Behind the masks in coastal ocean color

Henry F Houskeeper[1] and Raphael M Kudela[1]

[1]Ocean Sciences Department
University of California, Santa Cruz
1156 High St.
Santa Cruz, CA 95064

Introduction: During standard processing of level 3 (L3) NASA Ocean Color data, suspect retrievals are masked prior to spatial and temporal binning in order to provide high-quality composites. Default L3 masks may be triggered, for example, when satellite or sun remote sensing geometries are poor (i.e. high sunglint or high sensor view zenith angle), or when atmospheric correction errors are likely (i.e. derivation of negative water-leaving radiances). For biological remote sensing of the coastal ocean, masking low-quality retrievals in a manner independent of sought after environmental parameters is challenging due to strong biological modification of remote sensing signals, such as the generation of non-negligible near infrared reflectances by dense algal concentrations (Siegel et al. 2000). Here we evaluate the phenology of default L3 masks in Monterey Bay, CA and find that climatology of flags indicating likely atmospheric correction errors resembles seasonal cycles in phytoplankton biomass rather than climatology of corresponding atmospheric constituents. We find that flags indicating likely atmospheric correction errors are more frequently assigned to pixels with high normalized fluorescence line height (nFLH; 60.6% of pixels above upper quartile, 30.8% of pixels below lower quartile); nFLH is a good proxy for red tide events in coastal water masses, and nFLH is relatively robust to atmospheric correction errors. Standard removal of high nFLH observations before L3 binning decreases the mean (7.1%) and variance (22.1%) of the level 2 (L2) nFLH measurements in our study area and hides observations from biologically important regions, such as dinoflagellate-rich upwelling retention zones. Our results suggest that removal of biologically-important observations using standard L3 processing may distort satellite perspectives of change in coastal ecosystems.

Background: Although ocean-observing satellites allow near-daily coverage of most ocean regions, accuracy of retrieved geophysical parameters is low in coastal and inland watermasses where terrestrial aerosols, resuspended particles, and organic constituents confound ocean color methods designed for the global ocean. Flags which identify low-quality retrievals for masking allow a basic level of quality control of ocean color data, and different flag criteria are chosen to mask, or remove, observations from sequential data tiers. Data processed to the L2 tier, or individual atmospherically-corrected derived geophysical parameters, is masked by default when the derivation of meaningful products is severely inhibited, for example when the sensor is viewing land or clouds, or when the sensor saturates. L2 datasets contain shifting pixel coordinates, frequent data gaps (i.e. from clouds), and large file sizes inconvenient for users requiring continuous or less computationally expensive products (Campbell et al., 1995). To satisfy these user needs, statistical composites of geophysical variables binned in space and time (L3 data tier) are provided by the NASA Ocean Biology Processing Group (OBPG) and are of high importance to users beyond the ocean color community, for example as inputs into biogeochemical models. As a result, the default masks applied during L3 processing are more rigorous than during L2 processing in order to provide higher quality composites

for a larger end-user community, and default L3 masks remove observations flagged for likely or known errors in atmospheric correction.

Conventional approaches for atmospheric correction take advantage of the strong light absorption by water at longer wavelengths to estimate that the water-leaving radiance (L_w) in the near-infrared (NIR) is negligible (Gordon & Wang, 1994). Thus, after removal of glint and white capping effects, the derived top-of-atmosphere (TOA) radiance (L_{TOA}) in the NIR region is attributed to the atmospheric contributions by Rayleigh scattering, aerosol scattering, and multiple interactions between Rayleigh and aerosol scattering.

The approximation that L_W is zero within the NIR region, termed the "Black Pixel Assumption" (Siegel et al., 2000), is frequently incorrect in coastal waters where high near-surface particle loads (organic or inorganic) can strongly backscatter light, such that L_W in the NIR domain is appreciably non-zero. Because the TOA NIR signal is attributed to atmospheric contributions, NIR L_W contributions to L_{TOA} cause overestimation of backscattering by atmospheric aerosols, and thus result in incorrect (often negative) derivation of L_W in the visible domain, particularly in the blue bands which are used for (e.g.) chlorophyll derivation (IOCCG Report 10, 2010). As a result, high particle loads, including of algal cells, increase the likelihood of poor atmospheric correction among high biomass pixels.

In this study, we evaluate the extent to which the relationship between biology and success of atmospheric retrieval affects coastal ocean color datasets, and we assess whether the removal of high biomass data biases satellite perspectives of biology in coastal zones. We examine default flags applied during standard L3 processing that indicate likely or definite errors in atmospheric correction, hereafter cumulatively referred to as AC flags. Our selection includes flags triggered when the maximum allowable iterations of the aerosol model have run without reaching model agreement (MAXAERITER), when low L_W is derived (LOWLW), and when likely errors are detected in the atmospheric correction (ATMWARN). We report that AC flag assignments in Monterey Bay, CA track biological rather than atmospheric constituents, and we assess the impacts on L2 and L3 ocean color datasets from removing observations in a manner dependent on the underlying biology.

Approach: *Remote sensing data:* L2 MODIS *Aqua* data was obtained from the NASA Ocean Color level 2 browser (oceancolor.gsfc.nasa.gov) for 2002 – 2018 within Monterey Bay, CA (36.5N – 37.0N; 121.75W – 122.25W), a productive marine sanctuary within the coastal upwelling ecosystem of the central California Current. A baseline L2 dataset was developed with observations removed if assigned LAND, CLDICE, HIGLINT, HILT, HISATZEN, HISOLZEN, NAVFAIL, ATMFAIL, and NAVWARN flags with standard (OBPG) flag thresholds. Climatology of the remaining flags was calculated using the monthly mean fraction of flagged observations for each flag. A secondary L2 dataset was developed with the additional removal of observations assigned any of the AC flags (LOWLW, MAXAERITER, and ATMWARN). L3 (8-day, 9km) MODIS *Aqua* composites were derived using the SeaDAS (version 7.5) scripts l2bin and smigen (available at https://seadas.gsfc.nasa.gov) for two identical L2 datasets with differing L3 processing options to retain versus exclude observations assigned AC flags.

Atmospheric Ground Data: Climatological data for Angstrom exponent, aerosol optical depth (500nm), and fraction precipitable water was obtained for the Monterey site of the Aerosol Robotic Network (AERONET; http://aeronet.gsfc.nasa.gov) in southern Monterey Bay, CA (36.59255 N, 121.85487 W). Although the site is located roughly 2 kilometers from the coast, the region's prevailing winds blow predominantly onshore, and coastal mountains to the east separate the station from a nearby agricultural region. Comparison with satellite data was achieved using a matchup domain bounded by the northern extent of the adjacent coastal mountain range (36.650N), the western limit of the Monterey headlands (-121.920 E), and the Monterey coastline to the south and east.

Biological Ground Data: Weekly fluorometric chlorophyll *a* measured at the Santa Cruz Wharf (SCW), in northern Monterey Bay, CA, was obtained from the Southern California Coastal Ocean Observing system portal (http://www.sccoos.org/data/habs/). Climatology of chlorophyll *a* was derived as the mean of $\log_{10}$ transformed values for each month. Matchups with satellite data were selected from pixels within 9.6 kilometers from SCW, as this distance was found to provide the maximum Pearson coefficient between satellite chlorophyll (OCI) and *in situ* chlorophyll *a* measurements.

Results: *Climatology of atmospheric correction flags and comparison with ground data:* After removal of pixels flagged for obstruction by clouds, high glint, saturating radiances, poor viewing geometry, or likely navigation errors, the MAXAERITER, LOWLW, and ATMWARN flag assignments each show distinct spring and fall maxima. Cumulatively, AC flag assignments within the AERONET matchup domain reach 47.7% of April observations and 38.2% of September observations (figure 1a). At Monterey AERONET, Angstrom exponent climatology suggests seasonal fluctuations in particle size, with maximum Angstrom exponent values occurring in late summer / early fall, and minimum values occurring in December (figure 1b). Precipitable water records at Monterey AERONET show maximum values occurring in late summer / early fall, and minimum values occurring in winter, albeit with high variability (monthly standard deviations on the order of 50% of mean observations; figure 1c). Aerosol optical depth (500nm) at Monterey AERONET is elevated and highly variable during summer, while lower and more stable during fall, winter, and spring (figure 1d). No parameter obtained from the Monterey AERONET site suggests seasonal patterns in atmospheric structure that would be responsible for the spring and late fall maxima in AC flag coverage.

Within the SCW matchup domain in northern Monterey Bay, AC flag climatology closely resembles that described for the southern domain, with maxima for AC flag assignments occurring in spring (45.7%; April) and fall (37.1%; November; figure 2a). Climatology of $\log_{10}$ transformed chlorophyll *a* measured at SCW shows May and November peaks corresponding with the region's spring and fall phytoplankton blooms (figure 2b). Climatology of AC flags and chlorophyll both appear highly sensitive to a handful of large phytoplankton blooms. For example, if the analysis omits 2007, a year with a massive November dinoflagellate red tide (Jessup et al, 2009), fall maxima for both *in situ* chlorophyll *a*, as well as AC flag assignments shift to October.

Impact of atmospheric correction masks on distribution of nFLH: Heightened assignments of AC flags among high biomass observations would be expected to suppress the dataset biomass distribution of default L3 products. Unfortunately, standard ocean color chlorophyll algorithms are highly sensitive to atmospheric correction errors

because of their reliance on blue-green wavelengths where the potential errors in modeled aerosol contributions are high. Instead, we compare normalized fluorescence line height (nFLH), a spectral shape algorithm using red and NIR wavelengths, which measures the fluorescence emitted by chlorophyll *a* molecules. Although susceptible to changes in phytoplankton physiology and prone to errors related to peak fluorescence wavelength relative to sensor band placement, nFLH has been shown to be a more robust biomass estimate in coastal waters than ocean color chlorophyll products and has been validated for use as a top of atmosphere product, or without application of an atmospheric correction step (Gower & King 2012, Gower 2016).

Comparison of nFLH between L2 datasets with AC flagged observations retained versus removed shows that the removal of AC flagged observations from the dataset results in reduction of the mean by 7.1% and reduction in the variance of the dataset by 22.1%. Although the overall impact to the mean is negative, we observe that very low nFLH observations are also frequently removed, presumably because the lowest nFLH values occur during winter, when particle resuspension from waves and river discharge may increase the likelihood of poor retrievals. Probability density functions of the distribution of nFLH before and after masking by AC flags illustrates the tail trimming effect of the AC flag masks (figure 3a). Comparison between the distributions shows that the likelihood that observations are assigned AC flags steadily increases with higher nFLH (figure 3b).

Propagation of bias from atmospheric correction masks to L3 dataset: The nFLH datasets with and without AC flags masked were processed to L3 composites to assess the impacts of AC flag masks at lower spatial and temporal resolutions. Root mean square difference (RMSD) calculated between nFLH datasets with AC-flagged observations rejected or retained shows that the difference between the two datasets is dependent on distance from the coast, with RMSD approaching 10% within standard map grid composites adjacent to the coast. L3 composites from the northeastern corner of Monterey Bay, a biologically productive region deemed the Monterey Bay red tide incubator (Ryan et al., 2008), were frequently absent in the dataset with AC-flagged observations removed, and thus no sensible RMSD could be derived within this region.

Discussion: Climatology of flag assignments indicating likely atmospheric correction errors were compared to *in situ* measurements of aerosols, water vapor, and phytoplankton. No atmospheric climatology record from the Monterey AERONET site shows similar seasonality to the AC flag assignments or suggests that atmospheric constituents are primarily responsible for the spring and fall mask increases. Instead, AC flag assignments closely track underlying seasonal phytoplankton cycles in Monterey Bay, and large red tide events coincide with heightened assignment of AC flags.

The climatological similarities between phytoplankton biomass and AC flags, which are not reflected in independent measurements of atmospheric structure, suggest that AC flags across the Monterey Bay region may preferentially remove observations of high biomass. Qualitative observations comparing nFLH and AC flags for MODIS *Aqua* scenes show similar shape and location between filaments with high nFLH and AC flag assignments. Similarly, we note that during large scale red tide events, daily L3 MODIS *Aqua* scenes of Monterey Bay may be fully masked despite clear-sky conditions.

Relevance to remote estimates of biomass: Traditional blue-green band ratio chlorophyll algorithms were not preferred for assessing potential biases in standard flags for this study because of their sensitivity to poor atmospheric correction procedures, which may be most pronounced at shorter wavelengths of the visible spectrum. Similarly, red band ratio chlorophyll algorithms were not favored because of the importance of accurate atmospheric retrievals, and because of the lack of a MODIS band near 700nm (Yacobi et al, 2010).

Criteria that trigger AC flag assignment often indicate that derivation of chlorophyll using standard NASA algorithms would be impossible or meaningless. For example, no band ratio chlorophyll product may be derived from negative reflectances resulting from the overestimation of atmospheric aerosols. To demonstrate, 23% of the L2 pixels observed in this study and assigned the 'LOWLW' flag failed to produce a standard chlorophyll product. Therefore, comparison of standard chlorophyll products between datasets with relaxed mask selections would minimize the apparent impact of re-introducing masked observations.

Comparison of nFLH before and after removal of AC-flagged pixels shows that poor atmospheric correction within more turbid watermasses causes underestimation of chlorophyll fluorescence and suggests that biologically important events such as red tides are heavily removed from ocean color time series. However, nFLH is an imperfect proxy for biomass, and interpreting nFLH introduces other uncertainties. Most fundamentally, chlorophyll *a* fluorescence is not solely a function of chlorophyll *a* concentration, although the two are expected to covary. Factors that may alter the relationship between chlorophyll *a* fluorescence and concentration include phytoplankton species composition and pigment packaging effects, physiology, and limitation of nutrients or light. Sensor-specific response functions also complicate the use of nFLH as a biomass proxy. For example, we expect to underpredict line height for high biomass observations with MODIS *Aqua* because of a corresponding shift in the red peak to a region between existing ocean color bands (~700nm; Gower & Borstad, 2004).

Despite the caveats described when equating biomass with nFLH, our results suggest that careful analysis of mask defaults must be considered when deriving time series in coastal waters, as results may differ based on whether or not AC-flagged pixels are masked. Although not significant, we find that MODIS *Aqua* nFLH observations within Monterey Bay (36.5, -37.0 N; -122.25, -121.75E) decreased roughly 3 times more quickly between 2002 and 2018 when using a time series with AC-flagged observations masked (-7.25%), compared to a time series with AC-flagged observations retained (-2.62%).

We also note that range thresholds defined for ocean color products do not match *in situ* ranges recorded in Monterey Bay. For example, standard ocean color chlorophyll algorithms set upper limits on valid products at 100mg/m^3, a value considered unlikely to occur in most parts of the ocean, while *in situ* measurements from SCW show that chlorophyll *a* may exceed the ocean color expected range. For example, on November 14th, 2007, a major dinoflagellate bloom in northern Monterey Bay, CA, resulted in fluorometric chlorophyll *a* measurements of 289 mg/m^3 (nearly three times the maximum accepted algorithm output). Because chlorophyll *a* at SCW never approached the lower ocean color algorithm threshold, low ranges for satisfactory algorithm performance are also expected to lead to underestimation of coastal biomass.

The results of our L3 composite comparison show that coastal L3 datasets may strongly differ from underlying L2 datasets because of default masks applied during standard processing, in particular from the AC flag group. Despite the smoothing effect from decreased spatial and temporal resolution, comparison of L3 nFLH datasets processed with and without inclusion of AC flagged observations shows that the effects of removing AC flagged pixels persist into L3 products. Normalized RMSD of the two datasets increases with proximity to the coast, with nearshore pixels approaching 10% (figure 4). The high removal of the red tide incubator region also suggests that default processing for L3 products may hide many biologically important regions.

Conclusion: Standard ocean color protocols are not meant for application in coastal waters. However, these regions are expected to undergo rapid change in the coming century, and ocean color satellites remain our best option for monitoring across wide spatial scales. As a result, standard satellite products designed for the global ocean are disseminated to researchers and modelers working on coastal zones. Careful consideration of how to quality control ocean color datasets without skewing the distribution of the remaining 'good' observations is required to provide a wide variety of end users with the best, unbiased estimate of the environment. Arguably, lower accuracy is preferable over systematic bias when assessing changes in biological regimes, and some past analyses of the California Current System have opted to not invoke AC flags before analyzing phytoplankton trends (Kahru et al. 2014, 2015).

Relaxing AC flags for ocean color processing does not in itself solve the problem described here in which high biomass observations such as red tides are absent from ocean color datasets because the AC flag criteria detect valid irregularities indicative of low quality data. However, our results suggest that the data remaining after removal of AC flags may also be considered suspect in regions such as Monterey Bay, where AC flags can remove nearly half of the otherwise valid observations during high biomass seasons. In effect, ocean color assessments of biological change in coastal zones should be wary of the amount of data missing due to poor retrievals. Products that are relatively insensitive to atmospheric correction errors allow comparison between datasets before and after removal of low quality retrievals, and can test whether regional datasets are robust to the removal of biologically important data. Such comparisons should be consulted when observing coastal ocean ecosystems from space.

Figure. 1. Climatology of combined MAXAERITER, LOWLW, and ATMWARN flags compared with atmospheric climatology from the Monterey AERONET site.

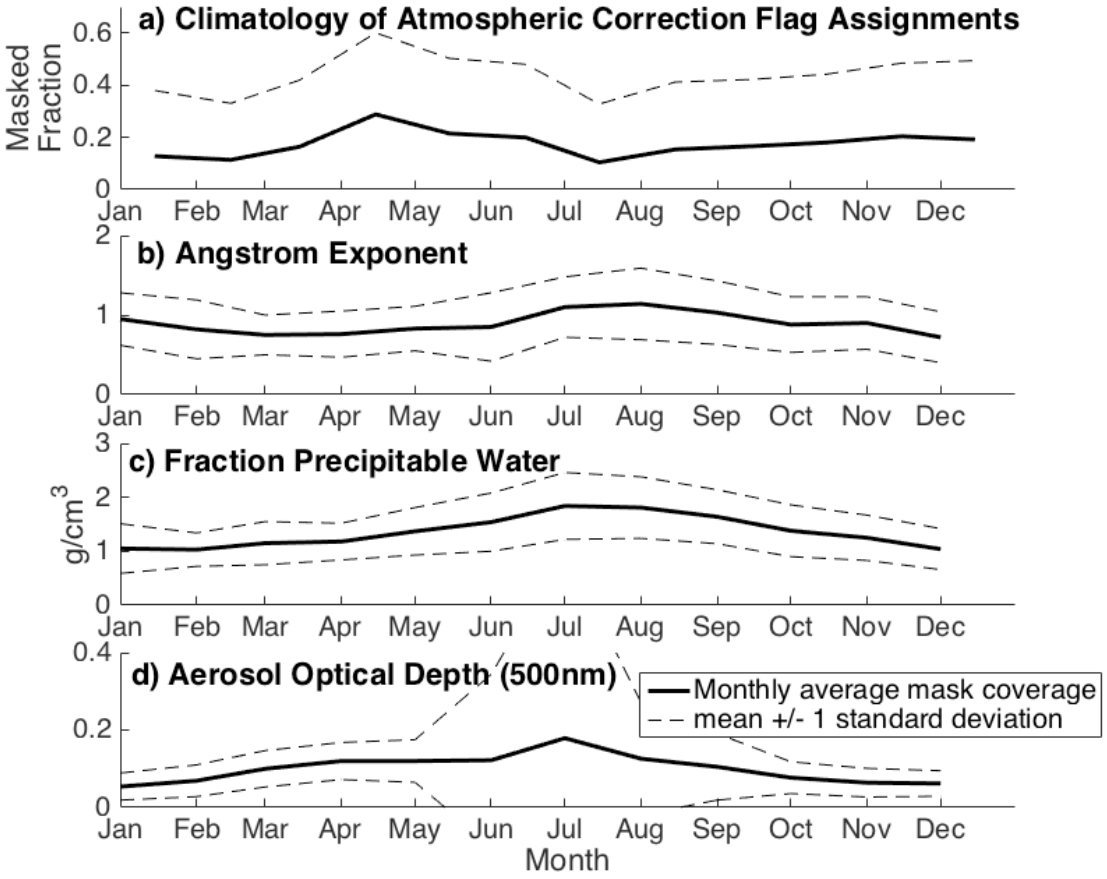

Fig 1. Monthly mean climatology shown in solid black with one standard deviation above and below the mean monthly values shown in dashed black for (a) MODIS *Aqua* L2 MAXAERITER, LOWLW, and ATMWARN flags in southern Monterey Bay, (b) Angstrom exponent (440nm-870nm), (c) precipitable water, and (d) aerosol optical depth (500nm) at Monterey AERONET site. AERONET climatology provided by the Aerosol Robotic Network (aeronet.gsfc.nasa.gov) data access portal.

Figure 2. Climatology of combined MAXAERITER, LOWLW, and ATMWARN flags compared with climatology of chlorophyll *a* at the Santa Cruz Wharf

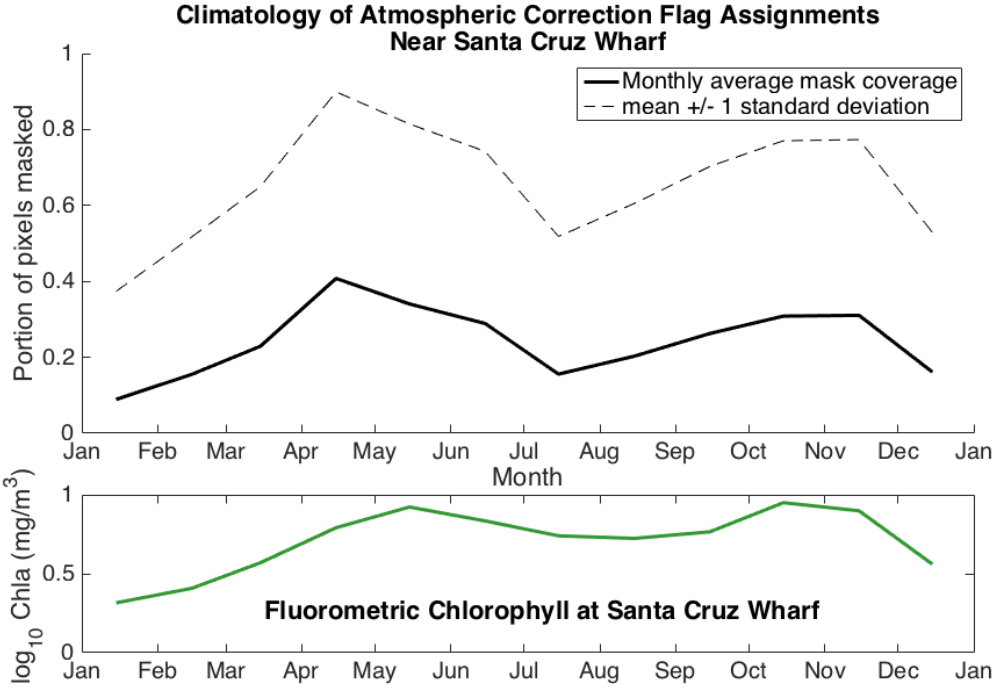

Fig 2. a) Monthly climatology of percent coverage of the MAXAERITER flag in MODIS *Aqua* L2 data from 2002 to 2018 within 9 km of Santa Cruz Wharf after removal of pixels flagged for land, clouds, glint, poor sensor geometry and saturating radiances. Mean coverage plotted in solid black, with one standard deviation above the mean shown with dashed black line.

b) Monthly climatology of mean $\log_{10}$ fluorometric chlorophyll *a* from weekly measurements at SCW (2006 – 2015).

Figure 3. Probability density function of log-transformed nFLH with and without atmospheric correction flags, and the percent removal of data as a function of log-transformed nFLH.

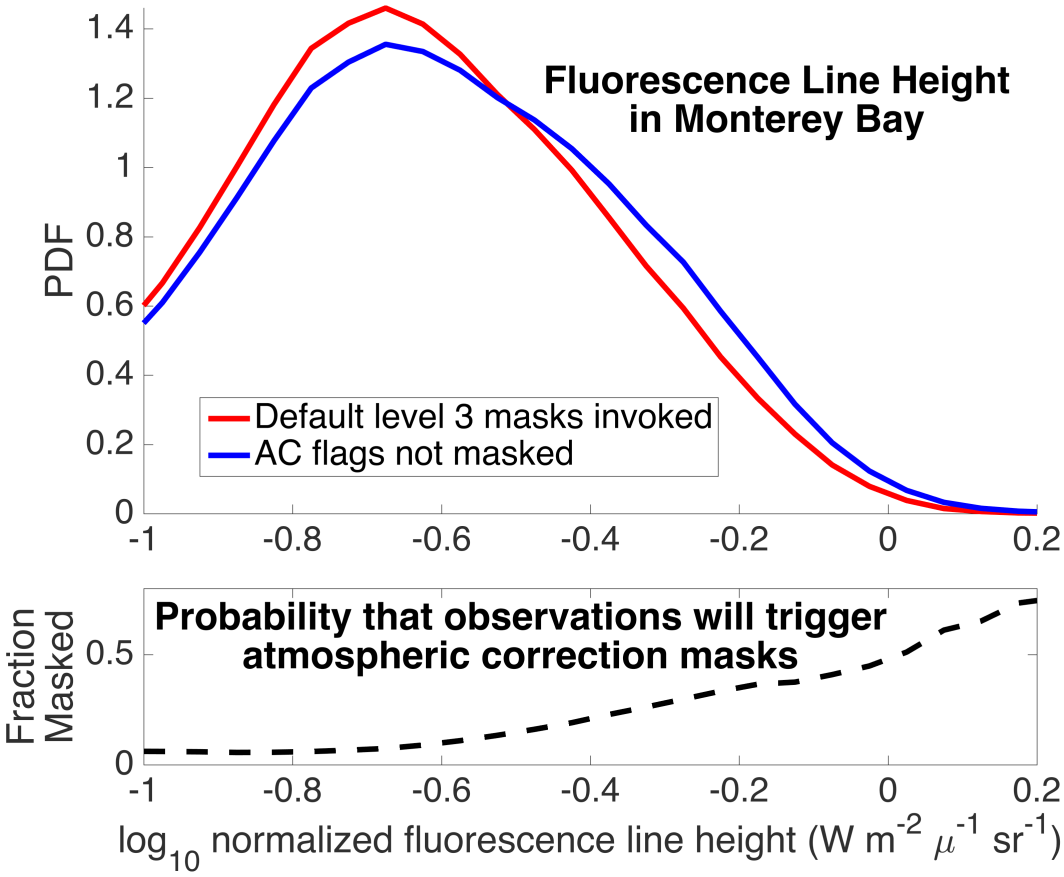

Fig 3. Top Panel: Probability density functions of log-transformed normalized fluorescence line height from L2 MODIS *Aqua* data from Monterey Bay, CA, 2002-2018. Distribution of data masked for glint, poor viewing geometries, navigation failures, and saturating radiances shown in blue. Distribution of data additionally masked for likely atmospheric correction errors (MAXAERITER, LOWLW, and ATMWARN) shown in red.

Bottom Panel: Percent removal of L2 MODIS *Aqua* data by the MAXAERITER, LOWLW, and ATMWARN flags (after other default L3 flags removed) as a function of log-transformed normalized fluorescence line height.

Figure 4. Normalized RMSD of L3 data with and without MAXAERITER, LOWLW, and ATMWARN flags applied during L3 processing

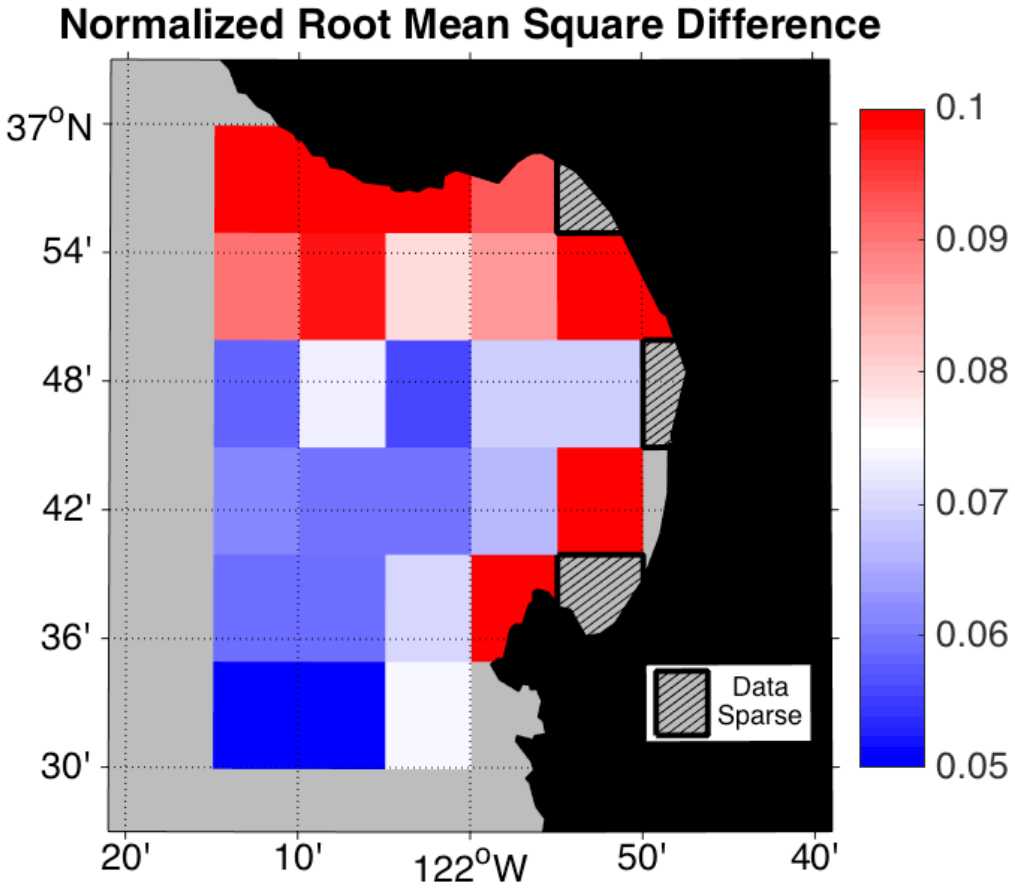

Fig 4. Normalized root mean square difference between L3 MODIS *Aqua* (9km, 8 day) datasets produced with default L3 processing versus L3 processing with MAXAERITER, LOWLW, and ATMWARN flag masks turned off. Normalized root mean square difference was not calculated for standard map grids if data was present when AC flagged observations were retained but not when AC flagged observations were removed (N < 40), and are indicated by diagonal hatch marks.

References:

Campbell, J.W., J.M. Blaisdell, & M. Darzi. (1995). Level-3 SeaWiFS Data Products: Spatial and Temporal Binning Algorithms. *NASA Technical Memorandum 104566.* 32, 73 pp.

Gordon, H. R., & Wang, M. (1994). Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. *Applied optics*, *33*(3), 443-452.

Gower, J. (2016). On the use of satellite-measured chlorophyll fluorescence for monitoring coastal waters. *International Journal of Remote Sensing*, *37*(9), 2077-2086.

Gower, J. F. R., & Borstad, G. A. (2004). On the potential of MODIS and MERIS for imaging chlorophyll fluorescence from space. *International Journal of Remote Sensing*, *25*(7-8), 1459-1464.

Gower, J., & King, S. (2012). Use of satellite images of chlorophyll fluorescence to monitor the spring bloom in coastal waters. *International journal of remote sensing*, *33*(23), 7469-7481.

IOCCG (2010). Atmospheric Correction for Remotely-Sensed Ocean-Colour Products. Wang, M. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 10, IOCCG, Dartmouth, Canada.

Jessup, D. A., Miller, M. A., Ryan, J. P., Nevins, H. M., Kerkering, H. A., Mekebri, A., Crane, D. B., Johnson, T. A., & Kudela, R. M. (2009). Mass stranding of marine birds caused by a surfactant-producing red tide. *PLoS One*, *4*(2), e4550.

Kahru, M., Kudela, R. M., Anderson, C. R., Manzano-Sarabia, M., & Mitchell, B. G. (2014). Evaluation of satellite retrievals of ocean chlorophyll-a in the California Current. *Remote Sensing*, *6*(9), 8524-8540.

Kahru, M., Lee, Z., Kudela, R. M., Manzano-Sarabia, M., & Mitchell, B. G. (2015). Multi-satellite time series of inherent optical properties in the California Current. *Deep Sea Research Part II: Topical Studies in Oceanography*, *112*, 91-106.

Ryan, J. P., Gower, J. F., King, S. A., Bissett, W. P., Fischer, A. M., Kudela, R. M., Kolber, Z., Mazzillo, F., Rienecker, E.V., & Chavez, F. P. (2008). A coastal ocean extreme bloom incubator. *Geophysical Research Letters*, *35*(12).

Siegel, D.A., M. Wang, S. Maritorena, & W. Robinson. (2000). Atmospheric correction of satellite ocean color imagery: the black pixel assumption. *Applied Optics*. 39, 3582-3591. doi:10.1364/AO.39.003582

Yacobi, Y.Z., Moses, W.J., Kaganovsky, S., Sulimani, B., Leavitt, B.C. and Gitelson, A.A., 2010, June. Chlorophyll a in turbid productive waters: testing the limits of NIR-Red algorithms. In *Proc. 'ESA Living Planet Symposium', Bergen, Norway* (Vol. 28).