proxy variable wind speed (W). After subtracting the reflected sky radiance from L_t , L_w can be normalized by E_d at the surface to obtain R_{rs} using the following equation:

$$R_{rs}(\lambda) = L_w(\theta, \varphi, \lambda) / E_d(\lambda). \quad (2)$$

This calculated R_{rs} measurement is needed to generate nearly all higher-level ocean color products (IOCCG 2019).

One major challenge in this calculation is the estimation of ρ as a function of several environmental variables and which varies across the visible spectrum. Many previous estimation methods (Morel 1980; Mobley 1999) assume that ρ is constant across the spectrum, which has been shown to be incorrect (Lee et al. 2010; Zhang et al. 2017). In addition, most published examples either (1) obtain measurements from point spectrometers with a relatively small ($< 10^\circ$) field of view (FOV), yielding a single spectrum, rather than an image where each pixel represents a spectrum; or (2) obtain measurements from satellites where the pixel scale and contributors to image noise are materially different. Key assumptions for these sensors do not apply to wide-FOV imaging sensors that are commonly used with UAS platforms: a single, fixed ρ value is especially erroneous because the variance in viewing geometry across a single image can yield dramatically different contributions of sky reflection both in magnitude and shape of the spectrum. The amount and spectral shape of skylight reflected by the water's surface into the sensor can range widely from regions below the sky's zenith toward regions at the horizon—for example, around midday with clear skies and typical marine haze, the spectrum of reflected skylight captured near the horizon is substantially whiter (spectrally even) than the blue dominated spectrum captured from near the sky's zenith (Gilerson et al. 2018).

While Mobley (1999) determines a ρ value of 0.028 for the viewing geometry of $\varphi = 135^\circ$ and $\theta = 40^\circ$, and this approach works well in low wind and clear sky conditions, in many cases this can lead to large errors in L_w . Other estimation methods suggest using near-infrared (NIR; $\lambda = \sim 700\text{--}800\text{ nm}$) radiances to subtract out surface-reflected light under the assumption that all NIR is reflected (Siegel et al. 2000), but given that the intensity of reflected skylight varies across the spectrum (Carrizo et al. 2019), subtracting all measured bands by any single fixed value does not accurately correct measurements and can introduce errors, especially when using band-ratio algorithms (e.g., O'Reilly and Werdell 2019). Methods using NIR measurements are especially unsuitable for estuarine or near-shore waters that contain high concentrations of particles which enhance backscattering in the NIR. Though it is worth noting bio-optical modeling techniques can estimate NIR in these regimes (Bailey et al. 2010) and spectral optimization approaches using estimates of inherent optical properties can help derive more accurate ρ values across the spectrum (Lee et al. 2010).

Wide-FOV imaging sensors can outperform point spectrometers in many oceanographic applications because image-type data can provide major insight into spatial properties and can be filtered to remove unwanted measurements, such as glint, foam or sargassum, that are often integrated into single-point spectrum measurements. Given that a fixed ρ value is especially

inappropriate for a wide-FOV imager, we propose a simple approach that includes a step using the skylight-blocked approach (SBA) to measure L_w (Tanaka et al. 2006; Lee et al. 2013, 2018), here called L_w^* , which is used as an empirical baseline for correcting L_t measurements. Given that L_w^* can be easily measured manually using the SBA during most operational scenarios employing UAS, we show that $L_t - L_w^*$ can solve for $\rho \times L_{sky}$ (the surface reflected skylight, commonly called L_{sr}) across all measured bands for a set of environmental conditions during a short flight ($< 1\text{ h}$)—with fewer assumptions than methods that model ρ and measure L_{sky} .

Study objectives

This study applies and compares multiple methods of deriving ocean color from UAS imagery to measure submesoscale variance in Chl a across the Gulf Stream front, with methods that can be generalized to study diverse oceanographic systems. We test the ideal viewing geometry for a low-altitude wide-FOV sensor (Fig. 3), we implement a method for accurate removal of sky reflectance from imagery from such a sensor (Eq. 2), we quantify uncertainty in R_{rs} measured by such a sensor, we propagate this uncertainty to Chl a estimates, and we demonstrate the implementation of this full workflow on a survey across the Gulf Stream front. We share this work to demonstrate and advance the use of UAS to study oceanographic systems at new spatial scales, and to provide new tools for a broad array of techniques that depend on above-water radiometry.

Materials and procedures

Unoccupied aircraft system

Ocean color imagery was captured using a Micasense Altum (<https://micasense.com/altum/>) multispectral camera payload gimbal-mounted on a Freefly Alta 6 commercial grade hexacopter (<https://freeflysystems.com/alta>). This UAS has a “motion boot” mode that allows for boot-up and calibration on a moving vessel—a scenario that prevents at-sea operation for many UAS. The Alta 6 has a high payload capacity (6.8 kg) and can be easily adapted to carry different payloads. The Micasense Altum payload has five optical sensors with filters centered at 475, 560, 668, 717, and 842 nm, respectively, with an approximate 20-nm full width at half maximum (FWHM) response. The FOV on this sensor is 48° (horizontal) $\times 36.8^\circ$ (vertical). This instrument also has a wideband thermal sensor centered at $11\text{ }\mu\text{m}$ that was not used in this study.

The Micasense Altum payload was mounted on a two-axis gimbal to stabilize its pitch and roll axes (Supporting Information Fig. S1); Yaw was determined by the UAS heading. The gimbal and sensor were controlled using a custom Arduino microcontroller that allowed for precise pitch angle and timing of image capture (code available at <https://github.com/marrs-lab/AltumGimbalCtrl>). With this approach, viewing geometry is held consistent to within 1° —roll is always zero,

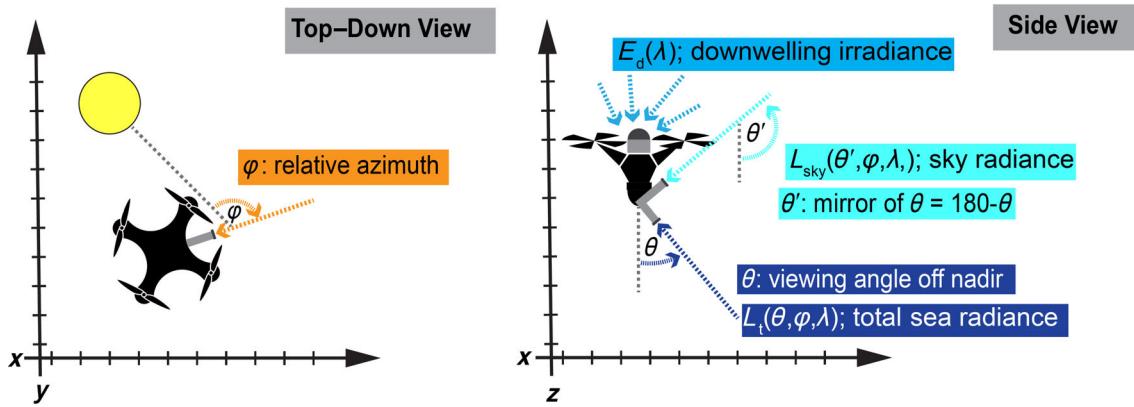


Fig. 3. Three measurements are needed to derive remote-sensing reflectance (R_{rs}) from above-water sensor measurements: total sea radiance (L_t), sky radiance (L_{sky}), and downwelling irradiance (E_d). The UAS sensor viewing geometry relative to the sun and the water's surface are optimized to collect these measurements with the minimal amount of glare and reflected skylight (Mobley 1999; IOCCG 2019).

and pitch and yaw are variable depending on the environmental context. For all ocean color surveys on the Gulf Stream front, we held pitch at 25° off-nadir and yaw at 135° azimuthally off the sun. This UAS did not send real-time positional telemetry, but the autopilot did record a detailed internal flight log that was used for the analysis. To provide a live map as a navigational aid, we used a redundant Pixhawk autopilot, GPS module (<https://ardupilot.org/copter/docs/common-pixhawk-overview.html>) and long-range RFD900 (<http://rfdesign.com.au/products/rfd900-modem/>) telemetry radio that we installed on the aircraft. Map data displayed the real-time aircraft and ship positions on a Samsung Galaxy Tab7 tablet. E_d measurements were taken before and after each flight with the Micasense Altum using the manufacturer's reflectance panel.

We collected images at an altitude of 100 m, yielding footprints of approximately 80 m × 60 m and pixel sizes of ~5 cm × 5 cm though varying across the image depending on sensor viewing angle. As a first step after launch and concurrent with the SBA measurements, images are taken to provide the data for $L_t - L_w^*$. Multiple images are taken to remove waves and glint by averaging across images and smoothing spatially, with a minimum of 5–10 images needed for removal. Environmental conditions were stable throughout the duration of flights.

Gulf Stream front study area

The region off North Carolina's coast from Cape Lookout to Cape Hatteras is a highly dynamic confluence of the warm, rapid, northbound Gulf Stream, the cooler, relatively slow circulation of the Mid Atlantic Bight, and shelf water from the South Atlantic Bight entrained into the Gulf Stream. This area spans productive waters at the continental shelf break and oligotrophic deep open ocean, with water depth dropping sharply from 100 to 1000 m within a lateral span of just a few dozen kilometers (Fig. 4).

In situ Chl *a* measurement

In situ Chl *a* was measured by filtering 100 mL of seawater onto 0.7 µm GF/F filters, in 100% methanol and reading fluorescence on a chlorophyll calibrated Turner 10 AU fluorometer equipped with Welschmeyer filters (Johnson et al. 2010).

Satellite data

We used multiple ocean-observing satellites to provide context and comparison with our UAS measurements. We used R_{rs} and Chl *a* products from OLCI, MODIS, and VIIRS. For R_{rs} comparisons all imagery was acquired was within 75 min of UAS imagery acquisition. To compare between satellite and UAS data, UAS images were averaged to a point measurement and these points were then compared to the pixels they fall within. For spatial context, we used sea surface temperature products from the Geostationary Operational Environmental Satellite 16 (GOES-16). MODIS and VIIRS Chl *a* products were Level 2, processed using the standard NASA algorithm and obtained from <https://oceancolor.gsfc.nasa.gov/cgi/browse.pl>, OLCI Chl *a* was Level 2 and processing using the OC4ME algorithm and obtained from <https://coda.eumetsat.int>, and contextual GOES-16 hourly SST was acquired from <https://cwcgom.aoml.noaa.gov/erddap/griddap/goes16SSThourly.html>.

In situ radiometry

We collected L_t , L_{sky} , and E_d every 10 s using a set of common hyperspectral radiometers (TriOS RAMSES, Rastede DE) that were mounted on the starboard bow of the ship. These sensors were set for auto integration with a time range between 10 ms and 8 s. These three measurements were used to calculate L_w (Eq. 1) and R_{rs} (Eq. 2), using a ρ of 0.028 (Mobley 1999). The TriOS radiometers collect 256 wavelength bands at 3.3-nm intervals within the 320–950 nm range. All in situ hyperspectral measurements were interpolated to 1-nm intervals. For

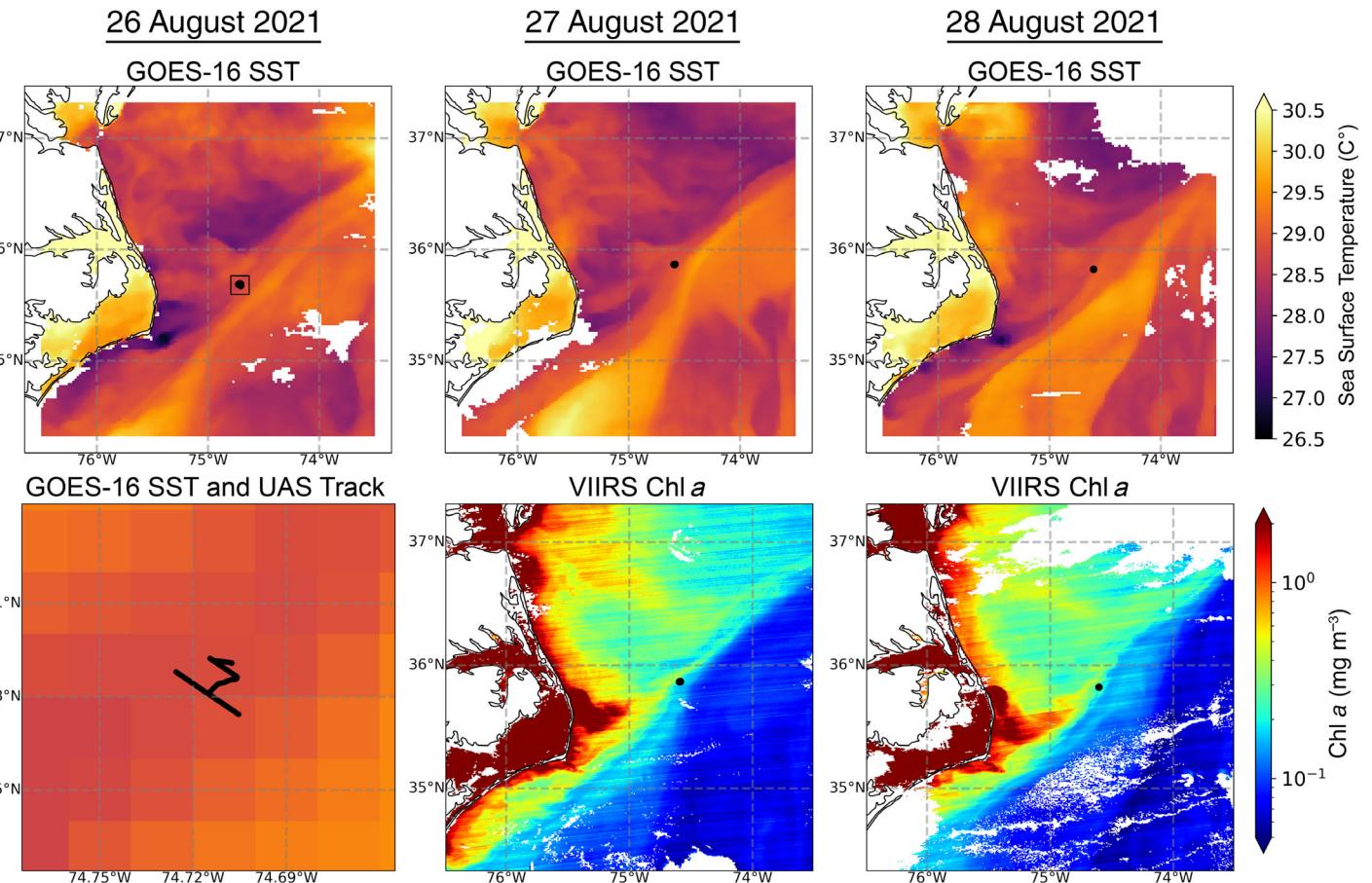


Fig. 4. An overview of the Gulf Stream front study area from ocean-observing satellites. NOAA's GOES-16 platform describes the sea surface temperature of the region in 1-km resolution during the primary days of this study (top row), with the UAS survey location overlaid (black). A 6-km zoom of the GOES-16 imagery illustrates the scale of the UAS survey (black) against 1-km pixels (bottom row, left); and VIIRS products describe Chl *a* in 1-km resolution during the latter 2 d (bottom row, center, and right). Note that VIIRS Chl *a* is shown with a logarithmic color stretch given the extreme range in this area.

comparisons between TriOS and UAS, all UAS images were averaged to points and we subset TriOS data to measurements collected within 15 min and 100 m of the UAS samples. When directly comparing R_{rs} , 10 nm around the Altum band centers were averaged in the TriOS data to be more comparable.

Chl *a* algorithm

We used the Color Index algorithm (Hu et al. 2012) for all UAS Chl *a* retrievals; this algorithm is used in the NASA Chl *a* retrieval for values below 0.15 mg m^{-3} and is appropriate for the relatively oligotrophic environment of the Gulf Stream (Chl *a* predominantly $< 0.3 \text{ mg m}^{-3}$). The Color Index algorithm is a simple index relating the reflectance at 547 nm to a straight line from the reflectance at 443 nm to the reflectance at 670 nm. This approach, a conceptually simple change from NASA's OCx blue/green ratio (O'Reilly and Werdell 2019), is less sensitive to instrument noise and imperfect atmospheric correction because it includes the red band, which has negligible reflectance in oligotrophic waters (Hu et al. 2012). We did not fine-tune this algorithm for the slightly different band

centers and spectral responses of our sensor, and therefore expect some minor bias in derived measurements.

Imagery preprocessing

Data collected by the MicaSense sensors were radiometrically calibrated to convert digital number (DN) values to total spectral radiances (L_t) using a Python workflow provided by MicaSense (<https://github.com/micasense/imageprocessing>), based on the following equation:

$$L_t = V(x, y) \times a_1/g \times (DN - DN_{BL}) / (t_e + a_2y - a_3t_e y), \quad (3)$$

where the dark pixel value (DN_{BL}), sensor gain (g), exposure settings (t_e), and lens vignette effects ($V(x, y)$) are combined to derive radiance of each pixel based on the measured digital number (DN) with radiometric calibration coefficients (a_1, a_2, a_3) and the column and row number of the pixel (x and y , respectively). Lens distortion effects, such as band-to-band image alignment, were removed from image captures by an unwarping procedure. We converted L_t to R_{rs} using a

calibrated reflectance panel with known reflectance (R_{plaque}) that allows accurate measurement of E_d :

$$E_d(\lambda) = L(\lambda)^* \pi / R_{\text{plaque}}(\lambda). \quad (4)$$

We then normalized other measurements of L to R_{rs} using E_d (Eq. 2).

Skylight-blocked approach

We measured L_w^* with the SBA by mounting a Micasense RedEdge multispectral camera on a 1.5-m long, 10-cm diameter black tube (Supporting Information Fig. S1). The optical characteristics of the Micasense RedEdge are the same as the Altum, except that it does not feature a thermal sensor. We captured images with this sensor while placing the open end of the tube into the water and holding the opening just below the surface to bypass surface reflections. We then measured pixels from the section of the image that represented the end of the tube, since most of each image consisted of the tube itself. We averaged the spectra from these pixels to measure L_w^* . We measured R_{plaque} with the RedEdge sensor at the same time as the Altum sensor, using the same plaque.

Removal of surface reflection and sun glint

We implemented a sequence of processing steps to accurately derive L_w and resolve spatial patterns across an individual image, which otherwise could be biased and obscured by reflected skylight, due to the dependency of reflected skylight on viewing geometry (Fig. 5). Our approach uses L_w^* from our SBA measurement and L_t from images collected by the UAS at the same time and location where L_w^* was measured. Using the assumption that L_w^* is the correct L_w for each pixel in the co-located L_t image set (ignoring the viewing angle dependence of L_w), and the assumption that the sky conditions did not change substantially during the flight, we averaged a set of at least 10–20 L_t images over the same location, smoothed this result using a 45×45 -pixel gaussian window, and then subtracted the L_w^* value from each pixel (Eq. 5) to get the value of $\rho \times L_{\text{sky}}$ (i.e., L_{sr} , Fig. 5).

$$\rho \times L_{\text{sky}} = L_t - L_w^*. \quad (5)$$

Having calculated this correction factor at each pixel, for the images throughout the flight we subtract the correction matrix from the L_t image to obtain L_w for each pixel.

$$L_w = L_t - \rho \times L_{\text{sky}}. \quad (6)$$

We applied a threshold of $2 \times$ the standard deviation of the image pixels greater than the median L_w values to mask out extreme values representing sun glint. We selected this threshold empirically and evaluated its success visually, finding that it appears to remove nearly all sun glint.

Comparison of viewing geometries

We compared different viewing geometries to estimate a suitable compromise between near-nadir angles that receive high amounts of sun glint and further off-nadir angles that receive high amounts of reflected skylight off the surface and oblique wave facets. We tested the effects of different geometries on measured reflectance by (1) sampling different regions of image data from a wide-FOV sensor (Fig. 6), and (2) permuting our viewing angles through different gimbal pitches, centering the sensor view vertically from 0° to 40° off-nadir, and different aircraft yaws, centering the sensor view horizontally from 90° to 180° off-azimuth from the sun (Fig. 7).

Comparison of Chl a measurements

We compared our measurements of Chl a from UAS images to both in situ measurements and measurements from satellite-imagery (Fig. 11). To compare image-based measurements from our wide-FOV sensor to single-value Chl a measurements, we used the median value of each band in the image and plotted this value at the location of the center pixel in each UAS image. Given the rapid flow of the Gulf Stream in situ measurements collected within 15-min and 1.5 km of a UAS image were used. When comparing measurements from satellite imagery, we selected satellite imagery captured at the closest time available to our survey, and we used the values of the pixels that overlapped our measurement points. This did lead to time differences of up to 4 h (max 4 h for MODIS, max 3.5 h for VIIRS, max 2 h for OLCI) though on average the difference in time was around 2 h.

Uncertainty analysis

We measured the uncertainty of these single-point UAS measurements by flying a short variance estimating mission to obtain replicate measurements over the same parcel of water. During a period of consistent environmental conditions with low wind ($< 2 \text{ m s}^{-1}$), no cloud cover, and 2 km away from the front we placed the UAS in a stationary flight for 2 min, while imaging the same footprint of water every 1.5 s; resulting single-point UAS measurements were assumed to describe uncertainty in the measurement approach and not changes in the measured R_{rs} values, and we report this uncertainty as the standard deviation of those replicate measurements. We then propagated this uncertainty in R_{rs} to retrieved Chl a using linear error propagation (Taylor 1996). We compared this uncertainty to TriOS radiometric measurements sampled while the R/V floated passively with a parcel of water. We also conducted duplicate survey flights along the same flight paths, 10–15 min apart, for five flights to verify approximate measurement uncertainty across repeat surveys—expecting some additional variability due to the movement of water between duplicates. While these two approaches to uncertainty estimation are relatively specific for the sky and sea conditions during collection, they provide a general magnitude of the full workflow variability in moderately good

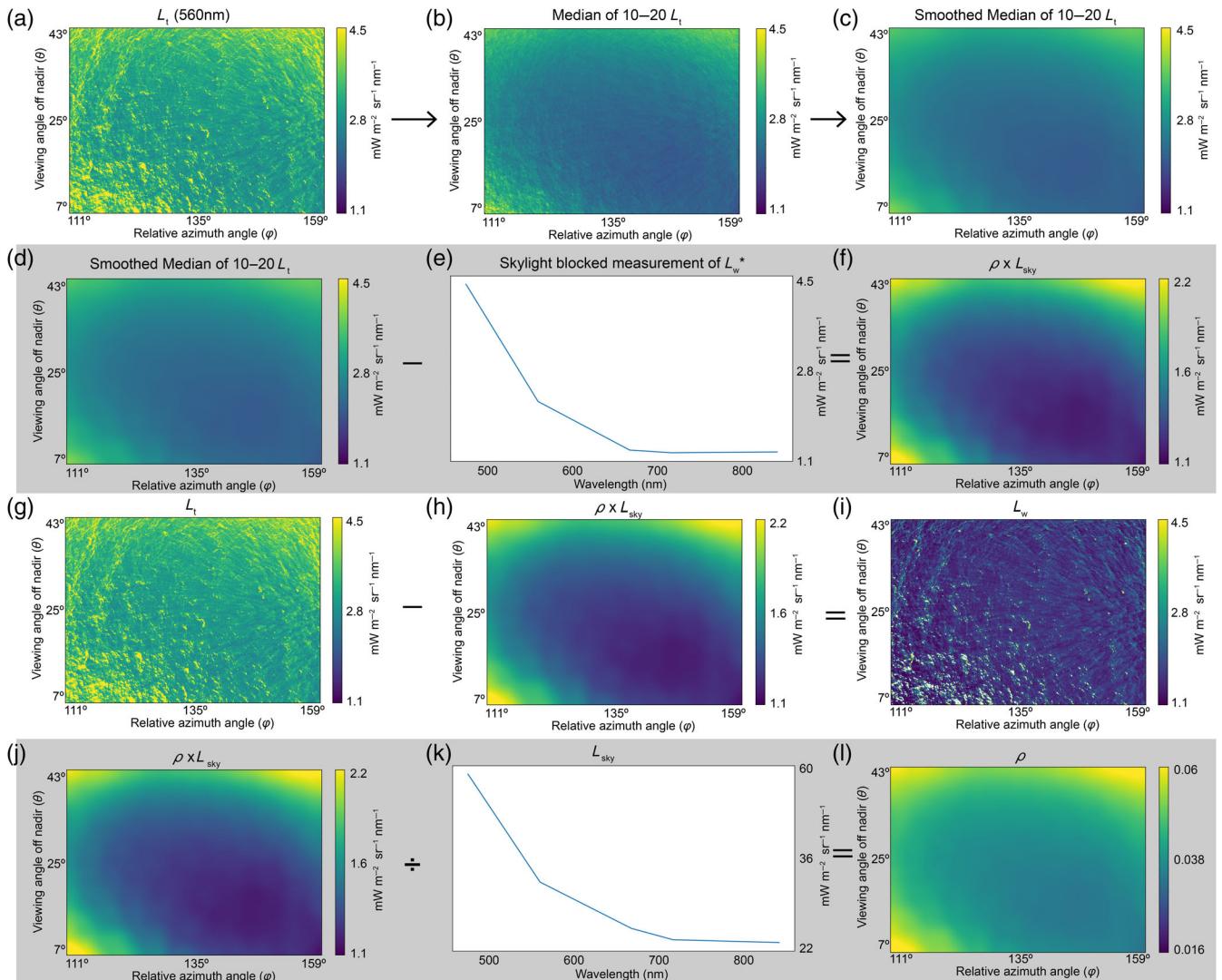


Fig. 5. The sequence of data processing steps needed to derive $\rho \times L_{\text{sky}}$ (i.e., L_{sr}) and L_w from a wide-FOV image sensor, here using an example image of the green band, L_t (560 nm), taken by a UAS at 100 m altitude with a viewing geometry of $\varphi = 135^\circ$ and $\theta = 25^\circ$, yielding a spatial footprint of $\sim 80 \times 60$ m. All variables here (L_t , L_{sky} , L_w , and ρ) are spectral, and in this figure refer to the green band at 560 nm. Given an unprocessed image (**A**), waves are removed by taking median pixel values from 10 to 20 images of the same area (**B**), which are then smoothed using a 45×45 pixel window (**C**). From this smoothed median L_t (**C,D**), we subtract the SBA measurement of L_w^* (**E**), according to Eq. 5, yielding the surface reflected light ($\rho \times L_{\text{sky}}$) across all image locations (**F**). From the original images (e.g., **A, G**), we then subtract these $\rho \times L_{\text{sky}}$ values (**F,H,J**), according to Eq. 6, yielding estimates of L_w across each image (**I**). While not a part of the workflow, the derived $\rho \times L_{\text{sky}}$ image (**F,H,J**) can be divided by a measurement of L_{sky} (**K**) to show image-wide estimates of ρ (**L**) which range well beyond the commonly used value of 0.028.

conditions where there is low wind < 5 m s⁻¹, minor cloud cover or fully cloud covered, and sun $> 20^\circ$ off the zenith.

Assessment

Our methodology was able to provide repeatable and accurate measurements of ocean color with a reasonable uncertainty that could be implemented by a wide variety of aquatic scientists. Using this methodology, we surveyed ocean color across the Gulf Stream at fine-scale; these surveys described variability and frontal features occurring

at spatial scales too large to be captured by in situ measurements and too fine to be resolved in satellite imagery. Measurements from in situ sampling and satellite imagery validated our UAS measurements and correction process.

Influence of viewing geometry on measurements

As viewing angle increased off-nadir within an image, reflectance increased across all image regions, and this rate of increase was spectrally variable (Fig. 6). Blue radiance increased proportionally more than green radiance at higher viewing

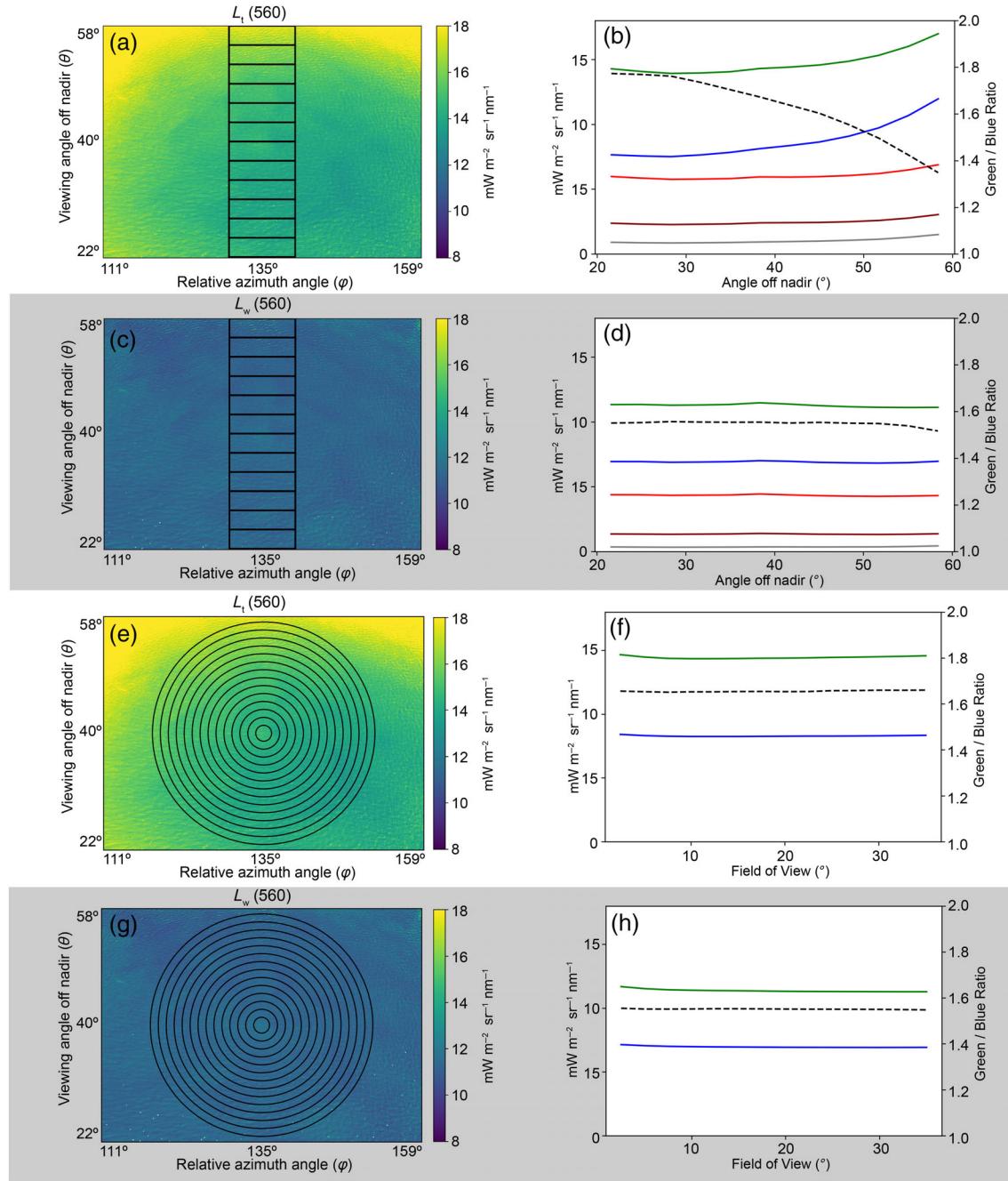


Fig. 6. Effects of viewing angle and FOV on measurements of L_t compared to L_w . Regional subsamples from a single image, representing different viewing angles (**A–D**) and different fields of view (**E–H**), illustrate how these sampling choices affect reflectance. As viewing angle increases off-nadir, L_t increases across the spectrum at spectrally variable rates (**B**), contributing angle-dependent biases to the green : blue ratio (dashed line in **B**). This change is minimal in the L_w measurements (**D**) until the very highest angle off nadir $\sim 60^\circ$ where either obliquely viewing wave facets or actual changes in water content appear to drive a minor change in the ratio. As FOV increases radially, the radially averaged L_t (**F**) shows only marginal increases that appear spectrally consistent which is similar to L_w (**H**).

angles. If not accurately corrected for in the calculation of L_w , this could negatively bias Chl a estimates from a band-ratio algorithm using measurements at these angles. This effect of viewing angle manifests within single images, where a wide-FOV sensor captures surface reflections from a range of viewing

angles (Fig. 6), and in the directed viewing geometry of the sensor on its gimbal mount (Fig. 7).

We confirmed that a viewing geometry with φ of 135° – 180° relative to the sun and θ of 10 – 30° off-nadir is ideal, but a higher angle off-nadir may be appropriate at times to reduce

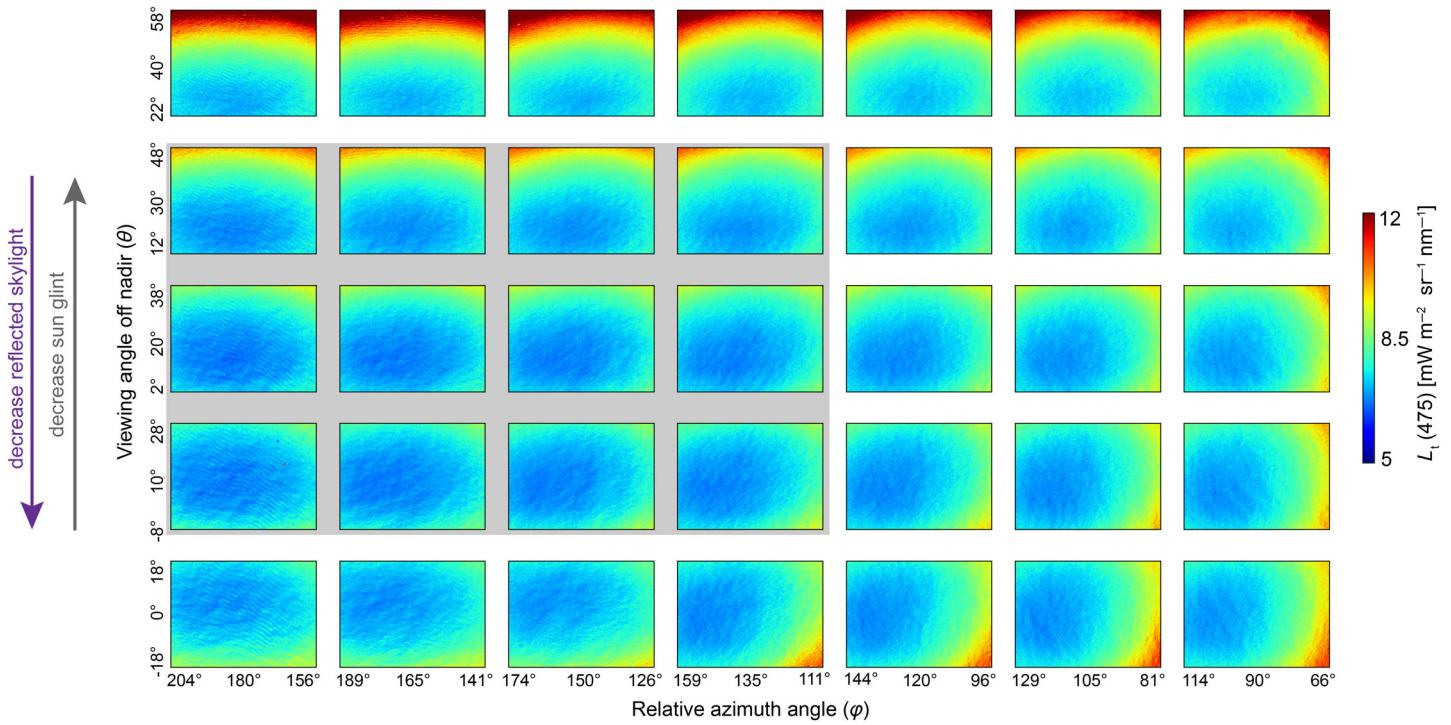


Fig. 7. An array of different combinations of viewing angle off nadir (θ) and azimuth angle relative to the sun (ϕ), illustrating how viewing geometry affects the amounts of glint and reflected skylight that are captured by an airborne wide-FOV sensor. Survey conditions consisted of 100% clear sky, minor haze on the horizon and the sun at 105° azimuth and 35° elevation. Optimal viewing geometry was achieved at yaw angles of $135\text{--}180^\circ$ with pitch angles of $10\text{--}30^\circ$ (gray background box), yielding the smallest and most uniform effects of reflected skylight. Images were collected by a Micasense Altum and show $L_t(475)$.

glint from a high solar elevation, high wind speeds, or to avoid self-shading, if this is a concern. Conversely, under conditions with minimal sun glint—low solar elevation, low wind speed, fully overcast—a θ closer to nadir can minimize capture of reflected skylight.

Reflected skylight removal method

Our method for removing reflected skylight (Fig. 5) appears to successfully align our measured spectra values with expected values, yielding R_{rs} (NIR) near 0 for our pelagic measurements (Supporting Information Fig. S2) and matching closely with our validation measurements (Fig. 10). The R_{rs} derived from our reflected skylight removal and workflow is slightly lower than TriOS and VIIRS measurements, possibly from a more complete removal of reflected skylight. In this comparison (Fig. 10), it should be noted for TriOS data the reflected skylight corrected used the Mobley (1999) approach and a NIR subtraction, so the NIR bands are forced to be effectively zero and thus UAS data are 2.5 times higher in the NIR. Our method for removing reflected skylight also reveals features that occur in all parts of an image (Figs. 8, 9) and yields an image-wide distribution of measured values that more closely approximate the Gaussian distribution that is expected for small random wave impacts on the data (Supporting Information Fig. S3).

Comparison to in situ and satellite measurements

Our corrected UAS measurements of R_{rs} closely matched measurements from TriOS and VIIRS satellite data (Fig. 10). Between these techniques, VIIRS yielded greater differences from UAS measurements, which we attribute to the difference of spatial scale that each sample described. Derived Chl a measurements also closely matched to measurements from VIIRS, MODIS, and OLCI satellite data, but with some offset in the retrieved values and deviations from a 1 : 1 linear relationship (Fig. 11). OLCI data had the best linear relationship with UAS measurements, as judged by R^2 , and notably this product featured the highest spatial resolution of the three considered (OLCI 300 m, VIIRS 750 m, and MODIS 1000 m) and had the closest match in time to the UAS flights (~ 1 h difference on average as opposed to ~ 2), though the UAS did have an underestimation compared to OLCI. VIIRS data compared to UAS had the lowest mean absolute percentage error (MAPE).

Uncertainty in measurements and Chl a

The uncertainty of our UAS R_{rs} measurements from this method was 1.5% (475 nm), 0.97% (560 nm), 6.7% (668 nm), 8.4% (717 nm), and 23% (842 nm) of the total R_{rs} (Supporting Information Fig. S5). When propagated through the Chl a algorithm (Hu et al. 2012), this yields $0.17 \pm 0.0034 \text{ mg m}^{-3}$

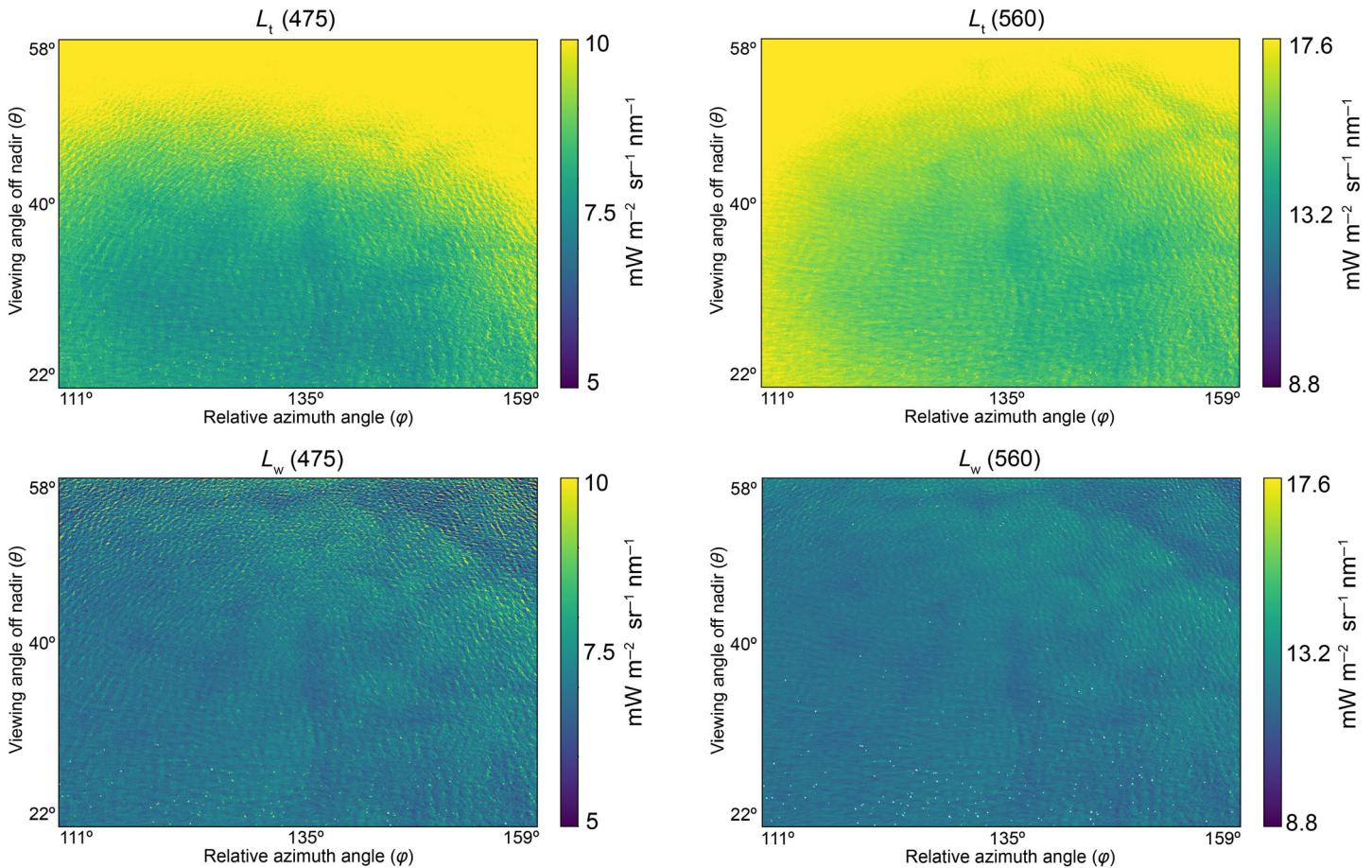


Fig. 8. Example measurements of L_t (top row) and L_w (bottom row) derived from our method for removing reflected skylight using an image of an estuarine environment. The image was taken from a viewing geometry centered at 40° off-nadir with 100% clear sky on 04 March 2021.

Chl *a*, or an uncertainty of 1.9% of the mean. R_{rs} measurements from UAS had substantially lower uncertainty than measurements from our ship-based TriOS (Supporting Information Fig. S5) which should be interpreted largely as a product of the UAS gimbal holding a fixed viewing geometry rather than a difference in sensor capabilities or our skylight reflection approach. The equivalent uncertainty in blue (475 nm), green (560 nm), and red (668 nm) bands from the TriOS was 5%, 6%, and 10% of the mean, respectively. When propagated to Chl *a*, this led to an uncertainty at 20% of the mean.

Identical duplicate flight paths ($n = 5$) yielded uncertainty estimates close to the value calculated from stationary measurements (± 0.000085 or 1.5% difference from the mean), though differences between duplicates ranged up to 3%, twice our estimated uncertainty (Supporting Information Fig. S6).

To complement our uncertainty calculations with a more broadly used metric, we calculated the signal-to-noise ratio (SNR) of a single image via the mean divided by the standard deviation. This was done over a consistent area of ocean water

and for 55 images (the same data reported in Supporting Information Fig. S5) the averaged spectral SNR was 1.25 (475 nm), 0.57 (560 nm), 0.28 (668 nm), 0.28 (717 nm), and 0.25 (842 nm). Once each image was averaged to a single point and SNR was calculated from the 55 point samples (via the mean of the points over their standard deviation) SNR = 111.17 (475 nm), 35.07 (560 nm), 23.22 (668 nm), 18.08 (717 nm), 17.80 (842 nm). In another test where a large calibration plaque took up a 175 × 175 pixel component of the image during flight, the pixelwise SNR of that plaque was > 60 per band for 40 images, indicating that while the sensor SNR is not extremely high, most of the noise is from the environment and that this environmental noise is substantial. This also demonstrates that averaging each image to a single point substantially increases the SNR as expected.

Gulf Stream frontal mapping

UAS surveys of Chl *a* at the Gulf Stream front off North Carolina clearly describe the frontal boundary, smaller areas of Chl *a* enhancement, and continuous variability extending

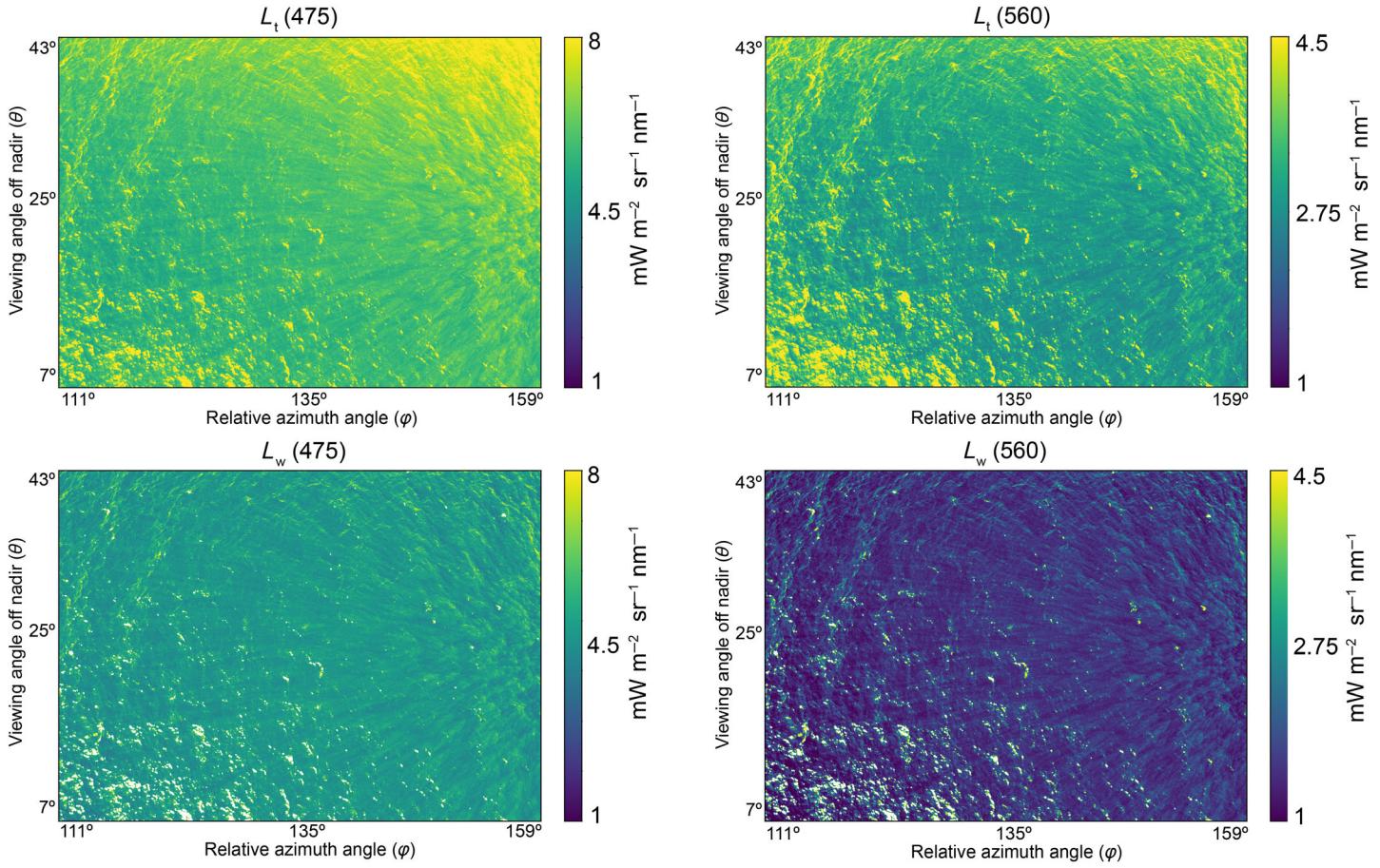


Fig. 9. Example measurements of L_t (top row) and L_w (bottom row) derived from our method for removing reflected skylight using an image of a pelagic environment. The image was taken from a viewing geometry centered at 25° off-nadir with 100% clear sky on 27 August 2021.

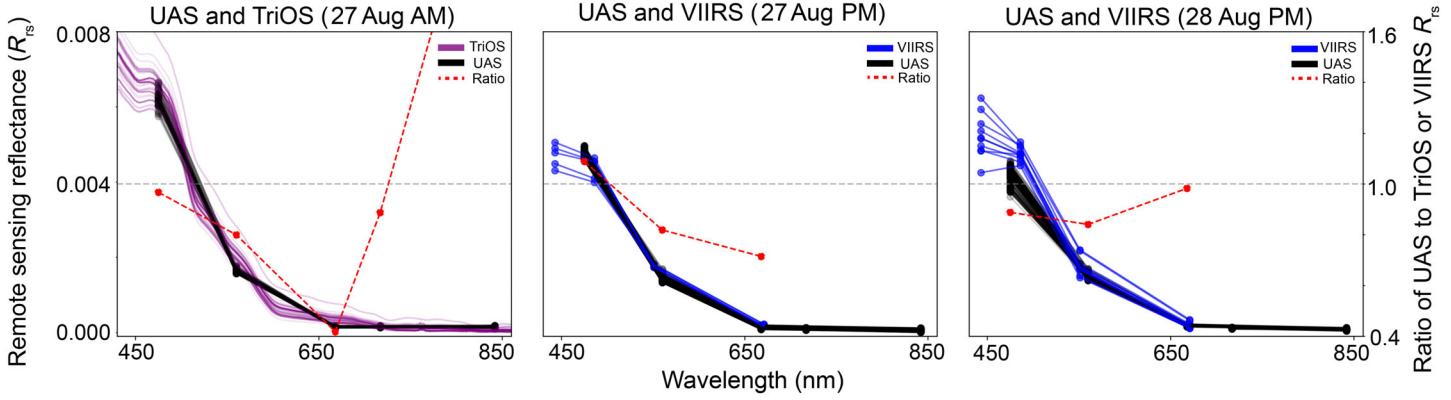


Fig. 10. Comparisons between R_{rs} measurements from our UAS survey and TriOS samples (left, purple) and images from VIIRS (center and right, blue). The ratio of UAS to TriOS and VIIRS R_{rs} is shown in the red dashed line with the gray dashed line representing unity. TriOS data were collected within 15 min and 100 m of the UAS data. VIIRS data represent pixels that overlapped with our UAS data. VIIRS images were captured within 75 min of the UAS data.

offshore (Fig. 12). Structuring of sargasso and breaking waves are visible in RGB-composite images collected at the front (Supporting Information Fig. S7). Substantial variability is

present among Chl a measurements from UAS imagery, occurring within the spatial scale of a single pixel of VIIRS imagery (Fig. 13).

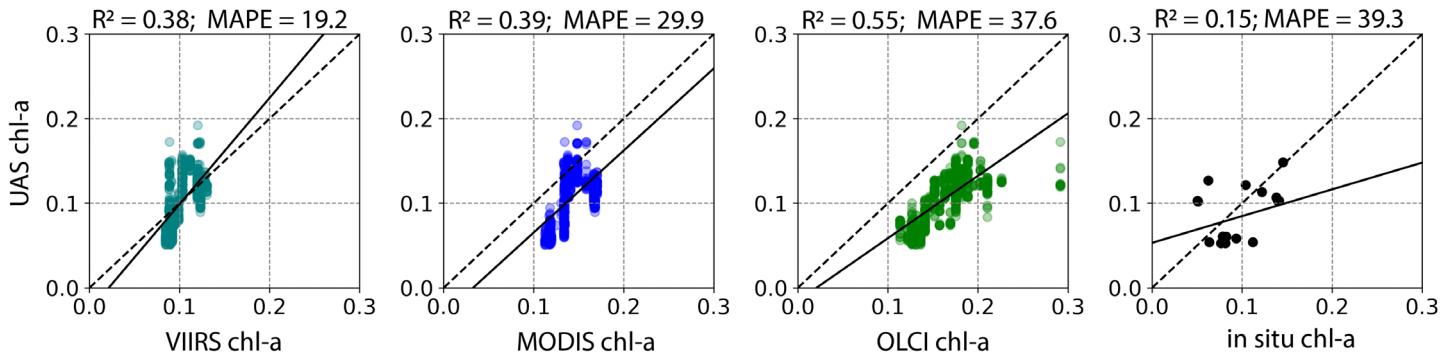


Fig. 11. Match-ups between Chl a measurements derived from UAS imagery ($n = 934$ images), VIIRS imagery ($n = 20$ pixels), MODIS imagery ($n = 24$ pixels), OLCI imagery ($n = 19$ pixels) and in situ (extracted and read on fluorometer) data ($n = 20$ samples). Data show roughly linear relationships though with some bias and spread as might be expected from different Chl a algorithms and slight differences in spatial scale and time mismatch between methods. Coefficient of determination (R^2) and MAPE are reported above each plot. Map of the UAS and in situ data overlaid on OLCI available in Supporting Information Fig. S4.

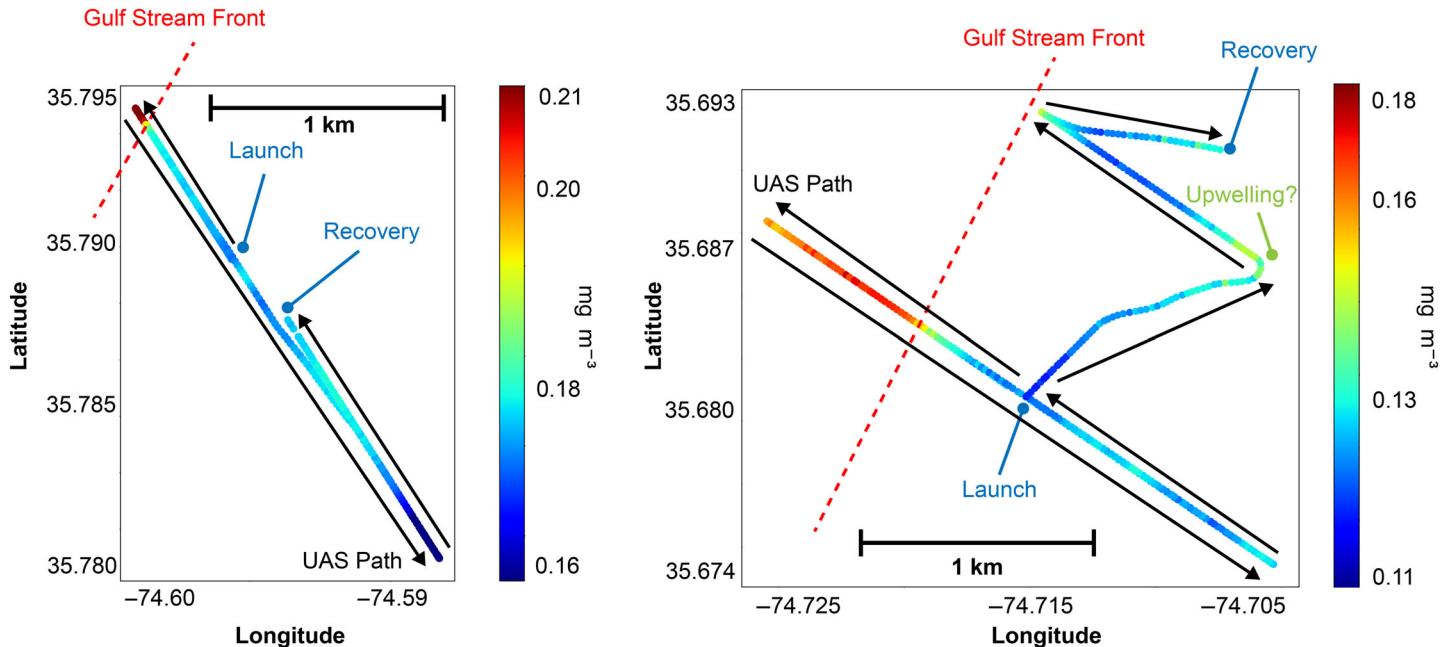


Fig. 12. Two UAS surveys, on 27 August 2021 and 28 August 2021, capturing Chl a gradients at the Gulf Stream front (red dashed line) and some smaller regions of higher productivity.

Discussion

The results of this study reveal that UAS-based sampling of ocean color is a robust and maturing technique for observing fine-scale patterns in a rapidly moving environment. Amid ongoing advancements to the technology, multi-UAS fleets and integrated sorties may capture even larger domains of observation in closer synchrony—though such methods would more critically require precise calibration and management of uncertainty.

Proper viewing geometry critically affects ocean color measurements from a UAS-borne sensor, and there is a tradeoff

between solar glint and reflected skylight. As its pitch angle is increased off-nadir, a sensor captures less solar glint but more reflected skylight (Figs. 6, 7). During 100% cloud-covered conditions or sun elevations $< 30^\circ$ above the horizon, glint is minor, and a sensor can effectively pitch to nadir and minimize reflected skylight. In typical midday conditions with a sun elevation angle $> 45^\circ$, glint disrupts measurements, and thus 25° off-nadir— $25^\circ \pm 20^\circ$ with a sensor with a vertical FOV of $\sim 40^\circ$ —approaches an ideal middle-ground, capturing the majority of the image from a zone

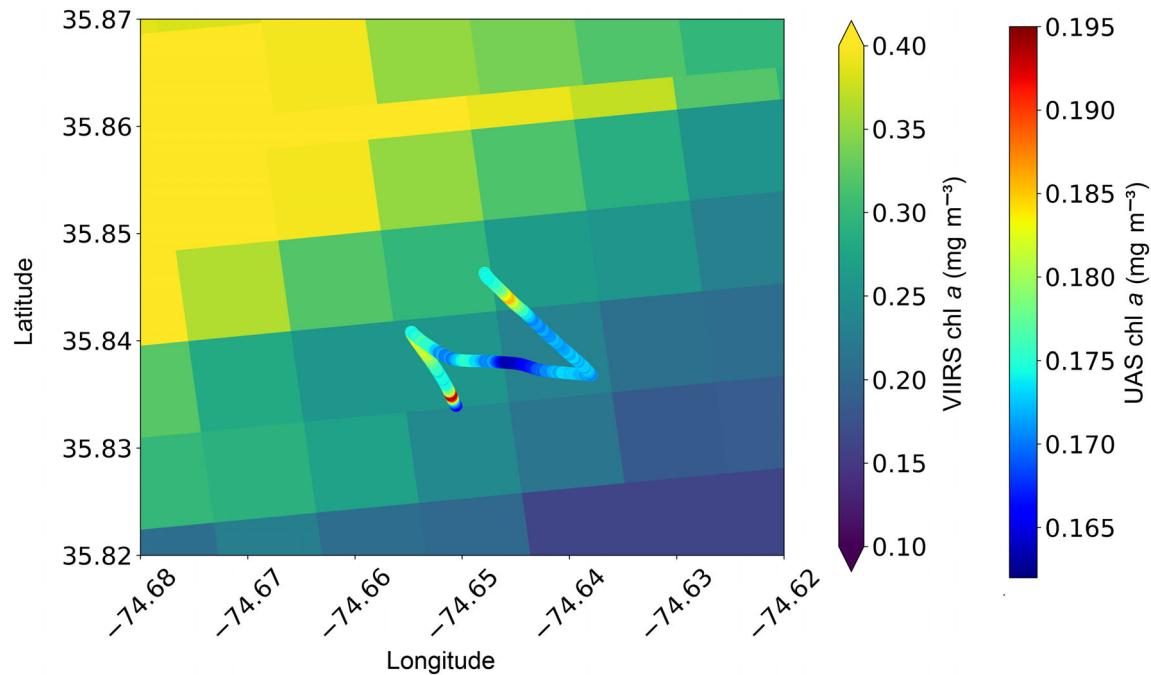


Fig. 13. Measured Chl α values from UAS imagery (red–blue color ramp) overlaid on Chl α values from VIIRS imagery (yellow–violet color ramp), describing Chl α around the Gulf Stream front off North Carolina. Separate color ramps illustrate the different ranges and resolved variability of Chl α measured by each platform.

with reduced glint and low skylight reflection (Fig. 7). With a wide-FOV sensor, a pitch angle centered above 25° off-nadir drastically increases reflected skylight in the top of the image and is so oblique that wave facets become much more visible. This understanding matches well Mobley's findings (1999) that captured sky reflection increases substantially as sensor viewing angle increases from 60° off-nadir and upward. This can be seen in our back-calculated ρ values (Fig. 5L) where ρ varies from just below the modeled value of 0.028 to a factor of 2–3 higher near 60° off-nadir. When examining different yaw angles, we did not find a substantial difference between 135° and 180° off-azimuth from the sun (Fig. 7). In a clear pelagic environment, the sun beam pattern within the water radiates outward from 180° opposite the sun and can disrupt measurements (Fig. 9). This noise (e.g., sun beams within the water column) challenges observations of within image patterns and may limit some of the benefits of an imager vs a point radiometer in pelagic environments. We detected no noticeable UAS shadow in imagery, even at 180° off-azimuth from the sun. These findings agree with recent work using wide-FOV hyperspectral imagers, which showed that L_t and L_{sky} depend upon viewing geometry, wind speed, and FOV (Carrizo et al. 2019).

Should a workflow use consistent viewing geometry out of convenience or necessity, we recommend ~25° off-nadir and 180° off-azimuth from the sun. Our surveys employed a 135° azimuthal angle to avoid a possible anhelion (sundog opposite the sun) in the imagery. In estuarine environments or on fully

overcast days, this is an unlikely concern, and 180° off-azimuth is likely preferable to reduce glint and variation across the image. We suggest that UAS without a gimbal can still collect robust ocean color measurements if the sensor is mounted at a fixed pitch angle of ~25° off-nadir, with a fixed heading to maintain the yaw angle. The absence of a gimbal introduces more variance into the sensor's viewing geometry, being tethered to the aircraft's pitch and roll movements amid varying flight speeds and wind; this would contribute more uncertainty to measurements, but such a fixed optimal viewing geometry still reduces glint and reflected skylight relative to a sensor orientation at nadir. In the simplest case where a sensor is integrated onto a UAS in a fixed nadir orientation, we recommend avoiding survey times near solar noon to reduce specular glint and otherwise following the procedures of this study. We believe this constitutes a useful, practical solution for many UAS practitioners to make reasonable ocean color measurements.

Our procedure for removing skylight reflection (Fig. 5) successfully allowed us to visualize and measure features within an image that would otherwise be obscured by reflected skylight (Fig. 8), yielding R_{rs} values near those of conventional techniques (TriOS) and satellite sensors (Figs. 10, 11). The one comparison that is poorest (28 August VIIRS) is well matched spatially but was directly on the front and in the time between the VIIRS acquisition and UAS survey the front was moving southeast, thus the VIIRS image (taken earlier) was more pelagic and thus bluer. It is also worth noting that within our Chl α comparison, given the use of the Color Index algorithm

our Chl a retrievals may be saturating at higher values ($> 0.15 \text{ mg m}^{-3}$) and thus the relationship may not be completely linear (Fig. 11). Our SBA-based skylight reflection removal approach reduced the skew of the distribution of measured R_{rs} within an image to a roughly normal distribution (Supporting Information Fig. S3). The remaining variability matches what would be expected due to wave facets, which are not corrected for in this work. Accurate R_{rs} from a wide-FOV sensor, combined with latitude, longitude, yaw, pitch, and roll of the sensor, may soon enable ultra-high-resolution mosaics of the water surface that are not typically possible from common UAS structure-from-motion methods (Westoby et al. 2012) due to the absence of persistent key points in aquatic imagery.

Our UAS measurements entail considerable uncertainty (Supporting Information Figs. S5, S6) from various sources of noise in the environment and sensor. Uncertainty from the sensor is likely higher due to the small signal coming from the water—which likely occurred near the bottom of our sensor’s sensitivity range given that it was designed for terrestrial targets with one to two orders of magnitude higher radiances. We successfully achieved repeat R_{rs} in the blue and green spectral region $< 5\%$ of the measurement, in line with goals from the IOCCG (2019). Uncertainty in Chl a due to R_{rs} was calculated as 1.9%, and while the actual Chl a uncertainty is certainly higher, this is an estimate of the uncertainty contribution from the ocean color measurement itself. In comparison to the ship-based TriOS, which was moving through a range of viewing geometries due to waves, UAS uncertainty is substantially lower, primarily demonstrating the benefit of a gimbal and fixed viewing geometry rather than differences in instruments. Uncertainty constitutes critical information that must accompany optical measurements: many valuable data products and proxies can be derived from optical observations but require proper calculation and propagation of uncertainty through retrieval algorithms to enable useful oceanographic interpretation.

Across the Gulf Stream front there is a major step-change in temperature, Chl a and other oceanographic properties that are shown by satellite and in situ measurements, occurring over the span of a few dozen meters, which may not be well described by kilometer-scale satellite imagery or model products. Fine-scale sargassum patterns may provide well-structured habitat for zooplankton and larger fauna, and at times appear to be structured by Langmuir cells (personal observation). The variability that we describe within a 1-km pixel footprint (Fig. 12) has been discussed previously (Moses et al. 2016), but these findings add to the growing body of work suggesting important submesoscale dynamics structure the biogeochemistry around frontal features.

Assumptions

Our procedure for removing reflected skylight entails two major assumptions: (1) that L_w^* is representative of the

smoothed average of the stationary image series that is taken nearby and (2) that sky and wind conditions stay the same during the length of the flight. The first assumption may not hold if the initial image set has persistent features that remain after averaging and smoothing, or if the image set was captured too far from the site where its estimate is being applied. The second assumption could easily be violated by a rapid change in weather, atmospheric conditions, or sea state. However, many UAS surveys occur over limited spans of time and space, so this assumption should often hold for such applications. Our approach for removing sun glint assumes that glint can be resolved by the sensor and does not contribute radiance to other pixels; this appears to be the case in our study, which resolved centimeter-scale resolutions.

Our empirical estimation of uncertainty likely represents a baseline. We assumed that the standard deviation of R_{rs} is a proper representation of uncertainty, and that the short period that was used to calculate standard deviation is representative of other periods of data collection. In this case, our estimation may be reasonable for similar instruments, sky conditions, water conditions, and sea state, but is not applicable to other water bodies, wind speeds, sky conditions, and instruments.

We expected that a wide-FOV image sensor would provide advantages over a point radiometer for detecting and describing spatial patterns in ocean color at the scale of a few meters in a pelagic environment. Our findings and previous work support this for estuarine and coastal environments (Windle and Silsbe 2021), but our study suggests that in pelagic environments at this spatial scale the high-frequency signal from sun beams within the water and surface gravity waves can overwhelm some of the variation in optical properties from phytoplankton and other optically active constituents in the water. This finding led us to take the median of all pixels in each image for our analysis of the flight-long transects, which allowed us to identify and describe features within the span of the transects. Image-type data do convey a variety of advantages in this environment over point radiometers, such as the ability to spatially resolve sun glint that would otherwise be more complex to remove from L_t , and to inspect the front itself and visualize and spatially characterize the accumulation and structuring of sargassum. But, in a pelagic environment ultra-high-resolution images ($\sim 5 \text{ cm}$) may not be quite so advantageous over single-point measurements for ocean color analysis. In an estuarine environment, where absorption and backscattering are higher and vary at ultra-fine scale, high resolution UAS images can reveal a great deal of information concerning water constituents and their spatial distribution. Along the continuum from estuarine (Fig. 8) to pelagic (Fig. 9) environments, the variation within in image becomes substantially less, and the ocean color signal can be obscured by optical noise (e.g., surface waves and sun beams) both due to less variation, and less absorption and backscattering which fundamentally provide the ocean color signal. However, the environmental threshold at which this obscuring occurs, and

the scale of spatial resolution at which this becomes relevant, was not explored in this work, nor were other methods that may help resolve these sources of optical noise. Even in open-ocean environments, we find that having the image-type data conveys advantages that make it thoroughly worth collecting.

Caveats

Our method's requirement of an *in situ* SBA measurement for reflected skylight removal may not be practical in future missions using autonomous, long endurance platforms, especially given its assumption that sky conditions and sun angle do not change significantly during a survey. For UAS with longer endurance, it would be more appropriate to mount two sensors to simultaneously collect L_{sky} and L_t and use Eq. 1 to calculate L_w . Our described approach does not resolve the effect of wave facets on both glint and reflected skylight; this effect challenges many approaches, limiting the most accurate measurements to low wind speeds (Groetsch et al. 2020). However, image data from a wide-FOV sensor may be averaged to potentially reduce this effect, where a narrow-FOV radiometer would potentially be dominated by a few wave facets.

Early attempts at this work used the gimbal to change the pitch of the sensor and take an L_{sky} image measurement once every 10 L_t images, but this was challenged by our hardware configuration, because our sensor was partly blocked by the propellers when skyward, and our sensor's firmware forced a recalibration to integration time and gain because the L_{sky} image was so much brighter than L_t images. These challenges could be fixed by using two sensors, though this solution entails a higher cost and payload weight.

While not a caveat of this method in particular, an issue with the intercomparisons throughout this study, and a persistent challenge in intercomparison between methods in ocean color, was that this area was highly variable. The Gulf Stream front is a step change that moves in space over time, and there is substantial turbulence and micro-upwelling such that water properties could vary substantially within a few hundreds of meters. The current speed was 1.5 m s^{-1} , and we attempted to match up *in situ* samples in space with the UAS measurements in the moving current, but this was quite challenging. This applies also to the satellite imagery comparison where unless the timing is near synchronous, the advection, upwelling, and evolution of water properties can make intercomparisons especially noisy.

The “synopticity problem” continues to challenge oceanographic research, but new tools such as UAS erode its scale and impact. The potential contributions of a UAS method must be placed in proper context: aerial ocean color imagery can complement, but not fully replace, many oceanographic data streams, and often the value of a dedicated UAS survey is dwarfed by the scientific value and operational costs of an underway research vessel and its many instruments. But where small UAS can complement these operations—for example, if a vessel is at a planned stop—a UAS survey can add a

meaningful spatial footprint and provide insight into regional features and variability. A UAS flying a typical speed 10 m s^{-1} for a common flight time of 30 min can fly 18 km, a footprint that can be deployed judiciously over a study region with a variety of potential flight paths. The use of multiple UAS simultaneously, of longer endurance platforms, and of new and additional sensor payloads will all increase the scientific return on time and investment with UAS methods. Oceanographic regions of high spatiotemporal variability, such as at fronts, eddies, and coastal areas, justify the costs of these localized and rapidly deployable UAS methods more readily than areas of low variability.

Comments and recommendations

From estuaries and coastal waters to open-ocean fronts, eddies, and blooms, there are ocean regions that are rich with spatiotemporally fine-scale features that unfold within pixels of satellite imagery and beyond the ability of *in situ* methods to sample synoptically. UAS can help address this blindspot (Gray et al. 2022) if proper methods are used to calibrate measurements of ocean color and constrain uncertainty from these platforms. This work demonstrates that small UAS with medium- to low-cost sensors can resolve fine-scale features, even in pelagic environments, and are a useful new option in the oceanographer's toolkit.

As the ocean science community more broadly adopts multispectral or hyperspectral imagers and UAS, this work (1) describes a simple *in situ* approach for removing reflected skylight, a key challenge in above-water radiometry; (2) demonstrates the challenges, inherent to high-resolution imaging spectroscopy of aquatic surfaces, that are associated with myriad viewing geometries; (3) constrains uncertainty of these data in an ideal environment and propagates this to derived Chl a measurements, and (4) demonstrates the implementation of this workflow with surveys of the Gulf Stream front.

The ocean color community needs consistent approaches for making UAS-based radiometric measurements and a thorough framework for recording and understanding uncertainty in these measurements. It is critical that the community converge on a small set of robust sensors, open-source code for processing, and consistent calibration and validation metrics in order to add these potentially gap-filling tools into the optical oceanography toolkit. If we succeed, UAS-based sampling of ocean color can address a heretofore inaccessible scale that is critical for our understanding of marine ecosystems.

Data availability statement

Data are available on the Duke Data Repository at <https://research.repository.duke.edu/> and code is available at https://github.com/marrs-lab/drone_ocean_color.

References

- Bailey, S. W., B. A. Franz, and P. J. Werdell. 2010. Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. *Opt. Express* **18**: 7521–7527. doi:[10.1364/oe.18.007521](https://doi.org/10.1364/oe.18.007521)
- Behrenfeld, M. J., and others. 2009. Satellite-detected fluorescence reveals global physiology of ocean phytoplankton. *Biogeosciences* **6**: 779–794. doi:[10.5194/bg-6-779-2009](https://doi.org/10.5194/bg-6-779-2009)
- Bower, A. S., H. T. Rossby, and J. L. Lillibridge. 1985. The Gulf Stream—Barrier or blender? *J. Phys. Oceanogr.* **15**: 24–32 doi:[10.1175/1520-0485\(1985\)015%3C0024:TGSOB%3E2.0.CO;2](https://doi.org/10.1175/1520-0485(1985)015%3C0024:TGSOB%3E2.0.CO;2)
- Carrizo, C., A. Gilerson, R. Foster, A. Golovin, and A. El-Habashi. 2019. Characterization of radiance from the ocean surface by hyperspectral imaging. *Opt. Express* **27**: 1750–1768. doi:[10.1364/oe.27.001750](https://doi.org/10.1364/oe.27.001750)
- Clarke, G. L., G. C. Ewing, and C. J. Lorenzen. 1970. Spectra of backscattered light from the sea obtained from aircraft as a measure of chlorophyll concentration. *Science* **167**: 1119–1121. doi:[10.1126/science.167.3921.1119](https://doi.org/10.1126/science.167.3921.1119)
- D’Ovidio, F., S. De Monte, S. Alvain, Y. Dandonneau, and M. Lévy. 2010. Fluid dynamical niches of phytoplankton types. *Proc. Natl. Acad. Sci. USA* **107**: 18366–18370. doi:[10.1073/pnas.1004620107](https://doi.org/10.1073/pnas.1004620107)
- Gilerson, A., C. Carrizo, R. Foster, and T. Harmel. 2018. Variability of the reflectance coefficient of skylight from the ocean surface and its implications to ocean color. *Opt. Express* **26**: 9615. doi:[10.1364/oe.26.009615](https://doi.org/10.1364/oe.26.009615)
- Gray, P. C., G. D. Larsen, and D. W. Johnston. 2022. Drones address an observational blind spot for biological oceanography. *Front. Ecol. Environ.* **20**: 413–421. doi:[10.1002/fee.2472](https://doi.org/10.1002/fee.2472)
- Groetsch, P. M. M., R. Foster, and A. Gilerson. 2020. Exploring the limits for sky and sun glint correction of hyperspectral above-surface reflectance observations. *Appl. Optics* **59**: 2942–2954. doi:[10.1364/ao.385853](https://doi.org/10.1364/ao.385853)
- Gula, J., M. J. Molemaker, and J. C. McWilliams. 2015. Gulf Stream dynamics along the southeastern U.S. seaboard. *J. Phys. Oceanogr.* **45**: 690–715. doi:[10.1175/JPO-D-14-0154.1](https://doi.org/10.1175/JPO-D-14-0154.1)
- Hu, C., Z. Lee, and B. Franz. 2012. Chlorophyll *a* algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *J. Geophys. Res. Ocean.* **117**: 1–25. doi:[10.1029/2011JC007395](https://doi.org/10.1029/2011JC007395)
- IOCCG. 2008. IOCCG report number 07: Why ocean colour? The societal benefits of ocean-colour technology. International Ocean Colour Coordinating Group.
- IOCCG. 2019. IOCCG ocean optics and biogeochemistry protocols for satellite ocean colour sensor validation. IOCCG Protocol Series 3. International Ocean Colour Coordinating Group. doi:[10.25607/OBP-691](https://doi.org/10.25607/OBP-691)
- Johnson, Z. I., R. Shyam, A. E. Ritchie, C. Mioni, V. P. Lance, J. W. Murray, and E. R. Zinser. 2010. The effect of iron-and light-limitation on phytoplankton communities of deep chlorophyll maxima of the western Pacific Ocean. *J. Mar. Res.* **68**: 283–308. doi:[10.1357/002224010793721433](https://doi.org/10.1357/002224010793721433)
- Kim, M., J. Y. Park, Y. Kopilevich, G. Tuell, and W. Philpot. 2013. Correction for reflected sky radiance in low-altitude coastal hyperspectral images. *Appl. Optics* **52**: 7732–7744. doi:[10.1364/AO.52.007732](https://doi.org/10.1364/AO.52.007732)
- Kirk, J. T. O. 2010. Light and photosynthesis in aquatic ecosystems, 3rd Edition. Cambridge University Press.
- Klymak, J. M., and others. 2016. Submesoscale streamers exchange water on the north wall of the Gulf Stream. *Geophys. Res. Lett.* **43**: 1226–1233. doi:[10.1002/2015GL067152](https://doi.org/10.1002/2015GL067152)
- Lee, Z., Y.-H. Ahn, C. Mobley, and R. Arnone. 2010. Removal of surface-reflected light for the measurement of remote-sensing reflectance from an above-surface platform. *Opt. Express* **18**: 26313–26324. doi:[10.1364/oe.18.026313](https://doi.org/10.1364/oe.18.026313)
- Lee, Z., N. Pahlevan, Y. H. Ahn, S. Greb, and D. O’Donnell. 2013. Robust approach to directly measuring water-leaving radiance in the field. *Appl. Optics* **52**: 1693–1701. doi:[10.1364/AO.52.001693](https://doi.org/10.1364/AO.52.001693)
- Lee, Z., J. Wei, Z. Shang, R. Garcia, H. Dierssen, J. Ishizaka, and A. Castagna. 2018. On-water radiometry measurements: Skylight-blocked approach and data processing. IOCCG Protocol Series. 7. International Ocean Colour Coordinating Group.
- Lévy, M., P. Klein, and A. M. Treguier. 2001. Impact of submesoscale physics on production and subduction of phytoplankton in an oligotrophic regime. *J. Mar. Res.* **59**: 535–565. doi:[10.1357/002224001762842181](https://doi.org/10.1357/002224001762842181)
- Lévy, M., P. J. S. Franks, and K. S. Smith. 2018. The role of submesoscale currents in structuring marine ecosystems. *Nat. Commun.* **9**: 4758. doi:[10.1038/s41467-018-07059-3](https://doi.org/10.1038/s41467-018-07059-3)
- Mahadevan, A. 2016. The impact of submesoscale physics on primary productivity of plankton. *Ann. Rev. Mar. Sci.* **8**: 161–184. doi:[10.1146/annurev-marine-010814-015912](https://doi.org/10.1146/annurev-marine-010814-015912)
- Mann, K. H., and J. R. N. Lazier. 2005. Dynamics of marine ecosystems. Blackwell Publishing.
- Martin, A. P. 2003. Phytoplankton patchiness: The role of lateral stirring and mixing. *Prog. Oceanogr.* **57**: 125–174. doi:[10.1016/s0079-6611\(03\)00085-5](https://doi.org/10.1016/s0079-6611(03)00085-5)
- McClain, C. R. 2009. A decade of satellite ocean color observations. *Ann. Rev. Mar. Sci.* **1**: 19–42. doi:[10.1146/annurev.marine.010908.163650](https://doi.org/10.1146/annurev.marine.010908.163650)
- McWilliams, J. C., J. Gula, and M. J. Molemaker. 2019. The Gulf Stream North Wall: Ageostrophic circulation and frontogenesis. *J. Phys. Oceanogr.* **49**: 893–916. doi:[10.1175/JPO-D-18-0203.1](https://doi.org/10.1175/JPO-D-18-0203.1)
- Mobley, C. D. 1999. Estimation of the remote-sensing reflectance from above-surface measurements. *Appl. Optics* **38**: 7442–7455. doi:[10.1364/ao.38.007442](https://doi.org/10.1364/ao.38.007442)
- Morel, A. 1980. In-water and remote measurements of ocean color. *Boundary Layer Meteorol.* **18**: 177–201. doi:[10.1007/BF00121323](https://doi.org/10.1007/BF00121323)

- Moses, W. J., S. G. Ackleson, J. W. Hair, C. A. Hostetler, and D. W. Miller. 2016. Spatial scales of optical variability in the coastal ocean: Implications for remote sensing and in situ sampling. *J. Geophys. Res. Ocean.* **121**: 1–15. doi:[10.1002/2016JC011767](https://doi.org/10.1002/2016JC011767). Received
- O'Reilly, J. E., and P. J. Werdell. 2019. Chlorophyll algorithms for ocean color sensors—OC4, OC5 & OC6. *Remote Sens. Environ.* **229**: 32–47. doi:[10.1016/j.rse.2019.04.021](https://doi.org/10.1016/j.rse.2019.04.021)
- O'Shea, R. E., S. R. Laney, and Z. Lee. 2020. Evaluation of glint correction approaches for fine-scale ocean color measurements by lightweight hyperspectral imaging spectrometers. *Appl. Optics* **59**: B18–B34. doi:[10.1364/ao.377059](https://doi.org/10.1364/ao.377059)
- Pahlevan, N., S. Sarkar, B. A. Franz, S. V. Balasubramanian, and J. He. 2017. Sentinel-2 multispectral instrument (MSI) data processing for aquatic science applications: Demonstrations and validations. *Remote Sens. Environ.* **201**: 47–56. doi:[10.1016/j.rse.2017.08.033](https://doi.org/10.1016/j.rse.2017.08.033)
- Ruddick, K. G., K. Voss, E. Boss, and others. 2019. A review of protocols for fiducial reference measurements of water-leaving radiance for validation of satellite remote-sensing data over water. *Remote Sens.* **11**. doi:[10.3390/rs11192198](https://doi.org/10.3390/rs11192198)
- Ruiz, S., and others. 2019. Effects of oceanic mesoscale and submesoscale frontal processes on the vertical transport of phytoplankton. *J. Geophys. Res. Ocean.* **124**: 5999–6014. doi:[10.1029/2019JC015034](https://doi.org/10.1029/2019JC015034)
- Siegel, D. A., M. Wang, S. Maritorena, and W. Robinson. 2000. Atmospheric correction of satellite ocean color imagery: The black pixel assumption. *Appl. Optics* **39**: 3582–3591. doi:[10.1364/ao.39.003582](https://doi.org/10.1364/ao.39.003582)
- Tanaka, A., H. Sasaki, and J. Ishizaka. 2006. Alternative measuring method for water-leaving radiance using a radiance sensor with a domed cover. *Opt. Express* **14**: 3099–3105. doi:[10.1364/oe.14.003099](https://doi.org/10.1364/oe.14.003099)
- Taylor, J. R. 1996. An introduction to error analysis: The study of uncertainties in physical measurements, 3rd Edition. University Science Books.
- Werdell, P. J., and others. 2019. The plankton, aerosol, cloud, ocean ecosystem mission status, science, advances. *Bull. Am. Meteorol. Soc.* **100**: 1775–1794. doi:[10.1175/BAMS-D-18-0056.1](https://doi.org/10.1175/BAMS-D-18-0056.1)
- Westoby, M. J., J. Brasington, N. F. Glasser, M. J. Hambrey, and J. M. Reynolds. 2012. “Structure-from-Motion” photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **179**: 300–314. doi:[10.1016/j.geomorph.2012.08.021](https://doi.org/10.1016/j.geomorph.2012.08.021)
- Windle, A. E., and G. M. Silsbe. 2021. Evaluation of unoccupied aircraft system (UAS) remote sensing reflectance retrievals for water quality monitoring in coastal waters. *Front. Environ. Sci.* **9**: 1–15. doi:[10.3389/fenvs.2021.674247](https://doi.org/10.3389/fenvs.2021.674247)
- Wulder, M. A., and others. 2019. Current status of Landsat program, science, and applications. *Remote Sens. Environ.* **225**: 127–147. doi:[10.1016/j.rse.2019.02.015](https://doi.org/10.1016/j.rse.2019.02.015)
- Zhang, X., S. He, A. Shabani, P.-W. Zhai, and K. Du. 2017. Spectral sea surface reflectance of skylight. *Opt. Express* **25**: A1–A13. doi:[10.1364/oe.25.0000a1](https://doi.org/10.1364/oe.25.0000a1)

Acknowledgments

Funding support was provided by National Aeronautics and Space Administration (NASA) Future Investigators in NASA Earth and Space Science and Technology (FINESST) #80NSSC19K1366 in conjunction with the Terrestrial Ecology and Ocean Biology and Biogeochemistry programs, and the American Society for Photogrammetry and Remote Sensing's Robert N. Colwell Memorial Fellowship. Ship time was supported by Office of Naval Research—Naval Research Lab WU# 72-1R25 and the Duke Marine Lab R/V Shearwater student research fund. The authors thank the crew of the R/V Shearwater, Matt Dawson, Tina Thomas, Zach Swaim, and John Wilson.

Conflict of interest

None declared.

Submitted 23 March 2022

Revised 01 July 2022

Accepted 17 August 2022

Associate editor: Ivona Cetinic