Bayesian Inference on Factors Influencing Rental Bikes Demand

1 Introduction

Rental bikes are a current popular method for sharing economic. It is important for government or companies to supply the number of bikes efficacious. In this analysis project, we focus on the bike sharing demand in Seoul, Korea and try to explore the relationship between hourly rented bikes demand and some influencing factors. The data to be used for analysis is "SeoulBikeData.csv", which can be found on UCI Machine Learning Data Repository (2020). We wish to build a Poisson regression model on all observed data and use Bayesian methods to find the posterior distribution of the coefficients. We will try to discover how differnt priors on β will affect the posterior distribution. We suggest three priors, a flat prior, a normal prior with Inverse-Gamma variance, and a normal prior with Half-Cauchy vairance. In selection 2 and 3, hyperpriors are involved and therefore hierarchical models need to be implemented. In the methods section, we introduce our three models, state the likelihood function, and derive conditional posterior distributions. In the results section, we provide summaries for the coefficient posterior estimates for all three models. In the discussion section, we comment on the results of the three different priors and suggest a best recommendation.

2 Methods

We consider the following model. Let $y_i, i = 1, 2, \cdots, n$ represent rental bike demand each day where n is the total number of days. We assume y_i's are independent Poisson with parameter λ_i. That is, $y_i \sim Poisson(\lambda_i)$. Since rental bike demands are count data, we may suggest a Poisson Regression Model to predict the rental demand. That is, $\lambda = \mathbf{X}\beta + \mu$, where λ is the vector of all parameters λ_i, $\mathbf{X}$ is the design matrix, β is the vector of coefficients, and μ is the error term centered at 0 and distributed normally. In our analysis, we focus on choosing priors for β. And since the error term is relatively small, we drop μ from the model. Our likelihood is

$$\mathcal{L}(\mathbf{y}|\beta, \mathbf{X}) \propto \prod_{i=1}^{n} (X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta\right)$$

We explore the effect on the coefficents β using different priors on β.

Model 1: Flat prior

Using a flat prior β, $\pi(\beta) \propto 1$, we have posterior as

$$p(\beta|\mathbf{X}, \mathbf{y}) \propto \prod_{i=1}^{n} (X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta\right)$$

Model 2: Normal prior with Inverse-Gamma variance

Using a multivariate normal prior $\beta \sim MVN(0, \sigma^2 I_{n\times n}), \sigma^2 \sim IG(a, b)$, we have the full conditionals as

$$p(\beta|\sigma^2, \mathbf{X}, \mathbf{y}) \propto \prod_{i=1}^{n} (X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta - \frac{1}{2\sigma^2}\beta'\beta\right)$$

$$\sigma^2|\beta, \mathbf{X}, \mathbf{y} \sim IG\left(a + \frac{1}{2}, \frac{\beta'\beta}{2} + b\right)$$

Model 3: Normal prior with Half-Cauchy variance

Using a multivariate normal prior$\beta \sim MVN(0, \sigma^2 I_{n\times n}), \sigma^2 \sim HC(A)$. We use a mixture of Inverse Gammas in this case, $\sigma^2 \sim IG(\frac{1}{2}, \frac{1}{\alpha}), \alpha \sim IG(\frac{1}{2}, \frac{1}{A^2})$. We have the full conditionals as

$$p(\beta|\sigma^2, \alpha, \mathbf{X}, \mathbf{y}) \propto \prod_{i=1}^{n} (X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta - \frac{1}{2\sigma^2}\beta'\beta\right)$$
$$\sigma^2|\beta, \alpha, \mathbf{X}, \mathbf{y} \sim IG\left(1, \frac{\beta'\beta}{2} + \frac{1}{\alpha}\right)$$
$$\alpha|\beta, \sigma^2, \mathbf{X}, \mathbf{y} \sim IG\left(1, \frac{1}{\sigma^2} + \frac{1}{A^2}\right)$$

For each parameter in the models, we draw Monte Carlo chains of 50000 samples each (See Appendix B for all code). In model 1, we use M-H sampler to draw samples from the posterior distribution for parameter β. We use normal distribution for approximation. For model 2 and 3, we use a Gibbs sampler with M-H step to draw samples from conditional posterior distribution for parameters. We use M-H algorithem with normal distribution to approximate parameter β. We use Gibbs sampler to draw samples from the conditional posterior distributions for σ^2 (model 2) and σ^2, α (model 3). The conditional posterior distributions are calculated as inverse gamma distributions (See Appendix A for derivation). Finally, we compare the distributions of the posterior samples of the coefficients to understand how the choice of prior affect the estimates of coefficient. Remeber in the dataset, variables "Seasons", "Holiday", "Functioning Day" are categorical instead of numerical. We first need to prepossess the data and change them to dummy variables.

3 Result

From the results, we know that the acceptance rates of β in three models are 57.78%, 47.95% and 52.66% accordingly, which are all close to the ideal acceptance rate. All coefficients for beta converge except the intercept (β_0) and functioning day β_{13} in model1 (see Appendix C). However, the σ^2 and α are not converge. The summary statistics are displayed in the table below. The plots of the posterior densities for β show that the density is unimodal and roughly symmetric (see Appendix C). All three models produce similar results where the coefficient estimates are nearly equal and the credible intervals are generally comparable except β_0 and functioning day β_{13} in model1. Therefore, the choice of prior for β does not appear to influence the posterior distributions.
What's more, the credible intervals for β doesn't contains 0 except "visibility" in all three models. We conclude that the "visibility" don't affect the prediction and we should not include this variable in the model. In all three models, covariates that have a negative relationship with response variable bikes rented per hour include: "humidity", "solar radiation", "rainfall", "snowfall", and "holiday", and the covariates that have a positive relationship include: "temperature", "wind speed", "dew point temperature", "seasons", and "functioning day".

	model1			model2			model3		
	50.000%	2.500%	97.500%	50.000%	2.500%	97.500%	50.000%	2.500%	97.500%
(Intercept)	-497.579	-788.947	-196.485	-10.916	-60.570	-5.756	-15.933	-147.071	-6.144
Temperature	0.036	0.036	0.037	0.036	0.036	0.037	0.036	0.036	0.037
Humidity	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016
Wind_speed	0.095	0.094	0.096	0.095	0.094	0.096	0.095	0.094	0.096
Visibility	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Dew_point_temperature	0.008	0.007	0.008	0.008	0.007	0.008	0.008	0.007	0.008
Solar_Radiation	-0.164	-0.165	-0.163	-0.164	-0.165	-0.163	-0.164	-0.165	-0.163
Rainfall	-0.490	-0.494	-0.486	-0.490	-0.494	-0.486	-0.490	-0.494	-0.486
Snowfall	-0.073	-0.077	-0.070	-0.073	-0.077	-0.070	-0.073	-0.077	-0.069
Autumn	0.881	0.877	0.885	0.881	0.877	0.885	0.881	0.877	0.885
Spring	0.665	0.661	0.669	0.665	0.661	0.669	0.665	0.661	0.669
Summer	0.561	0.556	0.566	0.561	0.556	0.566	0.561	0.556	0.566
holiday	-0.209	-0.213	-0.205	-0.209	-0.214	-0.205	-0.209	-0.214	-0.205
functioning day	503.917	202.833	795.294	17.252	12.101	66.916	22.274	12.483	153.406

	Model1	Model2	Model3
β_0: (Intercept)	3.04	1.04	1.02
β_1: Temperature	1	1	1
β_2: Humidity	1	1	1
β_3: Wind_speed	1.02	1.01	1
β_4: Visibility	1	1.01	1
β_5: Dew_point_temperature	1	1	1
β_6: Solar_Radiation	1	1	1.01
β_7: Rainfall	1	1	1.01
β_8: Snowfall	1	1	1
β_9: Autumn	1.02	1	1
β_{10}: Spring	1.02	1.01	1.01
β_{11}: Summer	1.01	1.01	1
β_{12}: holiday	1.01	1.02	1
β_{13}: functioning day	3.04	1.04	1.02
σ^2	-	1.25	1.24
α	-	-	1.27

4 Discussion

The choice of prior for the variance parameter does not appear to affect the coefficient estimates in this Seoul bike data. All three models produced almost identical point estimates for the coefficients and credible intervals. If we use non-informative prior, the beta doesn't converge. If we use mean zero normal prior with Half-Cauchy variance,the model is more complex than the other two models. Thus, for the simplicity and convergence, we should choose mean zero normal prior with Inverse-Gamma variance (model2).
For the future work, it could be helpful if we delete the "visibility" , and since the posterior distribution of the "functioning day" is skewed, by deleting it might also help the performance in model. Moreover, our model missed the μ part in poisson model,however in reallife, it might not be zero, therefore in the future work, we can also try to assume that the μ follows different prior and do the sampler. Despite these limitations, our results are significantly better than random.

Reference

UC Irvine Machine Learning Repository. (2020). https://archive.ics.uci.edu/ml/datasets/Seoul+Bike+Sharing+Demand.

Appendix A: Derivations

Poisson Regression Model

Let $y_i, i = 1, 2, \cdots, n$ represent rental bike demand each day where n is the total number of days observed in dataset. Let $\mathbf{y}$ be a vector of all observed rental bike demand, $i.e., \mathbf{y} = (y_1, y_2, \cdots, y_n)'$. We assume y_i's are independent Poisson with parameter λ_i

$$y_i \sim Poisson(\lambda_i), i = 1, 2, \cdots, n$$

Let λ be a vector of all parameters $\lambda_i, i = 1, 2, \cdots, n$. Since rental bike demands are count data, we may suggest a Poisson Regression Model to predict the rental demand. That is,

$$\lambda = \mathbf{X}\beta + \mu$$

where $\mathbf{X}$ is the design matrix. β is the vector of coefficients, and μ is the error term centered at 0 and distributed normally. In our analysis, we focus on choosing priors for β.
We drop μ from the model

$$\lambda = \mathbf{X}\beta$$

Likelihood

For each $\lambda_i, i = 1, 2, \cdots, n$,

$$\lambda_i = X_i\beta$$

where X_i is the ith row of the design matrix.
Substituing the above into the distribution of y_i,

$$p(y_i) = \frac{(X_i\beta)^{y_i} \exp -(X_i\beta)}{y_i!}$$

Likelihood

$$\begin{aligned}\mathcal{L}(\mathbf{y}|\beta, \mathbf{X}) &\propto \prod_{i=1}^{n}(X_i\beta)^{y_i} \exp -(X_i\beta) \\ &= \prod_{i=1}^{n}(X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta\right)\end{aligned}$$

Posteriors

Model 1: Flat prior

We suggest a flat prior on β, $\pi(\beta) \propto 1$.
Posterior

$$\begin{aligned}p(\beta|\mathbf{X}, \mathbf{y}) &\propto \mathcal{L}(\mathbf{y}|\beta, \mathbf{X}\pi(\beta) \\ &\prod_{i=1}^{n}(X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta\right)\end{aligned}$$

This distribution is not recognizable.

Model 2: Normal prior with Inverse-Gamma variance

We suggest a multivariate normal prior on the coefficents $\beta \sim MVN(0, \sigma^2 I_{n\times n}), \sigma^2 \sim IG(a, b)$.
Priors:

$$\begin{aligned}\pi(\beta|\sigma^2) &\propto (\sigma^2)^{-1/2} \exp\left[-\frac{1}{2}\beta'(\sigma^2 I_{n\times n})^{-1}\beta\right] \\ &= (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}\beta'\beta\right)\end{aligned}$$

$$\pi(\sigma^2) \propto (\sigma^2)^{-(a+1)} \exp\left(-\frac{b}{\sigma^2}\right)$$

Posterior

$$\begin{aligned}
p(\beta, \sigma^2|\mathbf{X}, \mathbf{y}) &\propto \mathcal{L}(\mathbf{y}|\beta, \sigma^2, \mathbf{X}\pi(\beta|\sigma^2)\pi(\sigma^2) \\
&= \prod_{i=1}^{n} (X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta\right) (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}\beta'\beta\right) (\sigma^2)^{-(a+1)} \exp\left(-\frac{b}{\sigma^2}\right)
\end{aligned}$$

Full conditionals

$$\begin{aligned}
p(\beta|rest) &\propto \prod_{i=1}^{n} (X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta\right) \exp\left(-\frac{1}{2\sigma^2}\beta'\beta\right) \\
&= \prod_{i=1}^{n} (X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta - \frac{1}{2\sigma^2}\beta'\beta\right)
\end{aligned}$$

This distribution is not recognizable.

$$\begin{aligned}
p(\sigma^2|rest) &\propto (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}\beta'\beta\right) (\sigma^2)^{-(a+1)} \exp\left(-\frac{b}{\sigma^2}\right) \\
&= (\sigma^2)^{-(a+\frac{1}{2}+1)} \exp\left[-\frac{1}{\sigma^2}\left(\frac{\beta'\beta}{2} + b\right)\right]
\end{aligned}$$

We recognize it as a kernel of an Inverse Gamma distribution.

$$\sigma^2|rest \sim IG\left(a + \frac{1}{2}, \frac{\beta'\beta}{2} + b\right)$$

Model 3: Normal prior with Half-Cauchy variance

We suggest a multivariate normal prior on the coefficents $\beta \sim MVN(0, \sigma^2 I_{n\times n}), \sigma^2 \sim IG(\frac{1}{2}, \frac{1}{\alpha}), \alpha \sim IG(\frac{1}{2}, \frac{1}{A^2})$. This is equivalent to a Half Cauchy prior with scale A.
Priors:

$$\begin{aligned}
\pi(\beta|\sigma^2) &\propto (\sigma^2)^{-1/2} \exp\left[-\frac{1}{2}\beta'(\sigma^2 I_{n\times n})^{-1}\beta\right] \\
&= (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}\beta'\beta\right)
\end{aligned}$$

$$\pi(\sigma^2|\alpha) \propto \alpha^{-\frac{1}{2}} (\sigma^2)^{-(\frac{1}{2}+1)} \exp\left(-\frac{\frac{1}{\alpha}}{\sigma^2}\right)$$

$$\pi(\alpha) \propto \alpha^{-(\frac{1}{2}+1)} \exp\left(-\frac{\frac{1}{A^2}}{\alpha}\right)$$

Posterior

$$\begin{aligned}
p(\beta, \sigma^2, \alpha|\mathbf{X}, \mathbf{y}) &\propto \mathcal{L}(\mathbf{y}|\beta, \sigma^2, \mathbf{X}\pi(\beta|\sigma^2)\pi(\sigma^2) \\
&= \prod_{i=1}^{n} (X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta\right) (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}\beta'\beta\right) \alpha^{-\frac{1}{2}} (\sigma^2)^{-(\frac{1}{2}+1)} \exp\left(-\frac{\frac{1}{\alpha}}{\sigma^2}\right) \alpha^{-(\frac{1}{2}+1)} \exp\left(-\frac{\frac{1}{A^2}}{\alpha}\right)
\end{aligned}$$

Full conditionals

$$\begin{aligned}
p(\beta|rest) &= \prod_{i=1}^{n} (X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta\right) \exp\left(-\frac{1}{2\sigma^2}\beta'\beta\right) \\
&= \prod_{i=1}^{n} (X_i\beta)^{y_i} \exp\left(-\sum_{i=1}^{n} X_i\beta - \frac{1}{2\sigma^2}\beta'\beta\right)
\end{aligned}$$

This distribution is not recognizable.

$$
\begin{aligned}
p(\sigma^2|rest) &\propto (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}\beta'\beta\right) (\sigma^2)^{-(\frac{1}{2}+1)} \exp\left(-\frac{\frac{1}{\alpha}}{\sigma^2}\right) \\
&= (\sigma^2)^{-(1+1)} \exp\left[-\frac{1}{\sigma^2}\left(\frac{\beta'\beta}{2} + \frac{1}{\alpha}\right)\right]
\end{aligned}
$$

We recognize it as a kernel of an Inverse Gamma distribution.

$$
\sigma^2|rest \sim IG\left(1, \frac{\beta'\beta}{2} + \frac{1}{\alpha}\right)
$$

$$
\begin{aligned}
p(\alpha|rest) &\propto \alpha^{-\frac{1}{2}} \exp\left(-\frac{\frac{1}{\alpha}}{\sigma^2}\right) \alpha^{-(\frac{1}{2}+1)} \exp\left(-\frac{\frac{1}{A^2}}{\alpha}\right) \\
&= \alpha^{-(1+1)} \exp\left[-\frac{1}{\alpha}\left(\frac{1}{\sigma^2} + \frac{1}{A^2}\right)\right]
\end{aligned}
$$

We recognize it as a kernel of an Inverse Gamma distribution.

$$
\alpha|rest \sim IG\left(1, \frac{1}{\sigma^2} + \frac{1}{A^2}\right)
$$

Appendix B: Code

```
model1 <- function(B,tau,seed,x,y){

  ################regression###################
  func1 <- glm(y ~Temperature+Humidity+Wind_speed+ Visibility+Dew_point_temperature+Solar_Radiation+Rainfall+Snowfall+Autumn
+Spring+Summer+holiday+functioningday,data=df,family = poisson(link='log')) #
  bhat <- coef(func1)
  vbeta <- vcov(func1)

  ################parameters####################
  beta      <- matrix(0, nrow = B, ncol = ncol(x))
  ar            <- vector('numeric', length = B)

  beta[1,]  <- bhat

  tdens <- function(b, X, Y){
    sum(Y*(X%*%b) - exp(X%*%b))
  }

  tau   <- tau # need to tune this
  B= B
  ################M-H steps###################
  set.seed(seed)
  for(t in 2:B){

    bstar   <- rmvnorm(1, beta[t-1,], tau*vbeta)
    r       <- exp(tdens(t(bstar), x, y)-tdens(beta[t-1,], x, y))
    U       <- runif(1)
    if(U < min(1,r)){
      beta[t,]  <- bstar
      ar[t]       <- 1
    } else{
      beta[t,]  <- beta[t-1,]
      ar[t]       <- 0
    }
  }
  out <- NULL
  out$ar <- ar
  out$beta <-beta
  return(out)
}
```

```r
model2 <- function(B,tau,seed,x,y,a,b,sigma_start){

  ################regression##################
  func1 <- glm(y ~Temperature+Humidity+Wind_speed+ Visibility+Dew_point_temperature+Solar_Radiation+Rainfall+Snowfall+Autumn
+Spring+Summer+holiday+functioningday,data=df,family = poisson(link='log')) #
  bhat <- coef(func1)
  vbeta <- vcov(func1)

  ################parameters###################
  beta      <- matrix(0, nrow = B, ncol = ncol(x))
  ar            <- vector('numeric', length = B)
  sigma2   <- vector('numeric', length = B)

  beta[1,]  <- bhat
  sigma2[1] <- sigma_start

  tdens <- function(b, X, Y,sigma2){
    likelihood <-sum(Y*(X%*%b) - exp(X%*%b))
    prior <- (-1/(2*sigma2)*(t(b)%*%b))
    post <- likelihood + prior
    return(post)
  }

  tau   <- tau # need to tune this
  B= B
  ################M-H steps###################
  set.seed(seed)
  for(t in 2:B){

    #sample beta
    bstar   <- rmvnorm(1, beta[t-1,], tau*vbeta)
    r       <- exp(tdens(t(bstar), x, y,sigma2[t-1])-tdens(beta[t-1,], x, y,sigma2[t-1]))
    #print(r)
    U       <- runif(1)
    if(U < min(1,r)){
      beta[t,]  <- bstar
      ar[t]       <- 1
    } else{
      beta[t,]  <- beta[t-1,]
      ar[t]       <- 0
    }

    #sample sigma2 for beta
    sigma2[t] <- rinvgamma(1,a+0.5,0.5*(t(beta[t,])%*%beta[t,])+b)
  }
  out <- NULL
  out$ar <- ar
  out$beta <-beta
  out$sigma2 <-sigma2
  return(out)
}
```

```r
model3 <- function(B,tau,seed,x,y,sigma_start,A,alpha_start){

  ################regression##################
  func1 <- glm(y ~Temperature+Humidity+Wind_speed+ Visibility+Dew_point_temperature+Solar_Radiation+Rainfall+Snowfall+Autumn
+Spring+Summer+holiday+functioningday,data=df,family = poisson(link='log')) #
  bhat <- coef(func1)
  vbeta <- vcov(func1)

  ################parameters##################
  beta      <- matrix(0, nrow = B, ncol = ncol(x))
  ar            <- vector('numeric', length = B)
  sigma2   <- vector('numeric', length = B)
  alpha <- vector('numeric', length = B)

  beta[1,]  <- bhat
  sigma2[1] <- sigma_start
  alpha[1] <- alpha_start

  tdens <- function(b, X, Y,sigma2){
    likelihood <-sum(Y*(X%*%b) - exp(X%*%b))
    prior <- (-1/(2*sigma2)*(t(b)%*%b))
    post <- likelihood + prior
    return(post)
  }

  tau   <- tau # need to tune this
  B= B
  ################M-H steps##################
  set.seed(seed)
  for(t in 2:B){

    #sample beta
    bstar   <- rmvnorm(1, beta[t-1,], tau*vbeta)
    r       <- exp(tdens(t(bstar), x, y,sigma2[t-1])-tdens(beta[t-1,], x, y,sigma2[t-1]))
    #print(r)
    U       <- runif(1)
    if(U < min(1,r)){
      beta[t,]  <- bstar
      ar[t]       <- 1
    } else{
      beta[t,]  <- beta[t-1,]
      ar[t]       <- 0
    }

    #sample sigma2 for beta
    sigma2[t] <- rinvgamma(1,1,0.5*(t(beta[t,])%*%beta[t,])+1/alpha[t-1])

    alpha[t] <- rinvgamma(1,1,1/A^2+1/sigma2[t]^2)
  }
  out <- NULL
  out$ar <- ar
  out$beta <-beta
  out$sigma2 <-sigma2
  out$alpha <-alpha
  return(out)
}
```

Appendix C: Results Analysis

Coefficient

	model1			model2			model3		
	50.000%	2.500%	97.500%	50.000%	2.500%	97.500%	50.000%	2.500%	97.500%
(Intercept)	-497.579	-788.947	-196.485	-10.916	-60.570	-5.756	-15.933	-147.071	-6.144
Temperature	0.036	0.036	0.037	0.036	0.036	0.037	0.036	0.036	0.037
Humidity	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016
Wind_speed	0.095	0.094	0.096	0.095	0.094	0.096	0.095	0.094	0.096
Visibility	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Dew_point_temperature	0.008	0.007	0.008	0.008	0.007	0.008	0.008	0.007	0.008
Solar_Radiation	-0.164	-0.165	-0.163	-0.164	-0.165	-0.163	-0.164	-0.165	-0.163
Rainfall	-0.490	-0.494	-0.486	-0.490	-0.494	-0.486	-0.490	-0.494	-0.486
Snowfall	-0.073	-0.077	-0.070	-0.073	-0.077	-0.070	-0.073	-0.077	-0.069
Autumn	0.881	0.877	0.885	0.881	0.877	0.885	0.881	0.877	0.885
Spring	0.665	0.661	0.669	0.665	0.661	0.669	0.665	0.661	0.669
Summer	0.561	0.556	0.566	0.561	0.556	0.566	0.561	0.556	0.566
holiday	-0.209	-0.213	-0.205	-0.209	-0.214	-0.205	-0.209	-0.214	-0.205
functioning day	503.917	202.833	795.294	17.252	12.101	66.916	22.274	12.483	153.406

Running mean plot

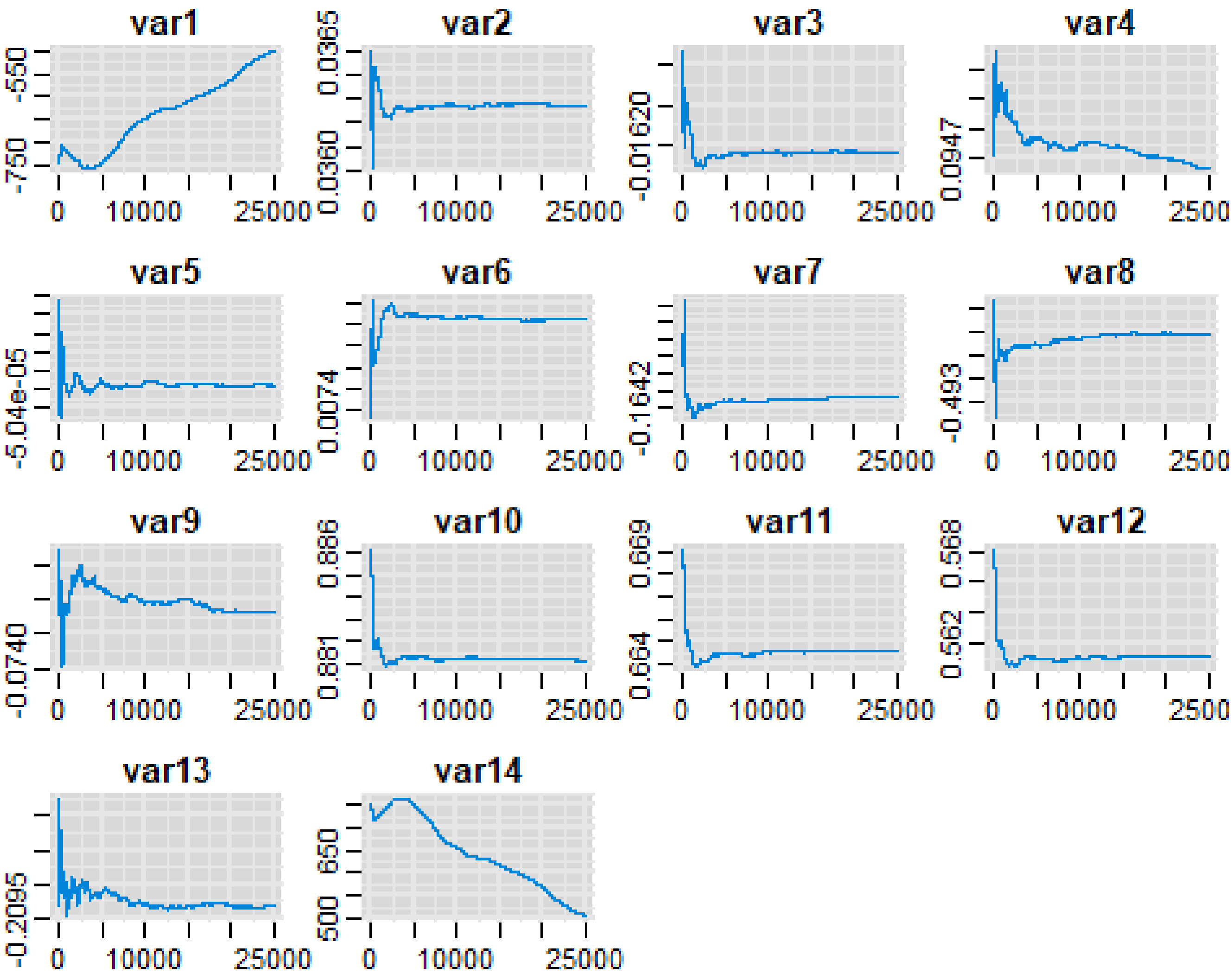


Figure 1: Running mean plot for β in model1

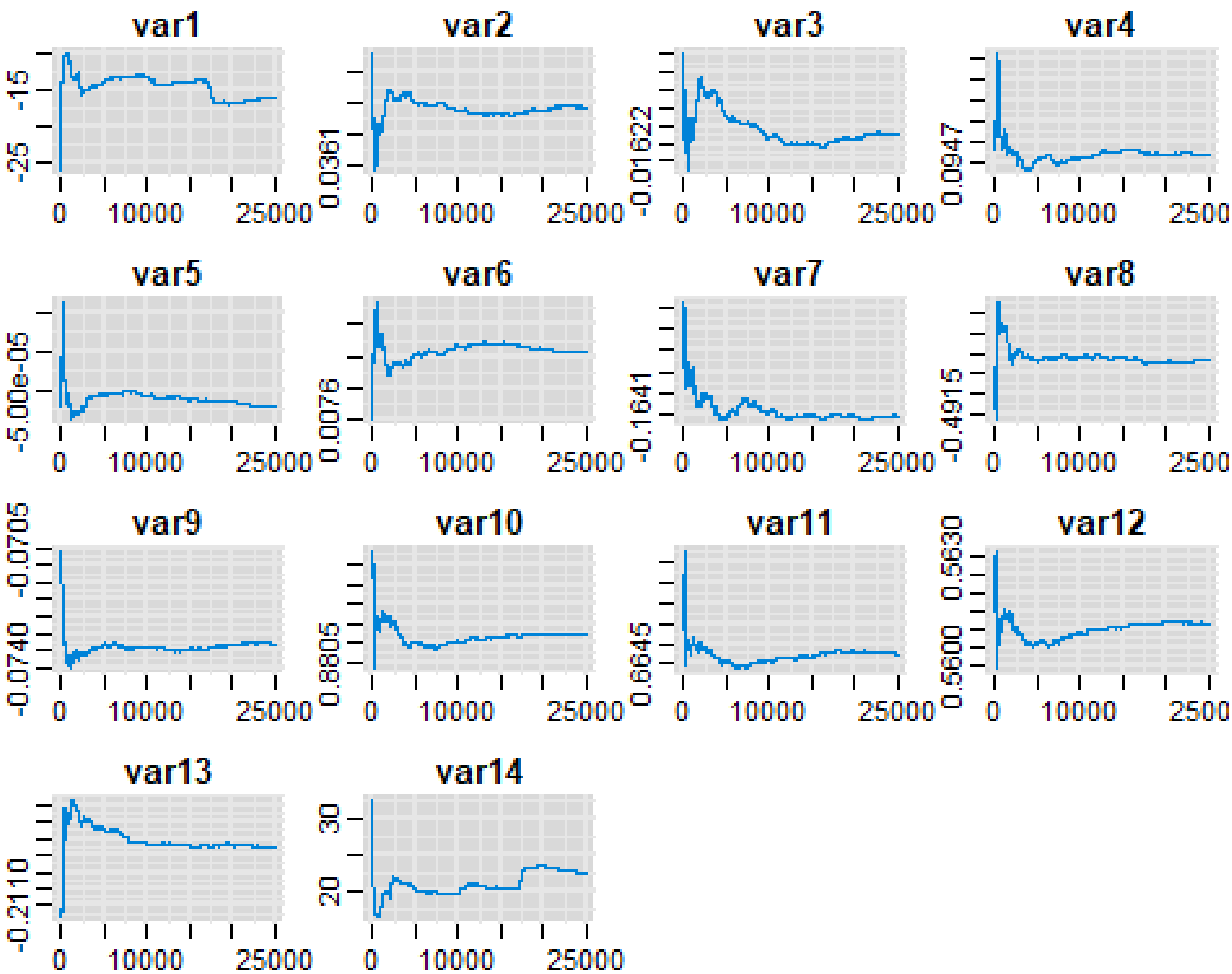


Figure 2: Running mean plot for β in model2

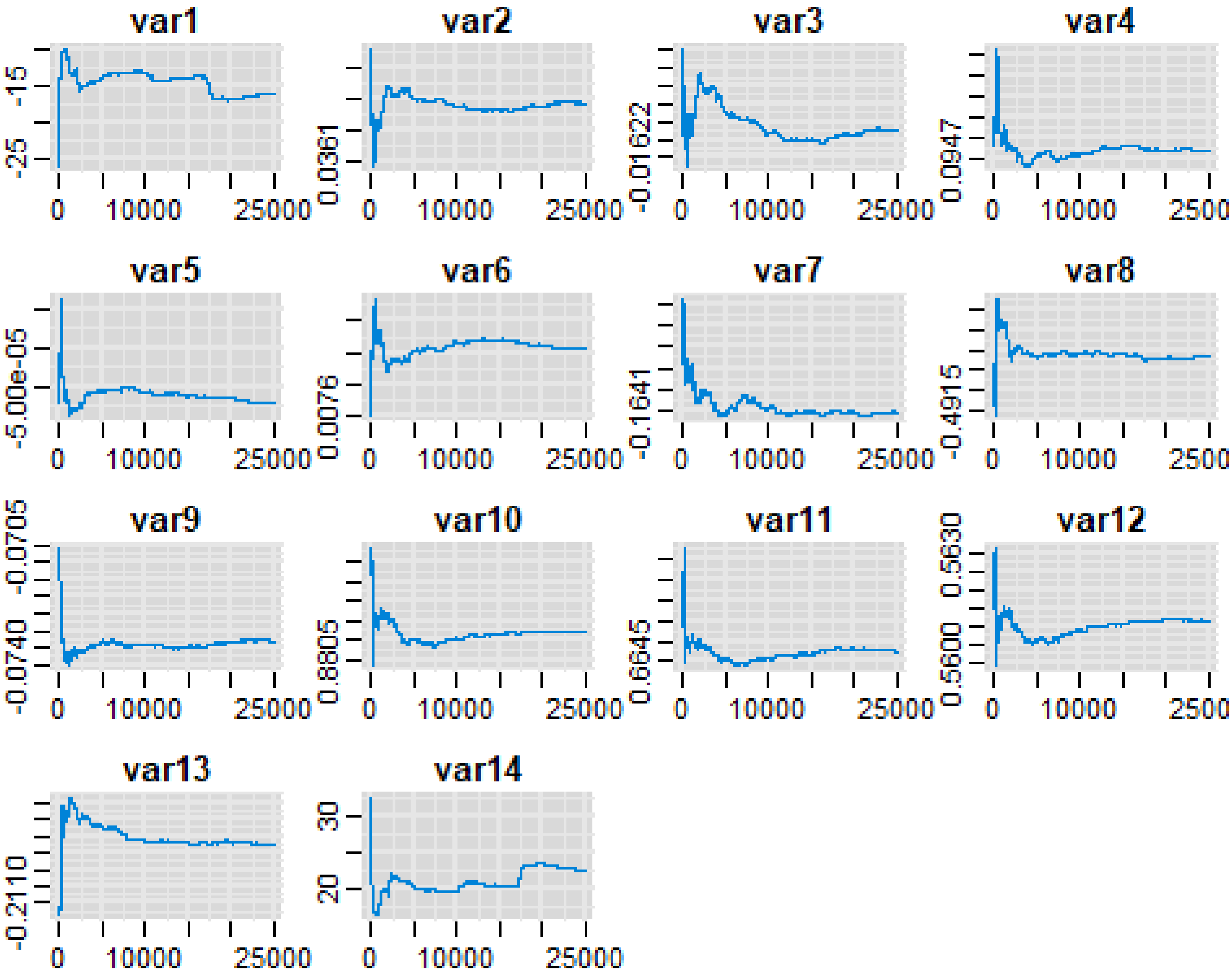


Figure 3: Running mean plot for β in model3

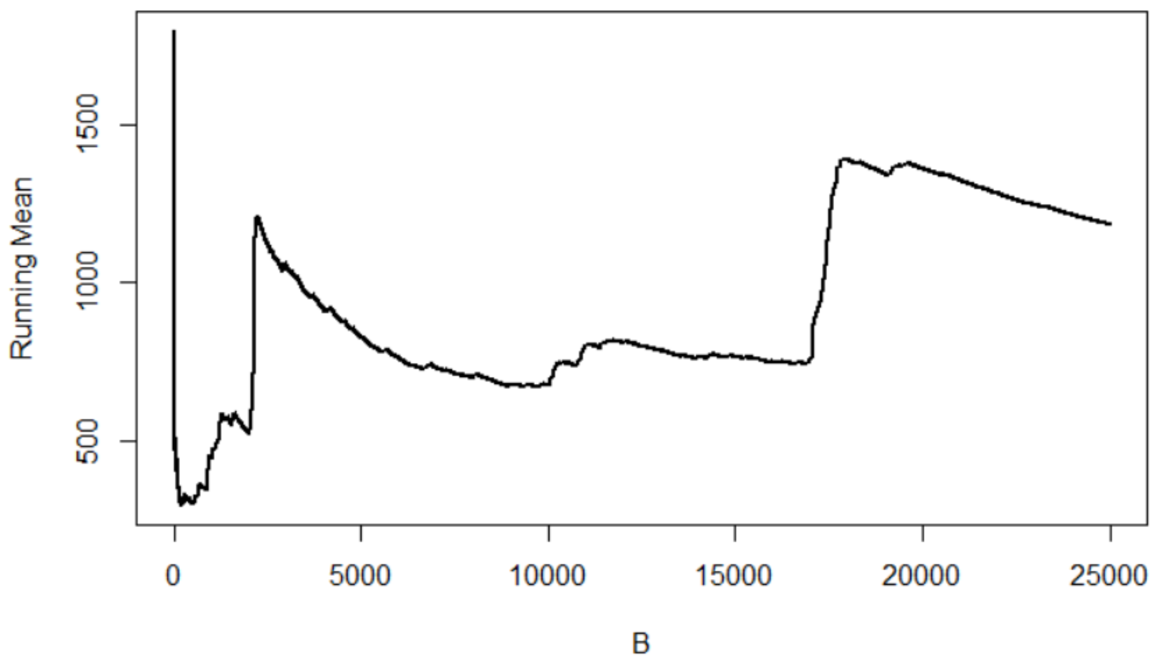


Figure 4: Running mean plot for σ^2 in model2

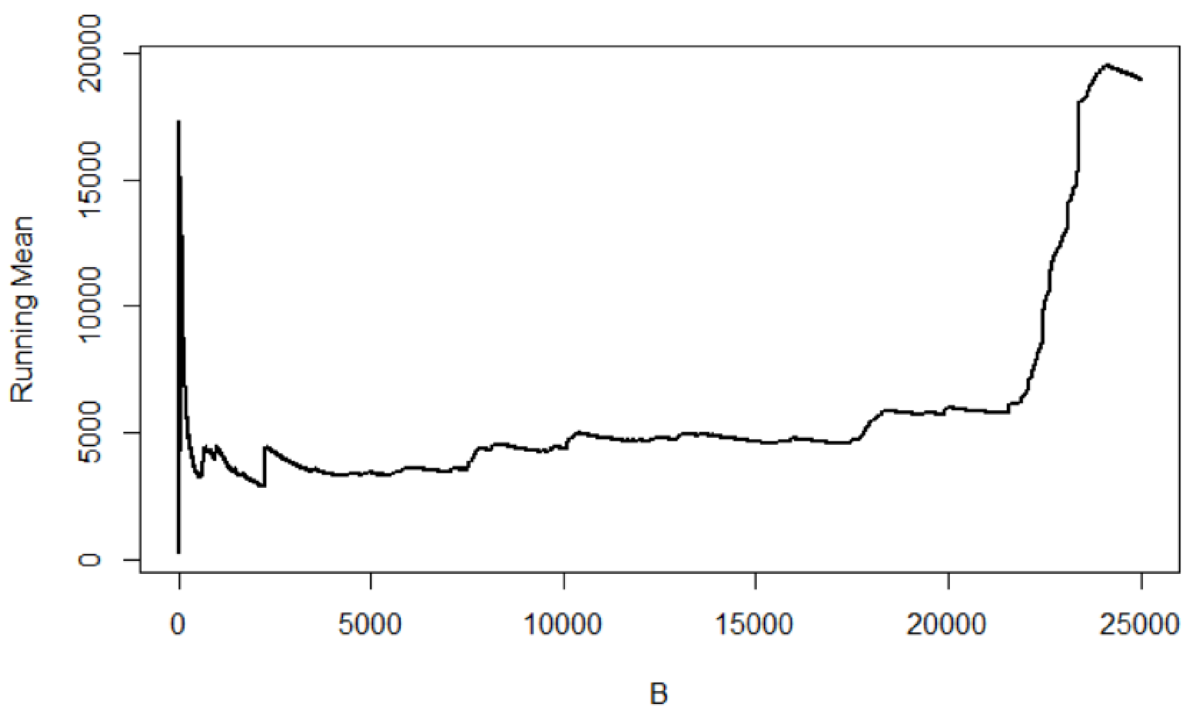


Figure 5: Running mean plot for σ^2 in model3

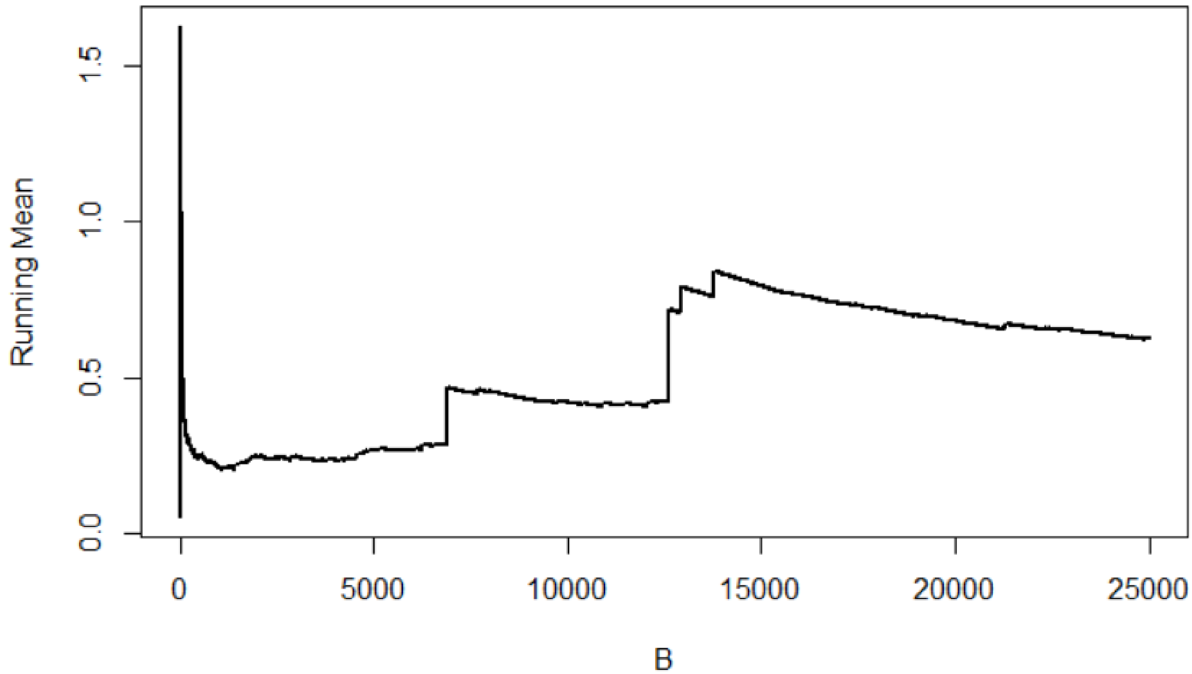
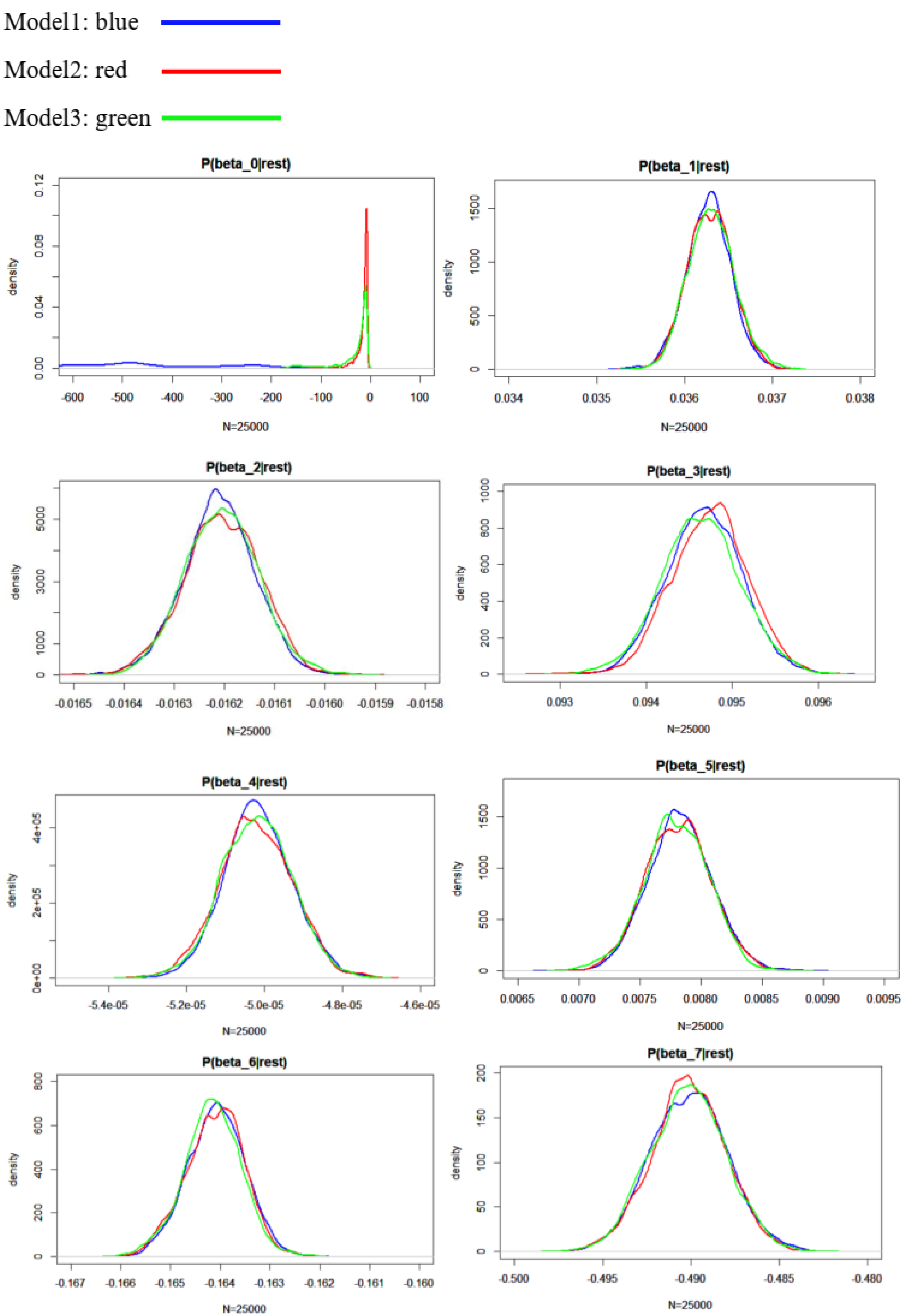


Figure 6: Running mean plot for α in model3

Compare density plot



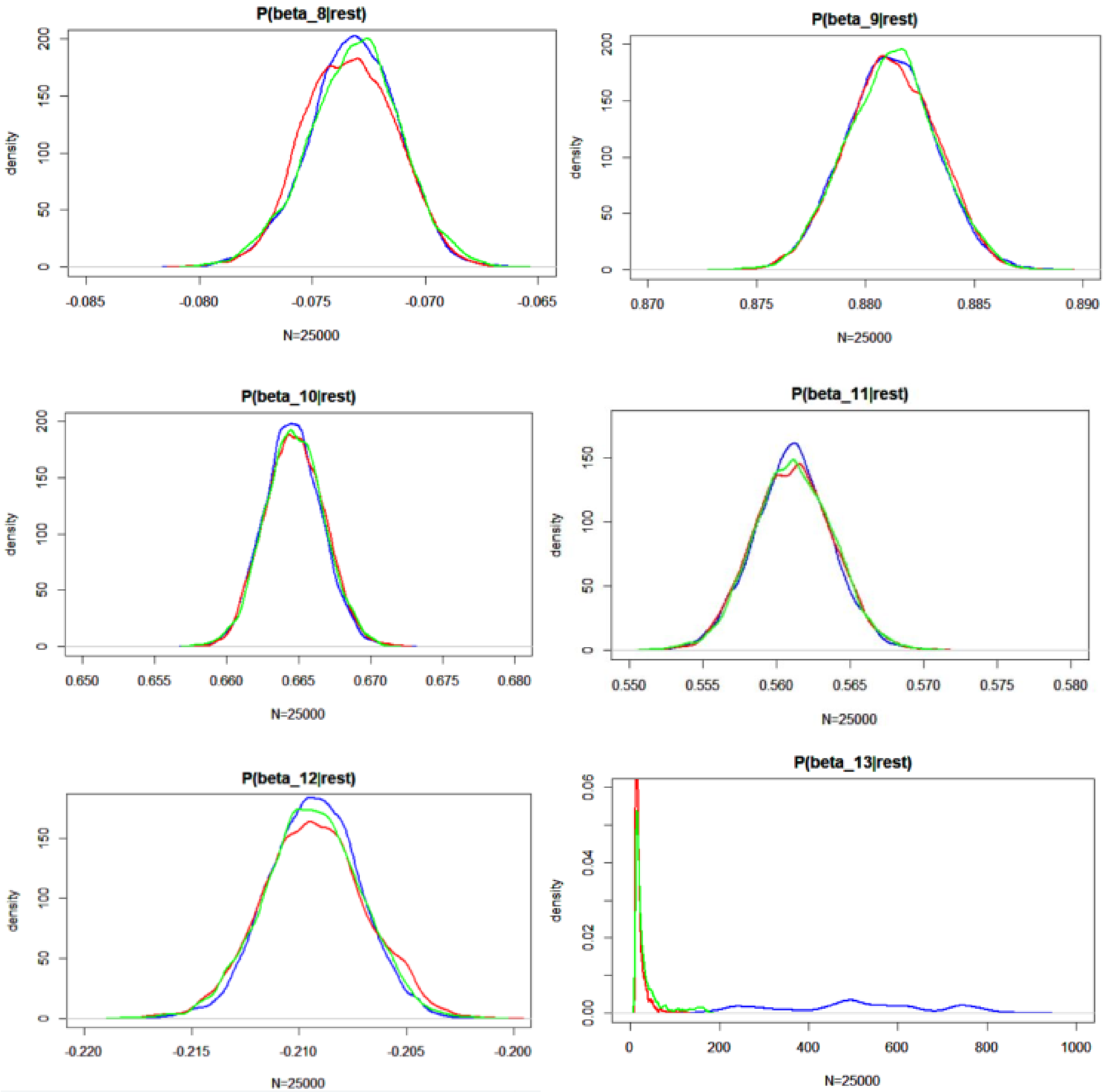


Figure 7: Compare density plot for β

ACF plot

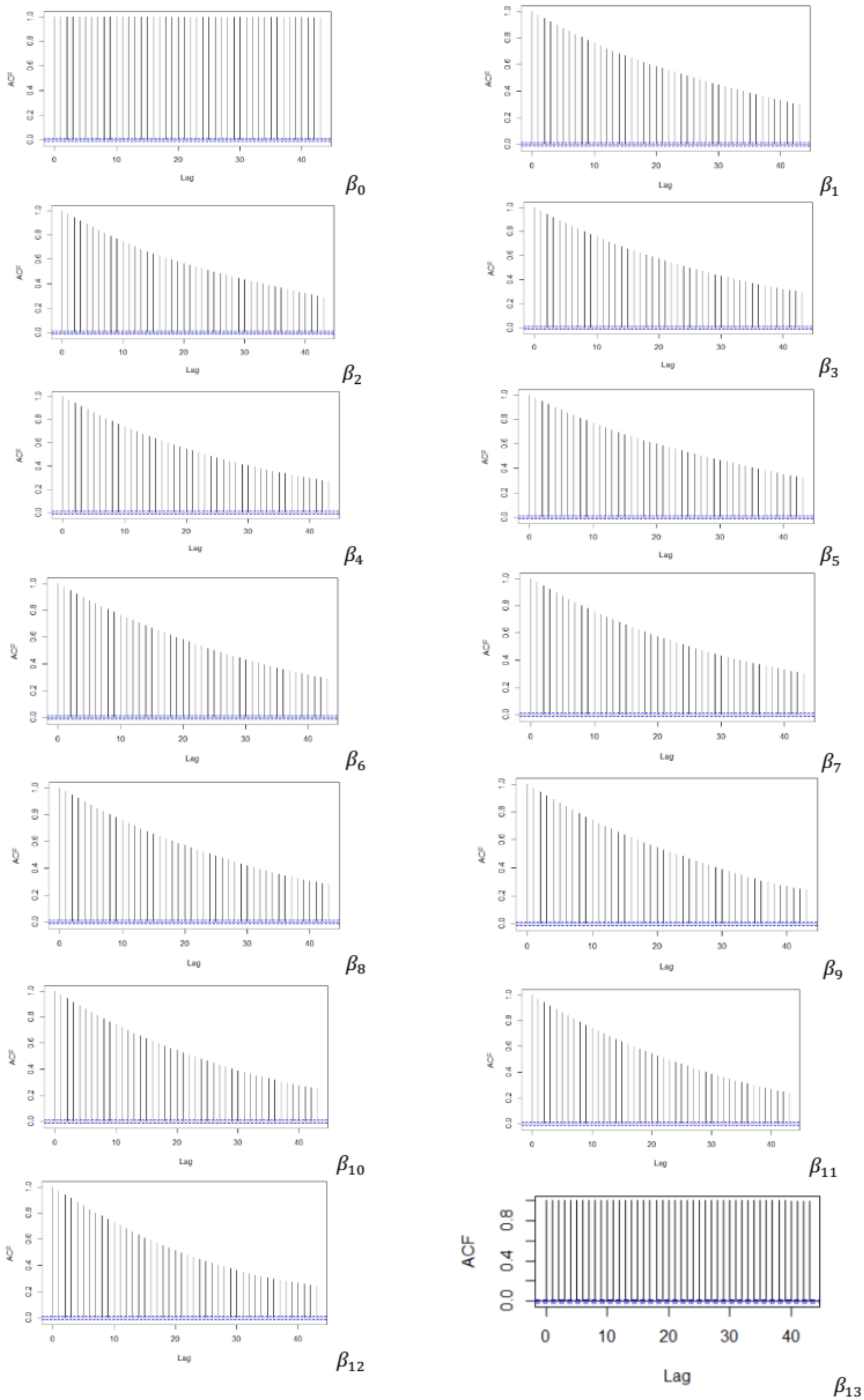


Figure 8: ACF plot for β in model1

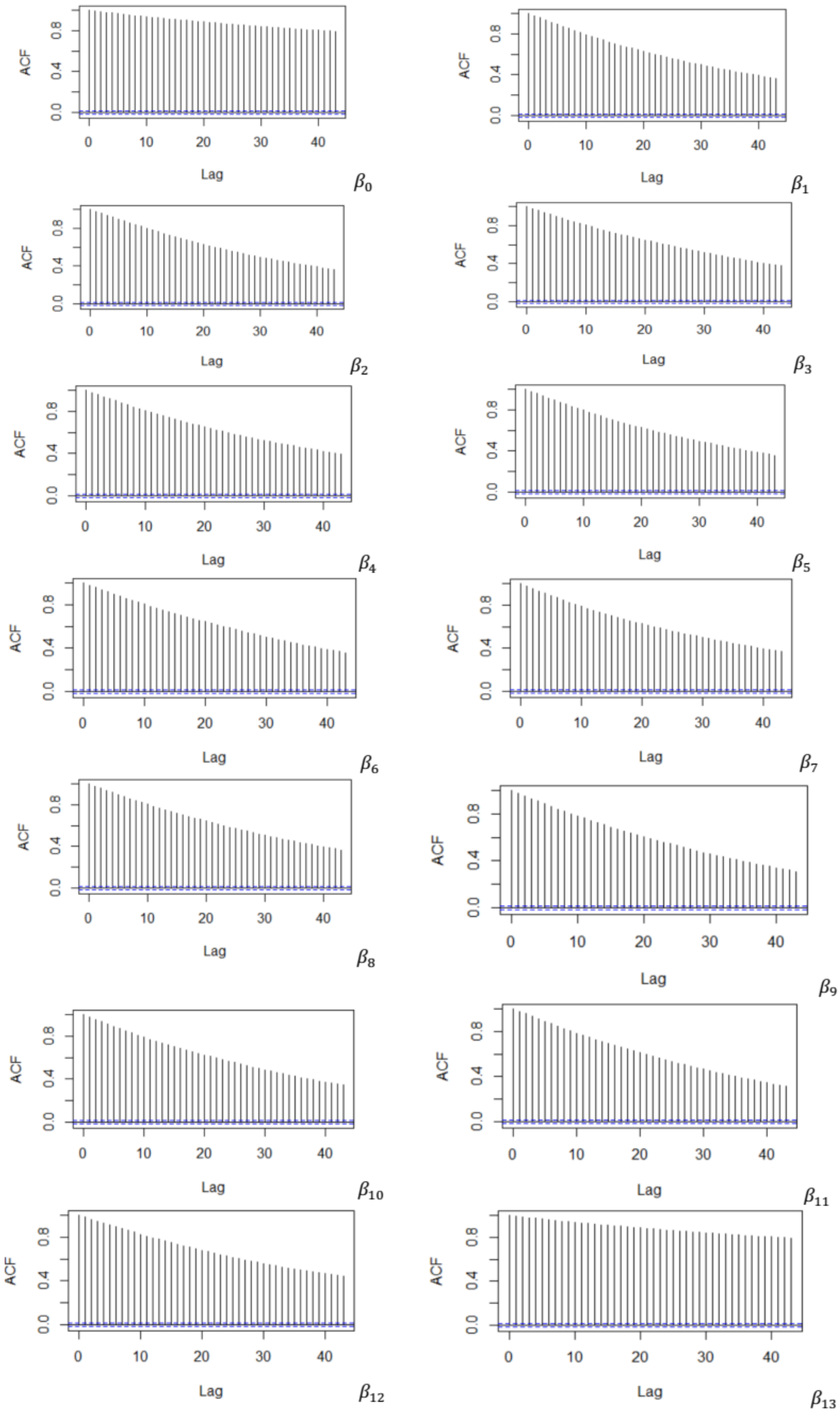


Figure 9: ACF plot for β in model2

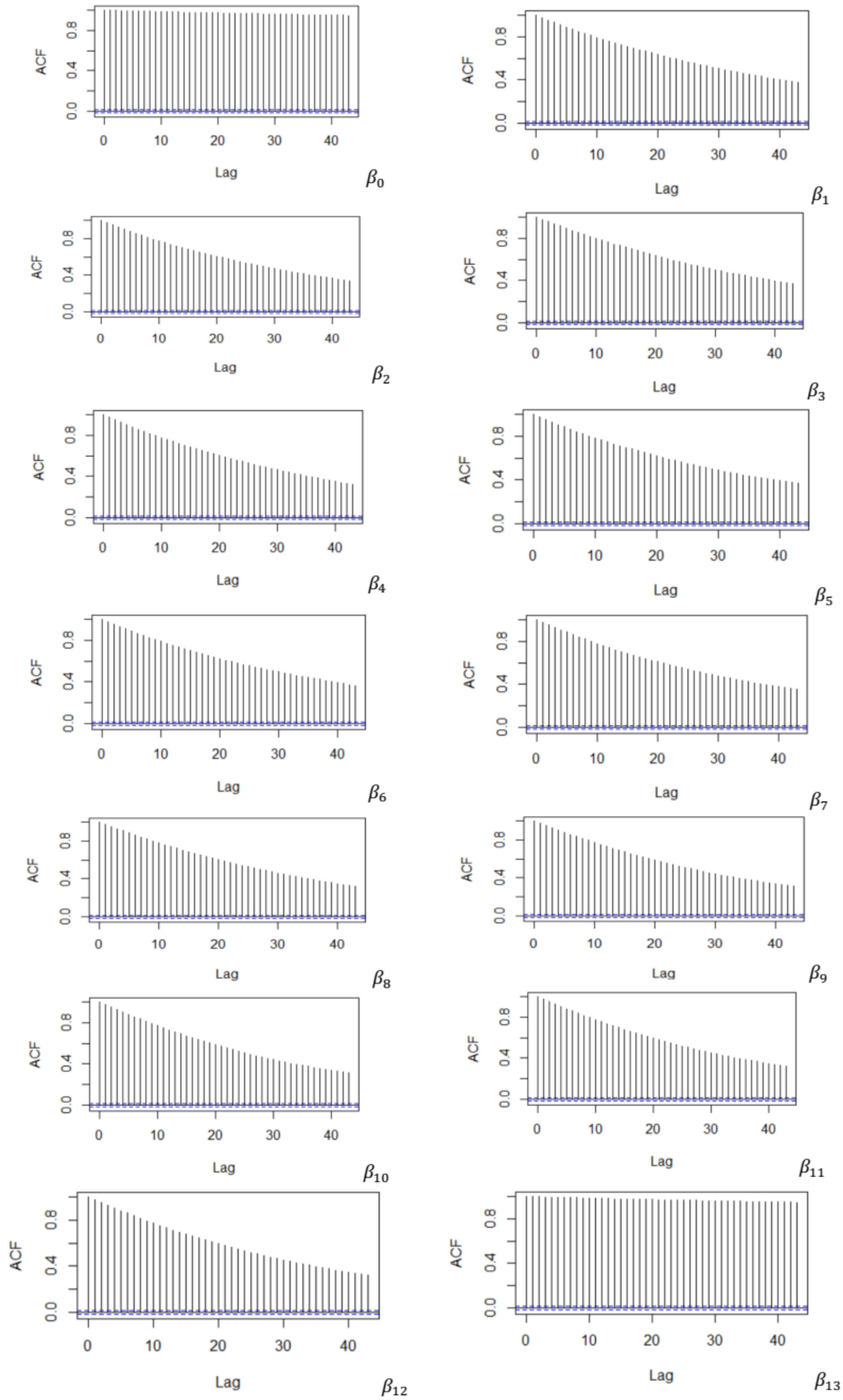


Figure 10: ACF plot for β in model3

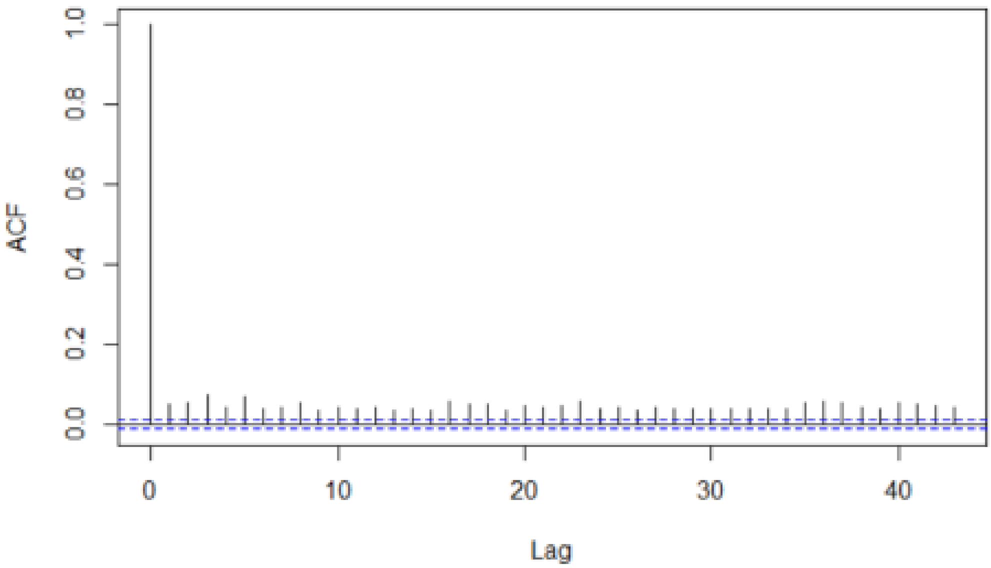


Figure: ACF plot for sigma2 in model2

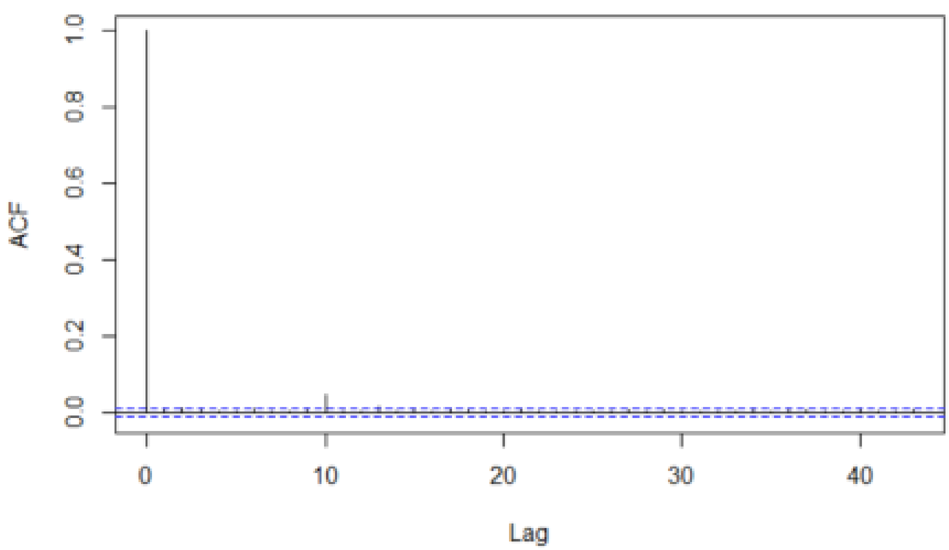


Figure: ACF plot for sigma2 in model3

Figure 11: ACF plot for σ^2

4.0.1 Gelman-Rubin diagnostic

	Model1	Model2	Model3
β_0: (Intercept)	3.04	1.04	1.02
β_1: Temperature	1	1	1
β_2: Humidity	1	1	1
β_3: Wind_speed	1.02	1.01	1
β_4: Visibility	1	1.01	1
β_5: Dew_point_temperature	1	1	1
β_6: Solar_Radiation	1	1	1.01
β_7: Rainfall	1	1	1.01
β_8: Snowfall	1	1	1
β_9: Autumn	1.02	1	1
β_{10}: Spring	1.02	1.01	1.01
β_{11}: Summer	1.01	1.01	1
β_{12}: holiday	1.01	1.02	1
β_{13}: functioning day	3.04	1.04	1.02
σ^2	-	1.25	1.24
α	-	-	1.27