**1. Introduction**: Gulf Stream and NAO Components – Azores High and Icelandic Low

The response of the Gulf Stream (GS) system to atmospheric forcing is generally linked to a combined effect of the basin-scale winds on the subtropical gyre and the buoyancy forcing from the Labrador Sea.

A recent overview of the response mechanism is provided by Gangopadhyay et al. (2016) and explained in their Figure 6 reproduced below (Fig. 1). The Figure caption highlights the overall concept of the force-response system.



Figure 1. A synergistic perspective on the coupling between the Gulf Stream and

the NAO. The GS (dark red line) is situated at the intergyre boundary between the subtropical and subpolar gyres. Its response is thus partly basin-scale wind-driven through long baroclinic Rossby waves (BRW) and partly influenced by buoyancy advection of LC and LSW from the Labrador Sea region. While the Azores high directly sets up the wind-driven gyre, the buoyancy impact from the subpolar gyre is centered on the Icelandic low, the two of which are components of the NAO. Note the spread of NESC (blue wiggly line) across the GS in the transition zone (65°–60W), probably impacting/contributing to the changeover of temporal variability of the GS from the west to the east. (Reproduced from Gangopadhyay et al. 2016; their Figure 6).

A number of earlier studies (Taylor and Stephens, 1996; Taylor et al., 1998; Taylor and Gangopadhyay, 2001; Hameed and Pointovosky (2004), Benway and Rossby (2004), Sanchez-Franks et al. (2016) and Bisagni et al. (2016)) have indicated that the inter-annual variability of the latitudinal excursion of the path of the Gulf Stream can be explained partly by the variations of the components of the North Atlantic Oscillation, i.e., the sea level pressures of the Azores High (AH) and the Icelandic Low (IL), and such relationship might have a significant lag-correlation with time-scales of up to three years.

With that in mind we set out to explore the Bayesian Modeling approach to develop a posterior distribution of the Gulf Stream path (defined by a Gulf Stream Index – GSI), given the pressure distribution of the AH and IL for the current year, and two years before present. This first-order model is presented in Section 2 with its variables and parameters. The Bayesian model is then developed with a Graph and its prior distributions. For the purpose of testing, we then present a simpler model with minimal parameter set for better understanding of the details of the numerical complexity of the JAGS model. The JAGS codes are presented in Section 3 with results in Section 4. Section 5 outlines the future steps.

Before we delve into the Bayesian Model, it is worth looking at the distribution of the different series of NAO (AH-IL), AH and IL along with the distribution of the GSI (from Taylor, PML, UK Website). Figure 2 shows that each of these four variables reasonably follows the normal distribution.

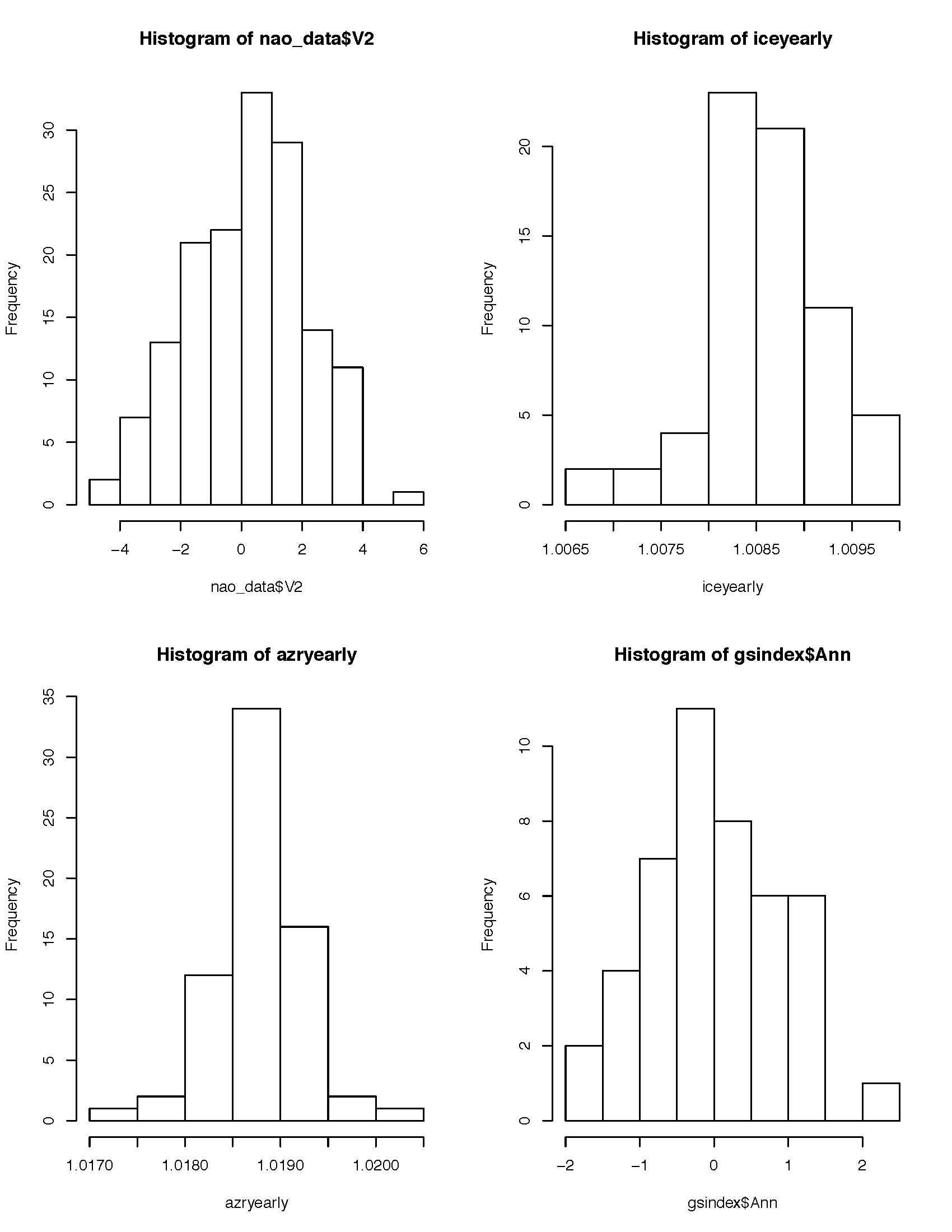


Figure 2. Histogram of annual values for NAO, Icelandic Low and Azores High over 1948 to 2015.

The GS Index time series covers 1966 through 2010.

**2. The Bayesian Model**

We have chosen the following deterministic model to express the first-order relationship between the GSI and the three years of AH and IL data.

As a first order simplification, the following graph omits the first term (the dependence on the GSI of the year before present).

The hand-drawn graph is included below. – This will be modified in the next version of this document.

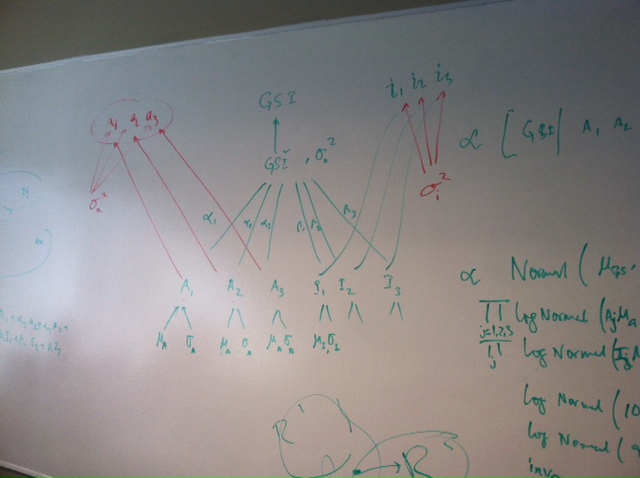
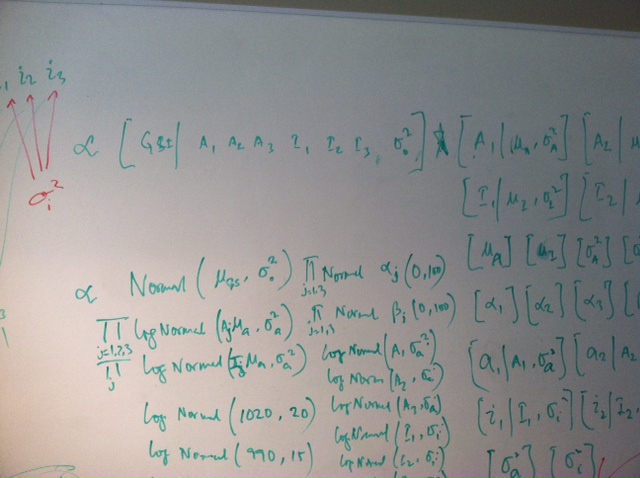


Figure 3. The Bayesian Graph of the GSI-AH-IL model.

Next, based on the graph in Figure 3, the posterior and the joint probability is given in terms of likelihood and priors as in the figure below (Figure 4).

Posterior,

 🡨 Likelihood

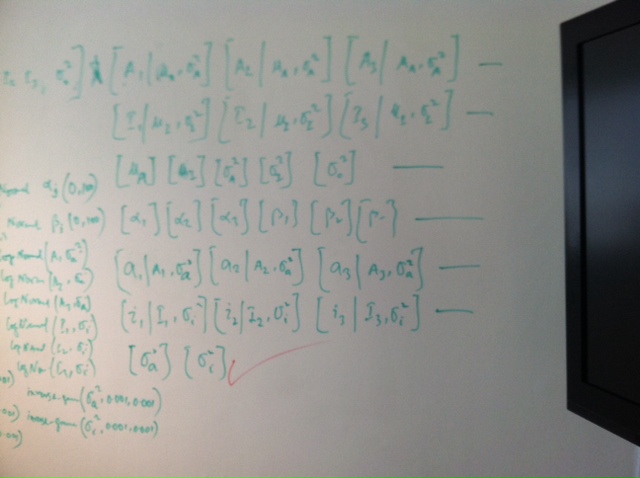
 🡨 Priors

Figure 4. The posterior and joint probabilities with priors for the selected model.

Next, we present our best guess distribution functions for each of the joint and prior distributions below (Figure 5).

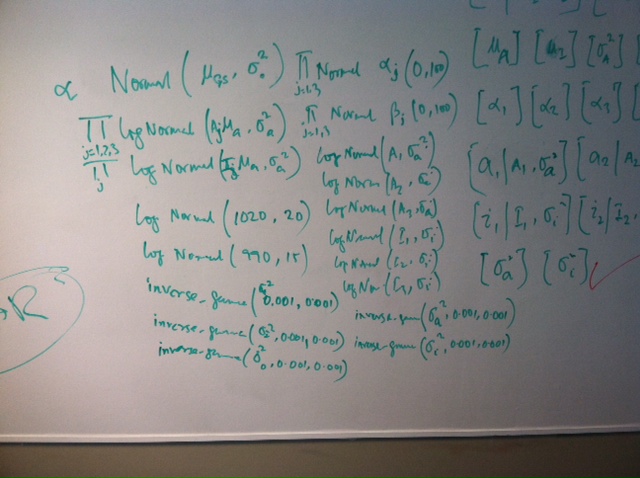


Figure 5. First-guess distributions for likelihood and Priors.

**3. A Simple Model for implementation in JAGS**

Before we implement the full Bayesian model developed above, we started with a simpler version of the model. The deterministic equation for the simpler model is given below.

The JAGS code ( is given below. Note that we are assuming normal distribution for α and β. We assumed gamma distribution for the precision τ.

**3.1 The Code for the Simple Model (R\_GSNAO\_Simple.R)**

rm(list=ls())

gsindex=read.table("~/Desktop/u\_Jags/GSI\_Annual\_text.txt",header=T)

head(gsindex)

ice=read.table("~/Desktop/u\_Jags/IceLow\_text.txt",skip=1,header=T)

head(ice)

iceyearly <- tapply(ice$Pressure,ice$Year,FUN=mean)

ice\_data\_full <- data.frame(Year=names(iceyearly),mean=iceyearly)

ice\_data=ice\_data\_full[19:63,] # selected for 1966 to 2010

#

azr=read.table("~/Desktop/u\_Jags/Azr\_High\_text.txt",skip=1,header=T)

azryearly <- tapply(azr$Pressure,azr$Year,FUN=mean)

azr\_data\_full <- data.frame(Year=names(azryearly),mean=azryearly)

azr\_data= azr\_data\_full[19:63,] # selected for 1966 to 2010

azr\_data

#

gs\_data=gsindex$Ann

ice\_data1=ice\_data$mean

azr\_data1=azr\_data$mean

# Lets keep the same numbers for the Markov Chain below

inits=list(

list( alpha=1, beta=1, tau=.01),# , #chain 1

list( alpha=2, beta=1, tau=.5),#, #chain 2

list( alpha=3, beta=1, tau=.01) #chain 3

)

data1=list(

n=45,

y=as.double(gsindex$Ann),

AH=as.double(ice\_data1),

IL=as.double(azr\_data1)

)

library(rjags)

n.adapt=5000

n.update=10000

n.iter=10000

##call to JAGS

set.seed(1)

jm=jags.model("~/Desktop/u\_Jags/GSNAO\_Simple\_model.txt",data=data1, inits,

n.chains=length(inits), n.adapt = 5000)

update(jm, n.iter=n.update)

zm=coda.samples(jm,variable.names=c("alpha", "beta", "sigma", "tau"),

n.iter=n.iter, n.thin=1)

summary(zm)

plot(zm)

traceplot(zm)

densplot(zm)

gelman.diag(zm)

heidel.diag(zm)

raftery.diag(zm)

**3.2 The model Text file for the simple case – GSNAO\_Simple\_model.txt**

##Gulf Stream and NAO (Iceland-Azores) model

# This is a simple version

model{

#priors

#

# Just alpha and beta

#

alpha ~ dnorm(0,100)

beta ~ dnorm(0,100)

tau ~ dgamma(0.001,0.001) #precision

sigma <- 1/sqrt(tau) # calculate sd from precision

#likelihood -

for(i in 1:n){

mugs[i] <- alpha \* AH[i] + beta \* IL[i]

y[i] ~ dnorm(mugs[i],tau)

}

} #end of model

**3.2 Results from the Simple Model:**

summary(zm)

Iterations = 10001:20000

Thinning interval = 1

Number of chains = 3

Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,

plus standard error of the mean:

Mean SD Naive SE Time-series SE

alpha 0.0005096 0.08585 0.0004956 0.0005711

beta -0.0002142 0.08584 0.0004956 0.0005736

sigma 0.8996720 0.09757 0.0005633 0.0005679

tau 1.2782596 0.27072 0.0015630 0.0015744

2. Quantiles for each variable:

2.5% 25% 50% 75% 97.5%

alpha -0.1669 -0.05845 0.0009081 0.05860 0.1686

beta -0.1693 -0.05805 -0.0004609 0.05767 0.1694

sigma 0.7334 0.83134 0.8914018 0.95905 1.1127

tau 0.8077 1.08721 1.2584993 1.44691 1.8591

> traceplot(zm)

> gelman.diag(zm)

Potential scale reduction factors:

Point est. Upper C.I.

alpha 1 1

beta 1 1

sigma 1 1

tau 1 1

Multivariate psrf

1

> heidel.diag(zm)

[[1]]

Stationarity start p-value

test iteration

alpha passed 1 0.530

beta passed 1 0.498

sigma passed 1 0.910

tau passed 1 0.974

Halfwidth Mean Halfwidth

test

alpha failed -1.92e-05 0.00190

beta failed -3.11e-04 0.00192

sigma passed 9.01e-01 0.00194

tau passed 1.28e+00 0.00536

[[2]]

Stationarity start p-value

test iteration

alpha passed 1 0.396

beta passed 1 0.444

sigma passed 1 0.396

tau passed 1 0.253

Halfwidth Mean Halfwidth

test

alpha failed 0.001442 0.00195

beta failed -0.000436 0.00199

sigma passed 0.899181 0.00194

tau passed 1.278809 0.00538

[[3]]

Stationarity start p-value

test iteration

alpha passed 1 0.955

beta passed 1 0.611

sigma passed 1 0.913

tau passed 1 0.802

Halfwidth Mean Halfwidth

test

alpha failed 0.000106 0.00197

beta failed 0.000105 0.00193

sigma passed 0.899031 0.00191

tau passed 1.279979 0.00530

> raftery.diag(zm)

[[1]]

Quantile (q) = 0.025

Accuracy (r) = +/- 0.005

Probability (s) = 0.95

Burn-in Total Lower bound Dependence

(M) (N) (Nmin) factor (I)

alpha 2 3897 3746 1.040

beta 2 3865 3746 1.030

sigma 2 3680 3746 0.982

tau 2 3650 3746 0.974

[[2]]

Quantile (q) = 0.025

Accuracy (r) = +/- 0.005

Probability (s) = 0.95

Burn-in Total Lower bound Dependence

(M) (N) (Nmin) factor (I)

alpha 2 3929 3746 1.05

beta 2 3962 3746 1.06

sigma 2 3710 3746 0.99

tau 2 3865 3746 1.03

[[3]]

Quantile (q) = 0.025

Accuracy (r) = +/- 0.005

Probability (s) = 0.95

Burn-in Total Lower bound Dependence

(M) (N) (Nmin) factor (I)

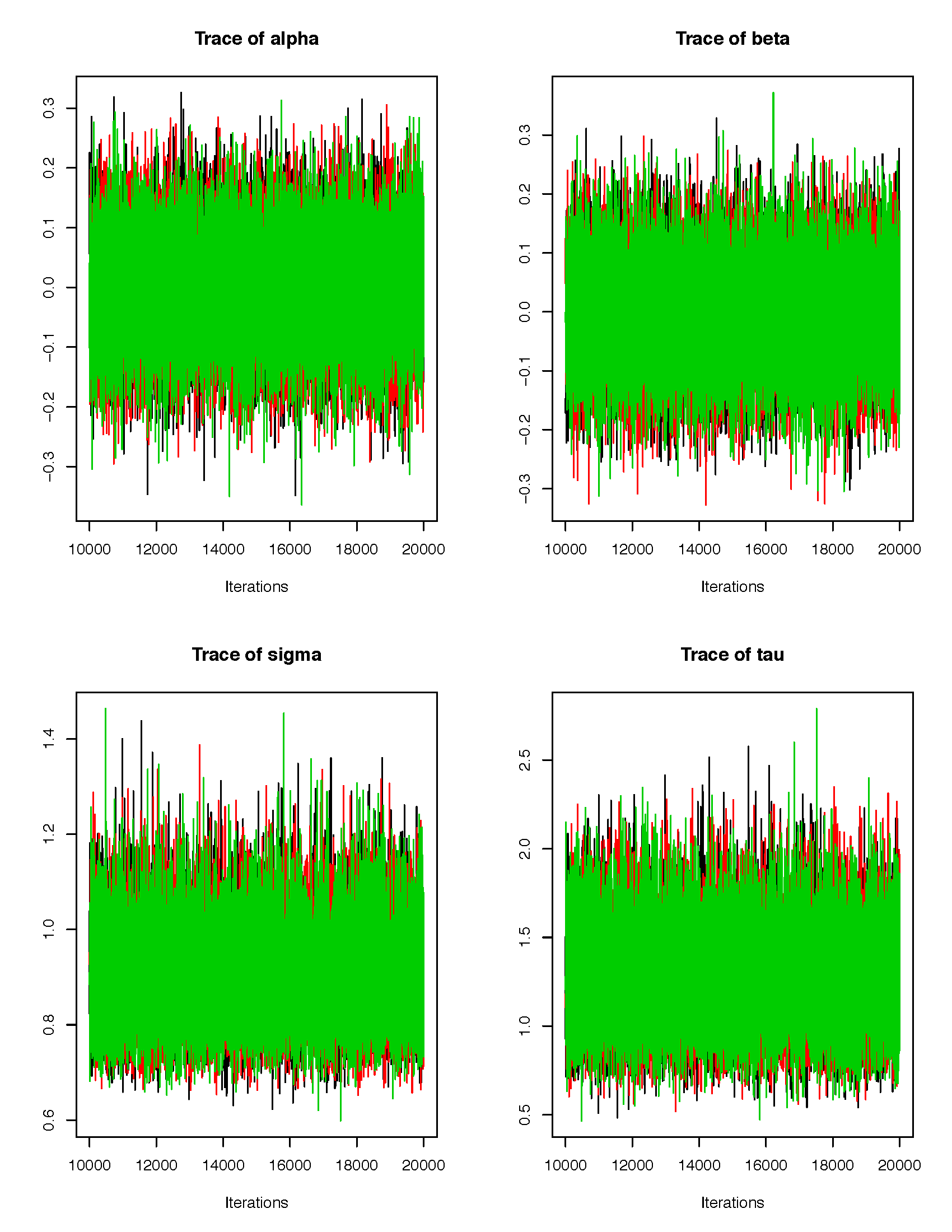
alpha 3 4028 3746 1.080

beta 2 3882 3746 1.040

sigma 2 3680 3746 0.982

tau 2 3680 3746 0.982

The traceplot is shown below.



**Figure 6. Traceplot of alpha, beta, sigma and tau from the simple model**.

**4. A First-order (slightly complex) Model for implementation in JAGS**

Since the lags (one-to-three years) are important to understand and part of the full model, we next developed a first-order model with the following deterministic equations.

Here, ypm1 (i) denotes the Gulf Stream Index (y) for last year (i.e., present minus 1 – pm1); ILPr is the Icelandic Pressure distribution (a known prior – “Normal” for this case, will be linked to data for Icelandic pressure later), ILP (i) is the Icelandic Pressure for the present year (i). The above equation is adapted from a simplification of Sanchez-Franks et al. (2016).

**4.1 The Jags code for a first-order complex model is here (R\_GSNAO\_Complex01.R)**

rm(list=ls())

gsindex=read.table("~/Desktop/u\_Jags/GSI\_Annual\_text.txt",header=T)

head(gsindex)

ice=read.table("~/Desktop/u\_Jags/IceLow\_text.txt",skip=1,header=T)

head(ice)

iceyearly <- tapply(ice$Pressure,ice$Year,FUN=mean)

ice\_data\_full <- data.frame(Year=names(iceyearly),mean=iceyearly)

ice\_datapm2=ice\_data\_full[19:63,] # selected for 1966 to 2010

ice\_datapm1=ice\_data\_full[20:63,]

ice\_datap=ice\_data\_full[21:63,]

#

azr=read.table("~/Desktop/u\_Jags/Azr\_High\_text.txt",skip=1,header=T)

azryearly <- tapply(azr$Pressure,azr$Year,FUN=mean)

azr\_data\_full <- data.frame(Year=names(azryearly),mean=azryearly)

azr\_datapm2= azr\_data\_full[19:63,] # selected for 1966 to 2010

azr\_datapm1= azr\_data\_full[20:63,]

azr\_datap= azr\_data\_full[21:63,]

gsdatap=gsindex$Ann[3:45]

gsdatapm1=gsindex$Ann[2:45]

gsdatapm2=gsindex$Ann[1:45]

# Lets keep the same numbers for the Markov Chain below

inits=list(

list( alpha=1, beta=1, ILPr=1000, tau=.01),# , #chain 1

list( alpha=2, beta=1, ILPr=1005, tau=.5),#, #chain 2

list( alpha=3, beta=1, ILPr= 995, tau=.01) #chain 3

)

data1=list(

n=43,

yp=as.double(gsdatap),

ypm1=as.double(gsdatapm1),

ypm2=as.double(gsdatapm2),

AHp=as.double(azr\_datap$mean),

ILp=as.double(ice\_datap$mean),

AHpm1=as.double(azr\_datapm1$mean),

ILpm1=as.double(ice\_datapm1$mean),

AHpm2=as.double(azr\_datapm2$mean),

ILpm2=as.double(ice\_datapm2$mean)

)

library(rjags)

n.adapt=500

n.update=1000

n.iter=1000

##call to JAGS

set.seed(1)

jm=jags.model("~/Desktop/u\_Jags/GSNAO\_Complex01.txt",data=data1, inits,

n.chains=length(inits), n.adapt = 500)

update(jm, n.iter=n.update)

zm=coda.samples(jm,variable.names=c("ILPr", "alpha", "beta", "sigma", "tau"),

n.iter=n.iter, n.thin=1)

summary(zm)

plot(zm)

traceplot(zm)

densplot(zm)

gelman.diag(zm)

heidel.diag(zm)

raftery.diag(zm)

**4.2 The first-order complex model (GSNAO\_Complex01.txt)**

##Gulf Stream and NAO (Iceland-Azores) model

# This is a simple version

model{

#priors

# Some definitions

# Just AH and IL

#AH ~ dnorm(1024,20)

ILPr ~ dnorm(1001,15)

# Means of azores and iceland are set to 1024 and 1001

#

# alphas and betas

alpha ~ dnorm(0,100)

beta ~ dnorm(0,100)

tau ~ dgamma(0.001,0.001) #precision

sigma <- 1/sqrt(tau) # calculate sd from precision

#likelihood --

for(i in 1:n){

mugs[i] <- 0.40 \* ypm1[i] + alpha \* ILPr + beta \* ILp[i]

yp[i] ~ dnorm(mugs[i],tau)

}

} #end of model

**4.3 Results**

> summary(zm)

Iterations = 1001:2000

Thinning interval = 1

Number of chains = 3

Sample size per chain = 1000

1. Empirical mean and standard deviation for each variable,

plus standard error of the mean:

Mean SD Naive SE Time-series SE

ILPr 1.001e+03 0.2567412 4.687e-03 4.674e-03

alpha -1.285e-06 0.0001538 2.808e-06 4.475e-06

beta -2.134e-03 0.1005603 1.836e-03 2.968e-03

sigma 7.506e-01 0.0843823 1.541e-03 1.541e-03

tau 1.841e+00 0.4050713 7.396e-03 7.395e-03

2. Quantiles for each variable:

2.5% 25% 50% 75% 97.5%

ILPr 1.000e+03 1.001e+03 1.001e+03 1.001e+03 1.002e+03

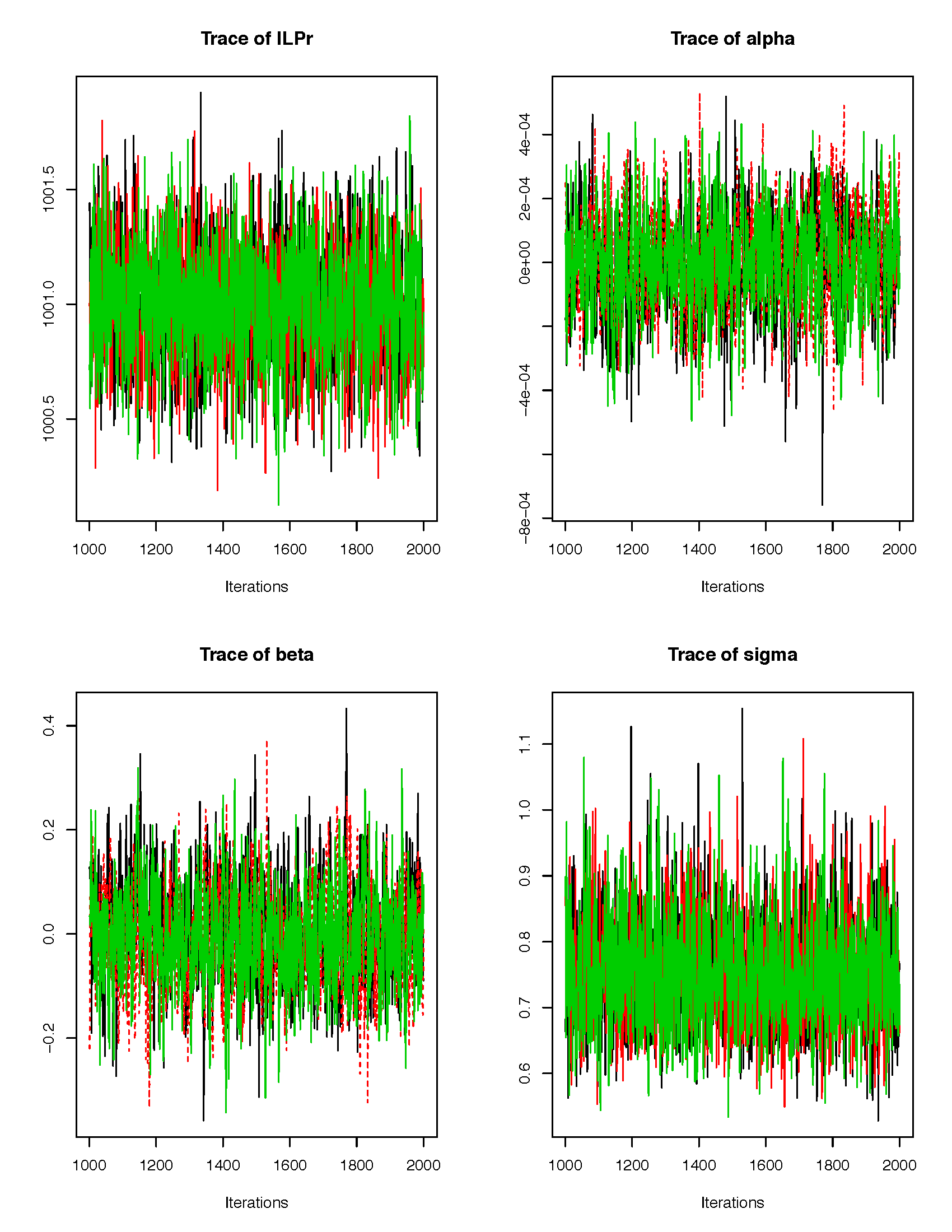
alpha -3.037e-04 -1.043e-04 3.375e-06 9.900e-05 2.959e-04

beta -1.964e-01 -7.075e-02 -3.447e-03 6.639e-02 1.948e-01

sigma 6.070e-01 6.905e-01 7.443e-01 8.017e-01 9.312e-01

tau 1.153e+00 1.556e+00 1.805e+00 2.098e+00 2.714e+00

**The traceplot is shown below.**

****

**Figure 7. Traceplot from the first-order complex model of ILPr, alpha, beta and sigma.**

**5. Future Steps**

5.1 Understand the convergence process of the simple and complex models

5.2 Develop the full model

5.3 Maybe develop a series of alternate models to understand the sensitivity of adding one-year, two-year and three-year lags.

5.4 Is it possible to find the optimal number of years of lag from the sensitivity analysis of the Bayesian modeling?

5.5 How do we interpret this for a predictive system for the GSI?

5.6 Many other questions will arise – longitudinal dependence of GSI; increasing the number of years in the dataset? Using a different measure of index (from altimeter data), adding the Labrador current index in the mix; relate to WCRs.