Project Number: 901707

Project Start: Jan. 1, **2017**[1] **Project Duration (in years):** **Year 1 of 1** [1]

Please indicate if proposal includes requests for any of the following (mark all that apply).

[ ]	[ ]	[ ]	[ ]	[ ]
Ship Time	**ROV**	**AUV**	**Wave Glider**	**Mini ROV**

Please indicate what type of project is being proposed (mark only one).

[X]	[ ]	[ ]	[ ]	[ ]
Feasibility/Scoping [Less than 100 days effort]	**Development**[2]	**Research**	**Infrastructure**[2]	**Education**
	[ ]	[ ]		
	Time Series	**Institute Initiative**		

Project Manager[3]: Monique Messié

Principal Investigator: Monique Messié

Lead Scientist: Monique Messié, Francisco Chavez

Lead Engineer: Mathieu Kemp

Project Title: Phytoplankton taxa proxies from satellite and autonomous sensors

[1] This date and number should remain the same for the duration of the project. If your project start date for 2017 is not January 1, please indicate the start date in the first cell of the milestone chart. Format for duration example: “Year 2 of 5”.
[2] Development and infrastructure projects must complete the risk matrix (see item #5 at the end of the proposal form).
[3] Not all projects will require all four titles to be filled in. Please indicate individuals only for those titles that are appropriate for your project.

1) ABSTRACT (new projects only)

We currently rely on *in situ* sampling to get taxonomic information on phytoplankton communities, whether from ship rosette, Dorado gulper, LRAUV sipper or ESP. The goal of this feasibility study is to examine whether reliable proxies for phytoplankton taxonomic groups (i.e. diatoms, dinoflagellates, etc.) can be derived from alternate measurements, namely (1) autonomous fluorescence/backscatter sensors currently onboard most AUVs, ships CTD/rosette and gliders, and (2) satellite ocean color in the California Current. If this is proven feasible it opens up an incredible opportunity to investigate decades of local historical data and would provide an unprecedented view of phytoplankton ecology in the region. Moreover, in the future real-time processing of phytoplankton taxa proxies onboard AUVs would enable more advanced targeted sampling.

2) PROPOSAL

a) <u>New Projects</u>

Statement of the problem, why it is important

Phytoplankton is the base of the oceanic food web, and as such drives the structure and productivity of bottom-up ecosystems such as along the West Coast. Total phytoplankton biomass is relatively well monitored from satellite and autonomous sensors, using chlorophyll concentration as a proxy for biomass. However, ecosystems health and ecological status depend not only on the total amount of phytoplankton available to the food web, but also on the phytoplankton species. Communities dominated by diatoms vs. dinoflagellates may support very different food webs, even if their total biomass is similar. Certain species are toxic and result in harmful algal blooms with devastating consequences for marine life. More generally, understanding biodiversity patterns is a central issue in ecology, and phytoplankton biodiversity patterns remain poorly known.

We currently rely on *in situ* sampling to get taxonomic information on phytoplankton communities, whether from ship rosette, Dorado gulpers, LRAUV sipper or ESP. Other specialized instruments currently onboard Dorado (LISST 100x, LISST-HOLO, LOPC, bathyphotometer) can provide taxonomic information but remain limited to one platform. By contrast, commonly used optical fluorescence/backscatter sensors are available onboard most AUVs, gliders, ship CTDs and the expanding fleet of biogeochemical Argo floats. Taxonomy-specific phytoplankton optical properties provide a path for identifying high-level taxonomy (functional groups) using these autonomous sensors, as well as satellite ocean color sensors. While specialized instruments may be able to better resolve taxonomy (possibly to the species level), fluorescence/backscatter sensors and satellite data provide a much higher coverage and functional groups already contain precious information on major phytoplankton players. Moreover, while true measures of biodiversity require an extensive knowledge of the community, information on spatial heterogeneity of functional groups has been shown to be highly correlated to local diversity and can thus be considered as a good proxy for overall plankton biodiversity (Soccodato et al., 2016).

The goal of this feasibility study is to examine whether it is possible to derive reliable proxies for phytoplankton taxonomic groups such as diatoms, dinoflagellates, picoplankton, or haptophytes, from (1) autonomous fluorescence/backscatter sensors, and (2) satellite ocean color in the California Current. If this is proven feasible, reprocessing the MBARI AUVs dataset alone would increase our information on phytoplankton taxonomy by close to 2 orders of magnitude and provide full profiles. MODIS-based proxies would provide a synoptic 4-km data set of dominant phytoplankton groups with >80% monthly coverage. Combined, these products would provide an unprecedented view of phytoplankton ecology in the region. A preliminary investigation of Dorado proxies for diatoms and dinoflagellates based on fluorescence and bioluminescence sensors for the 2003-2008 time period by summer intern Alaina Smith is encouraging (see project 901103), and suggests that interesting discoveries could be made by deriving similar proxies from all AUVs. Moreover, in the future real-time processing of phytoplankton taxa proxies onboard AUVs would enable more advanced targeted sampling by providing the capability to differentiate between a diatom- and dinoflagellate-dominated patch, for example.

Reference cited:

Soccodato, A., d'Ovidio, F., Lévy, M., Jahn, O., Follows, M. J., & De Monte, S. (2016). Estimating planktonic diversity through spatial dominance patterns in a model ocean. Marine genomics, in press, doi:10.1016/j.margen.2016.04.015.

How does it relate/contribute to MBARI's mission and strategic plan?

Assuming the feasibility study is successful and phytoplankton taxonomic proxies can be derived from fluorescence/backscatter sensors and/or satellite, this would open several research paths tightly coupled with MBARI's research themes:

- *ocean visualization:* these proxies would provide information on the abundance and distribution of different phytoplankton groups, and support studies of biodiversity patterns in the California Current.
- *ecosystem processes:* phytoplankton taxonomy impacts higher trophic levels and the carbon cycle, thus expanded coverage of phytoplankton taxonomic information would provide valuable insight into ecosystems functioning.
- *exploration and discovery:* new analytical tools related to machine learning will be investigated and developed for this project.

Taken together, these advances would directly contribute to MBARI's overarching goal to "develop or adapt innovative technologies that allow researchers to identify and resolve important questions and advance our understanding of the ocean".

How does it relate/contribute to the Packard Foundation's Ocean Visions Grantmaking policies?
(See http://mww.mbari.org/oed/OED_Documents.htm#Packard for documents.)

The Packard Foundation's long-term goal as described in their U.S. West Coast strategic plan is to enhance the ability of U.S. West Coast marine ecosystem to sustain biodiversity and productivity. Because phytoplankton taxonomic proxies would provide invaluable information on phytoplankton high-level biodiversity patterns as well as provide means to monitor its status and change, this project contributes to the Packard Foundation goals.

Does it relate to the Technology Roadmap? If so, please describe how.
(See http://mww.mbari.org/oed/TechnologyRoadmap.pdf.)

This project directly supports several of MBARI's technological priorities (TR) put forward in the Technology Roadmap:

- *taking the laboratory into the ocean:* we are "developing new tools and techniques for identifying, quantifying and characterizing species" (TR quote).
- *enabling targeted sampling:* assuming it is possible to derive proxies from fluorescence/backscatter sensors, these could be computed in real-time onboard the AUVs. The current targeted sampling algorithms developed by Yanwu, notably patch tracking and peak capture, could be modified to specifically target a given type of organism.

Approach to the problem

This project is divided into 2 parts: (1) proxies from fluorescence/backscatter sensors, and (2) proxies from satellite. In both cases, the potential for proxies will be evaluated against *in situ* phytoplankton counts (BOG time series). These parts are independent but can inform each other: matching satellite and fluorescence/backscatter profiles should yield the same phytoplankton taxonomic group, which can be used to validate the proxies. The figure below summarizes the approach to the problem, then the approach for each part is detailed.

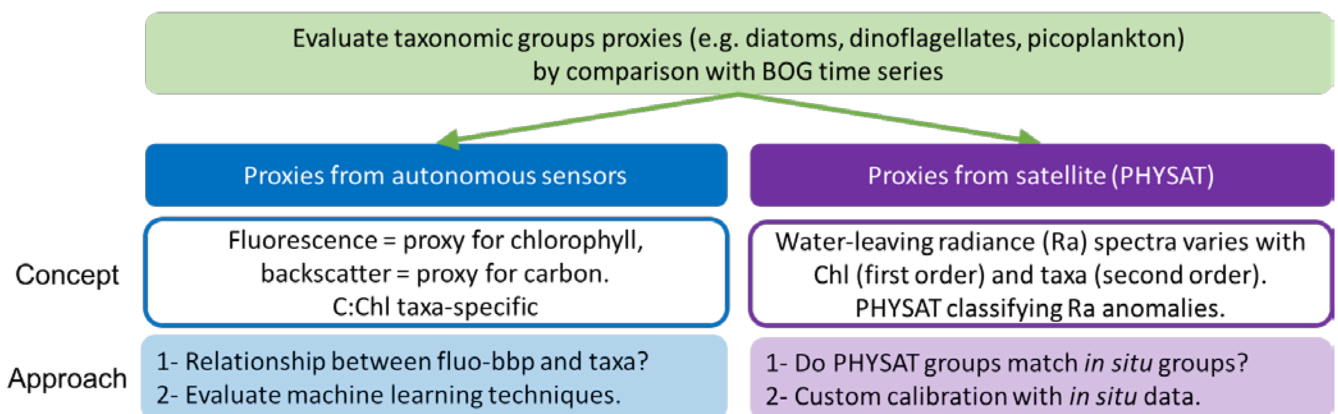

Figure 1: Approach to the problem. There is currently no proxy developed from autonomous sensors; the first step is to evaluate whether fluo-bpp is indeed related to phytoplankton taxa. Proxies are already available from satellite (PHYSAT outputs) and would need to be validated.

(1) Deriving proxies from fluorescence/backscatter sensors

<u>Background:</u> The theory that fluorescence-backscatter (hereafter fluo-bbp) signatures may be indicative of phytoplankton taxa is based on taxonomy-specific phytoplankton C:Chl: for example, diatoms have lower C:Chl than dinoflagellates. Fluorescence is a proxy for chlorophyll (Chl) and backscatter for carbon (C) so that scatter plots of the two properties may delineate communities in a manner similar to flow cytometry. In support of this concept, Whitmire et al. (2010) found linear backscatter-chlorophyll relationships with species-specific slopes from culture dilution series experiments. However, increasing

light or nutrient stress can significantly reduce cellular Chl and increase C:Chl, complicating the relationship. Photoacclimation in particular may explain >90% of the C:Chl variability at the global scale (Behrenfeld et al., 2016). Daytime fluorescence quenching near the surface will also be an issue and needs to be corrected (Sackmann et al., 2008), or surface daytime data discarded. The relationship between fluo-bbp and taxonomy is thus not trivial.

Results to date: Previous MBARI effort focused primarily on the Fall 2013 CANON dataset. We demonstrated that the fluo-bbp signatures are variable in time and space but coherent between multiple AUVs and gliders (Mike McCann, using STOQS), and that it was possible to “tag” sediments and populations based on their fluo-bbp signal (Monique Messié). Results from classification using unsupervised learning methods (Danelle Cline) suggested the use of concurrent plankton data may be required. A test using a type of neural network (Self-Organizing Maps, SOM) combined with population tags appears promising. These efforts have relied solely on fluo-bbp, including those using population tags that were artificially labelled A-F with no link to plankton taxa. There has been no attempt yet to look at available phytoplankton counts to see whether there is indeed a relationship between plankton taxa and fluo-bbp, except for work by MBARI collaborator Brad Penta (NRL) on the 2010 CANON experiment. He was able to derive proxies for dominance by diatoms vs. dinoflagellates from fluorescence/backscatter in agreement with concurrent HPLC samples; these proxies were however specific to the CANON 2010 dataset.

Approach: We propose to investigate relationships between (a) BOG phytoplankton counts at C1, M1 and M2 and (b) concurrent fluo and beam-c (another carbon proxy) datasets from the CTD/rosette, as well as fluo-bbp from the M1 mooring. Counts from gulper samples are also available for a few campaigns and can be directly compared with Dorado fluo-bbp. We would investigate the presence/absence and robustness of a relationship between phytoplankton taxa and fluo-bbp, as well as examine the potential of different machine learning techniques in capturing this relationship, for example SOM or mixed-membership models (eg. topic models).

(2) Evaluating proxies from satellite ocean color

Background: The main parameter derived from satellite ocean color is chlorophyll, a proxy for phytoplankton biomass. While satellite-measured normalized water leaving radiances vary with Chl at the first order, their second-order variations are characteristic of phytoplankton taxa. The reason is that normalized water leaving radiance anomalies (Ra) depend on specific phytoplankton absorption, colored dissolved organic matter absorption, and particles backscattering, which are taxa-specific for a given Chl concentration (Alvain et al., 2012).

Current status: The PHYSAT algorithm, developed by Séverine Alvain (France), relates Ra to the dominant phytoplankton group at each pixel based on a 2-step process (Ben Mustapha et al., 2013). First, the Ra were classified into 100 “neuron” groups using an automatic unsupervised classification approach (SOM). Then, these 100 neuron groups were labeled with phytoplankton taxa when possible (about half of the groups), based on concurrent pigments samples. Phytoplankton taxa currently identified by PHYSAT are diatoms, nanoeukaryotes, *Synechococcus*-like cyanobacteria, *Prochlorococcus* and *Phaeocystis*-like.

Approach: This second part of the project would evaluate the potential of PHYSAT in the California Current by comparison with BOG phytoplankton counts at two levels. First, the 5 plankton taxa identified by PHYSAT can be directly compared to *in situ* plankton counts (e.g., do PHYSAT "diatoms" capture *in situ* diatom blooms). Second, the original 100 neuron groups themselves (corresponding to anomalies in the satellite signal) can be mapped onto *in situ* phytoplankton taxa using a labeling technique similar to the one used by Ben Mustapha et al. (2013). It is possible that spectral signatures may be region-specific, and it might be possible to refine the labeling by identifying other / more detailed phytoplankton taxa beyond the 5 original PHYSAT taxonomic groups. Evaluating the potential of PHYSAT should be straightforward as the full MODIS dataset (100 groups and 5 taxa outputs) would be available as part of a collaboration with Séverine Alvain. This effort would be very timely as researchers from France (Hervé Demarcq) and Morocco (Aicha Berrada) are beginning to work with PHYSAT in the Peru and Northwest Africa upwelling systems, which could enable some interesting collaborations and intercomparisons.

References cited

Alvain, S., Loisel, H., & Dessailly, D. (2012). Theoretical analysis of ocean color radiances anomalies and implications for phytoplankton groups detection in case 1 waters. *Optics express*, 20(2), 1070-1083.

Behrenfeld, M. J., O'Malley, R. T., Boss, E. S., Westberry, T. K., Graff, J. R., Halsey, K. H., ... & Brown, M. B. (2015). Revaluating ocean warming impacts on global phytoplankton. *Nature Climate Change*, 6, 323–330.

Mustapha, Z. B., Alvain, S., Jamet, C., Loisel, H., & Dessailly, D. (2014). Automatic classification of water-leaving radiance anomalies from global SeaWiFS imagery: Application to the detection of phytoplankton groups in open ocean waters. *Remote Sensing of Environment*, *146*, 97-112.

Sackmann, B. S., Perry, M. J., & Eriksen, C. C. (2008). Seaglider observations of variability in daytime fluorescence quenching of chlorophyll-a in Northeastern Pacific coastal waters. *Biogeosciences Discussions*, *5*(4), 2839-2865.

Whitmire, A. L., Pegau, W. S., Karp-Boss, L., Boss, E., & Cowles, T. J. (2010). Spectral backscattering properties of marine phytoplankton cultures. *Optics Express*, *18*(14), 15073-15093.

Publications (from other projects over the last 5 years)

Peer-reviewed publications co/authored by the project team in the last 5 years:

2016

Kavanaugh, M.T., Oliver, M.J. Chavez, F. P., Letelier, R.M., Muller-Karger, F.E.,Doney, S.C. (2016) Seascapes as a new vernacular for pelagic ocean monitoring, management and conservation. ICES Journal of Marine Science, doi: 10.1093/icesjms/fsw086.

Kroeker, K.J., E. Sanford, J. M. Rose, C. A. Blanchette, F. Chan, F. P. Chavez, B. Gaylord, B. Helmuth, T. M. Hill, G. E. Hofmann, M. A. McManus, B. A. Menge, K. J. Nielsen, P. T. Raimondi, A. D. Russell, and L. Washburn (2016) Interacting environmental mosaics drive geographic variation in mussel performance and species interactions. Ecology Letters, doi: 10.1111/ele.12613

Radenac, M.-H., Léger, F., Messié, M., Dutrieux, P., Menkes, C. & Eldin, G. (2016). Wind-driven changes of surface current, temperature, and chlorophyll observed by satellites north of New Guinea. Journal of Geophysical Research, 121(4), 2231–2252, doi:10.1002/2015JC011438.

2015

Aylward, F.O, J. M. Eppley, J. M. Smith, F. P. Chavez, C. A. Scholin and E. F. DeLong (2015), Microbial community transcriptional networks are conserved in three domains at ocean basin scales. Proceedings of the National Academy of Sciences, 112 (17) 5443-5448; doi:10.1073/pnas.1502883112.

Hughes, B.B., M.D. Levey, M.C. Fountain, A. B. Carlisle, F. P. Chavez and M.G.Gleason (2015). Climate mediates threats to fish diversity and nursery function at the land-sea interface. Proceedings of the National Academy of Sciences USA. 112:8025-8030.

Messié, M., & Chavez, F. P. (2015). Seasonal regulation of primary production in eastern boundary upwelling systems. *Progress in Oceanography*, *134*, 1–18. http://doi.org/10.1016/j.pocean.2014.10.011

Miljutin, D., Miljutina, M., & Messié, M. (2015). Changes in abundance and community structure of nematodes from the abyssal polymetallic nodule field, Tropical Northeast Pacific. *Deep Sea Research Part I: Oceanographic Research Papers*, *106*, 126–135.http://doi.org/10.1016/j.dsr.2015.10.009

Pennington, J. T., M. Blum, and F. P. Chavez (2015). Seawater sampling by an autonomous underwater vehicle: “Gulper” sample validation for nitrate, chlorophyll, phytoplankton, and primary production. Limnology and Oceanography: Methods. DOI: 10.1002/lom3.10065.

Raanan, B. Y., Bellingham, J., Zhang, Y., Kieft, B., Stanway, M. J., McEwan, R., & Hobson, B. (2015). A real-time vertical plane flight anomaly detection system for a long range autonomous underwater vehicle (pp. 1–6). Presented at the Proceedings of the Marine Technology Society/Institute of Electrical and Electronics Engineers Oceans Conference.

Sevadjian, J. C., McPhee-Shaw, E. E., Raanan, B. Y., Cheriton, O. M., & Storlazzi, C. D. (2015). Vertical convergence of resuspended sediment and subducted phytoplankton to a persistent detached layer over the southern shelf of Monterey Bay, California. *Journal of Geophysical Research: Oceans*, *120*, 3462–3483. http://doi.org/10.1002/2015jc010785

Smith Jr, K., Messié, M., Sherman, A., Huffard, C., Hobson, B., Ruhl, H., & Boetius, A. (2015). Navigating the Uncertain Future of Global Oceanic Time Series. *EOS*, *96*. http://doi.org/10.1029/2015eo038095

Smith, J.M., J. Damashek, F. P. Chavez and C.A. Francis (2015) Factors influencing nitrification rates and the abundance and transcriptional activity of ammonia-oxidizing microorganisms in the dark northeast Pacific Ocean. Limnology and Oceanography, doi: 10.1002/lno.10235.

Sudek,S, Everroad, RC, Gehman A, Smith JM, Poirier C, Chavez FP & AZ Worden (2015). Cyanobacterial distributions along a physico-chemical gradient in the northeastern Pacific Ocean. Environmental Microbiology. doi: 10.1111/1462-2920.12742.

Xu, Y., K.A. Rose, F. Chai, F. P. Chavez, and P. Ayon (2015): Does spatial variation in environmental conditions affect recruitment? A study using a 3-D model of Peruvian anchovy. Progress in Oceanography, 2015, doi:10.1016/j.pocean.2015.04.013

2014

Berdalet, E, MA McManus, ON Ross, H Burchard, FP Chavez, JS Jaffe, IR Jenkinson, R Kudela, I Lips, U Lips, A Lucas, D Rivas, MC Ruiz-de la Torre, J Ryan, JM Sullivan, H Yamazaki (2014) Understanding harmful algae in stratified systems: Review of progress and future directions. Deep Sea Research II 101: 4-20. doi: 10.1016/j.dsr2.2013.09.042.

Cheriton, O. M., McPhee-Shaw, E. E., Storlazzi, C. D., Rosenberger, K. J., Shaw, W. J., & Raanan, B. Y. (2014). Upwelling rebound, ephemeral secondary pycnoclines, and the creation of a near-bottom wave guide over the Monterey Bay continental shelf. *Geophysical Research Letters*, *41*(23), 8503-8511.

Doney, S., A. Rosenburg, M. Alexander, F. Chavez, C.D. Harvell, G. Hoffman, M. Orback, and M. Rucklehaus (2014). Chapter 24. Oceans and Marine Resources. Climate Change Impacts in the United States: The Third National Climate Assessment, Melillo, Jerry M., Terese (T.C.) Richmond, and Gary W. Yohe, Eds., U.S. Global Change Research Program, 557-578. doi:10.7930/J0RF5RZW.

Fiechter, J., E. N. Curchitser, C. A. Edwards, F. Chai, N. L. Goebel, and F. P. Chavez (2014) Spatiotemporal variability of air-sea CO2 exchange in the California Current. Global Biogeochemical Cycles, 28, 371-385.

Lin, G., F. Chai, P. Xiu, Y. Liu, S. Rao, and F.P. Chavez (2014) Seasonal dynamics of physical and biological processes in the central California Current System: A modeling study. Ocean Dynamics. http://dx.doi.org/10.1007/s10236-014-0721-x.

Muller-Karger, F.E., M.T. Kavanaugh, E. Montes, W.M. Balch, M. Breitbart, F.P. Chavez, S.C. Doney, E.M. Johns, R.M. Letelier, M.W. Lomas, H.M. Sosik, and A.E. White (2014) A framework for a marine biodiversity observing network within changing continental shelf seascapes. Oceanography 27(2): 18–23, http://dx.doi.org/10.5670/oceanog.2014.56.

Smith, J. M., K. L. Casciotti, F. P. Chavez and C. A Francis (2014) Differential contributions of archaeal ammonia oxidizer ecotypes to nitrification in coastal surface waters. The ISME Journal. doi:10.1038/ismej.2014.11.

Smith, J.M., F.P. Chavez and C.A. Francis (2014) Ammonium uptake by phytoplankton regulates nitrification in the sunlit ocean. PLoS ONE 9(9):e108173. doi:10.1371/journal.pone.0108173.

Sutton, A.J., R. A. Feely, C. L. Sabine, M. J. McPhaden, T. Takahashi, F. P. Chavez, G. E. Friederich, and J. T. Mathis (2014) Natural variability and anthropogenic change in equatorial Pacific surface ocean pCO2 and pH. Global Biogeochemical Cycles, 28, 131–145. doi:10.1002/2013GB004679.

Ryan, J. P., McManus, M. A., Kudela, R. M., Artigas, M. L., Bellingham, J. G., Chavez, F. P., … Zhang, Y. (2014). Boundary influences on HAB phytoplankton ecology in a stratification-enhanced upwelling shadow. *Deep Sea Research Part II: Topical Studies in Oceanography*, *101*, 63–79. http://doi.org/10.1016/j.dsr2.2013.01.017

2013

Bakker, D. C. E., Pfeil, B., Smith, K. , Hankin, S. , Olsen, A. , Alin, S. R. , Cosca, C. , Harasawa, S. , Kozyr, A. , Nojiri, Y. , O'Brien, K. M. , Schuster, U. , Telszewski, M. , Tilbrook, B. , Wada, C. , Akl, J. , Barbero, L. , Bates, N. , Boutin, J. , Cai, W. J. , Castle, R. D. , Chavez, F. P. , et al.. (2013): An update to the Surface Ocean CO2 Atlas (SOCAT version 2) , Earth System Science Data Discussions, 6 (2), pp. 465-512 . doi: 10.5194/essdd-6-465-2013.

Fishman, J., L. T. Iraci, J. Al-Saadi, K. Chance, F. Chavez, et al. 2013: The United States' Next Generation Of Atmospheric Composition And Coastal Ecosystem Measurements NASA's Geostationary Coastal And Air Pollution Events (GEO-CAPE) Mission. Bulletin of the American Meteorological Society 93, 1547–1566. doi:http://dx.doi.org/10.1175/BAMS-D-11-00201.1

Gierach, M. M., Messié, M., Lee, T., Karnauskas, K. B., & Radenac, M.-H. (2013). Biophysical responses near equatorial islands in the Western Pacific Ocean during El Niño/La Niña transitions. *Geophysical Research Letters*, *40*, 5473–5479.http://doi.org/10.1002/2013GL057828

Gomes, K., Cline, D., Edgington, D., Godin, M., Maughan, T., McCann, M., … Rajan, K. (2013). ODSS: A decision support system for ocean exploration (p. 12 pp.). Presented at the Proceedings of the Workshop on Data-Driven Decision Guidance and Support Systems (DGSS), 29th Institute of Electrical and Electronics Engineers International Conference on Data Engineering 2013.

Messié, M., & Chavez, F. P. (2013). Physical-biological synchrony in the global ocean associated with recent variability in the central and western equatorial Pacific. *Journal of Geophysical Research*, *118*, 3782–3794. http://doi.org/10.1002/jgrc.20278

Ottesen, E. A., C. R. Young, J. M. Eppley, J. P. Ryan, F. P. Chavez, C. A. Scholin, and E. F. DeLong (2013) Pattern and synchrony of gene expression among sympatric marine microbial populations. PNAS www.pnas.org/cgi/doi/10.1073/pnas.1222099110.

Radenac, M.-H., Messié, M., Léger, F., & Bosc, C. (2013). A very oligotrophic zone observed from space in the equatorial Pacific warm pool. *Remote Sensing of Environment*, *134*, 224–233. http://doi.org/10.1016/j.rse.2013.03.007

Sabine, C. L., Hankin, S., Koyuk, H., Bakker, D. C. E., Pfeil, B., Olsen, A., Metzl, N., Kozyr, A., Fassbender, A., Manke, A., Malczyk, J., Akl, J., Alin, S. R., Bellerby, R. G. J., Borges, A., Boutin, J., Brown, P. J., Cai, W.-J., Chavez, F. P., et al. (2013) Surface Ocean CO2 Atlas (SOCAT) gridded data products, Earth Syst. Sci. Data, 5, 145-153, doi:10.5194/essd-5-145-2013.

Santora, J. A., Sydeman, W. J., Messié, M., Chai, F., Chao, Y., Thompson, S. A., … Chavez, F. P. (2013). Triple check: Observations verify structural realism of an ocean ecosystem model. *Geophysical Research Letters*, *40*, 1367–1372.http://doi.org/10.1002/grl.50312

Wells, B.K., I. D. Schroeder, J. A. Santora, E. L. Hazen, S. J. Bograd, E. Bjorkstedt, V. J. Loeb, S. McClatchie, E. D. Weber, W. Watson, A. R. Thompson, W. Peterson, R. D. Brodeur, J. Harding, J. Field, K. Sakuma, S. Hayes, N. Mantua, W. J. Sydeman, M. Losekoot, S. A. Thompson, J. Largier, S. Y. Kim, F. P. Chavez, et al. (2013) State of the California Current 2012-2013: No such thing as an average year. CalCOFI Reports, 54, 37-71.

Xu Y., F. Chai, K. A. Rose, M. Ñiquen C. and F. P. Chavez (2013) Environmental influences on the interannual variation and spatial distribution of Peruvian anchovy (Engraulis ringens) population dynamics from 1991 to 2007: a three-dimensional modeling study. Ecological Modelling,http://dx.doi.org/10.1016/j.ecolmodel.2013.01.009.

2012

Bjorkstedt Eric, Ralf Goericke, Sam McClatchie, Ed Weber, William Watson,Nancy Lo, William Peterson, Richard Brodeur, Steven Bograd, Toby Auth, Jennifer Fisher, Cheryl Morgan, Jay Peterson, Reginaldo Durazo, Gilberto Gaxiola-Castro, Bertha Lavaniegos, Francisco Chavez, et al. (2012) State of the California Current 2011-2012: Ecosystems respond to local forcing as La Niña wavers and wanes. CalCOFI Report, 53, 41-76.

Buitenhuis, E.T., W. K. W. Li, D. Vaulot, M. W. Lomas, M. Landry, F. Partensky, D. M. Karl, O. Ulloa, L. Campbell, S. Jacquet, F. Lantoine, F. Chavez, D. Macias, M. Gosselin, and G. B. McManus (2012) Picophytoplankton biomass distribution in the global ocean. Earth System Science Data Discussions, 5, 221-242.

Chavez, F. P. (2012) Climate Change and Marine Ecosystems. Proceedings of the National Academy of Sciences, 47, 19045-19046, doi: 10.1073/pnas.1217112109.

Das, J., Py, F., Maughan, T., O'Reilly, T., Messié, M., Ryan, J., … Rajan, K. (2012). Coordinated sampling of dynamic oceanographic features with underwater vehicles and drifters. *International Journal of Robotics Research*, *31*, 626–646.http://doi.org/10.1177/0278364912440736

Das, J., Py, F., Maughan, T., O'Reilly, T., Messié, M., Ryan, J., … Rajan, K. (2012). Simultaneous tracking and sampling of dynamic oceanographic features with AUVs and drifters. *Usc.academia.edu*, 3 pp.

Frolov, S., J.P. Ryan and F.P. Chavez (2012) Predicting euphotic-depth-integrated chlorophyll-a from discrete-depth and satellite-observable chlorophyll-a off central California. Journal of Geophysical Research, 117, C05042, doi:10.1029/2011JC007322

Gomez, F., A. Montecinos, S. Hormazabal, L.A. Cubillos M. Correa-Ramirez, and F.P. Chavez (2012) Impact of spring upwelling variability off southern-central Chile on common sardine (Strangomera bentincki) recruitment. Fisheries Oceanography, 21, 405-414, DOI: 10.1111/j.1365-2419.2012.00632.x.

Hales, B., P.G. Strutton, Saraceno, Letelier, Takahashi, Sabine, Feely and F. Chavez (2012). Satellite-based prediction of pCO2 in coastal waters of the eastern North Pacific. Progress in Oceanography, 103, 1-15, doi: 10.1016/j.poocean.2012.03.001

Jolliff, J.K., R.W. Gould Jr., B. Penta, W.J. Teague, S. deRada, F.P. Chavez, R.A. Arnone (2012) Water mass bio-optical properties in the Monterey Bay region: Fluorescence-based inference of shifts in phytoplankton photophysiology Journal of Geophysical Research, 117, DOI: 10.1029/2011JC007568.

Kemp, M., Palanza, M., Skibski, C., Ormsby, J., & Estaphan, M. (2012, September). Persistence at full ocean depth. In *2012 IEEE/OES Autonomous Underwater Vehicles (AUV)* (pp. 1-7). IEEE.

Messié, M., & Chavez, F. P. (2012). A global analysis of ENSO synchrony: The oceans' biological response to physical forcing. *Journal of Geophyiscal Research*, *117*. http://doi.org/10.1029/2012JC007938

Pfeil, B. , Olsen, A. , Bakker, D. C. E. , Hankin, S. , Koyuk, H. , Kozyr, A. , Malczyk, J. , Manke, A. , Metzl, N. , Sabine, C. L. , Akl, J. , Alin, S. R. , Bellerby, R. G. J. , Borges, A. , Boutin, J. , Brown, P. J. , Cai, W. J. , Chavez, F. P. , et al. (2012) A uniform, quality controlled Surface Ocean CO2 Atlas (SOCAT), Earth System Science Data Discussions, Copernicus Publications, 5 (2), pp. 735-780. doi: 10.5194/essdd-5-735-201.

Sabine, C. L. , Hankin, S. , Koyuk, H. , Bakker, D. C. E. , Pfeil, B. , Olsen, A. , Metzl, N. , Kozyr, A. , Fassbender, A. , Manke, A. , Malczyk, J. , Akl, J. , Alin, S. R. , Bellerby, R. G. J. , Borges, A. , Boutin, J. , Brown, P. J. , Cai, W. J. , Chavez, F. P. , et al. (2012) Surface Ocean CO2 Atlas (SOCAT) gridded data products , Earth System Science Data Discussions, Copernicus Publications, 5 (2), pp. 781-804 . doi: 10.5194/essdd-5-781-2012.

2011

Bertrand , A., Chaigneau, A., Peraltilla S., Ledesma J. , Graco M., Monetti, F. and F.P. Chavez (2011) Oxygen: A Fundamental Property Regulating Pelagic Ecosystem Structure in the Coastal Southeastern Tropical Pacific. PLoS ONE 6(12): e29558. doi:10.1371/journal.pone.0029558

Bjorkstedt. E., R. Goericke, S. McClatchie, F.P. Chavez and 27 other authors (2011) State of the California Current 2010-2011. Regionally variable responses to a strong (but fleeting?) La Niña. CalCOFI Reports 52: 36-68.

Castro, C.G., C.A. Collins, J.T. Pennington, D. Zuniga and F.P. Chavez (2011) Spatial distribution and offshore export of total organic carbon along the eastern boundary of the Subtropical North Pacific." Advances in Oceanography and Limnology 2(2): 93-106. Doi: 10.1080/19475721.2011.626941.

Chavez, F. P., Messié, M., & Pennington, J. T. (2011). Marine primary production in relation to climate variability and change. *Annual Review of Marine Science*, *3*, 227–260. http://doi.org/10.1146/annurev.marine.010908.163917

Das, J., Maughan, T., McCann, M., Godin, M., O'Reilly, T., Messié, M., … Rajan, K. (2011). Towards mixed-initiative, multi-robot field experiments: Design, deployment, and lessons learned. Presented at the Proceedings of the International Conference on Intelligent Robots and Systems.

Field, D. B., F. P. Chavez, C. B. Lange and P. E. Smith (2011) Variations in fisheries and complex ocean environments. In Shifting Baselines. J.B.C. Jackson, K.E. Alexanfer and E. Sala (eds.) Island Press, Washington, DC.

Kemp, M., Richardson, B., Love, L., Rodgers, T., Nelson, K., & Dixon, T. (2011). Analysis of manipulator arm steadiness during ship-hull mine neutralization by a UUV. *OCEANS 2011*, 1-5.

Messié, M., & Chavez, F. P. (2011). Global modes of sea surface temperature variability in relation to regional climate indices. *Journal of Climate*, *24*, 4314–4331. http://doi.org/10.1175/2011JCLI3941.1

Paerl, R. W., Turk, K. A., Beinart, R. A., Chavez, F. P. and Zehr, J. P. 2011. Seasonal change in the abundance of Synechococcus and multiple distinct phylotypes in Monterey Bay determined by rbcL and narB quantitative PCR. Environmental Microbiology. doi: 10.1111/j.1462-2920.2011.02594.x

b) Continuing Projects

N/A

c) New and Continuing Projects: Plans for 2017

Research Activities

The overall plan for research activities was described above (Figure 1). Here we detail the different steps associated with each part, and evaluation metrics for answering the question "Can we derive reliable proxies for phytoplankton taxonomic groups from autonomous sensors and/or satellite?". In the following we are using diatoms and autotrophic dinoflagellates (hereafter adinos) as examples for the phytoplankton groups; the method will be expanded to include additional phytoplankton groups such as monitored by the BOG group.
We did note the MT suggestion to collaborate with Ed Delong on this project and agree that expanding this work to the tropical areas would be very interesting. However, this 2017 feasibility study is intended to be a very small and focused effort, and we believe it will be more efficient to first assess the potential of the method in Monterey Bay before expanding it to other regions. We will consider such collaborations in a subsequent proposal if the feasibility study is successful.
We will collaborate with Brad Penta (NRL) on the fluorescence/backscatter question since he has experience working on this topic and was able to derive proxies for dominance by diatoms vs. dinoflagellates from fluorescence/backscatter for the 2010 CANON experiment in Monterey Bay. While these proxies were specific to one campaign and thus cannot be readily applied to the AUVs dataset, his experience on the subject will be very valuable.
We will also collaborate with Yanwu Zhang & LRAUV team (project 901513) as they will be simultaneously working on implementing very simple proxies calculations in LRAUV. Even though we do not anticipate developing reliable proxies by 2017, our work may identify variables of interest for early LRAUV testing.

(1) Deriving proxies from fluorescence/backscatter sensors

The first step will be to generate datasets matching phytoplankton counts with corresponding fluorescence/backscatter and fluorescence/beam-c profiles ("combined dataset"). These datasets will consist of:

- C1, M1 and M2 counts and corresponding CTD profiles (fluorescence/beam-c). We will test the Sackmann et al. (2008) quenching correction.
- M1 counts and corresponding mooring fluorescence/backscatter data (both closest in time and at nighttime, to avoid quenching effects).

If time allows, we will add gulper counts (available for a few campaigns) and corresponding Dorado profiles. We may also attempt a general comparison of Dorado fluorescence/backscatter with diatoms/adinos proxies derived from bioluminescence as part of project 901103.

These datasets will be provided to Ben Raanan for testing of topic modeling methods (see below). Additional variables will be included, such as PAR and temperature, that may be valuable in correcting for the photoacclimation and nutrient stress effects.

Then the relationship between phytoplankton taxonomy and fluo-bbp or fluo-beamc (hereafter fluo-bbp for simplification) will be explored. Methods to test the existence of a relationship can include:

- Partition the fluo-bbp data from the combined dataset into "diatoms-dominated", "adinos-dominated" (etc) situations and test whether these are statistically different from each other.
- Look at the seasonal and interannual variations of the entire fluo-bbp dataset and test whether there are significant differences between periods known to be dominated by adinos vs. diatoms, as identified by the BOG time series (e.g. adinos: fall, 2004-2006; diatoms: spring, other years).
- Compute and remove the photoacclimation effect based on an adaptation of the Beherenfeld et al. (2015) method and look at the residuals.

Last, machine learning techniques will be tested to evaluate their potential at capturing the relationship between taxonomy and fluo-bbp, and at defining proxies. To assess the potential of machine learning on taxa classification, the input data will be split into a training set, a validation set, and a test set, and the classification performance will be measured using cross-validation. Two methods will be considered:

- *Self Organizing Maps (SOM):* SOM is an unsupervised learning method that trains a neuron network on the predictors, here fluo-bbp, effectively clustering the predictors. Then matching data, here BOG counts, are used to label the neurons. This "neurons labelled map" represents a transfer function between the predictors and the taxonomic groups, and can be applied to new datasets. A preliminary investigation of the method on artificial tags from CANON 2013 was promising.
- *Topic modeling:* Topic modeling is a generative mixed-membership unsupervised learning method that identifies patterns – aka topics – in the data and represents observations as probability distributions over patterns. Once trained, each topic should map one-to-one with the fluo-bbp pattern of the individual taxa, and each sample/observation will be assigned a probability distribution that should correspond to the relative concentration of the taxonomic groups.

The primary objective of the machine learning part is to evaluate the potential for these two techniques (see "(3) Evaluation metrics" below). Developing reliable proxies would require more time, testing and calibration and is not the goal of the feasibility study. However, some preliminary proxies can be quickly derived by the team who is already familiar with both methods, and evaluated – hence the relatively short duration of this feasibility study. The development of reliable proxies would take place in a subsequent proposal.

(2) Evaluating proxies from satellite ocean color

Similarly, a database matching *in situ* phytoplankton counts with the concurrent PHYSAT output (both phytoplankton dominant group and neuron group) will be generated. Because PHYSAT only captures the dominant species, we will restrict the BOG samples to those clearly dominated by one functional group.

As a first step, PHYSAT functional groups will be validated. While the data currently available to us are very preliminary (the algorithm was developed for SeaWiFS and is currently being ported to MODIS), a look at spatial and seasonal patterns suggests that but these may not readily work in the region, to be confirmed with a subsequent release planned for later this year.

The most important part of this work will use the neurons outputs (MODIS radiance anomalies classified into 100 groups using the SOM technique). We will apply a neuron-labelling method similar to the one used in PHYSAT, where neurons were labelled with functional groups based on concurrent HPLC samples, using a random subset of BOG samples. The rationale is that plankton optical properties may be region-specific, and that plankton groups not yet captured by PHYSAT (such as dinoflagellates) may be related to certain neurons. The resulting labelled neurons map will be used to produce a new, custom product of satellite-based phytoplankton taxonomic groups that will be validated against the remaining BOG counts.

(3) Evaluation metrics

Several relationships will be tested and quantified, that will together define the success of the feasibility study:

a- There are significant differences in fluo-bbp between diatoms-dominated and dinoflagellates-dominated situations (possibly after removal of photoacclimation, quenching and nutrient effects).

b- Preliminary proxies defined on a random subset of *in situ* counts accurately predict the remaining *in situ* counts. The same metrics will be applied to proxies defined from SOM and topic modeling on fluo-bpp, PHYSAT “official” dominant phytoplankton groups and PHYSAT “customized” dominant phytoplankton groups.

c- If time allows: proxies reproduce known seasonal and interannual patterns. This will likely only be tackled after reliable proxies are derived (i.e. b- is successful) which may not happen in 2017.

Engineering/technology development

None.

Field programs (detail ship requirements/deployments/diver requirements)

None.

Isotope Usage

Project teams must provide details regarding any planned isotope use on shore, on ships, in the Sanctuary, etc. Please review the guidance on isotope usage at http://www.mbari.org/at-sea/cruise-planning/hazardous-materials/ and policy at *http://mww.mbari.org/safety/SafetyManual/Isotope_Use_Policy.pdf, and then provide details on plans below.*

None.

Specific deliverables

The main deliverable of this project is the result of the "evaluation metrics" described above.

Milestones (see personnel classification list)

N/A

2017 Project Start Date: 01/01/2017	**Milestone 1**	**Milestone 2**	**Milestone 3**	**Milestone 4**	**Milestone 5**	**Total Resources**[4] for each resource category for all milestones listed here
Milestone Date						
Milestone Description						
EE						
SE						
ME						
EE Tech						
SE Tech						
ME Tech						
Machine Shop						
Divers						

Budget justification

N/A (no request besides Ben's and Monique's time).

[4] This number should be the same number you entered on the resource requests for this particular resource.

Project team

Monique Messié (PI) – datasets preparation, SOM, PHYSAT evaluation (300h)
Ben Raanan – topic modeling (160h)
Mathieu Kemp
Francisco Chavez

Project lab and office space requirements
In addition to any changes to current space usage, please also include anticipated office and lab needs for students, postdocs, collaborators, etc.

None.

Data storage requirements (see Todd Ruston for assistance)

None.

Data availability outside project team, uniqueness

Regarding proxies from autonomous sensors, the goal of this feasibility study is to evaluate whether it is possible to generate phytoplankton proxies from fluo-bbp. We do not anticipate reaching a point where reliable proxies are derived and can be made available to other teams by the end of 2017. Developing and refining proxies, as well as making them available to other teams (notably for targeted sampling) would be the main goal of a possible subsequent proposal.
Regarding PHYSAT, the outputs are not publicly available and will be shared with us as part of a collaboration with S. Alvain, so that there is no plan to make them available outside the project team in 2017.

This work is related to a study of Dorado proxies conducted under 901103 by M. Messié based on fluorescence and bioluminescence sensors. This feasibility study remains however unique because of the much larger number of fluo-bbp profiles (all LRAUVs, ship CTD/rosette, gliders). It is possible that methods developed here may be of use to the fluorescence/bioluminescence project and vice versa. MBARI is uniquely positioned to tackle such a study because we possibly hold the largest database of concurrent plankton counts and fluorescence/backscatter profiles in the world.

R/V *Paragon* Requests
If the project team will be requesting time for the R/V Paragon, please specify the operators in the group and provide estimates of their availability to operate the vessel.

None.

Export Regulation
Does the project team anticipate any foreign collaborations or international travel? If so, please identify collaborators by name, institution, and country.

Foreign collaborator: Séverine Alvain, LOG/CNRS, France.

Does the project team plan or anticipate export controlled purchases, shipments, or transfers of items (e.g., research instruments, proprietary data, hand-carried tools or samples) outside the U.S. temporarily or permanently? If yes, please identify country or countries and indicate whether these items will remain under MBARI control while abroad or controlled by others. Will the items be exported to more than one country? Please include foreign collaborator information as described above.

None.

Monterey Bay National Marine Sanctuary (MBNMS)

General Permit
Projects teams need to indicate if any work will be performed that is either: (1) currently authorized under our general permit; or (2) will require that we obtain authorization. The current permit can be viewed at http://mww.mbari.org/pco/permits/mbnms-2015-002_feb2015.pdf. Please see Mandy Allen for questions.

Is a permit with MBNMS required for your work?

No.

SESAs (General map)
*Projects will need to indicate whether or not work will be conducted within one of the Monterey Bay National Marine Sanctuary's Ecologically Significant Areas (SESAs). An interactive map can be found on the Sanctuary's SIMoN webpages at http://sanctuarysimon.org/maps/sesa/. (*Note: It may take a moment for the page to load, so please be patient. Also, if prompted for a login, hit "Cancel" and then you will be able to access page.*) To determine if your work will occur in a SESA, specific coordinates can be entered using the locator function (click on the red "target" button in the toolbar at the top of the map). If your work will be conducted in one of these areas, please indicate which area, by number.*

None.

Authorizations/Permits from other Agencies or Organizations
Is the project team aware of any other permits or authorizations that will be necessary to perform the proposed work? If so, please list below along with status of authorizations (obtained, pending, need to apply, etc.). Please see Mandy Allen if you have questions.

None.

Active Acoustics
We are compiling information on active acoustic use at MBARI. If the project involves using active acoustic instrumentation, please describe plans.

None.

Risk matrix (development and infrastructure projects only)

	Technical Risk	Organizational Risk	Tech Transfer Risk	Labor Estimate Risk
High				
Medium				
Low				