INT. J. REMOTE SENSING, 1999, VOL. 20, NO. 17, 3423–3429

Empirical chlorophyll algorithm and preliminary SeaWiFS validation for the California Current

M. KAHRU† and B. G. MITCHELL‡

Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093-0218, USA

†Tel: + 1-858-534-8947. Fax: + 1-858-822-0562.
E-mail address: mkahru@ucsd.edu

‡Tel: + 1-858-534-8947. Fax: + 1-858-822-0562.
E-mail address: gmitchell@ucsd.edu

(*Received 10 November 1998; in final form 17 May 1999*)

Abstract. A new empirical chlorophyll algorithm is proposed for SeaWiFS (Sea-viewing Wide Field-of-view Sensor) and other ocean colour sensors. The CAL-P6 algorithm uses a sixth-order polynomial of the ratio of normalized water leaving radiances (L_{WN}) at 490 nm and 555 nm and is based on 348 measurements of L_{WN} and chlorophyll-*a* in the California Current. Validation of the SeaWiFS-derived chlorophyll values with 27 concurrent *in situ* measurements showed high correlation ($r^2 = 0.93$ in the log–log space) but significant overestimation by SeaWiFS at high chlorophyll-*a* concentration. The problem was traced to significant underestimation of the SeaWiFS-derived L_{WN} (490) at high chlorophyll-*a* concentration (3–5 mg m^{-3}). Further refinement of the atmospheric correction is needed for SeaWiFS to attain its goal of 35% accuracy for chlorophyll retrieval in the coastal zone.

1. Introduction

O'Reilly *et al.* (1998) have described the OC2 chlorophyll algorithm that is used by the National Aeronautics and Space Administration (NASA) in the operational processing of SeaWiFS data (Fu *et al.* 1998). The OC2 algorithm uses the ratio of remote sensing reflectances (R_{rs}) at 490 nm and 555 nm (Aiken *et al.* 1995) to estimate chlorophyll-*a* concentration. The OC2 coefficients were derived by a statistical fit to a dataset of 919 bio-optical measurements (the SeaBAM dataset). The SeaBAM data are heterogeneous due to both methodological and natural variability. The methodological variability is caused by merging measurements from different instruments, both below surface and above surface, processed with variable procedures. Uncertainty in the adjustments to convert different spectral bands to SeaWiFS bands, a lack of correction for instrument self-shading, variable radiometric calibrations and surface extrapolation procedures, as well as differences in methods for chlorophyll-*a* determination, contribute to the methodological variability of the dataset. The SeaBAM dataset included chlorophyll-*a* measurements by both fluorometric and high-performance liquid chromatography methods that have systematic differences (see references in Aiken *et al.* (1995)). Natural variability due to merging

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/JNLS/res.htm
http://www.taylorandfrancis.com/JNLS/res.htm

measurements from different bio-geographic domains may not be desirable when developing a regional algorithm.

The sigmoid pattern of the OC2 algorithm is very sensitive to small variations at low $R_{rs}(490)/R_{rs}(555)$, producing unrealistically high chlorophyll estimates in cases of high gelbstoff, detrital and/or accessory pigment absorption. For example, in a red tide off California (Kahru and Mitchell 1998) the OC2 algorithm predicted $536\,\mathrm{mg\,m^{-3}}$ of chlorophyll instead of the measured $32.5\,\mathrm{mg\,m^{-3}}$, i.e. and overestimation by more than 16 times. The SeaBAM dataset had only nine points in the high chlorophyll region above $10\,\mathrm{mg\,m^{-3}}$. Six of those were from high latitudes (Cota 1997) with known differences in the bio-optical relationships (Mitchell 1992). In the California Current, very high chlorophyll ($>15\,\mathrm{mg\,m^{-3}}$) is almost invariably due to red tides that have high soluble absorption (Kahru and Mitchell 1998). In the intermediate chlorophyll region between 1 and $10\,\mathrm{mg\,m^{-3}}$, the OC2 algorithm significantly underestimates chlorophyll-*a* in the California Current data. These considerations warranted the development of a regional chlorophyll algorithm. In August 1998 NASA announced a revised version of the OC2 algorithm (C. McClain 1998, personal communication). The new OC2-v2 algorithm (S. Maritorena 1998, personal communication) eliminates the dramatic overestimation at high concentrations but accentuates the underestimation in the intermediate chlorophyll range for the California Current area. Here we present an empirical algorithm developed with our California Current data, and compare it to the OC2 and OC2-v2 algorithms. A preliminary validation is performed for SeaWiFS-derived L_{WN} and chlorophyll-*a* values.

2. Data and methods

Vertical profiles of downwelling spectral irradiance and upwelling radiance were measured with underwater radiometers (Biospherical Instruments MER-2040 and MER-2048) as part of the California Cooperative Oceanic Fisheries Investigations (CalCOFI) bio-optical program (Mitchell and Kahru 1998). About one-third of the SeaBAM dataset originated from CalCOFI. Since the submission of the CalCOFI data to SeaBAM, more measurements have been acquired ($n = 461$) and improved processing procedures have been implemented. Instrument self-shading correction (Gordon and Ding 1992, Kahru and Mitchell 1998) was routinely applied and profiles with the ship shadow and/or variable illumination were eliminated. Near-shore stations with increased reflectance due to suspended sediments were also excluded. The remaining dataset consisted of 348 stations. The advantage of this dataset compared to the heterogeneous SeaBAM dataset is that it has been collected with the same well calibrated instruments and processed using similar procedures. Measurements of chlorophyll-*a* were taken in the CalCOFI program using the fluorometric method (Venrick and Hayward 1984) and consistent calibration protocols. At the same time our combined CalCOFI and red tide dataset encompasses diverse conditions ranging from the oligotrophic offshore stations to highly eutrophic red tide events (chlorophyll-*a* between 0.05 and $32.5\,\mathrm{mg\,m^{-3}}$).

The CAL-P6 algorithm described here is formulated as a function of the ratio of the standard SeaWiFS products $L_{WN}(490)$ and $L_{WN}(555)$. The OC2 and OC2-v2 algorithms use $R_{rs}(490)/R_{rs}(555)$. The difference between ratios $L_{WN}(490)/L_{WN}(555)$ and $R_{rs}(490)/R_{rs}(555)$ is approximately 4%. The statistical and graphical criteria used to evaluate the agreement between the measured and modelled values were similar to those described by O'Reilly *et al.* (1998).

3. Results

Formulations of the OC2, OC2-v2 and CAL-P6 algorithms are given in table 1. When applied to the CalCOFI data (figure 1) OC2 overestimates chlorophyll-*a* at high chlorophyll and underestimates elsewhere (with the exception of the extremely low chlorophyll-a concentration). The revised OC2-v2 algorithm performs better at high chlorophyll but underestimates at intermediate concentrations. The rms error of OC2-v2 actually increased compared to OC2 (table 2) when applied to our dataset.

The CAL-P6 algorithm uses a sixth-order polynomial to approximate the observed pattern of the chlorophyll-*a* concentration versus $L_{WN}(490)/L_{WN}(555)$ relationship in a log–log space. Although we used our CalCOFI data to develop the CAL-P6 algorithm, it does not provide a statistical best fit to this dataset. We added constraints at the data extremes with the aim of making the algorithm applicable to data slightly outside the observed CalCOFI data range. In the very low chlorophyll-*a* range, the modelled $L_{WN}(490)/L_{WN}(555)$ ratio was forced to approach clear water values (5.6 at 0.03 mg Chl-a m^{-3}) produced by semi-analytical models (Gordon *et al.* 1988). At high chlorophyll concentrations, in order to achieve a compromise between red tides and upwelling blooms, the chlorophyll-*a* versus $L_{WN}(490)/L_{WN}(555)$ curve was forced to higher chlorophyll-*a* values, reducing the effect of a few very high chlorophyll-*a* values from a red tide. Variations in the relative amounts of soluble, detrital and phytoplankton absorption and the magnitude of pigment packaging render a single variable like the $L_{WN}(490)/L_{WN}(555)$ ratio insufficient to determine bio-optical properties accurately at high chlorophyll-*a* concentrations. Due to these manipulations at both low and high concentration extremes, it is expected that the CAL-P6 algorithm is applicable to chlorophyll-*a* concentrations between 0.02 and 50 mg m^{-3} and $L_{WN}(490)/L_{WN}(555)$ ratios greater than 0.26. When applied to the SeaBAM dataset, CAL-P6 produces r^2 values comparable to those from OC2 and OC2-v2 but the rms error is slightly higher (table 2). Quantile–quantile plots of the difference between the measured and modelled chlorophyll-*a* values (figure 1) accentuate the differences between the algorithms. It is evident that while the OC2-v2 algorithm performs better for high chlorophyll-*a* concentrations, it underestimates over most of the chlorophyll *a* concentration range.

Validation of the chlorophyll algorithms was carried out by comparing L_{WN} and chlorophyll-*a* derived from SeaWiFS images with *in situ* data collected concurrently (± 4 hours). The SeaWiFS data were processed to L_{WN} and chlorophyll-*a* using

Table 1. Formulations of the empirical chlorophyll algorithms OC2, OC2-v2 and CAL-P6. The band ratio R is defined as $R = \log[R_{rs}(490)/R_{rs}(555)]$ for OC2 and OC2-v2 and $R = \log[L_{WN}(490)/L_{WN}(555)]$ for CAL-P6.

Algorithm	Type	Equation coefficients (a)	Reference
OC2	Modified cubic polynomial	$\mathrm{Chl} = 10^{(a0 + a1R + a2R^2 + a3R^3)} + a^4$ $a = [0.341, -3.001, 2.811, -2.041, -0.04]$	O'Reilly *et al.* 1998
OC2-v2	Modified cubic polynomial	$\mathrm{Chl} = 10^{(a0 + a1R + a2R^2 + a3R^3)} + a^4$ $a = [0.2974, -2.2429, 0.8358, -0.0077, -0.0929]$	Maritorena, pers. comm.
CAL-P6	Sixth-order polynomial	$\mathrm{Chl} = 10^{(a0 + a1R + a2R^2 + a3R^3 + a4R^4 + a5R^5 + a6R^6)}$ $a = [0.565, -2.561, -1.051, -0.294, 5.561, 3.130, -10.816]$	This work

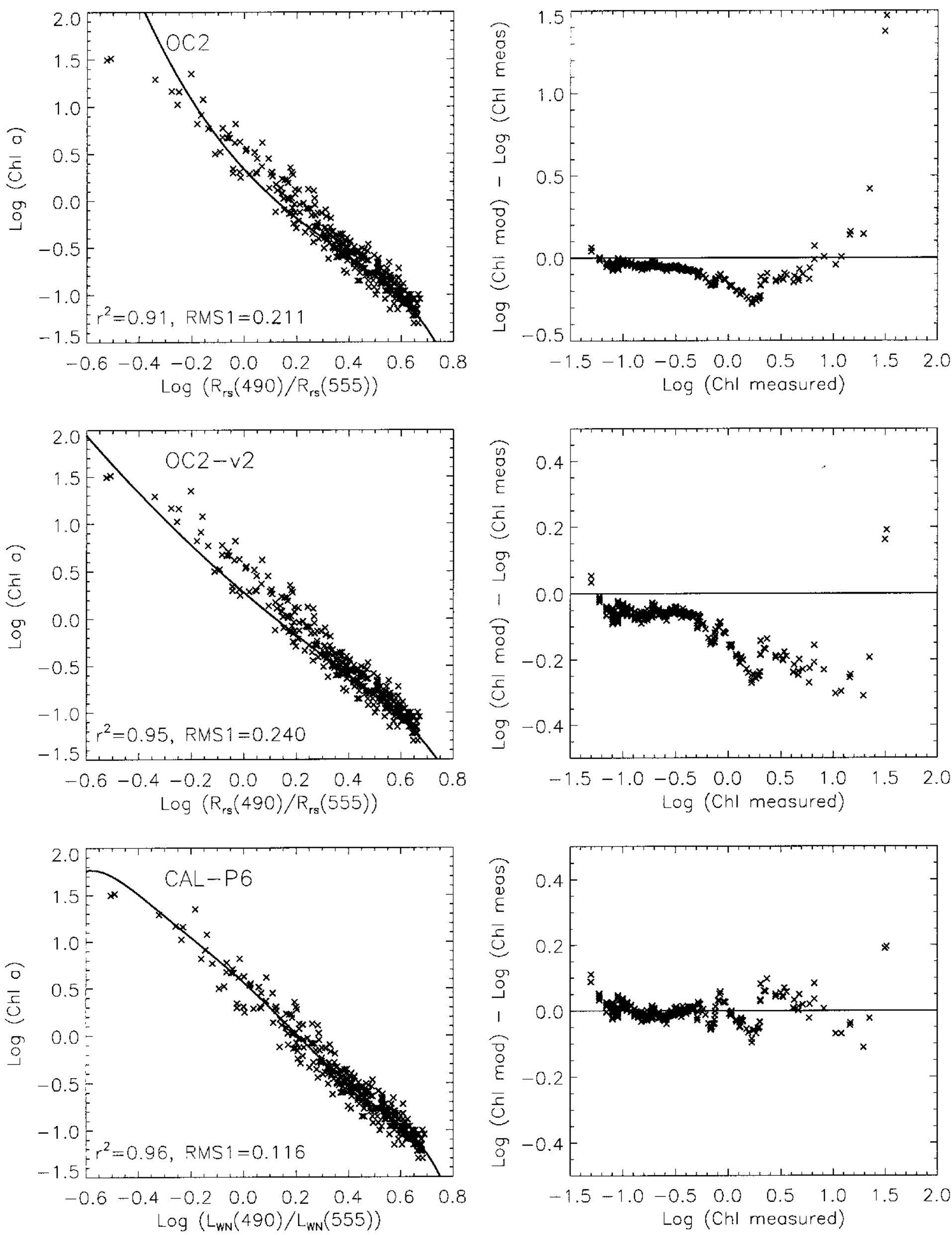


Figure 1. A comparison of the SeaWiFS at-launch chlorophyll algorithm OC2 (O'Reilly *et al.* 1998, upper panels), the revised OC2-v2 (centre panels) and CAL-P6 algorithms (lower panels) as applied to our CalCOFI *in situ* data. The left panels show scatter plots of the measured chlorophyll-*a* concentration versus band ratio. The right panels are quantile–quantile plots of the differences between modeled and measured chlorophyll values (note the difference in scale of the OC2 panel).

SeaDAS 3.2 software (Fu *et al.* 1998). A total of 27 matching chlorophyll-*a* values and 19 matching sets of L_{WN} were found between 2 October 1997 and 24 September 1998. Satellite values were derived as averages over 3 pixel × 3 pixel areas centred at

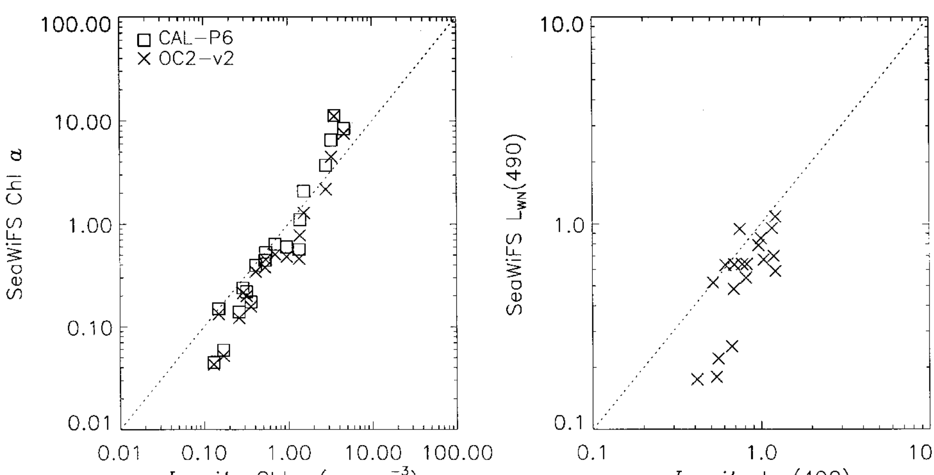
Table 2. Comparison between the OC2, OC2-v2 and the CAL-P6 chlorophyll algorithms as applied to the CalCOFI and SeaBAM datasets.

Algorithm	Dataset	*n*	Intercept	Slope	r^2	rms	Bias
OC2	CalCOFI	348	0.04	0.96	0.91	0.21	– 0.12
OC2-v2	CalCOFI	348	0.14	1.09	0.95	0.24	– 0.18
CAL-P6	CalCOFI	348	0.00	1.00	0.96	0.12	0.00
OC2	SeaBAM	919	0.00	1.00	0.92	0.17	0.00
OC2-v2	SeaBAM	919	0.06	1.06	0.91	0.19	– 0.04
CAL-P6	SeaBAM	919	– 0.10	0.95	0.92	0.23	0.14

the *in situ* measurement. All three chlorophyll algorithms produced approximately similar results (table 3). In general, they overestimated chlorophyll-*a* at higher concentrations and underestimated it at lower concentrations (figure 2(*a*)), resulting in regression slopes over 1.3. Whereas the CAL-P6 algorithm produced slightly higher r^2 values and lower rms error, it also had the highest slope. The origin of these deviations was traced to the bias in SeaWiFS-derived L_{WN} (490) (figure 2(*b*)) and L_{WN} (555) that was typically lower than in the *in situ* values. Taking a ratio of the two radiances eliminated most of the bias, but did not correct for the most severe

Table 3. Results of the match-ups between SeaWiFS chlorophyll-*a* concentration derived with the OC2, OC2-v2 and CAL-P6 algorithms versus *in situ* surface chlorophyll-*a* (both log-transformed).

Algorithm	*n*	Intercept	Slope	r^2	rms	Bias
OC2	27	– 0.07	1.38	0.90	0.31	– 0.15
OC-v2	27	– 0.05	1.32	0.92	0.27	– 0.11
CAL-P6	27	0.05	1.38	0.93	0.26	– 0.02



(a) (b)

Figure 2. A comparison of the SeaWiFS-derived and *in situ* measurements of (*a*) chlorophyll-*a* and (*b*) normalized water-leaving radiance L_{WN} at 490 nm. The chlorophyll-*a* values were calculated with the OC2-v2 (×) and CAL-P6 (□) algorithms.

($\sim$60%) underestimation of L_{WN} (490) at high chlorophyll-*a* concentrations (3–5 mg m^{-3}). The underestimation in the SeaWiFS-derived L_{WN} was more evident at high chlorophyll and shorter wavelengths (figure 3).

4. Discussion and conclusions

It is not surprising that the CAL-P6 algorithm developed with the CalCOFI dataset outperforms OC2 and OC2-v2 algorithms when evaluated with the CalCOFI dataset. The OC2 algorithms were tuned to a heterogeneous dataset (including a limited version of the CalCOFI dataset). However, the OC2 algorithms seem to be biased when applied to the CalCOFI data. The original OC2 algorithm severely overestimates chlorophyll-*a* at high concentrations which is probably due to a bias towards high-latitude data. In the California Current, chlorophyll-*a* concentrations exceeding 15 mg m^{-3} occur predominantly in red tide blooms. Red tides can occur every year and in numerous locations, and therefore a regional algorithm should be able cope with these cases and not produce unrealistically high values. Based on our *in situ* data, NASA's OC2-v2 algorithm is expected to underestimate over most of the chlorophyll *a* range. The CAL-P6 algorithm proposed here can be used to calculate chlorophyll-*a* from the L_{WN} values of SeaWiFS or a similar sensor. The source code compatible with SeaDAS can be downloaded at website ftp://spg.ucsd.edu/pub/chl_algorithm/chlor_calp6.c.

When comparing satellite-derived chlorophyll with *in situ* measured chlorophyll data, other parts of the processing scheme (notably atmospheric correction) become dominant over the bio-optical algorithm. As most of the radiance measured by a satellite sensor at the top of the atmosphere originates from the atmosphere and not from the ocean (see e.g. Gordon *et al.* 1988), small errors in the atmospheric correction can result in large errors in the derived in-water properties. A preliminary validation

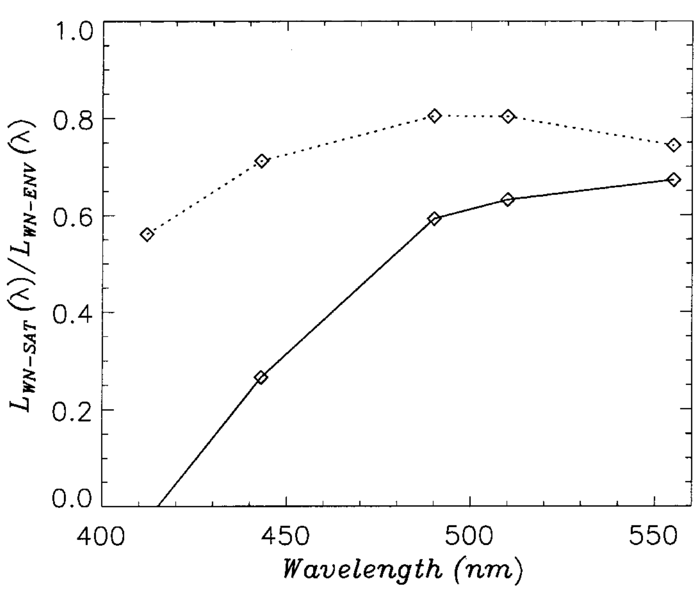


Figure 3. Average spectral shapes of the ratio SeaWiFS-retrieved L_{WN} to *in situ* L_{WN} at chlorophyll-*a* < 1 mg m^{-3} (dotted curve, $n = 12$) and at chlorophyll-*a* ≥ 1 mg m^{-3} (full curve, $n = 7$).

using CalCOFI data revealed significant underestimation of the SeaWiFS-retrieved L_{WN} at high chlorophyll-*a* stations compared to *in situ* measurements. Those errors were more severe at shorter wavelengths. The overall correlation between SeaWiFS-derived and *in situ* chlorophyll-*a* values was good and similar for different algorithms. More accurate estimates of L_{WN} will be required before the benefits of the CAL-P6 algorithm can be realized in actual SeaWiFS applications.

Acknowledgments

This work was supported by the NASA Ocean Biogeochemistry Program (NAGW-3665), NASA GSFC and NASA SIMBIOS (NAS5-97130) programs and NOAA (NA37GP0526). We thank the CalCOFI technicians for their provision of the CalCOFI data, Brian Schieber for help in implementing the match-up software, the SeaWiFS Project and the Distributed Active Archive Center at the Goddard Space Flight Center, NASA, for the production and distribution of the SeaWiFS data.

References

Aiken, J., Moore, G. F., Trees, C. C., Hooker, S. B., and Clark, D. K., 1995, The SeaWiFS CZCS-type pigment algorithm. NASA Technical Memorandum 104566, SeaWiFS Technical Report Series, Vol. 29, NASA, Greenbelt, MD, USA.

Cota, G. F., 1997, Bio-optical properties of Arctic coastal waters. *Aquatic Sciences Meeting, Santa Fe, New Mexico, 10–14 February 1997* (Waco, Texas: The American Society of Limnology and Oceanography), p. 133.

Fu, G., Settle, K., and McClain, C. R., 1998, SeaDAS: The SeaWiFS Data Analysis System. *Proceedings of the Fourth Pacific Ocean Remote Sensing Conference, Qingdao, China, 28–31 July 1998* (Beijing, China: PORSEC'98 Secretariat), pp. 73-77.

Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., and Baker, K. S., 1988, A semianalytic radiance model of ocean color. *Journal of Geophysical Research*, **93**, 10 909–10 924.

Gordon, H. R., and Ding, K., 1992, Self-shading of in-water optical instruments. *Limnology and Oceanography*, **37**, 491–500.

Kahru, M., and Mitchell, B. G., 1998, Spectral reflectance and absorption of a massive red tide off southern California. *Journal of Geophysical Research*, **103**, 21 601–21 610.

Mitchell, B. G., 1992, Predictive bio-optical relationships for polar oceans and marginal ice zones. *Journal of Marine Systems*, **3**, 91–105.

Mitchell, B. G., and Kahru, M., 1998, Algorithms for SeaWiFS standard products developed with the CalCOFI bio-optical data set. *California Cooperative Oceanic Fisheries Investigations Reports,* **39**, 133–147.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., Kahru, M, and McClain, C. R., 1998, Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research*, **103**, 24 937–24 953.

Venrick, E. L., and Hayward, T. L., 1984, Determining chlorophyll on the 1984 CalCOFI surveys. *California Cooperative Oceanic Fisheries Investigations Reports,* **25**, 74–79.