ABSTRACT

Title of Dissertation: CLIMATE-SCALE CHLOROPHYLL PATTERNS IN THE TROPICAL PACIFIC FROM A 51 YEAR STATISTICAL RECONSTRUCTION

Stephanie Schollaert Uz, Doctor of Philosophy, 2015

Dissertation directed by: Professor Antonio J. Busalacchi
Department of Atmospheric and Oceanic Sciences

Understanding large, slow biological changes in the oceans has been hindered by a lack of spatial coverage by direct measurements and a lack of temporal coverage by satellite remote-sensing observations. Global ocean surface chlorophyll, a proxy for phytoplankton standing stock, has been derived from satellites for over a decade. With these measurements, the strong connection between ocean physics and biology has become clear and provided new insights about what drives seasonal and interannual biological processes. At longer time scales, however, there are many unanswered questions about the variability of phytoplankton in the ocean that plays a critical role in the carbon cycle as well as the marine food web. Statistical reconstructions have been used by others to extend physical climate variables in space and time. Taking advantage of the fact that physical forcing has been found to be the primary driver of biological primary production in the tropical Pacific, especially during El Niño, the most closely

correlated physical variables are used as predictors in a statistical reconstruction to extend monthly chlorophyll anomalies from just over a decade to just over five decades between 1958-2008. The reconstructed chlorophyll is evaluated through leave-one-out-cross-validation, compared to several independent data sets: *in situ* samples, another ocean color satellite data set, model output from a dynamic, fully-coupled ocean circulation-biogeochemistry model. Highest skill in the tropical Pacific reconstruction is away from the coast and within 10^{o} of the equator, including areas known as Niño 3/3.4/4. Over the half-century of chlorophyll anomalies, the most dominant climate pattern apparent in the reconstruction is associated with the interannual El Niño followed by the Pacific Decadal Oscillation. Biological distinctions emerged between the east Pacific El Niño events and those that only extend to the central Pacific. Chlorophyll anomalies were compared between regimes to ascribe physical forcing mechanisms. While the overall patterns were consistent with what is known about the impact of ENSO on biology, with the PDO primarily serving to amplify or damp ENSO, a narrow equatorial band consistently displayed an inverse response to the rest of the equatorial cold tongue: lower values during the PDO cool phase between 1958-1976, higher values during the PDO warm phase between 1977-1995. A likely explanation for this anomaly is linked to variability in the depth of the Equatorial Undercurrent that transports iron to the high-nutrient, low-chlorophyll east Pacific. These and other ideas are explored to demonstrate the feasibility and utility of reconstructing ocean color chlorophyll to address open questions about large-scale, low frequency primary production that forms the base of the marine food web and plays an important role in Earth's climate system.

CLIMATE-SCALE CHLOROPHYLL PATTERNS IN THE TROPICAL PACIFIC FROM A 51 YEAR STATISTICAL RECONSTRUCTION

by

Stephanie Schollaert Uz

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:

Professor Antonio J. Busalacchi, Chair
Professor James A. Carton
Professor Kayo Ide
Dr. Christopher W. Brown
Dr. Thomas M. Smith
Professor Michael N. Evans, Dean's Representative

Acknowledgements

Many people have helped or encouraged me along the way. First of all, I have appreciated all of the insights I have gained about the tropical Pacific from my advisor, Tony Busalacchi. This study was much improved by his critical thinking, sharp focus, and enviable efficiency. Tom Smith was generous with his fortran codes and his scholarly advice, which enabled the reconstructions. Chris Brown, Jim Carton, Kayo Ide and Mike Evans served on my committee, shared their expertise and gave detailed suggestions that strengthened the interpretation of this study. Wendy Wang provided the output of her fully-coupled biogeochemical model to validate the reconstruction. Phil Arkin wisely steered me toward completing my doctorate. Eric Hackert has been a fellow graduate student, sharing helpful references and background information about the region and model. Before my time at UMD, I had two notable mentors at the University of Rhode Island's Graduate School of Oceanography: Tom Rossby, whose curiosity about physical processes in the ocean and his tenacity to measure them with floats, ships of opportunity and other sensors was inspirational. Jim Yoder introduced me to ocean color remote sensing and its potential for synoptic scale biological oceanography. And finally, my family has been incredibly supportive and encouraging through this process. My mother and father set great examples to follow my interests and have always been there for me and helpful however they could be. Unfortunately my father did not live to see the end of my doctoral studies, but his spirit of curiosity in the natural world lives on. My boys, Andrew, Henry and Oliver, all stepped up in their own ways and remind me daily about the big picture and why it all matters. My husband, Mete, has given me unflagging moral support as both a sounding board and honest critic. He is a true partner and best friend.

Table of Contents

List of Tables

List of Figures

Figure 1.5 CHL-SST combined EOF time amplitude (black, left axis) with climate indices (red, right axis) for CZCS (1979-1983) and SeaWiFS subsets (1998-2002) over a) Pacific Ocean, r_{PDO}=0.66; b) Indian Ocean, r_{PDO}=0.6; c) Atlantic Ocean, r_{AMO}=0.75. Figure from Martinez et al. [2009].

Figure 1.6 Mean winter (DJF) SST (left) and CHL (right) variation with longitude averaged between 1°N-1°S for neutral (black), Central Pacific El Niño (red), Eastern Pacific El Niño (blue) phases. Figure adapted from Turk et al. [2011].

Figure 2.1 NODC Ocean Station CHL casts per 1° bin collected since 1959 and archived within the World Data Center for Oceanography (previously World Ocean Atlas).

Figure 2.2 CZCS observations. Mission average number of observations per bin (top), mission standard deviation of the number of observations per bin (middle), and total observations per month over the entire mission (bottom). Note that the mission was primarily coastal and focused on US territorial waters, thus largely ignoring the open ocean and other regions outside of the US domain.

Figure 2.3 Average number of valid data each month for SeaWiFS and MODIS/Aqua missions.

Figure 2.4 SeaWiFS observations in each bin. Mission average number of observations per bin (top) and standard deviation of the number of observations per bin over the mission (bottom).

Figure 2.5 CHL average time series over extra-tropical (left column) and tropical regions for SeaWiFS (—) and MODIS/Aqua (----). Correlation coefficients for overlapping times are annotated.

Figure 2.6 CHL deseasoned anomalies for all 5^o bins: MODIS/Aqua plotted against SeaWiFS, with their regression line superimposed. The dotted diagonal line demonstrates a slope of 1.

Figure 2.7 Distribution of SeaWiFS/MODIS/Aqua CHL: slope (top) and intercept (bottom).

Figure 2.8 Decorrelation scales for monthly, 5^o SeaWiFS chlorophyll averaged over three regions (subpolar, subtropical, equatorial): temporal decorrelations (top); zonal decorrelations (bottom).

Figure 2.9 Monthly deseasoned CHL anomaly standard deviations between 1998-2008, with calculations performed on log-transformed CHL.

Figure 2.10 Temporal cross-correlations between monthly anomalies of CHL and SST (top) and between monthly anomalies of CHL and MLD (bottom) for the period 9/1997-2008.

Figure 2.11 Cross-correlations (r) between monthly anomalies of wind speed and CHL at 25km resolution for the period 11/1996-2008. Black curves indicate r = 0. Western

equatorial Pacific and eastern equatorial Pacific are delineated. Figure adapted from Kahru et al. [2010].

Figure 2.12 Zonal velocity (cm/s) for a transect at 140^{o}W from Ocean Data View.

Figure 2.13 Vertical, zonal distribution of CHL on the equator at the eastern edge of the warm pool collected by the Flupac cruise in September-October 1994 and contoured at 0.05 mg m-3 intervals. Figure adapted from Le Borgne et al. [2002].

Figure 2.14 Top: monthly averaged CHL at 2^{o} where ◊ is roughly the western edge of equatorial upwelling and * marks the eastern edge of the warm pool. Bottom: monthly values at each spot.

Figure 2.15 Standard deviation of deseasoned monthly, 2^{o} CHL anomalies. As in Figure 2.14, ◊ is toward the western edge of the equatorial upwelling and * is near the eastern edge of the warm pool. The bottom panel shows the timeseries for these two spots as a dashed line and *'s, respectively.

Figure 2.16 Interannual CHL variance (top), seasonal CHL variance (middle), estimated error (bottom) is the difference between the observed autocorrelation at zero lag and a Gaussian estimate.

Figure 2.17 Signal-to-noise ratio (SNR) where signal is defined as the combination of interannual and seasonal variance, noise is estimated from the difference between the temporal auto-correlation function and a Gaussian fit for monthly CHL: 9/1997-12/2008.

Figure 3.1 Cross-correlations for the period 9/1997-2008 between: CHL and SST (top), CHL and MLD (middle), CHL and SSH (bottom). Boxes delineate areas for Figure 3.2, 3.3, 3.4.

Figure 3.2 Western equatorial Pacific time series (165-175^{o}E): CHL (red, all 3 panels) with SST (top), MLD (middle), and SSH (bottom). Average cross-correlations for each area noted.

Figure 3.3 Central equatorial Pacific (160-140^{o}W): variables as in Figure 3.2.

Figure 3.4 Eastern equatorial Pacific (120-110^{o}W): variables as in Figure 3.2.

Figure 3.5 Cross-correlations between normalized anomalies of CHL (9/1997-2008) and zonal wind stress, τ_x (top); meridional wind stress, τ_y (middle); magnitude of wind stress, $|\tau|$ (bottom).

Figure 3.6 Longitude-time distribution of averages (5^{o}S-5^{o}N) of deseasoned CHL (left), SST (2nd from left), SSH (3rd from left) and MLD (right). Niño4 is between the dotted and solid lines; Niño 3 is eastward of the solid line. Note the 1997/98 east Pacific extreme El Niño predominately in Niño3.

Figure 3.7 CHL/SST/SSH combined EOF mode 1 explains 30% of the total variance: CHL (top map), SST (middle map), SSH (bottom map) and the time-varying amplitude (bottom).

Figure 3.8 CHL/SST/SSH combined EOF mode 2 explains 15% of the total variance: CHL (top map), SST (middle map), SSH (bottom map) and the time-varying amplitude (bottom).

Figure 3.9 CHL/SST/SSH combined EOF mode 3 explains 8% of the total variance: CHL (top map), SST (middle map), SSH (bottom map) and the time-varying amplitude (bottom).

Figure 3.10 CHL/SST/SSH combined EOF mode 4 explains 4.4% of the total variance: CHL (top map), SST (middle map), SSH (bottom map) and the time-varying amplitude (bottom).

Figure 3.11 RMSE and r for SeaWiFS and cross-validated CHL reconstructed with 9-15 modes (color) for all bins (.), Nino area averages (*), and tropical Pacific averages (◊).

Figure 3.12 CCA mode 1 explains about 13% of the total variance with spatial functions: SST (top map), SSH (middle map), CHL (bottom map) and the time-varying amplitude (bottom).

Figure 3.13 CCA mode 2 explains about 12% of the total variance with spatial functions: SST (top map), SSH (middle map), CHL (bottom map) and the time-varying amplitude (bottom).

Figure 3.14 CCA mode 3 explains about 12% of the total variance with spatial functions: SST (top map), SSH (middle map), CHL (bottom map) and the time-varying amplitude (bottom).

Figure 3.15 CCA mode 4 explains about 10% of the total variance with spatial functions: SST (top map), SSH (middle map), CHL (bottom map) and the time-varying amplitude (bottom).

Figure 3.16 Fraction of Variance Resolved (FVR) by the dominant EOF modes (top) calculated by dividing the variance of the dominant ten EOF modes (middle) by total variance in CHL (bottom).

Figure 3.17 Deseasoned CHL for December, 1997: SeaWiFS (top), CCA (middle), MEOF(bottom).

Figure 3.18 Correlations between the original and reconstructed CHL from 9/1997-12/2008: for CCA (top, left map) and MEOF (top, right map). Colored bins have significant correlations. Boxes show Niño areas. Lower four panels are time series

averaged over the Niño areas for original (black), MEOF (red), CCA (blue) with average correlations to original CHL annotated for the latter two.

Figure 3.19 Leave five out cross-validations using SeaWiFS training data and reconstructing CHL by CCA (top) and MEOF (bottom) during the dependent period (9/1997-2008). Colored bins show significant correlations. Boxes show Niño areas that correspond to the average time series in Figure 3.20.

Figure 3.20 Comparisons of SeaWiFS CHL (black) to leave-five-out-cross-validations of CHL reconstructed by CCA (blue) and MEOF (red) and averaged over four Niño regions. Average correlations with the original SeaWiFS CHL are annotated for each area.

Figure 3.21 Reconstructed CHL (--), original CHL (--), NODC observations (+) and observations reconstructed using first 5 EOF modes (◊), with correlation coefficients annotated for each area.

Figure 3.22 Reconstructed CHL (-) and CZCS CHL (-x-) averaged over four Niño regions.

Figure 3.23 Model CHL cross-correlated with original from 9/1997-2007 (top); model CHL cross-correlated with reconstructions over the independent period from 1988-8/1997 using CCA (middle) and MEOF (bottom). The Niño regions delineated correspond to the time series in Figure 3.24.

Figure 3.24 Time series corresponding to areas delineated in Figure 3.23: Niño 1&2 (top), Niño 3 (2nd), Niño 3.4 (3rd), Niño 4 (bottom): model CHL (grey), compared to CCA CHL (blue), MEOF CHL (red), SeaWiFS CHL (black). Area average cross-correlation coefficients are annotated.

Figure 3.25 Reconstructed CHL (blue) +/- RMSE (dotted blue) and original CHL (black). RMSE during the dependent period is calculated between the original and CCA CHL; RMSE during the independent period is estimated using the leave-five-out-cross-validations, as described in the text.

Figure 3.26 La Niña – El Niño CHL differences for December 1998 minus December 1997 averaged for the original CHL (top) and reconstructed CHL (bottom). Note the end points of the color scale represent values greater than 0.15 mg m^{-3} and less than -0.15 mg m^{-3} respectively.

Figure 3.27 La Niña – El Niño differences in physical variables: December 1998 minus December 1997 for SSH (top) and SST (bottom).

Figure 4.1 Standard deviation of the deseasoned anomalies normalized by the mean for SSH (top), SST (middle), and CHL (bottom).

Figure 4.2 Longitude–time plots of reconstructed CHL (left), SST (2^{nd}), SSH (3^{rd}), MLD (right) between 1958-2008. Niño 4 is between dotted and solid lines; Niño 3 is east of solid line.

Figure 4.3 Annual average CHL anomalies during very strong El Niño events (ONI > 2) that formed during 1972 (top), 1982 (middle), 1997 (bottom). Averages centered on December +/- 6 months.

Figure 4.4 Average CHL anomalies during central Pacific El Niño events that formed during 1958 (top), 1987 (second), 1994 (third), 2004 (bottom), centered on December +/- 6 months. 1958 started as a strong El Niño (2 >ONI > 1), but turned weak by February while the other three remained strong.

Figure 4.5 Annual average CHL anomalies during La Niña events (ONI < -2) that peaked during 1973 (top), 1988 (middle), 1999 (bottom). Averages centered on December +/- 6 months.

Figure 4.6 Decadal averages of reconstructed CHL anomalies, as annotated for 1958-2007.

Figure 4.7 Top: monthly variation in ONI (red) and PDO (blue) indices smoothed over 12 months. Middle: CHL anomalies correlated with DJF ONI (1958-2008). Bottom: CHL correlated with PDO.

Figure 4.8 Top: monthly variation in EP (red) and CP (blue) indices smoothed over 12 months. Middle: CHL correlated with EP index (1958-2008). Bottom: CHL correlated with CP index.

Figure 4.9 Cool phase, 1958-1976, average CHL anomalies with Niño regions delineated (top); time series averaged over Niño regions (2nd panel).

Figure 4.10 Warm phase, 1977-1995, average CHL anomalies with Niño regions delineated (top); time series averaged over Niño regions (bottom).

Figure 4.11 Niño 4 (top left), Niño 3.4 (top right), Niño 3 (bottom left), Niño 1&2 (bottom right). CHL anomaly averaged for cool (—) and warm (—) PDO phases and each month (x and +).

Figure 4.12 CHL difference over a PDO cycle: cool phase (1958 – 1976) subtracted from warm phase (1977 – 1995). Niño regions delineated.

Figure 4.13 Average differences in physical variables between the warm and cool regimes: SST (top), SSH (2nd from top), zonal wind stress (3rd from top), wind stress curl (bottom). Niño areas are delineated. Missing data are shaded white.

Figure 4.14 Ekman pumping averaged over the PDO cool phase from 1958-1976 (top), with time series averaged over each Nino area as indicated (bottom): positive is upward, negative is downward.

Figure 4.15 Ekman pumping averaged over the PDO warm phase from 1977-1995 (top), with time series averaged over each Nino area as indicated (bottom): positive is upward, negative is downward.

Figure 4.16 Niño 3 and 3.4 divided into inner (3^{o}S-3^{o}N) and outer (3-10^{o}N) areas: CHL averaged over cool regime from 1958-1976 (top) and corresponding time series averaged over delineated areas (2^{nd} panel). CHL averaged over warm regime from 1977-1995 (3^{rd} panel) and corresponding time series (bottom).

Figure 4.17 Zonal wind stress anomalies for same time periods and areas as Figure 4.17: Niño 3 and 3.4 divided into inner (3^{o}S-3^{o}N) and outer (3^{o}-10^{o}N): Cool regime average anomalies (top) and time series over inner areas (2^{nd} panel). Warm regime average (3^{rd} panel) and times series (bottom).

Figure 4.18 Quarterly zonal wind stress anomalies and CHL anomalies for the equatorial band (3^{o}S-3^{o}N, 110-150^{o}W) during the cool era (+) and warm era (**x**).

Figure 4.19 Nitrate climatology, World Ocean Atlas 2009. Top: Nitrate map at 100m with zonal, meridional transects (white lines). Middle: Nitrate profile at 0^{o}N. Bottom: Nitrate profile at 110^{o}W.

Figure 4.20 Zonal velocity (cm s^{-1}) from the TAO/Triton equatorial ADCP at 140^{o}W from April, 1990 through 2008 with the depth (left axis) and strength of the EUC (color bar). Positive velocities are eastward. Figure courtesy of TAO Project Office/PMEL/NOAA.

Figure 4.21 Depth of the EUC (m) in SODA at 140^{o}W between 1958 and 2008. Individual data points are shown (·) and plotted as a line that was smoothed over 12 months.

Figure 4.22 EUC depth averaged over the PDO cool and warm eras: lines indicate average values at each longitude over the PDO era; median values are included for the cool (x) and warm (+) phases.

Figure 4.23 Reconstructed CHL anomalies averaged 1988-2008 (top) with blue box on the equator corresponding to the blue (x) in the scatter plot (below) of the monthly CHL against modeled iron [Wang et al., 2008] integrated over the euphotic zone (r~0.7). The red box on the equator in the west corresponds to the red (□) in the scatter plot and smaller cross-correlation (r~0.5).

Figure 4.24 Left: Monthly CHL where the contour demarcates 0.1 mg m^{-3}. Niño 4 is between dotted and solid lines; Niño 3 is east of solid line. Right: Monthly mean tuna catch per unit effort (CPUE) where the pink line indicates the longitudinal center of gravity of CPUE, the blue line indicates the 29^{o}C SST isotherm, the dashed red line indicates the Southern Oscillation Index. All variables smoothed over 3 months. Figure on right from Lehodey et al. [1997].

Figure 4.25 Left: Monthly CHL where contour demarcates 0.1 mg m^{-3}. Niño 4 is between dotted and solid lines; Niño 3 is east of solid line. Right: Oceanic Niño Index (ONI) in black and Pacific Decadal Oscillation (PDO) in red, both smoothed over 12 months.

Figure 4.26 Percent of peak value of sardine and anchovy populations in Japan-Kuroshio system (top) and Peru-Humboldt system (bottom). Figure adapted from Lehodey et al. [2006].

List of Abbreviations

AMO – Atlantic multi-decadal oscillation
BATS – Bermuda Atlantic ocean time series
CCA - canonical correlation analysis
CHL – chlorophyll
CP – Central Pacific El Niño
CPUE – catch per unit effort
CZCS – Coastal Zone Color Scanner (ocean color satellite mission)
ENSO – El Niño Southern Oscillation
EOF – empirical orthogonal function
EP – East Pacific El Niño
ESTOC – European station for time series in the ocean Canary Islands
EUC – equatorial under current
FVR – fraction of variance resolved
HOT – Hawaii ocean time series
HPLC – high performance liquid chromatography
MEI – multivariate ENSO index
MEOF – multi-variate empirical orthogonal function
MERIS – medium resolution imaging spectroradiometer (ocean color satellite mission)
MLD – mixed layer depth
MOBY – marine optical buoy off of Hawaii
MODIS – moderate resolution imaging spectroradiometer (ocean color satellite mission)
NAO – North Atlantic Oscillation
NODC – National Oceanographic Data Center
NPGO – North Pacific Gyre Oscillation
OCTS – ocean color temperature scanner (Japanese ocean color satellite mission)
ONI – Ocean Niño Index
PDO – Pacific Decadal Oscillation
RMSE – root-mean-square error
SeaWiFS – Sea-viewing Wide Field-of-view Sensor (ocean color satellite mission)
SNR – signal-to-noise ratio
SODA – simple ocean data assimilation
SSH – sea surface height
SST – sea surface temperature
TAO – tropical atmosphere ocean mooring
T_{20} – 20°C isotherm depth, a proxy for the thermocline

1. Introduction

1.1 Statement of the Problem

Marine phytoplankton accounts for about half of all primary production on Earth and contributes to the ocean's uptake of CO_2 and long-term carbon sequestration [Field et al., 1998]. Since the beginning of the industrial age, the atmospheric CO_2 concentration has increased from about 280ppm to more than 400ppm today, but the observed annual rate of increase has been half of that expected from fossil fuel and cement-manufacturing emissions [Sabine et al., 2004]. Without the ocean, atmospheric CO_2 would be much higher than that currently observed. The ocean absorbs CO_2 through physical solubility and biological mechanisms, known as pumps for their ability to transfer carbon down through the water for short and long-term storage.

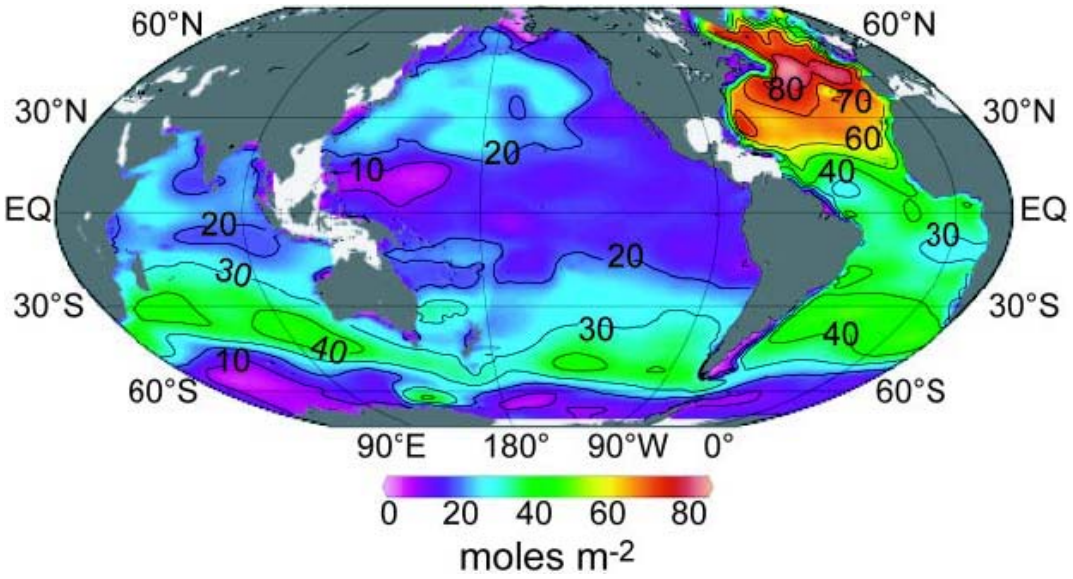

Figure 1.1 Column inventory of anthropogenic CO_2 in the ocean. High inventories are associated with deep water formation in the North Atlantic (deepest sequestration) and intermediate and mode water formation between 30°N and 50°S. Total inventory of colored region is 106 ± 17 Pg C or 1/3 of anthropogenic carbon emitted over the past two centuries. Figure from Sabine et al. [2004].

Chemical concentration data collected by global oceanographic cruises during the 1990s and a tracer-based reconstruction technique combined with modeling studies indicate that one-third of the anthropogenic CO_2 produced over the past two centuries is stored in the ocean [Sabine et al., 2004]. Most of the carbon inventory has accumulated in the ocean's interior, delivered there from the ventilated surface layers by either chemical or biological mechanisms.

The solubility pump describes the oceanic uptake of CO_2 during high latitude deep water formation and winter convection as well as subtropical mode water formation. Lower temperatures increase the solubility of CO_2 and higher winds enhance its air-sea flux. Once it enters the ocean, carbon transport to the interior is largely determined by mixing and convection [Gnanadesikan et al., 2007; Gruber et al., 2009]. The solubility pump has been well observed for the past two decades. Coupled physical/biogeochemical models run to the end of the 21st century indicate global warming reduces the ocean's capacity to absorb carbon in two ways: lowering the solubility of CO_2 at higher temperatures and slowing convective overturning which transports carbon to the deep ocean interior for long-term storage [Sarmiento et al., 1998]. Natural climate variability also impacts the ocean's uptake of carbon, as has been observed in the North Atlantic: CO_2 uptake during subtropical mode water formation was greatly enhanced during the 1990s positive phase of the North Atlantic Oscillation but decreased during the 2000s neutral and negative phases [Bates, 2012].

The biological pump describes the uptake of CO_2 at the ocean's surface by phytoplankton that are consumed by higher trophic levels and eventually expelled, enabling the small fraction not remineralized in the surface layer to fall below the

pycnocline. An even smaller fraction that settles to the ocean's bottom and becomes part of the sediment, representing an important sink over long time scales. Although understanding of the biological pump is still in its infancy, modeling studies indicate that it may account for half of carbon uptake in the surface layer [Sarmiento et al., 1998].

Accounting for phytoplankton variability at low frequencies and basin scales is critical to understanding its role in the Earth's climate system. The distribution of phytoplankton in the oceans is dependent upon the availability of nutrients and sunlight. Ecological changes related to climate have been seen in long-term regional sampling and modeling studies [Raitsos et al., 2005; Henson et al., 2009; Saba et al., 2010; Cianca et al., 2012], yet large-scale, long-term studies have been limited by the scope of available biological records [Behrenfeld et al., 2006; Martinez et al., 2009; Yoder et al., 2010].

The first continuous global record of ocean biology collected by the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) ocean color satellite mission from late 1997 through 2008 enabled an unprecedented decadal view of the spatial distribution of chlorophyll in the ocean's surface layer. This synoptic perspective revealed physical mechanisms that transport nutrients to fertilize blooms. In general, chlorophyll distributions are high along coasts, subpolar regions, eastern boundaries and along the equator, where nutrients are supplied through coastal run-off, aeolian deposition or upwelling from the light-limited deep ocean. Wind-driven mixing and cyclonic circulation cause vertical transport of nutrients to the surface waters from depth; winter convective overturning also supplies new nutrients to the surface as deep nutrient-rich water is ventilated. Chlorophyll concentrations are low in the central subtropical gyres as nutrients are depleted from surface waters and wind-driven anticyclonic circulation

causes downwelling, sending nutrients below the mixed layer. The distribution of blooms are apparent in seasonal averages calculated over the SeaWiFS mission (Figure 1.2).

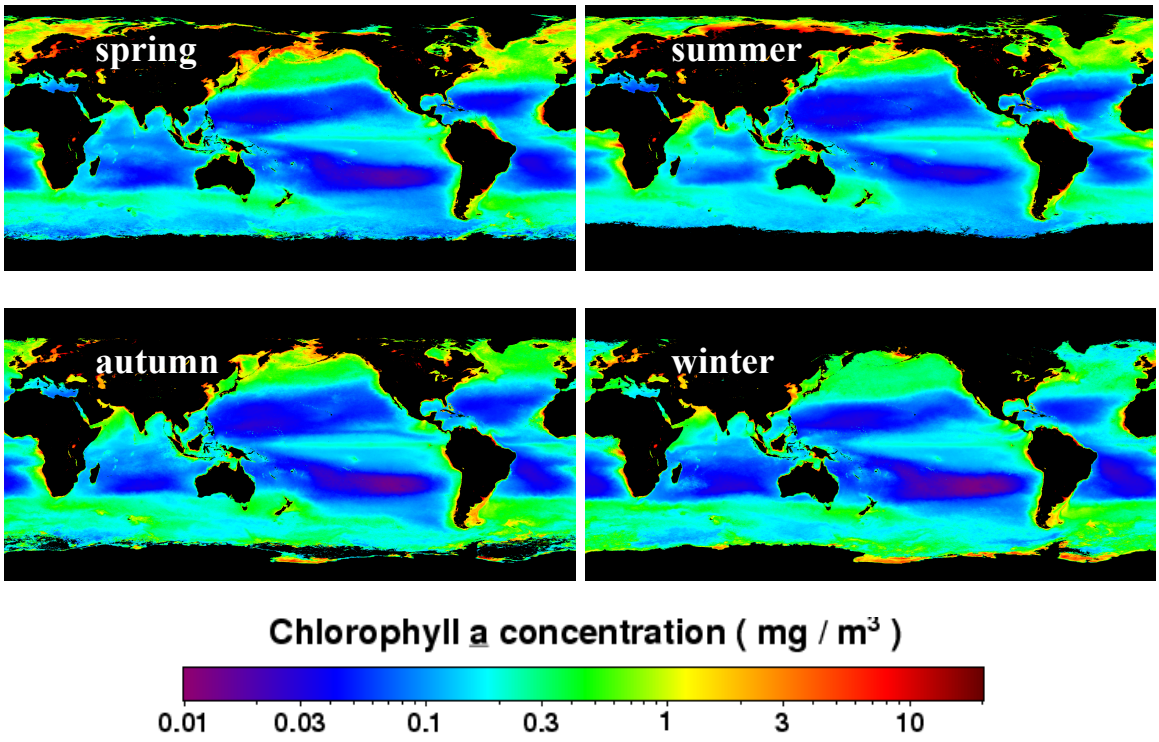

Figure 1.2 Seasonal distribution of chlorophyll-a averaged over the SeaWiFS mission (9/1997-2009) for Northern Hemisphere spring (top left), summer (top right), autumn (bottom left), winter (bottom right). Figure courtesy of NASA Ocean Biology Processing Group.

For many years a simple view of vertical nutrient supply leading to blooms predominated. Spring bloom was defined as the time when surface mixing shoals to a critical depth where phytoplankton growth exceeds losses (Figure 1.3, top). However, new insights facilitated by a decade of satellite imagery have resulted in a three dimensional picture and much more nuanced understanding of the 'critical depth hypothesis' (Figure 1.3, bottom). The spatial and temporal resolution of ocean color

remote sensing combined with high resolution physical variables, such as temperature and winds, plus in situ observations from floats and buoys has confirmed and clarified basin scale nutrient transport that fuels primary productivity [Sverdrup, 1953; Williams and Follows, 2003; Palter et al., 2005; Ayers and Lozier, 2010; Behrenfeld, 2010].

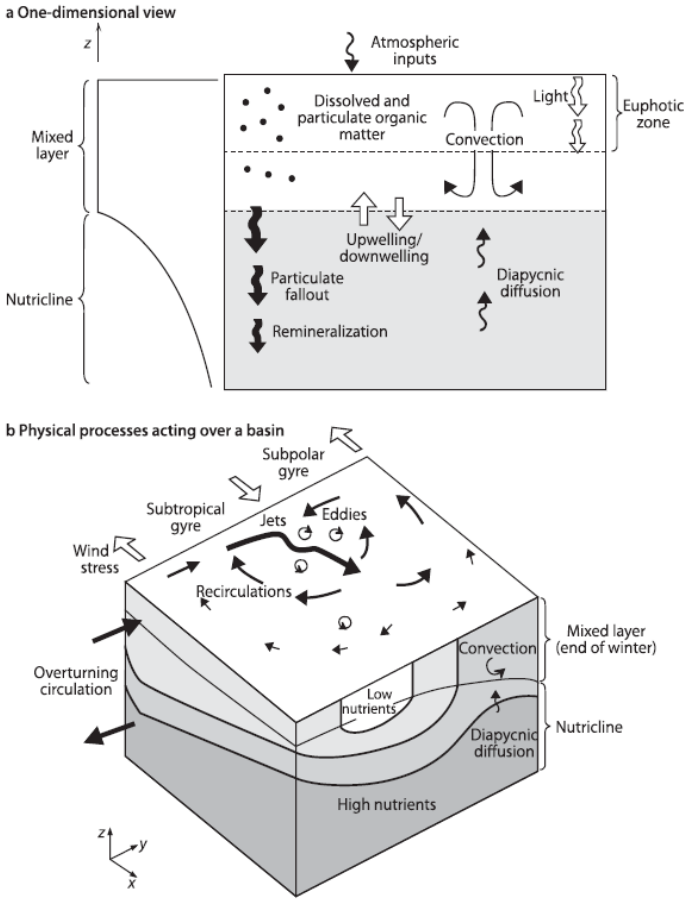

Figure 1.3 Physical processes affecting nutrient supply that fuels primary production: a) one-dimensional view shows biomass export to depth is partly transferred back to the mixed layer through advection, convection, and diapycnic diffusion; and b) three-dimensional view is more complete depicting nutrient supply by gyres, eddies, fronts, jets, and convective overturning. Figure adapted from Williams and Follows [2003].

Advection, convection, cross-frontal exchange, and diapycnal diffusion are processes now known to supply nutrients to phytoplankton in the euphotic zone.

The connection between physical forcing and ocean biology has been elucidated by the synoptic view in ocean color satellite data records, yet the length of these records has limited their range from daily to interannual temporal scales. Lower frequency phytoplankton bloom variability is not well understood. Having an ocean color record span several decades would make it feasible to address an open question in climate science about how phytoplankton bloom patterns change in space and time over decadal and longer time scales.

1.2 Motivation and background

The ultimate goal of this study is to determine what large-scale chlorophyll patterns over 50 years reveal about the interaction of ocean biology with climate-scale variability. Ocean biology data collection lags ocean physics: there is not a continuous, well-calibrated direct measurement of any biological variable that spans both space (e.g. ocean basins) and time (e.g. multiple decades). This data paucity has limited the scope of our understanding of ocean biology at long length and time scales. A century-long study of phytoplankton inferred from global transparency data collected from secchi disks found nearly universal declines in biomass [Boyce et al., 2010], however it was shown by others that their technique of combining transparency data to more modern sampling introduced bias [Mackas, 2011; Rykaczewski and Dunne, 2011; McQuatters-Gollop et al., 2011]. There is a clear need for validated, multi-decadal ocean biological data sets.

In contrast, climate variability is much in evidence in physical variables, which are significantly better sampled by comparison. Prominent spatial patterns in temperature, atmospheric pressure and winds have been identified that are known to affect weather and climate around the world. The most dominant and truly periodic mode of climate variability is the seasonal cycle, with other modes being quasi-periodic or having a spectrum of temporal variability [Kaplan, 2011]. The first climate-scale pattern to be discovered, the Southern Oscillation, is associated with the surface air pressure difference between the eastern and western tropical Pacific Ocean. Normally, easterly equatorial trade winds push warm water toward the west with cold water upwelling along South America [Walker, 1924; Bjerknes, 1969; Philander, 1999]. During the negative phase of the oscillation, the zonal surface pressure difference is small, trade winds slack, the surface warm water pool normally in the west Pacific moves eastward, the thermocline deepens along the coast and upwelling diminishes off of South America, an effect termed *El Niño* by the Peruvian fishermen because it typically occurs at Christmas time. During the positive phase of the oscillation, the pressure difference is great, trade winds increase, the warm water pool is pushed farther west than normal, and upwelling increases, called *La Niña* as the opposite of El Niño. The combined effect of the Southern Oscillation in the atmosphere and El Niño in the ocean is known as El Niño Southern Oscillation (ENSO). The interannual ENSO cycle is recognized as the strongest mode of climate variability over the past century [Philander, 1990, McPhaden et al., 2006; Deser et al., 2010; Messie and Chavez, 2011]. Whether this pattern is intensifying or changing is an area of active research [Federov and Philander, 2000; McPhaden et al., 2011; L'Heureux et al., 2013]. Lesser oscillations in sea-level pressure happen in other ocean basins: the

North Atlantic Oscillation and the North Pacific Oscillation [Walker, 1924]. The advent of satellite remote sensing and autonomous *in situ* sensors has greatly increased the observation network in space and time. Data assimilation techniques have evolved to handle larger volumes of data, to fill in missing data and to merge entire records from multiple sources. With these advances in data collection and computing power, climate oscillations at lower frequencies have been identified in the oceans (e.g. multidecadal Atlantic Multidecadal Oscillation [Schlesinger and Ramankutty, 1994]; multicentennial to millennial Atlantic Meridional Overturning Circulation [Schulz et al., 2007]).

Large changes in atmospheric and ocean time series are used to define climate indices. In the Pacific, ENSO is defined using sea-surface temperature (SST) anomalies [Cane et al., 1986; Trenberth, 1997]. One index is the Ocean Niño Index (ONI), typically used from Niño 3.4 during December-February (DJF) when El Niño events peak [L'Heureux et al., 2012]. The Pacific Decadal Oscillation (PDO) is defined by the first principal component of the North Pacific SST anomaly field after subtracting the global mean [Mantua et al., 1997; Nigam et al., 1999]. The North Pacific Gyre Oscillation (NPGO) is the second principal component of SST and sea-surface height (SSH) [DiLorenzo et al., 2008]. Positive NPGO indicates an intensification of the geostrophic circulation between the Alaskan Gyre and the Subtropical Gyre, which increases transport of both the northward Alaskan Coastal Current and the southward California Current. These and other climate indices have been defined by raw or statistically analyzed observations using multi-decadal and century-long physical records.

In order to understand the connection between low frequency forcing and biological response, climate indices have been correlated to chlorophyll and net primary

production, but results have all been limited by the length or extent of the biological observations. For example, *in situ* chlorophyll collected in the Canary Islands since 1994 reveal anomalies sensitive to the low-frequency climate modes of ENSO and NAO [Cianca et al., 2012]. Deep ocean time series, such as BATS in Bermuda and HOT in Hawaii, show that ecosystem measurements collected over 20 years relate to the NAO and the NPGO. That the NPGO index has the highest correlation with net primary productivity at both sites implies a teleconnection between ocean basins [Müller et al., 2008]. In the Pacific, NPGO is linked to ecosystem dynamics through wind-driven upwelling and horizontal advection of nutrients to fuel phytoplankton blooms [DiLorenzo et al., 2008], while the link to the Atlantic is harder to explain [Saba et al., 2010].

Several studies have noted that primary production at BATS is inversely correlated to the NAO: negative NAO is associated with stronger winds, more mixing, deeper MLD, cooler SST and increased primary production [Lomas et al., 2010; Saba et al, 2010]. When the NAO is in neutral or negative phase, the CO_2 sink is short-term (less than 10 years). Conversely, during positive phases of the NAO, long-term storage of anthropogenic CO_2 is enhanced because stronger stratification results in less mixing at the surface, but more diapycnal and isopycnal diffusion at depth during subtropical mode water formation [Bates, 2012]. Understanding how low-frequency physics impacts ocean biogeochemical cycling is still in its infancy.

The regular, large-scale view afforded by ocean color records have led to new insights about the synchrony between ocean biology and physics. One study found that chlorophyll declines over the subtropical stratified ocean during 1999-2004 were strongly correlated to the Multivariate ENSO Index (MEI), $r^2 = -0.77$, shown in Figure 1.4, caused

by enhanced stratification that suppressed vertical mixing and nutrient exchange [Behrenfeld et al., 2006]. Another study used the first principal components of global biological and physical data from satellites and other sources between 1993-2010 combined with a phytoplankton model and found that interannual biological change can be explained directly from physics [Messie and Chavez, 2012]. Specifically, thermocline depth is the primary driver of biological variability during El Niño [Barber and Chavez, 1983; Christian et al., 2002], with a high correlation (r = 0.89) between the MEI and the first mode of global chlorophyll. They found other factors to be secondary during El Niño (e.g. community structure) [Messie and Chavez, 2012].

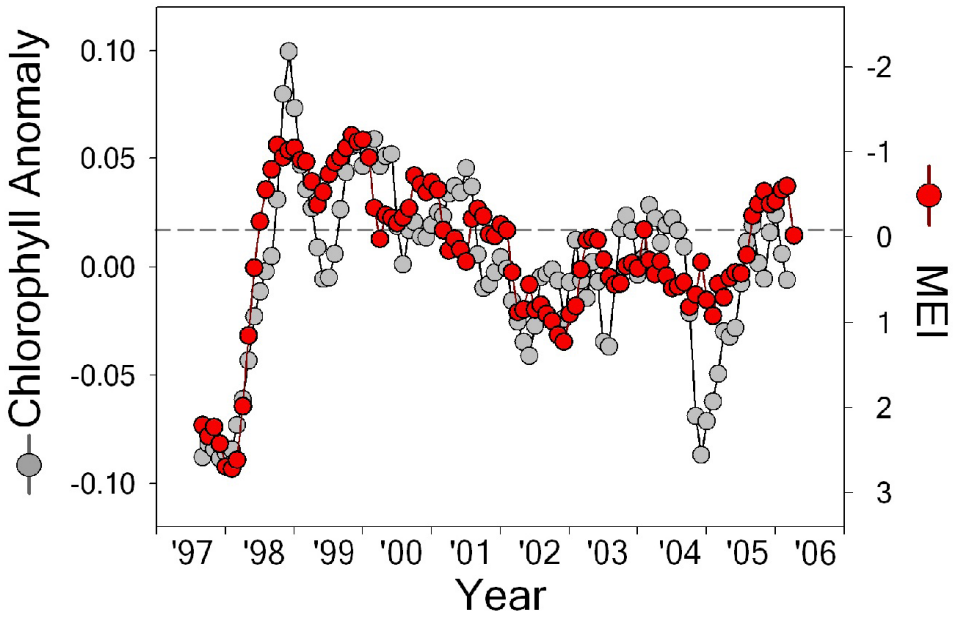

Figure 1.4 **SeaWiFS monthly water-column integrated chlorophyll anomaly for the global permanently stratified ocean, defined as annual average SST > 15C roughly between 40°N - 40°S, (grey symbols, left axis) to multivariate ENSO, or MEI index (red symbols, right axis). Note reversed MEI axis. Figure from Behrenfeld et al. [2006].**

Multi-decadal changes in global phytoplankton abundances have been related to basin-scale forcing, namely the PDO and the Atlantic Multi-decadal Oscillation (AMO), shown in Figure 1.5, by utilizing the intermittent record of the Coastal Zone Color Scanner (CZCS) between late 1978 and mid 1986 in addition to the SeaWiFS chlorophyll record from late 1997 through 2006 [Martinez et al., 2009]. Because CZCS was a shared mission, not dedicated to ocean color, its CHL record is limited and primarily focused on the coastal United States, as detailed further in section 3.2.3. The data gap between 1983-1997 made it impossible to directly compare CZCS and SeaWiFS, but they were each correlated to climate oscillations.

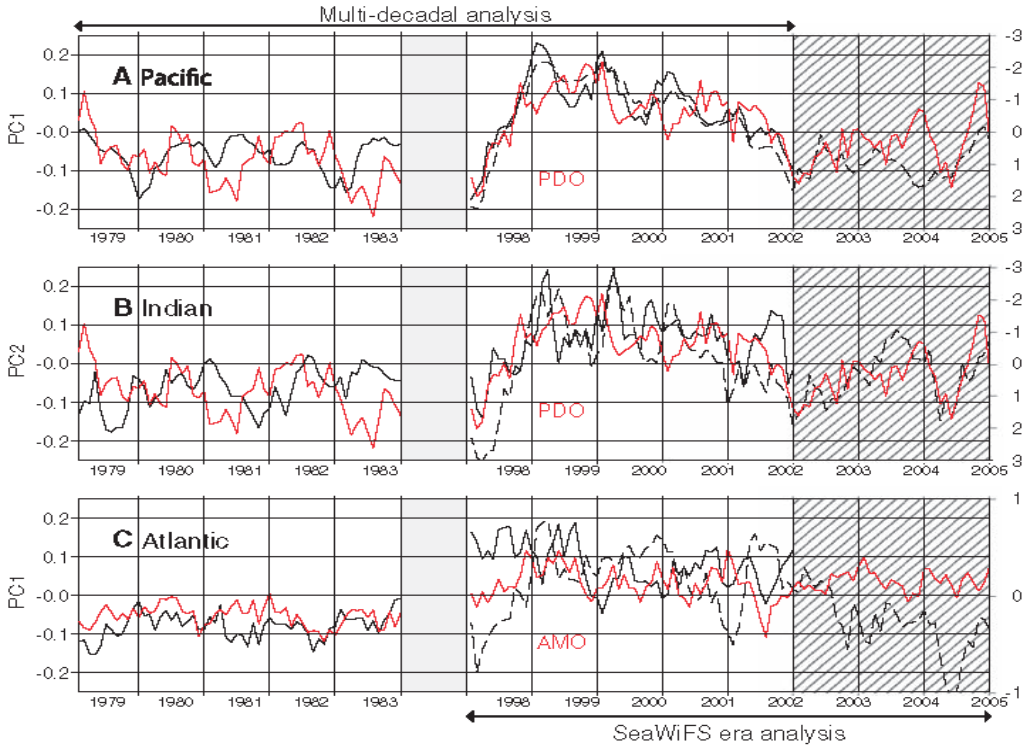

Figure 1.5 CHL-SST combined EOF time amplitude (black, left axis) with climate indices (red, right axis) for CZCS (1979-1983) and SeaWiFS subsets (1998-2002) over a) Pacific Ocean, r_{PDO}=0.66; b) Indian Ocean, r_{PDO}=0.6; c) Atlantic Ocean, r_{AMO}=0.75. Figure from Martinez et al. [2009].

Despite such efforts to tease large-scale, low frequency information from the ocean color data collected, we remain sorely limited in our understanding of how climate-scale variability affects phytoplankton bloom patterns. One technique that several studies have used to study climate-scale variability is statistical reconstruction to extend data spatially and temporally. Canonical Correlation Analysis (CCA) is one method to statistically reconstruct a sparse data set using longer records as proxies. CCA is a powerful statistical tool that takes advantage of high density data to produce mathematical relationships that can be used to reconstruct periods when one data set is sparse or non-existent. The success of the CCA method depends upon the selection of high quality proxies that have close correspondence to the target variable. Using covariance patterns between correlated variables, a more extensive modern data set can be used to predict or reconstruct an earlier data set in space or time. The high data density era is used as the training period to establish the correlation patterns with the more extensive data sets. Major patterns from the dominant modes are then projected back in time using the more extensive data sets to reconstruct the target variable. Over the past 15 years, this and similar reconstruction methods have been applied to reconstruct an increasing number of physical climate variables backward in time. Historical sea-surface temperatures have been extended back to the 1860s [Kaplan et al., 1998; Smith et al., 1996]; marine sea level pressure has been reconstructed back to the 1850s [Kaplan et al, 2000]; global land-air-sea surface temperatures have been merged and reconstructed back to 1880 [Smith and Reynolds, 2005]; global oceanic precipitation has been extended back to 1900 [Smith et al., 2009]; satellite altimetry combined with historical tide gauge data were used to reconstruct sea level rise since 1950 [Church et al., 2004]. Reconstructions

identify large-scale, long-term variability and provide useful information on interannual climate variability. CCA has been successfully applied to extend physical climate data back more than a century, but has not previously been used to extend biological data.

Reconstructing a biological variable presents challenges because ocean biology does not necessarily have a linear response to forcing mechanisms. Chlorophyll estimates have been extended through modeling studies [e.g. Follows et al, 2007; Henson et al., 2009; Wang et al., 2009] but have not been statistically reconstructed. Because of the strong dependence of phytoplankton community structure upon physical forcing to supply the nutrients needed to bloom, and because physical measurements in the ocean generally have longer records with more complete coverage than those for ocean biology, well-correlated physical proxies could potentially be used to extend the biological record back in time. Statistical reconstruction using proxies is well-established in other disciplines and is valuable for comparison with other methods. Models include known mechanisms, while statistical reconstructions have the potential to capture large, complex signals including unknown mechanisms. It is conceivable that a model would exclude an important process that is not yet understood, a situation where a statistical reconstruction might have an advantage. Although the variance captured in a CCA reconstruction will only be as good as the covariance between the reconstructed variable and the proxies, selecting the most relevant proxies could reproduce the essential patterns.

The advantage of a multi-decadal ocean color chlorophyll reconstruction over a long ocean time series, such as BATS and HOT, is its spatial information: the former are point source measurements which may or may not be representative of large-scale patterns. Using two-dimensional information collected by ocean color satellites over a

decade combined with a longer time series of proxy data, a statistical method can be applied to reconstruct the past few decades of ocean surface chlorophyll prior to ocean color. A two-dimensional ocean color reconstruction will elucidate how slowly varying patterns change in space and time with low-frequency oscillations.

1.3 Objectives of this research

Taking advantage of high resolution ocean color satellite data and its correlation to physical proxies, CCA will be used to extend the chlorophyll record during the entire 50 years of a physical proxy data set. Reconstructing chlorophyll over five decades will bridge data collected during disparate ocean color eras, identify spatiotemporal patterns and links to climate phenomena, and isolate regional differences.

Reconstructing the chlorophyll record back in time will permit a multi-decadal study at the relatively high data density of recent ocean color remote sensing observations. The gap between various ocean color satellite sensors will also be bridged by the reconstruction: CZCS (1979-1983), SeaWiFS (9/1997-2010), MODIS/Aqua (2002-2010). The 50 year reconstruction will highlight slow changes in spatial patterns, including internal variability in the ocean-atmosphere system which impacts nutrient supply to fuel primary production. The chlorophyll reconstruction will be used to identify low frequency patterns rather than trends. This is still very valuable information. For example, variability in the spatial extent of a regime and bloom timing can impact the efficiency of the biological carbon pump as in the North Atlantic mode-water formation region where autumn/winter primary production can be rapidly exported while spring bloom primary production cannot [Levy et al, 2005].

Climate-scale forcing can have global ramifications, but differences in biological provinces and their variation in response make it most practical to observe climate effects on biology in regional studies. Many studies have expounded upon the variation in marine phytoplankton bloom phenological characteristics and have defined various regions qualitatively [Longhurst, 1995] or quantitatively [Racault et al., 2012; Sapiano et al., 2012]. A chlorophyll reconstruction can address how climate system dynamics impact regional ocean biology. The tropical Pacific lends itself well to the first regional reconstruction due to its large extent, accounting for roughly 20% of global ocean productivity [Behrenfeld et al., 2006], biologically productive equatorial region, and relatively high data sampling as well as being a favorite site for modeling studies. The most prominent climate-scale oscillation is ENSO [Philander, 1990; McPhaden et al., 2006; Messie and Chavez, 2011] which originates in the equatorial Pacific but has world-wide teleconnections. It makes sense to focus the initial discussion about the effect of climate-scale oscillations on ocean biology in the tropical Pacific where ENSO is strongest.

In the past few decades, understanding of the dynamics of Earth's atmosphere and how it interacts with the ocean has increased dramatically with respect to ENSO, although the physical processes are not completely understood. Recent studies of ocean biology in the tropical Pacific have indicated that sea surface chlorophyll is sensitive to the El Niño structure, which some say has shifted westward over the past decade [Turk et al., 2011; Radenac et al., 2012; Gierach et al., 2012]. The eastward advection of the west Pacific warm pool during El Niño events sometimes stops in the Central Pacific and sometimes extends all the way across the basin to the eastern Pacific [Trenberth and Stepaniak, 2001; Ashok et al., 2007; Lee and McPhaden, 2010; McPhaden et al., 2011;

Karnauskas, 2013]. As shown in the left panel of the figure from Turk et al. [2011] (Figure 1.6a), during neutral conditions equatorial sea surface temperatures (SST) are highest west of 160^{o}E; during an eastern Pacific El Niño, the maximum shifts eastward toward the dateline with a minimum in the east that is about 3^{o}C warmer as the east-west thermocline slope is flattened and coastal upwelling diminishes; during a central Pacific El Niño the SST maximum is higher near the dateline but the eastern minimum only warms about 1^{o}C and coastal upwelling continues. SeaWiFS chlorophyll (CHL) concentrations experience maxima and minima locations and shifts that are roughly inverse to SST, as shown in the right panel (Figure 1.6c). During neutral phases, the CHL concentrations are less than 0.1 mg m^{-3} in the warm pool west of 160^{o}E and gradually increase to 0.25 mg m^{-3} east of the dateline due to the upwelling of nutrients along the equator and along the west coast of South America. During both east and central El Niño events: the warm pool is advected eastward, the nutricline lifts behind it and the chlorophyll minimum rises above 0.1 mg m^{-3} in the west. During an east Pacific El Niño, the warm pool propagates all the way to the east side of the basin and suppresses coastal upwelling along South America causing decreased CHL concentrations in the eastern Pacific. During a central Pacific El Niño, the chlorophyll minimum extends to around 170^{o}E rising gradually east of 160^{o}W. The total surface CHL integrated across the equatorial Pacific decreases much less during a central Pacific El Niño than during an eastern Pacific El Niño, as seen in Figure 1.6c [Turk et al., 2011]. Whether the recent trend of more frequent central Pacific El Niños is part of a natural periodic cycle or not is an area of active research and debate [Yeh et al., 2009; Lee and McPhaden, 2010;

McPhaden et al., 2011]. Regardless of the forcing, the eastward extent of the El Niño event impacts primary production and higher trophic levels.

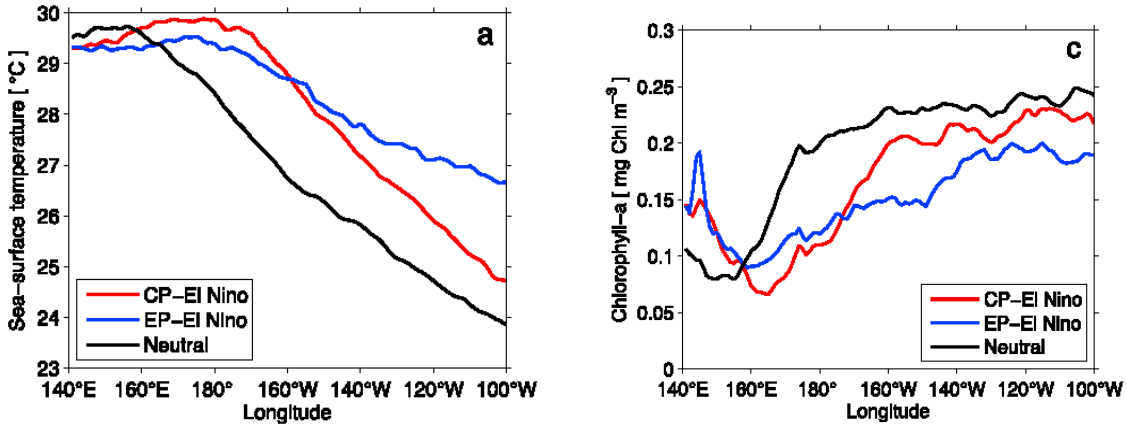

Figure 1.6 Mean winter (DJF) Optimally Interpolated SST (left) and SeaWiFS CHL (right) variation with longitude averaged between 1°N-1°S for neutral (black), Central Pacific El Niño (red), Eastern Pacific El Niño (blue) phases. Figure adapted from Turk et al. [2011].

One of the limitations of a chlorophyll reconstruction is that it can only be used to identify significant changes in spatial and temporal patterns, rather than any slow, steady change in total abundance. For example, periods of negative temperature anomaly may correspond to upwelling and enhanced productivity with reasonable confidence, but a slow warming trend does not necessarily mean a decrease in nutrients since a temperature change could also be due to change in cloudiness or latent heat flux. There is a fairly robust inverse relationship between SST and NO_3 because nutrients upwell with colder water [Kamykowski and Zentara, 1986]. This relationship varies slightly with location and season, as does the nitrate-depletion temperature that indicates which types of phytoplankton can thrive in any location [Carder et al., 1999], but is otherwise relatively constant. Increasing atmospheric CO_2 levels may change the SST-NO_3 relationship toward warmer values, just as glacial episodes changed the relationship toward colder

values in the geologic record [Walsh, 1980]. A low frequency trend in primary production could restructure the marine ecosystem and alter the equatorial Pacific carbon cycle [Gierach et al., 2012]. Because a chlorophyll reconstruction using proxy data relies upon the short-term covariance between variables, any long-term trend in the proxy data could be reflected as a long-term trend in the reconstructed chlorophyll, yet such a trend would be speculative without *in situ* validation. Subtle long-term biological trends should be determined from the data directly, not by proxy.

1.4 Organization of the dissertation

The remainder of this dissertation is organized as follows: Chapter 2 explores various ways biological processes are observed in the global ocean, their relationship to physical oceanography and the prominence of the tropical Pacific when observing physical-biological processes. Chapter 3 leverages the correlation between phytoplankton chlorophyll and physical proxies in a 51 year statistical reconstruction, which is then validated and compared to *in situ* observations and CZCS ocean color data as well as to chlorophyll calculated by a fully-coupled three-dimensional physical-biogeochemical phytoplankton model. The transition between the largest El Niño event of the century and the La Niña that followed in 1998 is discussed for its exceptional characteristics. Chapter 4 explores low frequency chlorophyll anomalies in the tropical Pacific and characterizes the climate-scale forcing and phytoplankton response that impacts higher trophic levels through bottom-up controls in the food web. Finally, Chapter 5 summarizes my research and highlights the significance of these conclusions with ideas for future research.

2. Observing biological processes in the ocean

There is evidence of large-scale changes in marine biology as noted in the fisheries record over several centuries, yet the reason fish stocks shift in abundance and location and what happens to the rest of the marine food web in conjunction with these changes remain open questions [Llucha-Belda et al., 1989; Klyashtorin, 1998; Hare and Mantua, 2000; Overland et al., 2008]. Low frequency synchrony has been observed between climate-scale oscillations and ocean biology, especially ENSO [Chavez et al., 2003; Messie and Chavez, 2012], but our mechanistic understanding of changes from the base of the food web on up has been constrained by sparse, sometimes incomplete, biological records. Direct measurement of marine life has traditionally been limited by the vastness and opacity of the ocean, not to mention its corrosive environment and propensity to biofoul long-term instruments. There is a need for intensive, regularly sampled observations of primary production to put our current understanding in context and address unanswered questions. From *in situ* to remotely sensed observations, this chapter gives an overview of what is known about phytoplankton variability from current measurements, discusses the capabilities and limitations of observing methods, elaborates on the data sets used in this study, explores major global spatial and temporal patterns coincident between physical forcing and ocean surface chlorophyll maps, then focuses on the tropical Pacific where physical forcing has been found to be the primary driver of biological anomalies. A discussion of the physical oceanography impacting biology in the tropical Pacific concludes the chapter, highlighting major pertubations to the normal or neutral state there.

2.1 Phytoplankton measurements and modeling

In addition to its role at the base of the food web, phytoplankton is recognized as serving an important function in moderating Earth's climate through processes such as the biological pump. Various research efforts seek to quantify and model phytoplankton by its biomass, primary production, carbon export and other characteristics. One of the easiest ways to observe phytoplankton from space is through its dominant photosynthetically active pigment, *chlorophyll a*. Typically called simply *chlorophyll*, phytoplankton pigments absorb more blue and red light than green, causing the spectrum of backscattered sunlight, or color of seawater, to shift from blue to green as the concentration of phytoplankton increases [Yentsch, 1960]. Methods to quantify biomass range from *in situ* measurements of single cells to synoptic scale satellite ocean color remote sensing of their signature in the visible portion of the electromagnetic spectrum [O'Reilly et al, 1998]. Different phytoplankton species vary in their nutrient and light requirements, life cycles, ranges, photoadaptation and package effect, making them difficult to quantify using standard criteria. For example, the covariation between light absorption and chlorophyll concentration is related to their photoadaptation: phytoplankton pigments adapt to low light levels by increasing their photosynthesis rate per unit of chlorophyll meaning two areas having the same chlorophyll concentration could have different rates of primary production [Bricaud et al., 1995]. Nonetheless, it is useful to try to quantify their abundance at varying spatial and temporal resolutions to better understand what controls ocean biology. Single point *in situ* measurements perform certain functions, such as identifying species composition. Averaging over large-scales serves other purposes, such as understanding the global biogeochemical cycle.

2.2 In situ observing methods: advantages and limitations

There are several methods to determine phytoplankton concentrations through *in situ* measurements: spectrophotometry, fluorometry, and chromatography. The classical method of determining chlorophyll amount as a representation of phytoplankton concentration was by the labor-intensive spectrophotometry process of collecting a water sample, filtering it, and extracting chlorophyll for analysis by its absorbance [Vernon, 1960]. Fluorometry also requires extraction to a filter and then uses a fluorometer to measure chlorophyll fluorescence [Yentsch and Menzel, 1963]. Underway fluorometry yields a lot more data, but relies on discrete samples for calibration due to the differences in local species' fluorescence [Herman and Denman, 1977]. Despite refinement of spectrophotometry and fluorometry over the years, neither of them handle the complex range of pigments found in seawater [Holm-Hansen et al., 1965; Murray et al., 1986]. High-performance liquid chromatography (HPLC) is considered the most precise method of measuring multiple pigments through extraction and analytic chemistry [O'Reilly et al., 1998]. These methods are time-consuming and require an experienced analyst.

Time series projects, including Bermuda Atlantic Time series Study (BATS) Hawaii Ocean Time series (HOT) and the European Station for Time series in the Ocean Canary Islands (ESTOC) have been collecting monthly chlorophyll concentration (CHL) measurements over the past two decades through fluorometry and HPLC [Ducklow et al., 2009; Cianca et al, 2012]. Additionally, a couple of moored optical buoys (e.g. Marine Optical BuoY or MOBY off of Hawaii) have been deployed to measure *in situ* radiometric quantities for the purpose of vicariously calibrating ocean color satellite sensors [Clark et al., 2003], but optical buoys are expensive to maintain to the desired

accuracy. The National Oceanographic Data Center's World Ocean Database includes a comprehensive archive of CHL measured *in situ* since 1957 from ships or buoys [Conkright and Gregg, 2003]. These data are high quality and accurate but extremely limited in space and time due to the expense of ship time and the size of the ocean (Figure 2.1). In more than 50 years of sampling, the most frequently sampled regions in the ocean have been sampled less than 50 times (i.e. not even enough to quantify the seasonal signal). Although not a direct *in situ* measurement, the World Ocean Database also includes phytoplankton color index (e.g. Continuous Plankton Recorder surveys in the North Atlantic) and secchi disk observations which have both been correlated to CHL [Raitsos et al., 2005; Boyce et al., 2010]. More recently, the new Bio-Argo float program has begun sampling vertical profiles to help quantify the small fraction of carbon that sinks to depth by the ocean's biological pump [IOCCG, 2011].

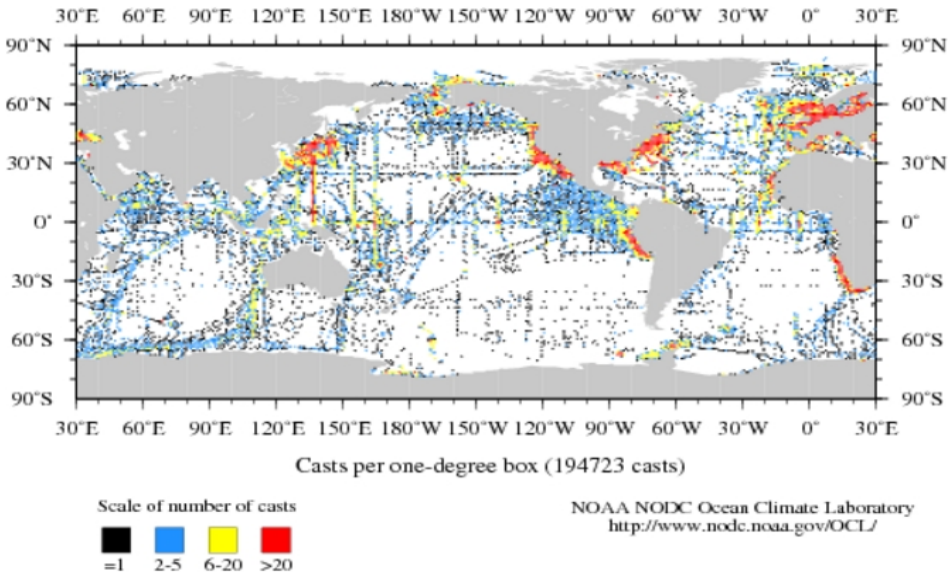

Figure 2.1 NODC Ocean Station CHL casts per 1° bin collected since 1959 and archived within the World Data Center for Oceanography (previously World Ocean Atlas).

In summary, historical *in situ* sampling of marine phytoplankton lacks spatial and temporal coverage and has not been capable of characterizing the global ocean biology. Advances in bio-optical measurement methods, such as underway fluorometry and spectroscopy, have enabled the recent development of autonomous observation of biogeochemical processes in the ocean by profiling floats and gliders that, provided proper calibration and validation, have the potential operate on the temporal and spatial scale of the Argo physical profilers [Uz, 2006; Boss et al., 2008; Johnson et al., 2009].

2.3 Ocean color remote sensing: advantages and limitations

Qualitative characteristics of water at the ocean's surface, including chlorophyll concentrations (CHL), can be seen by bio-optical instruments on aircraft, satellites, and recently the International Space Station. These provide the potential for greater areal coverage than *in situ* measurements ever could, with satellite remote sensing providing the only global data set of ocean surface biology. Satellite ocean color remote sensing began with the limited Coastal Zone Color Scanner (CZCS) proof-of-concept mission in 1978, which established the feasibility of deriving ocean surface CHL from satellite. CHL at the ocean's surface can be derived from satellite observations using a band ratio algorithm of the visible light reflected back to the sensor from Earth after subtracting out the intervening atmosphere. Since the light backscattered to the satellite is primarily from the atmosphere (~90%) with the remainder received from the ocean, atmospheric correction is a critical and ongoing area of active research [Gordon and Clark, 1980; Gordon and Wang, 1994; Gordon, 1997; Wang, 2003]. Thick clouds block the satellite's view of the ocean and result in missing ocean color data. Additionally, small amounts of

dust and sulfate aerosols were shown to bias ocean color retrievals in early processing [Moulin et al., 2001; Schollaert et al., 2003]. Subsequent reprocessing has continually improved methods to remove the aerosol signal [Ziauddin et al., 2010]. After subtracting the atmospheric component, CHL is estimated from the ratio of the green to blue bands based upon a statistical, empirical fit to high quality *in situ* data [O'Reilly et al., 1998].

Although ocean color remote sensing represents CHL at the surface, this information has been extrapolated to the rest of the water column by various studies. Some have used vertical pigment profiles to define relationships between surface CHL and its vertical distribution [Morel and Berthon, 1989; Uitz et al., 2006]. Column-integrated CHL amounts can then be used to quantify biomass and, ultimately, primary production through complex algorithms [Saba et al., 2011]. Others have used ocean color in empirical algorithms to identify phytoplankton species composition [Brown et al., 1994; Subramaniam et al., 2002; Alvain et al., 2005; Hirata et al., 2011]. Phytoplankton community structure shifts have been observed in ocean color due to climate oscillations. For example, in the equatorial Pacific ENSO is linked to shifts from the normal dominance by cyanobacteria (i.e. synechococcus) to fewer of them and increasing nanoeukaryotes during El Niño to larger phytoplankton dominance during La Niña (i.e. diatoms, chlorophytes) [Mackey et al., 2002; Masotti et al., 2011; Rousseaux and Gregg, 2012]. As progress has been made utilizing ocean color data to make connections between biological parameters and their physical forcing mechanisms in disparate regions around the world, the growing length of the record has also enabled discoveries at lower temporal frequencies.

The global ocean color record now includes several satellites with sporadic coverage over several decades. Although CZCS enabled an unprecedented view of biology at the ocean's surface between October, 1978 and June, 1986, the sensor was shared with other missions resulting in intermittent sampling of CHL [Feldman et al., 1989]. The irregular distribution of observations derived from CZCS is displayed in Figure 2.2. The average number of monthly observations in each bin had its maximum for U.S. coastal waters, followed by mid-latitude coastal areas elsewhere, with practically no coverage over the open ocean and at high latitudes. Over the length of the mission, the average number of valid pixels in each scene gradually increased between 1979-1982. The 1982 eruption of El Chicón volcano in Mexico blew more material into the stratospheric aerosol layer than any time since Krakatau in 1883 [Hartmann and Mouginis-Mark, 1999] and interfered with CZCS ocean color retrievals. Following El Chicón, CZCS had reduced coverage (Figure 2.2, bottom), from which it was gradually recovering during its remaining four years before it abruptly quit in 1986.

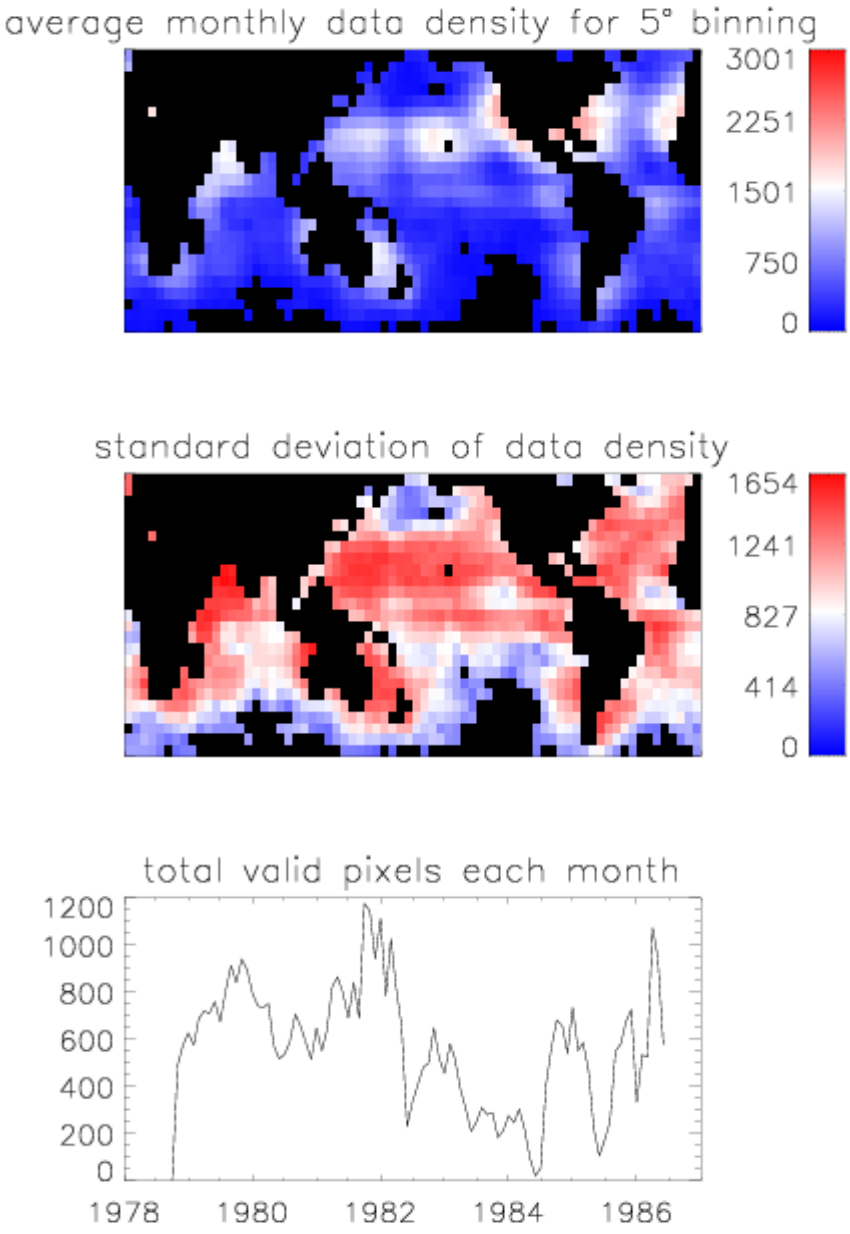

Figure 2.2 CZCS observations. Mission average number of observations per bin (top), mission standard deviation of the number of observations per bin (middle), and total observations per month over the entire mission (bottom). Note that the mission was primarily coastal and focused on US territorial waters, thus largely ignoring the open ocean and other regions outside of the US domain.

After the proof-of-concept CZCS demonstration, several subsequent international missions were launched to provide global ocean color data. Eleven years after CZCS

ended, the first dedicated ocean color sensor, Sea-viewing Wide Field-of-view Sensor (SeaWiFS) was launched in the United States in September, 1997. SeaWiFS provided continuous ocean color data for more than ten years followed by intermittent coverage for another three. The SeaWiFS mission provided stable, dedicated ocean color detection focused on a few key products, including CHL. Stringent accuracy requirements combined with a comprehensive quality control program including regular lunar sensor calibrations and characterization, bio-optical algorithm development, a high quality field measurement program with moored buoys, ships and time series stations, resulted in a data product which surpassed its accuracy goals and remains the highest quality ocean color data set to date [McClain et al., 2004; McClain, 2009].

Several other ocean color sensors have experienced mixed results in longevity or quality since the launch of SeaWiFS. The Japanese Ocean Color Temperature Scanner (OCTS) was operational for less than nine months between 11/1996-6/1997 and suffered temporal degradation which could not be successfully corrected because there was no on-board or stable external calibration source [Shimada et al., 1998]. Similarly, the Japanese Global Imager (GLI) flew for less than an annual cycle between 2002-2003. Two United States' Moderate Resolution Imaging Spectroradiometers (MODIS) multi-disciplinary missions have had different successes. MODIS/Terra was launched in 2000 and continues to the present, but suffers significant sensor degradation and other problems that cannot be adequately corrected for ocean color retrievals [Franz et al., 2005]. MODIS/Aqua began in July, 2002 and continues to the present. Although MODIS/Aqua has suffered from being a multiple-purpose mission and having sensor degradation over time, it is the ocean color sensor that has produced high quality data most similar to that of SeaWiFS

[Meister et al., 2012]. The European Space Agency launched the European Medium Resolution Imaging Spectrometer (MERIS) ocean color sensor in March, 2002, which flew for 10 years until it failed in April, 2012. Several studies have concluded that MERIS needs to be vicariously calibrated for better correlation with the SeaWiFS data [Antoine et al., 2008; Hu et al., 2012]. Additionally, the MERIS swath was much narrower than those of SeaWiFS making global coverage less frequent [McClain, 2009].

Only the SeaWiFS sensor was designed specifically for biological ocean science and is considered the highest quality, attaining the requirements necessary for climate research, with MODIS/Aqua data considered the next best [McClain, 2009]. For this thesis, both SeaWiFS and MODIS/Aqua level 3 9km, monthly CHL were obtained from NASA's Ocean Color web site (http://oceancolor.gsfc.nasa.gov). The 2010 Ocean Color reprocessing includes corrections and inter-calibrations for both SeaWiFS and MODIS/Aqua [NASA GSFC Ocean Color Documents, 2011].

SeaWiFS CHL data through 2008 are used in this study because coverage was increasingly degraded after that, shown in Figure 2.3. Additionally, the physical variables, detailed in section 3.1.1, were available through 2008. For the entire 11+ year time series (9/1997-2008), the 9km resolution CHL were initially averaged over 5^{o} bins with a landmask applied to exclude any pixel within 5^{o} of the coast. Averaging the number of valid data each month shows the seasonal cycle and the drop-off in SeaWiFS coverage toward the end of the mission, with several months having sparse or totally missing data: January-March, 2008, July, 2008, May, 2009, September-October, 2009 before the satellite finally failed in December, 2010. Despite these lapses, the SeaWiFS

mission provided an unprecedented view of global ocean color and the longest, highest climate-quality ocean surface CHL data record.

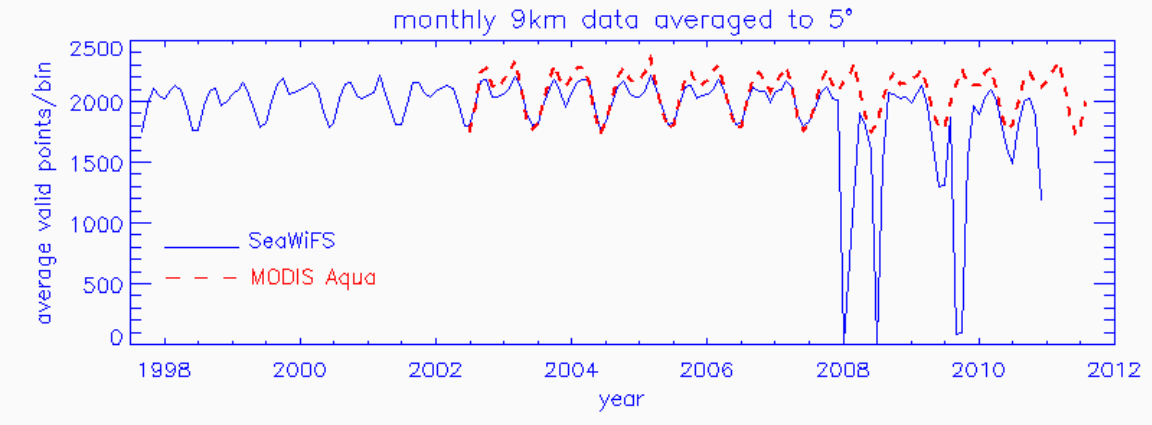

Figure 2.3 **Average number of valid data each month for SeaWiFS and MODIS/Aqua missions.**

The number of valid CHL data in each 5^{o} bin were tallied, averaged, and plotted in Figure 2.4. Areas with the most data coverage generally receive the most sun and the least clouds: the subtropics, where bins could include up to 3600 valid data points in a month. Areas of persistent cloudiness, such as the Intertropical Convergence Zone, and seasonally low solar zenith angle toward the poles have the least coverage. The standard deviation of data density is basically the inverse of the average: areas with lower data density have higher variability. Because the CCA method requires a complete data set, polar regions north of 60^{o}N and south of 60^{o}S were excluded from this study as data there are missing during winter months due to extensive cloud cover and low solar zenith angle. Any other missing data were filled using a three dimensional median filter.

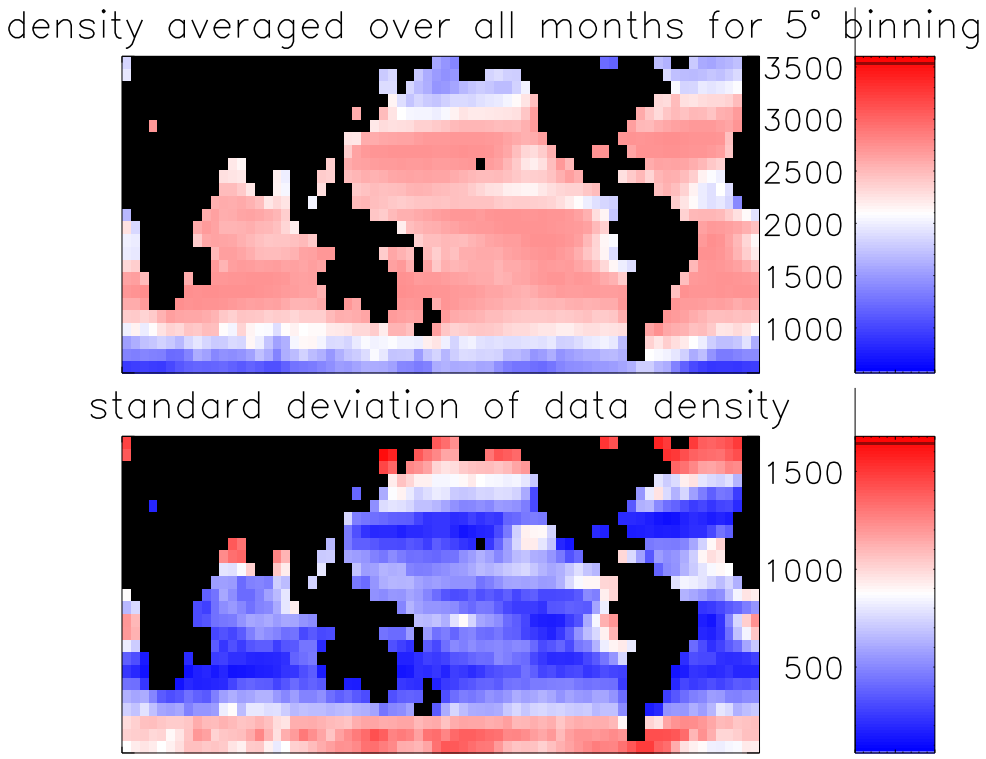

Figure 2.4 **SeaWiFS observations in each bin. Mission average number of observations per bin (top) and standard deviation of the number of observations per bin over the mission (bottom).**

The MODIS/Aqua ocean color data set started in July, 2002 and continues to the present at a relatively stable collection rate, as shown in Figure 2.3. Unfortunately MODIS/Aqua is a more complex system than SeaWiFS and its characterization and calibration has not been at the precision required for ocean-color applications [McClain et al., 2004; Franz, 2011]. MODIS/Aqua has been plagued with signal drift across all channels due to unstable sensor degradation. The SeaWiFS period can be extended during its intermittent coverage using the MODIS/Aqua CHL, which has been tuned and validated against SeaWiFS through 2010 with increased uncertainty after that [Franz, 2011]. Other groups have merged the SeaWiFS and Aqua data sets, but they only include data when both sensors are available and leave gaps whenever SeaWiFS data are missing

[Maritorena and Siegel, 2005; Bryan Franz, personal communication]. There are 99 months when both SeaWiFS and MODIS/Aqua are available: July, 2002 through December, 2010 (minus months missing SeaWiFS coverage). Time series from the two satellites indicate fairly good consistency between the sensors (Figure 2.5).

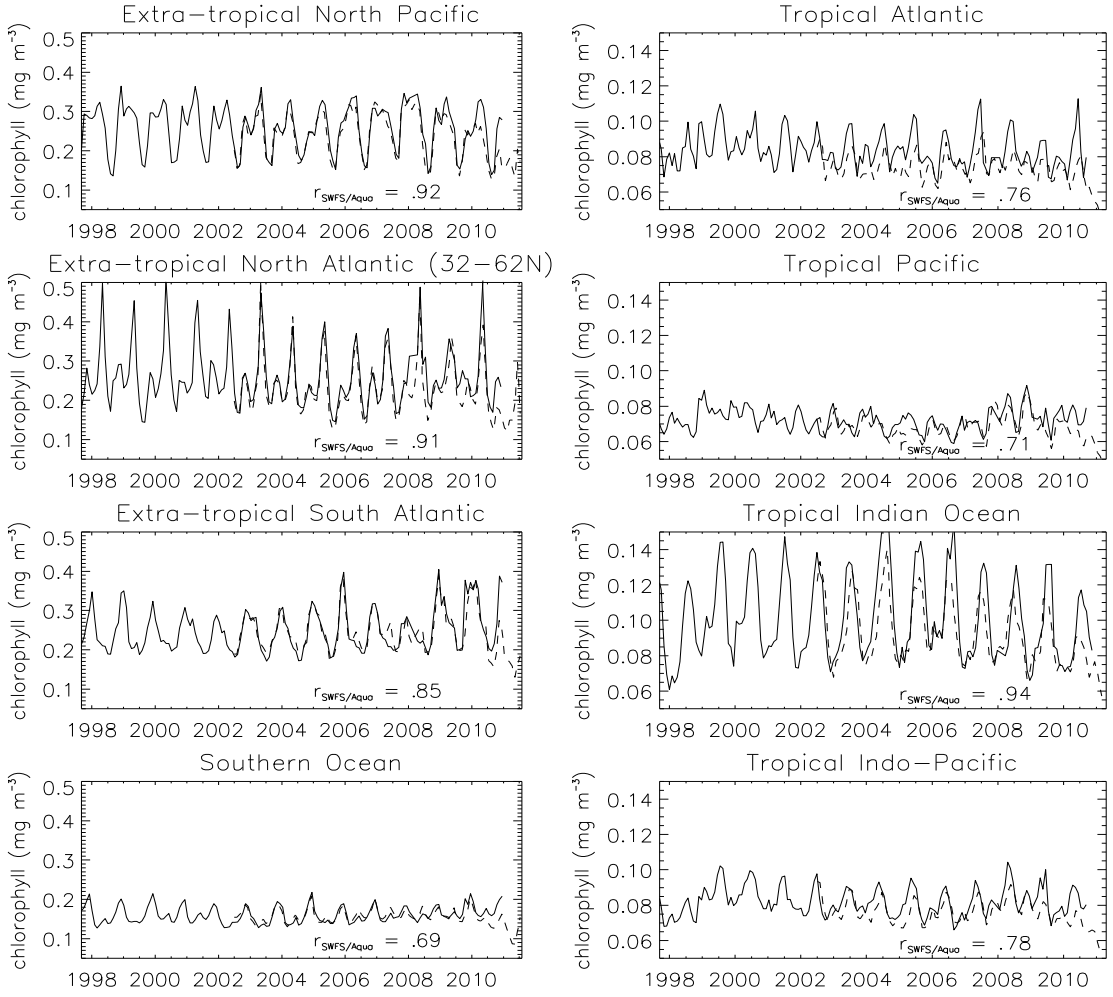

Figure 2.5 CHL average time series over extra-tropical (left column) and tropical regions for SeaWiFS (—) and MODIS/Aqua (----). Correlation coefficients for overlapping times are annotated.

The CHL annual cycle is evident in Figure 2.5, including a pronounced spring bloom in most locations and a secondary fall bloom in many regions, most notably in the extra-tropical North Pacific and Atlantic. SeaWiFS and MODIS/Aqua generally match-

up best where CHL are highest (e.g. Indian Ocean, r = 0.94, North Pacific, r = 0.92, and North Atlantic, r = 0.91) and worst in areas of low CHL or sparse coverage (e.g. Southern Ocean, r = 0.69). The Southern Ocean has low CHL due to its iron limitation [de Baar et al., 1995; Boyd et al., 2007] and is notorious for ubiquitous, fast-moving clouds making satellite coverage of the ocean infrequent. Imperfect match-ups between the two satellites are likely caused by seeing different areas and days within a monthly, 5^{o} bin. CHL anomalies were calculated by removing the seasonal climatology. A linear relationship between SeaWiFS and Aqua was calculated for co-located bins (Figure 2.6). While the bias is negligible, there is an average slope of approximately 0.77 (SeaWiFS/Aqua).

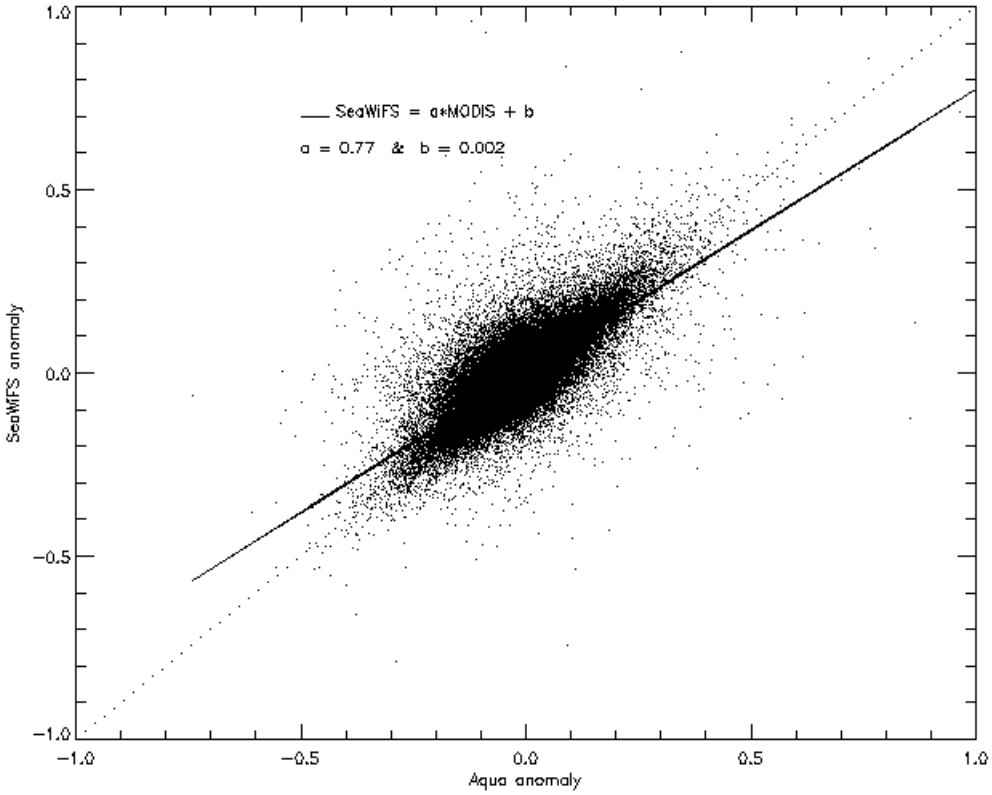

Figure 2.6 CHL deseasoned anomalies for all 5^{o} bins: MODIS/Aqua plotted against SeaWiFS, with their regression line superimposed. The dotted diagonal line demonstrates a slope of 1.

Since the sensors cover slightly different eras, CHL concentrations compared with climatology not removed indicate an average slope of 0.9 (SeaWiFS/Aqua) and a global distribution shown in Figure 2.7. For most of the global ocean, SeaWiFS values are slightly lower than Aqua, except in the western tropical Pacific warm pool and small areas west of Africa and South America where Aqua had up to 40% lower CHL concentrations. SeaWiFS is the preferred data set, but Aqua can be used to fill its gaps.

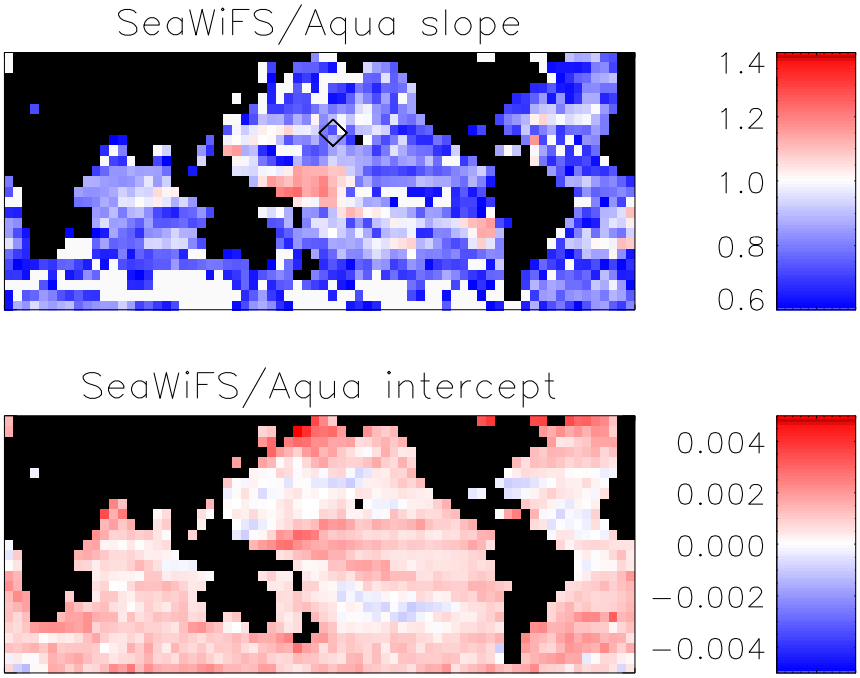

Figure 2.7 Distribution of SeaWiFS/MODIS/Aqua CHL: slope (top) and intercept (bottom).

The slope and intercept of theses regressions were used to scale the Aqua time series. Combining the data using scaled or unscaled Aqua indicates little difference in most regions, but scaling appears to slightly damp CHL concentrations in the subtropical North Pacific and North Atlantic and to slightly amplify them in the western tropical Pacific and Indian Ocean. Empirical Orthogonal Functions calculated on both yielded insignificant differences in mode partitioning and times of greatest variability.

2.4 Biological variability in space and time

To determine optimal binning in space and time to filter out mesoscale features, autocorrelations were first calculated on 5° bins (Figure 2.8). Temporal decorrelation scales are approximately 1.5 months for subpolar waters; 2 months, subtropical; 3.5 months near the equator. The zonal spatial decorrelation scales are about 20° at subpolar latitudes and about 30° at subtropical and equatorial latitudes. Meridional spatial decorrelation scales are about 5° for the subpolar region and 5-10° for subtropical and tropical regions. That zonal autocorrelations are larger than meridional indicates greater zonal homogeneity. Meridional autocorrelations are not shown because there are only three or four bins per region. Autocorrelations of the monthly, 2° CHL are comparable.

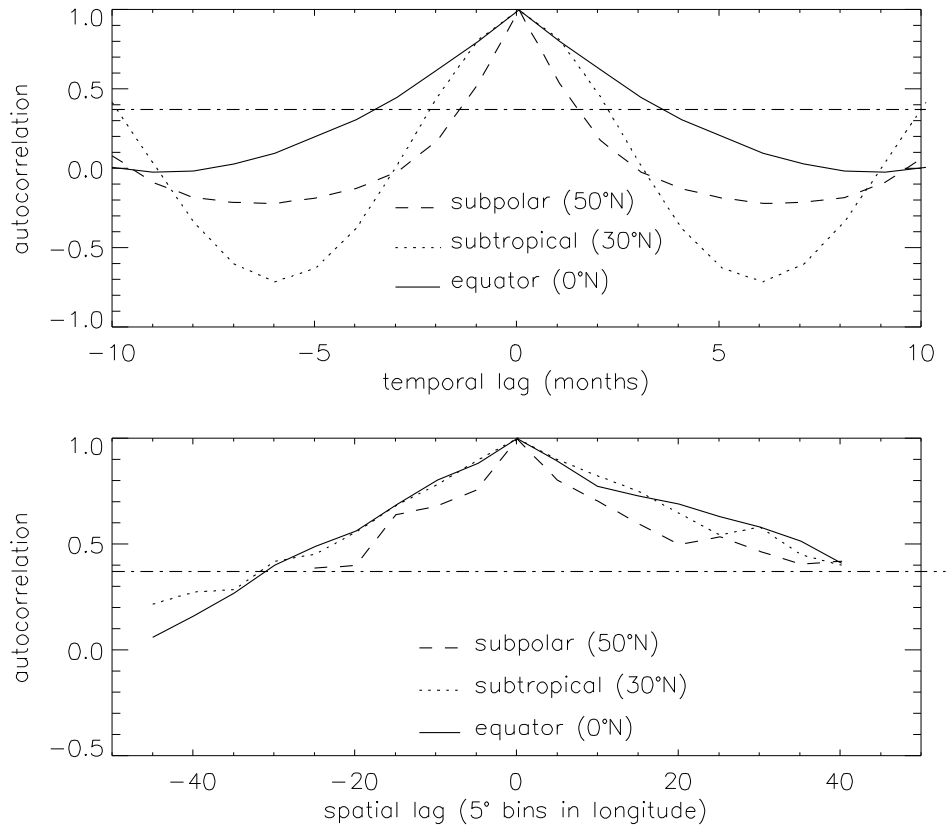

Figure 2.8 Decorrelation scales for monthly, 5° SeaWiFS chlorophyll averaged over three regions (subpolar, subtropical, equatorial): temporal decorrelations (top); zonal decorrelations (bottom).

Monthly CHL anomalies were calculated by removing the seasonal cycle. All calculations were first log-transformed because CHL approximates a log-normal distribution [Campbell, 1995], with its global variability shown in Figure 2.9. Greatest CHL variability is in areas of high productivity where ocean circulation brings nutrients from depth into the euphotic zone: subpolar gyres and upwelling regions along the equator and coasts. There is also high variability in areas of coastal run-off or where storm-induced mixing provides intermittent nutrients leading to episodic blooms, such as the Indian Ocean. Areas of very low variability include the central subtropical gyres, also areas of low CHL concentration (recall Figure 1.2). The eastern equatorial Pacific upwelling region has relatively constant productivity, but is known to be iron-limited [Martin et al. 1994]. Regions of high CHL values typify an exponential distribution, however the exponential distribution has been found to be unsuitable for the oligotrophic waters of the tropics and subtropics [Sapiano et al., 2012]. Because the remainder of this study focuses on the tropical Pacific, the CHL calculations are not transformed.

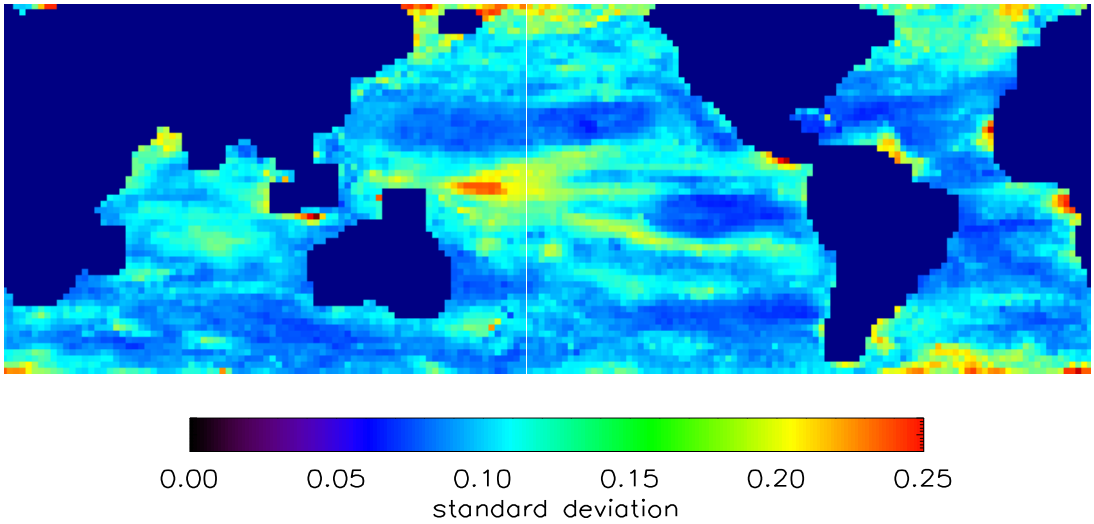

Figure 2.9 Monthly deseasoned CHL anomaly standard deviations (mg m^{-3}) between 1998-2008, with calculations performed on log-transformed CHL. Hereafter CHL are not log-transformed.

2.4.1 Physical oceanography impacting biology

Cross-correlating CHL to physical variables provides indications about the physical-biological interactions leading to phytoplankton blooms. Simple Ocean Data Assimilation (SODA) version 2.1.6 includes temperature and salinity profiles, ocean currents, wind stress and SSH available 1958 through 2008 [Carton and Giese, 2008]. The MLD was calculated as the level at which there is a 0.2C difference from the temperature at 5m depth. CHL anomalies cross-correlated against SST anomalies (Figure 2.10, top) yield predominately inverse values in the tropics and subtropics, which depend upon cooler, deeper water to fertilize the surface, while the relationship is more positive in subpolar regions that tend to be light-limited and require stratification so that the nutrient-rich surface layer does not extend deeper than the euphotic zone. The cross-correlation of CHL to MLD (Figure 2.10, bottom) is generally positive in areas that are nutrient-limited (e.g. subtropics); negative in areas that are light limited (e.g. subpolar gyres), with more complicated dynamics along the equator.

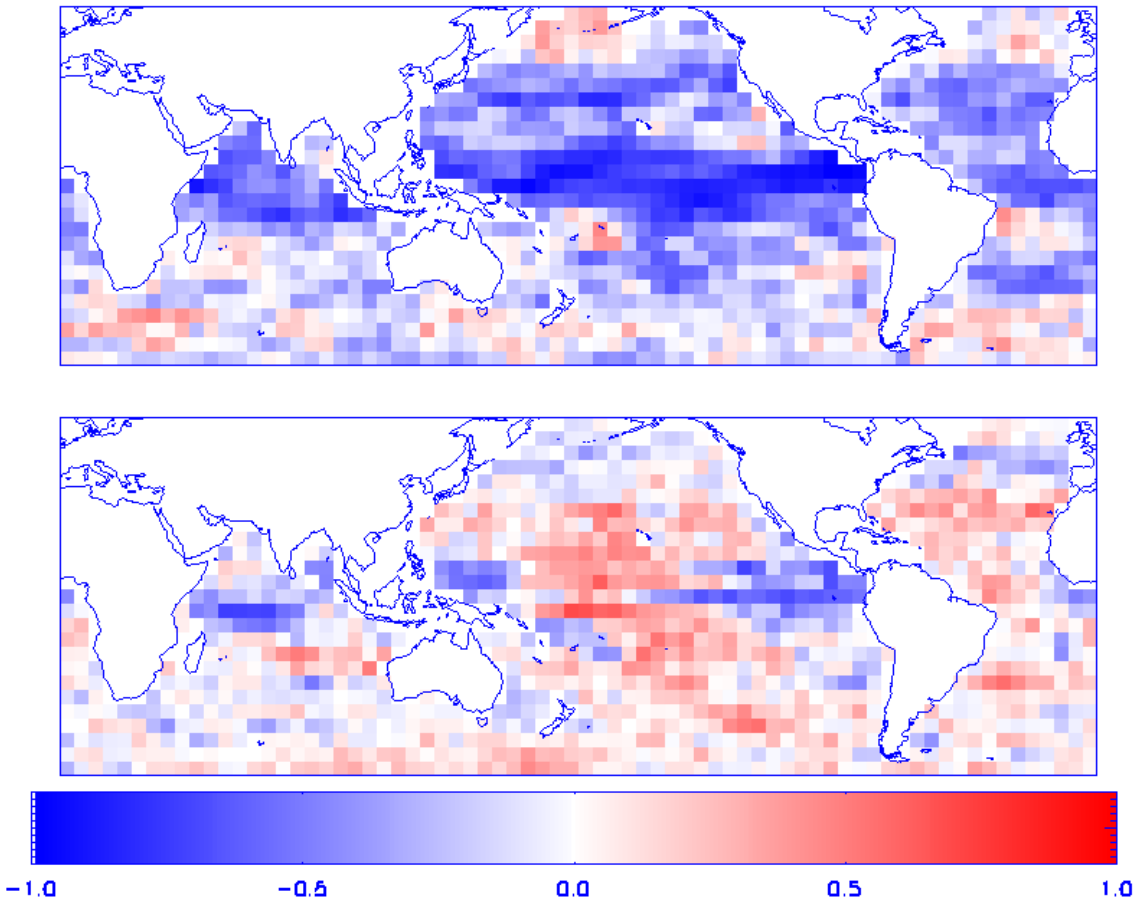

Figure 2.10 **Temporal cross-correlations between monthly anomalies of CHL and SST (top) and between monthly anomalies of CHL and MLD (bottom) for the period 9/1997-2008.**

The other physical variable correlated with nutrient supply to the surface layer is wind, which can cause mixing, divergence and Ekman pumping. Off the equator, the Earth's rotation causes the surface currents to veer to the right of the winds in the Northern Hemisphere and to the left in the Southern Hemisphere. In the tropics, this means easterly trade winds cause northward (southward) flow north (south) of the equator and divergence at the surface that causes upwelling along the equator. Planetary vorticity (f) is positive to the right in the Northern Hemisphere and negative to the left in the Southern. North of the equator, positive wind stress curl ($\nabla \times \tau$) results in upward

motion, called Ekman suction (+w_{ek}); negative wind stress curl (-$\nabla \times \tau$) results in downward pumping (-w_{ek}):

$$w_{ek} = \frac{1}{\rho_{ref}} \hat{z} \cdot \nabla \times \left(\frac{\tau_{wind}}{f} \right) \quad (1)$$

Kahru et al. [2010] cross-correlated CHL with winds from the assimilation of satellite-derived winds and conventional winds. Figure 2.11 shows these have a very similar distribution to that between MLD and CHL, not surprising since the MLD results from wind-driven ocean circulation and mixing. They find more complex relationships in equatorial regions. Specifically in the equatorial Pacific, they note that high winds in the west correspond to easterly or westward winds that induce equatorial divergence and upwelling combined with other possible mechanisms including the island effect of the Kiribati Islands [Messie et al., 2006], changes in the New Guinea coast currents that supply iron to the equatorial undercurrent [Ryan et al., 2006], and the erosion of the west Pacific warm pool barrier layer [Murtugudde et al., 1999]. Warm and cool anomalies generally appear in the western and central basin and propagate eastward as equatorial Kelvin waves [Carton and Giese, 2008]. Kahru et al. [2010] ascribe the negative correlation between wind speed and CHL in the east to remote forcing: westerly wind bursts in the west Pacific warm pool region generate equatorial Kelvin waves that propagate eastward and deepen the nutricline in the east, reducing biological productivity. Conversely, low winds in the west precede shoaling of the nutricline in the east and enhanced productivity. Despite many interesting regional variations in physical-biological processes, it is clear from the global maps of cross-correlations between

physical variables and CHL that the strongest correlations are found in the tropical Pacific, which is known for having strong physics that dominates the biology there.

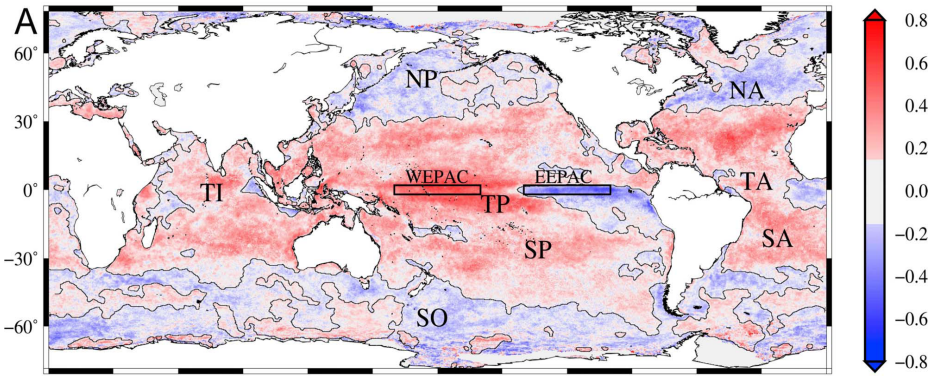

Figure 2.11 **Cross-correlations (r) between monthly anomalies of wind speed and CHL at 25km resolution for the period 11/1996-2008. Black curves indicate r = 0. Western equatorial Pacific and eastern equatorial Pacific are delineated. Figure adapted from Kahru et al. [2010].**

2.4.2 Importance of the tropical Pacific

Understanding large-scale, low frequency biological patterns in the tropical Pacific, the largest ocean region on Earth, will deepen our understanding of the impact of the climate system upon primary production. Many phenomena experienced there have global repercussions: upwelling in the eastern and equatorial Pacific provides nutrients to fuel roughly 20% of new biological production in the world's oceans [Chavez and Togweiller, 1995; McPhaden and Zhang, 2002] and causes the largest oceanic source of CO_2 to the atmosphere [Feely et al., 1999]; the west Pacific warm pool is the location of greatest oceanic heat source to the atmosphere [Lewis et al., 1990]. The tropical Pacific ocean drives ENSO, the strongest climate-scale oscillation with a quasi-periodic frequency typically between 2-7 years, whose effects are wide-ranging and central to

understanding the global ocean-atmosphere system. Equatorial physical-biological patterns have been observed to be more complex due to regional interactions, making isolating and understanding them a high priority [Kahru et al., 2010]. Recent efforts have begun addressing complex biophysical patterns through observations and modeling studies in the vast expanse of the equatorial Pacific [Murtugudde et al., 2002; Chavez et al., 2003; Marzeion et al., 2005; Messie and Radenac, 2006; Wang et al., 2009; Kahru et al., 2010; Jin et al., 2013].

The equatorial Pacific is characterized by two main features at the surface: a mesotrophic, upwelling cold tongue that extends from the east toward the center and an oligotrophic, low salinity, surface warm pool in the west. Geostrophic flow is anticyclonic around the subtropical gyres in the North and South Pacific. Near the equator, easterly trade winds prevail and are strongest east of the dateline. West of the dateline, winds are lighter and decrease almost to zero near the warm pool. The easterly winds cause zonal flow to the west, where the western boundary causes water to pile up setting up a west-east pressure gradient. On the equator, the apparent force caused by the Earth's rotation goes to zero and forcing mechanisms that determine basin scale fluid motion are a balance between the zonal pressure gradient and wind-induced vertical frictional forces:

$$0 = -\frac{1}{\rho_{ref}}\frac{\partial p}{\partial x} + \frac{1}{\rho_{ref}}\frac{\partial \tau_x}{\partial z} \tag{2}$$

The most general description of oceanic flow in the tropical Pacific is that it is primarily zonal. A cross-section at 140°W highlights permanent currents in the region (Figure 2.12).

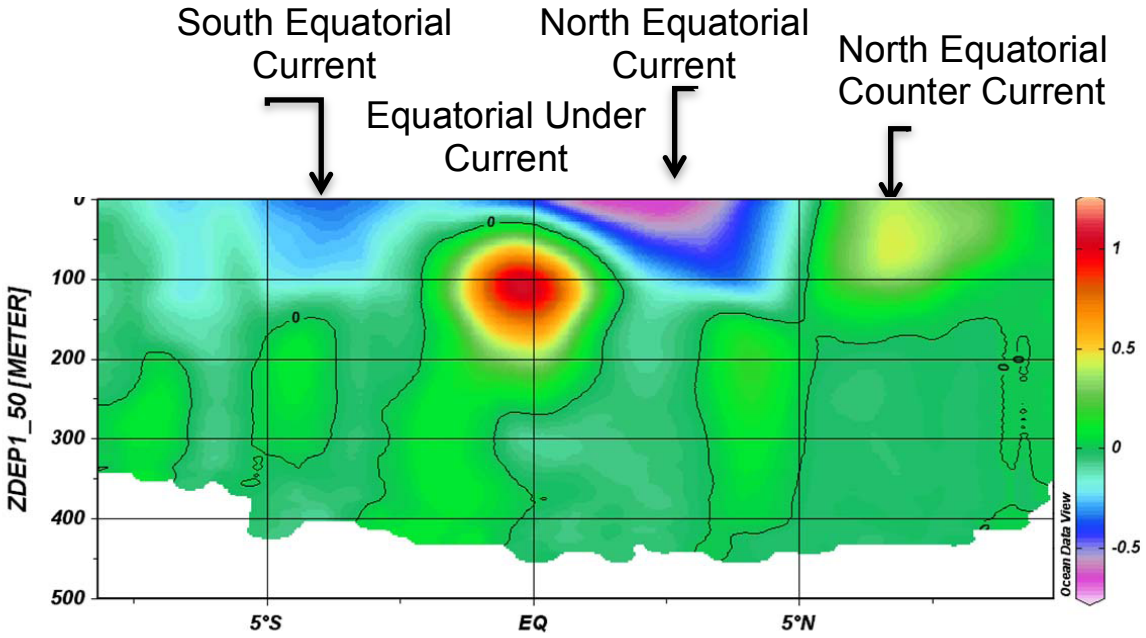

Figure 2.12 Zonal velocity (m/s) for a TAO array transect at 140°W plotted by Ocean Data View.

The westward flowing North and South Equatorial Currents (~2°N and ~4°S) converge in the surface waters of the west Pacific warm pool and become the North Equatorial Counter Current flowing eastward around 7°N, driven by the Intertropical Convergence Zone winds [Talley et al., 2011]. Below the surface on the equator, around 100m at 140°W, the eastward strong (>100cm sec^{-1}) Equatorial Undercurrent (EUC) follows the pycnocline, deepest in the west about 200m, sloping upward toward the eastern boundary and the upwelling zone in the east [Knauss, 1960]. Surface CHL decreases in the east during El Niño and increase during La Niña have been linked to the shut down and recommencement of the iron-rich EUC [Wilson and Adamec, 2001]. Coincident with the EUC, sub-surface CHL sampled along the equatorial upwelling region has revealed that

the deep CHL maximum slopes upward from west to east, sampled at 50-60m in the central equatorial Pacific to 30-40m in the east [Pennington et al., 2006]. Below the EUC lies a weaker westward flow extending down to 1000m, with eastward subsurface countercurrents on either side of the equator [Tsuchiya, 1972].

At the surface, easterly trade winds along the equator combine with Ekman transport north and south of the equator to cause divergence at the surface resulting in equatorial upwelling. Observations indicate that nearly 80% of the variance in near-surface current in the tropical Pacific is caused by the climatological geostrophic field (63%) and wind-driven currents in the presence of stratification (15%) [Ralph and Niiler, 1999; Lagerloef et al., 1999]. Seasonal variation in Ekman pumping (i.e. equation 1 in section 2.4.1) and planetary Rossby waves affect current positions. For example, the North Equatorial Counter Current shifts northward (southward) of its annual mean on the western (eastern) boundary during the first half of the year, reversing the pattern during the second half of the year [Hsin and Qiu, 2012].

The boundary is dynamic between the oligotrophic, low salinity, surface warm pool in the west and the mesotrophic, upwelling cold tongue that extends from the east toward the central equatorial Pacific. This boundary extends vertically deeper than 40m impacting seasonal-to-interannual climate variations by its effect upon large-scale atmospheric convection [Maes et al., 2010]. The eastern edge of the strongly stratified warm pool has distinct hydrological features and ecosystem dynamics [Le Borgne et al., 2002], yet has no clear physical front and is thus problematic to identify at the surface using temperature. Surface CHL is a more robust, reliable indicator of the warm pool's eastward extent [Maes et al., 2010], shown in a CHL profile (Figure 2.13).

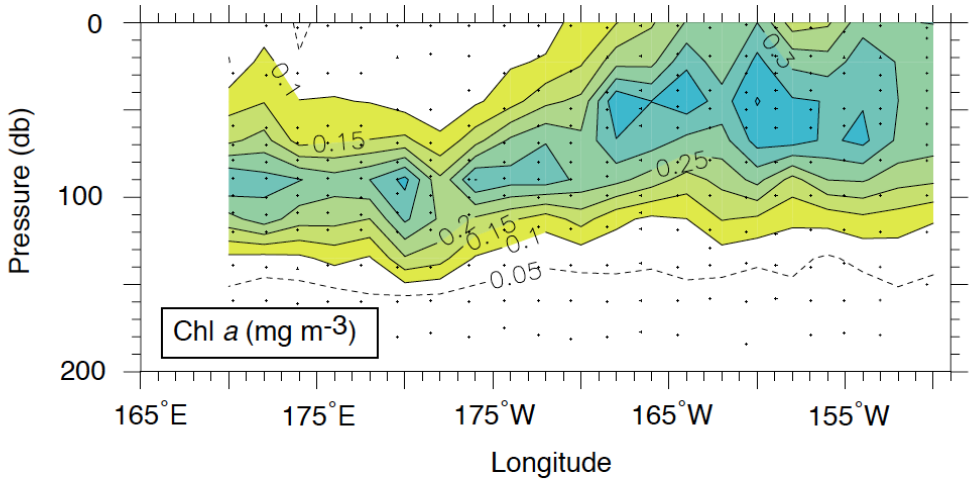

Figure 2.13 **Vertical, zonal distribution of CHL on the equator at the eastern edge of the warm pool collected by the Flupac cruise in September-October 1994 and contoured at 0.05 mg m^{-3} intervals. Figure adapted from Le Borgne et al. [2002].**

There is evidence that slow changes have been taking place in physical and biological processes in the tropical Pacific, caused by some combination of natural decadal-scale oscillations and anthropogenic effects [McPhaden and Zhang, 2002].

2.4.3 Characteristics of surface chlorophyll in the tropical Pacific

Variability in tropical Pacific CHL is primarily due to the basin-scale modification of the west-east thermocline slope during ENSO that impacts nutrient supply and primary production across the basin [Barber and Chavez, 1983; Radenac and Rodier, 1996]. In the neutral phase, the thermocline and nutricline are extremely deep in the western equatorial Pacific (approaching 200m). The surface layer is strongly stratified with a warm SST (~ 30C), which has been noted as the strongest oceanic heat source to the atmosphere resulting in convection and a fresher surface caused by precipitation [Gill and Rasmusson, 1983]. The thermocline and nutricline slope upward toward the surface,

rising to less than 50m in the east. Upwelling results in cooler SST along the coast and equator in the eastern tropical Pacific as well as higher biological productivity due to the nutrients it lifts into the sunlit surface waters. Zonal displacement of the warm pool has been identified as an indicator of the onset of the El Niño phase of ENSO [Picaut et al., 1996; Picaut et al., 1997]. At the initiation of an El Niño, a Kelvin wave displaces the stratified, oligotrophic warm pool eastward. Deeper water then rises in the west behind it, evident in temperature profiles as well as the depth of the Equatorial Undercurrent, lifting nutrients toward the euphotic zone. During El Niño conditions, the trade winds weaken and the NECC strengthens [Hsin and Qui, 2012], the west-east slope of the thermocline flattens and the thermocline deepens in the east. With the suppression of upwelling and its supply of cool, nutrient-rich deep water, primary production drops and SST increases in the east. The maximum SST anomaly is sometimes in the east (Niño3 region: 150^{o}W-90^{o}W) and sometimes in the central equatorial Pacific (Niño4 region: 160^{o}E – 150^{o}W) [Ashok et al., 2007; Lee and McPhaden, 2010; Radenac et al., 2012].

For this regional study of tropical Pacific biology, the CHL data were binned to 2^{o} because the finer resolution provides useful meridional information, particularly in the vicinity of the equator. The top panel of Figure 2.14 shows tropical Pacific surface CHL at 2^{o} monthly resolution averaged between 9/1997-2008. Upwelling along the west coast of the Americas causes average CHL values greater than 0.4 mg m^{-3}; equatorial upwelling fuels CHL concentrations greater than 0.3 mg m^{-3} with a gradual decrease westward away from the coast. The west Pacific warm pool is evidenced by CHL values less than 0.1 mg m^{-3} [Maes et al., 2010]. In the lower panel, time series for the two pixels marked as * and ◊ on the map roughly represent the east side of the west Pacific warm

pool and the westward extension of the cold tongue respectively. Overall, the ◊ location had higher CHL. The * location normally had less than 0.1 mg m^{-3}; CHL values more than 0.1 mg m^{-3} indicate times when the warm pool was displaced away from the * location. Months when the dashed line fell below 0.1 mg m^{-3} and the * values were greater coincide with El Niño during DJF: 1997/98, 2002/03, and 2006/07 as the warm pool moved east. The 1997/98 El Niño event completely disrupted the biogeochemical properties in the equatorial Pacific [Chavez et al., 1999] and resulted in an extreme spike in surface CHL values at the * pixel (> 0.4 mg m^{-3}) as the EUC and nutricline shoaled [McClain et al., 1999], raising nutrients into the euphotic zone there.

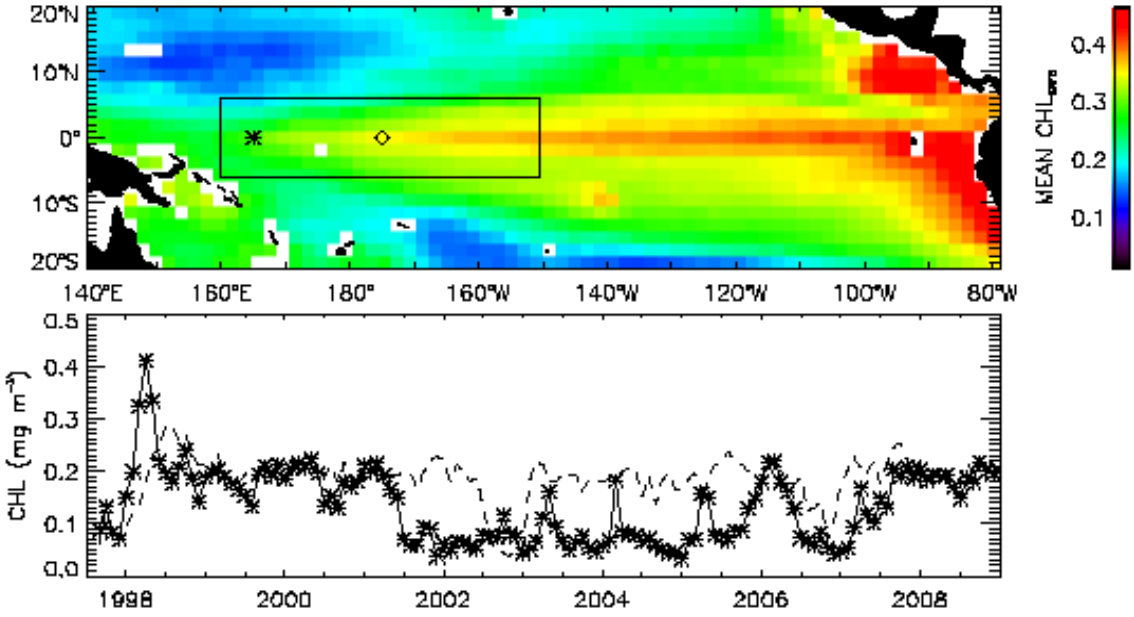

Figure 2.14 **Top: monthly averaged CHL (mg m^{-3}) where ◊ is roughly the western edge of equatorial upwelling and * marks the eastern edge of the warm pool. Bottom: monthly values at each spot.**

The standard deviations (σ) of the monthly CHL anomalies (Figure 2.15) generally approximate the average map with greater variability in areas of higher values. A notable exception is the * location, which has higher σ than other oligotrophic areas. Positioned at the east side of the warm pool, there the normally deep nutricline is maintained by a salinity barrier layer, except when the warm pool propagates east during El Niño and west during La Niña. With the reduction in barrier thickness, a shoaling EUC and nutricline [McClain et al., 1999] combine with upwelling-favorable winds and enhanced vertical mixing to supply nutrients to the euphotic zone [Ryan et al., 2002], as evidenced in early 1998 with the extremely high CHL anomaly at * which was twice the maximum value of any anomaly during the rest of the time series that otherwise ranged +/-0.1 mg m^{-3} at both locations. Anomalies at the ◊ cold tongue location fell lower than at * during El Niño events of 1997/98 and 2002/03 when the warm pool shifted eastward.

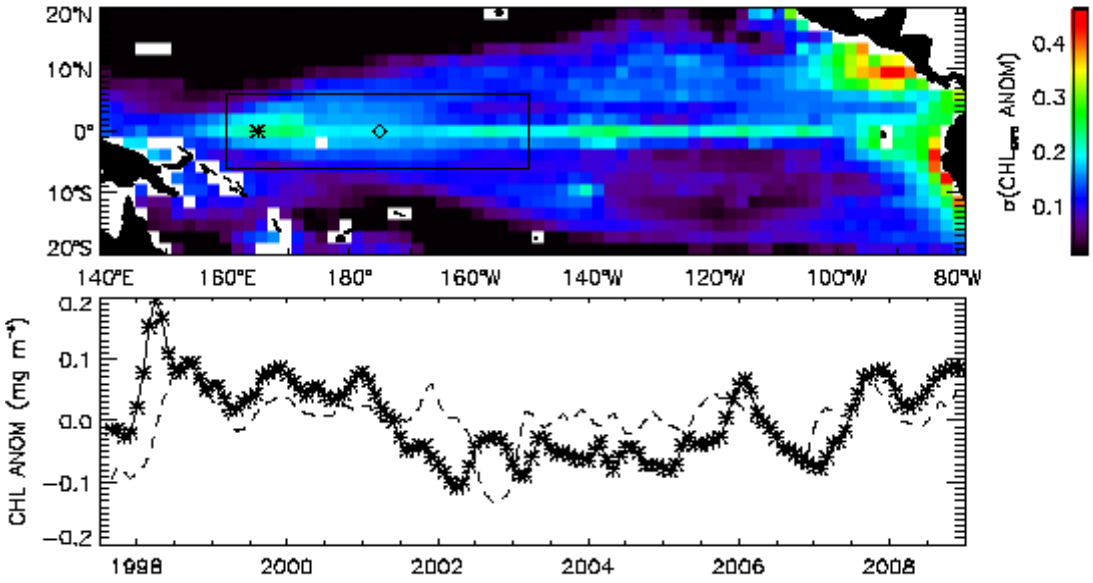

Figure 2.15 Standard deviation of deseasoned monthly CHL (mg m^{-3}). As in Figure 2.14, ◊ is toward the western edge of the equatorial upwelling and * is near the eastern edge of the warm pool. The bottom panel shows the timeseries for these two spots as a dashed line and *'s, respectively.

While the standard deviation map indicates areas where CHL varies, a signal-to-noise ratio (SNR) map gives information about changes that are likely to be detected and those that may be missed by an observation method. Here the SNR was calculated by equation (3) following the method of Ballabrera-Poy et al. [2003]:

$$SNR = (\sigma_a^2 + \sigma_s^2)/\sigma_e^2 \quad (3)$$

where the 'signal' is the sum of the interannual and seasonal variance ($\sigma_a^2 + \sigma_s^2$) and the 'noise' is the estimated error (σ_e^2), all shown in Figure 2.16.

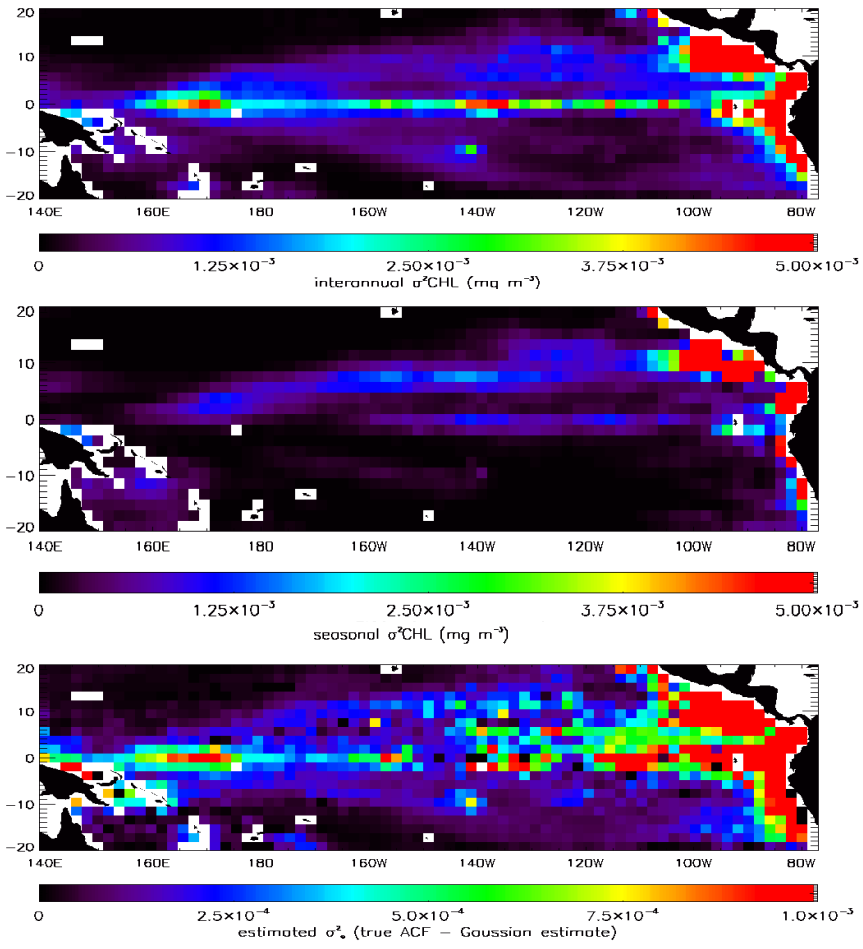

Figure 2.16 Interannual CHL variance (top); seasonal CHL variance (middle); estimated temporal error (bottom).

At basin-scale, the signal in monthly, 2^o ocean color CHL is primarily due to interannual and seasonal variance while the noise is due to features that cannot be adequately sampled at this resolution. The interannual variance, σ_a^2 (Figure 2.16, top), is the variance in the deseasoned CHL (September, 1997 - December, 2008). The seasonal variance, σ_s^2 (Figure 2.16, middle), is the variance in the seasonal climatology. Interannual variance is dominated by changes caused by ENSO primarily along the equator and along the coast of the Americas. Seasonal variance is smaller, due to the near-absence of seasons in the tropics, with some seasonality along the coast in the east, especially in the vicinity of the Costa Rica dome, and some weaker signal along the equator and slightly north in the vicinity of the seasonal migration of the Intertropical Convergence Zone which reaches its maximum northward extent in July and southward extent in January.

The estimated temporal error, σ_e^2 (Figure 2.16, bottom), is the difference at the origin between the autocorrelation function of the CHL anomalies and a Gaussian curve fitted to the data for lags of 1 to 5 months, by the method of Ballabrera-Poy et al. [2003]. Unresolved variability or noise is greatest in areas that do not conform to the Gaussian curve fit to the temporal autocorrelation, this could be caused by fast or transient features that are apparent within one month but negligible in the months before and after. Low noise means there is a good fit between the autocorrelation function and the Gaussian curve implying little or slow changes between zero and five months. Both the greatest signal and greatest noise are in the east toward the coast of the Americas as well as along the equator. As shown in Figure 2.17, the SNR is greatest where signal is greater than estimated noise: within ~10^o of the equator from the west Pacific eastward to about

140°W, due to the higher interannual and seasonal variance there as well as the relative low error or good fit between the autocorrelation at zero lag and its Gaussian estimate.

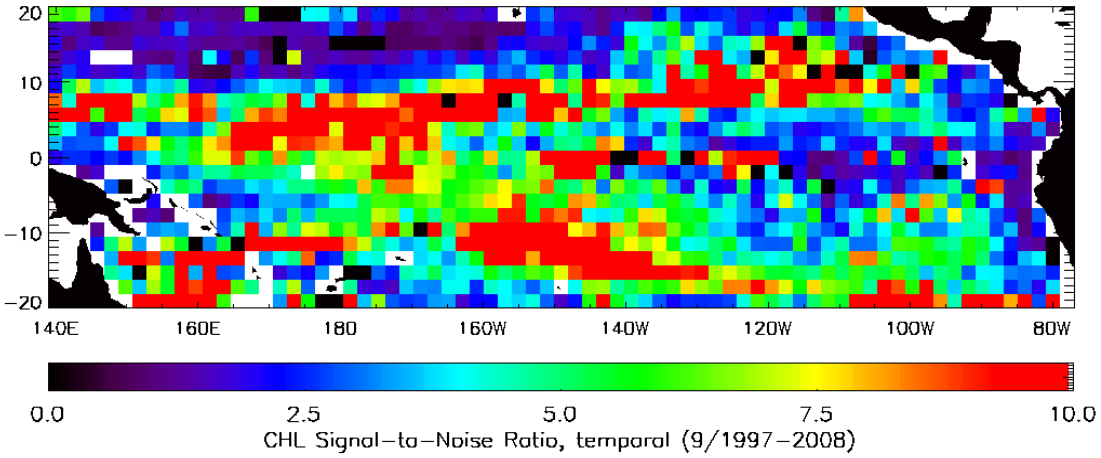

Figure 2.17 Signal-to-noise ratio (SNR) where signal is defined as the combination of interannual and seasonal variance, noise is estimated from the difference between the temporal auto-correlation function and a Gaussian fit for monthly CHL: 9/1997-12/2008.

Oligotrophic gyres generally have low signal-to-noise ratios except in active bloom areas with large contrast [Uz and Yoder, 2004; Yoder et al., 2010; Beaulieu et al., 2013]. The higher SNR north of the equator toward the east Pacific is due to a combination of greater seasonal variance combined with lower estimated error, implying slow patterns that vary gradually between months (e.g. seasonal migration of the Intertropical Convergence Zone). The low SNR in the productive east Pacific cold tongue along the equator east of 140°W occurs where estimated noise exceeds signal, due in part to fast-moving and transient features that are present one month but gone the next, such as tropical instability waves [Evans et al., 2009].

To summarize this chapter on observing biological processes in the ocean, there is a lack of complete spatial coverage by direct measurements and limited range of temporal

coverage by the ocean color satellite record to determine climate-scale variability of biological features in the ocean directly. Existing measurements have enabled an understanding of physical forcing that leads to nutrient entrainment and blooms. The most prominent biological-physical patterns in the existing data sets are in the tropical Pacific, due to its size and the dominance of ENSO as an interannual perturbation. Across the equatorial Pacific, El Niño is associated with weaker easterly trade winds, reduced upwelling, warmer temperatures and reduced CHL in the east and increased CHL in the west Pacific warm pool. Its opposite phase, La Niña, is associated with stronger easterly trade winds, enhanced upwelling, an extended cold tongue, and greater equatorial divergence leading to enhanced Ekman pumping north and south of the equator that causes upwelling, supplying nutrients leading to higher CHL anomalies in the east. In this area of the tropical Pacific, where climate-scale forcing is great and CHL variability is tightly coupled to physical variability, longer physical records have good potential to serve as proxies to statistically reconstruct and extend the CHL record.

3. Reconstructing chlorophyll over 51 years

Continuous coverage by satellite ocean color sensors over a decade confirmed that primary production is forced by basin-scale physical oceanography, which is likewise linked to atmospheric circulation. Winds force underlying ocean currents, transporting nutrients that fuel primary production and ultimately higher trophic levels in the food web. As wind-driven patterns in the ocean reflect the shifting circulation of the atmosphere, it is increasingly evident that the atmosphere oscillates at various frequencies. How the atmosphere drives or is driven by the ocean at multi-decadal and lower frequencies is an open question. Climate-scale oscillations have been noticed in fish catch records, although the link between climate forcing and ocean biology is not clearly understood. As discussed in the previous chapter, an observational record that resolves climate-scale ocean biology does not currently exist. Numerical modeling alone is also unlikely to resolve climate-scale ocean biology because many key parameters describing interactions between biological variables remain poorly constrained. A long-term view of ocean color CHL over several decades or more would enhance our understanding of low frequency changes happening at the base of the food web. This chapter compares physical variables best correlated to ocean biology, gives an overview of statistical reconstruction methods used to extend limited data sets with longer proxy records, uses them to reconstruct CHL in the tropical Pacific, and clarifies what can and cannot be reconstructed by a statistical reconstruction including its uncertainties. The reconstruction is evaluated by comparison with several independent fields. The training CHL data set includes the largest El Niño of the century in 1997/98, followed by its transition to La Niña in 1998/99, which merits additional discussion as a unique case.

3.1 The statistical reconstruction

When data sets are correlated, analytical techniques may be applied to determine scales of variability and a denser record may be used to increase the resolution of a sparser record through statistical reconstruction by one of two methods: 1) interpolation of an existing data set that has gaps [e.g. Alvera-Azcárate et al., 2007] or 2) using the covariance between variables to extend the sparser data set [e.g. Smith et al., 1996]. An example of the latter, canonical correlation analysis (CCA) is a statistical method for quantifying linear relationships among two or more variables in order to reconstruct a target variable [Barnett and Preisendorfer, 1987; Bretherton et al., 1992]. Other statistical methods exist, such as multi-variate empirical orthogonal functions (MEOF) [Alvera-Azcárate et al., 2007]. Because the connection between ocean and atmospheric dynamics is well-established at seasonal and interannual scales, a growing body of physical climate variables have been extended back in time using such statistical reconstructions: sea-surface temperatures, back to the 1860s [Smith et al., 1996; Kaplan et al., 1998]; marine sea level pressure, back to the 1850s [Kaplan et al, 2000]; sea level, back to 1950 [Church et al., 2004]; land-air-sea surface temperatures, back to 1880 [Smith and Reynolds, 2005]; oceanic precipitation, back to 1900 [Smith et al., 2009]. While statistical reconstructions have successfully extended physical variables by more than a century, the method has not previously been used to extend biological data back in time.

In the current observational records, ocean biology has been correlated to physics over interannual, basin scales. Understanding the connection between lower frequency physical forcing and ocean biology is less well-understood because the longest complete high resolution data set, sea surface CHL, spans just over a decade. Here a statistical

reconstruction of CHL is developed, exploiting its correlation with longer-record physical variables that relate to the depth of the nutricline. Physical variables such as temperature have been much better sampled in space and time (i.e. more than five decades) and can be considered for use as proxies. Reconstruction skill depends partly upon how well CHL covaries with the physical predictors used in the reconstruction. The reconstruction is performed in EOF-space to filter out higher frequency patterns. The skill of the reconstruction depends partly upon how much CHL variance is captured in the dominant modes.

3.1.1 Selection of physical predictors

Simple Ocean Data Assimilation (SODA) version 2.1.6 is a reanalysis of ocean climate that provides complete fields of physical variables from 1958 to 2008. This version uses the Geophysical Fluid Dynamics Laboratory (GFDL) modular Parallel Ocean Program (POP) model version 2.1, forced by surface wind stress from ERA40/QSCAT winds [Carton, 2013; Carton and Giese, 2008]. SODA is constrained by the constant assimilation of observed temperatures, salinities and altimetry using an optimal data assimilation technique. Observations come from the World Ocean Database 2009. Extending the SODA record beyond 2008 requires winds from the NOAA-CIRES 20th Century Reanalysis, which were still being examined at the time of this study [Grodsky, personal comm.]. SODA includes temperature profiles from which mixed layer depth (MLD) was defined as the depth of 0.2^{o}C difference from the temperature at 5m, salt, zonal and meridional currents (u, v), zonal and meridional wind stress (τ_x, τ_y),

and sea surface height (SSH). The 20°C isotherm (T_{20}) was used to characterize the depth of the thermocline. The SODA data were binned consistently with the CHL data.

Monthly 9km SeaWiFS (R2010.0) CHL data between September, 1997 through December, 2008 were averaged over 2° bins centered on the equator in the tropical Pacific (20°N-20°S). CHL was initially log-normally transformed, as that is standard pretreatment procedure in productive regimes. In the oligotrophic waters of the tropics and subtropics, however, the lognormal distribution has been found to be unsuitable [Sapiano et al., 2012]. For this study, despite the large upwelling region, log-normally transforming CHL gave too much weight to the edge of the west Pacific warm pool. Thus, linear calculations of CHL were used instead. The spatial and temporal resolutions were selected from the decorrelation scales of monthly CHL in the equatorial region (5-8° meridionally, 3.5 months). Zonal decorrelation scales decrease to about 200km in the lee of islands in the tropical Pacific [Matthews et al., 2011]. Gaps toward the end of the record were filled with MODIS Aqua CHL and any remaining missing data were filled using a 3x3 median filter, repeated three times. An expanded land mask excludes pixels within 2° of the coast. Annual cycles were removed. CHL were smoothed over three months using a 1-2-1 weighted moving average to retain important low frequency information while minimizing mesoscale features.

CHL was cross-correlated with physical variables to find the most suitable proxies to use in a CCA reconstruction. CHL in the tropical Pacific shows the strongest anti-correlation with SST (Figure 3.1, top) due to being nutrient-limited rather than light-limited and having the primary nutrient source from cooler, deep water. Inverse cross-correlation with MLD (Figure 3.1, middle) in the east Pacific corresponds to upwelling, a

shallow nutricline and high production. In the west Pacific warm pool, the mixed layer is decoupled from the thermocline and nutricline by the salinity barrier layer. Sea level has been found to be a robust indicator of the thermocline depth in this region [Turk et al., 2001] and, by extension, an indicator of nutricline depth and nutrient availability to the euphotic zone. The correlation between CHL and SSH (Figure 3.1, bottom) is strongly inverse along equatorial upwelling areas and in the warm pool as a thicker surface layer causes a deeper nutricline and negative CHL anomaly. The depth of the 20^{o}C isotherm is also considered a proxy for the thermocline depth, but its relationship with CHL was not as robust as SSH and is not shown here.

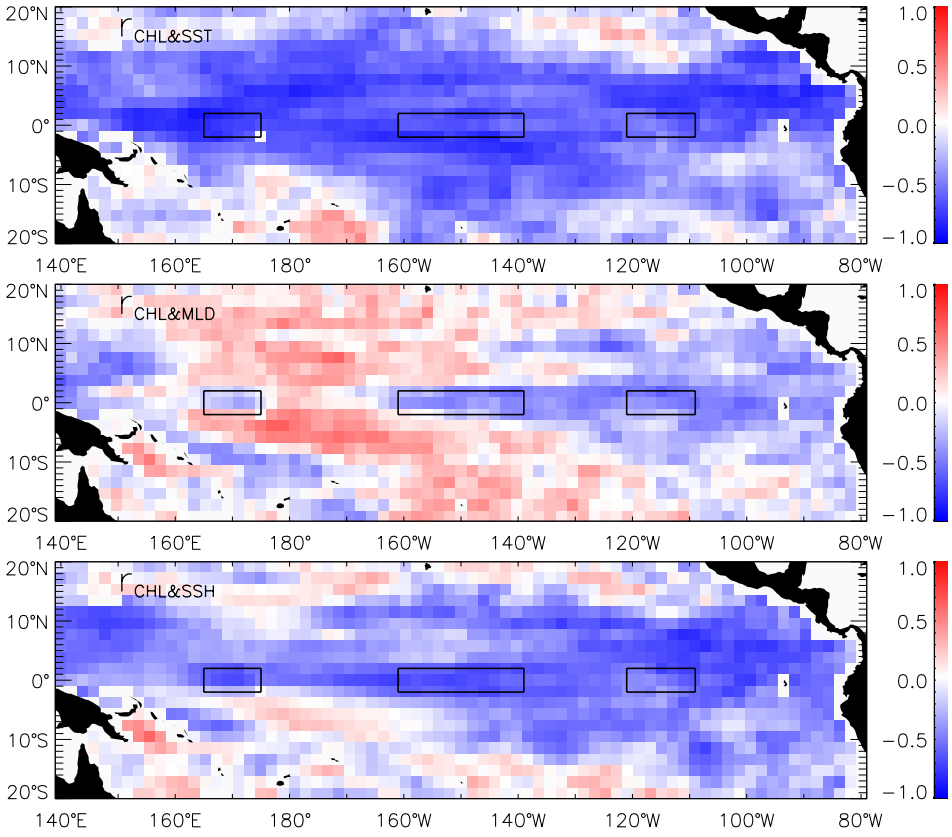

Figure 3.1 Cross-correlations for the period 9/1997-2008 between: CHL and SST (top), CHL and MLD (middle), CHL and SSH (bottom). Boxes delineate areas for Figure 3.2, 3.3, 3.4.

The student's t-test (4) [Fisher, 1921] was applied to test for significance:

$$t = r\sqrt{\frac{n-2}{1-r^2}} \qquad (4)$$

For a 136 month-long time series filtered over 3 months, n = 45 and r > 0.45 is highly statistically significant ($p < 0.001$). Dark blue and red indicate significantly correlated areas (Figure 3.1). Time series plots show the variation of each pair for the delineated boxes. In the oligotrophic west Pacific area (Figure 3.2), the strongest correlation is with CHL and SST: r = -0.80, next with SSH: r = -0.53, and insignificant with MLD: r = -0.04. The extreme 1998/99 La Niña caused twice the typical CHL while SST cooled and SSH decreased as stratification eroded when the warm pool propagated east. The moderate central Pacific El Niños in 2002/03, 2004/05, 2006/07 are evident in decreased CHL, increased SST and SSH. A moderate La Niña in 2007/08 had high CHL and cool SST.

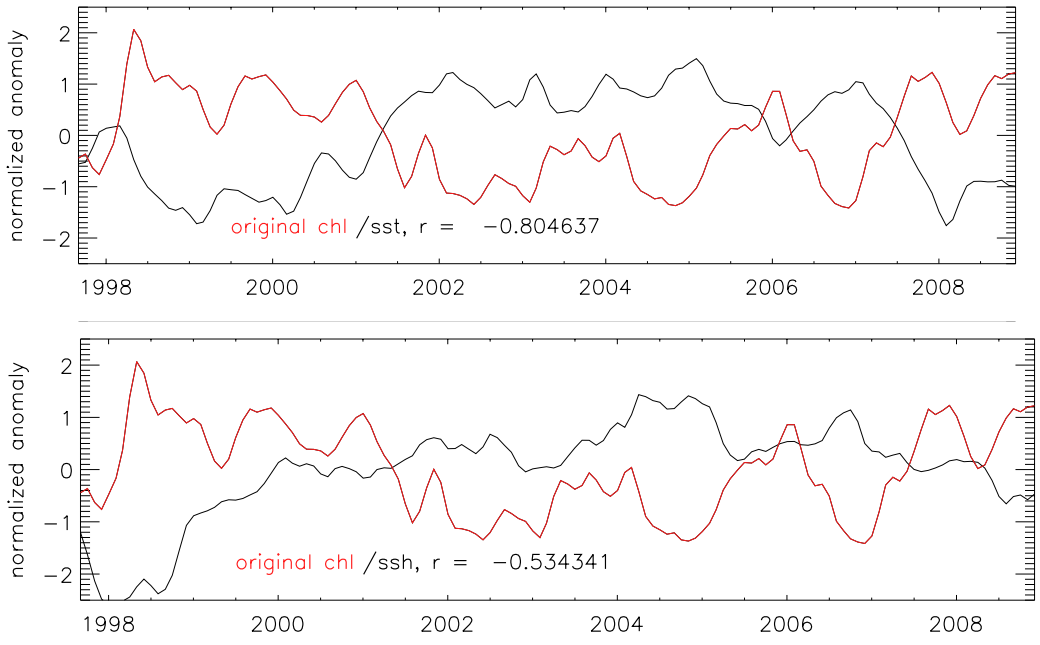

Figure 3.2 Western equatorial Pacific time series (165-175°E): CHL (red, both panels) with SST (top) and SSH (bottom). Average cross-correlations for each area noted.

In the central Pacific area (Figure 3.3) there is a strong correlation between CHL and SST, r = -0.75, while SSH has its strongest correlation with CHL here: r = -0.67. The extreme 1997/98 El Niño and 1998/99 La Niña events are more pronounced than in the west and east areas. Lesser central Pacific El Niños are evident in 2002/03 and 2006/07, while the 2004/05 El Niño barely registers.

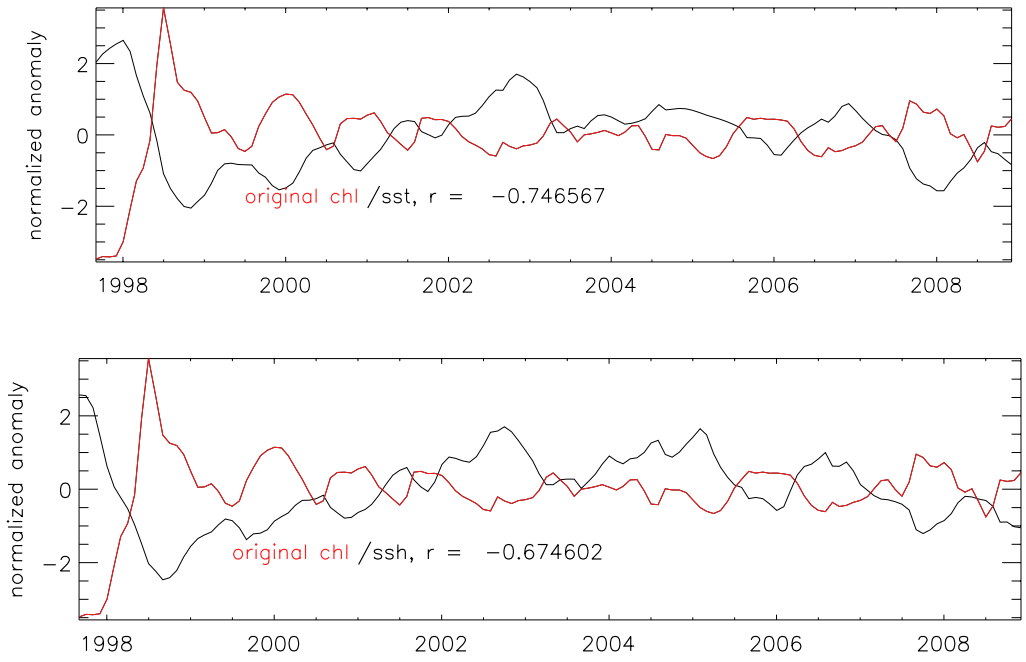

Figure 3.3 **Central equatorial Pacific (160-140°W): variables as in Figure 3.2.**

In the east Pacific, SST has the strongest correlation with CHL, r = -0.64, though less than for the two areas to the west (Figure 3.4, top). SSH and CHL are correlated slightly less, r = -0.58 (Figure 3.4, bottom). Here MLD and CHL are significantly correlated in the absence of the salinity barrier layer, r = -0.49, though not as strongly as CHL covaries with the other two variables. During El Niño, the basin-wide west-east thermocline slope relaxes, shoaling in the west and deepening in the east along with the nutricline resulting in and less productivity and less CHL. During La Niña events,

easterly winds and upwelling resume, the basin-wide thermocline slope is reestablished, the nutricline likewise shoals bringing more nutrients toward the surface leading to more CHL. The extreme 1997/98 El Niño and 1998/99 La Niña are most pronounced. Weaker El Niños in 2002/03 and 2006/07 register as low CHL anomalies. A positive CHL peak in 2003 during a weak La Niña barely registered as a negative Oceanic Niño Index.

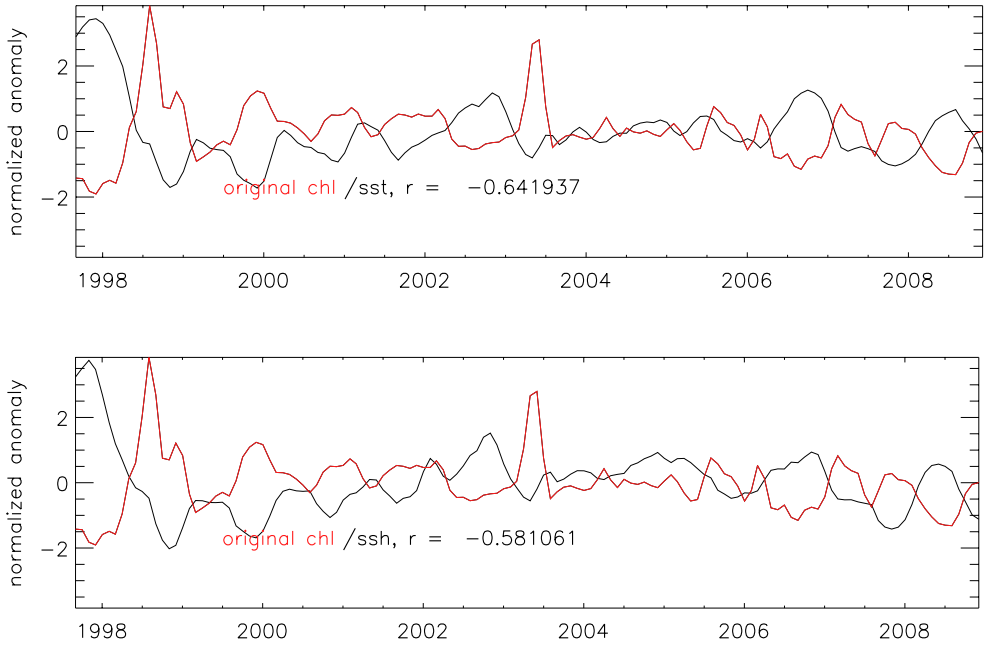

Figure 3.4 Eastern equatorial Pacific (120-110°W): variables as in Figure 3.2.

Wind-driven upwelling and vertical mixing can supply nutrients to the euphotic zone, thus CHL anomalies were cross-correlated with wind (Figure 3.5). Zonal wind stress is weakly correlated with CHL toward the east where remotely forced thermocline and nutricline depths have the greatest control upon CHL anomalies. West of 140°W the response to local forcing is greater and the relationship between zonal wind stress and CHL is inverse (i.e. easterly winds enhance upwelling and nutrients supplied from depth to increase CHL). Meridional wind stress correlations with CHL are strongest just north

of the equator where southerly (i.e. off equator) wind has a direct relationship with CHL through vertical Ekman pumping (Figure 3.5 middle). Wind stress magnitudes and CHL covary directly west of 180^{o} (Figure 3.5, bottom) meaning stronger winds correlate with more CHL there, primarily through turbulent vertical mixing [Ryan et al., 2002]. East of the dateline, the horseshoe pattern of inverse correlation and the weak positive correlation in the narrow area between 120-140^{o}W indicate blooms in the cold tongue are primarily controlled by non-local forcing, i.e. large-scale thermocline depth variation. Wind-driven vertical mixing and upwelling can secondarily contribute to blooms between 120-140^{o}W.

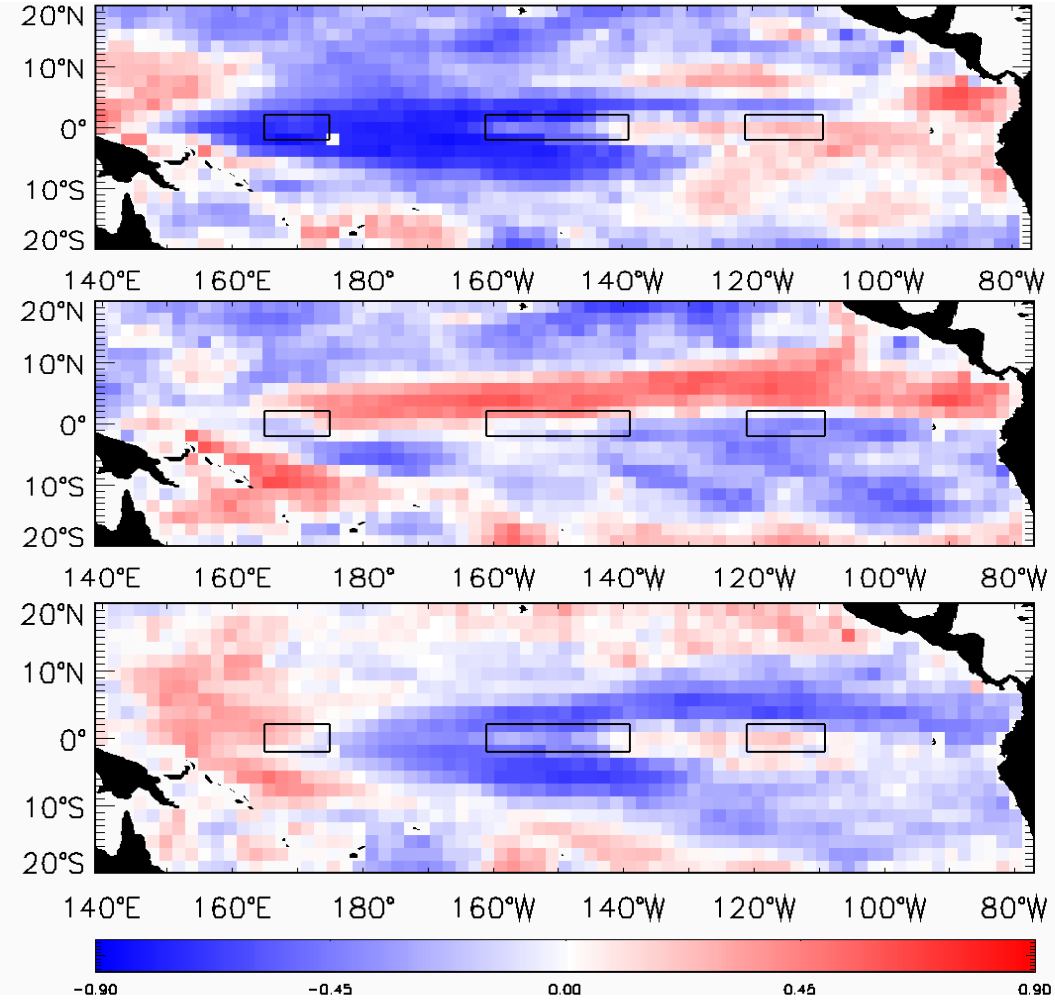

Figure 3.5 **Cross-correlations between normalized anomalies of CHL (9/1997-2008) and zonal wind stress, τ_x (top); meridional wind stress, τ_y (middle); magnitude of wind stress, $|\tau|$ (bottom).**

Figure 3.6 displays the zonal distribution of CHL, SST, SSH and MLD over time where the 1997/98 El Niño stands out in all variables. East of 150°W, SST is anomalously high in 1997/98 as are SSH and MLD, while CHL is depressed. By contrast, the El Niños in 2002/03, 2004/05, 2006/07 display smaller warm anomalies and less pronounced CHL anomalies. When conditions enter a La Niña phase, trade winds intensify and the east-west thermocline slope and upwelling are exaggerated. Equatorial upwelling expands westward causing positive CHL anomalies along the equator. During the training period, a La Niña phase began strongly in 1998/99, returned in 1999/2000, then weakened in 2000/01 and returned in 2007/08. These can be seen in Figure 3.6 as positive CHL anomalies. The 1998/99 La Niña corresponded with huge blooms that yielded a 40% increase in primary production over normal years [Chavez et al., 1999].

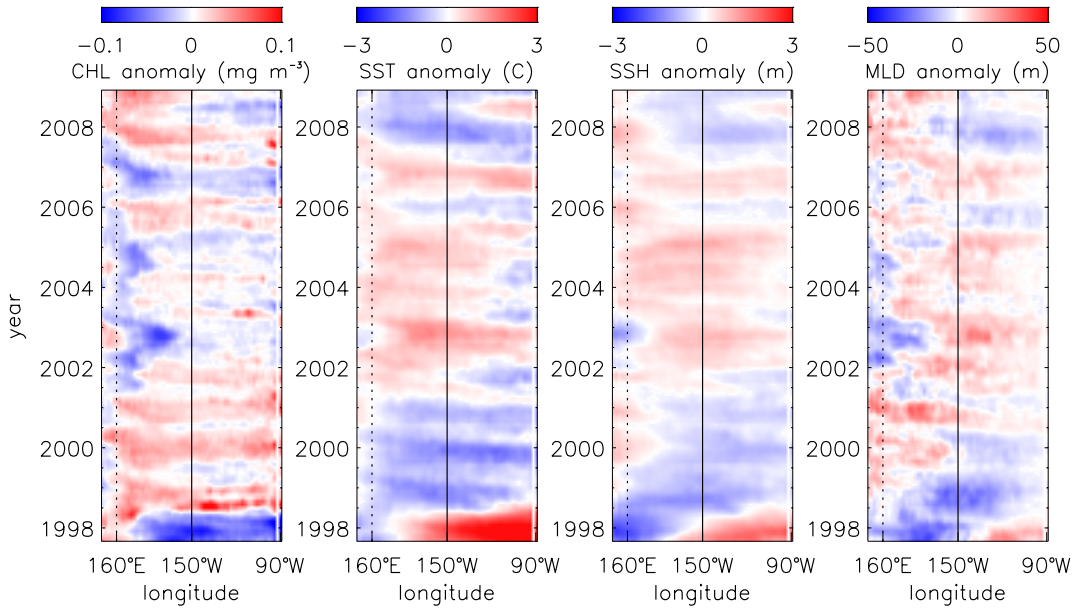

Figure 3.6 Longitude-time distribution of averages (5°S-5°N) of deseasoned CHL (left), SST (2nd from left), SSH (3rd from left) and MLD (right). Niño4 is between the dotted and solid lines; Niño 3 is eastward of the solid line. Note the 1997/98 east Pacific extreme El Niño predominately in Niño3.

There is ongoing debate about whether El Niños are changing from predominately east Pacific types to more frequent central Pacific types [Lee and McPhaden, 2010]. The physical forcing mechanisms of the two types of El Niño are not completely understood but have slightly different manifestations. Both experience eastward advection of the warm pool, with less eastward extent for the central Pacific El Niños. Only the east Pacific El Niño experiences the basin-wide flattening of the west-east thermocline slope that causes the cessation of upwelling in the east [Radenac et al., 2012]. Fewer east Pacific El Niños would have implications for biological productivity because central Pacific El Niños are more geographically confined and do not experience the basin-wide thermocline shift [Turk et al., 2011; Radenac et al., 2012]. During a central Pacific El Niño (e.g. 2002/03, 2004/05, 2006/07), the eastward advection of the warm pool causes a deeper thermocline in the central Pacific and a decrease in biological productivity near the dateline, while the east Pacific continues to be productive. During an east Pacific El Niño (e.g. 1997/98), the warm pool extends all the way east depressing the thermocline there and causing the collapse of upwelling off of South America, resulting in the loss of vertical nutrient input to the equatorial Pacific. Several studies have argued that shifting spatial patterns in the extent of El Niño events could contribute to a restructuring of marine ecosystems [Boyce et al., 2010; Turk et al., 2011], yet whether the shift toward more central Pacific El Niño events is a long-term trend or part of a lower frequency cycle is still an open question [McPhaden et al., 2011].

Because of the consistently strong inverse correlation between CHL and SST caused by stratification in the western warm pool and upwelling in the cold tongue, SST is an obvious choice for a predictor. Additional predictors were compared with the goal

of finding one that best complements SST as an indicator of nutrient entrainment that could fuel phytoplankton blooms. Another factor to consider is how physical forcing changes over time. For example: whether a cool anomaly is getting cooler or warmer could indicate continued deep water entrainment or the onset of stratification which could serve to cut-off the supply of new nutrients or allow nutrients already entrained to remain in the sunlit euphotic zone. The time derivative of SST was calculated by taking the three point difference between the preceding month, current month, and following month. For the four Niño regions, the relationship between CHL and d(SST)/dt was always positive but never significant, with the highest correlation coefficient of 0.25 for Niño 3 region. Time derivatives of other physical variables showed less correlation to CHL. The skill of each proxy was defined by the fraction of area significantly correlated to CHL (Table 3.1), with 1 being perfect skill and 0 being no skill. After testing many variables as the second predictor, the best choice became SSH due to its consistency across most of the tropical Pacific (evident in Figure 3.1, bottom). Highest skill means that changes in SST and SSH are more representative of processes leading to blooms than other variables. Although these two reflect similar dynamics, they are somewhat independent as SSH represents thermocline depth while SST represents how quickly cooler water upwelled from depth gains heat as it is transported westward and away from the equator at the surface. Their average cross-correlation with each other is 0.41, with greater values toward the equator meaning that they are slightly less independent of one another there.

Table 3.1 Skill (0-1) or fraction of total area significantly correlated to CHL in the tropical Pacific

variable	SST	SSH	MLD	T_{20}	$\frac{d(SST)}{dt}$	τ_x	τ_y	$\lvert\tau\rvert$	τ^2
skill	0.51	0.31	0.14	0.16	0	0.09	0.01	0.02	0.05

3.1.2 Selection of statistical method

Several statistical reconstruction methods exist. These generally involve multiple regression of correlations between spatial fields of the most predictively important variables in a predictor data set to explain the variance in a predictand data set [Barnett and Preisendorfer, 1987]. This is in contrast to ordinary multiple regression, which is a point-to-point regression. The training data consisting of predictors and predictand are first orthogonalized, or converted into empirical orthogonal function (EOF) spectral space. The dominant modes are used to reduce the number of degrees of freedom and capture major variations while filtering out much of the noise [Davis, 1976; Barnett and Preisendorfer, 1987; Smith et al., 2009]. An additional advantage of filtering a time series by projecting it onto a subset of the EOFs is that it can make the statistical reconstruction less susceptible to sampling fluctuations [Barnett and Preisendorfer, 1987].

According to Barnett and Preisendorfer [1987], canonical correlation analysis (CCA) is at the top of the hierarchy of regression modeling approaches. Bretherton et al. [1992] compared different reconstruction methods to determine the advantages of each toward detecting coupled patterns in climate data and found that CCA in EOF spectral space is superior to CCA of non-orthogonalized data. In their study, the multi-variate EOF (MEOF) method extracted a coupled pattern more accurately than the CCA method, but MEOF exhibited a bias toward the leading EOF and a potential to identify a coupled signal that is actually orthogonal to the true coupled signal. Bretherton et al. [1992] concluded that the CCA applied in EOF spectral space is superior for longer time series or when the coupled signal does not resemble the EOFs of the individual fields. Both MEOF and CCA are tested here to determine which performs best for this reconstruction.

The percentage of the combined CHL/SST/SSH variance explained by each of the first 15 modes is apportioned as follow: 29.9, 15.0, 8.0, 4.4, 3.0, 2.7, 2.6, 2.3, 2.1, 1.9, 1.7, 1.6, 1.5, 1.4, 1.4 for a total of 80% of the variance. After mode five, there is little coherence between the spatial functions for the three variables and no basin scale patterns. After mode seven, the values in the spatial functions approach zero. The spatial and temporal distributions of the first four modes indicate the dominant basin-scale patterns (Figures 3.7, 3.8, 3.9, 3.10). Normalized fields of predictors and predictands are used in the MEOF analysis to find combined relationships. The predictors are used to estimate the time-series weights for the MEOF and to weight the predictand component [Jolliffe, 2002; Weare and Nasstrom, 1982].

Modes do not necessarily correspond to physical processes, however, some combination of features can be inferred by the spatial and temporal distribution of their patterns. The first mode (Figure 3.7) looks like the main ENSO pattern. The second mode (Figure 3.8) appears to be a low frequency variation to the ENSO pattern, perhaps part of the NPGO. The third and fourth modes (Figure 3.9 and Figure 3.10) contain higher frequency variation to the dominant modes, likely due to wind variability but are less and less coherent and increasingly impacted by mesoscale features.

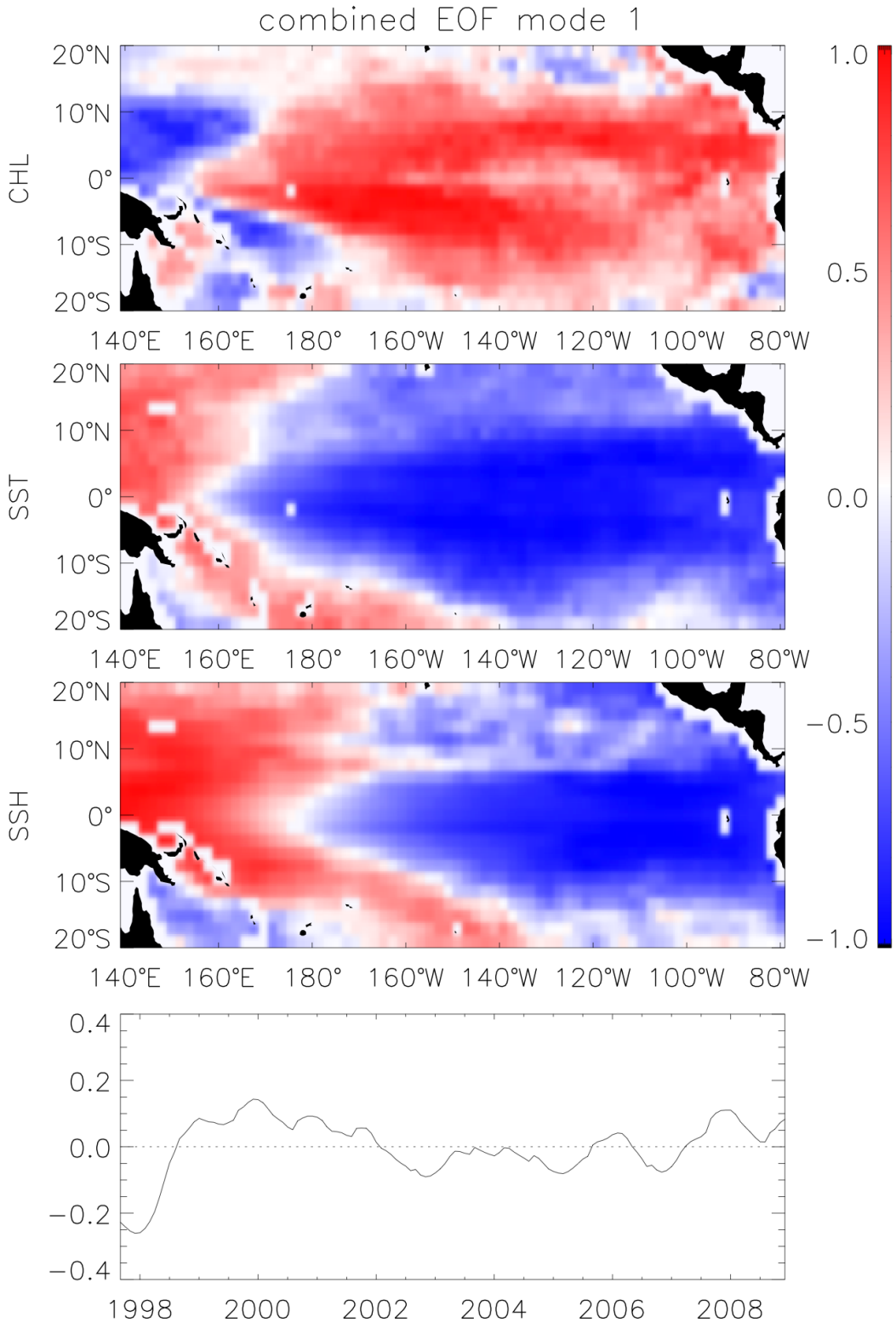

Figure 3.7 **CHL/SST/SSH combined EOF mode 1 explains 30% of the total variance: CHL (top map), SST (middle map), SSH (bottom map) and the time-varying amplitude (bottom).**

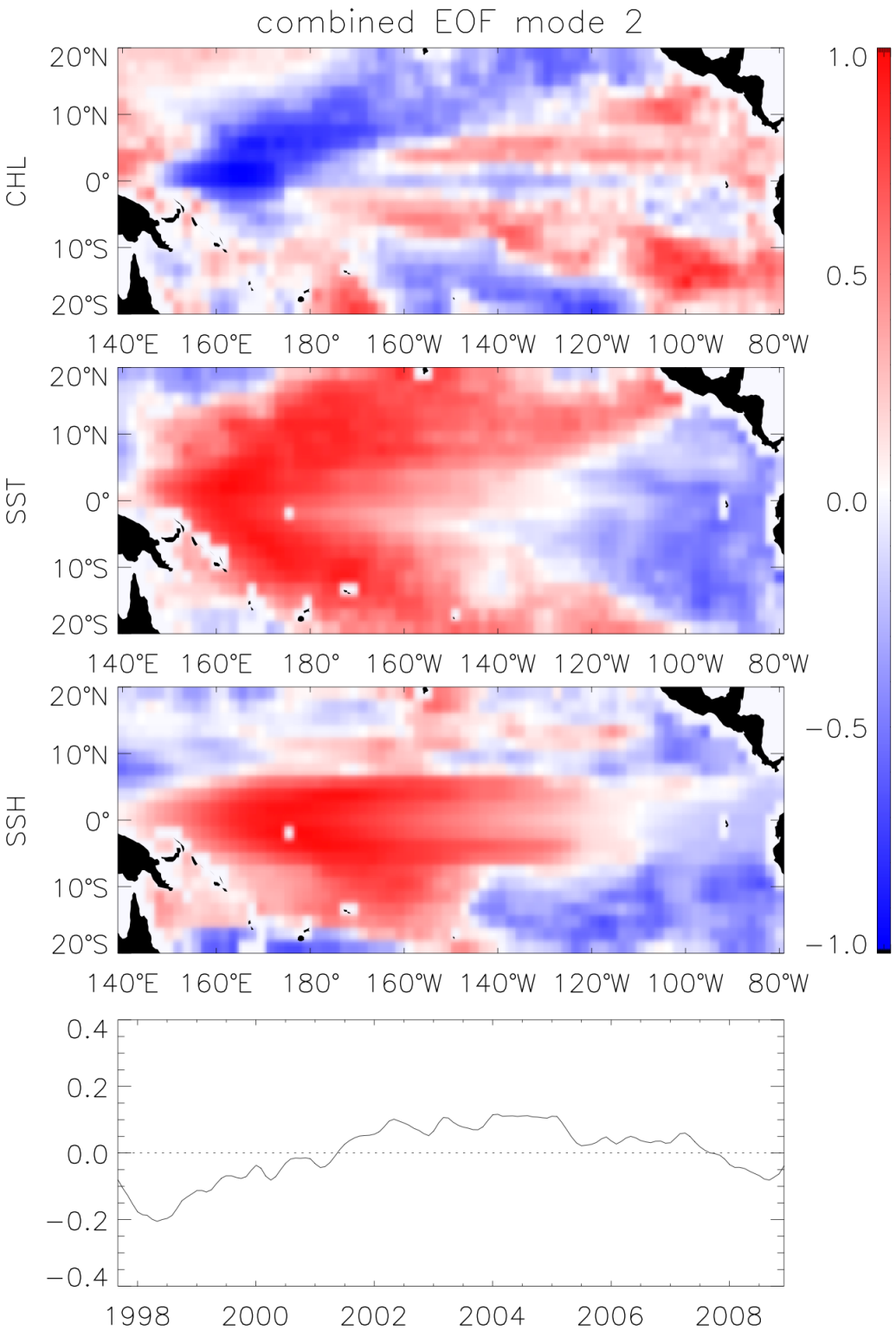

Figure 3.8 **CHL/SST/SSH combined EOF mode 2 explains 15% of the total variance: CHL (top map), SST (middle map), SSH (bottom map) and the time-varying amplitude (bottom).**

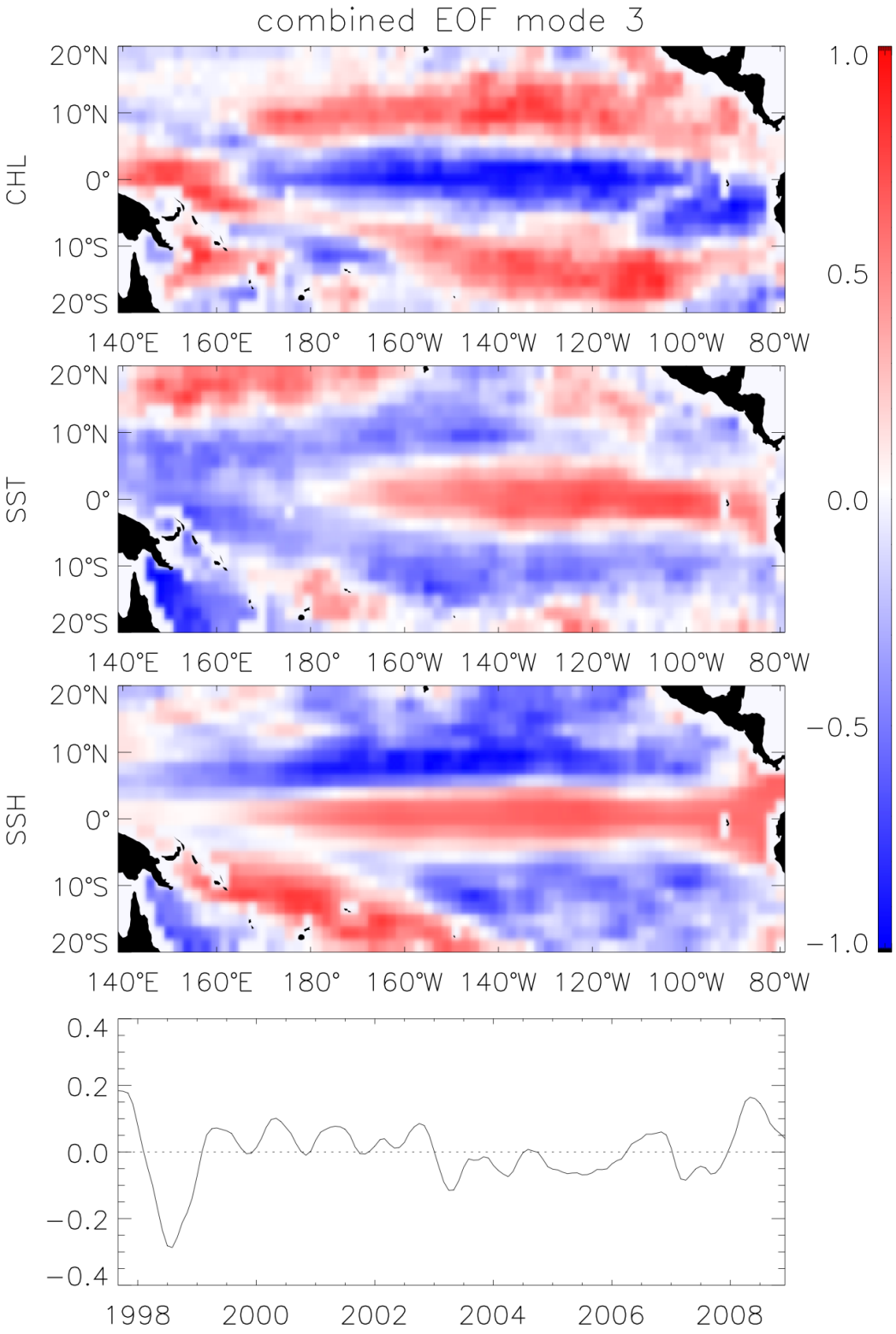

Figure 3.9 **CHL/SST/SSH combined EOF mode 3 explains 8% of the total variance: CHL (top map), SST (middle map), SSH (bottom map) and the time-varying amplitude (bottom).**

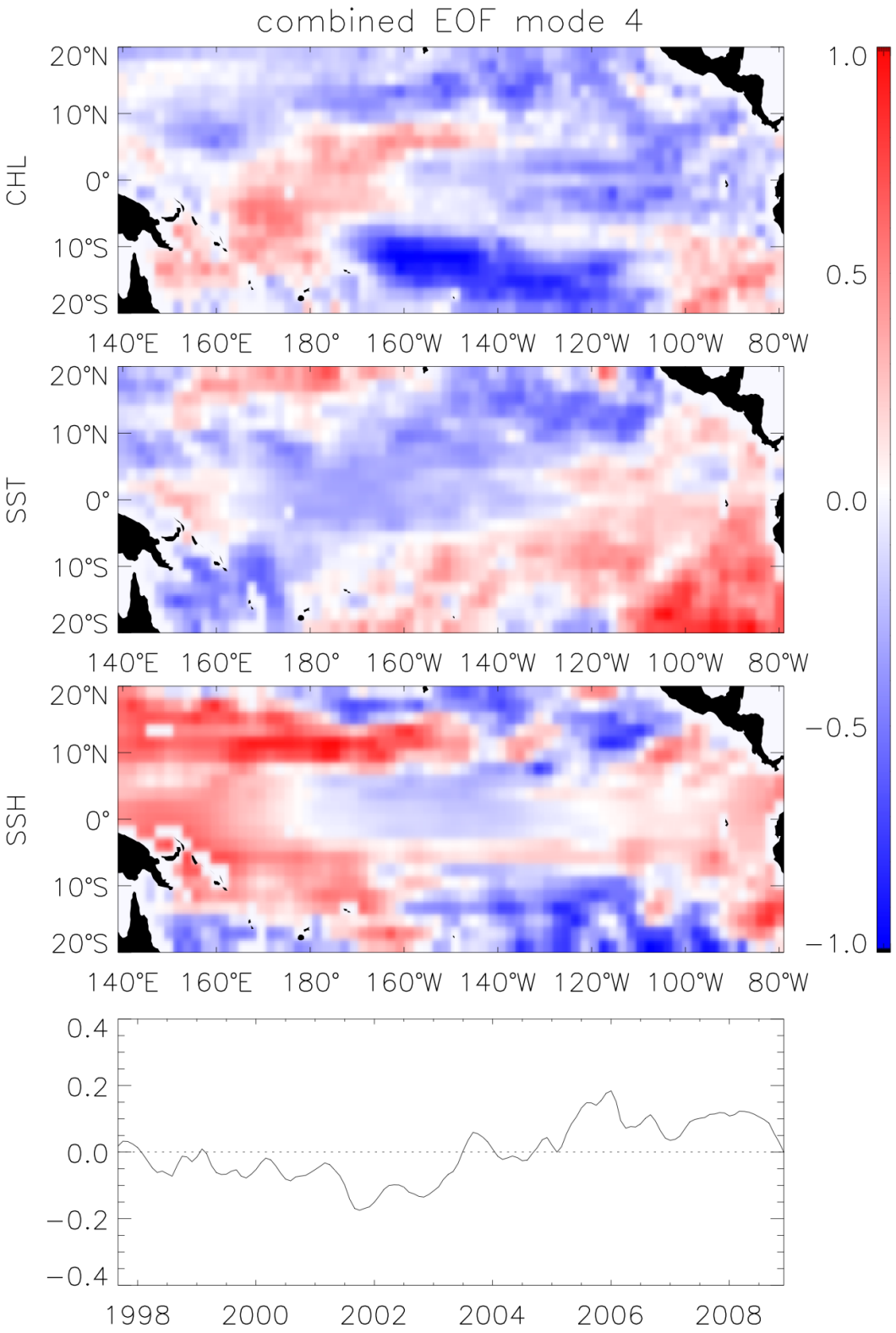

Figure 3.10 **CHL/SST/SSH combined EOF mode 4 explains 4.4% of the total variance: CHL (top map), SST (middle map), SSH (bottom map) and the time-varying amplitude (bottom).**

The CCA uses normalized fields to define a coordinate system that optimizes the correlation between the predictor and the predictand [Barnett and Preisendorfer, 1987]. In EOF spectral space, the correlation in the leading modes is partitioned between the predictor (SST and SSH) and predictand (CHL) data sets to find the space and time evolution of SST and SSH that best predicts CHL variability (see Appendix).

The number of modes applied to a statistical reconstruction is determined by how much variance is contained in the modes. To avoid artificial predictability, the number of parameters used as statistical estimators is limited [Davis, 1976]. Prior to taking EOFs for both types of reconstruction, the deseasoned anomalies were normalized by their standard deviations for nondimensional data with signals of comparable magnitude. All bins were weighted by cos(lat) to account for distortion away from the equator in cylindrical projection. For both the EOF on CHL alone and the combination of CHL, SST and SSH, 15 modes were significantly different than noise by Rule N and the bootstrapped eigenvalue tests [Overland and Priesendorfer, 1982; Jackson, 1993]. To test the best combination of modes to use in this study, up to 15 modes were compared by their root-mean-square errors (RMSE) and cross-correlations between the original SeaWiFS CHL and the cross-validated CHL reconstruction. The cross-validation method is described in section 3.2.1. As shown in Figure 3.11, using ten modes yielded the lowest RMSE with highest correlations over the Nino areas and the entire tropical Pacific (e.g. RMSE=0.26, r=0.93 in Niño 4).

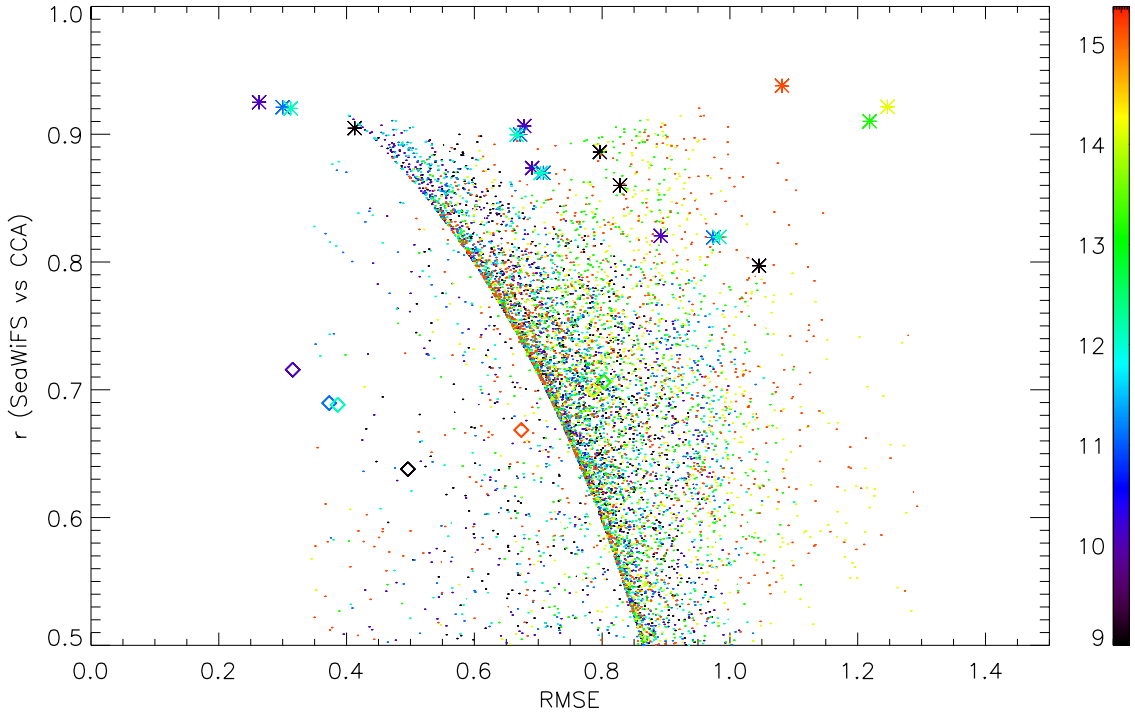

Figure 3.11 RMSE and r for SeaWiFS and cross-validated CHL reconstructed with 9-15 modes (color) for all bins (.), Nino area averages (*), and tropical Pacific averages (◊).

Ten modes account for 71% of the variance in the EOF of CHL alone and 72% of the variance for combined SST and SSH. The percentage of the total variance contained in each of the first ten modes is 30, 15, 8, 4.4, 3, 2.7, 2.6, 2.3, 2.1, 1.9, 1.7, 1.6, 1.5, 1.4, 1.4. The spatial and temporal distributions of the first four CCA modes are shown (Figures 3.13, 3.14, 3.15, 3.16). Mode 1 (Figure 3.12) primarily reflects basin-wide ENSO patterns with El Niños in 1997/98, 2002/03, weakly in 2004/05, 2006/07 and La Niñas in 1998/99 – 2000/01, 2005/06, 2007/08. The second mode (Figure 3.13) yields a pattern similar to the PDO [Mantua and Hare, 2002]. The third mode (Figure 3.14) and fourth mode (Figure 3.15) highlight smaller, higher-frequency fluctuations, either from atmospheric effects or adjustment to the two main modes. Patterns past mode four are increasingly noisy and show decreased spatial coherence, so they are not displayed here.

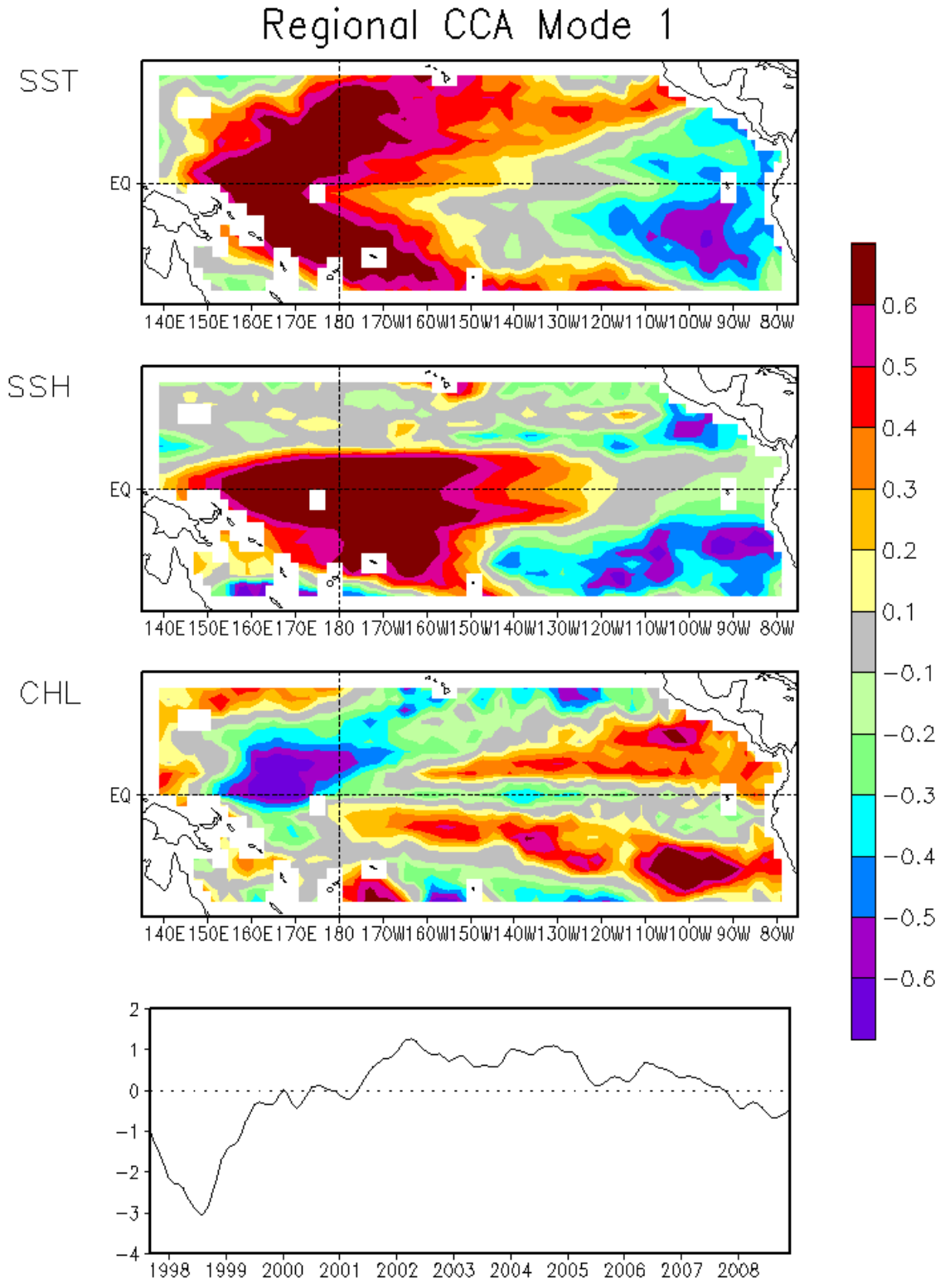

Figure 3.12 **CCA mode 1 explains about 13% of the total variance with spatial functions: SST (top map), SSH (middle map), CHL (bottom map) and the time-varying amplitude (bottom).**

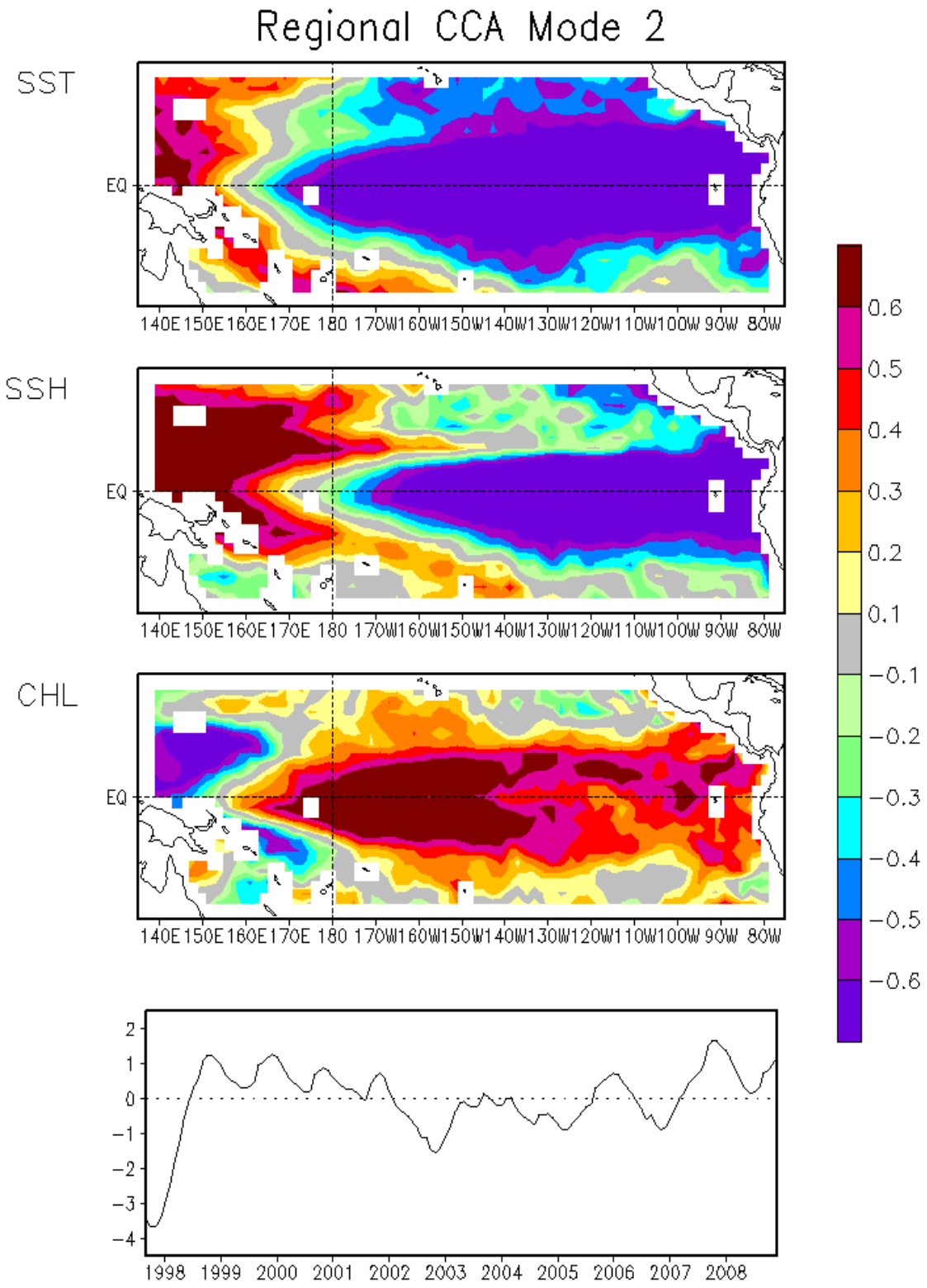

Figure 3.13 **CCA mode 2 explains about 12% of the total variance with spatial functions: SST (top map), SSH (middle map), CHL (bottom map) and the time-varying amplitude (bottom).**

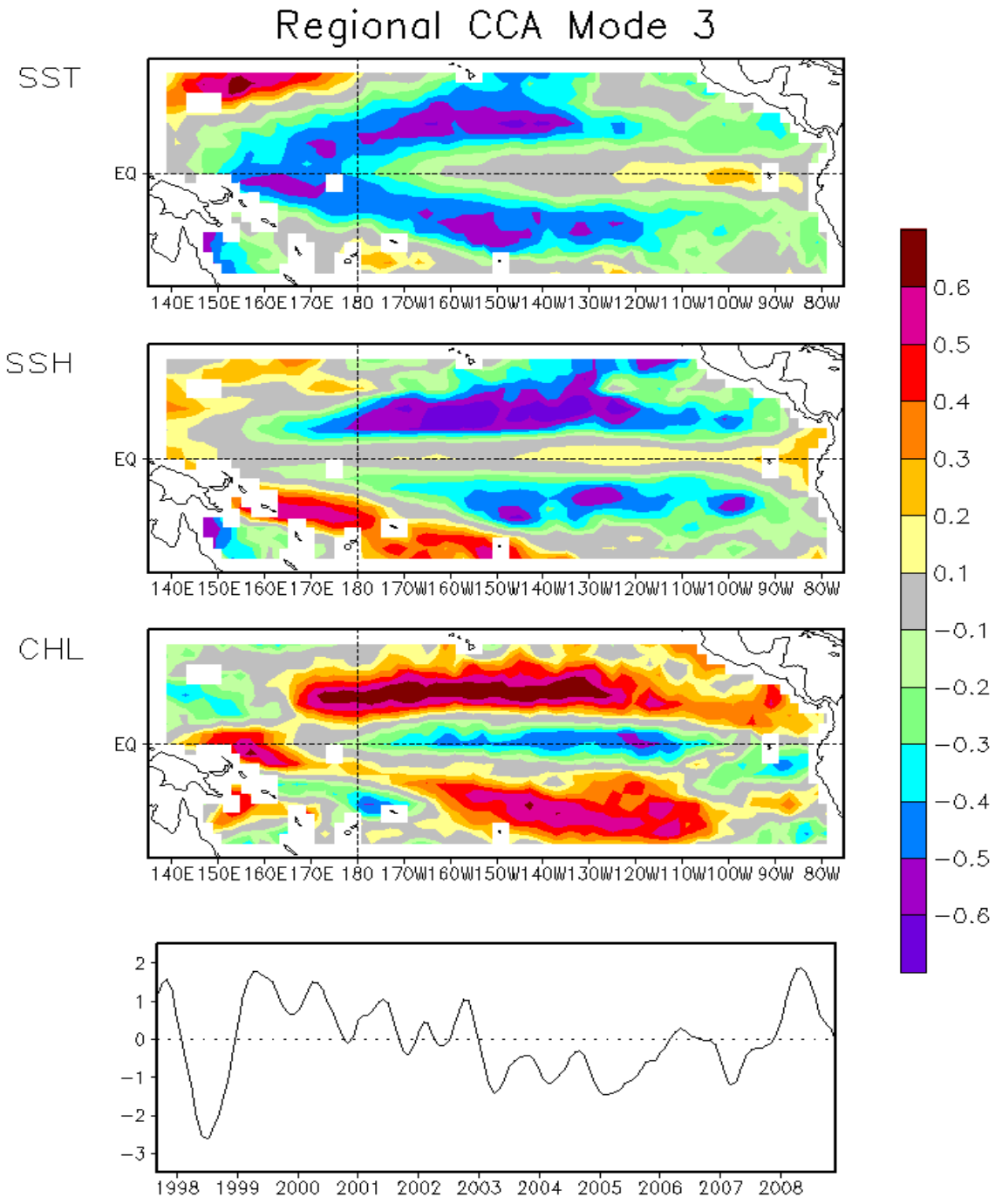

Figure 3.14 **CCA mode 3 explains about 12% of the total variance with spatial functions: SST (top map), SSH (middle map), CHL (bottom map) and the time-varying amplitude (bottom).**

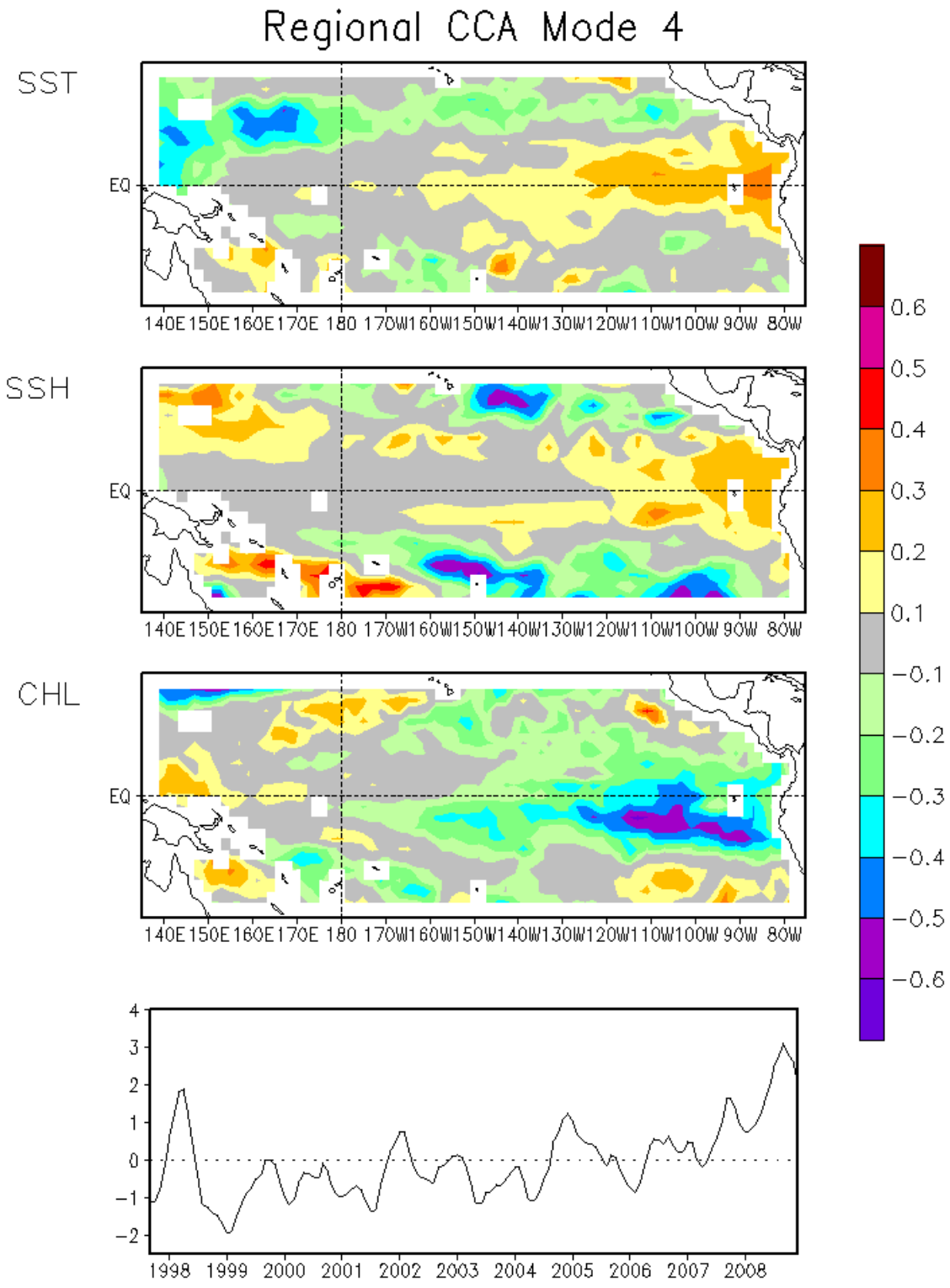

Figure 3.15 **CCA mode 4 explains about 10% of the total variance with spatial functions: SST (top map), SSH (middle map), CHL (bottom map) and the time-varying amplitude (bottom).**

An *a priori* estimate can be made about the potential of a reconstruction that uses the dominant EOF modes to reproduce the signal in the original data. The variance of the first ten EOF modes is divided by the variance of the original CHL data to yield the Fraction of Variance Resolved, FVR (Figure 3.16).

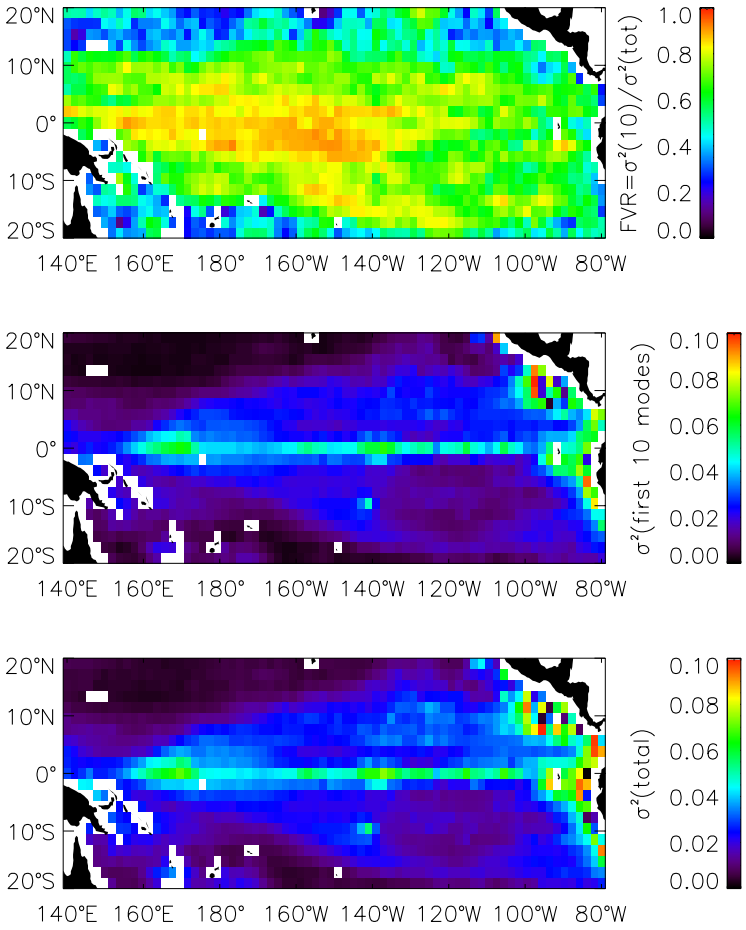

Figure 3.16 Fraction of Variance Resolved (FVR) by the dominant EOF modes (top) calculated by dividing the variance of the dominant ten EOF modes (middle) by total variance in CHL (bottom).

A reconstruction using ten modes is thus expected to capture the majority of the CHL variance within 10^{o} of the equator and more than 80% of the total variance in the equatorial area west of 120^{o}W. The first ten modes do not adequately capture the higher

variance near coastal margins and the low variance in the oligotrophic subtropical gyres. The EOF method selects time functions that explain as much of the variability as possible over the whole domain, emphasizing modes of variability that are in phase over large areas. Propagating localized features (e.g. eddies, waves and coastal features) can be well-resolved in the ocean color imagery and appear as high SNR (by the Ballabrera-Poy et al. method discussed in section 2.4.3) but break up into multiple higher EOF modes that do not get included in the reconstruction. An EOF in real number space does not account for phase changes as a function of location as a complex EOF would. Propagating features such as planetary waves, captured in monthly composites, are resolved by a Gaussian fit to the temporally lagged autocorrelation function due to their smoothly propagating nature, contributing to a high SNR (Figure 2.17) but are not included in the dominant EOF modes because of being localized. The spatial distribution of the FVR indicates that using ten EOF modes in the reconstruction will capture basin-scale features impacting CHL variability, but localized features and those close to the coasts or having very low variability will not be resolved.

3.1.3 Chlorophyll reconstruction

After both CCA and MEOF reconstructions were calculated using ten modes with the same predictors, SST and SSH, an inspection of individual months showed that the methods are consistent. Both reconstruct major spatial patterns, with slight differences in smoothing and smaller scale patterns, evident in any given month and shown here for December, 1997 (Figure 3.17).

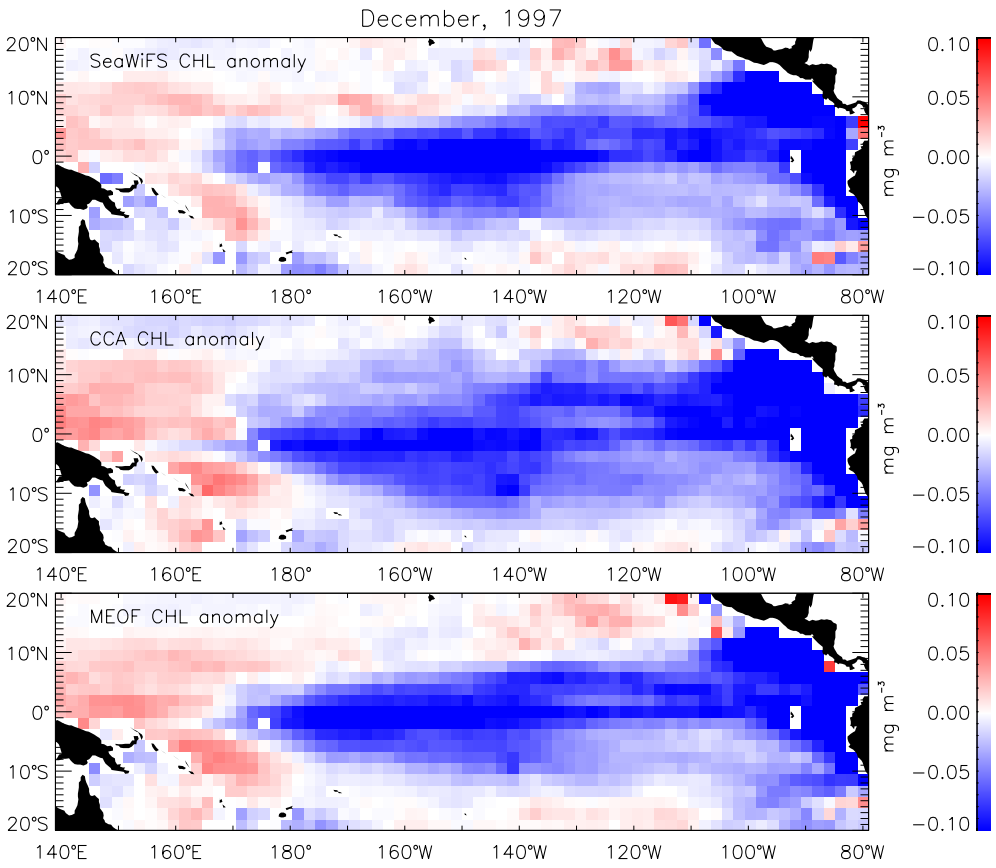

Figure 3.17 Deseasoned CHL for December, 1997: SeaWiFS (top), CCA (middle), MEOF(bottom).

Understanding where the reconstructions perform best is important for qualifying their strengths and limitations. Comparing the CCA and MEOF reconstructions against the original SeaWiFS CHL (Figure 3.18, top panels) shows that the MEOF is comparable and often out-performs the CCA reconstruction over the Niño areas during the dependent period. As expected, locations of high FVR match-up best between reconstructed and original CHL. Both reconstructions correlate highest with the original along the equator: Niño 4 (r_{MEOF}=0.97; r_{CCA}=0.93), Niño 3.4 (r_{MEOF}=0.94; r_{CCA}=0.91), Niño 3 (r_{MEOF}=0.90; r_{CCA} = 0.87). The match-ups are still significant ($r > 0.45$, $p < .001$) but lower for coastal Niño 1&2 (r_{MEOF}=0.88; r_{CCA}=0.82), where FVR is lower and variance is higher.

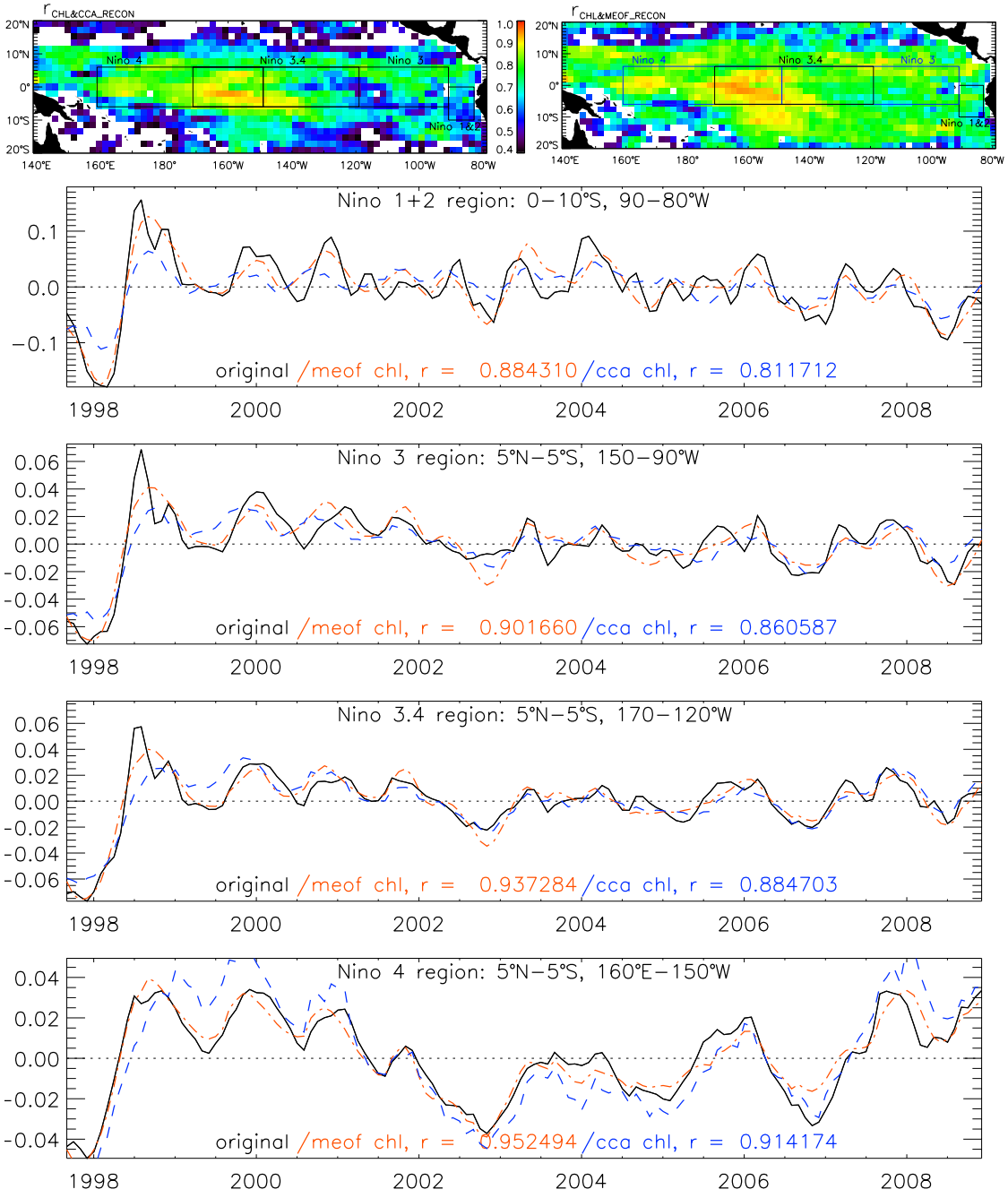

Figure 3.18 **Correlations between the original and reconstructed CHL from 9/1997-12/2008: for CCA (top, left map) and MEOF (top, right map). Colored bins have significant correlations. Boxes show Niño areas. Lower four panels are time series averaged over the Niño areas for original (black), MEOF (red), CCA (blue) with average correlations to original CHL annotated for the latter two.**

The skill of both methods is associated with how well the ENSO events were captured by the reconstructions. In all four Niño regions, the reconstructions tracked the original CHL anomalies fairly well. The main differences between the original and reconstructed CHL fields are in the timing and magnitude of the ENSO events. It is expected that some of the magnitude would be damped in the reconstructions because they only used the first ten modes. Damping is notable during the 1998/99 La Niña in the central and eastern Pacific (Niño 1&2, 3, 3.4). The reconstructions were slow to register the CHL peak during the 2005/06 La Niña in all four areas, though the transition out of that La Niña was accurately reconstructed in timing and magnitude. In Niño 4, the CCA reconstruction was slow to register the 1998/99 and 2007/08 La Niña peaks and exaggerated their magnitudes. The Niño 1&2 coastal area includes more variability than the dominant EOF modes capture, as demonstrated by the lower SNR there. Nevertheless, the correlations between the original and reconstructed CHL are significant over all four Niño areas and 86% of the entire tropical Pacific indicating that the statistical reconstructions successfully reproduced the majority of the variability in the CHL time series. That reconstructed CHL matches the original so closely confirms that large-scale biological variability can be reproduced by physical proxies in the tropical Pacific where the physical signals are large. Before the reconstructed data are interpreted for their scientific content, it is necessary to validate them and quantify their uncertainties to understand where and when they successfully represent CHL anomalies and where and when they do not.

3.2 Validation and quantification of uncertainties

To determine skill and where the method is successful, CHL reconstructions were cross-validated against the training data by excluding the SeaWiFS CHL for each month plus and minus some number of months during the model training and then comparing the reconstruction of the central excluded month to the original SeaWiFS CHL over all 136 months of training data to determine how well they were reproduced. Several combinations of months were tested for exclusion to determine the best number to maintain a robust reconstruction while minimizing the chance that the central excluded month would have any input to the reconstruction. Comparing the CHL reconstruction to independent records during the pre-SeaWiFS period, the reconstructed CHL were validated against any and all *in situ* sea surface CHL in the neighborhood, estimates derived by an independent ocean color sensor, CZCS, which operated between 1978-1986 as a proof-of-concept ocean color mission, and to output from a fully-coupled physical-biogeochemical ocean model. Finally, uncertainties in the observations and reconstruction model are quantified and discussed.

3.2.1 Cross-validation against training data

Cross-validation is a method commonly used to assess how a statistical model performs by holding out data from the training period, reconstructing the target variable and then comparing the original data left out to its reconstruction [Barnett and Preisendorfer, 1987; Smith et al., 1996; Mann et al., 1998; Smerdon et al., 2011]. Leaving out an increasing number of months was tested. By seven months, the relatively small length of the training record caused the reconstruction to diverge

markedly from the original data, thus five months was determined to give the best balance between enough months retained to train a robust reconstruction and enough months excluded to ensure the central month had minimal input to the reconstruction. The cross-validation reconstruction was performed for each month while holding out that month +/-2 months and then the central excluded month was compared to the original training data and repeated for all 136 months of the training data set. As shown in Figure 3.19 and Figure 3.20, the cross-validation of the CCA and MEOF reconstructions most closely fit the left-out training data within 5-10^{o} of the equator.

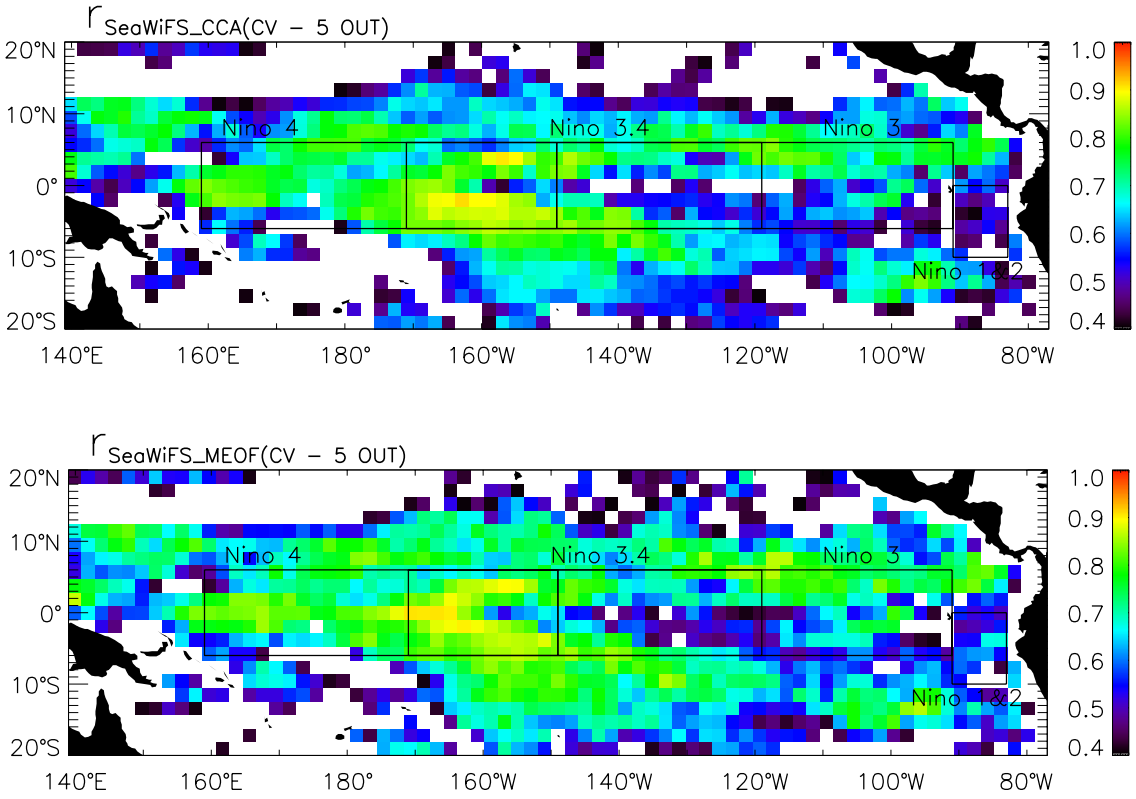

Figure 3.19 Leave five out cross-validations using SeaWiFS training data and reconstructing CHL by CCA (top) and MEOF (bottom) during the dependent period (9/1997-2008). Colored bins show significant correlations. Boxes show Niño areas that correspond to the average time series in Figure 3.20.

Cross-validations show highest fidelity by both models in the equatorial areas (Niño 3, 3.4, and 4) with r~0.8-0.9, while Niño 1&2 has r ~ 0.7 for the MEOF method and r~0.6 for the CCA. Both models fare worst during the extreme ENSO 1997-1999, slightly out of phase and damping the amplitude of the La Niña toward the east Pacific. Otherwise, cross-validations indicate the methods capture the major patterns in timing and magnitude. Where both MEOF and CCA cross-validate well gives confidence in using statistical reconstructions there.

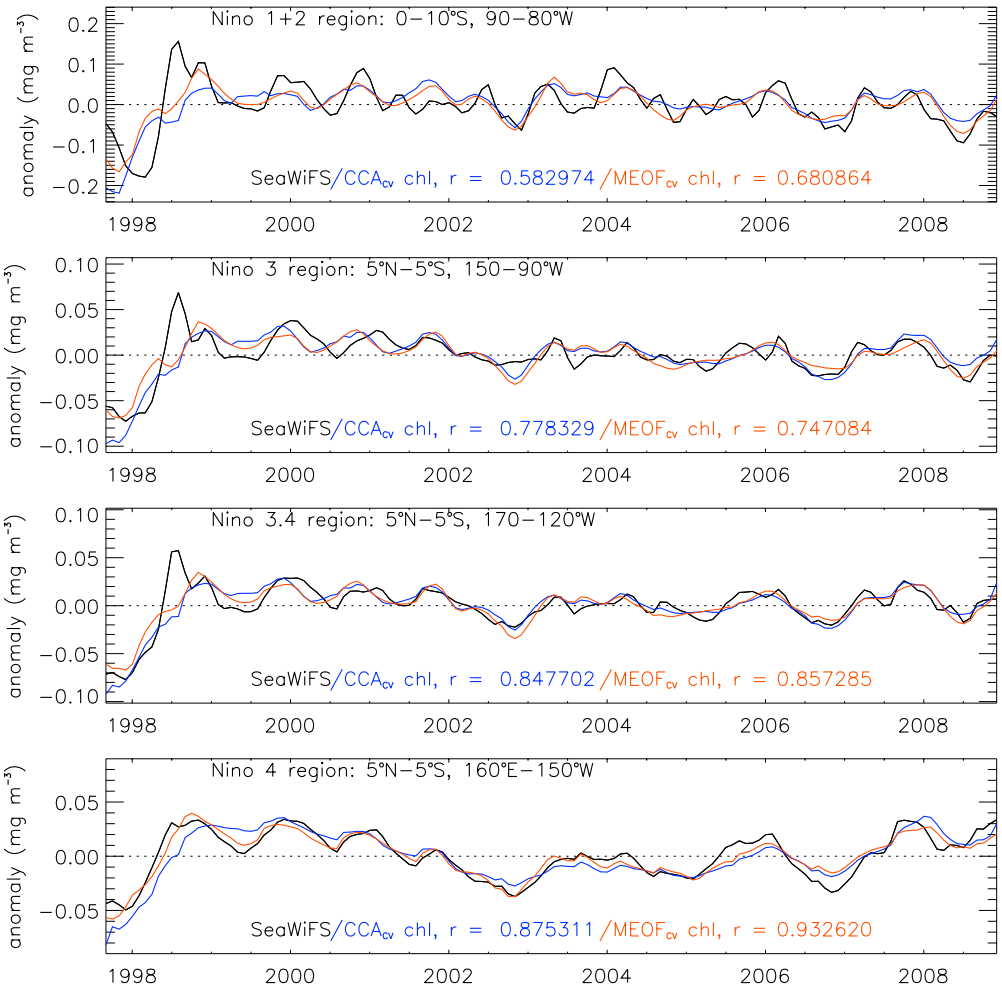

Figure 3.20 Comparisons of SeaWiFS CHL (black) to leave-five-out-cross-validations of CHL reconstructed by CCA (blue) and MEOF (red) and averaged over four Niño regions. Average correlations with the original SeaWiFS CHL are annotated for each area.

3.2.2 Comparison to independent data sets

Independent data are important for validation, as a close match-up between a reconstruction and training data could be caused by overfitting rather than true skill by the method. The reconstructed CHL were compared to available *in situ* data, *in situ* data mapped to EOF spectral space, CZCS ocean color estimates and output from a fully-coupled biogeochemical model.

In situ data comparison

As detailed in Chapter 2, *in situ* data collection of biological parameters lags data collection of physical parameters such as temperature. National Oceanographic Data Center (NODC) climate data records include about 10,000 casts containing CHL in the tropical Pacific between 1958 and 2008 (recall Figure 2.1). Limiting surface observations to depths less than 20m yields 27,540 data points. Averaging these observations over the four Niño areas typically yielded less than 100 months of match-ups with the reconstruction (~16%) and weak, insignificant correlations ($r \sim 0.1$ for three areas and $r = 0.2$ for Niño 4). By comparison, the *in situ* observations and training CHL also show weak, insignificant correspondence ($r < 0.1$ for three areas and $r = -0.2$ for Niño 4). It has been shown that EOFs can be used to regress *in situ* data onto spatial patterns and improve their skill compared to traditional *in situ* only analysis [Smith et al., 1996; Smith and Livezey, 1998]. When there were at least three observations in a bin, the sparse data were interpolated to a regular 2^{o} grid using the first five EOF modes which explain 66% of the total variance of the *in situ* data set. Original and reconstructed *in situ* observations were compared to the CHL reconstruction, as shown in Figure 3.21. After interpolating

the sparse data using the EOF method, there is still too much scatter to yield statistically significant match-ups with the reconstructions. Only a rough correspondence is evident between the *in situ* and reconstructed CHL (r~0.3 in Niño 3 and 3.4). The absence of *in situ* data for each bin in the area contributes to the discrepancy. Also, the *in situ* data are point source measurements collected over discreet days, usually with too few of them in a month to truly represent the monthly average. The training data used in the reconstruction are large area averages sampled by satellites as often as daily and then averaged over the month. The difference in the two collection methods and amount of data being averaged makes their comparison a rough approximation. That they are within an order of magnitude of each other is encouraging, though cannot be called validation.

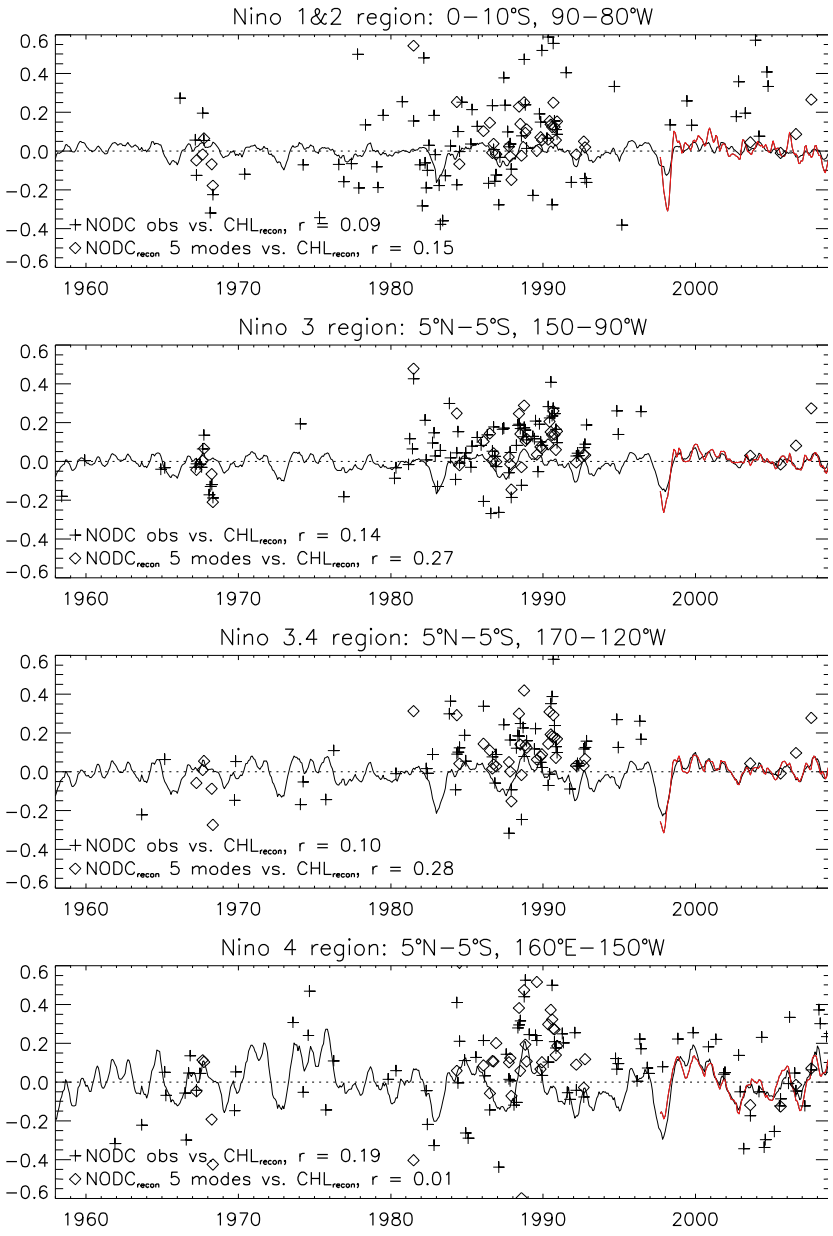

Figure 3.21 Reconstructed CHL (--), original CHL (--), NODC observations (+) and observations reconstructed using first 5 EOF modes (◊), with correlation coefficients annotated for each area.

CZCS ocean color data comparison

The CZCS ocean color sensor operated between 1978-1986 as part of a shared mission whose primary focus was the coastal zone of the United States. Elsewhere, data

coverage was sparse, as detailed in Chapter 2. The July 2011 reprocessed CZCS CHL were acquired and compared to the CHL reconstruction in the tropical Pacific. Match-ups were best near the coast where there were 50% more observations due to the sensor's mission: Niño 1&2, the best covered area, had about 1000 valid CZCS observations per bin each month compared to 700 in Niño 4, the least covered area. Despite widely varying coverage, all available data were averaged over the Niño regions and compared to the reconstructed CHL, shown in Figure 3.22. The cross-correlation between CZCS and reconstructed CHL is best in Niño 1&2 with r = 0.6 and degrades away from land: Niño 3, r = 0.5; Niño 3.4, r = 0.3; Niño 4, r = 0.2. CZCS observations in all areas indicate a high CHL anomaly in 1979/80 that is absent in the reconstruction. Since that year was unremarkable in the Niño indices, the CZCS data may have experienced an artifact, such as an atmospheric correction problem. In 1982, there were positive CHL anomalies in CZCS, notably in the equatorial Niño areas, corresponding to a weak La Niña in 1981/82. A strong El Niño followed in 1982/83 and a pronounced decrease in CHL is apparent in 1982/83 in Niño 1&2. In 1982 El Chicón volcano in Mexico erupted: its gases and particles encircled the Earth between the equator and 30^{o}N for more than six months and then spread more widely [Robock, 2002], likely biasing the CZCS measurements since such extreme aerosol events that obscure the ocean are difficult to completely correct [Gordon, 1997]. CZCS CHL matches up closer to the reconstruction than *in situ* CHL, but data paucity and possible atmospheric correction issues render it problematic.

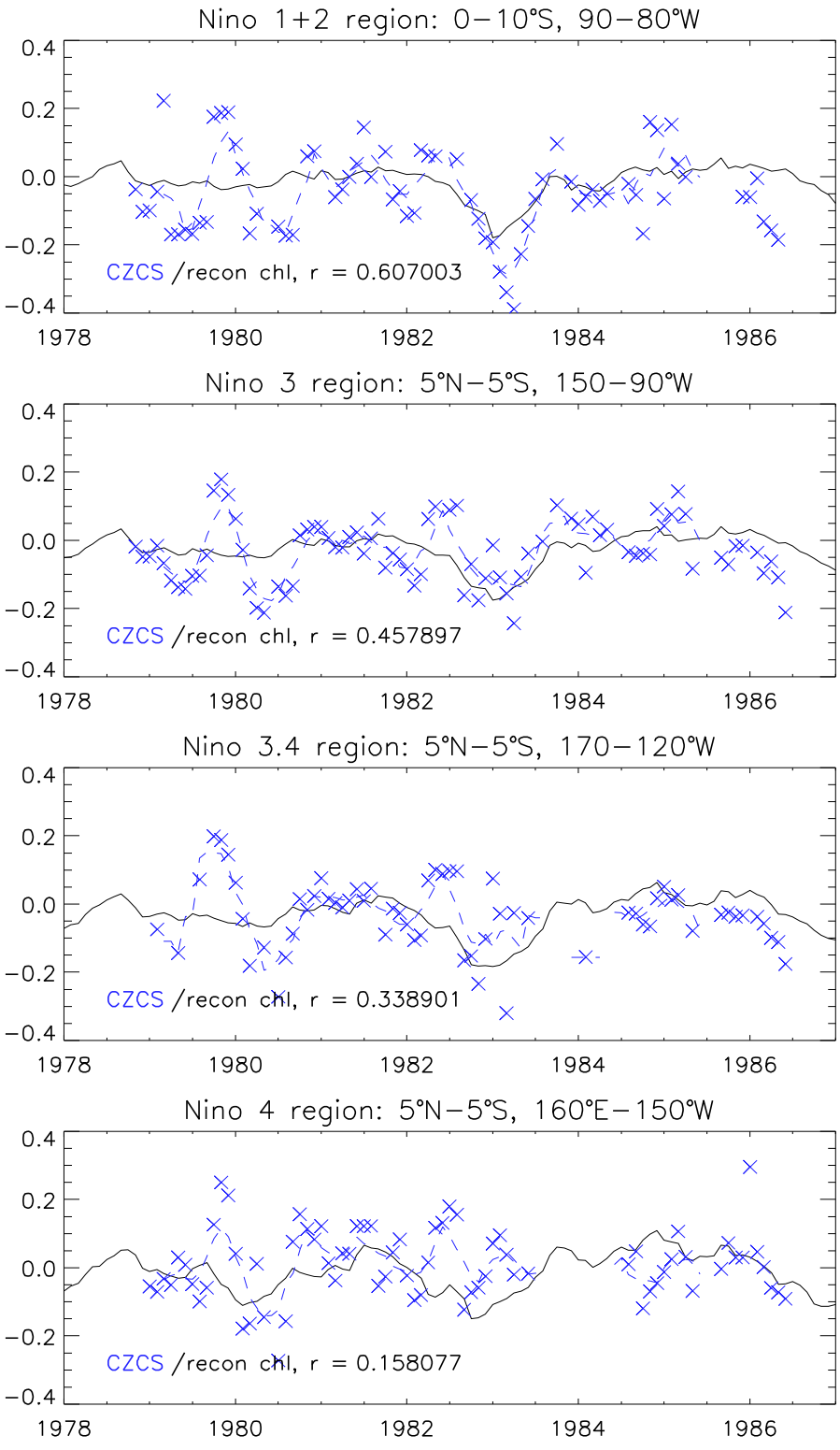

Figure 3.22 **Reconstructed CHL (-) and CZCS CHL (-x-) averaged over four Niño regions.**

Fully-coupled physical-biogeochemical ocean model comparison

Output from a fully coupled three dimensional physical-biogeochemical model developed at basin-scale for the tropical Pacific between 30^{o}S-30^{o}N was used as an independent data set for validation [Wang et al., 2009]. The biogeochemical model includes three nutrients (nitrate, ammonium, dissolved iron) and seven biological categories (small and large phytoplankton, zooplankton and detritus, and dissolved organic nitrogen) [Wang et al., 2008]. This model has been shown to have good fidelity for simulating ecosystem dynamics [Wang et al., 2005] and spatial and temporal variation in biogeochemical fields [Wang et al., 2008]. Its ocean general circulation model is based on a primitive-equation, sigma-coordinate model coupled to an advective atmospheric mixed layer model [Gent and Cane, 1989; Murtugudde et al., 1996] forced by 6-day mean surface wind stress from the National Centers for Environmental Prediction reanalysis [Kalnay et al., 1996], and climatological monthly mean solar radiation, cloudiness, and precipitation. CHL was calculated using a dynamic model with a non-steady C:Chl ratio covarying with four environmental parameters, i.e. irradiance, nitrate and iron concentrations and temperature. Modelled CHL was cross-correlated against the SeaWiFS CHL used to train the reconstructions (Figure 3.23, top), the CHL reconstructed against the CCA method (Figure 3.23, middle) and the CHL reconstructed against the MEOF method (Figure 3.23, bottom). The modeled CHL showed best validations within 10^{o} of the equator and better in the central equatorial Pacific than in the eastern equatorial Pacific. Over the entire tropical Pacific, 22% of model vs. SeaWiFS bins are significantly correlated ($r > 0.45$, $p < 0.001$); 20% for model versus CCA CHL; 19% for model versus MEOF CHL.

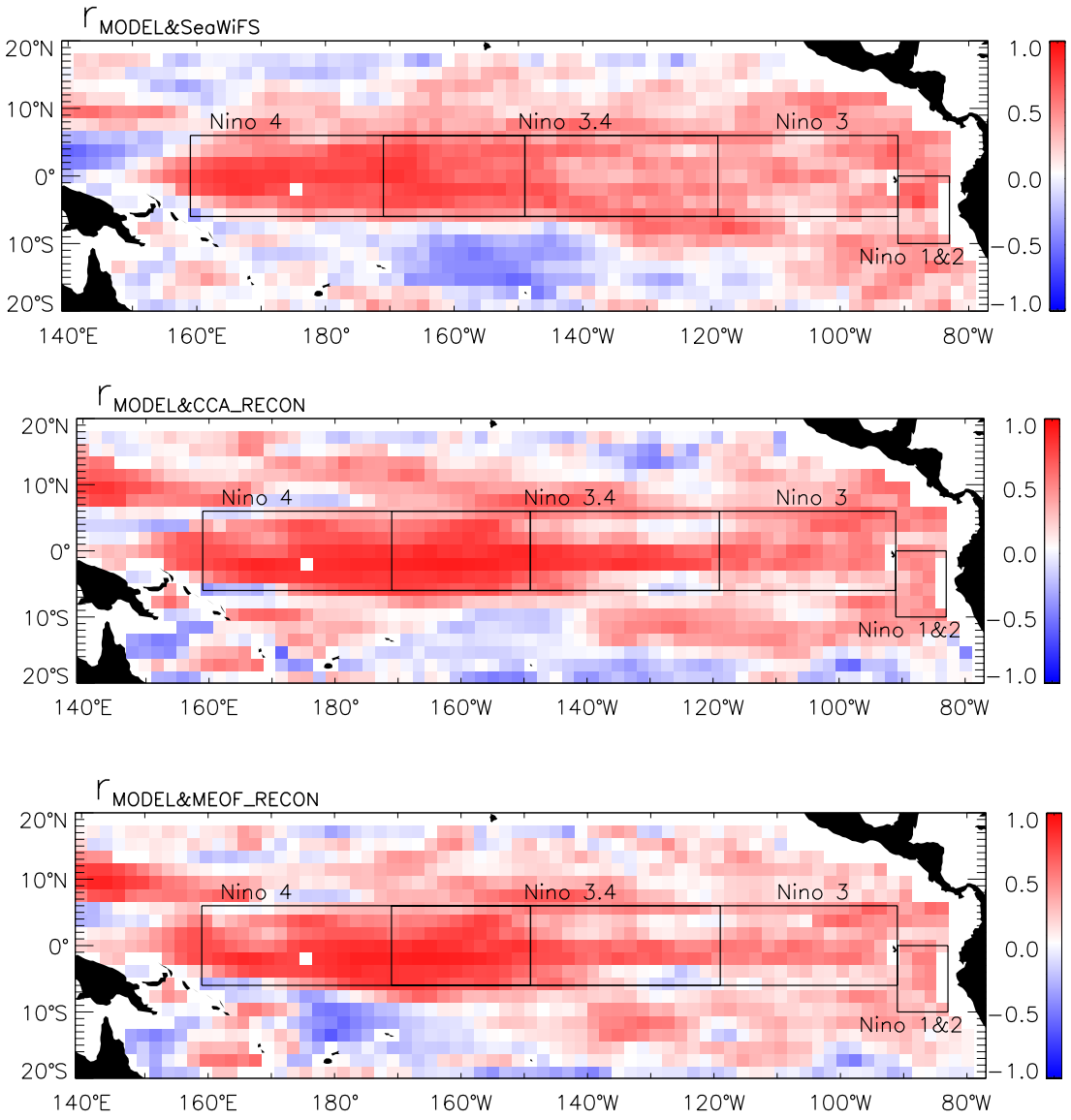

Figure 3.23 **Model CHL cross-correlated with original from 9/1997-2007 (top); model CHL cross-correlated with reconstructions over the independent period from 1988-8/1997 using CCA (middle) and MEOF (bottom). The Niño regions delineated correspond to the time series in Figure 3.24.**

The time series were averaged over each Niño area and plotted for the modeled CHL and original CHL during the dependent period and the two reconstructions with the model for the independent period (Figure 3.24). The model is closer to the original CHL and the reconstructions in the west, i.e. Niño 4 ($r \sim 0.9$) than in the east ($r \sim 0.5$).

Comparisons between the model and reconstructions are best over the equatorial areas. In the east Pacific, Nino 1&2, the complexities of the biological system appear to be captured less by the model.

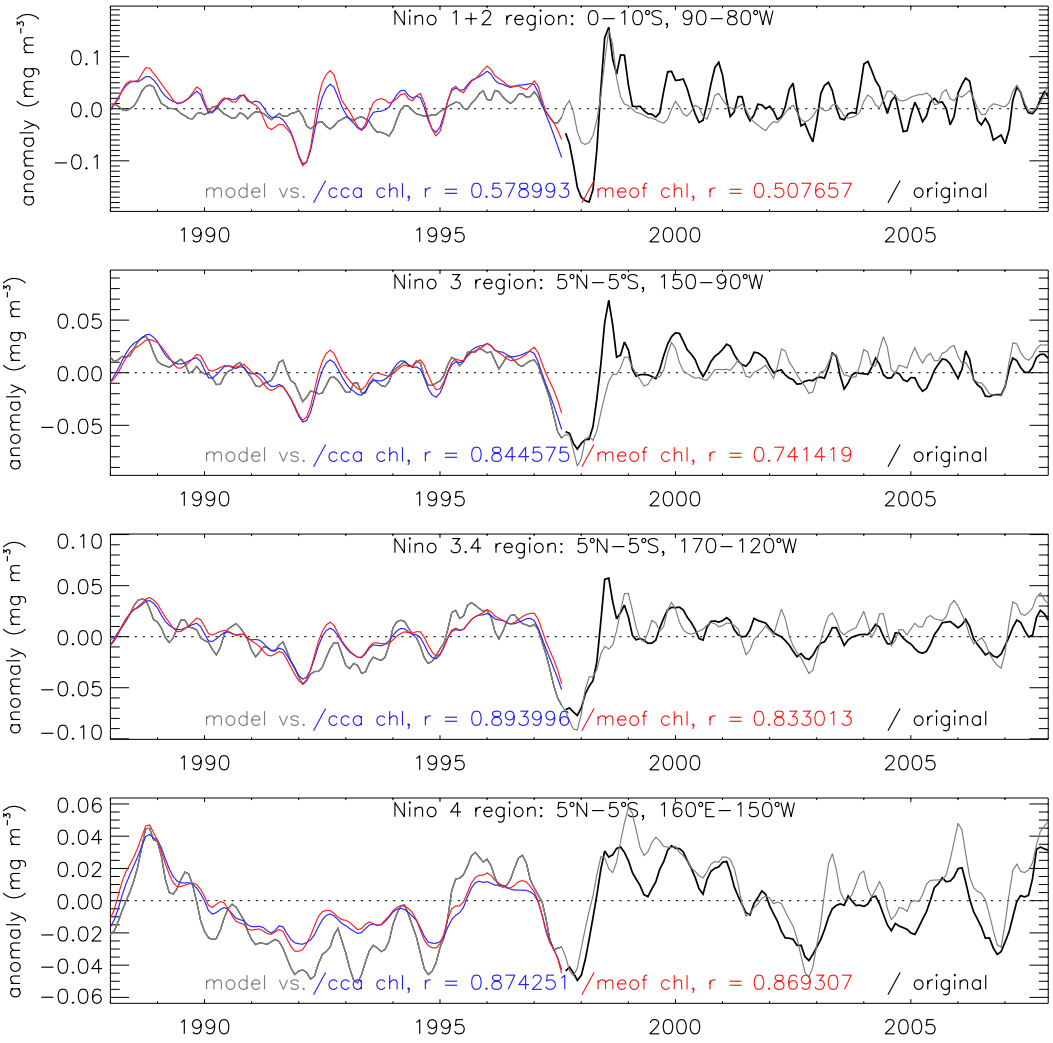

Figure 3.24 Time series corresponding to areas delineated in Figure 3.23: Niño 1&2 (top), Niño 3 (2nd), Niño 3.4 (3rd), Niño 4 (bottom): model CHL (grey), compared to CCA CHL (blue), MEOF CHL (red), SeaWiFS CHL (black). Area average cross-correlation coefficients are annotated.

Table 3.2 Cross-correlations of CHL reconstructed using CCA or MEOF with SeaWiFS during the dependent period, model and CZCS CHL during the independent period.

	Nino 1&2	Nino 3	Nino 3.4	Nino 4
SeaWiFS/CCA* 9/1997-2008	0.58	0.78	0.85	0.88
SeaWiFS/MEOF* 9/1997-2008	0.68	0.75	0.86	0.93
Model / CCA 1988-8/1997	0.58	0.84	0.89	0.87
Model / MEOF 1988-8/1997	0.51	0.74	0.83	0.87
CZCS / CCA 1978-1986	0.61	0.46	0.34	0.16
CZCS / MEOF 1978-1986	0.50	0.27	0.19	0.07

* Comparisons with SeaWiFS are leave five out cross-validations.

Summarizing the performance of the two comparable reconstruction methods: cross-validations against the original SeaWiFS CHL indicate MEOF out-performs CCA over three out of four Niño areas during the dependent period. When compared to the biogeochemical model and CZCS CHL during the independent period, reconstruction by CCA out-performs the MEOF method in three out of four Niño areas. Independent data are important for validation, as a close match-up with training data could be caused by overfitting rather than true skill by the method. Furthermore, Bretherton et al. [1992] compared different methods to determine the advantages of each toward finding coupled patterns in climate data and found that the MEOF method extracted the coupled pattern more accurately than the CCA method used in this study but MEOF exhibited a bias toward the leading EOF and may not be as stable over the independent part of a reconstruction. They found that the CCA method applied here is superior for longer time

series or when the coupled signal does not resemble the EOFs of the individual fields. For all of these reasons, the CCA method was selected for use by this study.

3.2.3 Uncertainty estimates

Uncertainty in the final results stems from the cumulative uncertainty in the observations, including their pretreatment, and the uncertainty in the method, which includes the ability of the proxies to represent the variability in the target field and that due to the number of modes applied in the statistical reconstruction and how much of the total variance they can never resolve.

Observational uncertainty

Bailey and Werdell [2006] found a root-mean-square error (RMSE) of 0.406 on log-transformed individual daily scenes of SeaWiFS CHL at native resolution of 300m to 1.1km over the deep ocean compared to the highest quality in situ samples collected within +/- 3 hours of the satellite overpass. This value decreases with greater data density, grid box size and smoothing. By binning to 2^o monthly resolution and smoothing over three months, random errors could be reduced by two orders of magnitude for the higher number of degrees of freedom of each data point. The number of degrees of freedom of each data point is conservatively estimated as $\sqrt{N}$ ~100, where N~200x30x3 assuming the 2^o bin averages include at least 25% cloud-free pixels, 30 days in a month, and smoothing over 3, however, in case all pixels going into each average are not completely independent the estimate of degrees of freedom can be pessimistically reduced to one order of magnitude. Thus binned, smoothed CHL concentrations in the tropical Pacific

used in this study are estimated to have an observational uncertainty no more than 0.04 RMSE in log-transformed units or less than 1% of the CHL value.

SODA 2.1.6 includes winds from ERA40 and QSCAT, and hydrographic data from World Ocean Database 2009 and the International Comprehensive Ocean-Atmosphere Data Set [Woodruff et al., 2011]. Atmospheric and oceanic observations used within SODA provide complimentary information that improve each other's accuracy, such that it may be reasonable to make a simplifying assumption that the ocean observations are perfect [Carton et al., 2012]. SODA has many screening and filtering steps to include only the highest quality observations, however, the input observations undoubtedly include some systematic errors (e.g. due to changes in observing systems such as the introduction of satellite observations in 1981, the increase in the height of shipboard anemometers, scarcity of data during earlier times) and random errors [Carton and Giese, 2008]. Random errors cancel out when large numbers of observations are averaged together, so this type of error can be minimized by using well-sampled eras and averaging over relatively large bins and are thus of minor importance in this large scale study. Systematic uncertainty in the global SST record has been less than 0.2^{o}C since the 1950s and then decreased further in the 1980s with increased fleet observations, satellite measurements and the initiation of drifting and moored buoys [Kennedy, 2014]. Over the tropical Pacific, SST has a seasonal range of approximately 5^{o}C, thus systematic error could be up to 4%. According to the online Satellite Altimetry Data User Handbook [http://www.aviso.altimetry.fr/fileadmin/documents/data/tools/hdbk_duacs.pdf], SSH has approximately 2.5cm noise level. Average SSH values vary up to one meter in the tropical Pacific, thus systematic error could be almost 3%.

Data pretreatment

The number of modes used in the reconstruction was optimized for this region such that each mode contains significant variance, meaning at least 1% of the dominant mode. For the CCA method, CHL variance explained in the first ten modes is 71% and 72% for combined SST and SSH. Thus, almost a third of the variability is excluded prior to calculating a reconstruction. Since this study is focused on climate scale patterns, the loss of small-scale, high frequency 'noise' is generally desirable, except in cases where the high frequency pattern signals the initiation of a climate-scale event, such as westerly wind bursts that initiate El Niño events. Analyses suggest that excluding high frequency patterns adds insignificant uncertainty to the climate-scale patterns [Liu et al., 2015].

Method uncertainty

All components of method uncertainty can be quantified by comparing the reconstruction to the original data, including how well the proxies represent the target variable and how well the dominant modes capture its low frequency signal. The reconstruction had the closest match-ups to the original data away from the coasts and toward the equator where physical signals dominate the biology. Differences between the original data and the reconstruction during the dependent period were used to calculate RMSE in normalized units, with uncertainty ranging from RMSE~0.27 for the equatorial Niño areas and RMSE~0.6 for Niño 1&2. For the independent period, cross-validations were used to estimate the average normalized RMSE = $\sqrt{2(1-r)}$ [Smith et al., 1995], which ranges from RMSE~0.5 in Niño 4 to RMSE~0.9 in Niño 1&2. RMSE

calculations were unnormalized using the standard deviations and plotted with the CHL anomalies (Figure 3.25).

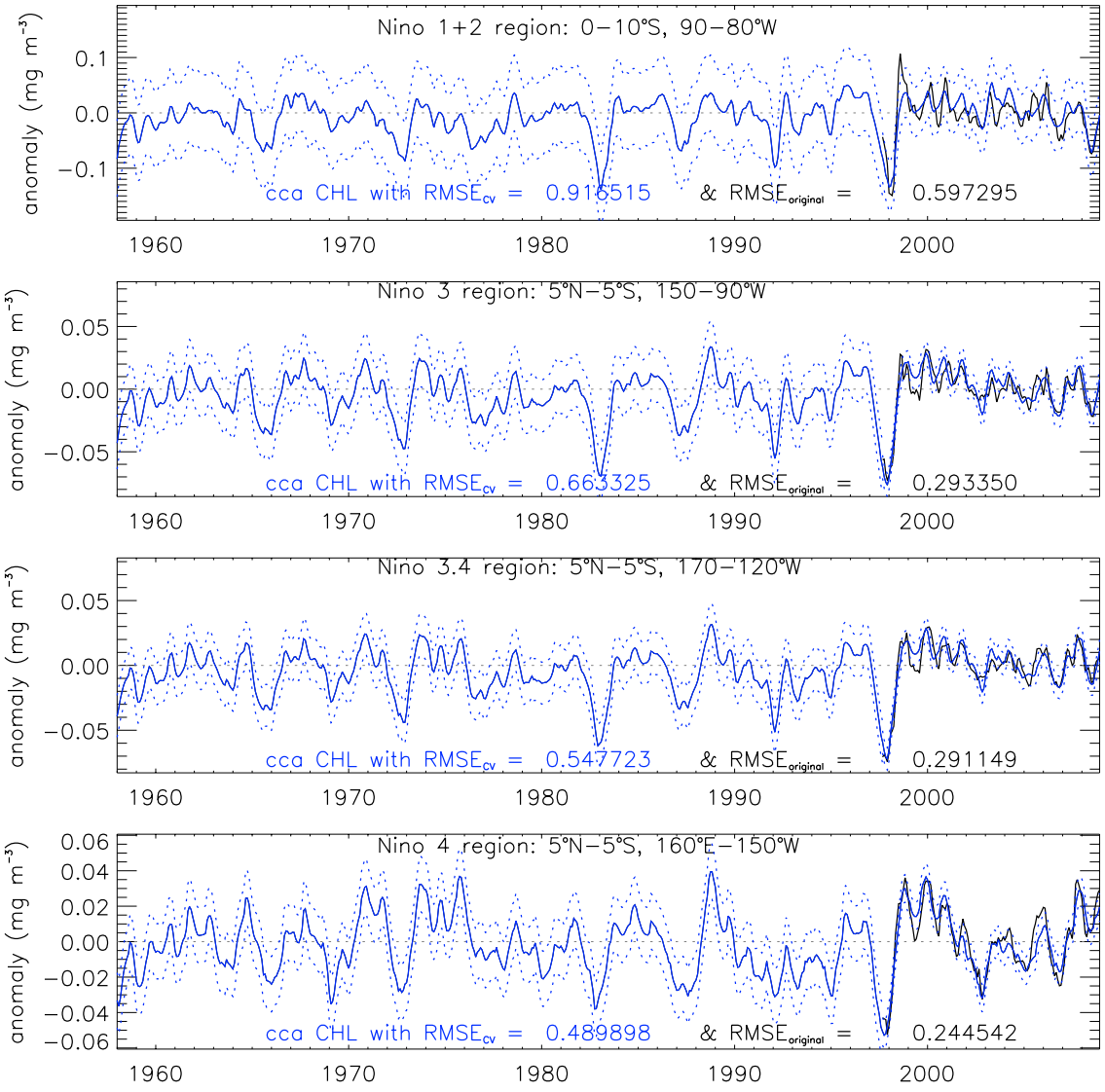

Figure 3.25 Reconstructed CHL (blue) +/- RMSE (dotted blue) and original CHL (black). RMSE during the dependent period is calculated between the original and CCA CHL; RMSE during the independent period is estimated using the leave-five-out-cross-validations, as described in the text.

Uncertainty in the final results

In this kind of statistical reconstruction, the main uncertainty is due to the method. Observational uncertainty is demonstrated to have a minimal contribution to the final results by testing it. If systematic errors are uniform, their effect can be ameliorated by deseasoning and analyzing anomalies. If, however, systematic errors are random and non-uniform, they could potentially introduce some error that would propagate through to the final results. To illustrate the relative contribution of the data error, several extremes of hypothetical observational errors were tested to characterize the sensitivity of the final results to observational errors. Random months and bins of the SST data set were perturbed by 0.2^{o}C and CHL was reconstructed with all other parameters being the same (i.e. same training data and EOFs): a systemic bias applied to the entire basin and a random bias over a 10^{o}x4^{o} box on the equator at 140^{o}W was tested for a single month up to five years. Perturbations of the SST record caused less than 1% difference to the CHL reconstruction in every case, with the exception of the last case (equatorial box over five years) that spiked a 0.01 mg m^{-3} difference in Nino 1&2 during the 1982/83 ENSO or almost 5% difference to CHL concentrations there during that time. The equatorial Nino areas had CHL differences that remained around 1% or less for all of the perturbations of the SST records. Likewise, SSH was perturbed by 2.5cm over the same box for five years and yielded consistent magnitudes with different temporal variability. Only the 1997/98 El Nino caused a spike of about 5% of the CHL signal in Nino 1&2, while the equatorial Nino areas were always within 1-2%.

Because the reconstruction is performed in EOF spectral space, the final results are much less sensitive to observational errors than they would be with other computational techniques. Observations are first mapped to EOF spectral space using the orthogonality property of the spatial functions. Therefore the uncertainty introduced by an observational error at one location does not impact the results only at that place, but instead influences the amplitude for the entire domain, so the propagation of errors is not a local problem. Furthermore, any observational errors that do not project onto one of the dominant modes have no appreciable effect on the results, so the sensitivity of the results to observational errors cannot be fully characterized without specifying the errors in space and time. The observational error contributions estimated here for extreme hypothetical systematic errors in the observations are still relatively small and the main uncertainty in the statistical reconstruction stems from the method of reducing the data to the dominant modes.

How well the EOF modes are defined during the training period and continue to capture the dominant variability in the reconstruction period determine the method uncertainty. Truncating the signal through EOF space has been found to be the most stable small-sample method [Bretherton et al., 1992]. Yet reducing the variance to ten modes immediately excludes about a third of it over the dependent period, resulting in essentially all of the uncertainty in the equatorial Niño 4 area and approximately half of the uncertainty in coastal Niño 1&2. The reduction of variance is primarily manifest in damping (Figure 3.25). The reconstructed CHL demonstrates high fidelity in capturing general patterns, notably ENSO, but damps their amplitudes, occasionally exceeding the margin of error (e.g. in 1998/99 in Nino 3, 3.4 and 1&2; 2006/07 in Nino 3.4 and 4).

Cross-validation is used to assess the skill of statistical reconstructions [Barnett and Preisendorfer, 1987; Smith et al., 1996; Mann et al., 1998; Smerdon et al., 2011] and provides the best analysis of the overall uncertainty in the final results. The spatial distribution of the errors is consistent with the fraction of variance resolved by the first ten modes (Figure 3.16) as all other sources of error are shown to be at least an order of magnitude smaller. The uncertainties in the dependent period were quantified through RMS differences with the training data and are lowest within 10^{o} of the equator (~24-30%); errors are double that or more away from the equator and toward the coasts. Uncertainties in the independent period were approximated through cross-validation and are within acceptable limits for open ocean, equatorial areas (e.g. Niño 3/3.4/4~50%), but unacceptably large toward the coasts (e.g. Niño 1&2 ~90%).

3.3 Extreme El Niño and La Niña in 1998

Data thinning experiments have shown that the SODA SST observations are adequate to capture the 1990s, including the extreme ENSO event, while the sampling density of the 1920s and 1940s would have been inadequate to catch it [Giese and Ray, 2010]. During the dramatic transition from the 1997/98 El Niño to the 1998/99 La Niña, the greatest difference in CHL was between December 1997 and December 1998 in the original and reconstructed CHL. In Figure 3.26, greater CHL in 1998 is shaded red; greater CHL in 1997 is blue. Note that the ends of the color scale indicate values greater than 0.15 mg m^{-3} and less than -0.15 mg m^{-3}, respectively, which is the range for the area of interest. Positive ΔCHL reflects higher CHL during La Niña in coastal and equatorial

upwelling areas. In the west Pacific warm pool, CHL was higher during El Niño and returned to oligotrophic conditions during La Niña. Overall, the original CHL (Figure 3.26, top) has more spatial variability than the reconstruction (Figure 3.26, bottom) since the reconstruction uses the dominant ten modes. The differences west of the cold tongue are due to the reconstruction being smoother and having less variability.

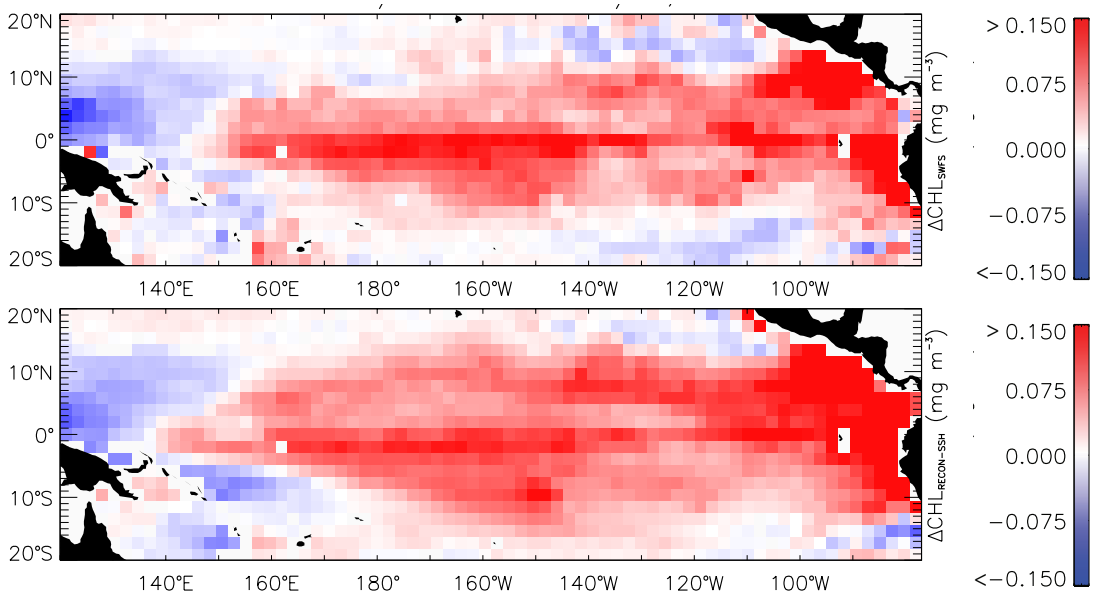

Figure 3.26 La Niña – El Niño CHL differences for December 1998 minus December 1997 averaged for the original CHL (top) and reconstructed CHL (bottom). Note the end points of the color scale represent values greater than 0.15 mg m^{-3} and less than -0.15 mg m^{-3} respectively.

The differences in the physical variables show similar spatial distributions (Figure 3.27) to the CHL. The most striking difference is that positive ΔCHL extends more than 20^{o} farther west along the equator than the negative ΔSSH in an area of very weak ΔSST. During El Niño, SST was warmer and SSH was higher in the east and central equatorial upwelling regions. During La Niña, SST was warmer and SSH was higher to the west while CHL concentrations decreased. Physical change appears consistent with changes in

the original CHL (Figure 3.26, top), thus the inability of the reconstruction changes (Figure 3.26, bottom) to exactly replicate the difference between the phases of ENSO is most likely caused by mesoscale features distributed to higher EOF modes. The dominant EOF modes give more weight to basin-scale phenomenon, leading to localized discrepancies between original and reconstructed CHL.

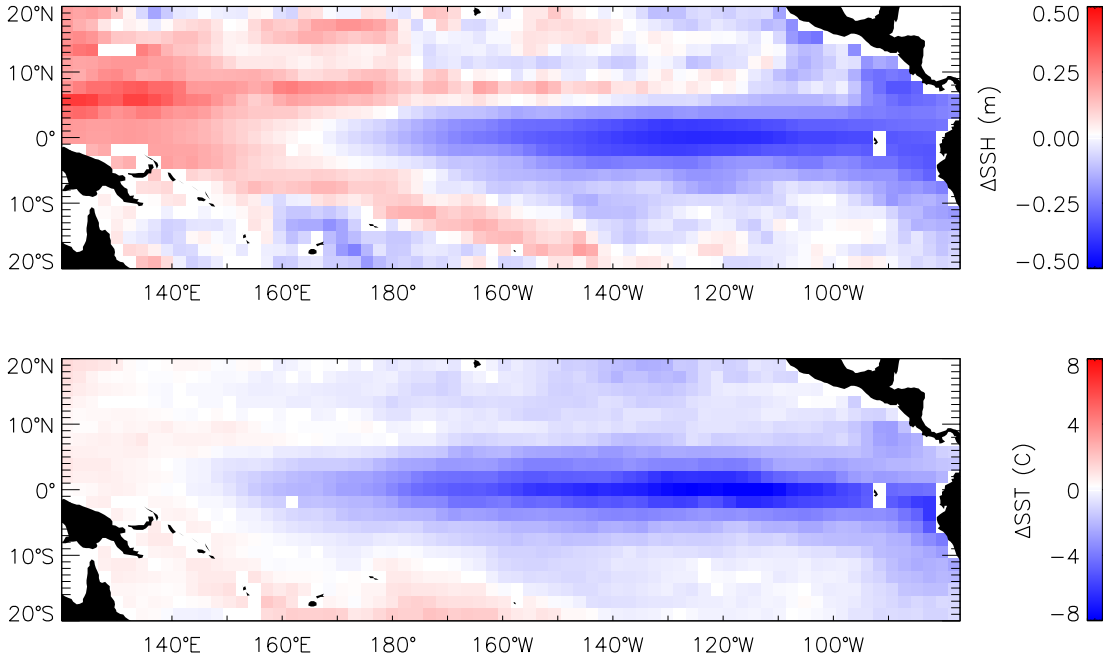

Figure 3.27 La Niña – El Niño differences in physical variables: December 1998 minus December 1997 for SSH (top) and SST (bottom).

Overall, both the fully-coupled physical-biogeochemical ocean model comparisons and the leave-five-out-cross-validations indicate the CHL reconstruction has good skill in the equatorial Pacific where the physical forcing is strong. Results were inconclusive for the sparse *in situ* samples and CZCS data set. Reconstruction skill is highest away from the coast in the western Pacific and decreases toward the east and the coasts. Because the statistical reconstruction uses the ten dominant modes out of 136, the reconstructed CHL has a smaller amplitude than the original CHL. Differences between

the reconstruction and the original data could be due to the absence of higher modes, which may help to describe the extreme transition from the large El Niño in 1997/98 to La Niña in 1998/99. Using the most closely correlated physical predictors, this statistical reconstruction extends 2^{o}, monthly CHL anomalies from just over a decade to just over five decades. The effect of climate-scale physical forcing on surface CHL patterns and related ocean biology in the tropical Pacific are explored in the next chapter.

4. Low frequency chlorophyll anomalies in the tropical Pacific

Quantifying slow changes in marine phytoplankton is important for understanding their impact upon ecosystems as well as their role in regulating climate. The ability to observe and model complex physical-biological patterns has improved, especially over the vast expanse of the equatorial Pacific. Yet even the most realistic models remain sensitive to unknown parameters. A multivariate statistical reconstruction has the advantage of being able to include phenomena, known and unknown, that are large enough to impact dominant modes of variability. As detailed in Chapter 2, many processes experienced in the tropical Pacific have global ramifications. There is evidence that slow changes have been taking place in biological features, either through natural low frequency oscillations or anthropogenic effects. This chapter discusses the interannual and decadal CHL variability captured in the training and reconstruction periods and how large-scale CHL patterns are impacted by low frequency climate oscillations.

4.1 Variability over the training period: 1997 - 2008

Understanding mechanisms influencing the spatial distribution of tropical Pacific CHL during the training period provides insight into what can be reconstructed. Continuous ocean color satellite derived CHL exists between September, 1997 and December, 2008. The correlation of physical variables to CHL during the training period established the linear relationships in the CCA method which were used to reconstruct CHL over the entire proxy record. A CHL bloom caused by a unique process independent of the physical oceanographic variables used in the training period

may not be fully reconstructed (e.g. fertilization by dry deposition of aerosols). Processes not explicitly represented in the training period could be represented in the reconstruction if a linear weighted sum of the reconstruction modes can account for the process. However, processes will always be better represented in a reconstruction when they are represented in the training data. Highest CHL is found in the east, near the coast, and along the equatorial cold tongue where high variance in CHL, SST and SSH corresponds to upwelling of nutrients along a shallow thermocline. High variance but low mean CHL is found in oligotrophic areas. Thermocline uplift in this region is associated with Kelvin, Rossby and tropical instability waves and has been shown to lift nutrients into the euphotic zone causing blooms [Turk et al., 2001]. The variability of CHL is made more visible by normalizing the standard deviations by the means to enable oligotrophic areas with high variability to stand out along with consistently high CHL areas (Figure 4.1).

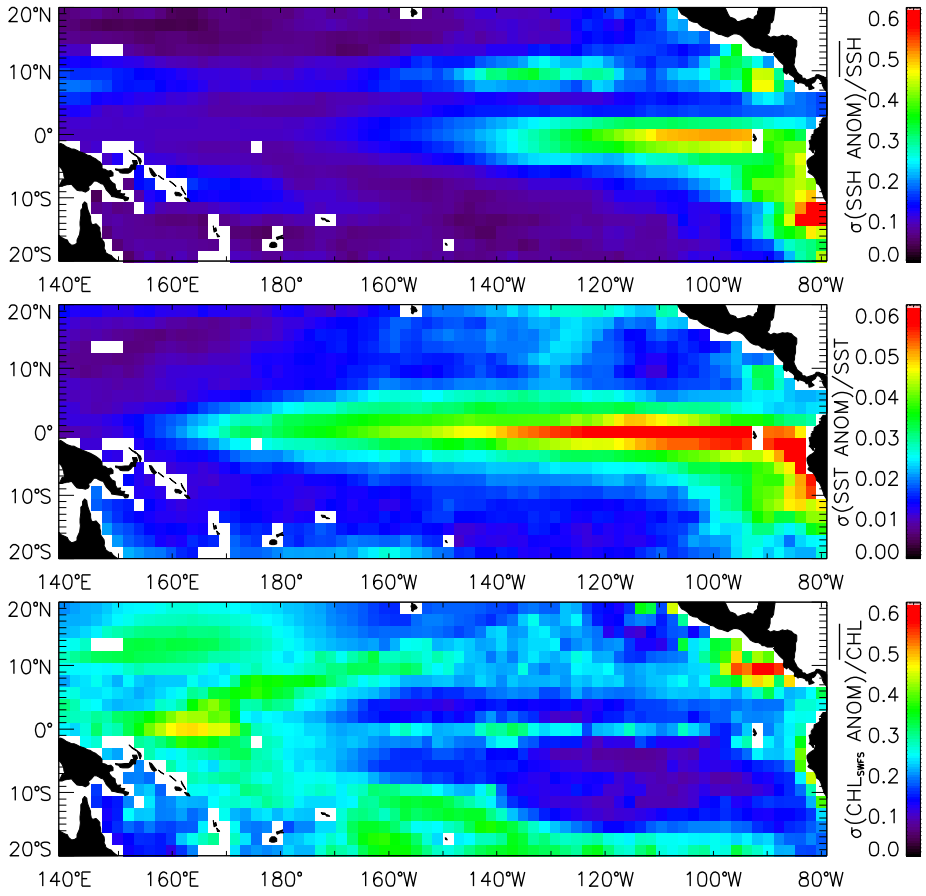

Figure 4.1 **Standard deviation of the deseasoned anomalies normalized by the mean for SSH (top), SST (middle), and CHL (bottom) between September, 1997 – December, 2008.**

The largest forcing mechanism in the tropical Pacific appears to be ENSO, with residual forcing by seasonal and other smaller effects. While the beginning of the period experienced the largest eastern Pacific El Niño of the century in 1997/98, subsequent El Niños only reached the central Pacific, centered near 150°W (2002/03, 2004/05, 2006/07) [Radenac et al., 2012]. The three central Pacific El Niño events registered in Niño 4, were barely discernable in Niño 3.4, and are not distinguishable in Niño 1&2 and 3. Five La Niñas during the training period (1998/99, 1999/2000, 2000/01, 2005/06, 2007/08) versus four El Niños contributed to a cooler training period that may have been more productive

than normal. Smaller, higher frequency variability is especially apparent in Niño 1&2, but most of the variability in the other three areas corresponds to ENSO events. El Niño and La Niña events cause the dominant variability as the oligotrophic warm pool shifts east during El Niño and west during La Niña.

4.2 Variability over the reconstruction period: 1958 - 2008

In the reconstruction, CHL patterns are most evident when viewed as anomalies from a mean seasonal cycle. The period over which the cycle is defined is somewhat arbitrary. Since there is the least uncertainty over the training period, its climatology could be used to define anomalies. However, low frequency variability causes differences in the long-term averages between the training and reconstruction periods. The training period is mostly during a decade of higher productivity. Because of this, reconstructed CHL was demeaned in reference to the reconstruction period seasonal climatology.

Changes in CHL and physical variables are plotted for each longitude averaged over 5^{o}S-5^{o}N between 1958-2008 (Figure 4.2). The covariation between CHL and the physical variables is strongest and most consistent with SST and more subtle with SSH. Although deseasoned anomalies were reconstructed, a weak annual cycle remains suggesting slight changes in the annual cycle over the reconstruction period. The most prominent feature in the Hovmöller plot is El Niño, where a less productive, warmer, thicker surface layer begins in the west Pacific and propagates eastward – sometimes as a very strong anomaly that reaches the east Pacific (e.g. 1982, 1997) and sometimes as a less obvious lower CHL, warmer, thicker surface layer that reaches the central Pacific and is weaker toward the east (e.g. 1958, 1994, 2004). After ENSO, a multi-decadal

signal is faintly evident as the cool phase prior to 1976 dominated by positive CHL anomalies, especially just before the end of the phase; the following warm phase until 1998 has less CHL; finally, during the cool phase through 2008, positive CHL anomalies are more common, with the exception of the western half of the equatorial Pacific where negative CHL anomalies predominate. Processes linking physical changes to ecosystem changes are complex and multivariate, but several that have been observed and reported in the literature are evident in the time-longitude plots, most obviously changes associated with El Niño and La Niña events [Radenac et al., 2012; Turk et al., 2011].

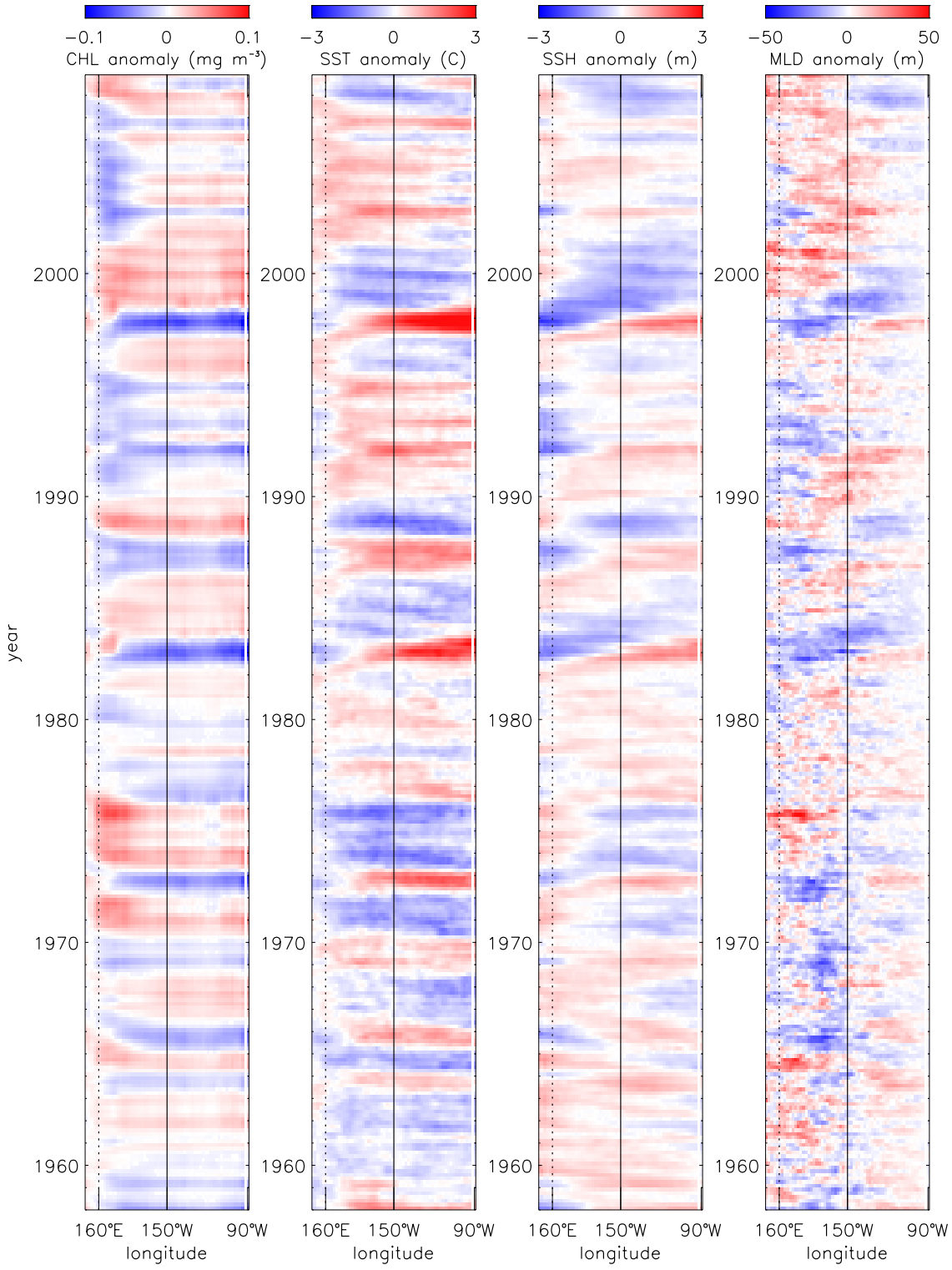

Figure 4.2 Longitude–time plots of reconstructed CHL (left), SST (2nd), SSH (3rd), MLD (right) between 1958-2008. Niño 4 is between dotted and solid lines; Niño 3 is east of solid line.

4.2.1 Interannual patterns in chlorophyll

The dominant signal in this region is ENSO, an interannual oscillation of roughly three to seven years that typically peaks in December but varies in intensity, extent and duration. During a very strong El Niño, when the Ocean Niño Index or ONI > 2, the west Pacific warm pool propagates eastward along the equator toward the coast of the Americas. A rising thermocline in the west results in higher CHL there; a depressed thermocline in the east, lower CHL east of 160°E. During the reconstruction period, there were three very strong El Niños that started in 1972, 1982, 1997 (Figure 4.3).

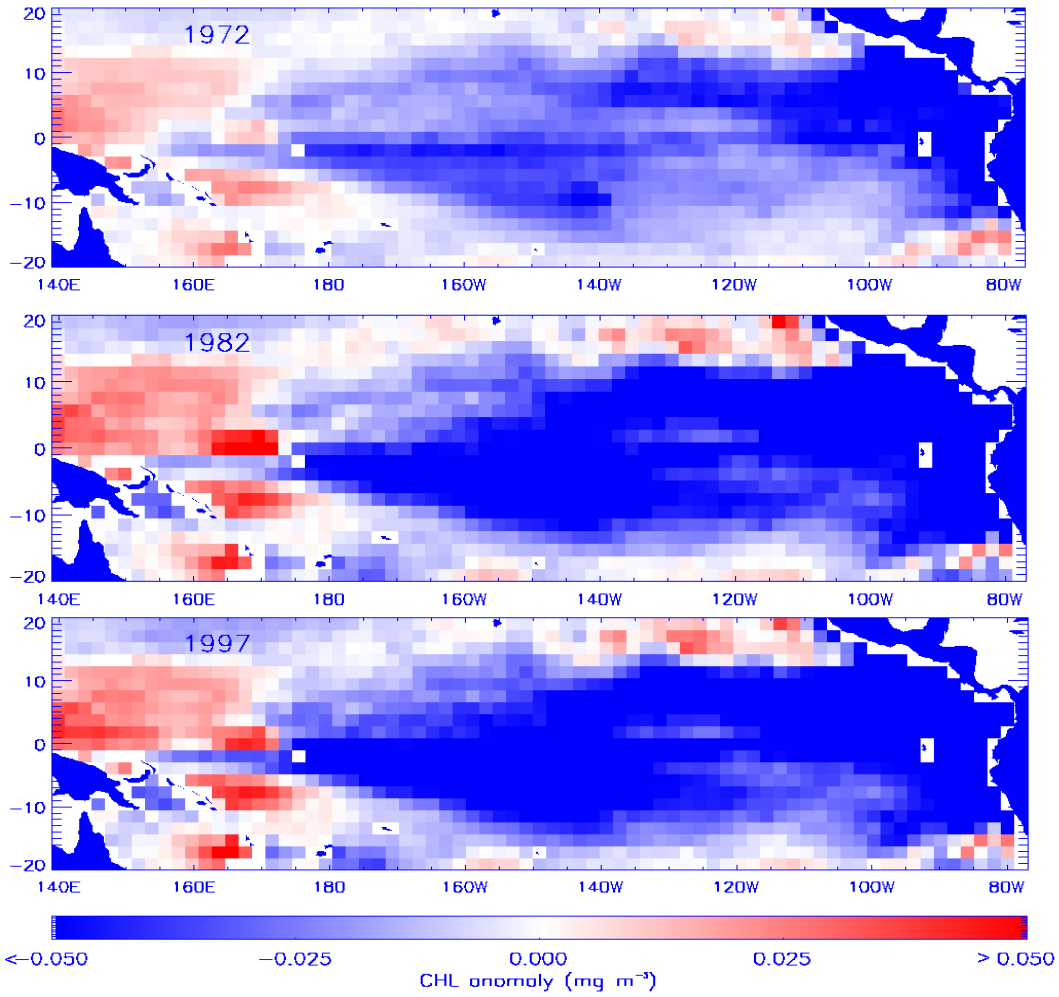

Figure 4.3 **Annual CHL anomalies during very strong El Niño events (ONI > 2) that formed during 1972 (top), 1982 (middle), 1997 (bottom). Averages are centered on December +/- 6 months.**

There are changes in CHL that correspond to canonical El Niño events such as these three very strong episodes. Below the oligotrophic west Pacific warm pool, the nutricline is normally deep (~200m). At the start of El Niño events, the easterly trade winds weaken allowing the west Pacific warm anomaly to propagate eastward along the equator through the eastern half of Niño 4 and Niño 3. As the warm anomaly moves eastward, its deeper surface layer, with positive SSH and MLD anomalies, causes the reduction of equatorial upwelling to the surface. Without its supply of nutrient-rich deep water, surface CHL decreases toward the east as the nutricline is pushed below the euphotic zone there. By contrast, in the west, a rising thermocline results in a cooler, shallower surface layer with negative SSH and MLD anomalies and positive CHL anomalies associated with the rise of the nutricline into the euphotic zone [Turk et al., 2001]. While El Niño events generally follow this basic pattern, there is much variability in their magnitude, eastward extent, and duration. The three very strong El Niños were the most pronounced over the 51 year time series. Strong El Niños ($2 > ONI > 1$) started in 1957, 1965, 1986, 1991, 1994, 2002, and almost in 2006. Stronger warm anomalies correspond to less CHL, as evidenced in Figure 4.2.

Although the effect of an El Niño event upon primary production is primarily related to its strength, its eastward extent also affects biology. CHL anomalies are more pronounced when the El Niño extends all the way to the east compared to those that stop mid-way. Years are classified as a central Pacific El Niño when the warm anomaly is greatest there, defined by the DJF Niño4 index greater than the DJF Niño3 index [McPhaden et al., 2011; Yu et al., 2012], as happened in 1958/59, 1963/64, 1968/69, 1977/78, 1987/88, 1994/95, 2002/03, 2004/05, with examples shown in Figure 4.4.

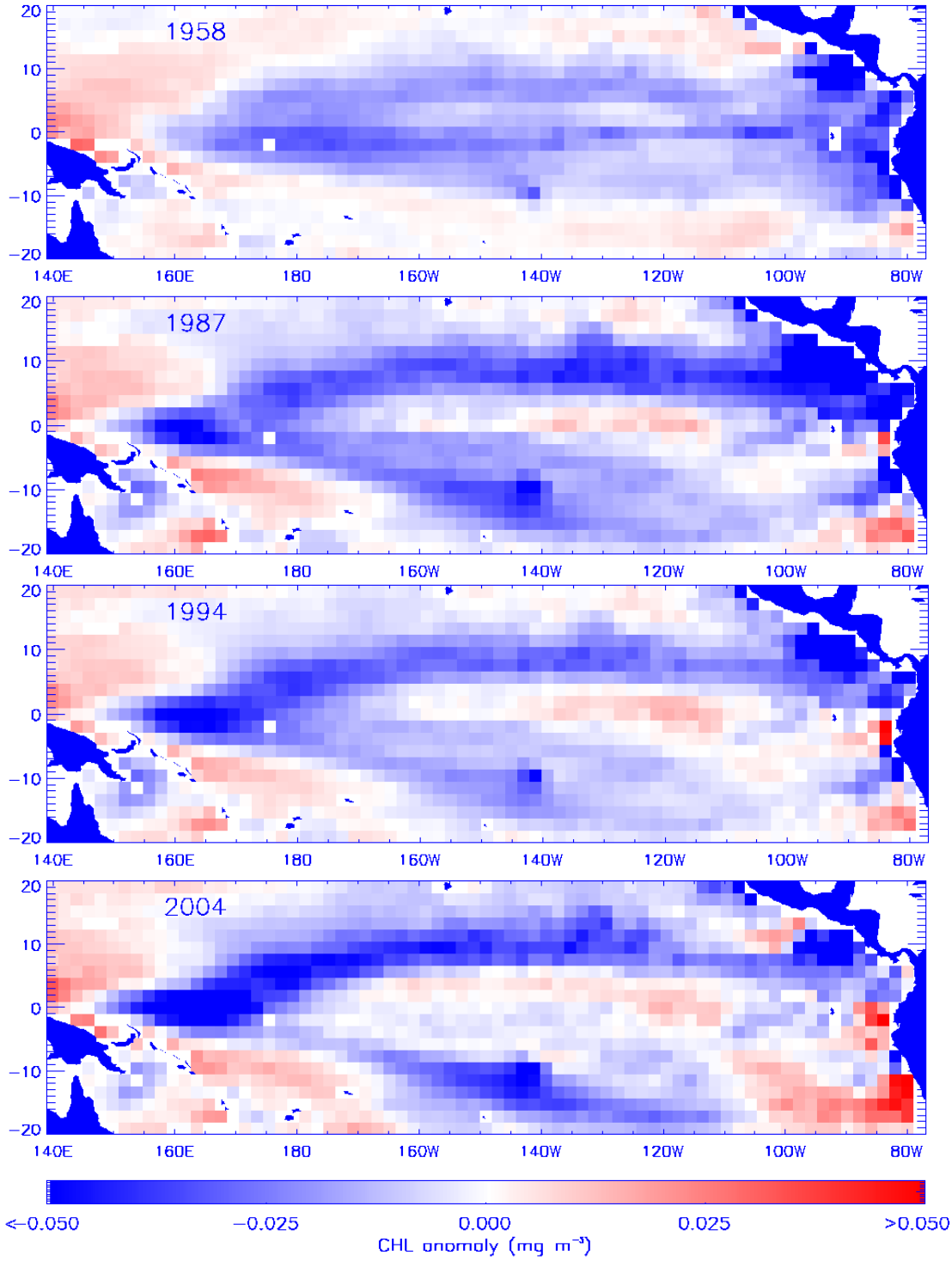

Figure 4.4 **Average CHL anomalies during central Pacific El Niño events that formed during 1958 (top), 1987 (second), 1994 (third), 2004 (bottom), centered on December +/- 6 months. 1958 started as a strong El Niño (2 >ONI > 1), but turned weak by February while the other three remained strong.**

Although the physical processes that cause central Pacific El Niño events are not completely understood, they appear to have different forcing and consequently result in a different effect upon biology [Turk et al., 2011; Radenac et al., 2012]. A central Pacific El Niño does not experience the basin-wide cessation of the trade winds nor the subsequent relaxation of the east-west thermocline slope [Ashok et al., 2007; McPhaden et al., 2011]. It does experience zonal advection of the warm pool that is limited in its eastward extent by westward equatorial surface currents from the eastern basin, thought to be a local response to wind forcing that may originate in the subtropics [Yu and Kao, 2007; Kao and Yu, 2009; Yu et al., 2012; Karnauskas, 2013]. The biological impact of a central Pacific El Niño is limited to decreased CHL in the central Pacific, while wind-driven equatorial and coastal upwelling farther east continue to support productivity. Because a central Pacific El Niño does not experience basin-wide relaxation of the thermocline slope, its effect on nutricline depth and biological processes is localized to 150-180^{o}E and less pronounced than that of an east Pacific El Niño, as evidenced in Figure 4.4 and the figure by Turk et al. [2011] shown in Figure 1.6 as well as other studies [DiLorenzo et al., 2010; Radenac et al., 2012; Messie and Chavez, 2012]. Coastal and equatorial upwelling in the east continue, thus a central Pacific El Niño has a smaller effect upon primary productivity over the entire region than that of an east Pacific El Niño. At the termination of any El Niño, the normal west-east thermocline slope returns or develops an exaggerated slope corresponding to La Niña.

The opposite ENSO phase, La Niña, manifests as enhanced equatorial upwelling (negative SST, MLD and SSH anomalies) lifting deep nutrients into the euphotic zone and resulting in positive CHL anomalies throughout the cold tongue. During the

reconstructed period, La Niña events (ONI < -1) began in 1964, 1970, 1973, 1975, 1988, 1998, and 2007, with examples shown in Figure 4.5. Although some studies have noted increasingly frequent El Niño events limited to the central Pacific and fewer La Niña events, in this study there is no clear trend toward more frequent central Pacific El Niños nor less frequent La Niñas during the reconstructed CHL time series, but rather evidence that lower frequency variability impacts interannual variability.

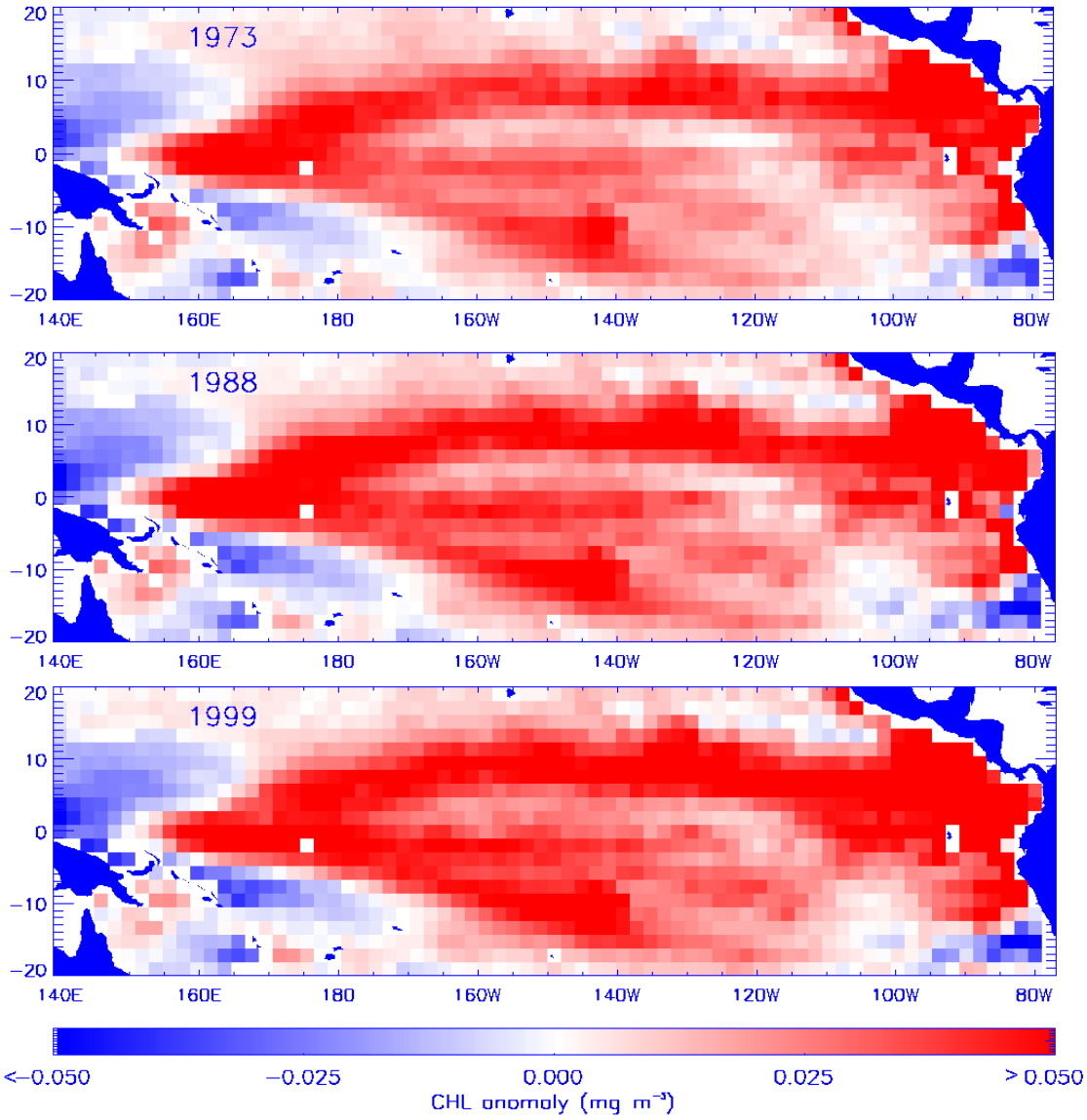

Figure 4.5 **Annual CHL anomalies during very strong La Niña events (ONI < - 2) that peaked during 1973 (top), 1988 (middle), 1999 (bottom). Averages centered on December +/- 6 months.**

The unique perspective provided by a 51 year ocean color CHL reconstruction in the tropical Pacific shows the dominance of the ENSO pattern on ocean biology, the difference between eastern and central Pacific El Ninos, and slower changes that impact the strength of the interannual patterns. The interannual spatial patterns in the multi-decadal CHL reconstruction are roughly consistent with those identified by several previous studies, namely the ocean color distributions associated with different phases and eastward extents of ENSO [e.g. Wilson and Adamec, 2001; Ryan et al., 2002]. If central Pacific El Ninos increase in frequency, as some have suggested, then prevailing equatorial biological dynamics would be expected to change [Turk et al., 2011; Gierach et al., 2012; Radenac et al., 2012]. During the reconstruction period, central Pacific El Nino events appeared more pronounced in 1994, 2002, 2004 than in 1958, 1968, 1986, meaning that negative CHL anomalies were increasingly negative between 150-180^{o}E with stronger positive anomalies farther east toward the end of this study. The last decade of the reconstruction (i.e. the SeaWiFS period used in previous studies) was more productive in the tropical Pacific overall than the preceding four decades, owing in part to lower frequency oscillations that will be discussed next.

4.2.2 Decadal patterns in chlorophyll

A multi-decadal CHL reconstruction enables the first synoptic view of slow changes in a tropical Pacific biological record. While the interannual ENSO patterns are clearly the dominant signal in the tropical Pacific biology, there are some interesting decadal patterns that appear to enhance or dampen interannual patterns. As shown in Figure 4.6, the first decade of the reconstruction period (1958-1967) had close to average

CHL amounts over most of the tropical Pacific with the exception of the Costa Rica Dome where CHL was higher than normal; elsewhere there was slightly higher CHL in the cold tongue except along the equator where CHL was lower. Between 1968-1977, there were much higher CHL values near Central America and the cold tongue except along the equator east of 160^{o}W where it was lower; over much of the warm pool in the west, CHL was lower than average. Between 1978-1997, the pattern reversed: lower CHL values around the Costa Rica Dome and in the cold tongue, with higher values in the warm pool and along the equator east of the dateline. For the last decade of the reconstruction, 1998-2007, the pattern nearly reversed again with higher CHL around the Costa Rica Dome and the cold tongue. However, this decade did not have an opposite CHL anomaly along the equator. The cold tongue did not extend as far west and north, likely due to several central Pacific El Niños between 2002-2005 that seemed to have a strong impact on the decadal average, giving it more similarity to the 2004 annual average (Figure 4.4, bottom). An inspection of monthly and annual anomalies shows that by 2008 the pattern resembled the 1968-1977 decadal average. Decadal variability in the Pacific is typically evaluated using the PDO index. The climate forcing mechanisms causing these interannual and decadal differences in ocean biology across the tropical Pacific are considered next.

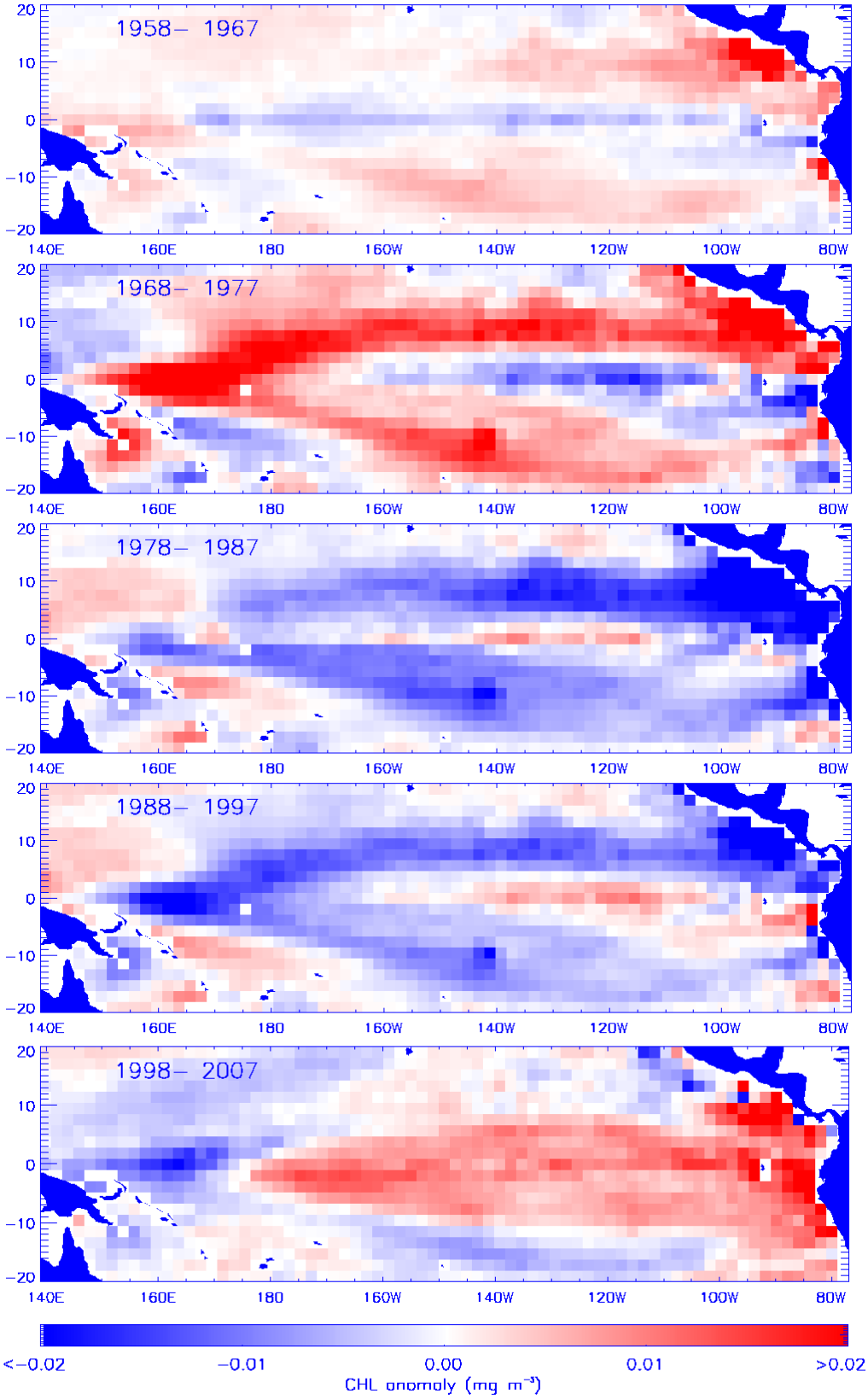

Figure 4.6 **Decadal averages of reconstructed CHL anomalies, as annotated for 1958-2007.**

4.3 The connection between large-scale climate oscillations and biology

Climate oscillations are recurring fluctuations in the ocean-atmosphere system that are evident in physical variables, such as temperature, and sometimes evident in biological records. One way to classify a quasi-periodic pattern that has been identified but not yet attributed to a known mechanism is through a climate index that uses several physical variables representative of the large-scale phenomena [Drinkwater et al., 2010]. The National Weather Service Climate Prediction Center calculates the Ocean Niño Index (ONI), which is one index among many used to represent the interannual ENSO signal. Although ENSO has been discussed extensively in many observational and modeling studies over the past 30 years, skill of its prediction has declined in the past decade indicating the limitations of current understanding [McPhaden, 2012].

The Pacific Decadal Oscillation (PDO) is an index created by Nate Mantua at the University of Washington's Joint Institute for the Study of the Atmosphere and Ocean that represents decadal forcing in the Pacific. PDO is the oceanic expression of the dominant ocean-atmosphere variability in the North Pacific and provides a link between the tropics and extra-tropics. Proxy and instrumental records support the idea that PDO variability originates in the tropics [Evans et al., 2001], with the greatest effect upon climate and biology outside the tropics [Hare and Francis, 1995; Mantua et al., 1997; Mantua and Hare, 2002]. The PDO cool phase is characterized by stronger tropical easterly trade winds, a steeper west-east pycnocline slope along the equator and enhanced upper ocean meridional overturning circulation [McPhaden and Zhang, 2002] which all support greater upwelling along the equator. Where the nutricline is shallow enough, this upwelling supplies nutrients to fuel phytoplankton blooms, similar to a La Niña pattern.

By contrast, during a warm PDO phase, easterly trade winds weaken, the normal west-east pycnocline slope relaxes, meridional overturning circulation slows and equatorial upwelling decreases. A PDO warm phase tends to enhance an El Niño pattern and suppress a La Niña pattern while the PDO cool phase tends to do the opposite. Reconstructed CHL cross-correlated to the 12-month smoothed ONI and PDO indices shows ENSO with the dominant correlation, but PDO significant as well with a similar pattern over most of the region (Figure 4.7).

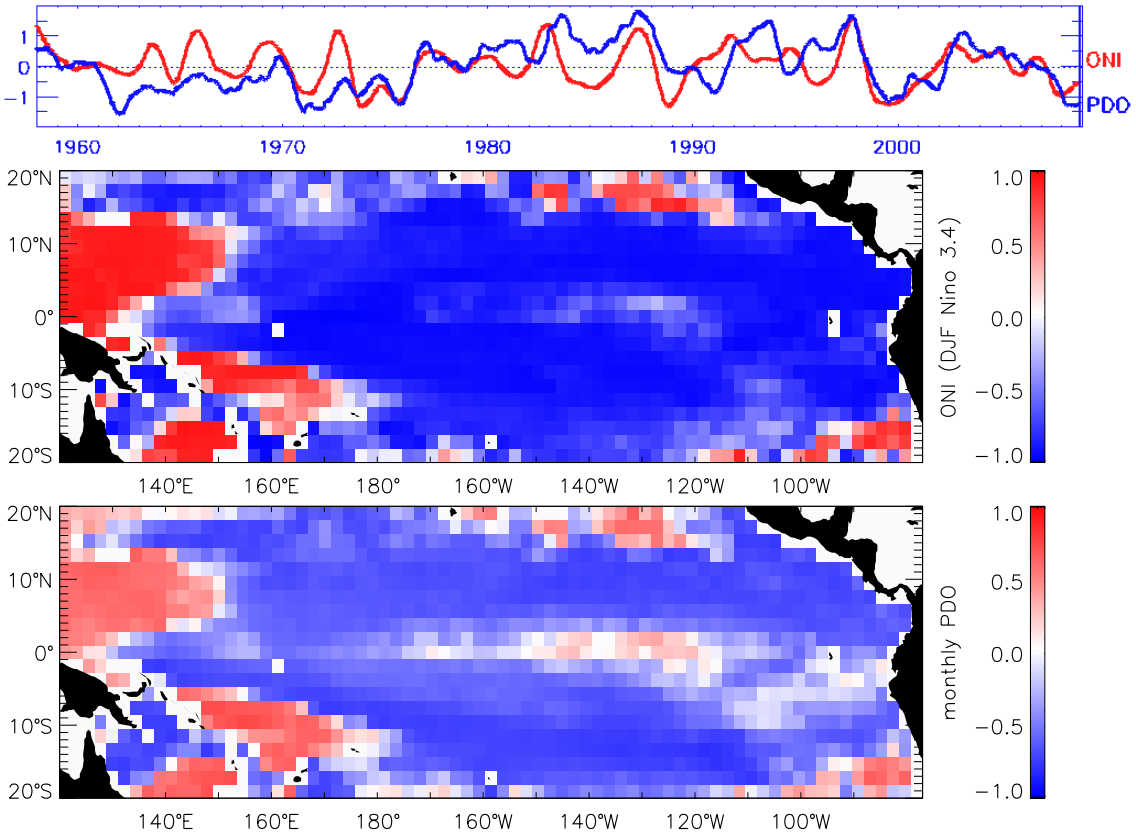

Figure 4.7 Top: monthly variation in ONI from NWS CPC (red) and PDO from UW JISAO (blue) indices smoothed over 12 months. Middle: CHL anomalies correlated with DJF ONI (1958-2008). Bottom: CHL correlated with PDO.

As evidenced in the plot of ONI and PDO (Figure 4.7, top), six years were classified as weak to strong La Niña events (ONI < -0.5) compared to five comparable El Niño events (ONI > 0.5) during the cool PDO phase between 1958-1976. During the warm phase between 1977-1998 there were seven El Niños, two very strong (ONI > 2), and four La Niña events, one of which was strong. The cross-correlation between ONI and CHL indicates an inverse relationship over most of the tropical Pacific with positive correlations in the west Pacific warm pool and around the edges of the tropics (Figure 4.7, middle). The correlation between PDO and CHL also has an inverse relationship over most of the tropical Pacific with a cool or negative PDO phase associated with greater CHL, except in the west Pacific warm pool and around the edges of the tropics and a narrow band along the equator between 120-150^{o}W, coincidentally where the correlation between CHL and ONI is weaker (Figure 4.7, bottom). The pattern between CHL and the PDO index in this narrow equatorial band implies the PDO warm phase is associated with greater CHL while the cool phase is associated with less CHL. Why the warm phase of the PDO is associated with greater CHL in this band and the cool phase with less will be explored later in this chapter. In addition to the PDO strengthening or weakening individual El Niño and La Niña events, modelling studies show these low frequency changes to ENSO have a significant impact on circulation at interdecadal and century timescales [Mann et al., 1995].

The ONI is centered on Niño 3.4 and characterizes all El Niño events, while the distinction between east Pacific (EP) El Niño and central Pacific (CP) El Niño is captured in the EP and CP indices created at UC Irvine (http://www.ess.uci.edu/~yu/2OSC/). The strength of the EP ENSO is determined by removing the Niño 4 SST index and then

performing a regression-EOF analysis on the SST anomalies. The strength of the CP ENSO is calculated likewise except by removing the Niño 1&2 index. These two types of El Niño are distinguished by their evolution and teleconnections: an EP El Niño is more strongly linked to thermocline variation and is generated through tropical atmospheric-oceanic coupling while a CP El Niño is more strongly associated with subtropical Pacific atmospheric forcing which has teleconnections to the Asian-Australian monsoonal variation [Yu and Kao, 2007; Kao and Yu, 2009; Yu et al., 2012].

The CHL reconstruction was cross-correlated against the UC Irvine EP and CP ENSO indices to further compare and contrast the discernable effect central and eastern Pacific extents of El Niño has upon biology (Figure 4.8). Both have significant correlations with CHL: the relationship with the EP indicates an unsurprising spatial distribution in contrast to the CP pattern. An EP El Niño deepens the thermocline across the basin, shutting down the vertical nutrient supply along the equator and resulting in less CHL eastward of 180^{o} as reflected in the negative CHL-EP cross-correlations (Figure 4.8, top). The spatial pattern displayed by the correlation of CHL with the CP El Niño is remarkably similar to the correlations with the PDO, most notably for the narrow band along the equator between 110-130^{o}W (Figure 4.8, bottom). The CP index correlates more closely with ONI ($r = 0.75$) than with PDO ($r = 0.46$). A CP El Niño extends to Niño 4 but not farther eastward than 180^{o}, interrupting the upwelling and advection of nutrients and causing negative CHL anomalies in the central Pacific, but allowing equatorial upwelling to continue to provide nutrients to surface waters in the east. The distinctions in the CHL patterns between PDO warm and cool regimes are explored in greater depth next.

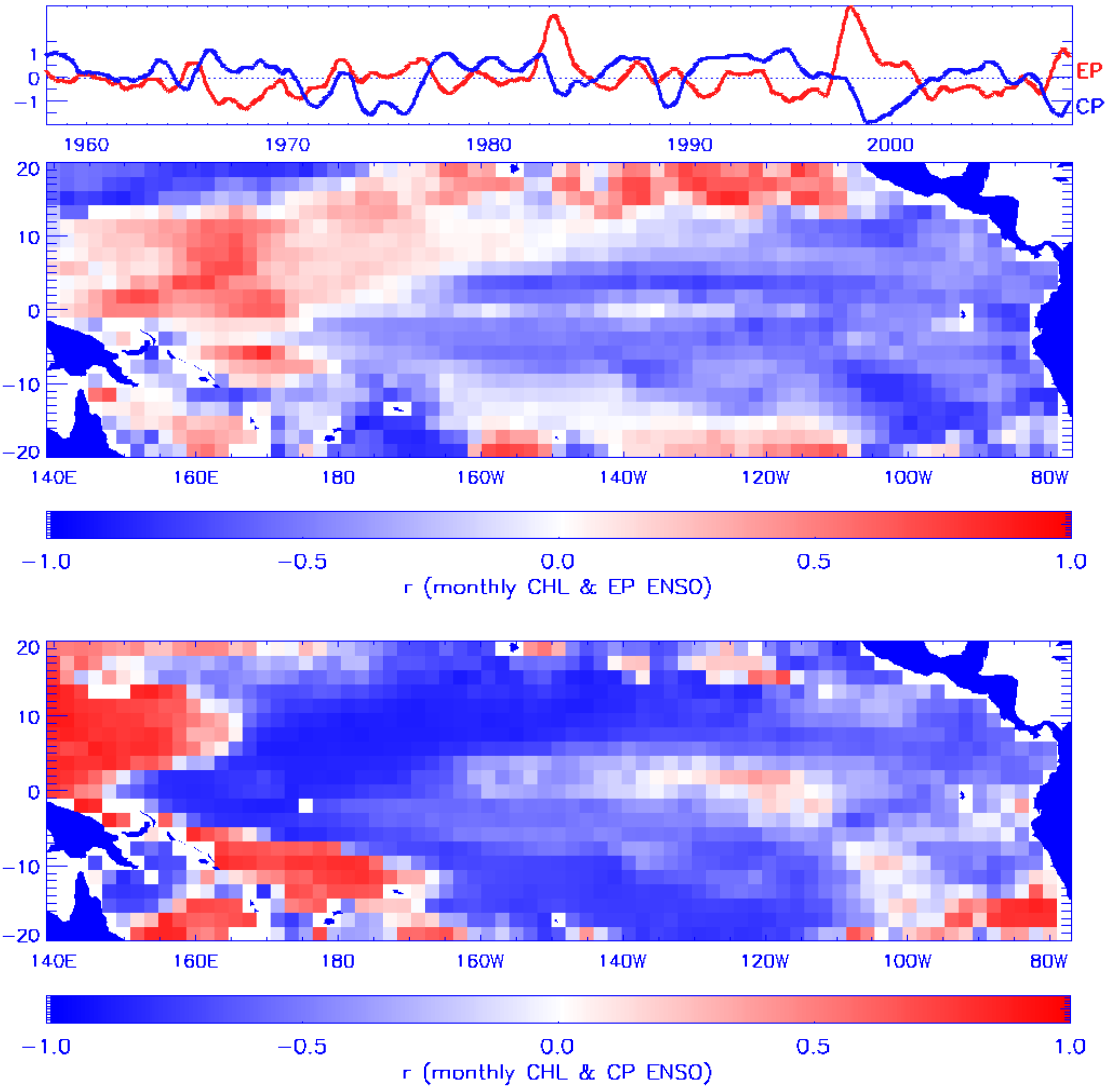

Figure 4.8 **Top: monthly variation in EP (red) and CP (blue) indices smoothed over 12 months. Middle: CHL correlated with EP index (1958-2008). Bottom: CHL correlated with CP index.**

The reconstruction starts and ends during a cool phase of the PDO. Between 1977 through the mid 1990s, a warm phase of the PDO dominated. Reconstructed CHL anomalies were averaged from 1958-1976 during the cool phase (Figure 4.9) and from 1977-1995 during the warm phase (Figure 4.10). The PDO cool phase had higher CHL over the tropical Pacific in a horseshoe pattern, with lower CHL along the equator east of the dateline, in the west Pacific warm pool and toward the edges of the tropics.

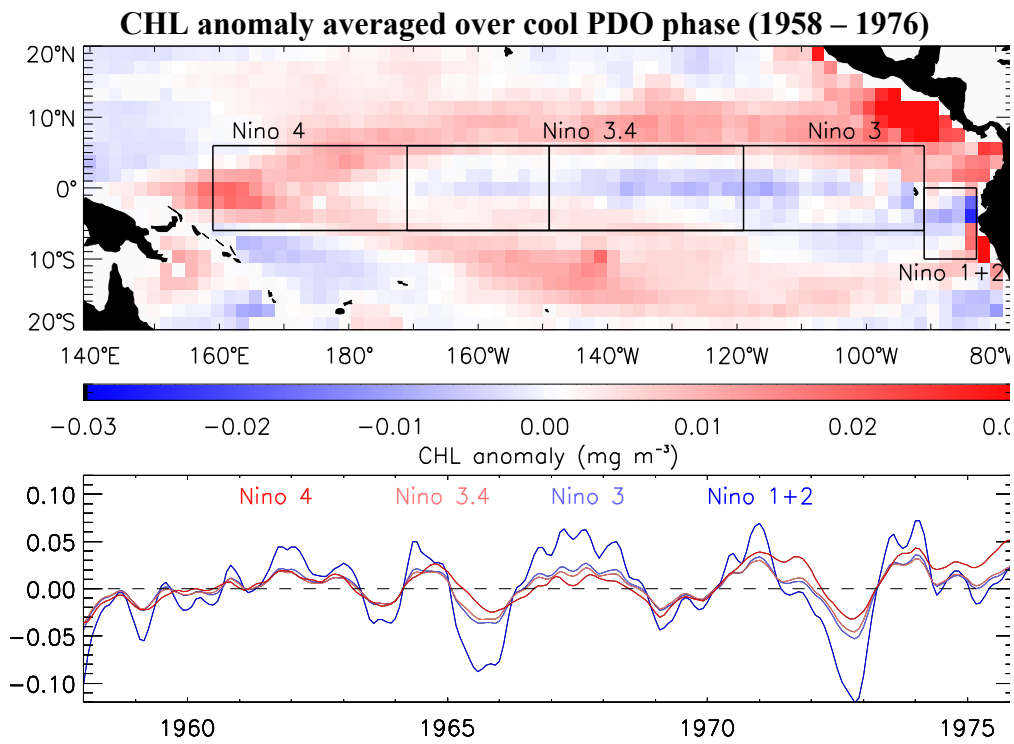

Figure 4.9 Cool phase, 1958-1976, average CHL anomalies with Niño regions delineated (top); time series averaged over Niño regions (bottom).

Positive CHL anomalies dominate most of the region during the cool phase, but negative anomalies dominate along the equator, with the exception of the west half of Niño 4 (Figure 4.9, top). Time series averaged over each Niño area (Figure 4.9, bottom) indicate large negative anomalies associated with the strong El Niños of 1957/58, 1965/66, and 1972/73. La Niña years had the opposite, not equal, effect as slight positive CHL anomalies in 1964/65, 1966-1968, 1970/71, 1973/74. Niño 1&2 had the largest variations. Equatorial Niño areas had damped amplitudes by comparison. El Niño was stronger in Niño 3 than Niño 4 during 1965/66, contributing to the lower CHL average eastward of the dateline. The La Niña events in 1970/71 and 1973/74 were stronger in Niño 4 than Niño 3, supporting higher mean CHL values in Niño 4 over the cool era.

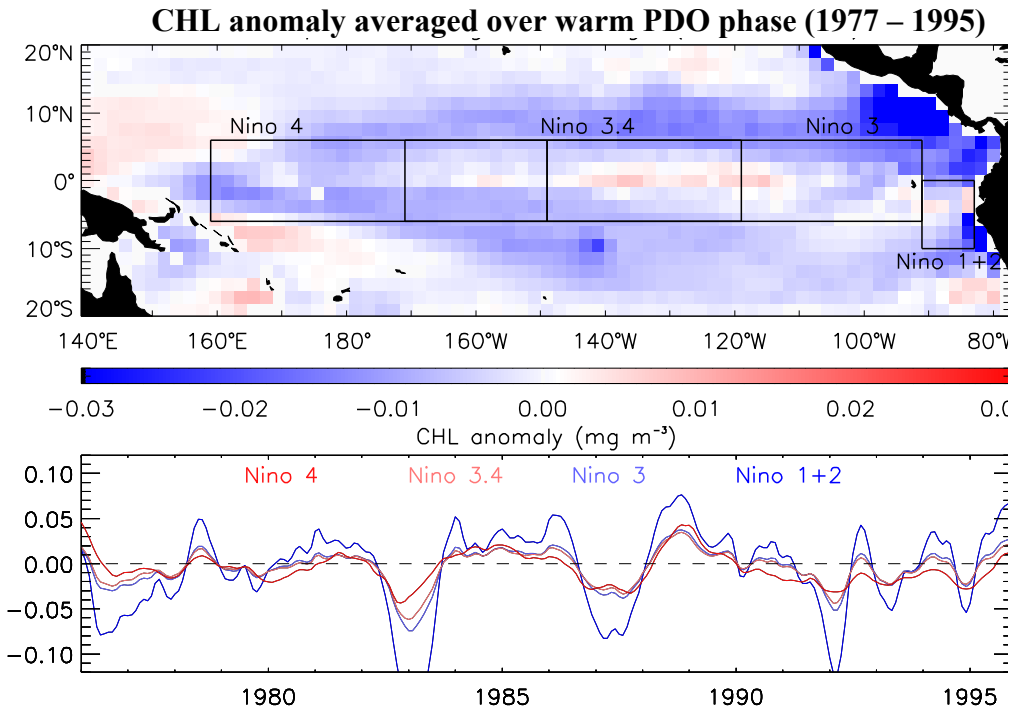

Figure 4.10 **Warm phase, 1977-1995, average CHL anomalies with Niño regions delineated (top); time series averaged over Niño regions (bottom).**

The warm phase of the PDO experienced the reverse pattern (Figure 4.10), averaged here between 1977-1995 to avoid the extreme El Niño of 1997/98. Although negative CHL anomalies predominated in most of the tropical Pacific, there was a band of positive anomalies along the equator between 110-160°W (Figure 4.10, top). The corresponding time series (Figure 4.10, bottom) demonstrate the largest decreases during strong El Niño years: 1982/83, 1991/92 and moderate in 1986/87 and 1994/95. The large La Niña in 1988/89 corresponded with increases. Again, greatest magnitudes were for coastal Niño 1&2, with equatorial Niño areas having damped amplitudes. Plotting CHL anomalies by month and averaging over warm and cool phases of the PDO confirms that the cool era had more CHL overall with different annual cycles for each PDO phase and Niño area despite large interannual scatter (Figure 4.11).

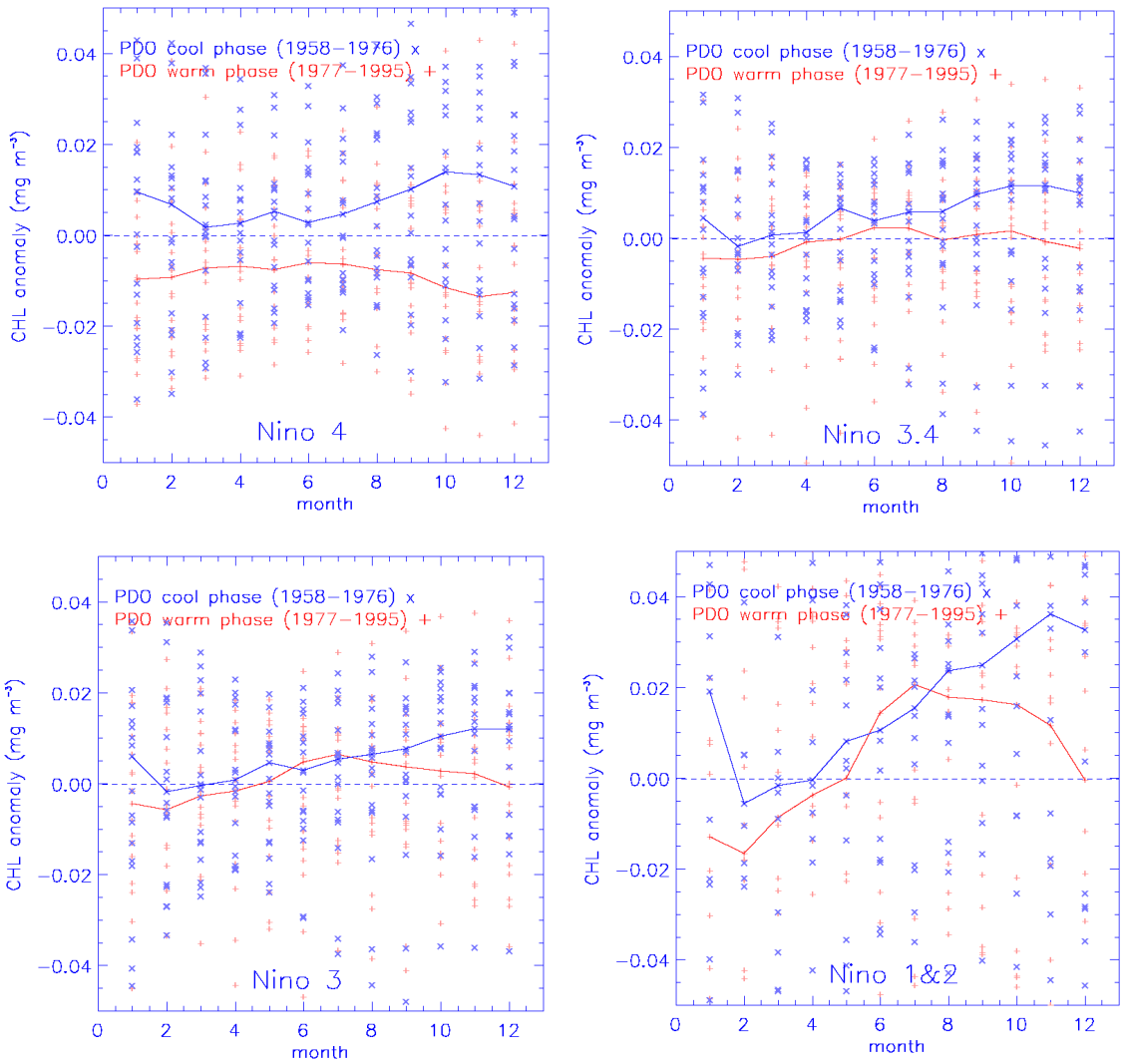

Figure 4.11 **Niño 4 (top left), Niño 3.4 (top right), Niño 3 (bottom left), Niño 1&2 (bottom right). CHL anomaly (mg m^{-3}) averaged for cool (——) and warm (——) PDO phases and each month (x and +). Months range from 0 (January) to 12 (December).**

During the PDO cool phase, average seasonal cycles in all areas had sharp decreases in CHL roughly between October and February with gradually increasing CHL during the rest of the year. The warm regime had increasing CHL between February and July, declining the rest of the year. In the difference between the two PDO eras, warm

phase minus cool shown in Figure 4.12, it is apparent that the warm phase of the PDO generally experienced less primary production except for the narrow band of positive anomalies along the equator between 110-160°W and in the west Pacific warm pool. Along the equator east of the dateline (Niño 3 and Niño 3.4), positive CHL anomalies were greater and negative anomalies smaller during the warm regime. West of the dateline (Niño 4) and away from the equator, the differences were reversed: the cool regime had positive CHL anomalies. Physical forcing differences between the PDO phases that could impact timing and location of phytoplankton blooms are explored next.

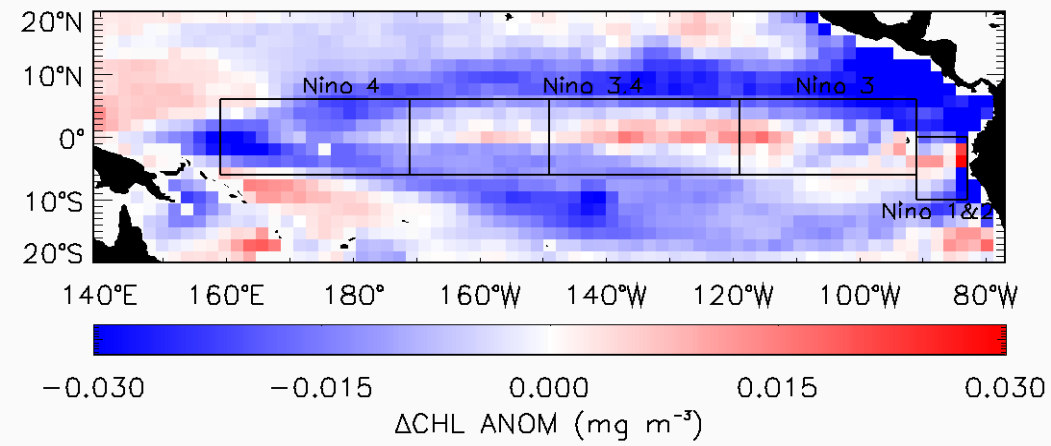

Figure 4.12 CHL differences over a PDO cycle: warm phase (1977 – 1995) minus earlier cool phase (1958 – 1976). Niño regions are delineated.

4.3.1 Physical forcing of biology evidenced at the ocean's surface

The physical variables were averaged for the same cool and warm PDO time periods and then differenced, warm minus cool, shown in Figure 4.13. SST differences confirm the warmer warm phase across the tropical Pacific, except for negligible differences in the west away from the equator (Figure 4.13, top). SST differences do not

have the same spatial distribution as CHL because the CHL shift only occurs in places where the nutricline rises into the euphotic zone in the equatorial upwelling zone. More importantly, the difference map demonstrates that using SST as a proxy within a CHL reconstruction does not merely reflect a change in SST. Likewise, SSH differences also do not have the same distribution as CHL, as well as being different from SST. For the Nino areas along the equator, SSH differences between the cool to the warm regime reflect that SSH decreased in the west Pacific and increased in the east Pacific (Figure 4.13, second). Away from the equator, SSH increased with greatest positive differences near 10^{o}N. Winds variables also indicate differences between regimes that may have contributed to the CHL patterns. Zonal wind stress (τ_X) had maximum increases between regimes along the equator between 110-140^{o}W where winds were anomalously easterly during the cool phase and westerly during the warm (Figure 4.13, third). Meridional wind stress (τ_Y, not shown) indicates a southerly anomaly between 0-5^{o}N during the warm phase that would tend to support equatorial convergence, with the opposite pattern supporting divergence during the cool phase. Wind stress magnitudes indicate that winds were stronger in the central Pacific during the warm phase, but stronger in the east Pacific during the cool, not shown here. Wind stress curl ($\nabla \times \tau$) was greater or more cyclonic toward the equator during the warm phase (Figure 4.13, bottom). Winds transfer momentum, cause vertical mixing of the ocean's surface layer, and may induce upwelling in combination with the Coriolis force and geography, but winds only lead to phytoplankton blooms if they act together with a nutricline shallow enough to lift nutrients into the euphotic zone.

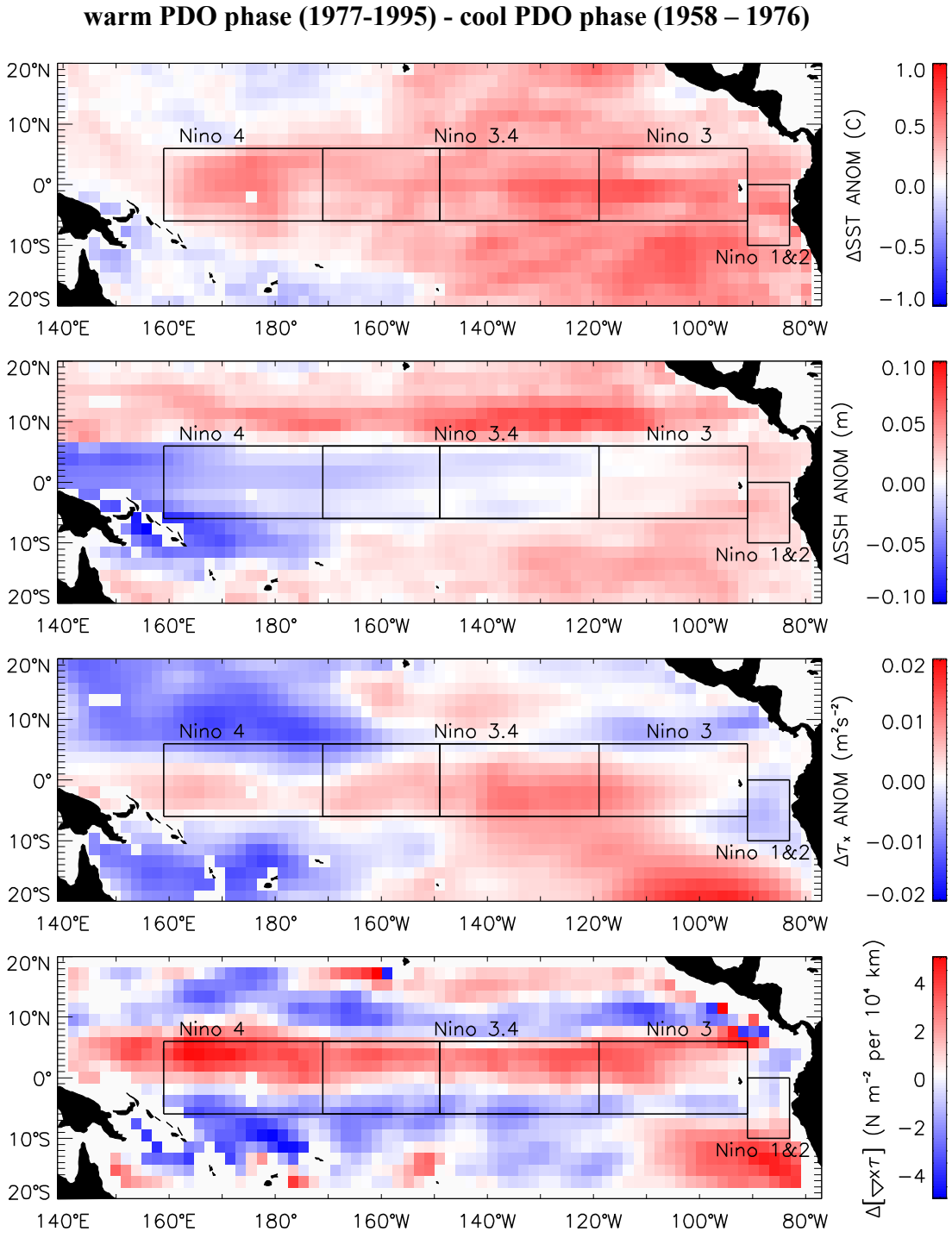

Figure 4.13 **Average differences in physical variables between the warm and cool regimes: SST (top), SSH (2nd from top), zonal wind stress (3rd from top), wind stress curl (bottom). Niño areas are delineated. Missing data are shaded white.**

A zonal band of greater wind stress curl between the equator and 10°N coincides with the North Equatorial Current (NEC) and Counter Current (NECC). Here there were stronger positive wind stress curl anomalies during the warm phase and more negative wind stress curl anomalies in the cool phase. Wind stress triggers westward propagating long baroclinic Rossby waves around 5-6°N that influence the intensity of the NECC and cause a southward shift during El Niño events with up to 25% larger transport [Johnson et al., 2002; Hsin and Qiu, 2012]. Since the NECC is a surface current, it has an indirect effect on CHL. Away from the equator, positive wind stress curl can cause Ekman pumping (w_{ek}), as shown in equation (1) from section 2.4.1, which could raise or lower the nutricline into or out of the euphotic zone. The cool phase experienced downward Ekman pumping within 5° of the equator (Figure 4.14), depressing the nutricline there.

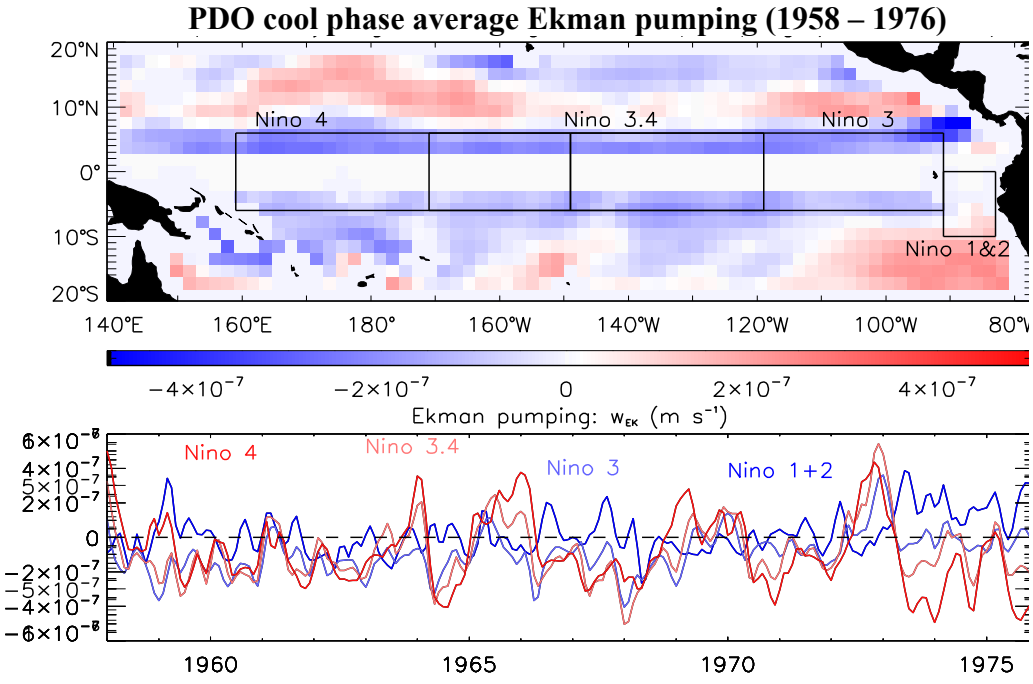

Figure 4.14 Ekman pumping averaged over the PDO cool phase from 1958-1976 (top), with time series averaged over each Nino area as indicated (bottom): positive is upward, negative is downward.

The warm regime experienced upward Ekman suction within 5° of the equator (Figure 4.15), which could have lifted nutrients into the euphotic zone if the nutricline was shallow enough. CHL anomalies might be correlated to Ekman pumping, with less CHL during downward pumping and more during upward suction, yet only the anomalous band of CHL along the equator is consistent with this generalization and it would be a neighborhood effect since Ekman pumping depends upon the Coriolis force that goes to zero on the equator. CHL in the rest of the region not being consistent with average Ekman pumping could indicate that the nutricline was too deep in those areas for Ekman pumping to supply nutrients to fuel phytoplankton blooms at the surface. Off the equator, a combination of waves, eddies and horizontal advection may have supplied nutrients to the cold tongue.

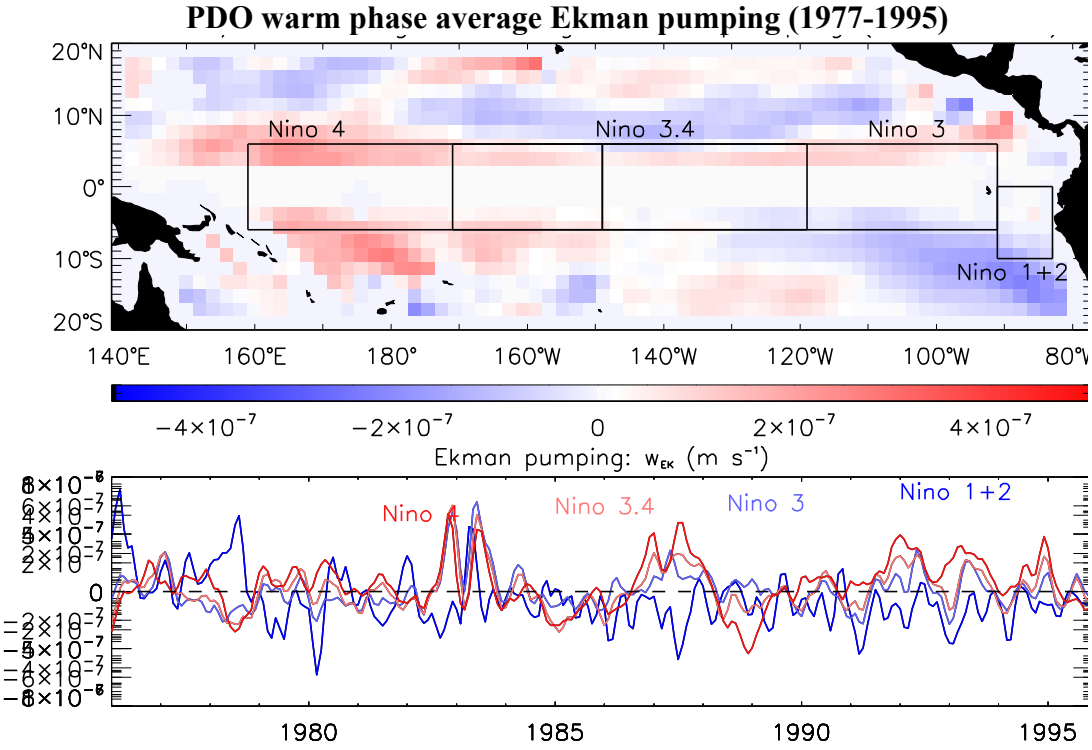

Figure 4.15 Ekman pumping averaged over the PDO warm phase from 1977-1995 (top), with time series averaged over each Nino area as indicated (bottom): positive is upward, negative is downward.

To further explore causes for the anomalous CHL band along the equator in Niño 3 and 3.4, inner areas were defined as +/- 3^{o} of the equator and outer areas were defined from 3-10^{o}N. As shown in Figure 4.16, El Niño events during the cool PDO phase had CHL decrease in the inner equatorial areas preceding the outer areas. During La Niña, however, CHL increased simultaneously for all areas, but the CHL anomalies away from the equator continued to increase longer to a greater magnitude while CHL anomalies in the inner areas decreased sooner to lower values. During the warm PDO regime, the areas away from the equator had larger negative CHL anomalies while CHL in the inner equatorial area recovered faster from El Niño and experienced positive anomalies during several years that the outer areas did not, namely in 1984, 1990, 1992, 1994 (Figure 4.16, bottom panel). Generally, CHL anomalies in the inner equatorial areas seemed to experience the less common ENSO phase more intensely (e.g. El Niño during the cool PDO era, La Niña during the warm PDO era), while CHL anomalies away from the equator were consistent with prevailing PDO and ENSO states.

CHL anomalies were cross-correlated against physical variables for the inner and outer areas within Niño 3 and 3.4 during the warm and cool PDO regimes to determine physical mechanisms impacting biology. Variables were cross-correlated for up to six month leads and lags; quarterly averages were compared. Only zero lag, plus and minus one quarter are included in Table 4.1 as no correlations were significant beyond that, with $r > 0.38$ the highly statistically significant threshold ($p<0.001$). SST had the strongest significant correspondence with CHL everywhere, inverse for all areas during both regimes. SSH was significantly correlated to CHL in Niño 3, declining slightly in Niño 3.4 where correlations between wind and CHL increase.

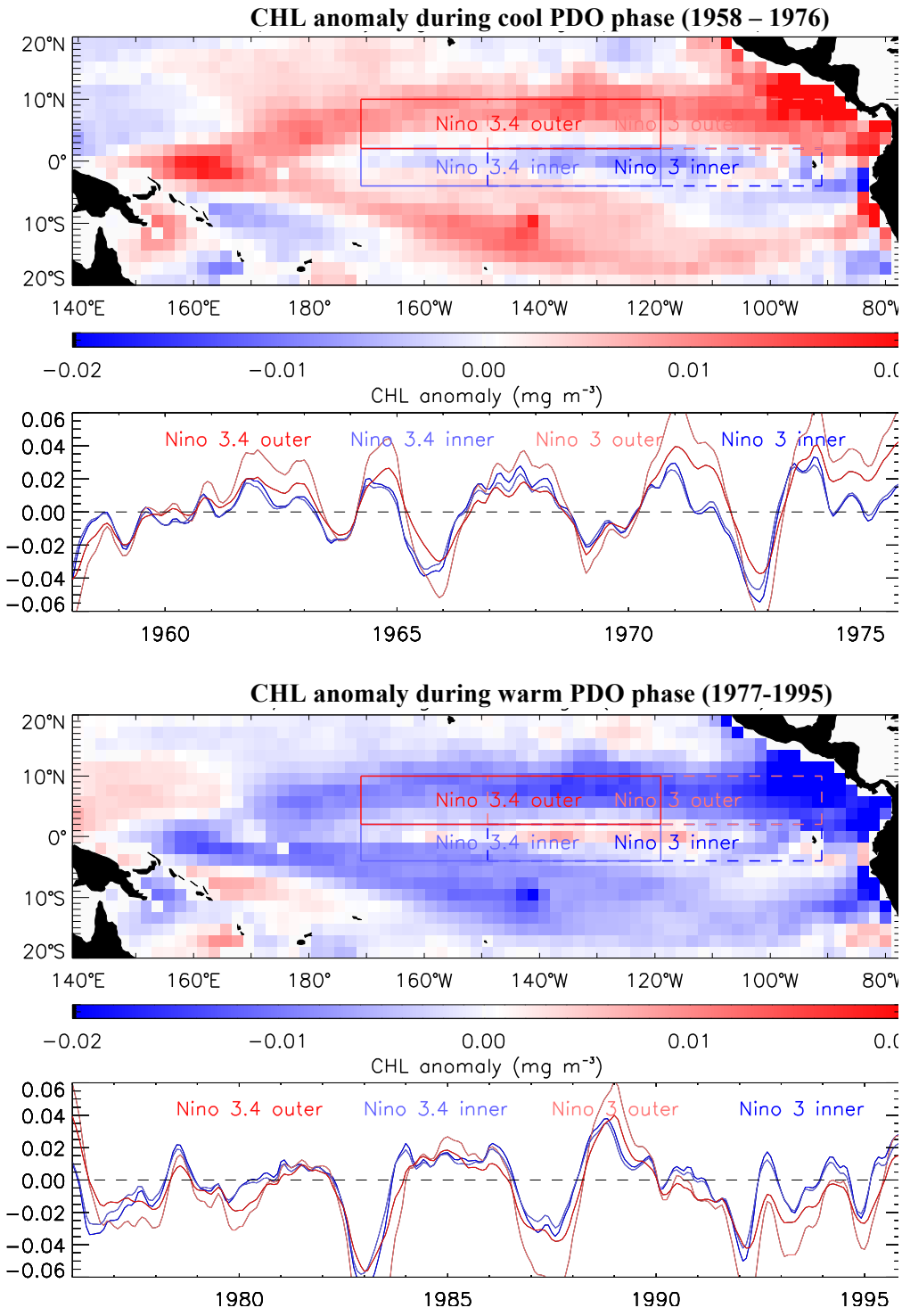

Figure 4.16 **Niño 3 and 3.4 divided into inner (3°S-3°N) and outer (3-10°N) areas: CHL averaged over the cool PDO regime from 1958-1976 (top) and corresponding time series averaged over the delineated areas (2nd panel). CHL averaged over the warm PDO regime from 1977-1995 (3rd panel) and corresponding time series for the delineated areas (bottom).**

Easterly trade winds support equatorial upwelling that can bring deep nutrient-rich water toward the surface to fuel phytoplankton blooms. Over both PDO eras, zonal wind stress (τ_X) anomaly correlations with CHL were inverse everywhere, except the Niño 3 outer area (Table 4.1). Zonal wind stress had the largest, inverse correlation to CHL anomalies for the inner area of Niño 3.4 ($r \sim -0.7$), second only to the inverse correlation between SST and CHL there ($r \sim -0.9$). Westerly τ_x anomalies are generally associated with El Niño events, particularly in Niño 4. Strong El Niños during both regimes were associated with positive τ_x anomalies and negative CHL anomalies. Positive τ_X anomalies over the inner areas would result in less surface divergence and equatorial upwelling, contributing to negative CHL anomalies. Indeed, during both regimes, the correlations between zonal wind stress and CHL were generally inverse, yet the averages over both PDO eras shows that during the warm, both τ_X and CHL were positive for the band along the equator; during the cool, they were negative.

The PDO warm regime experienced more westerly wind anomalies more frequently and to a greater magnitude, as shown in Figure 4.17. The warm era had four strong El Niño events; the cool regime had three, contributing to the average τ_X differences. Stronger, more frequent El Niños during the PDO warm era are consistent with the τ_x and SST PDO anomalies, but inconsistent with CHL PDO anomalies in the narrow band along the equator. To explore whether the anomalous CHL average along the equator could be due to a stronger response to El Niño during the cool regime and to La Niña during the warm one, quarterly and annual averages were compared over both eras to try to understand the variability in the relationship between zonal winds and CHL.

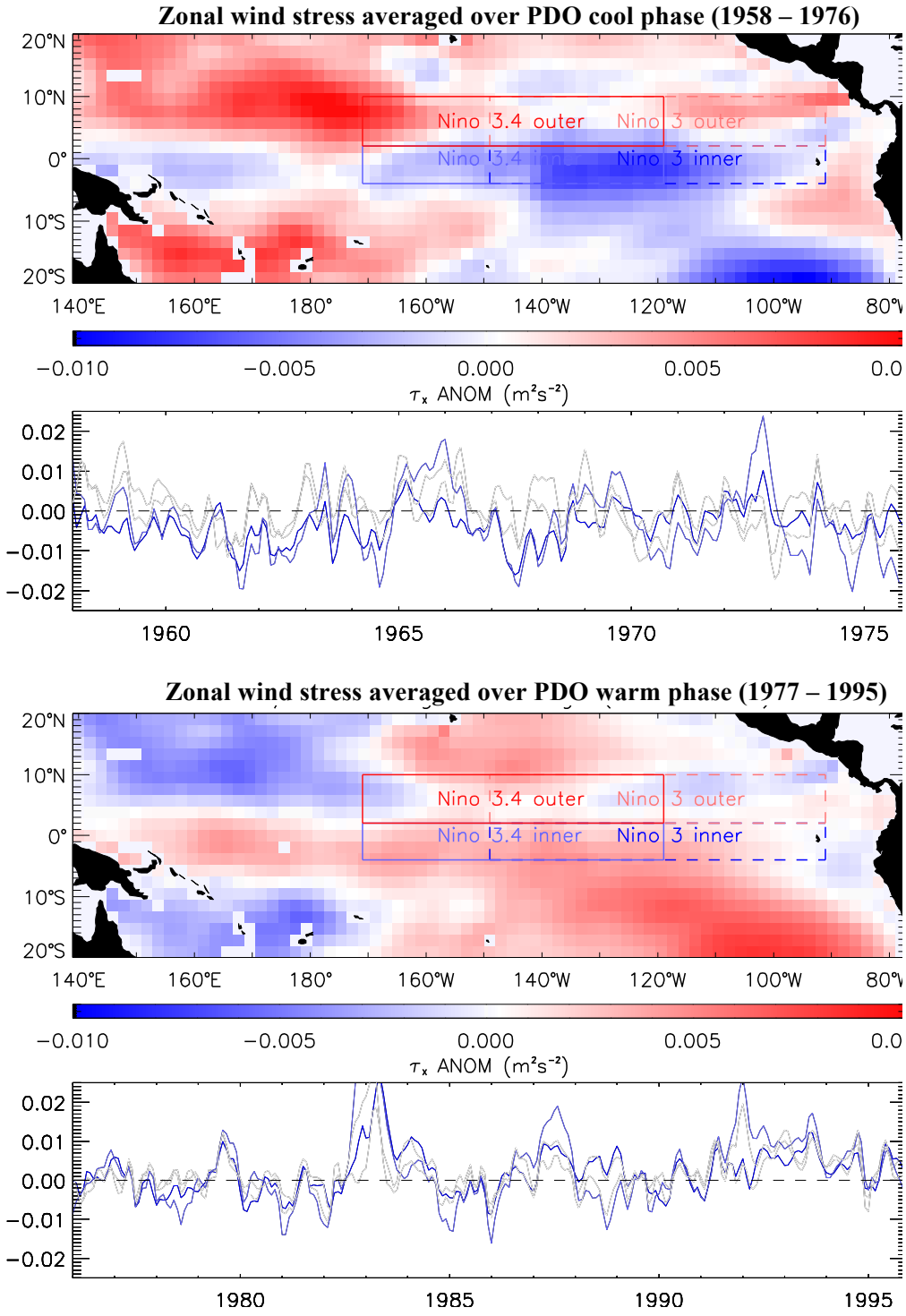

Figure 4.17 **Zonal wind stress anomalies for same time periods and areas as Figure 4.17: Niño 3 and 3.4 divided into inner (3°S-3°N) and outer (3°-10°N): Cool PDO era average anomalies (top) and time series over inner areas (2nd panel). Warm PDO era average (3rd panel) and times series (bottom).**

The variability in zonal wind stress quarterly and annual averages were compared to CHL anomalies over both eras for the narrow equatorial band with the opposite CHL response. Zonal wind variability (e.g. standard deviations or maximums normalized by the mean) had no discernable effect upon CHL response. Quarterly zonal wind anomalies plotted against quarterly CHL anomalies (Figure 4.18) have an inverse relationship during both eras, with more scatter during the warm era due to greater variability in zonal winds during the more frequent El Niño events then.

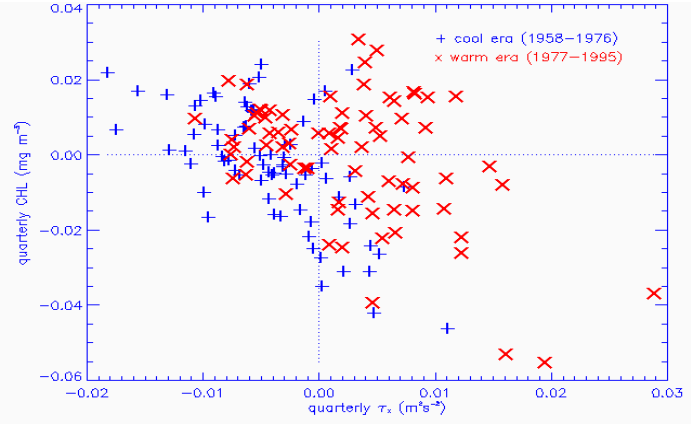

Figure 4.18 Quarterly zonal wind stress anomalies and CHL anomalies for the equatorial band (3°S-3°N, 110-150°W) during the cool PDO era (blue +) and warm PDO era (red x).

Whether the interannual El Niño pattern dominated the long-term average over the PDO regimes and whether the type of El Niño events, east or central, contributed to the anomalous equatorial band is now explored. Monthly zonal wind anomalies were plotted against the monthly CHL anomalies for the east Pacific El Niño events (EP > 1) and central Pacific El Niño events (CP > 1). While zonal winds were generally positive and CHL negative during all El Ninos, there was no consistent relationship between zonal winds and CHL for the central Pacific El Ninos, likely because central Pacific El Niños are associated with variation in the northern subtropical high pressure [Yu et al., 2012].

Conversely, east Pacific El Niños had some outlying negative zonal winds during the decay of the 1982/83 El Nino that lead to an overall positive correlation between zonal winds and CHL, inconsistent with the inverse correlation in the time series over the warm and cool eras and unlikely to indicate any meaningful tendency. Overall, there did not appear to be a consistent difference in zonal winds and CHL between east Pacific or central Pacific types of El Niño. Comparing observations of pycnocline flow over the past 50 years, McPhaden and Zhang [2002] found a 25% decrease in equatorial upwelling between the 1970s and 1990s as the PDO oscillated from cool to warm phase. Whether this CHL pattern is a manifestation of the slow down in meridional overturning circulation reported between PDO eras is investigated next with a focus on how such subsurface forcing could impact primary production at the ocean's surface.

4.3.2 Subsurface physical forcing and its impact on biology

In addition to responding to changes in wind-driven mixing, surface CHL anomalies may respond to deeper ocean dynamics at low frequencies. As introduced in section 2.4.2, the waters of the tropical Pacific contain two distinct nutrient regions due to the prevailing winds: macronutrient-rich waters in the eastern cold tongue where easterly winds predominate, and the oligotrophic surface waters of the warm pool in the west [Le Borgne et al., 2002]. In the nitrate climatology from the World Ocean Atlas, most nitrate at 100m is found in upwelling regions off the west coast of Central and South America and near the equator (Figure 4.19, top); the nutricline slopes west-east, as shown in the zonal cross-section (Figure 4.19, middle); upwelling along the equator and Costa Rica Dome to the north is seen in the meridional transect along 110^{o}W (Figure 4.19, bottom).

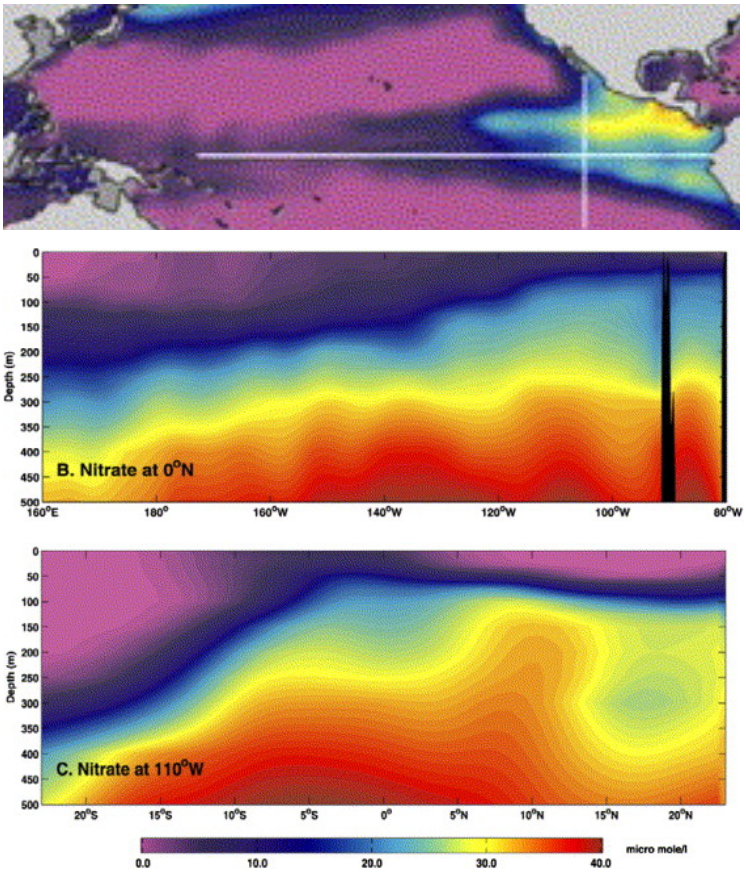

Figure 4.19 **Nitrate climatology, World Ocean Atlas 2009. Top: Nitrate map at 100m with zonal, meridional transects (white lines). Middle: Nitrate profile at 0°N. Bottom: Nitrate profile at 110°W.**

Despite an abundance of macronutrients, the east tropical Pacific is an iron-limited, high-nutrient, low-chlorophyll region. Zonal transects have confirmed that iron enters the eastward-flowing Equatorial Undercurrent (EUC, recall Figure 2.12) from sedimentary sources on the coastal Papua New Guinea slope in the west [Gordon et al., 1997; Mackey et al., 2002a; 2002b; Slemons et al., 2010]. The 1997/98 El Niño CHL minimum in the east Pacific was linked to EUC cessation; its recommencement preceded the end of El Niño by several months [Wilson and Adamec, 2001]. Subsurface current structure has been used to infer EUC iron transport [Ryan et al., 2002]. The EUC was observed by Acoustic Doppler Current Profilers (ADCP) from the Tropical Atmospheric-

Ocean/Triton Array nearly continuously between 1990 and 2010, revealing its intensity and vertical migration. The TAO/Triton ADCP record at 140°W began after the PDO cool phase that ended in 1976, but it shows some differences in the EUC between the warm phase that ended around 1998 and the cool phase that followed (Figure 4.20).

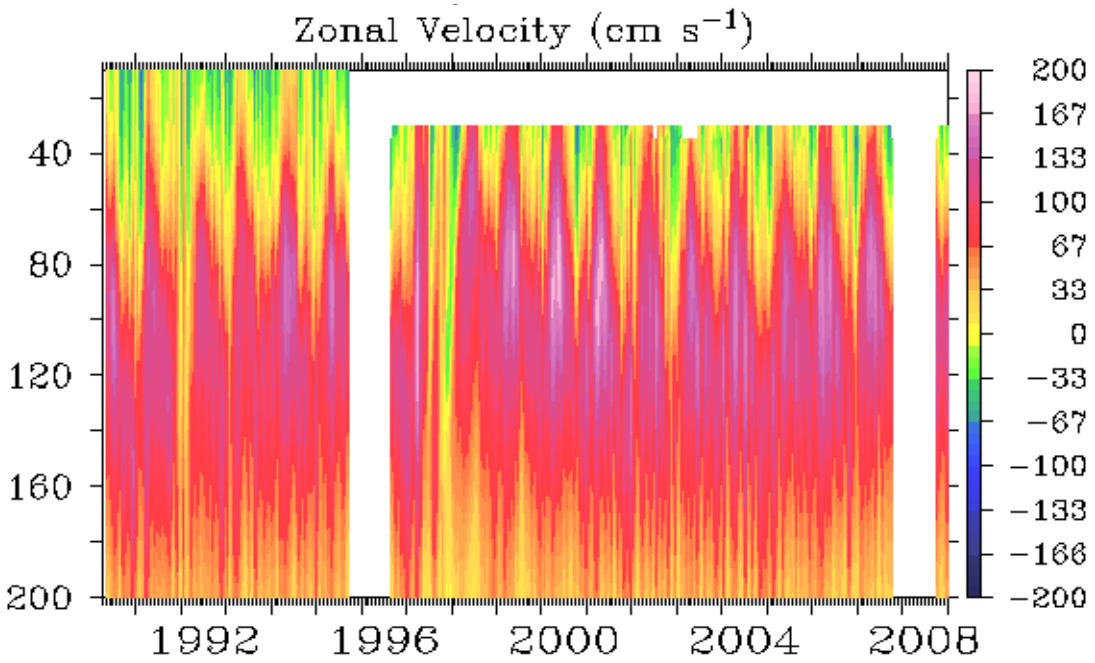

Figure 4.20 **Zonal velocity (cm s^{-1}) from the TAO/Triton equatorial ADCP at 140°W from April, 1990 through 2008 (bottom axis) with the depth (left axis) and strength of the EUC (color bar). Positive velocities are eastward. Figure courtesy of TAO Project Office/PMEL/NOAA.**

The eastward EUC tended to be deeper and weaker during the warm phase (1990-1998) and shallower and stronger during the cool (1999-2010). The ADCP readings at 165°E, 170°W, and 110°W confirm this shoaling tendency by the EUC, which at first seems consistent with the finding that there was a slowdown in the upper tropical Pacific circulation during the PDO warm phase [McPhaden and Zhang, 2002].

Recent modeling studies predict a strengthening of the EUC in response to anthropogenic forcing [Luo et al., 2009; Sen Gupta et al., 2012]. A long-term strengthening of the EUC since the mid-1800s has been observed in the SODA 2.2.6 assimilation of ocean general circulation models with quality-controlled observations of SST [Drenkard and Karnauskas, 2014]. Inspection of SODA zonal velocities since 1958 show that between 120-150°W the EUC shoaled most during La Niña events (~30m) compared to its background state, with a long-term trend of a ~20m shallower EUC over the 50 years of the reconstruction, as shown in Figure 4.21 for 130°W.

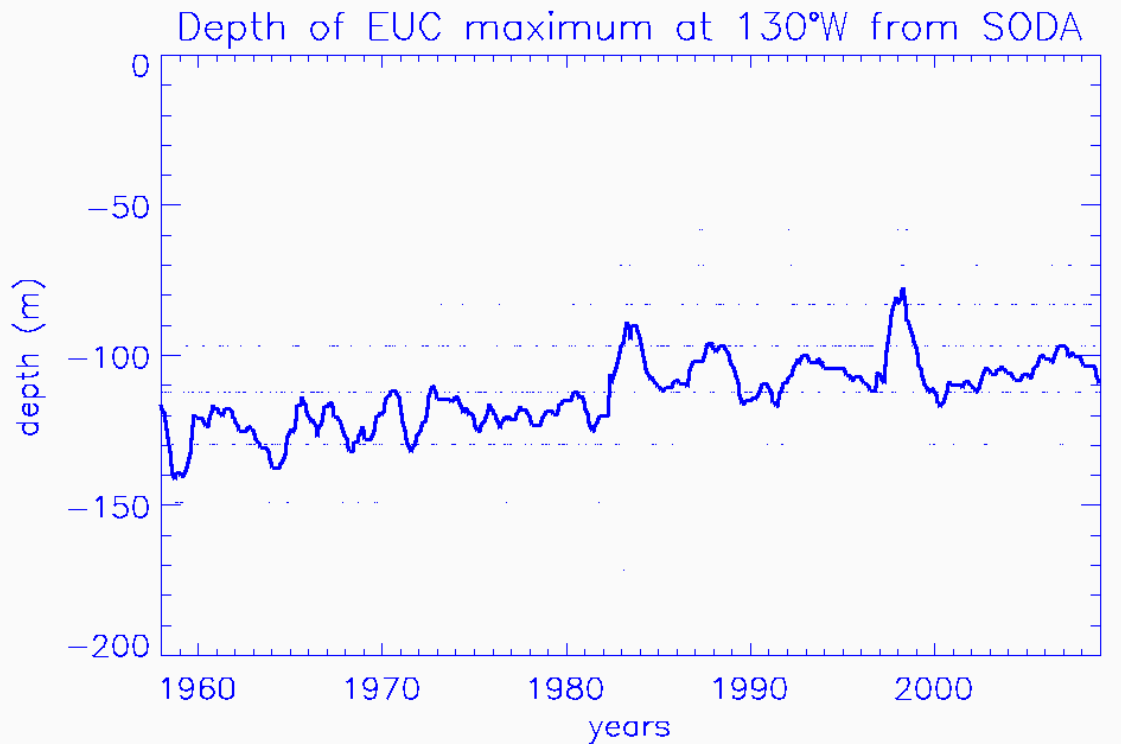

Figure 4.21 **Depth of the EUC (m) in SODA at 130°W between 1958 and 2008. Individual data points are shown (·) and plotted as a line that was smoothed over 12 months.**

While variations in the EUC depth due to ENSO are apparent, a prominent long-term shoaling trend obscures decadal patterns, as shown in Figure 4.22. Between 120-

150°W, the average EUC depth during the PDO cool phase (1958-1976) was 120m; then 110m during the PDO warm phase (1977-1995), and 102m during the beginning of the later cool phase included in this study (1998-2008). The average shoaling trend in EUC depth between 1958 and 2008 was more than 0.5 m yr^{-1}. A shoaling trend by the EUC is evident at varying magnitudes over most of the basin from a maximum of ~1m yr^{-1} near 150°E while east of 100°W the maximum zonal velocity had a deepening trend (Figure 4.22). The impact of these changes in the EUC speed and depth on biology over the long-term is explored next.

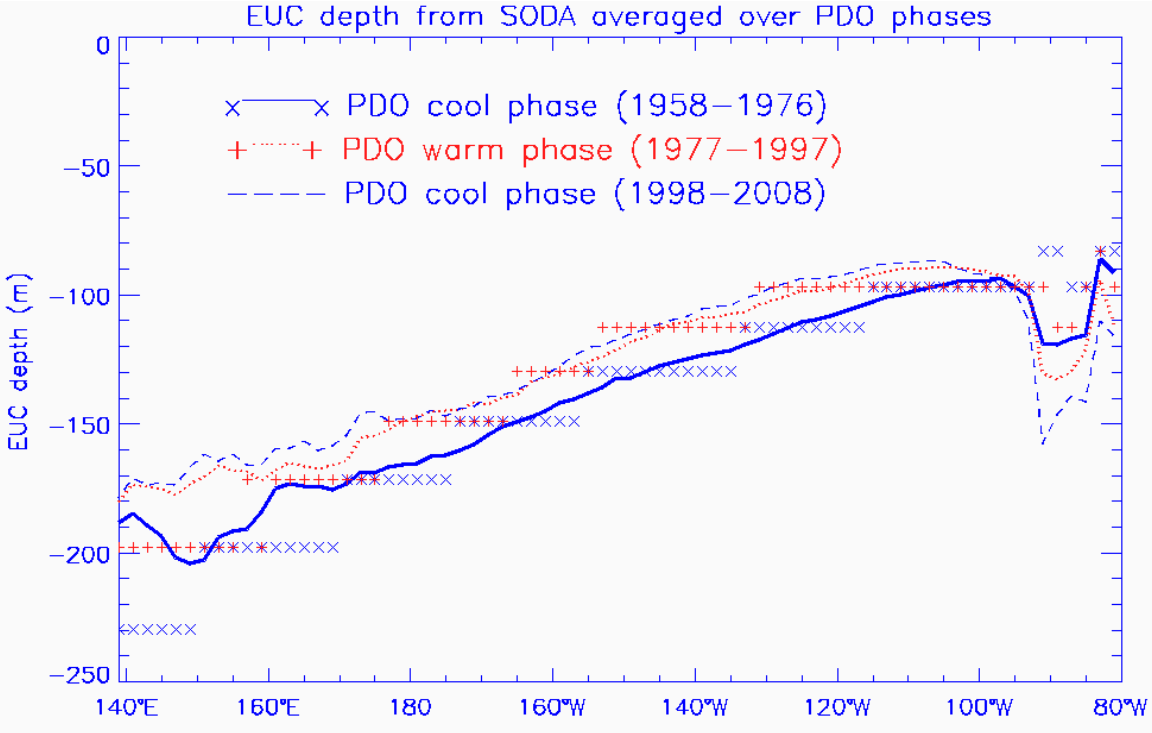

Figure 4.22 **EUC depth averaged over the PDO phases: lines indicate average values at each longitude over the PDO eras; median values are included for the cool (x) and warm (+) phases.**

The conceptual model of the equatorial Pacific is that the easterly trade winds typical along the equator cause the surface layer to pile up in the west, forcing the

pycnocline and EUC to slope downward there. A pycnocline that is deep in the west and shallow in the east induces a geostrophically balanced meridional inflow towards the equator in both hemispheres [McCreary and Lu, 1994]. Anomalous westerly winds near the equator cause the pycnocline tilt to relax, resulting in deepening in the east, shoaling in the west, and reduced equatorward geostrophic transport. McPhaden and Zhang [2002] found that the depth of the pycnocline became shallower by about 10–30m west of 160°W and deeper by about 10–30m between 120°–150°W in the 1990s relative to the 1970s. Consistent with the findings of McPhaden and Zhang [2002], from the cool phase of the PDO (1958-1976) to the warm (1977-1995), the EUC became 20m shallower west of 160°W and about 10m deeper east of 100^{o}W, indicating a steeper pycnocline tilt during the cool phase. Between 120-150^{o}W, the EUC was 20m shallower during the warm phase compared to the cool, a significant difference because it raised the EUC toward the euphotic zone, extending to a depth of approximately 100m [Chavez et al., 1999]. During the first cool phase, the EUC did not rise to that depth until east of 120^{o}W. This EUC depth difference could explain the anomalous CHL band between 120-150^{o}W during both eras: during the warm era, the EUC was shallower and supplied iron to the euphotic zone in this band. Furthermore, during the last decade of the reconstruction, the EUC was even shallower and the entire cold tongue displayed high CHL with no anomalous equatorial band between 120-150^{o}W (recall Figure 4.6), consistent with the EUC supplying iron to the otherwise high-nutrient, low-chlorophyll cold tongue and influencing low frequency CHL variability. While this theory is logical, it is only as good as the SODA reanalysis upon which it is based. Assuming that SODA zonal velocities are representative of the real ocean, this explanation is plausible. For further corroboration,

CHL was compared to surface layer iron from the fully-coupled physical biogeochemical model [Wang et al., 2008] detailed in section 3.2.4. The location of recurring anomalous CHL is delineated by the blue box in Figure 4.23 and plotted against corresponding iron integrated over the euphotic zone to the 1% light level, defined as 60m by Ryan et al. [2002]. Strong correlation (r~0.7) confirms that higher amounts of iron covary with greater CHL. By contrast, in the west the EUC is far below the euphotic zone; there is negligible surface layer iron and it has a weak correlation with CHL (r~0.5).

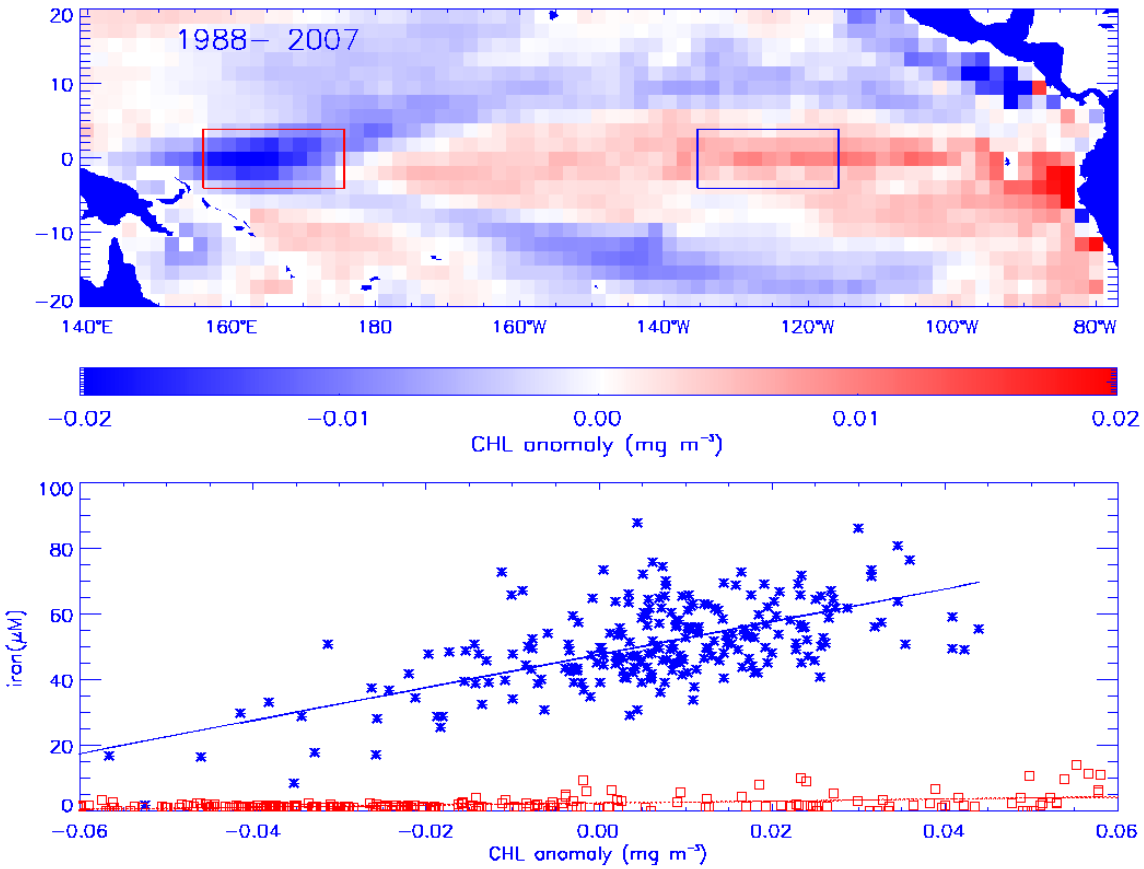

Figure 4.23 **Reconstructed CHL anomalies averaged 1988-2008 (top) with blue box on the equator corresponding to the blue (x) in the scatter plot (below) of the monthly CHL against modeled iron [Wang et al., 2008] integrated over the euphotic zone (r~0.7). The red box on the equator in the west corresponds to the red (□) in the scatter plot and smaller cross-correlation (r~0.5).**

Other studies have linked phytoplankton bloom episodes to the efficiency of the vertical processes that lift nutrients into the euphotic zone, namely the shoaling of the EUC and the strength of easterly trade winds [Wilson and Adamec, 2001; Ryan et al., 2002; Messie et al., 2006; Gierach et al., 2013]. The shoaling of the EUC is here shown to be the primary driver of low frequency patterns in phytoplankton blooms along the equator, which can differ from the prevailing ENSO or PDO pattern. The easterly trade winds influence larger-scale bloom patterns over the tropical Pacific in general, but have much less of a correspondence to the anomalous CHL band along the equator over the long-term average.

4.4 Implications for higher trophic levels

Climate oscillations affecting phytoplankton have implications for higher trophic levels. Although the mechanism for biological regime shifts are not completely understood, fluctuations of fish stocks in widely separated regions over long term records, on the order of 100-1000 years, support a view of global scale climate influence [Schwartzlose et al., 1999; Cury et al., 2007]. Two categories of mechanisms have been ascribed to fisheries regime shifts: continuous modification (e.g. trends in food or temperature) or episodic environmental events over a season that trigger ecosystem changes [Schwartzlose et al., 1999]. Changes in gyre circulation exert a major influence on interdecadal shifts. The direction and velocity of atmospheric circulation, ocean currents and upwelling affect the rate of nutrient transfer into the euphotic zone, which in turn affects primary production and has been correlated to the harvest of small pelagic fish [Bakun, 1990; Bakun and Cury, 1999]. Tuna populations across the tropical Pacific

respond to large climatic changes and are projected to be redistributed as the warm pool expands [Lehodey et al., 1997; Bell et al., 2013]. Multidecadal fluctuations of fish in the Japan-Kuroshio, California, Peru-Humboldt, and even Australia systems have been shown to have cross-basin synchrony driven by similar climatic forcing rather than fishing pressure [Kawasaki, 1983; Schwartzlose et al., 1999; Bakun, 2001]. It has been noted that biological regime shifts sometimes precede physical variability, suggesting the possibility of marine organisms as harbingers of climate shifts [Hare and Mantua, 2000; Chavez et al., 2003]. The CHL reconstruction is compared to fish data to look for low frequency patterns in basin-scale primary production that could impact higher trophic levels through bottom-up control.

Nearly 70% of the world's tuna harvest comes from the Pacific, with Skipjack the dominant species in the region, distributed throughout the equatorial and subtropical surface mixed layer, but predominately found around the eastern edge of the west Pacific warm pool [Lehodey et al., 1997]. This front is the boundary between eastward currents in the warm pool and westward currents in the cold tongue and its zonal extension is primarily influenced by ENSO variability [LeBorgne et al., 2002]. The tuna population has been observed to shift eastward with the convergent front during El Niño and westward during La Niña events. Because the front consists of a warm, fresh layer at the surface, the barrier layer, above more saline, warm waters above the thermocline, SST is known to be an unrealiable indicator of its eastern boundary. Cruises to sample the east edge of the warm pool hydrography, nutrients and biology concluded that the edge is coincident with the western limit of CHL during active upwelling, but may become separated if upwelling stops, for example in the event of a downwelling Kelvin wave in

which the salinity and CHL fronts separated by 2-5^{o} while zooplankton biomass did not separate from the hydrological boundary [Rodier et al., 2000]. Surface salinity and ocean color data have both been used to identify the eastern edge of the warm pool, where the frontal zone is defined as 0.1 mg m^{-3} CHL concentration [Maes et al., 2010]. To delineate this edge and compare it to tuna distributions collected during 1988-1995 by Lehodey et al. [1999], reconstructed CHL were averaged between 5^{o}N-5^{o}S at each longitude, smoothed over three months and six degrees (Figure 4.24).

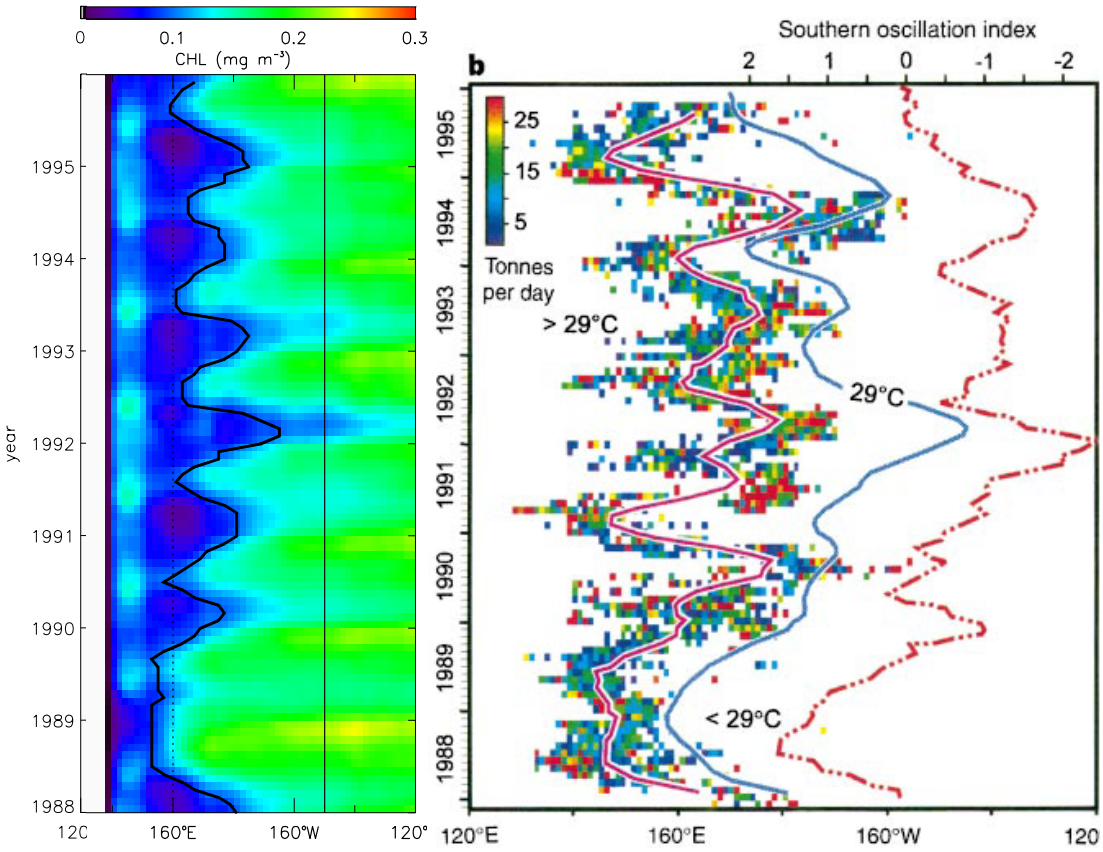

Figure 4.24 Left: Monthly CHL where the thick black contour demarcates 0.1 mg m^{-3}. Niño 4 is between dotted and solid lines; Niño 3 is east of solid line. Right: Monthly mean tuna catch per unit effort (CPUE) where the pink line indicates the longitudinal center of gravity of CPUE, the blue line indicates the 29^{o}C SST isotherm, the dashed red line indicates the Southern Oscillation Index. All variables were smoothed over three months. Figure on right from [Lehodey et al., 1997].

For six of eight years included in the tuna study, 1988-1993, the longitudinal variation of the 0.1 mg m^{-3} CHL contour aligns more closely with the distribution of tuna catch per unit effort (CPUE) over time than the SST front. The median separation between the tuna distribution indicated by the center of gravity of the CPUE and the CHL edge was 9^{o}, compared to 18^{o} for the SST definition. The RMS error is 12^{o} between the CPUE longitude and the CHL edge and 25^{o} between CPUE longitude and the SST edge. For example, in 1992 the tuna CPUE and 0.1 mg m^{-3} CHL contour reached their maximum eastward extent near the dateline and 170^{o}W, respectively while the 29^{o}C SST contour was located near 150^{o}W, a 30^{o} or >3,000km discrepancy. Likewise, from 1988-1990 the tuna CPUE and the CHL contour averaged within 10^{o} of each other, while the SST contour averaged 15^{o} farther east.

To compare the east edge of the warm pool defined by the 0.1 mg m^{-3} CHL contour to climate indices, the contour was smoothed over 12 months over the entire reconstruction time series, along with the ONI and PDO indices. The CHL contour most closely corresponds to the ENSO signal ($r = 0.85$), shifting eastward during El Niño and westward during La Niña, followed by the PDO with a lesser but still significant correlation ($r = 0.61$). Plotting the east edge next to the ONI index to represent ENSO and the PDO index demonstrates this correspondence through the span of the reconstruction (Figure 4.25). Only three of the strongest El Niños of the period pushed the edge east of the dateline (1957/58, 1982/83, 1997/98). The significant correlation of the CHL contour with ENSO and PDO is not surprising given the dominant effect physical forcing has on biology, yet there remains some variability in the CHL data that is not associated with typical climate indices. Smaller scale, localized forcing of primary production as well as

non-physical factors (e.g. fishing pressure, migration of the tuna or their prey) could influence the location of this species that occupies a niche near the top of the food web.

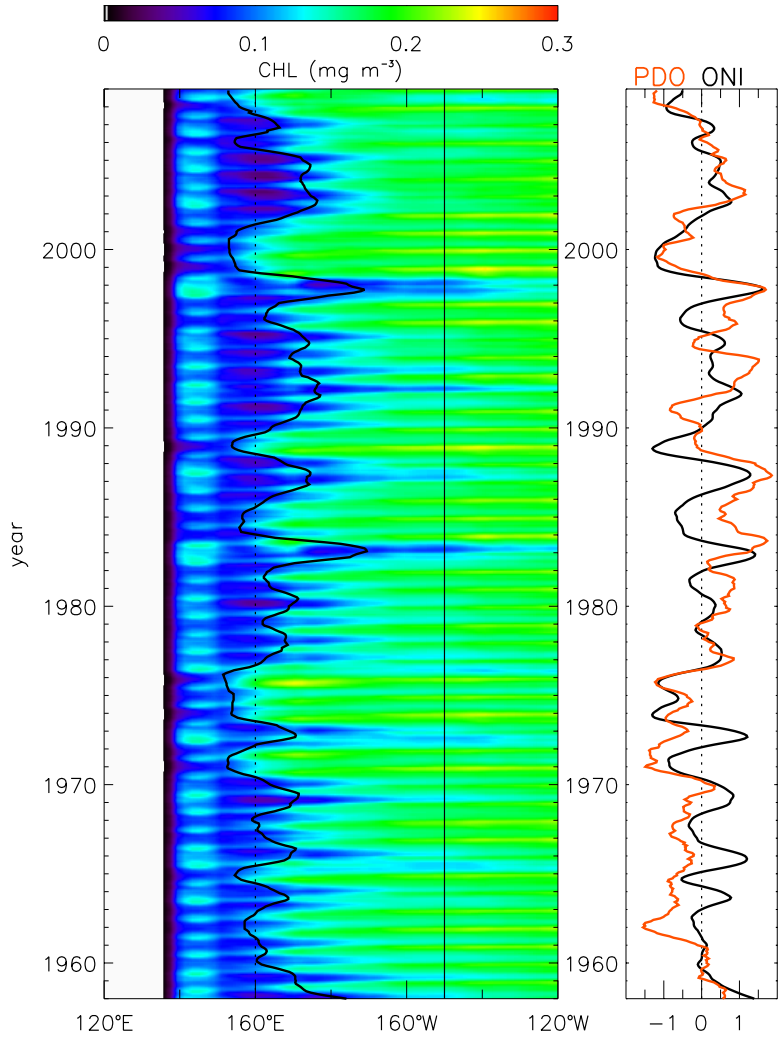

Figure 4.25 Left: Monthly CHL where contour demarcates 0.1 mg m^{-3} smoothed over 12 months. Niño 4 is between dotted and solid lines; Niño 3 is east of solid line. Right: Oceanic Niño Index (ONI) in black and Pacific Decadal Oscillation (PDO) in red, both smoothed over 12 months.

In addition to basin-scale climate patterns in marine predators such as tuna, several studies show cycles of Pacific salmon and other fish species. The PDO has been linked to shifts in Pacific fish regimes around 1950, 1975, and the late 1990s. As

illustrated in Figure 4.26, a fisheries regime has been identified that oscillates between anchovy-dominated for about 25 years and then sardine-dominated [Lluch-Belda et al., 1989; Chavez et al., 2003; Lehodey et al., 2006]. While fishing and predation play a role in fisheries abundance or collapse, the periodicity and synchrony of anchovies and sardines caught around the Pacific basin suggests a climatic cause.

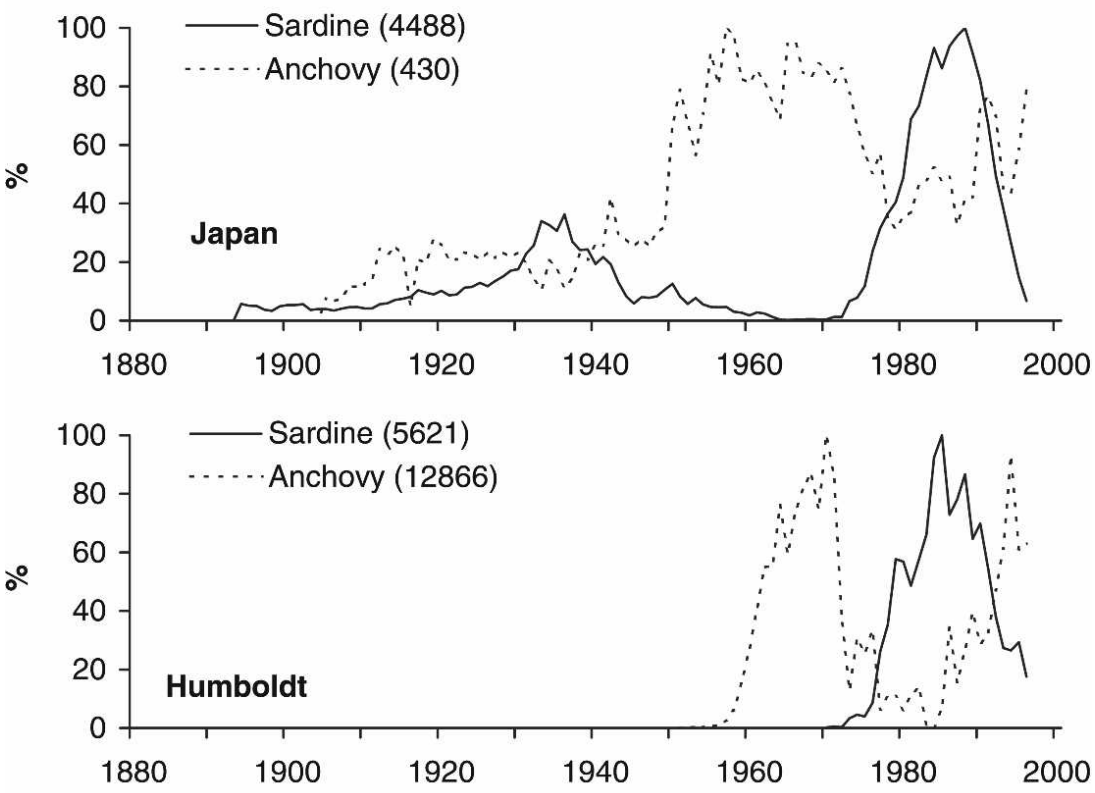

Figure 4.26 **Percent of peak value of sardine and anchovy populations in Japan-Kuroshio system (top) and Peru-Humboldt system (bottom). Figure adapted from [Lehodey et al., 2006].**

These pelagic fisheries vary over periods of about 50 years in coherence with large-scale variations in air and ocean temperatures, atmospheric CO_2, and climate indices such as the PDO [Chavez et al., 2003]. It has been noted in the literature that

fisheries landings in Peru were primarily anchovies during periods of stronger meridional overturning circulation, upwelling and higher primary production (1950-1975, then again after the late 1990s) or sardines when there was less meridional transport, weaker upwelling, and lower primary production (1975 - late 1990s) [Chavez et al., 2003]. Sardine regimes are characterized by a deeper thermocline and nutricline, weaker upwelling, warmer temperatures, lower nutrients, primary production, and CO_2 anomalies, the latter of which is measured at the Mauna Loa observatory with the annual cycle and long-term anthropogenic trend removed. By contrast, anchovy regimes experience a shallow thermocline and nutricline, stronger upwelling, cooler temperatures, higher nutrients, primary production, and CO_2 at Mauna Loa, implying possibly more CO_2 outgassing from equatorial upwelling assuming all other carbon sources were roughly constant [Chavez et al., 2003]. Anchovies feed on large zooplankton; sardines feed on phytoplankton and small zooplankton closer to the base of the food chain [Schwartzlose et al., 1999]. Sardines are more tropical and expand their range poleward during warmer periods. Anchovy thrive during cooler periods and do not expand far enough poleward during warm regimes to avoid increased temperatures, indicating an inability to adapt [Schwartzlose et al., 1999]. Only the regime shift in the late 1990s coincided with the ocean color era: the ocean ecosystem response was evident in both surface CHL as well as the higher trophic levels. All earlier biological regime shifts are missing a large-scale picture of primary production. This statistical reconstruction of CHL affords the opportunity to see the large-scale differences in primary production during anchovy and sardine regimes.

During the period from 1958 through 1975, an anchovy regime prevailed in the eastern Pacific coinciding with the PDO cool phase [Chavez et al., 2003; Lehodey et al., 2006]. In the CHL reconstruction, there was a pattern of very strong positive CHL anomalies (greater than 1 mg m^{-3}) along the coast of Central and South America, with the exception of small parts of Nino 1&2 (Figure 4.9). After the regime shifted to a warm phase in 1976-1977, the CHL reconstruction shows the CHL values switched as well. Strong negative anomalies, reaching lowest values of less than -1 mg m^{-3} off of Central America, were associated with the warm phase of the PDO during the sardine fisheries abundance (Figure 4.10). Although the CHL reconstruction does not yield information about phytoplankton community structure, it can be inferred that plentiful blooms support the larger zooplankton that feed anchovies while sparse blooms support the smaller species that feed sardines.

Low frequency variability in basin-scale biology as measured by fish catches has a known synchrony with interannual and decadal climate oscillations that is now shown to correspond to coherent spatial patterns in statistically reconstructed CHL maps. In the west Pacific warm pool, surface CHL has been demonstrated as a useful way to determine the eastern extent of tuna distributions which are primarily modified by ENSO and secondarily modified by the PDO. In the eastern tropical Pacific near the coast of South America, high CHL anomalies during the anchovy regime preceded and followed low CHL anomalies during the sardine regime. The CHL reconstruction across the tropical Pacific illuminates the effect of low frequency climate forcing on primary production, which has significant implications for ocean biology across trophic levels as well as the carbon cycle.

Table 4.1. Cross-correlations between CHL and physical variables for each regime

PDO cool regime (1958-1976)			
Niño 3 Inner	-1 quarter lag	zero lag	+1 quarter lag
CHL - SST	-0.65	-0.89	-0.69
CHL - SSH	-0.68	-0.83	-0.54
CHL - τ_x	-0.45	-0.36	-0.22
CHL - τ_Y	-0.05	-0.31	-0.29
Niño 3 Outer			
CHL - SST	-0.76	-0.92	-0.70
CHL - SSH	-0.69	-0.76	-0.53
CHL - τ_x	0.02	0.06	0.02
CHL - τ_Y	0.43	0.42	0.17
CHL - ($\nabla \times \tau$)	-0.39	-0.38	-0.19
Niño 3.4 Inner			
CHL - SST	-0.55	-0.87	-0.76
CHL - SSH	-0.64	-0.78	-0.53
CHL - τ_x	-0.66	-0.77	-0.55
CHL - τ_Y	0.01	-0.04	0.01
Niño 3.4 Outer			
CHL - SST	-0.77	-0.94	-0.73
CHL - SSH	-0.50	-0.49	-0.31
CHL - τ_x	-0.48	-0.34	-0.16
CHL - τ_Y	0.33	0.47	0.33
CHL - ($\nabla \times \tau$)	-0.43	-0.55	-0.40

PDO warm regime (1977-1995)			
Niño 3 Inner	-1 quarter lag	zero lag	+1 quarter lag
CHL - SST	-0.68	-0.90	-0.73
CHL - SSH	-0.77	-0.88	-0.58
CHL - τ_x	-0.19	-0.34	-0.43
CHL - τ_Y	-0.05	0.07	0.20
Niño 3 Outer			
CHL - SST	-0.78	-0.90	-0.69
CHL - SSH	-0.80	-0.87	-0.60
CHL - τ_x	0.26	0.23	0.05
CHL - τ_Y	0.68	0.74	0.55
CHL - ($\nabla \times \tau$)	-0.56	-0.60	-0.50
Niño 3.4 Inner			
CHL - SST	-0.64	-0.88	-0.71
CHL - SSH	-0.71	-0.73	-0.39
CHL - τ_x	-0.54	-0.73	-0.64
CHL - τ_Y	0.08	0.30	0.42
Niño 3.4 Outer			
CHL - SST	-0.78	-0.89	-0.67
CHL - SSH	-0.66	-0.65	-0.37
CHL - τ_x	-0.22	-0.08	-0.00
CHL - τ_Y	0.61	0.75	0.61
CHL - ($\nabla \times \tau$)	-0.55	-0.69	-0.57

5. Discussion

In addition to their impact upon fisheries and the marine food web, CHL spatial patterns also have biogeochemical implications, such as the ocean's uptake of atmospheric CO_2. While it is unclear how increasing levels of atmospheric CO_2 will impact primary production in the ocean over the long-term, it is clear from this study that slow changes caused by ocean circulation pattern shifts impact ocean surface CHL. It has been noted that the Walker circulation has strengthened since the 1950s in the atmosphere above the tropical Pacific, creating a positive feed-back loop [L'Heureux et al., 2013]: enhanced SST gradients strengthen Walker circulation, stronger equatorial winds enhance upwelling leading to cooler SST anomalies in the east and strengthening Walker circulation. Faster circulation also increases sea-level rise in the western tropical Pacific [Church et al., 2004] and strengthens surface and subsurface currents [Qiu and Chen, 2012]. Simultaneously, the EUC in the subsurface ocean has strengthened and shoaled since the mid-nineteenth century [Drenkard and Karnauskas, 2014]. Such long-term trends may reflect internal multi-decadal variability, but their coexistence with the increasing positive trend in global mean temperatures leads to the supposition that anthropogenic forcing plays a role. These reconstructed CHL were not used to detect any secular trend, as that would require a more robust *in situ* data set to anchor the beginning of the record with greater confidence. Nonetheless, the EUC variability observed and linked to CHL variability along the equator has a long-term trend of strengthening and shoaling that has been ascribed to anthropogenic changes. *In situ* data and modeling studies indicate that interannual climate variability affects phytoplankton community composition, with the equatorial Pacific experiencing radical taxonomic shifts in

response to ENSO [Uitz et al., 2010; Rousseaux and Gregg, 2012]. Whether phytoplankton abundance will increase, decrease or experience a change in species composition over long time scales are areas of active research.

Extending this reconstruction farther back in time would be challenging without a complete, well-validated observational record to provide useful skill. This study started in 1958 when the frequency of hydrographic observations greatly increased [Carton et al., 2012]. More recently, SODA 2.2.6 has assimilated observations of SST only to extend the record back to 1871. The reconstruction could be calculated back to 1871 using SST as the only predictor, with additional testing to determine the best method and the optimal number of modes for that application. Another option would be to add a second predictor from another record, such as sediment trap or corals, although these could have coarser resolution (i.e. annual or decadal) and would insert new uncertainties. This study demonstrates that there is the potential that a longer multi-decadal statistical reconstruction of CHL could be performed in the tropical Pacific, with the caveats noted, and could yield an even greater understanding of lower frequency patterns in ocean surface phytoplankton.

6. Concluding remarks

Interannual variability in ocean surface biology has been well-documented with ocean color remote sensing since 1997, yet the effect of decadal and multi-decadal climate oscillations on oceanic primary production has been unknown at basin scales due to the lack of a broad coverage data set spanning multiple decades. Here a method was developed to statistically reconstruct ocean color CHL in the tropical Pacific taking advantage of the strong physical forcing of biology in this region. Using the most highly correlated variables as proxies, SST and SSH from SODA, the ocean color CHL record was extended back to 1958. The reconstructed CHL were validated against a fully coupled physical-biogeochemical model [Wang et al., 2008] as well as sparse *in situ* observations. This reconstruction shows that tropical Pacific phytoplankton biomass oscillates at low frequencies due to physical forcing, both locally driven and originating outside the tropics. Changes in meridional overturning circulation between PDO regimes combined with changes in the strength, frequency and eastward extent of ENSO all impact CHL at basin-scale and have world-wide ramifications. Additionally, long-term trends in the depth of the EUC influence phytoplankton blooms for a narrow region along the equator toward the east.

The largest periodic variability across the tropical Pacific is interannual due to ENSO. While the impact of ENSO is observed in the existing ocean color record, this multi-decadal reconstruction adds to the record many more events of varying strength, duration and zonal extent. The differences between the east Pacific El Niño and central Pacific El Niño are observed with greater clarity over 51 years, revealing the contrast in CHL spatial patterns due to their distinct forcing mechanisms. The east Pacific El Niño

results from basin-scale thermocline flattening that causes nearly universal declines in phytoplankton standing stock, except in the west Pacific warm pool region which experiences more productivity. The central Pacific El Niño is a localized response to wind forcing that pushes the warm pool east toward the dateline, reducing CHL on the equator and slightly away from the equator in a wishbone pattern, while productivity continues as usual to the east along the equator and coast of South America.

Biological variation with the PDO has a smaller magnitude, but is clearly important because it strengthens or weakens ENSO, as well as leading to synchronous changes in higher trophic levels. For most of the tropical Pacific, extending from the Americas westward to the edge of the warm pool, the PDO cool phase is associated with greater CHL while the warm phase has less. This basin-scale pattern in ocean biology is generally consistent with conditions at the ocean's surface, namely zonal winds, wind strength and SST.

Contrary to this larger pattern is a narrow band along the equator between 110-150^{o}W that had the opposite productivity tendency over the cool PDO regime between 1958-1977 and the warm PDO regime between 1978-1997. This area had its own response in CHL often different from the rest of the region in annual, decadal and multi-decadal averages associated with the cool or warm phase of the PDO. The mechanism most likely to cause this narrow anomalous CHL band is the depth of the EUC, which supplies iron to the high-nutrient, low-chlorophyll, iron-limited region. Easterly winds along the equator that support divergence and upwelling do not lead to blooms if the nutricline is below the euphotic zone or if there is a decrease in iron supply by the EUC. Low frequency changes in the depth of the EUC impact CHL at the surface by shoaling

and providing iron to the euphotic zone. Ocean current velocities from the TAO array observations and SODA reanalysis indicate a shoaling trend by the EUC over the past five decades. During the PDO cool phase between 1958-1977, when the cold tongue had an overall positive CHL anomaly, the EUC was at its deepest for the reconstruction period and there was a negative CHL anomaly along the equatorial band. During the PDO warm phase between 1978-1997, the cold tongue had a predominately negative CHL anomaly, yet there was a positive anomaly along the equator where the EUC rose into the euphotic zone between 110-150^{o}W. By the end of the reconstruction, between 1998-2007, the PDO was in a cool phase yet the EUC was at its shallowest and the CHL anomaly along the equator was positive. These observations support a conclusion that over most of the tropical Pacific, winds are the predominant forcing mechanism controlling CHL variation, while over the equatorial strip the depth of the EUC determines productivity through iron supply.

The CHL spatial patterns identified here have important ecological implications. CHL can be used to demarcate the convergent front on the eastern edge of the west Pacific warm pool. The front's location in the CHL reconstruction corresponds well to the center of gravity in the skipjack tuna catch record. Previous studies linked low frequency climate oscillations to the anchovy and sardine fisheries shifts around the Pacific basin as well as fluctuations in tuna catches in the west and east tropical Pacific. This reconstruction now contributes a long-term, basin-scale view of primary production, as these low frequency patterns have an impact on higher trophic levels that is an area of active research.

One of the biggest questions in climate science today is: how much carbon can the ocean take up through biological and chemical processes and export to great depths for long-term carbon sequestration? Although high resolution global CHL measured by satellite over a decade revolutionized our understanding of physical controls of biological processes in the ocean, we cannot answer the carbon export question with surface CHL alone. We need a vertical view to link primary production to how much carbon sinks out of the surface layer to the deep ocean. Regular measurements of the vertical structure of CHL and other biological variables are in their infancy, but in a few years emerging technologies such as Bio-Argo floats will be able to sample subsurface biological processes. This work has established that CHL patterns can be reconstructed using physical proxies. With a well-resolved training data set, we should then be able to relate surface CHL patterns to vertical profiles and processes to develop a statistical reconstruction of carbon export.

Appendix

The CCA method in this study is described in detail by Barnett and Preisendorfer [1987]. Briefly, the analysis involves training spatial correlation functions of the predictand (CHL) and predictor (SST and SSH) fields during the training period. First, EOFs are calculated on deseasoned, normalized pairs of the predictand and predictors:

$$\mathrm{SST\&SSH(x,t)} = \sum_{j=1}^{p} \kappa_j^{\frac{1}{2}} \alpha_j(t) e_j(x) \qquad x = 1,2,\ldots,p$$

$$\mathrm{CHL(x',t)} = \sum_{j=1}^{q} \lambda_j^{\frac{1}{2}} \beta_j(t) f_j(x') \qquad x' = 1,2,\ldots,q \tag{A1}$$

Where κ_j and λ_j are eigenvalues; $\alpha_j(t)$ and $\beta_j(t)$, principal components; e_j and f_j are orthonormal eigenvectors. The number of predictor and predictand spatial points is p and q, respectively. Typically the number of EOF modes used is much less than the minimum of p and q to filter out noise and variations that cannot be reliably modeled.

Optimal representation of CHL in terms of SST&SSH is obtained by first forming the set of all linear combinations of the α_i and β_j in EOF spectral space:

$$\boldsymbol{u} = \sum_{j=1}^{p} \alpha_j r_j \qquad \text{and} \qquad \boldsymbol{v} = \sum_{k=1}^{q} \beta_k s_k \tag{A2}$$

where **r** and **s** are unit vectors in Euclidean vector space. For each choice of **r** and **s**, the correlation of $\boldsymbol{u}$ and $\boldsymbol{v}$ is maximized if r and s are the eigenvectors of the system:

$$[\mathbf{C}\mathbf{C}^T]\boldsymbol{r_j} = \boldsymbol{\mu_j^2 r_j} \qquad \boldsymbol{j} = \mathbf{1},\mathbf{2},\ldots,\boldsymbol{p}$$

$$[\mathbf{C}^T\mathbf{C}]\boldsymbol{s_k} = \boldsymbol{\mu_k^2 s_k} \qquad k = 1,2,\ldots,q \tag{A3}$$

where $\mathbf{r}_j = [r_{1j}, r_{2j}, \ldots r_{pj}]^T$ and similarly for $\mathbf{s}_k$, which form orthonormal sets of vectors in EOF spectral space by matrix theory, and μ_j are called canonical correlation coefficients.

Remembering (A2), the canonical component vectors are calculated:

$$\boldsymbol{u_j} = \sum_{i=1}^{p} \alpha_i r_{ij} \qquad \text{and} \qquad \boldsymbol{v_k} = \sum_{i=1}^{q} \beta_i s_{ik} \tag{A4}$$

All of which permits the representation of SST&SSH and CHL datasets as linear combinations of their canonical component vectors:

$$\mathrm{SST\&SSH(x,t)} = \sum_{j=1}^{p} u_j(t) g_j(x) \quad where\ g_j(x) = \langle SST\&SSH(x,t) u_j(t) \rangle_t$$

$$\mathrm{CHL(x',t)} = \sum_{k=1}^{q} v_k(t) h_k(x') \qquad where\ h_k(x') = \langle CHL(x',t) v_k(t) \rangle_t \tag{A5}$$

The components of the canonical map vectors, $\mathbf{g}_j$ and $\mathbf{h}_k$, show the correlation between SST&SSH and CHL and their respective canonical component time series (j or k) at a given location (x or x'). A linear combination of the canonical component vectors $\boldsymbol{u}_j$ of the predictor dataset (SST&SSH) can represent the n-dimensional predictand vector, CHL(x',·) = [CHL(x',1,…t(x',n)]^T, by projecting the CHL vectors onto the *q*-dimensional vector space spanned by the first q of the $\boldsymbol{u}_j$; j=1,…,q < p:

$$\widehat{\mathrm{CHL}}(\mathrm{x',t}) = \sum_{j=1}^{q''} \mu_j u_j(t) h_j(x') \qquad x' = 1,2,\dots,q'' \le q$$

$$t = 1,2,\dots,n \tag{A6}$$

where $q'' < q$ is the number of canonical modes retained for the reconstruction. The number of modes retained can be determined by observing the *q''* value which yields the maximum significant reconstruction skill.

References

Alvain, S., C. Moulin, Y. Dandonneau, F.M. Breon, 2005, Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery, *Deep Sea Res. I*, 52/11, 1989-2004, doi:10.1016/j.dsr.2005.06.015.

Alvera-Azcárate, A., A. Barth, J.-M. Beckers and R.H. Weisberg, 2007, Multivariate reconstruction of missing data in sea surface temperature, chlorophyll and wind satellite fields, *J. Geophys. Res.*, 112, C03008, doi:10.1029/2006JC003660.

Antoine, D., F. d'Ortenzio, S.B. Hooker, G. Becu, B. Gentili, D. Tailliez, and A.J. Scott, 2008, Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color sensors (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean Sea (BOUSSOLE project), *J. Geophys. Res.,* 113, C07013, doi:10.1029/2007JC004472.

Ashok, K., S. K. Behera, S.A. Rao, H. Weng, and T. Yamagata, 2007, El Niño Modoki and its possible teleconnection. *J. Geophy. Res.* 112, C11007, doi:10.1029/2006JC003798.

Ayers, J.M. and M.S. Lozier, 2010, Physical controls on the seasonal migration of the North Pacific transition chlorophyll front, *J. Geophys. Res.*, 115:C5, doi:10.1029/2009jc005596.

Bailey, S.W. and P.J. Werdell, 2006, A multi-sensor approach for the on-orbit validation of ocean color satellite data products, *Remote Sens. Environ*., 102, 12-23, doi:10.1016/j.rse.2006.01.015.

Bakun, A., 1990, Global climate change and intensification of coastal ocean upwelling. *Science*, *247*, 198–201.

Bakun, A. and P. Cury, 1999, The 'school trap': a mechanism promoting large-amplitude out-of-phase population oscillations of small pelagic fish species. *Ecology Letters*, *2*, 349–351.

Bakun, A., 2001, 'School-mix feedback': a different way to think about low frequency variability in large mobile fish populations, *Prog. in Oceano.*, 49, 485-511.

Bakun, A. and S. Weeks, 2008, The marine ecosystem off Peru: What are the secrets of its fishery productivity and what might its future hold?, *Prog. in Oceano.*, 79, 290-299.

Ballabrera-Poy, J., R.G. Murtugudde, J.R. Christian and A.J. Busalacchi, 2003, Signal-to-noise ratios of observed monthly tropical ocean color, *Geophys. Res. Lett.,* 30:12, doi:10.1029/2003GL016995.

Barber, R.T. and F.P. Chavez, 1983, Biological consequences of El Niño, *Science,* 222:4629, 1203-1210.

Barnett, T.P. and R. Preisendorfer, 1987, Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis, *Monthly Weather Review*, 115,1825-1850.

Beaulieu, C., S.A. Henson, J.L. Sarmiento, J.P. Dunne, S.C. Doney, R.R. Rykaczewski, and L. Bopp, 2013, Factors challenging our ability to detect long-term trends in ocean chlorophyll, *Biogeosciences*,10 ,2711-2724, doi:10.5194/bg-10-2711-2013.

Bell, J.D., A. Ganachaud, P.C. Gehrke, S.P. Griffiths, A.J. Hobday, O. Hoegh-Guldberg, J.E. Johnson, R. Le Borgne, P. Lehodey, J.M. Lough, R.J. Matear, T.D. Pickering, M.S. Pratchett, A. Sen Gupta, I. Senina and M. Waycott, 2013, Mixed responses of tropical Pacific fisheries and aquaculture to climate change, *Nature Climate Change*, doi:10.1038/NCLIMATE1838.

Behrenfeld, M.J., R.T. O'Malley, D.A. Siegel, C.R. McClain, J.L. Sarmiento, G.C. Feldman, A.J. Milligan, P.G. Falkowski, R.M. Letelier,E.S. Boss, 2006, Climate-driven trends in contemporary ocean productivity, *Nature*, 444, doi:10.1038/nature05317.

Behrenfeld, M.J., 2010, Abandoning Sverdrup's critical depth hypothesis on phytoplankton blooms, *Ecology*, 91, doi:10.1890/09-1207.1.

Bjerknes, J., 1969, Atmospheric teleconnections from the equatorial Pacific. *Mon. Weather Rev.*, 97, 163-172.

Boss, E.S., D. Swift, L. Taylor, P. Brickley, R. Zaneveld, S.C. Riser, M.J. Perry, and P.G. Strutton, 2008, Observations of pigment and particle distributions in the western North Atlantic from an autonomous float and ocean color satellite, *Limnology and Oceanography*, 53:2, 112-122.

Boyce, D.G., M.R. Lewis, B. Worm, 2010, Global phytoplankton decline over the past century, *Nature*, 466, doi:10.1038/nature09268.

Boyd, P.W., T. Jickells, C.S. Law, S. Blain, E.A. Boyle, K.O. Buesseler, K.H. Coale, J.J. Cullen, J.J.W. de Baar, M. Follows, M. Harvey, C. Lancelot, M. Levasseur, N.P.J. Owens, R. Pollard, R.B. Rivkin, J. Sarmiento, V. Schoemann, V. Smetacek, S. Takeda, A. Tsuda, S. Turner, A.J. Watson, 2007, Mesoscale iron enrichment experiments 1993-2005: synthesis and future directions, *Science,* 315, 612, doi:10.1126/science.1131669.

Bretherton, C.S., C. Smith and J.M. Wallace, 1992, An intercomparison of methods for finding coupled patterns in climate data*, J. Climate*, 5, 541-560.

Bricaud, A., M. Babin, A. Morel, and H. Claustre, 1995, Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: analysis and parameterization, *J. Geophys. Res,* 100: C7, 13,321-13,332.

Brown, C.W. and J.A. Yoder, 1994, Coccolithophorid blooms in the global ocean, *J. Geophys. Res.,* 99, 7467-7482, doi: 10.1029/93JC02156

Cane, M.A., S.E. Zebiak, S.C. Dolan, 1986, Experimental forecasts of El Niño, *Nature*, 321, 827-832.

Carder, K.L., F.R. Chen, Z.P. Lee, S.K. Hawes, D. Kamykowski, 1999, Semianalytic Moderate-Resolution Imaging Spectrometer algorithms for chlorophyll a and absorption with bio-optical domains based on nitrate-depletion temperatures, *J. Geophys. Res.*, 104:C3, pp 5403-5421.

Carton, J.A. and B.S. Giese, 2008, A reanalysis of ocean climate using Simple Ocean Data Assimiliation (SODA), *Monthly Weather Review*, 136, doi: 10.1175/2007MWR1978.1.

Carton, J.A., H.F. Seidel, B.S. Giese, 2012, Detecting historical ocean climate variability, *J. Geophys. Res*., 117, C02023, doi:10.1029/2011JC007401.

Carton, J.A., 2013, Table 1: list of eddy permitting SODA reanalysis experiments. Available at http://www.atmos.umd.edu/~carton/index2_files/soda.htm (accessed 30 January 2014).

Chavez, F P. and J.R. Toggweiler, 1995, Physical estimates of global new production: the upwelling contribution, In *Dahlem Workshop on Upwelling in the Ocean: Modern Processes and Ancient Records*, Chichester, UK, John Wiley & Sons, 313-320.

Chavez, F.P., P.G. Strutton, G.E. Friederich, R.A. Feely, G.A. Feldman, D. Foley, and M.J. McPhaden, 1999, Biological and chemical response of the equatorial Pacific Ocean to the 1997 and 1998 El Niño, *Science*, 286, 2126-2131.

Chavez, F.P., J. Ryan, S.E. Lluch-Cota, M. Niquen C., 2003, From anchovies to sardines and back: multidecadal change in the Pacific Ocean, *Science*, 299, 217-221.

Christian, J.R., M.A. Verschell, R.G. Murtugudde, A.J. Busalacchi, and C.R. McClain, 2002, Biogeochemical modeling of the tropical Pacific Ocean. I: Seasonal and interannual variability, *Deep Sea Res., Part II, 49*, 509-543, doi:10.1016/S0967-0645(01)00110-2.

Church, J.A., N.J. White, R. Coleman, K. Lambeck, J.X. Mitrovica, 2004, Estimates of the regional distribution of sea level rise over the 1950-2000 period, *J. Climate*, 2609-2625.

Cianca, A., J.M. Godoy, J.M. Martin, J. Perez-Marrero, M.J. Rueda, O. Llinas, S. Neuer, 2012, Interannual variability of chlorophyll and the influence of low-frequency climate modes in the North Atlantic subtropical gyre, *Global Biogeochem. Cycles*, 26:2, doi:10.1029/2010gb004022.

Clark, D.K. et al., 2003, MOBY, a radiometric buoy for performance monitoring and vicarious calibration of satellite ocean color sensors: measurement and data analysis protocols, *Ocean Optics Protocols for Satellite Ocean Color Sensor validation, Rev. 4*, vol VI, NASA/TM – 2003-211621, 3-34.

Conkright, M.E. and W.W. Gregg, 2003, Comparison of global chlorophyll climatologies: *In situ*, CZCS, Blended *in situ*-CZCS and SeaWiFS, *Int. J. Remote Sensing*, 24, 969-991.

Cury, P., L. Shannon, and Y.-J. Shin, 2007, The functioning of marine ecosystems, presentation at the 1-4 October, 2001 Reykjavik Conference on Responsible Fisheries in the Marine Ecosystem.

Davis, R.E., 1976, Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific ocean, *J. Phys. Oceano.*, 6:3, 249-266.

de Baar, H.J.W., J.T.M de Jong, D.C.E. Baaker, B.M. Loscher, C. Veth, U. Bathmann, and V. Smetacek,1995, Importance of iron for plankton blooms and carbon dioxide drawdown in the Southern Ocean, *Nature*, 373, 412-415.

DiLorenzo, E., N. Schneider, K.M. Cobb, P.J.S. Franks, K. Chhak, A.J. Miller, J.C. McWilliams, S.J. Bograd, H. Arango, E. Curchister, T.M. Powell, P. Riviere, 2008, North Pacific gyre oscillation links ocean climate and ecosystem change, *Geophys. Res. Lett*., 35, doi:10.1029/2007GL032838.

DiLorenzo, E., K.M. Cobb, J.C. Furtado, N. Schneider, B.T. Anderson, A. Bracco, M.A. Alexander, and D.J. Vimont, 2010, Central Pacific El Niño and decadal climate change in the North Pacific ocean, *Nature Geosci.*, doi: 10.1038/NGEO984.

Drenkard, E.J. and K.B. Karnauskas, 2014, Strengthening of the Pacific Equatorial Undercurrent in the SODA reanalysis: mechanisms, ocean dynamics, and implications, *J. Climate*, doi:10.1175/JCLI-D-13-00359.1.

Drinkwater, K.F., G. Beaugrand, M. Kaeriyama, S. Kim, G. Ottersen, R.I. Perry, H.-O. Portner, J.J. Polovina, A. Takasuka, 2010, On the processes linking climate to ecosystem changes, *J. Mar. Sys.*, 29, 374-388, doi:10.1016/j.jmarsys.2008.12.014.

Ducklow, H.W., S.C. Doney, and D.K. Steinberg, 2009, Contributions of long-term research and time-series observations to marine ecology and biogeochemistry. *Annual Review of Marine Science*, 1, 279-302.

Evans, M.N., M.A. Cane, D.P. Schrag, A. Kaplan, B.K. Linsley, R. Villalba, and G.M. Wellington, 2001, Support for tropically-driven Pacific decadal variability based on paleoproxy evidence, *Geophys. Res. Lett.*, 28:19, 3689-3692.

Evans, W., P.G. Strutton and F.P. Chavez, 2009, Impact of tropical instability waves on nutrient and chlorophyll distributions in the equatorial Pacific, *Deep Sea Res. I*, 56, 178-188, doi:10.1016/j.dsr.2008.08.008.

Federov, A.V. and S.G. Philander, 2000, Is El Niño Changing? *Science*, 288, 1997-2002.

Feely, R.A., R. Wanninkhof, T. Takahashi and P. Tans, 1999, The influence of the equatorial Pacific contribution to atmospheric CO_2 accumulation. *Nature*, 398, 597–601.

Feldman, G.C., D. Clark, and D. Halpern, 1984, Satellite color observations of the phytoplankton distribution in the eastern Equatorial Pacific during the1982-83 El Niño, *Science*, 226, 1069-71.

Feldman, G.C., 1989, Ocean Colors Imaged from Space Track the Distribution of Marine Phytoplankton. *Earth in Space,* 2:2.

Field, C.B., M.J. Behrenfeld, M.J., J.T. Randerson, P. Falkowski, 1998, Primary production of the biosphere: Integrating terrestrial and oceanic components, *Science*, 281, 237-240.

Fisher, R.A., 1921, On the 'probable error' of a coefficient of correlation deduced from a small sample, *Metron,* 1:4, 3-32.

Follows, M.J., S. Dutkiewicz, S. Grant, S.W. Chisholm, 2007, Emergent biogeography of microbial communities in a model ocean, *Science*, 315, 1843; doi:10.1126/science.1138544.

Franz, B., Werdell, J., Meister, G., Bailey, S., Eplee, R., Feldman, G., Kwiatkowska, E., McClain, C., Patt, F. and D. Thomas, 2005, The Continuity of Ocean Color Measurements from SeaWiFS to MODIS. *Earth Observing Systems X, SPIE,* Vol. 5882.

Franz, B., and the NASA Ocean Biology Processing Group, 2011, OBPG Satellite CalVal Update, presentation at the October 6, 2011 OB&B Program Meeeting.

Gierach, M.M., T. Lee, D. Turk, and M.J. McPhaden, 2012, Biological response to the 1997-98 and 2009-10 El Niño events in the equatorial Pacific Ocean, *Geophys. Res. Lett.*, 39, doi:10.1029/2012GL051103.

Giese, B.S. and S. Ray, 2011, El Niño variability in simple ocean data assimilation (SODA), 1871-2008, *J. Geophys. Res*., 116, doi:10.1029/2010JC006695.

Gill, A.E. and E.M. Rasmusson, 1983, The 1982-83 climate anomaly in the equatorial Pacific, *Nature*, 306, 229-234.

Gnanadesikan, A., J.L. Russell, Z. Fanrong, 2007, How does ocean ventilation change under global warming? *Ocean Sci*., 3, pp. 43-53, doi:10.5194/os-3-43-2007.

Gordon, H.R. and D.K. Clark, 1980, Atmospheric effects in the remote sensing of phytoplankton pigments, *Boundary-Layer Meteorology*, 18, 299-313.

Gordon, H. R. and M. Wang, 1994, Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: A preliminary algorithm, *Applied Optics,* 33, 443–452.

Gordon, H.R., 1997, Atmospheric correction of ocean color imagery in the Earth Observing System era, *J. Geophys. Res.*, 102: D14, 17081–17106*,* doi:10.1029/96JD02443.

Gregg, W.W. and N.W. Casey, 2007, Sampling biases in MODIS and SeaWiFS ocean color chlorophyll data, *Rem. Sens. Env*., 111, 25-35.

Gruber, N., M. Gloor, S.E. Mikaloff Fletcher, S.C. Doney, S. Dutkiewica, M.J. Follows, M. Gerber, A.R. Jacobson, F. Joos, K. Lindsay, D. Menemenlis, A. Mouchet, S.A. Muller, J.L. Sarmiento, T. Takahashi, 2009, Oceanic sources, sinks, and transport of atmospheric CO_2, *Global Biogeochem. Cycles*, 23, doi:10.1029/2008GB003349.

Hare, S.R. and N.J. Mantua, 2000, Empirical evidence for North Pacific regime shifts in 1977 and 1989, *Prog. in Oceano.*, 47, 103-145.

Hartmann, D. L. and P. Mouginis-Mark, 1999, Volcanoes and climate effects of aerosols, in *EOS Science Plan*, pp. 339-378.

Henson, S.A., D. Raitsos, J.P. Dunne, A. McQuatters-Gollop, 2009, Decadal variability in biogeochemical models: Comparison with a 50-year ocean colour dataset, *Geophys. Res. Lett.*., 36, 10.1029/2009gl040874.

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J. W., Aiken, J., Barlow, R., Suzuki, K., Isada, T., Howell, E., Hashioka, T., Noguchi-Aita, M., and Yamanaka, Y.: Synoptic relationships between surface Chlorophyll-*a* and diagnostic pigments specific to phytoplankton functional types, *Biogeosciences, 8*, 311-327, doi:10.5194/bg-8-311-2011, 2011.

Holm-Hansen, O., C.J. Lorenzen, R.W. Holmes, and J.D.H. Strickland, 1965, Fluorometric determination of chlorophyll, *J. Cons. Perm. Int. Explor. Mer*, 30: 3-15.

Hsin, Y.-C. and B. Qiu, 2012, Seasonal fluctuations of the surface North Equatorial Counter-Current (NECC) across the Pacific basin., *J. Geophys. Res.*, 117, C06001, doi:10.1029/2011JC007794.

Hu, C.M., L. Feng, Z. Lee, C.O. Davis, A. Mannino, C.R. McClain, B.A. Franz, 2012, Dynamic range and sensitivity requirements of satellite ocean color sensors: learning from the past. *Applied Optics*, 51(25), 6045-6062.

IOCCG, 2011, Bio-optical sensors on Argo floats, Claustre, H. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 11, IOCCG, Dartmouth, Canada.

Jackson, D., 1993, Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches, *Ecology,* 74(8), 2204-2214.

Jin, D., D.E. Waliser, C. Jones, and R. Murtugudde, 2013, Modulation of tropical ocean surface chlorophyll by the Madden-Julian Oscillation, *Clim. Dyn.*, 40: 39-58, doi:10.1007/s00382-012-1321-4.

Johnson, K.S., W.M. Berelson, E.S. Boss, Z. Chase, H. Claustre, S.R. Emerson, N. Gruber, A. Kortzinger, J.J. Perry, and S.C. Riser, 2009, Observing biogeochemical cycles at global scales with profiling floats and gliders: prospects for a global array, *Oceanography*, 22: 3, 216-225.

Jolliffe, I.T., 2002, *Principal Component Analysis, Second Edition*. Springer, New York, 487 pp.

Kahru, M., S.T. Gille, R. Murtugudde, P.G. Strutton, M. Manzano-Sarabia, H. Wang and B.G. Mitchell, 2010, Global correlations between winds and ocean chlorophyll, *J. Geophys. Res.*, 115, doi:10.1029/2010JC006500.

Kamykowski, D. and S.-J. Zentara, 1986, Predicting plant nutrient concentrations from temperature and sigma-t in the upper kilometer of the world ocean, *Deep Sea Res.*, 33, 89-105.

Kaplan, A., M.A. Cane, Y. Kushnir, A.C. Clement, M. B. Blumenthal and B. Rajagopalan,1998, Analyses of global sea surface temperature 1856-1991, *J. Geophys., Res.*, 103, C9, 18,567-18,589.

Kaplan, A., Y. Kushnir and M.A. Cane, 2000, Reduced space optimal interpolation of historical marine sea level pressure: 1854-1992, *J. Climate*, 13, 2987-3002.

Kaplan, A., 2011, Patterns and indices of climate variability, [in "State of the Climate in 2010"]. *Bull. Amer. Meteor.Soc*., 92 (1.1), S20–S25.

Karnauskas, K., 2013, Can we distinguish canonical El Niño from Modoki? *Geophys. Res. Lett.*, 40, doi:10.1002/grl.51007.

Kawasaki, T., 1983, Why do some pelagic fishes have wide fluctuations in their numbers? - biological basis of fluctuation from the viewpoint of evolutionary ecology., p. 1065-1080. *In* G.D. Sharp and J. Csirke (eds.), Reports of the Expert Consultation to Examine Changes in Abundance and Species Composition of Neritic Fish Resources. *FAO Fish. Rep. 291*:(2, 3), pp. 1224.

Kennedy, J.J., 2014, A review of uncertainty in in situ measurements and data sets of sea surface temperature, *Reviews of Geophys*., doi:10.1002/2013RG000434.

Klyashtorin, L.B., 1998, Long-term climate change and main commercial fish production in the Atlantic and Pacific, *Fisheries Res.*, 37, 115-125.

Knauss, J.A., 1960, Measurements of the Cromwell Current, *Deep-Sea Res*., 6, 265-286.

Lagerloef, G.S.E., G.T. Mitchum, R.B. Lukas, and P. P. Niiler, 1999, Tropical Pacific near-surface currents estimated from altimeter, wind and drifter data, *Marine Science Faculty Publication,* Paper 47.

Le Borgne, R., R.T. Barber, T. Delcroix, H.Y. Inoue, D.J. Mackey and M. Rodier, 2002, Pacific warm pool and divergence: temporal and zonal variations on the equator and their effects on the biological pump, *Deep Sea Res. II,* 49, 2471-2512.

Lee, T. and M.J. McPhaden, 2010, Increasing intensity of El Niño in the central-equatorial Pacific, *Geophys. Res. Lett.*, 37 (14): L14603, doi:10.1029/2010GL044007.

Lehodey, P., M. Bertignac, J. Hampton, A. Lewis and J. Picaut, 1997, El Niño Southern Oscillation and tuna in the western Pacific, *Nature*, 389, 715-718.

Lehodey, P., J. Alheit, M. Barange, T.Baumgartner, G. Beaugrand, K. Drinkwater, J.-M. Fromentin, S.R. Hare, G. Ottersen, R,I, Perry, C. Roy, C.D. Van der Lingen and F. Werner, 2006, Climate Variability, Fish, and Fisheries, *J. Clim*., 5009-5030.

Levy, M., Y. Lehahn, J.-M. Andre, L. Memery, H. Loisel, and E. Heifetz, 2005, Production regimes in the northeast Atlantic: a study based on Sea-viewing Wide Field-of-view Sensor (SeaWiFS) chlorophyll and ocean general circulation model mixed layer depth, *J. Geophys. Res*., 110, doi:10.1029/2004JC002771.

Lewis, M.R., M.-E. Carr, G.C. Feldman, W. Esaias, C.R. McClain,1990, Influence of penetrating solar radiation on the heat budget of the equatorial Pacific Ocean, *Nature,* 347, 543-545.

L'Heureux, M. L., Collins, D. C. and Hu, Z.-Z., 2012, Linear trends in sea surface temperature of the tropical Pacific Ocean and implications for the El Niño-Southern Oscillation, *Climate Dynamics*, 1–14. doi:10.1007/s00382-012-1331-2

L'Heureux, M.L., S. Lee, and B. Lyon, 2013, Recent multidecadal strengthening of the Walker circulation across the tropical Pacific, *Nature Climate Change*, 3, doi:10.1038/NCLIMATE1840.

Liu, W., B. Huang, P.W. Thorne, V.F. Banzon, H.-M. Zhang, E. Freeman, J. Lawrimore, T.C. Peterson, T.M. Smith, S.D. Woodruff, 2015, Extended reconstructed sea surface temperature version 4 (ERSST.v4): Part II. Parametric and structural uncertainty estimations, *J. Clim*., doi:10.1175/JCLI-D-14-00007.1.

Lluch-Belda, D., R.J.M. Crawford, T. Kawasaki, A.D. MacCall, R.H. Parrish, R.A. Schwartzlose, and P.E. Smith, 1989, World-wide fluctuations of sardine and anchovy stocks: the regime problem, *S. Afr. J. Mar. Sci.*, 8: 195-205.

Lomas, M. W., Steinberg, D. K., Dickey, T., Carlson, C. A., Nelson, N. B., Condon, R. H., and Bates, N. R., 2010, Increased ocean carbon export in the Sargasso Sea linked to climate variability is countered by its enhanced mesopelagic attenuation, *Biogeosciences*, 7, 57-70, doi:10.5194/bg-7-57-2010.

Longhurst, A.L., 1998, *Ecological Geography of the Sea*, Academic Press, San Diego, pp. 398.

Luo, Y., L.M. Rothstein, and R.-H. Zhang, 2009, Response of the Pacific subtropical-thermocline water pathways and transports to global warming, *Geophys. Res. Lett.,* 36, L04601, doi:10.1029/2008GL036705.

Lutz, M. J., K. Caldeira, R. B. Dunbar, and M. J. Behrenfeld, 2007, Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean, *J. Geophys. Res.*, 112, C10011, doi:10.1029/2006JC003706.

Mackas, D.L., 2011, Does blending of chlorophyll data bias temporal trend? *Nature*, 472, doi:10.1038/nature09951.

Maes, C., J. Sudre, and V. Garcon, 2010, Detection of the eastern edge of the equatorial Pacific warm pool using satellite-based ocean color observations, *SOLA*, 6, 129-132, doi:10.2151/sola.2010-033.

Mann, M.E., J. Park, and R.S. Bradley, 1995, Global interdecadal and century-scale climate oscillations during the past five centuries, *Nature*, 378, 266-270.

Mann, M.E., R.S. Bradley, and M.K. Hughes, 1998, Global-scale temperature patterns and climate forcing over the past six centuries, *Nature*, 392, 779-787.

Mantua, N.J., S.R. Hare, Y. Zhang, J.M. Wallace and R.C. Francis, 1997, A Pacific interdecadal climate oscillation with impacts on salmon production, *Bull. Amer. Meteor. Soc.*, 78, 1069-1079.

Mantua, N.J. and S.R. Hare, 2002, The Pacific-decadal oscillation, *J. of Oceano.*, 58, 35-44.

Maritorena, S. and D. Siegel, 2005, Consistent merging of satellite ocean color data sets using a bio-optical model, *Rem. Sens. Env.* 94, 429-440.

Marshall, J. and R.A. Plumb, 2008, Atmosphere, ocean and climate dynamics: an introductory text, Elsevier Academic Press, London.

Martin, J.H. and 43 co-authors, 1994, Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean, *Nature,* 371, 1994.

Martinez, E., D. Antoine, F. D'Ortenzio, B. Gentili, 2009, Climate-driven basin-scale decadal oscillations of oceanic phytoplankton, *Science*, 326, doi:10.1126/science.1177012.

Marzeion, B., A. Timmermann, R. Murtugudde, F.-F. Jin, 2005, Biophysical feedbacks in the tropical Pacific, *J. Climate*, 18, 58-70.

Masotti, I., Moulin, C., Alvain, S., Bopp, L., Tagliabue, A., and Antoine, D.: Large-scale shifts in phytoplankton groups in the Equatorial Pacific during ENSO cycles, *Biogeosciences, 8*, 539-550, doi:10.5194/bg-8-539-2011, 2011.

McClain, C.R., R. Murtugudde, and S. Signorini, 1999, A simulation of biological processes in the equatorial Pacific warm pool at 165^{o}E, *J. Geophys. Res*., 104, C8, 18,305-18,322.

McClain, C.R., G.C. Feldman, S.B. Hooker, 2004, An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series, *Deep Sea Research II, 51*, 5-42.

McClain, C.R., 2009, A decade of satellite ocean color observations, *Ann. Rev. Mar. Sci*, 1, 19-42.

McCreary, J.P., and P. Lu, 1994, On the interaction between the subtropical and equatorial ocean circulation: the Subtropical Cell. *J. Phys. Oceanogr*., 24, 466–497.

McPhaden, M.J. and D. Zhang, 2002, Slowdown of the meridional overturning circulation in the upper Pacific Ocean, *Nature*, 603-608.

McPhaden, M.J., S.E. Zebiak, and M.H. Glantz, 2006, ENSO as an integrating concept in Earth science, *Science, 314*, 1740-1745, doi:10.1126/science.1132588.

McPhaden, M.J., T. Lee, and D. McClurg, 2011, El Niño and its relationship to changing background conditions in the tropical Pacific Ocean, *Geophys. Res. Lett.*, 38, doi:10.1029/2011GL048275.

McQuatters-Gollop, P.C. Reid, M. Edwards, P.H. Burkill, C. Castellani, S. Batten, W. Gieskes, D. Beare, R.R. Bidigare, E. Head, R. Johnson, M. Kahru, J. A. Koslow, A. Pena, 2011, Is there a decline in marine phytoplankton?, *Nature*, 472, doi:10.1038/nature09950.

Meister, G., B.A. Franz, E.J. Kwiatkowska, and C. R. McClain, 2012, Corrections to the Calibration of MODIS Aqua Ocean Color Bands Derived From SeaWiFS Data, IEEE Geo. Rem. Sens., 50 (1). doi:10.1109/TGRS.2011.2160552.

Messié, M. and M.-H. Radenac, 2006, Seasonal variability of the surface chlorophyll in the western tropical Pacific from SeaWiFS data, *Deep-Sea Res. I*, 53: 1581-1600.

Messié, M. and F. P. Chavez, 2011, Global modes of sea surface temperature variability in relation to regional climate indices, *J. Climate, 24*, 4313-4330, doi:10.1175/2011JCL13941.1.

Messié, M. and F. P. Chavez, 2012, A global analysis of ENSO synchrony: The oceans' biological response to physical forcing, *J. Geophys. Res.*, 117: C09001, doi:10.1029/2012JC007938.

Messié, M. and F. P. Chavez, 2013, Physical-biological synchrony in the global ocean associated with recent variability in the central and western equatorial Pacific, *J. Geophys. Res.*, 118, 3782-3794, doi:10.1002/jgrc.20278.

Morel, A. and J.-F. Berthon, 1989, Surface pigments, algal biomass profiles, and potential production of the euphotic layer: relationships reinvestigated in view of remote-sensing applications, *Limnol. Oceanogr., 34(8)*: 1545-1562.

Moulin, C., H.R. Gordon, V.F. Banzon, and R.H. Evans, 2001, Assessment of Saharan dust absorption in the visible from SeaWiFS imagery, *J. Geophys. Res.,* 106, 18239-18250.

Murray, A.P., C.F. Gibbs, and A. R. Longmore, 1986, Determination of chlorophyll in marine waters: intercomparison of a rapid HPLC method with full HPLC, spectrophotometric and fluorometric methods, *Mar. Chem., 19*: 211-227.

Murtugudde, R., J. Beauchamp, and A. Busalacchi, 2002, Effects of penetrative radiation on the upper tropical ocean circulation, *J. Climate*, 15, 470-486.

NASA GSFC Ocean Color Documents, 2011, SeaWiFS Reprocessing 2010.0. Available at: http://oceancolor.gsfc.nasa.gov/WIKI/OCReproc20100SW.html (accessed 12 October 2011).

O'Reilly, J.E., S. Maritorena, B.G. Mitchell, D.A. Siegel, K.L. Carder, S.A. Garver, M. Kahru, and C. McClain, 1998, Ocean color chlorophyll algorithms for SeaWiFS, *Marine Science Faculty Publications*, Paper 6.

Overland, J.E. and R.W. Preisendorfer, 1982, A significance test for principal components applied to a cyclone climatology, *Mon. Wea. Rev*. 110, 1-4.

Overland, J., S. Rodionov, s.Minobe, and N. Bond, 2008, North Pacific regime shifts: definitions, issues and recent transitions, *Prog. in Oceano.*, 77, 92-102, doi:10.1016/j.pocean.2008.03.016.

Palter, J.B., M.S. Lozier, R.T. Barber, 2005, The effect of advection on the nutrient reservoir in the North Atlantic subtropical gyre, *Nature*, 437, doi:10.1038/nature03969.

Pennington, J.T., K.L. Mahoney, V.S. Kuwahara, D.D. Kolber, R. Calienes, and F.P. Chavez, 2006, Primary production in the eastern tropical Pacific: A review, *Prog. in Oceano.,* 69, 285-317, doi:10.1016/j.pocean.2006.03.012.

Philander, S.G., 1990, *El Niño, La Niña, and the southern oscillation,* Academic Press, New York, NY, 289 pp.

Philander, S.G., 1999, A review of tropical ocean-atmosphere interactions, *Tellus,* 51 A-B, 71-90.

Picaut, J., M. Ioualalen, C. Menkes, T. Delcroix, and M.J. McPhaden, 1996, Mechanism of the zonal displacements of the Pacific warm pool: implications for ENSO, *Science,* 274:5292, 1486-1489.

Picaut, J., F. Masia, and Y. du Penhoat, 1997, An advective-reflective conceptual model for the oscillatory nature of the ENSO, *Science*, 277:5326, 663-666.

Qiu, B. and S. Chen, 2012, Multidecadal sea level and gyre circulation variability in the northwestern tropical Pacific Ocean. *J. Phys. Oceanogr.*, 42, 193–206. doi: 10.1175/JPO-D-11-061.1.

Racault, M.-F., C. Le Quere, E. Buitenhuis, S. Sathyendranath, T. Platt, 2012, Phytoplankton phenology in the global ocean, *Ecological Indicators,* 14, 152-163.

Radenac, M.-H. and M. Rodier, 1996, Nitrate and chlorophyll distributions in relation to thermohaline and current structures in the western tropical Pacific during 1985-1989, *Deep-Sea Res. II*, 43:4-6, 725-752.

Radenac, M.-H., F. Leger, A. Singh, and T. Delacroix, 2012, Sea surface chlorophyll signature in the tropical Pacific during eastern and central Pacific ENSO events, *J. Geophys. Res.,* 117, C04007, doi:10.1029/2011JC007841.

Raitsos, D.E., P.C. Reid, S.J. Lavender, M. Edwards, A.J. Richardson, 2005, Extending the SeaWiFS chlorophyll data set back 50 years in the northeast Atlantic, *Geophys. Res. Lett.*, 32, L06603, doi:10.1029/2005GL022484.

Ralph, E.A. and P.P. Niiler, 1999, Wind-driven currents in the tropical Pacific, *J. Phys. Oceano.*, 29, 2121-2129.

Robock, A., 2002, Volcanic eruption, El Chichon, in Vol 1, The Earth system: physical and chemical dimensions of global environmental change, , pp 736-736, Edited by M.C. MacCracken and J.S. Perry in Encyclopedia of Global Environmental Change (ISBN 0-471-97796-9), Editor-in-Chief T. Munn, John Wiley & Sons, Ltd, Chichester.

Rodier, M. G. Eldin and R. Le Borgne, 2000, The western boundary of the equatorial Pacific upwelling: some consequences of climatic variability on hydrological and planktonic properties, *J. Oceano.,* 56, 463-471.

Rousseaux, C. S., and W. W. Gregg, 2012, Climate variability and phytoplankton composition in the Pacific Ocean, *J. Geophys. Res.,* 117, C10006, doi:10.1029/2012JC008083.

Rykaczewski, R.R. and J.P. Dunne, 2011, A measured look at ocean chlorophyll trends, *Nature*, 472, doi:10.1038/nature09952.

Saba, V.S., M.A.M. Friedrichs, M.-E. Carr, D. Antoine, R.A. Armstrong, I. Asanuma, O. Aumont, N.R. Bates, M.J. Behrenfeld, V. Bennington, L. Bopp, J. Bruggeman, E.T. Buitenhuis, M.J. Church, A.M. Ciotti, S.C. Doney, M. Dowell, J. Dunne, S. Dutkiewicz, W. Gregg, N. Hoepffner, K.J.W. Hyde, J. Ishizaka, T. Kameda, D.M. Karl, I. Lima, M.W. Lomas, J. Marra, G.A. McKinley, F. Melin, J.K. Moore, A. Morel, J. O'Reilly, B. Salihoglu, M. Scardi, T.J. Smyth, S. Tang, J. Tjiputra, J. Uitz, M. Vichi, K. Waters, T.K. Westberry, A. Yool, 2010, Challenges of modeling depth-integrated marine primary productivity over multiple decades: A case study at BATS and HOT, *Global Biogeochem. Cycles*, 24:3, doi:10.1029/2009gb003655.

Sabine, C.L., R.A. Feely, N. Gruber, R.M. Key, K. Lee, J.L. Bullister, R. Wanninkhof, C.S. Wong, D.W.R. Wallace, B. Tilbrook, F.J. Millero, T.H. Peng, A. Kozyr, T. Ono, A.F. Rios, 2004, The oceanic sink for anthropogenic CO_2, *Science*, 305, 5682, doi:10.1126/science.1109620.

Sapiano, M.R., C.W. Brown, S. Schollaert Uz, M. Vargas, 2012, Establishing a global climatology of marine phytoplankton phenological characteristics, *J. Geophys. Res. Oceans*, doi:10.1029/2012JC007958.

Sarmiento, J.L., T.M.C. Hughes, R.J. Stouffer, S. Manabe, 1998, Simulated response of the ocean carbon cycle to anthropogenic climate warming, *Nature*, 393, 245-249.

Schlesinger, M.E. and N. Ramankutty, 1994, An oscillation in the global climate system of period 65-70 years, *Nature,* 367, 723-726.

Schollaert, S.E., J.A. Yoder, J.E. O'Reilly, and D.L. Westphal, 2003, Influence of dust and sulfate aerosols on ocean color spectra and chlorophyll a concentrations derived from SeaWiFS off the U.S. east coast, *J. Geophys. Res*., 108:3191, doi:10.1029/2000JC000555.

Schollaert, S.E., T. Rossby, and J.A. Yoder, 2004, Gulf Stream cross-frontal exchange: possible mechanisms to explain interannual variations in phytoplankton chlorophyll in the Slope Sea during the SeaWiFS years, *Deep Sea Res. II*, 51, 173-188, doi:10.1016/j.dsr2.2003.07.017.

Schwartzlose, R.A., J. Alheit, A. Bakun, T.R. Baumgartner, R. Cloete, R.J.M. Crawford, W.J. Fletcher, Y. Green-Ruiz, E. Hagen, T. Kawasaki, D. Lluch-Belda, S.E. Lluch-Cota, A.D. MacCall, Y. Matsuura, M.O. Nevarez-Martinez, R.H. Parrish, C. Roy, R. Serra, K.V. Shust, M.N. Ward, and J.Z. Zuzunaga, 1999, Worldwide Large-scale Fluctuations of Sardine and Anchovy Populations. *South African Journal of Marine Science*, 21, 289-347.

Sen Gupta, A., A. Ganachaud, S. McGregor, J.N. Brown and L. Muir, 2012, Drivers of the projected changes to the Pacific Ocean equatorial circulations, *Geophys. Res. Lett.*, 39, L09605,doi:10.1029/2012GL051447.

Shimada, M., H. Oaku, Y. Mitomi, H. Murakami, A. Mukaida, Y. Nakamura, J. Ishizaka, H. Kawamura, T. Tanaka, M. Kishino, and H. Fukushima, 1998, Calibration and validation of the ocean color version 3 products from ADEOS OCTS, *J. Oceanogr.*, 54, 401-416.

Schulz, M., M. Prange and A. Klocker, 2007, Low-frequency oscillations of the Atlantic Ocean meridional overturning circulation in a coupled climate model, *Clim. Past*, 3, 97-107, doi:10.5194/cp-3-97-2007.

Slemons, L.O., J.W. Murray, J. Resing, B. Paul and P. Dutrieux, 2010, Western Pacific coastal sources of iron, manganese, and aluminum to the Equatorial Undercurrent, *Global Biogeochem. Cycles*, 24, GB3024, doi:10.1029/2009GB003693.

Smerdon, J.E., A. Kaplan, D. Chang and M.N. Evans, 2011, Corrigendum: a pseudoproxy evaluation of the CCA and RegEM methods for reconstructing

climate fields of the last millennium, *J. Climate*, 24, 1284-1309, doi:10.1175/2010JCL14110.1.

Smerdon, J.E., A. Kaplan, E. Zorita, J.F. Gonzalez-Rouco and M.N. Evans, 2011, Spatial performance of four climate field reconstruction methods targeting the Common Era, *Geophys. Res. Lett.*, 38, L11705, doi:10.1029/2011GL047372.

Smith, T.M., A.G. Barnston, M. Ji, and M. Chelliah, 1995, The impact of Pacific ocean subsurface data on operational prediction of tropical Pacific SST at the NCEP, *Weather and Forecasting*, 10, 708-714.

Smith, T.M., R.W. Reynolds, R.E. Livezey, D.C. Stokes, 1996, Reconstruction of historical sea surface temperatures using empirical orthogonal functions, *J. Climate*, 9, 1403-1420.

Smith, T.M. and R.E. Livezey, 1998, An improved method for analyzing sparse and irregularly distributed SST data on a regular grid: the tropical Pacific Ocean, *J. Climate*, 11, 1718-1729.

Smith, T.M. and R.W. Reynolds, 2005, A global merged land-air-sea surface temperature reconstruction based on historical observations (1880-1997), *J. Climate*, 18, 2021-2036.

Smith, T.M., P.A. Arkin, M.R.P. Sapiano, 2009, Reconstruction of near-global annual precipitation using correlations with sea surface temperature and sea level pressure, *J. Geophys. Res*., 114, doi:10.1029/2008JD011580.

Subramaniam, A., C.W. Brown, R.R. Hood, E.J. Carpenter and D.G. Capone, 2002, Detecting *Trichodesmium* blooms in SeaWiFS imagery, *Deep Sea Res. II*, 49/1-3, 107-121.

Sverdrup, H.U., 1953, On conditions for the vernal blooming of phytoplankton, *J. Cons. Int. Explor. Mer*, 18, 287-295.

Takasuka, A., Y. Oozeki, H. Kubota and S.E. Lluch-Cota, 2008, Contrasting spawning temperature optima: why are anchovy and sardine regime shifts synchronous across the North Pacific? *Prog. in Oceanogr.*, 77, 225-232.

Talley L.D., Pickard G.L., Emery W.J., Swift J.H., 2011, *Descriptive Physical Oceanography: An Introduction (Sixth Edition)*, Elsevier, Boston, 560 pp.

Trenberth, K.E., 1997, The definition of El Niño, *Bull. Amer. Meteor. Soc.*, 78, 2771-2777.

Trenberth, K.E. and D.P. Stepaniak, 2001, Indices of El Niño Evolution, *J. Climate*, 14, 1697-1701.

Tsuchiya, M., 1972, A subsurface north equatorial countercurrent in the eastern pacific ocean*, J. Geophys. Res.,* 77*(*30*),* 5981–5986, doi*:*10.1029/JC077i030p05981

Turk, D., M.J. McPhaden, A.J. Busalacchi, and M.R. Lewis, 2001, Remotely sensed biological production in the equatorial Pacific, *Science*, 293, 471 doi:10.1126/science.1056449.

Turk, D., C.S. Meinen, D. Antoine, M.J. McPhaden, and M.R. Lewis, 2011, Implications of changing El Niño patterns for biological dynamics in the equatorial Pacific Ocean, *Geophys. Res. Lett.*, 38, L23603, doi:10.1029/2011GL049674.

Uitz, J., H. Claustre, A. Morel, and S.B. Hooker, 2006, Vertical distribution of phytoplankton communities in open ocean: an assessment based on surface chlorophyll, *J. Geophys. Res., 111*, C08005, doi:10.10292005JC003207.

Uitz, J., H. Claustre, B. Gentili, and D. Stramski, 2010, Phytoplankton class-specific primary production in the world's oceans: seasonal and interannual variability from satellite observations, *Global Biogeochem. Cycles*, 24, doi:10.1029/2009GB003680.

Uz, B.M. and J.A. Yoder, 2004, High frequency and mesoscale variability in SeaWiFS chlorophyll imagery and its relation to other remotely sensed oceanographic variables, *Deep-Sea Res. II*, 51, 1001-1017, doi:10.1016/j.dsr2.2004.03.003.

Uz, B.M., 2006, Argo floats complement biological remote sensing, *Eos Trans. AGU*, 87: 32, 313-320, doi:10.1029/2006EO320002.

Vecchi, G.A. and D.E. Harrison, 2006, The termination of the 1997-98 El Niño. Part I: mechanisms of oceanic change, *J. Climate*, 19, 2633-2646.

Vernon, L.P.,1960, Spectrophotometric determination of chlorophylls and phaeophytins in plant extracts, *Anal. Chem.*, 32.

Walker, G.T. Correlation in seasonal variations of weather, IX. A further study of world weather, *Mem. India Meteor. Dept. 24*, 275-333.

Walsh, J.J., 1980, Concluding remarks: marine photosynthesis and the global carbon cycle. *In:* Primary productivity in the sea, (P.G. Falkowski, Ed.), Plenum Press, New York, pp. 497-506.

Wang, M. 2003, Correction of artifacts in the SeaWiFS atmospheric correction: removing discontinuity in the derived products. *Rem. Sens. Env.,* 84**,** 603-611.

Wang, X.J., J.R. Christian, R. Murtugudde, A.J. Busalacchi, 2005, Ecosystem dynamics and export production in the central and eastern equatorial Pacific: a modeling study of impact of ENSO. *Geophys. Res. Lett.*, 32, doi:10.1029/2004GL021538.

Wang, X.J., R. Le Borgne, R. Murtugudde, A.J. Busalacchi and M. Behrenfeld, 2008, Spatial and temporal variations in dissolved and particulate organic nitrogen in the equatorial Pacific: biological and physical influences, *Biogeosciences*, 5, 1705-1721.

Wang X.J., Behrenfeld, M., Le Borgne, R., Murtugudde, R. and Boss, E., 2009. Regulation of phytoplankton carbon to chlorophyll ratio by light, nutrients and temperature in the equatorial Pacific Ocean: a basin-scale model. Biogeosciences, 6, 391-404.

Wang, X.J., R. Le Borgne, R. Murtugudde, A.J. Busalacchi, and M. Behrenfeld, 2009, Spatial and temporal variability of the phytoplankton carbon to chlorophyll ratio in the equatorial Pacific: a basin scale model study, *J. Geophys. Res*., 114, C07008, doi:10.1029/2008JC004942.

Weare, B.C. and J.S. Nasstrom, 1982, Examples of extended empirical orthogonal function analyses. *Mon. Wea. Rev.*, 110, 481-485.

Werdell, P.J., C.W. Proctor, E. Boss, T. Leeuw, and M. Ouhssain, 2013, Underway sampling of marine inherent optical properties on the Tara Oceans expedition as a novel resource for ocean color satellite data product validation, *Methods in Oceanography*, doi:10.1016/j.mio.2013.09.001

Williams, R.G. and M.J. Follows, 2003, Physical transport of nutrients and the maintenance of biological production, 19-51, *In*: Ocean Biogeochemistry: The role of the ocean carbon cycle in global change. Edited by M.J.R. Fasham. Springer. ISBN: 3-540-42398-2.

Wilson, C. and D. Adamec, 2001, Correlations between surface chlorophyll and sea surface height in the tropical Pacific during the 1997-1999 El Niño-Southern Oscillation event, *J. Geophys. Res*., 106, C12, 31175-31188, 2000JC000724.

Yeh, S.-W., J.S. Kug, B. Dewitte, M.-H. Kwon, B.P. Kirtman, and F.-F. Jin, 2009, El Niño in a changing climate, *Nature*, 461, 511-514, doi:10.1038/nature08316.

Yentsch, C. S., 1960, The influence of phytoplankton pigments on the color of sea water, *Deep-Sea Res.*, 7, 1-9.

Yentsch, C.S. and D.W. Menzel, 1963, A method for the determination of phytoplankton chlorophyll and phaeophytin by flouresccence. *Deep-Sea Res.*, 10, 221-231.

Yentsch, C.S. and C.M. Yentsch, 1979, Fluorescence spectral signatures: The characterization of phytoplankton populations by the use of excitation and emission spectra, *J. of Mar. Res.,* 37, 471-483.

Yoder, J.A., M.A. Kennelly, S.C. Doney, I.D. Lima, 2010, Are trends in SeaWiFS chlorophyll time-series unusual relative to historic variability?, *Acta Oceanologica Sinica*, 29:2, doi:10.1007/s13131-010-0001-7.

Yu, J.-Y., Y. Zou, S.T. Kim, and T. Lee, 2012, The changing impact of El Niño on US winter temperatures, *Geophys. Res. Lett.*, 39:L15702, doi:10.1029/2012GL052483.

Ziauddin, A., B.A. Franz, C.R. McClain, E.J. Kwiatkowska, J. Werdell, E.P. Shettle, and B.N. Holben, 2010, New aerosol models for the retrieval of aerosol optical thickness and normalized water-leaving radiances from the SeaWiFS and MODIS sensors over coastal regions and open oceans. *Applied Optics*, 49: 29, 5545-5560, doi:10.1364/AO.49.005545.