

## Physiochemical properties influencing biomass abundance and primary production in Lake Hoare, Antarctica

Radu Herbei<sup>b,\*</sup>, W. Berry Lyons<sup>b</sup>, Johanna Laybourn-Parry<sup>a</sup>, Christopher Gardner<sup>b</sup>, John C. Priscu<sup>c</sup>, Diane M. McKnight<sup>d</sup>

<sup>a</sup> Institute for Antarctic and Southern Ocean Studies, University of Tasmania, Hobart, Australia

<sup>b</sup> The Ohio State University, Columbus, OH 43210, United States

<sup>c</sup> Montana State University, Bozeman, MT 59717, United States

<sup>d</sup> INSTAAR, University of Colorado, Boulder, CO 80309, United States

### ARTICLE INFO

#### Article history:

Received 22 July 2009

Received in revised form

14 December 2009

Accepted 18 December 2009

Available online 26 January 2010

#### Keywords:

Lake Hoare

MCM-LTER

Bayesian methods

Correlation analysis

Gaussian Markov random fields

### ABSTRACT

The perennially ice-covered, closed basin lakes in the McMurdo Dry Valleys respond rapidly to environmental changes, especially climate. For the past 15 years, the McMurdo Dry Valleys Long-Term Ecological Research (MCM-LTER) program has monitored the physical, chemical and biological properties of the lakes in Taylor Valley. In order to better assess the physiochemical controls on the biological process within one of these lakes (Lake Hoare), we have used vertical profile data to estimate depth-dependent correlations between various lake properties. Our analyses reveal the following results. Primary production rates (PPR) are strongly correlated to light (PAR) at 12–15 m and to soluble reactive phosphorus (SRP) at 8–22 m. Chlorophyll-a (CHL) is also positively correlated to PAR at 14 m and greater depths, and SRP from 15 m and greater. This preliminary statistical analysis supports previous observations that both PAR and SRP play significant roles in driving plant growth in Lake Hoare. The lack of a strong relationship between bacterial production (BP) and dissolved organic carbon (DOC) is an intriguing result of the analysis.

Published by Elsevier B.V.

### 1. Introduction

The McMurdo Dry Valleys of Antarctica (76°30' S to 78°30' S) from the largest ice-free area on the continent. In Taylor Valley, the mean annual temperature ranges from −14.8 °C to −30 °C with 10–100 days when average temperatures exceed 0 °C, see Doran et al. (2002). The average net annual precipitation are less than 5 cm of water equivalent, see Fountain et al. (1999) and Witherow et al. (2006). Yet Taylor Valley contains perennial ice-covered lakes that are maintained by the input of glacier melt flows from 4 to 10 weeks per year (Fountain et al., 1999). Taylor Valley is a major east–west trending valley within this landscape and has been the primary site of the McMurdo Dry Valleys Long-Term Ecological Research (MCM-LTER) program since 1993. Despite extremely climatic conditions, these ice-covered lakes contain a habitat where microbial life exists throughout the year (Roberts and Laybourn-Parry, 1999; Priscu et al., 1999; Marshall and Laybourn-Parry, 2002). Because

of their permanent ice covers, physical mixing is minimized and diffusion and the seasonal input of meltwater beneath the ice covers are the primary controls of chemical variation in these lakes (Spigel and Priscu, 1998; Foreman et al., 2004). Taylor Valley lakes have been termed “end-members” among the world’s lakes, in part because of their permanent ice covers (Fritsen and Priscu, 1999).

Lake Hoare is the freshest water lake of the Taylor Valley lakes and it has not been greatly impacted by cryoconcentration as the other Taylor Valley lakes have been (Doran et al., 1994; Lyons et al., 1998). Every year since 1993 limnological data have been collected from Lake Hoare by MCM-LTER scientists. These data include biological, chemical and physical information which are all part of the long-term monitoring program used to determine ecosystem change through time. It is evident from studies on these lake systems that an integrated knowledge of the biological, chemical and physical factors is required to understand the biogeochemical dynamics of these ecosystems. Previous research has demonstrated that these lakes are very oligotrophic and have important mixotrophic components (Priscu et al., 1999); however, gaps in knowledge remain, especially in linking how these ecosystems responds to physical drivers. In order to better assess the physiochemical controls on biological processes within Lake Hoare we have developed a statistical model which allows us to study the relationship between various environmental properties.

\* Corresponding author. Tel.: +1 614 247 7487; fax: +1 614 292 2096.

E-mail addresses: herbei@stat.osu.edu (R. Herbei), lyons.142@osu.edu (W. Berry Lyons), Jo.Laybourn-Parry@utas.edu.au (J. Laybourn-Parry), gardner.177@osu.edu (C. Gardner), jpriscu@montana.edu (J.C. Priscu), diane.mcknight@colorado.edu (D.M. McKnight).

The current research focuses on quantifying the response of the studied system (Lake Hoare) to several physiochemical properties. We focus on photosynthetically active radiation and soluble reactive phosphorus as major drivers and attempt to estimate their influence on biomass and primary production. In recent years, scientists have raised several hypotheses on this research problem, see for example Spigel and Priscu (1998), Priscu et al. (1999), Hawes and Schwarz (1999) and Takacs et al. (2001). With the availability of longer time-series data thanks to the MCM-LTER program (<http://www.mcmlter.org/>), we are now in position to provide a quantitative measure for the association between ecosystem response and its drivers. We do so in a formal probabilistic framework. Our results come to confirm hypothesized relationships. In addition, we provide measures of uncertainty in our estimates. These come as a natural byproduct of the Bayesian methodology which we employ. In Section 2 we describe the data available from the MCM-LTER project. Section 3 introduces the Bayesian methodology in general terms and its application to the current work. Our results are discussed in Section 5 followed by conclusions.

## 2. Data

Lake Hoare is 4.2 km in length, 1.0 km in width at its maximum and has a surface area of 1.94 km<sup>2</sup> (Spigel and Priscu, 1998). Ice thickness varied from 3.1 to 5.5 m with the percentage of light transmission through the ice cover varying from 0.5 to 2.8% (Spigel and Priscu, 1998). Ice thicknesses vary seasonally and annually, depending primarily on temperature and inflow differences (Doran et al., 2002). The maximum total dissolved solids is approximately 0.70 g L<sup>-1</sup> and occurs at the bottom of the lake and lake water temperatures range from 0 to 1.2 °C (Spigel and Priscu, 1998). The maximum depth of Lake Hoare is 34 m but the on-going LTER sampling program collects data in a region where the lake depth is roughly 28 m.

There are three major sources of meltwater to Lake Hoare: direct input from the Canada Glacier (that serves as the lake's eastern boundary); input from Andersen Creek that issues from the north-west portion of the Canada Glacier and flows ~2 km to the lake; and from the sporadic overflow of Lake Chad to the west (Fig. 1). Since the beginning of the MCM-LTER during the 1993–1994 austral summer, annual flows in Andersen Creek have varied from  $1.48 \times 10^4$  to  $4.13 \times 10^5$  m<sup>3</sup>. During the high flow seasons of 1982–1983, it was estimated that of the total water input into Lake Hoare, approximately 67%, 22%, and 11%, came from direct melt from the Canada glacier, Andersen Creek and Lake Chad overflow, respectively (Green et al., 1986). Fortner et al. (2005) have estimated through geochemical means that during high flow years such as 2001–2002, ~80% of the water flow into Lake Hoare from the Canada Glacier comes from direct glacier input and 20% comes from Andersen Creek, while during low flow years Andersen Creek may provide as much as 40% of the total flow from the Canada Glacier.

In this analysis we are utilizing samples that were collected beneath the ice-covered Lake Hoare between 1993 and 2004 from

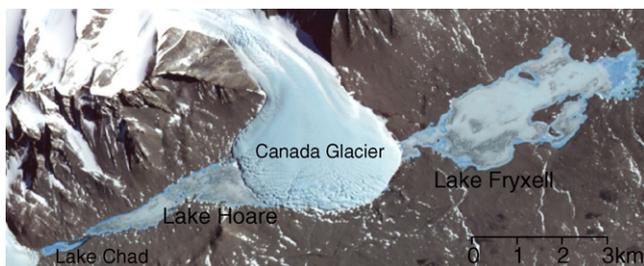


Fig. 1. Antarctica's Taylor Valley, from N Earth Observatory (2008).

depths between 4 and 24 m using standard limnological techniques (i.e. Niskin bottles). During any one austral summer 2–4 “limno runs” were conducted when vertical profiles of various constituents were either measured directly or samples were obtained for later analysis. Samples were collected and analyzed for primary production rate (PPR), bacterial production (BP), chlorophyll-a (CHL), dissolved organic carbon (DOC) and soluble reactive phosphorus (SRP). However, not all these measurements are available for every depth interval for every year. For example, small variations in lake sampling positions can lead to changes in lake depth. In addition, occasionally a sample is spoiled or lost. Photosynthetically active radiation (PAR) was measured, *in situ*, throughout the water column. All the measurement techniques as well as their analytical uncertainties are detailed in the Limnological Methods for the McMurdo Long-Term Ecological Research Program compilation found at [http://mcmlter.org/queries/lakes/lakes\\_home.jsp](http://mcmlter.org/queries/lakes/lakes_home.jsp) and will not be described here.

## 3. Methods

In Fig. 2 we display the data for PAR, PPR, CHL, BP, observed during the 1993–2004 period, where each austral summer is divided into four months (October through January). After an initial preprocessing stage, we standardize each variable and present it on a log scale, thus no units are reported. This transformation will normalize the unusually large values for most variables in the early 2000s and is required by the assumption we make later in this section, where we work with a Gaussian statistical model (see Section 3.1 for further details). Henceforth, all the statistical analyses are performed using the transformed data. The aim of the current study is to better understand the associations between these quantities. Formally, let  $X(z, t)$  and  $Y(z, t)$  denote two generic observable quantities. They depend both on depth  $z \in Z \equiv [-22 \text{ m}, -4 \text{ m}]$  and time  $t \in T \equiv [\text{Oct } 1993, \text{Jan } 2005]$ . The depth interval is discretized in steps of size 1 m while the time interval is discretized monthly, removing the austral winter months. We aim to estimate the functional correlation coefficient

$$\rho_{X,Y}(z) = \frac{\text{cov}(X(z), Y(z))}{\sqrt{\text{Var}(X(z)) \cdot \text{Var}(Y(z))}}, \quad (1)$$

which is a simple extension of the classical correlation coefficient, where  $X(z) = \{X(z, t), \text{fort } t \in T\}$  and  $Y(z) = \{Y(z, t), \text{fort } t \in T\}$ . Eq. (1) is thought of in the framework of *point-wise* linear correlation. The explanatory variable  $X$  will influence the response  $Y(z)$  only through its value  $X(z)$  at depth  $z$ . In what follows, the role of  $(X, Y)$  will be played by several pairs of the measured quantities described above. The drawback we encounter is that not all variables of interest are available simultaneously. For example in December 1997, PPR and CHL are available while BP is sparsely measured and PAR is not available at all. This implies a huge reduction in sample size when attempting to “match up” variables in order to estimate  $\rho$ . Because Antarctic lake data are difficult to obtain, we chose not to eliminate observed values for  $X$  that do not have a corresponding observation in  $Y$  or vice versa. Instead, we take a two-step approach. In a first stage, we augment the data by “filling in the blanks” in Fig. 2 using a formal statistical model. Afterwards, we perform the correlation analysis as described above using the newly obtained data.

### 3.1. Predicting the missing observations

#### 3.1.1. Background on Bayesian methods

In the first stage of our analysis we pursue a Bayesian approach. Ellison (1996) provides a brief introduction to Bayesian inference for ecological research. In recent years, stimulated by the exponen-

tial development of computing technology, Bayesian methodology has become increasingly popular in virtually any field, including ecology. For example, in Borsuk et al. (2001), the authors implement a hierarchical Bayesian model for predicting benthic oxygen demand from organic matter loading. For a formal in-depth treatment of this methodology we refer the reader to Gelman et al. (2004). In a nutshell, a Bayesian approach follows these steps. For every parameter of interest we specify a probability model based on *a priori* information. This model is subsequently changed by the available data, resulting in an updated probability model. This updated model can be used to answer any question of interest in a probabilistic framework. That is, we provide a quantitative measure for the probability of a certain event (hypothesis) to occur, given the data. This is in contrast with the more classical approach of providing the probability of observing the data, given a certain hypothesis.

Bayesian methodology is based on the simple fact from probability theory that the joint distribution of a collection of random

variables can be decomposed into a series of conditional probabilities. That is, if  $X, Y, Z$  are random variables, then we can factorize their joint distribution as  $[X, Y, Z] = [Z|X, Y][Y|X][X]$ . Note, throughout the current manuscript we will use square brackets notation for the probability distribution, so  $[X]$  denotes the distribution of the random variable  $X$  and  $[Y|X]$  denotes the conditional distribution of  $Y$  given  $X$ .

Initially we build an observational model which will lead us to a probability model for the available data conditionally on the parameters. In the second stage we account for uncertainty in the parameters, by specifying an *a priori* probability model. Using Bayes rule we can write the distribution of the parameters conditionally on the observed data as

$$[\text{parameters}|\text{data}] \propto [\text{data}|\text{parameters}] \times [\text{parameters}]. \quad (2)$$

The symbol  $\propto$  means “proportional to”, that is, the right hand side of (2) does not necessarily integrate to one with respect to the parameters (while the left hand side does). The normalizing constant

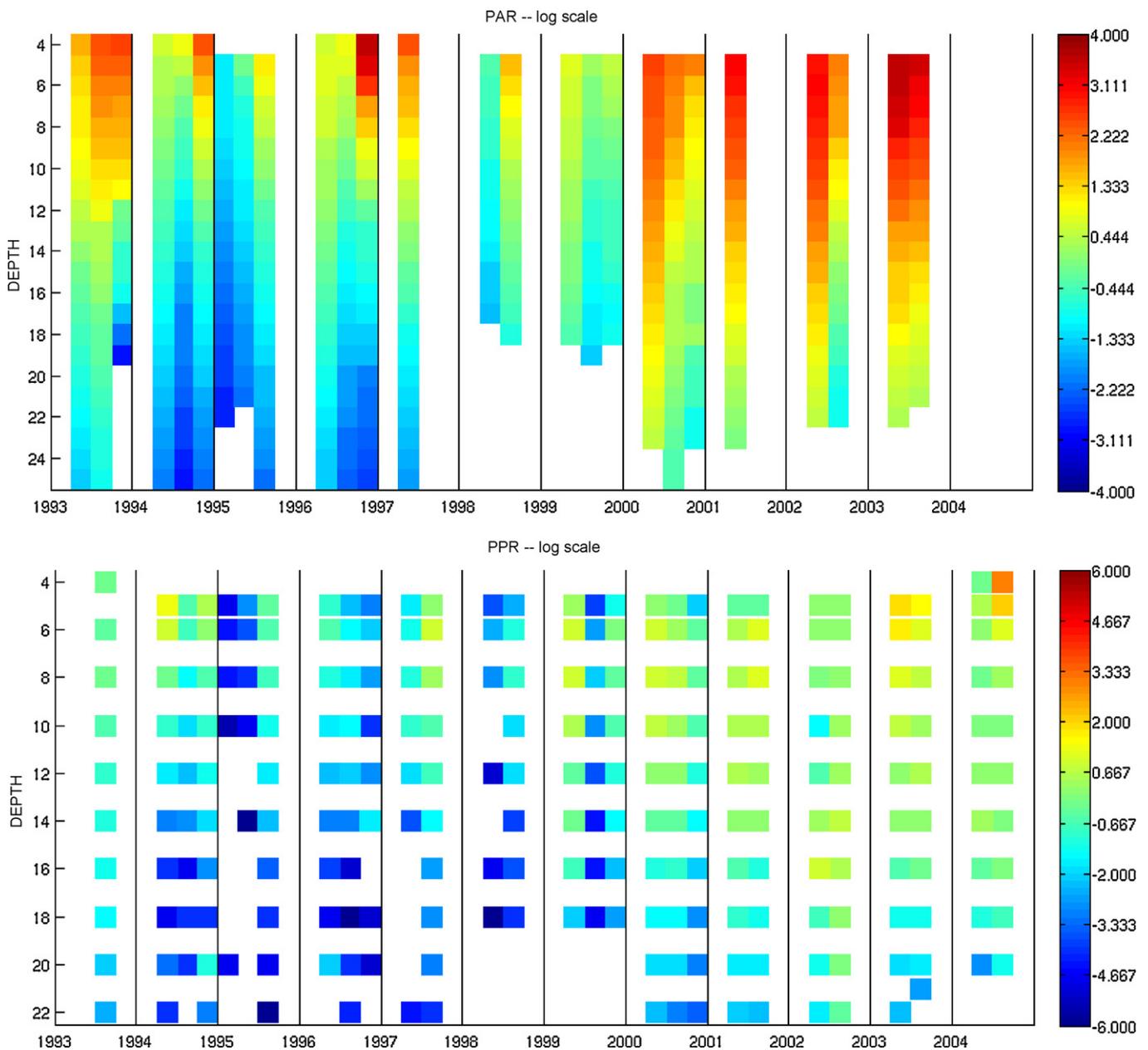


Fig. 2. Observed data: PAR (top left), PPR (top right), CHL (bottom left) and BP (bottom right), on a log scale.

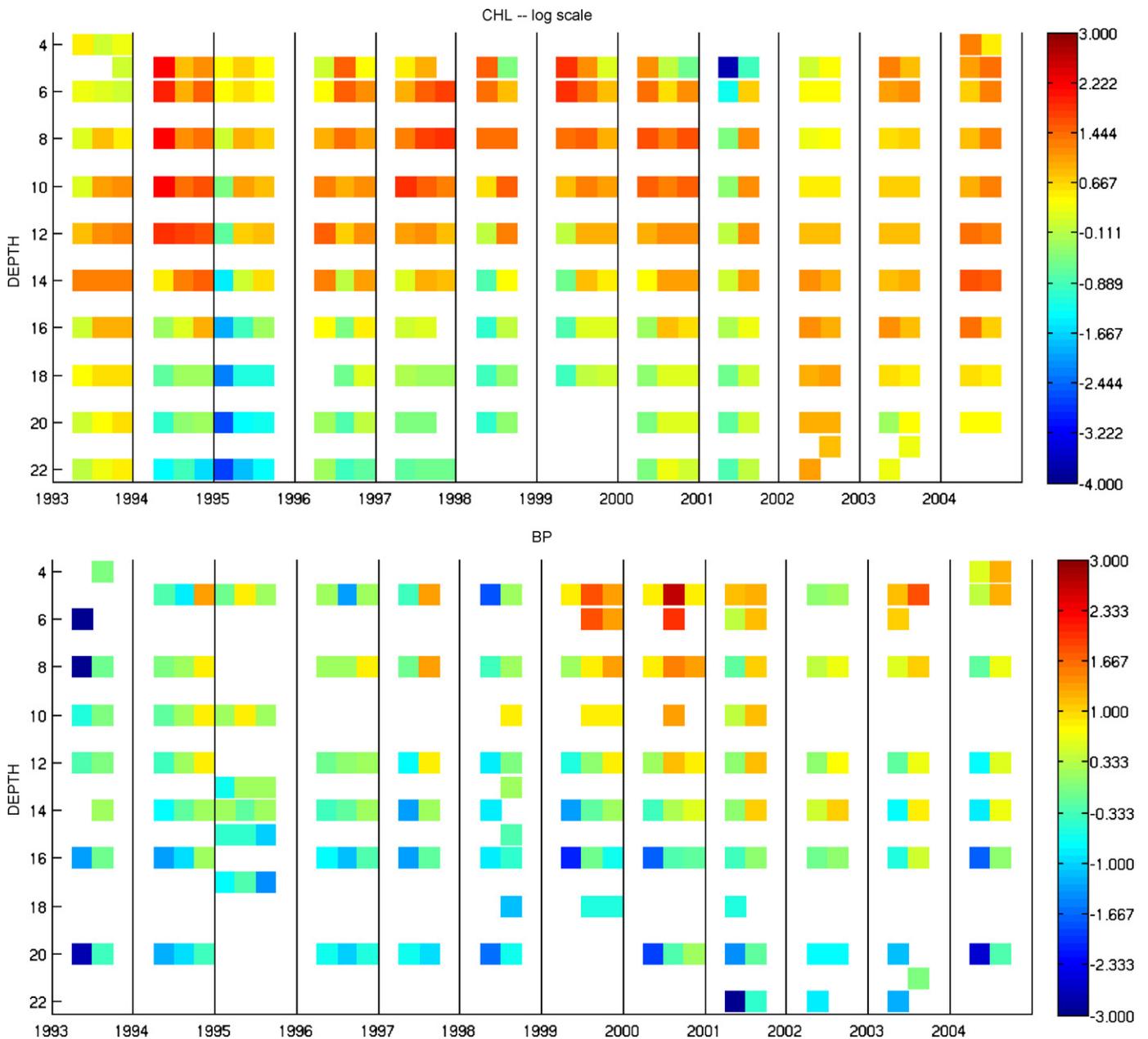


Fig. 2. (Continued).

will only depend on the observed data. This formula serves as the basis for the Bayesian hierarchical analysis. *The end-product of this approach is the left-hand side of (2)*. It represents the updated joint probability model of the parameters of interest, after observing the data. This probability model is further used to answer any question about hypotheses of interest. For more complicated processes, an intermediate step is inserted, where one would introduce a probability model for the physical process, conditioning on parameters, see Berliner (1996) for example.

One would typically select a few quantitative measures to summarize the distribution in displayed in (2), for example, representative values which are relevant to the scientist. The most common choices are the mean value and a measure of spread (standard deviation for instance), but quantiles (such as the median) or modes are also possible. Unfortunately, in most cases this updated distribution is too complicated to perform analytical integrations with, hence, one has to use Monte Carlo methods to perform such tasks. Another drawback may be that simulating independent

draws from this distribution is impossible, especially when the normalizing constant, which is ignored in the right hand side of (2), is unknown. However, in most cases Markov chain Monte Carlo (MCMC) methods can be used to simulate *dependent* draws from this updated distribution of the parameters, which can be further used to estimate features of interest. Different MCMC algorithms are useful for different situations (see Liu (2001) for an overview).

### 3.1.2. Application to Lake Hoare data

We use the PAR data to illustrate the Bayesian approach presented earlier. In what follows, by “site”  $s$  we mean a depth–time pair  $s = (z, t) \in Z \times T$ . The observational model assumes that

$$PAR_s^{obs} = PAR_s + \epsilon_i$$

where  $s = 1, \dots, n_D$  indexes a site where data are available,  $PAR_s$  is the underlying (true) photosynthetically active radiation at site  $s$  and  $\epsilon_i$  are independent Gaussian measurement errors having

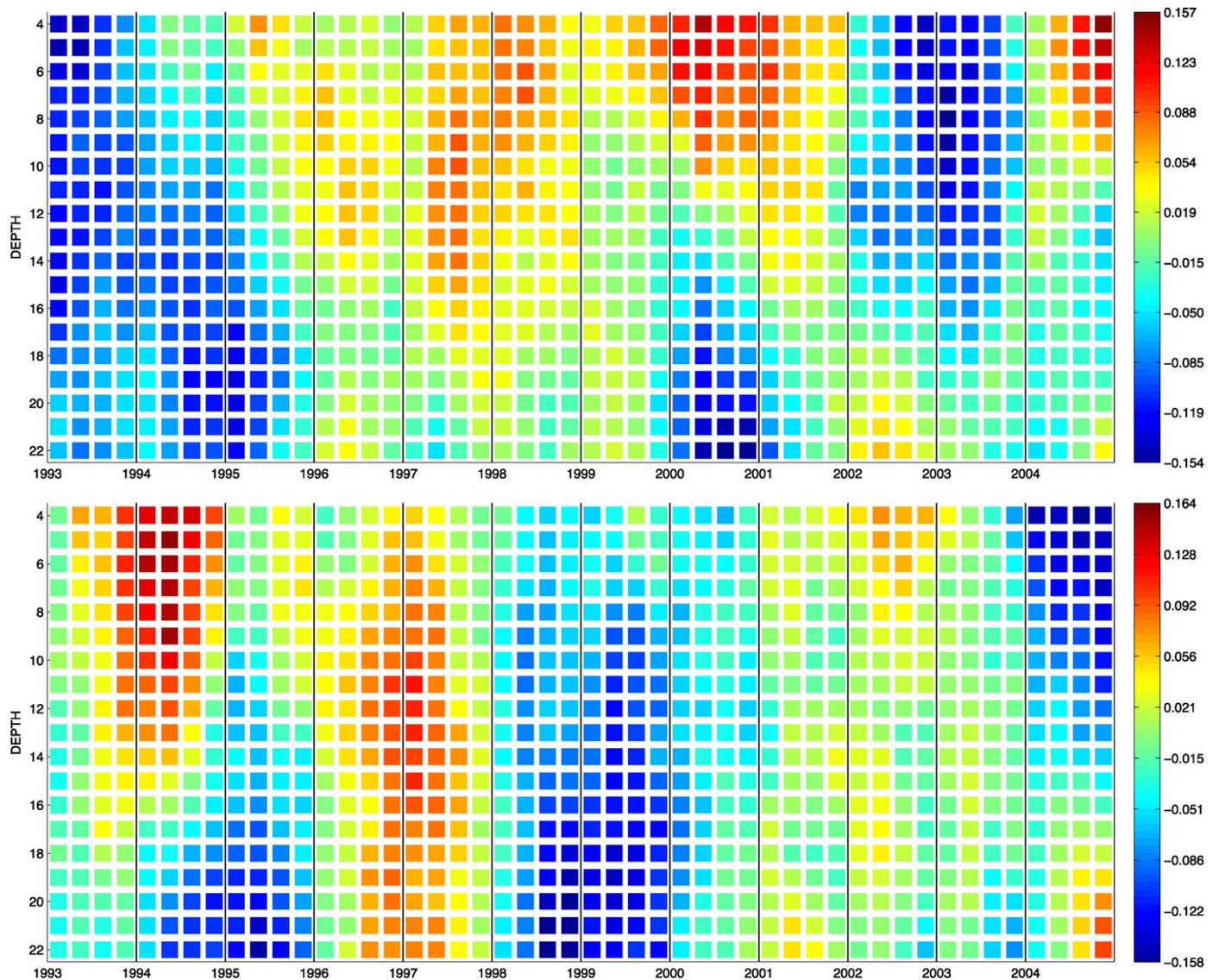


Fig. 3. Two random draws from the prior probability model (4).

zero mean and pre-specified variance  $\sigma_{PAR}^2$ . It follows that the conditional distribution of the data  $PAR^{obs} = \{PAR_s^{obs}, s = 1, \dots, n_D\}$ , given the underlying PAR field is multivariate Gaussian, having density

$$[data|parameters] = [PAR^{obs}|PAR] = \left( \frac{1}{\sigma_{PAR} \sqrt{2\pi}} \right)^{n_D} \prod_{s=1}^{n_D} \exp \left( -\frac{1}{2\sigma_{PAR}^2} (PAR_s - PAR_s^{obs})^2 \right). \quad (3)$$

To complete the Bayesian approach, we now specify a prior probability model for the underlying PAR field. We choose to model PAR using a two-dimensional Gaussian Markov random field (GMRF) having un-normalized probability density

$$[parameters] = [PAR] \propto \exp \left( -\sum_{s \sim s'} \delta_t(s, s') (PAR_s - PAR_{s'})^2 - \sum_{s \sim s'} \delta_z(s, s') (PAR_s - PAR_{s'})^2 - \sum_s \delta(s) PAR_s^2 \right). \quad (4)$$

Here the first sum runs over pairs of horizontally (time) adjacent sites, the second sum runs over pairs of vertically (depth) adjacent sites and the third sum runs over all sites. We pre-specify the positive parameters  $\delta_t(s, s')$ ,  $\delta_z(s, s')$  and  $\delta(s)$  to control the nearest neighbor interaction and the variance of PAR. For example, large

values of  $\delta_t$  will penalize discrepancies in the time-direction for two adjacent PAR values, thus imposing a time-smooth field. The values for these parameters are selected so that the model (4) will exhibit the variability and smoothness one expects to see in a typical PAR field. GMRF models are a popular choice in spatial statistics and we refer the reader to Rue and Held (2005) and Song et al. (2008), for an overview of GMRF and their statistical applications. For illustration purposes, in Fig. 3 we display two random draws from this prior probability model.

We note that display 4 provides us with a purely statistical model for PAR and no physical laws are being used nor we require PAR to satisfy any. For example, in Fig. 3 the reader will observe

that in our probability model, PAR does not necessarily decrease with depth.

Bayes formula (2) gives the updated probability model of PAR conditionally on the data  $PAR^{obs}$

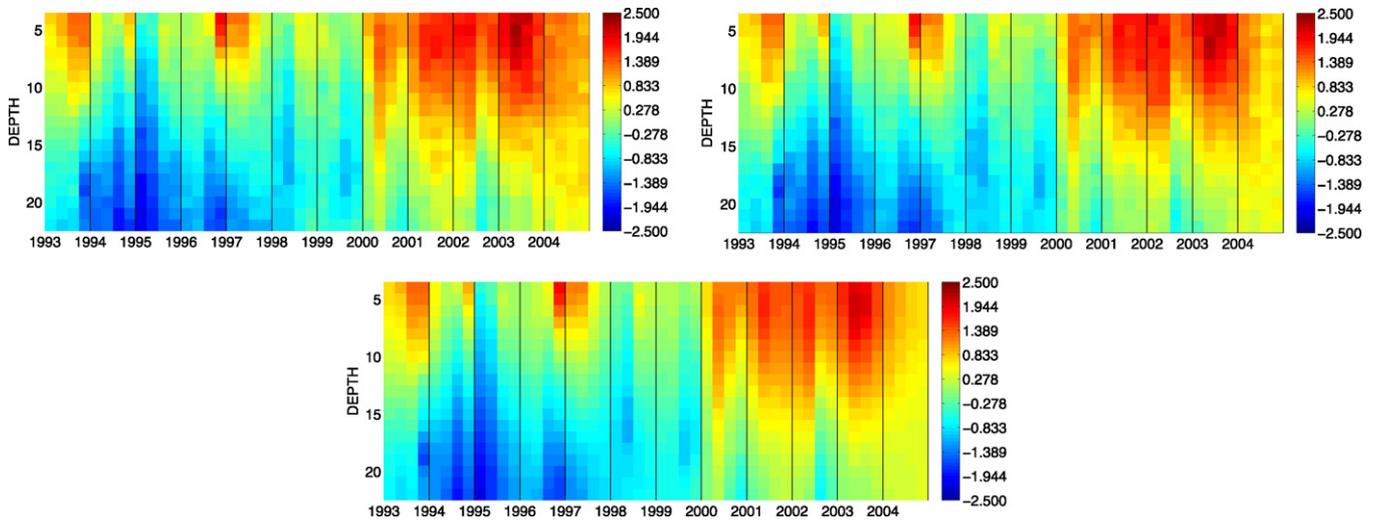


Fig. 4. Two random draws from the posterior (updated) probability model for PAR (top panels), and the mean PAR field (lower panel).

[ parameters|data]

$$= [\text{PAR}|\text{PAR}^{\text{obs}}] \propto \left\{ \prod_{s=1}^{n_D} \exp \left( -\frac{1}{2\sigma_{\text{PAR}}^2} (\text{PAR}_s - \text{PAR}_s^{\text{obs}})^2 \right) \right\} [\text{PAR}], \quad (5)$$

where we dropped the normalizing constant from (3). This is the so called *posterior* probability model for PAR, after observing  $\text{PAR}^{\text{obs}}$ . It allows us to quantify whether one particular PAR field is more probable than another, conditionally on the observations. We use the random-walk Metropolis algorithm as our MCMC technique of choice to draw dependent samples from the probability model (5). A detailed description of this popular algorithm can be found in Chapter 5 of Liu (2001). In Fig. 4 we display two such random draws from the probability model (5) and the average PAR field, based on 1000 samples.

We analyze PPR, CHL, BP, DOC and SRP in a similar manner. In Fig. 5 we display the mean values for their (corresponding) updated distributions, on the log scale. Henceforth we treat the augmented data displayed in Fig. 5 as “the data” and use them to estimate the correlation coefficient (1). There is no available DOC data during the 1993–1995 period, while SRP measurements are missing in the 1993 and 2004 austral summers. Consequently, for these two variables, the corresponding years are not estimated in Fig. 5.

### 3.2. Correlation analysis

At a first glance, in Fig. 5 we observe similar “patterns” in the PAR, PPR, BP maps as well as in the DOC and SRP maps. This suggests a significant relationship between these variables. Our analysis begins with the PAR–PPR relationship. For each depth value, we use the data displayed in Fig. 5 and compute the esti-

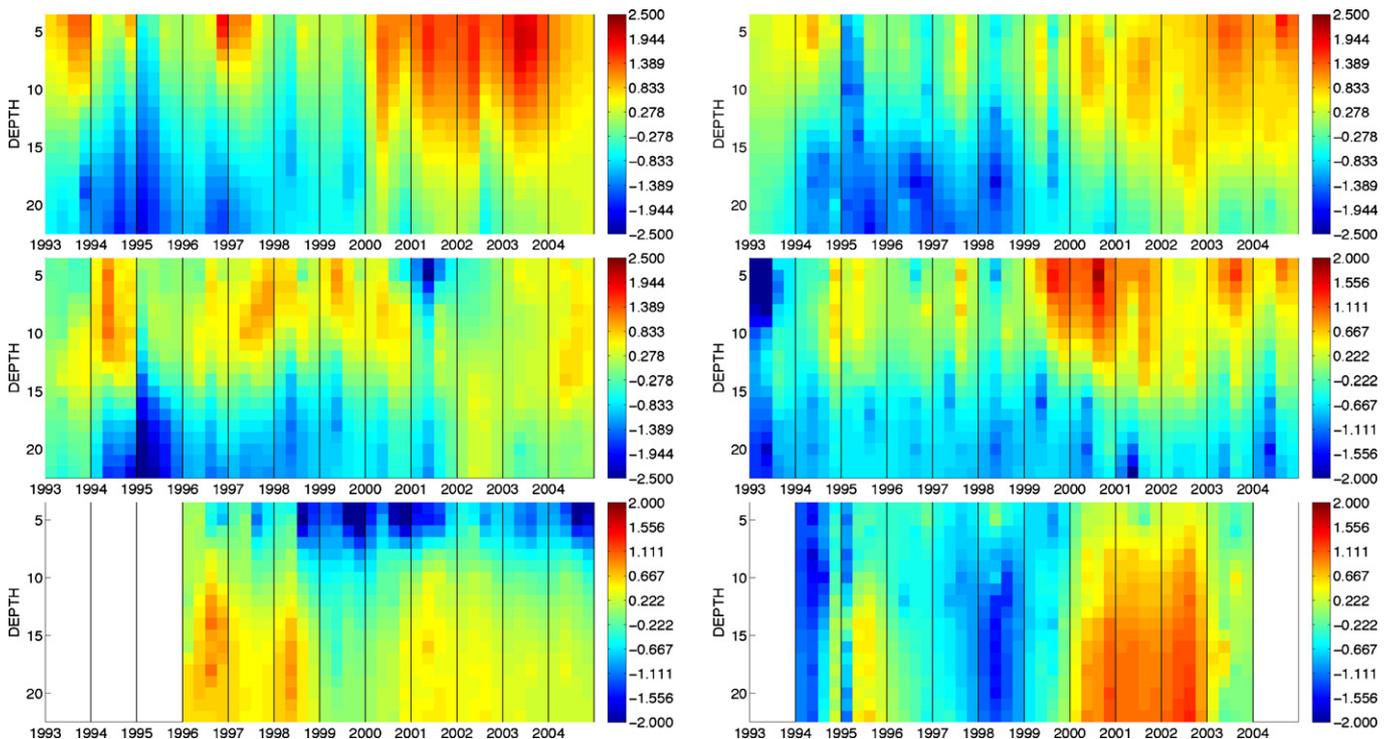
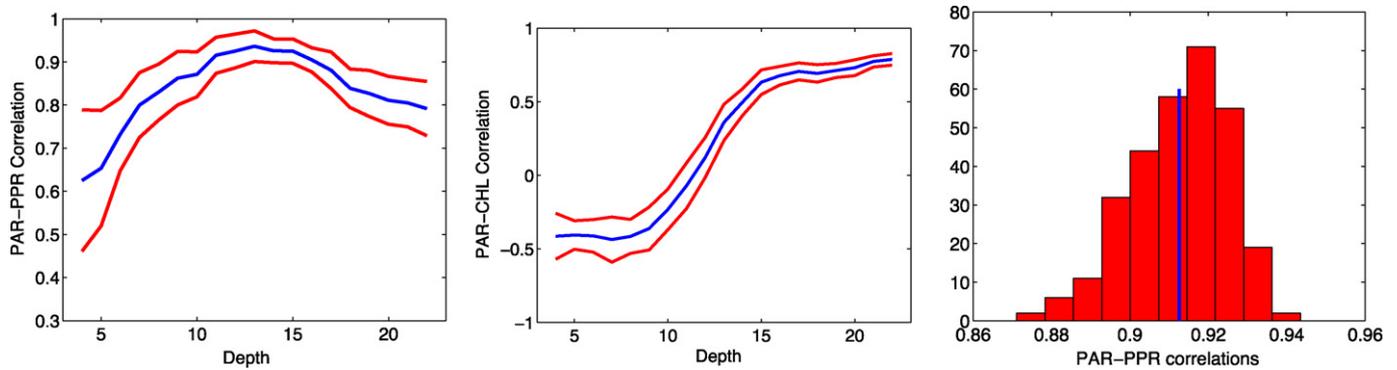


Fig. 5. The posterior (updated) mean fields for PAR, PPR, CHL, BP, DOC and SRP (top to bottom, row-wise).



**Fig. 6.** Depth dependent correlation (blue curve)  $\pm$  two standard deviations (red curves) for the PAR–PPR pair (left panel) and the PAR–CHL pair (middle panel). Histogram of PAR–PPR correlation values at 15 m depth (right panel). The blue line marks the mean value of 0.913. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

mate (1). The left panel of Fig. 6 displays the Pearson's correlation coefficient between these two variables as a function of depth (the blue curve). At 15 m depth, the correlation is as high as 0.913. Our Bayesian hierarchical approach allows us to quantify the uncertainty in this estimate as well, by considering several *draws* (not just the mean value) from the updated probability model for each variable. These draws are not independent; however, we can reduce the amount of dependence by thinning the output from the MCMC sampler (Liu, 2001). In the right panel of Fig. 6 we display a histogram of 300 such correlations, computed for the PAR–PPR pair at 15 m depth. Such histograms are obtained for each depth and thus we are able to estimate the standard deviations for all these probability distributions. The red curves in Fig. 6 are obtained by adding and subtracting two times the standard deviation to the mean correlation value represented by the blue curve. When the distribution of the correlation coefficient is approximately Gaussian, this corresponds to finding an interval which will roughly contain 95% of all correlation values. This is the analogue of the error bar (or confidence interval) in the Bayesian framework, and it is called a *credible interval*. It should be noted, however, that caution is needed in interpreting such intervals as the coverage does not hold simultaneously (at all depths). For example, if simultaneous 95%-coverage credible intervals are constructed for say 10 parameters, using standard probability rules, one can compute the probability that at least one interval does not contain its target parameter as  $1 - 0.95^{10} = 40.13\%$ .

An interesting finding of our analysis is that CHL is positively correlated with PAR only at depths of 10 m and higher, while close to the ice cover, the correlation is actually negative (see the middle panel of Fig. 6). This regime identifies two zones: upper under-ice and lower under-ice. Hawes and Schwarz (1999) interpret the two

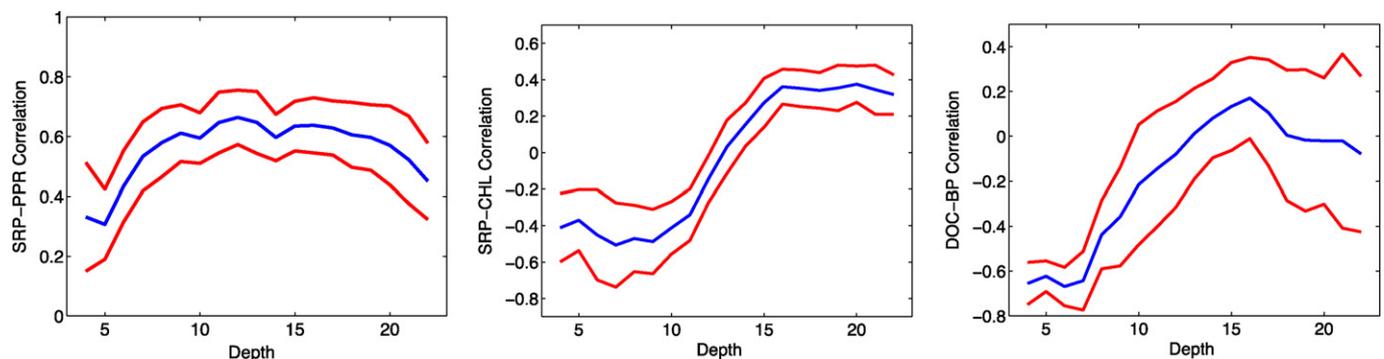
regimes as differences in biomass rather than acclimation to the light regime.

In the left panel of Fig. 7 we display the depth dependent correlation between PPR and SRP. This correlation appears to be stronger at mid-depths. Similar to the CHL–PAR association, the correlation between CHL and SRP, which we display in the middle panel of Fig. 7, takes positive values at depths of 15 m and higher, while close to surface, the correlations appears to be negative, confirming the hypothesis of shade-adaptivity of these organisms at shallower depths.

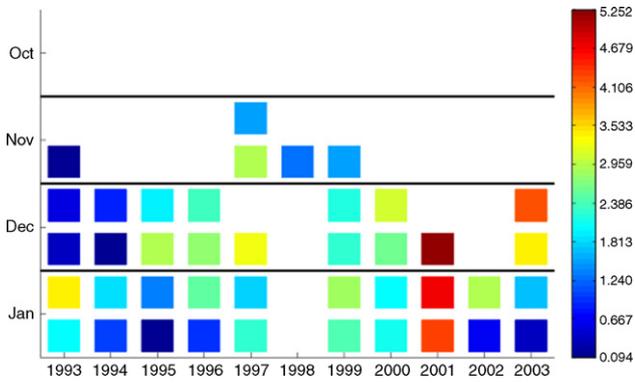
An intriguing result of our analysis is the relationship between DOC and BP (see the right panel of Fig. 7). Takacs et al. (2001) point out that in Lake Hoare, a major source of DOC required for BP is unaccounted for. Our analysis confirms their finding by estimating a negative correlation between DOC and BP close to surface, while at higher depths the correlation does not appear to be significant.

#### 4. Andersen Creek discharge

The primary source of liquid water to the MCM lakes is glacier melt (McKnight et al., 1999). The amount of this melt varies from austral summer to austral summer depending on the temperature. The generation of melt is particularly sensitive to warmer temperatures at higher elevations on the glacier surfaces brought about by increased down valley winds (Doran et al., 2008). In Fig. 8 we display bi-weekly averages of the discharge rate (on a log scale) for each austral summer from 1993 to 2003. Warmer summers can produce very high flow years, termed “flood years” (Foreman et al., 2004), and these flood years appear to be quasi-decadal events as demonstrated by Ebnet et al. (2005) and Doran et al. (2008). The most recent flood year was the 2001–2002 austral summer (see Fig. 8). During and after this event, the ecological conditions of the



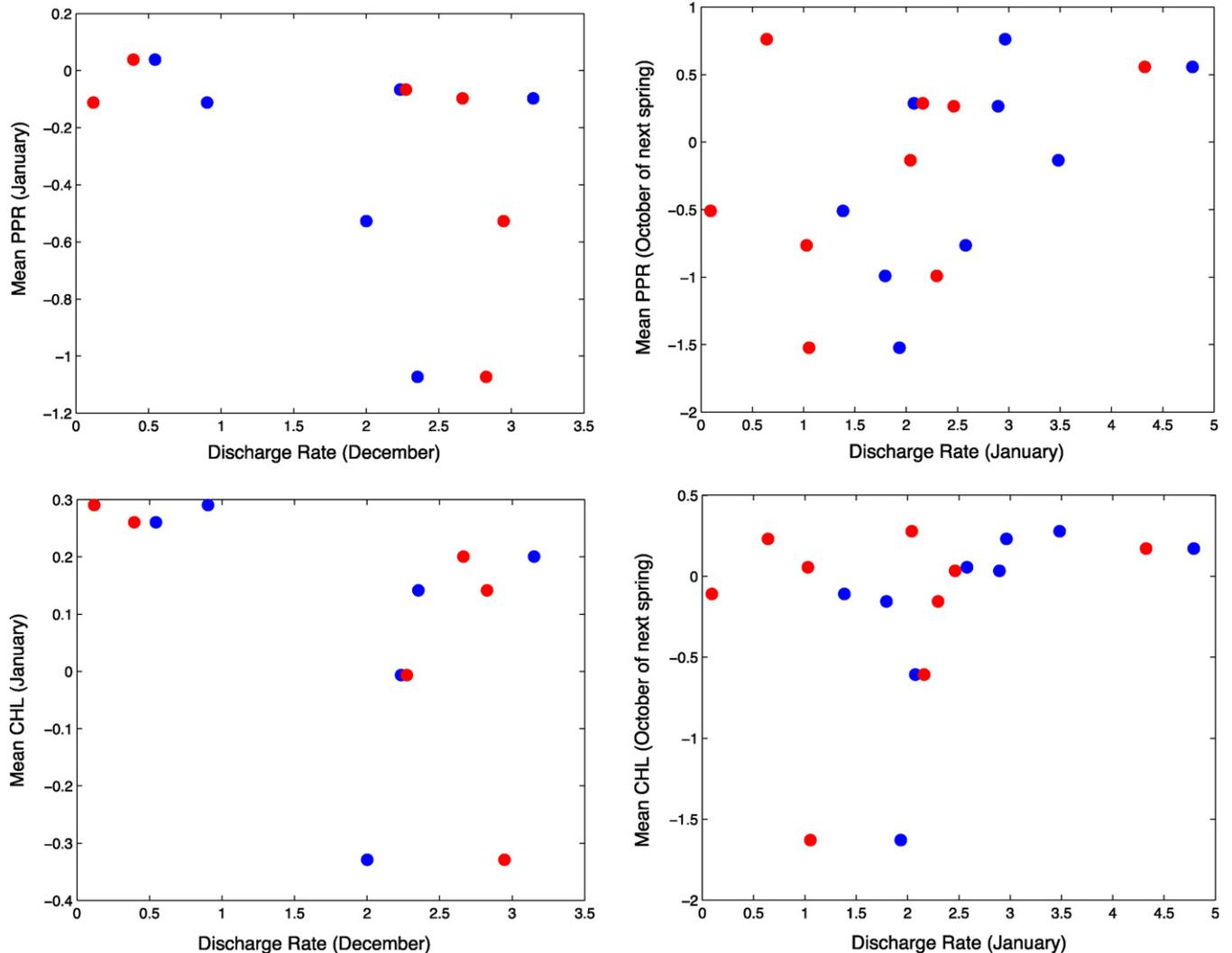
**Fig. 7.** Depth dependent correlation (blue curve)  $\pm$  two standard deviations (red curves) for the SRP–PPR pair (left panel), the SRP–CHL pair (middle panel) and the DOC–BP pair (right panel). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)



**Fig. 8.** Observed bi-weekly Andersen Creek discharge rates during 1993–2003 on a log scale.

various MCM lakes changed in response to the increased input of water, suspended sediments, and nutrients, but the lakes in Taylor Valley did not all respond in the same manner (Foreman et al., 2004). These large inflow events are thought to play a very important role in the longer term maintenance of water and nutrients in these closed-basin systems (Foreman et al., 2004).

In order to assess the longer term influences of meltwater inflow into Lake Hoare on the productivity and phytoplankton biomass of the lake, correlations were computed between the monthly discharge from Andersen Creek (the major source of stream input into the lake) and PPR and CHL. This was done both with a time lag of one month, i.e. using December’s discharge and January’s biological data, and what we have termed a seasonal lag, i.e. the discharge from January and the biological data from the next austral Spring. These relationships are shown in Fig. 9. In the monthly lagged data (left panels of Fig. 9), the PPR vs. discharge and the CHL vs. discharge scatter plots show little relationship for the entire data set. Similarly the CHL–discharge plot for the seasonally (October response vs. January discharge, right panels of Fig. 8) lagged data shows little association. However there is a better relationship between the PPR data from Lake Hoare and Andersen Creek discharge using the seasonal lag (upper right panel of Fig. 9). This relationship might suggest that increased water input from the previous austral summer melt season helps to drive primary production during the next year. The analysis of the 2001–2002 flood event by Foreman et al. (2004) demonstrated that increased PPR and CHL occurred in some of the lakes during the next Spring. Our analysis supports this idea that increased flow, bringing higher nutrient loadings, leads to enhanced biological activity in the photic zones of these lakes over time. The lack of an initial biological response may be due in



**Fig. 9.** Scatter plots of the discharge rate (L/s) vs. PPR (top panels) and CHL (bottom panels) at a one month lag (left panels) and one year lag (right panels). Blue dots correspond to the average discharge rate values for the first two weeks of the month and red dots correspond to average discharge rate values for the last two weeks of each month.

part to the increased suspended load input and, in some cases like Lake Hoare, increased aeolian deposition of soil/dust on the lake ice surface thereby decreasing PAR in the surface waters (Foreman et al., 2004). As noted above, PAR is a significant control on phytoplankton production in these lakes. However, these lakes have been demonstrated to have nutrient deficiencies as well (Priscu, 1995). Because phosphorus is a limiting nutrient in Lake Hoare (Priscu, 1995; Barrett et al., 2007), and the primary source of phosphorus is stream water input, increased stream flow should enhance production in the lakes over time. Priscu (1995) also suggests the nitrogen limitation on phytoplankton growth in Lake Hoare; however, this relationship was not investigated in the current study.

## 5. Discussion

These statistical analyses strongly support previously published work done at Lake Hoare and other Taylor Valley lakes. For example, experimental work and field studies have demonstrated that both light and SRP limit phytoplankton growth (Priscu, 1995; Fritsen and Priscu, 1999). Ice cover transparency which affects the penetration of light and hence under ice PAR strongly influences primary production in Lake Bonney (Fritsen and Priscu, 1999). These authors found a strong correlation between PPR and daily irradiance at 10 m in Lake Bonney. Fritsen and Priscu (1999) observed that beneath ice radiance had a seasonal pattern that was quite different than the incident irradiance, with a maximum in late November and another lesser maximum in late January. These differences are caused by changes in temperature and albedo within the ice covers themselves, which has been referred to as “ice-whitening”. Variations in both snow and sediment cover on the lake ice surface can also be important but it is very clear that the major variations in PAR relate to changes in the ice properties (Fritsen and Priscu, 1999). The maximum values of irradiance in late November were observed to be extremely low  $\sim 6 \mu \text{ mol photons m}^{-2} \text{ s}^{-1}$  so these autotrophs within the lakes are highly shade adaptive and because these organisms are below their saturation values of  $\sim 10$  to  $30 \mu \text{ mol photons m}^{-2} \text{ s}^{-1}$ , even small changes in irradiance can have significant impacts on PPR (Lizotte and Priscu, 1992). The feedback between PPR, ice conditions and nutrient limitation is equally significant as Fritsen and Priscu (1999) suggest. As ice whitening occurs, PPR decreases and hence nutrients such as SRP increase in the water column thereby extending the time period in the austral summer that SRP is not limiting for phytoplankton growth. Taken together, these results may imply that PAR is more limiting to growth at higher depths but SRP is limiting in the upper portions of the water column where light is more abundant.

The relationship between PAR and plant growth has been well established with field measurements and modeling. For example, Priscu et al. (1999) have used daily PAR records in Lake Bonney at 10 m and a hyperbolic tangent model to predict depth-integrated PPR. More recently Moorhead et al. (2005) have used a similar approach using the PAR observations above Lake Hoare to simulate net organic carbon production in the benthic algal mats. They calculate transmitted PAR as a constant fraction of the ambient radiation and found the highest annual production rates occurred during the times that PAR was the highest (December and January). They found little increase in carbon accumulation at transmittances greater than 5% ambient PAR except at the deepest depths.

As noted above during one sampling season Takacs et al. (2001) pointed that the observed BP in all three major Taylor Valley lakes, including Lake Hoare, could not be explained by the major inputs of DOC into the lakes. DOC from phytoplankton extracellular release, stream input and, upward diffusion into the trophogenic zone across the chemoclines of the lakes or from the pore waters can account for only about 10% of the BP in Lake Hoare. Additional sources to drive bacterioplankton production, not just in the austral

summer, but also on an annual basis must be provided by another, unknown source of DOC in Lake Hoare. It is speculated that this unknown source may come from the solubilization of particulate material and/or the slow consumption of bulk DOC within the lake. Our analysis indicates no relationship between BP and DOC also implying another source besides the bulk DOC driving BP in these systems. Part of this unaccounted for source of DOC may be due to viral activity. Viruses have a significant impact on the cycling of nutrients and carbon in aquatic environments. While they have not been widely studied in Antarctic lakes there is clear evidence that they can have a major impact at certain times. For example in a large freshwater lake in the Vestfold Hills (Crooked Lake) viral lysis of bacteria was estimated to contribute up to 69% of the DOC pool, and in Lake Bonney (Dry Valleys) it was estimated that around 23% of bacterial carbon demand could be met by viral lysis (Sawström et al., 2008).

## 6. Conclusion

The perennially ice-covered, closed basin lakes of the McMurdo Dry Valleys are unusual aquatic ecosystems where liquid water exists in an environment that is an extreme polar desert. The connection between geophysical parameters and the biogeochemical ones are closely coupled. The correlation analysis approach presented here provides another method to relate long-term geophysical parameters to biogeochemical ones in these environments. The results and their ecological interpretations are consistent with all the previous work, especially that of Priscu et al., regarding the physiochemical drivers of biological processes in the Taylor Valley lakes. Currently, we are extending our analysis to other Antarctic lakes that have different geophysical settings. In the future we aim to perform a meta-analysis and combine the available information from different Antarctic lakes to better understand these unique ecosystems. The results of the present work show a complex inter-dependent structure between the studied variables, suggesting that a multivariate approach could shed more light on the system.

## Acknowledgements

The field work portion of this work was supported by National Science Foundation grants OPP9211773, OPP9810219 and OPP0096250. We are very grateful to MCM-LTER personnel who collected and analyzed the samples. The initial data complication was supported through a Royal Society Traveling Fellowship grant to Johanna Laybourn-Parry and W. Berry Lyons.

## References

- Barrett, J.E., Virginia, R.A., Lyons, W.B., McKnight, D.M., Priscu, J.C., Doran, P.T., Fountain, A.G., Wall, D.H., Moorhead, D.L., 2007. Biogeochemical stoichiometry of Antarctic Dry Valley ecosystems. *Journal of Geophysical Research*, 112.
- Berliner, M., 1996. Hierarchical Bayesian time series models. In: Hanson, K.M., Silver, R.N. (Eds.). *Maximum Entropy and Bayesian Methods*, Kluwer Academic Publishers (Springer), New York, pp. 15–22.
- Borsuk, M., Higdon, D., Stow, C., Reckhow, K., 2001. A Bayesian hierarchical model to predict oxygen demand from organic matter loading in estuaries and coastal zones. *Ecological Modelling* 143, 165–181.
- Doran, P.T., Wharton, R.A., Lyons Jr., W.B., 1994. Paleolimnology of the McMurdo Dry Valleys, Antarctica. *Journal of Paleolimnology* 10, 85–114.
- Doran, P.T., McKay, C.P., Clow, G.D., Dana, G.L., Fountain, A.G., Nysten, T., Lyons, W.B., 2002. Valley floor climate observations from the McMurdo Dry Valleys, Antarctica, 1986–2000. *Journal of Geophysical Research (Atmospheres)* 107(D24), 4772.
- Doran, P.T., McKnight, D.M., Jaros, C., Fountain, A.G., Nysten, T.H., McKay, C.P., Moorhead, D.L., 2008. Hydrologic response to extreme warm and cold summers in the McMurdo Dry Valleys, East Antarctica. *Antarctic Science* 20, 5.
- Ebnet, A.F., Fountain, A., Nysten, T., McKnight, D., Jaros, C., 2005. An temperature-index model of stream flow at below freezing temperatures in Taylor Valley, Antarctica. *Annals of Glaciology* 40, 76–82.

- Ellison, A.M., 1996. An introduction to Bayesian inference for ecological research and environmental decision making. *Ecological Applications* 6, 1036–1046.
- Foreman, C.M., Wolf, C.F., Priscu, J.C., 2004. Impact of episodic warming events on the physical, chemical and biological relationships of lakes in the McMurdo Dry Valleys, Antarctica. *Aquatic Geochemistry* 10, 239–268.
- Fortner, S.K., Tranter, M., Fountain, A., Lyons, W.B., Welch, K.A., 2005. The geochemistry of supraglacial streams of Canada Glacier, Taylor Valley (Antarctica), and their evolution into proglacial waters. *Aquatic Geochemistry* 11, 391–412.
- Fountain, A., Lewis, K.J., Doran, P.T., 1999. Spatial climatic variation and its control on glacier equilibrium line altitude in Taylor Valley, Antarctica. *Global and Planetary Change* 22, 1–4.
- Fritsen, C.H., Priscu, J.C., 1999. Seasonal change in the optical properties of the permanent ice cover on Lake Bonney, Antarctica: consequences for lake productivity and phytoplankton dynamics. *Limnology and Oceanography* 44, 447–454.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. *Bayesian Data Analysis*. Chapman & Hall/CRC.
- Green, W.J., Ferdelman, T.G., Gardner, T.J., Varner, L.C., Angle, M.P., 1986. The residence times of eight trace metals in a closed-basin Antarctic lake: Lake Hoare. *Hydrobiologia* 134, 249–255.
- Hawes, I., Schwarz, A., 1999. Photosynthesis in an extreme shade environment: benthic microbial mats from Lake Hoare, a permanently ice-covered Antarctic lake. *J. Phycol.* 35, 448–459.
- Liu, J.S., 2001. *Monte Carlo Strategies in Scientific Computing*. Springer, New York.
- Lizotte, M., Priscu, J.C., 1992. Photosynthesis–irradiance relationships in phytoplankton from the physically stable water column of a perennially ice-covered lake (Lake Bonney, Antarctica). *Journal of Phycology* 28, 179–185.
- Lyons, W.B., Tyler, S.W., Wharton, R.A., McKnight, D.M., Vaughn Jr., B.H., 1998. A late holocene dessication of Lake Hoare and Lake Fryxell, McMurdo Dry Valleys, Antarctica. *Antarctic Science* 10, 247–256.
- Marshall, W., Laybourn-Parry, J., 2002. The balance between photosynthesis and grazing in Antarctic mixotrophic cryptophytes. *Freshwater Biology* 47, 2060–2070.
- McKnight, D.M., Niyogi, D.K., Alger, A.S., Bomblies, A., Conovitz, P.A., Tate, C.M., 1999. Dry Valley Streams in Antarctica: ecosystems waiting for water. *Bioscience* 49, 985–995.
- Moorhead, D., Schmeling, J., Hawes, I., 2005. Modelling the contribution of benthic microbial mats to net primary production in Lake Hoare, McMurdo Dry Valleys. *Antarctic Science* 17, 33–45.
- NASA Earth Observatory, 2008. url= <http://earthobservatory.nasa.gov/IOTD/view.php?id=35535>.
- Priscu, J.C., 1995. Phytoplankton nutrient deficiency in lakes of the McMurdo Dry Valleys, Antarctica. *Freshwater Biology* 34, 215–227.
- Priscu, J.C., Wolf, C.F., Takacs, C.D., Fritsen, C.H., Laybourn-Parry, J., Roberts, E.C., Sattler, B., Lyons, W.B., 1999. Carbon transformations in a perennially ice-covered Antarctic Lake. *BioScience* 49, 997–1008.
- Roberts, E.C., Laybourn-Parry, J., 1999. Mixotrophic cryptophytes and their predators in the Dry Valley lakes of Antarctica. *Freshwater Biology* 41, 737–746.
- Rue, H., Held, L., 2005. *Gaussian Markov Random Fields*. Chapman & Hall, London.
- Sawström, C., Lisle, J., Anesio, A.M., Priscu, J.C., Laybourn-Parry, J., 2008. Bacteriophage in polar inland waters. *Extremophiles* 12, 167–175.
- Song, H., Fuentes, M., Ghosh, S., 2008. A comparative study of Gaussian geostatistical models and Gaussian Markov random field models. *Journal of Multivariate Analysis* 99, 1681–1697.
- Spigel, R.H., Priscu, J.C., 1998. Physical limnology of the McMurdo Dry Valleys lakes. *Antarctic Research Series* 72, 153–187.
- Takacs, C.D., Priscu, J.C., McKnight, D.M., 2001. Bacterial dissolved organic carbon demand in McMurdo Dry Valley Lakes, Antarctica. *Limnology and Oceanography* 46, 1189–1194.
- Wetherow, R.A., Lyons, W.B., Bertler, N., Welch, K.A., Mayewski, P.A., Sneed, S.B., Nylén, T., Handley, M.J., Fountain, A., 2006. The aeolian flux of calcium, chloride and nitrate to the McMurdo Dry Valleys landscape: evidence from snow pit analysis. *Antarctic Science* 18, 497–505.