**Online Supplementary Material**

**Which environmental factors control phytoplankton populations? A Bayesian variable selection approach**

Crispin M. Mutshindaa,\*, Zoe V. Finkelb, Andrew J. Irwina

*aMathematics & Computer Science, Mount Allison University, 67 York Street, Sackville, NB E4L 1E6,* *Canada*

*bEnvironmental Science Program, Mount Allison University, Sackville, NB Canada*

*1. Description of the Data structure*

This file includes the BUGS code for fitting the model described in the MS “Bayesian variable selection on physical predictors of species abundances in a diverse phytoplankton community” to the data, along with details on the data structure.

The abundance data are in the 469x167 matrix *y*, where 469 is the number of observations per cruise (67 species x 7 observations (7depths)/species), and 167 is the number of cruises. The number of species abundance observations per cruise (469) and the number of cruises (167) are denoted by Nobs and NCruise. These values are specified in the data. The data on each environmental variable are presented in a 167 x7 matrix where the  entry is the observed value of the environmental variable at depth *k* during cruise *j*. The variable Species and Depth are 469x1 vectors of species identifying number (1 to 67) and depths (1 to 7) used to match species abundances with the environmental data.

*2. BUGS Code*

model{

########## Likelihood ##################################

for(i in 1: Nobs){

for (j in 1:NCruise){

y[i,j]~dnorm(mu[i,j], tau\_r[Species[i]])

mu[i,j]<-alpha[Species[i]]+ theta[Species[i],1]\*irrad[j,Depth[i]]+theta[Species[i],2]\*Temp[j,Depth[i]]+theta[Species[i],3]\*SiOH4[j,Depth[i]]+

theta[Species[i],4]\*PO4[j,Depth[i]]+theta[Species[i],5]\*NO3[j,Depth[i]]+theta[Species[i],6]\*Salinity[j,Depth[i]]+theta[Species[i],7]\*pH[j,Depth[i]]

}

}

########### Prior specification ###########################

# Joint prior for the environmental effects

for (s in 1:Nspecies){

beta[s, 1:Nvar]~dmnorm(m[Species[s],1:Nvar], taub[,])

}

# Precision matrix (inverse Covariance matrix) of environmental effects

taub[1 : Nvar,1 : Nvar]~dwish(R[ , ], Nvar)

# Covariance matrix of environmental effects

Omega[1:Nvar,1:Nvar]<-inverse(taub[,])

# Defining the regression coeffeicients and priors on the indicators

for(i in 1: Nspecies){

for(j in 1:Nvar){

theta[i,j]<-ind[i,j]\*beta[i,j]

ind[i,j]~dbern(0.5)

}

}

for(s in 1: Nspecies){

alpha[s]~dnorm(0, tau.a) # Intercept for species s

tau\_r[s]<-1/sigma2\_r[s]

sigma2\_r[s]~dgamma(a,b)

for (j in 1:Nvar){

m[s,j]<-0

}

}

tau.a~dgamma(0.1, 0.1)

a~dgamma(1,1)

b~dgamma(1,1)

# Correlation matrix of the environmental effects

for(k in 1:Nvar){

for(j in 1:Nvar){

rho[k,j]<-Omega[k,j]/sqrt(Omega[k,k]\*Omega[j,j])

}

}

}# END MODEL

##############################################################