Bayesian Navigation

Introduction:

Position resection method is a method for finding unknown position by measuring angles with respect to known positions (reference points). The intersection of the drawn lines from those reference points reveals the unknown position. In this project we will estimate our location on a map using the position resection method. The uncertainty of the positioning algorithm is studied, and the distribution of plausible locations is obtained.

Exploratory Analysis:

As mentioned, resection method depends on the intersection between the lines drawn from the reference points. Three reference points are used, and the bearing angles are measured. Figure 1 shows three reference points with the intersection of the three lines indicates our position. Due to the error in measuring the angles, the lines from reference points don't intersect in single point. Instead, the intersection results in a triangle as shown in Figure 1.

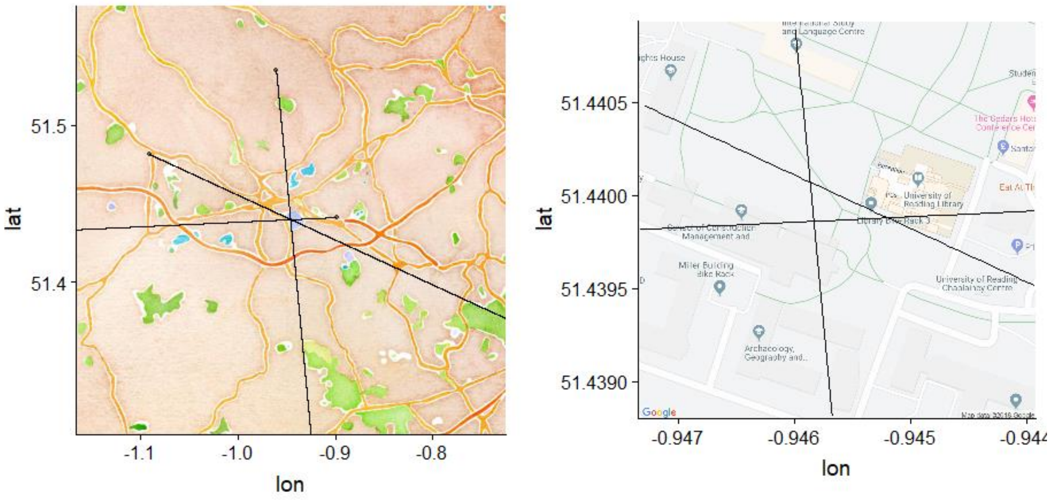

Figure 1:Position resection and zooming in our location

The error in angle measurements assumed to be normal with variance sigma. Our location longitude (lambda) and latitude (phi) are random variables. In the first part of the project, a wide uniform prior is assumed for our location. Also, the error variance has prior exponentially distributed (see Figure 2). In the second part the prior is the joint distribution of lambda and phi which is normal and depend on rho, the shortest distance to the middle of the nearest road (see Figure 3).

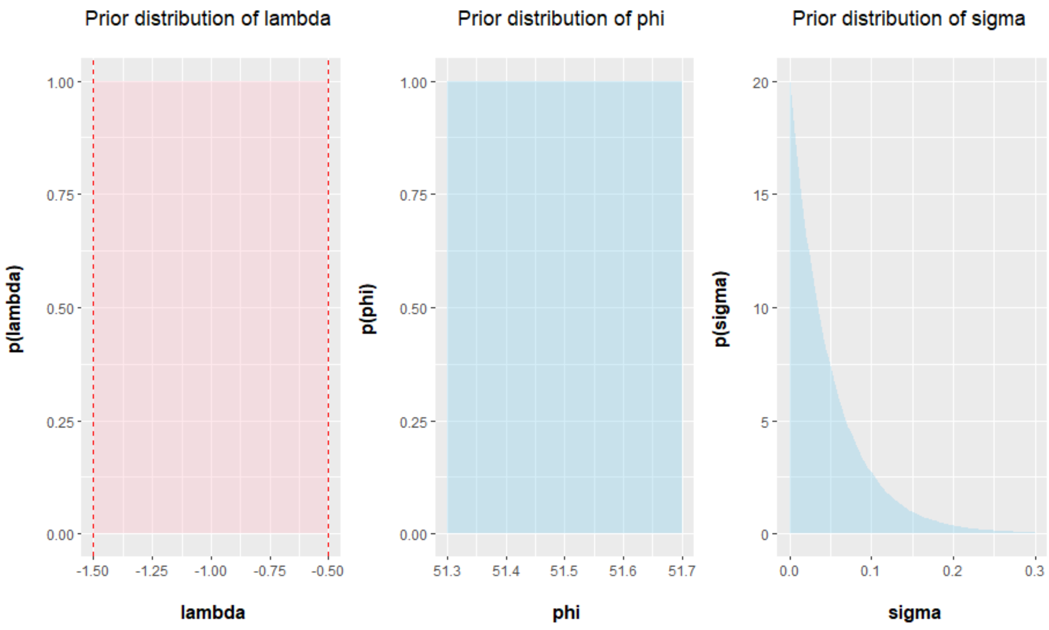

Figure 2: Part1 Prior Distribution

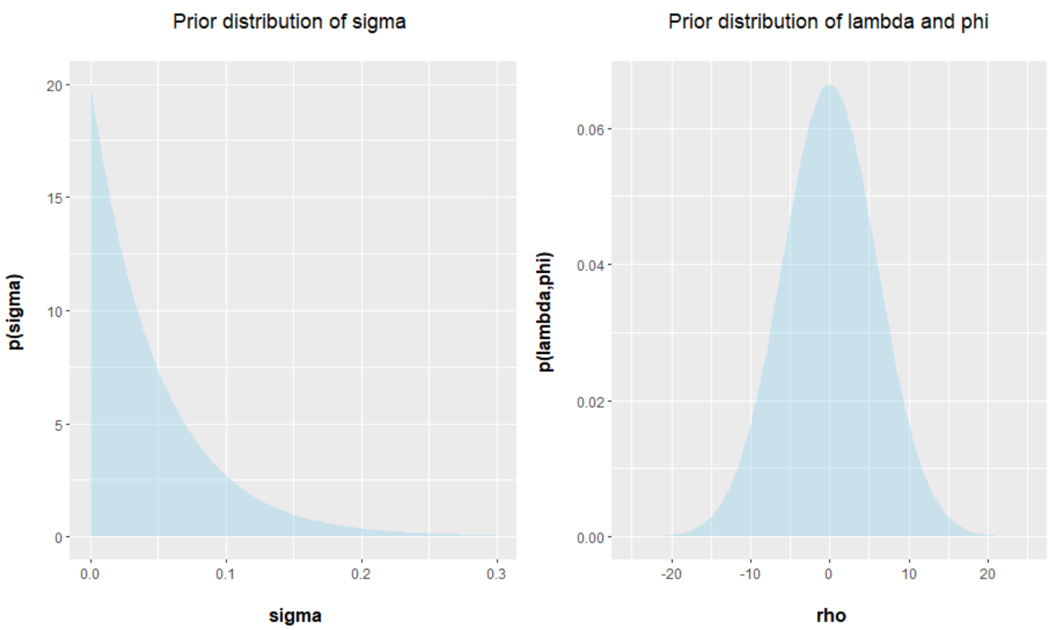

Figure 3:Part2 Prior Distribution

Description and Method used:

Three reference points (p1, p2, p3) on an online map around our location are collected and their coordinates are obtained. The bearing angles to our location (alpha, beta and gamma) are calculated. As there are many factors affect the quality of angles measurements, a randomly distributed error with mean zero and standard deviation sigma is added to the measurements to create a simulated data set. The closer our location to a reference point, the smaller the error corresponding to the measured angle to that reference point. The likelihood is formulated assuming that angles are measured independently. As mentioned before, two prior options are considered.

The posterior of lambda, phi and sigma is evaluated by sampling from it using Metropolis-Hastings. Three independent Markov chains are simulated. Multivariate normal distribution is used as a proposal distribution for lambda, phi and sigma. For sigma, the positive part is used .4000 iterations of the Metropolis-Hastings algorithm are performed for three chains independently. The algorithm is run in two phases: first phase, parameters are tuned to increase simulation efficiency, and second phase, the algorithm is run till approximate convergence. To find reasonable proposal, we used arbitrary proposal, collect some samples and use the samples covariance as the variance-covariance matrix for the multivariate normal proposal. Several tests are done both visual and statistical, to see if the chain appears to be converged. Gelman-Rubin convergence diagnostics are used. It measures whether there is a significant difference between the variance within several chains and the variance between several chains by a value that is called potential scale reduction factors (PSRF). The PSRF for the three parameters are calculated and chains are run long enough so that all PSRF are small (almost 1). Once convergence is achieved, the final samples are collected from the posterior.

Results and Discussion:

Part 1 Results and Convergence Diagnostics:

As mentioned before, in this part a wide uniform prior is assumed for our location. Also, the error variance has prior exponentially distributed. The results below show the posterior means, posterior standard deviations, and posterior quantiles for lambda, phi and sigma respectively. Also, it shows the naive standard error and the time series standard error. The naive standard error captures the standard error of the mean as a result of the simulation error, while the time series standard error adjusts the naive standard error for autocorrelation.

```
Iterations = 1:30000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 30000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

         Mean        SD  Naive SE Time-series SE
[1,] -0.94582 1.492e-04 8.617e-07      3.874e-06
[2,] 51.43988 7.349e-05 4.243e-07      1.646e-06
[3,]  0.07998 4.385e-02 2.531e-04      1.561e-03

2. Quantiles for each variable:

        2.5%      25%      50%      75%   97.5%
var1 -0.9461 -0.94589 -0.94582 -0.94575 -0.9455
var2 51.4397 51.43984 51.43988 51.43991 51.4400
var3  0.0290  0.04982  0.06804  0.09788  0.1947
```

One way to assess convergence is a traceplot for each variable as shown below (Figure 4). From the plots, chains are mixing well and not stuck in certain areas. Another way is to check the autocorrelations between the draws of Markov chain. It can be noticed that from the plots the kth lag autocorrelation is decreased as k increases and that gives an indicator of good mixing (see Figure 5)

The acceptance rate for the Metropolis-Hastings algorithm for each parameter is 28.7% which is acceptable as the ideal theoretical acceptance rate for multiparameter is between 20-40%. Also, the PSRF changes through the iterations till it reaches almost 1 so there is no significant difference between the variance within the three chains and the variance between them (see Figure 6). Finally, the visualisation of the posterior on the map is shown in Figure7. Our location is inside the triangle and closer to one of the corners. This is due to different distances from our location to the three reference points. In this case the posterior distribution is a mix between prior and data that’s why our estimated location is affected by how close the reference point is.

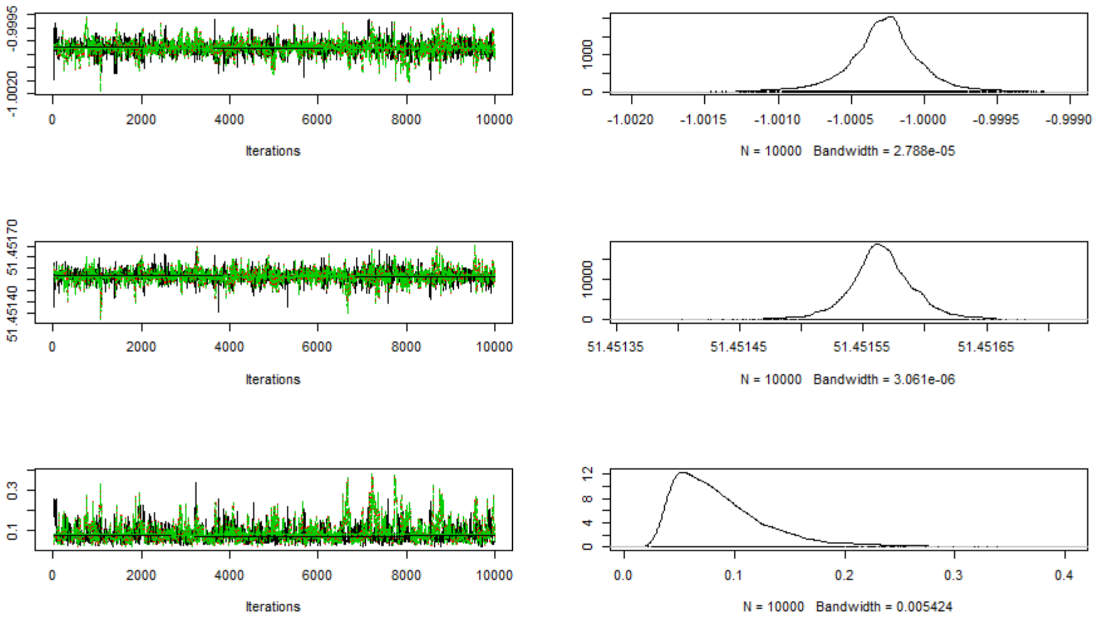

Figure 4:Traceplot and density plot for Lambda, Phi and Sigma

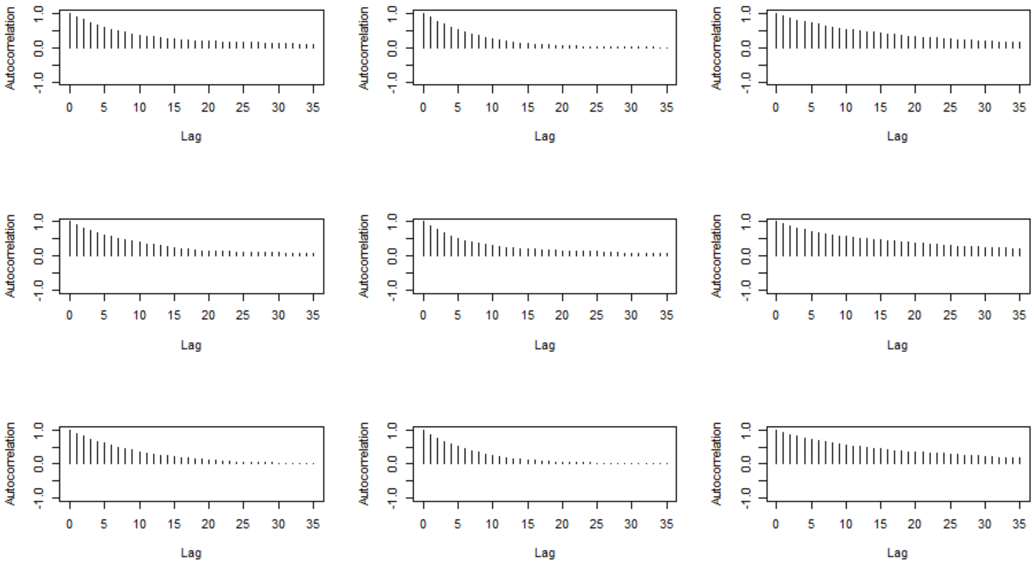

Figure 5:Autocorrelation plots for the three chains for Each Variable (λ, φ and σ)

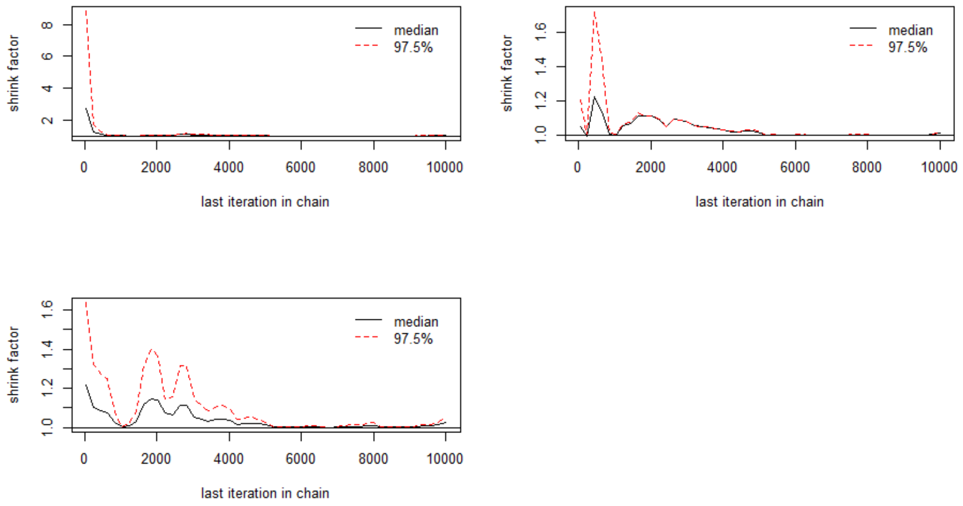

Figure 6: Potential Scale Reduction factor(PSRF) changes through the iterations

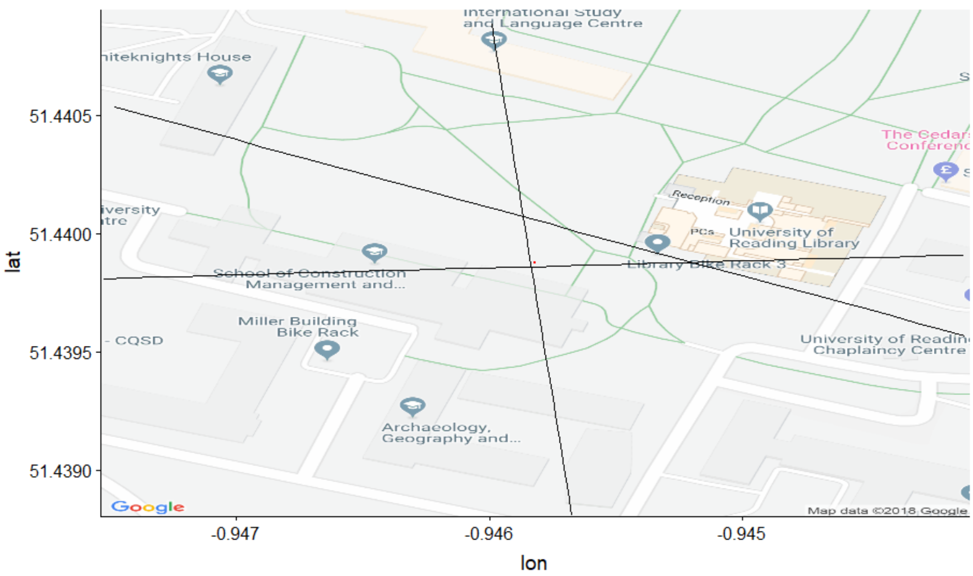

Figure 7:The posterior of the position assuming a wide uniform prior.

Part 2 Results and Convergence Diagnostics:

As mentioned before, in this part the prior is the joint distribution of lambda and phi which is normal and depend on rho, the shortest distance to the middle of the nearest road. Also, the error variance has prior exponentially distributed. The results below show the posterior means, posterior standard deviations, and posterior quantiles for lambda, phi and sigma respectively.

```
Iterations = 1:30000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 30000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:
```

```
          Mean        SD  Naive SE Time-series SE
[1,] -0.94572 1.783e-04 1.029e-06      6.743e-06
[2,] 51.43983 8.156e-05 4.709e-07      3.697e-06
[3,]  0.09779 3.517e-02 2.031e-04      2.376e-03

2. Quantiles for each variable:

         2.5%      25%      50%     75%   97.5%
var1 -0.94611 -0.94584 -0.94566 -0.9456 -0.9454
var2 51.43969 51.43978 51.43982 51.4399 51.4400
var3  0.04763  0.07136  0.09136  0.1182  0.1818
```

One way to assess convergence is a traceplot for each variable as shown below (see Figure 9). From the plots, chains are mixing well but at some points they stuck in certain areas of the parameter space. Mixing is influenced by the proposal function which may lead to one of two cases:

- High acceptance rate as the proposal function is narrower than the posterior distribution.
- Low acceptance rate as the proposal function is too wide in comparison to the posterior distribution.

The acceptance rate for the Metropolis-Hastings algorithm for each parameter is 21.3%. Although it's within the ideal theoretical acceptance rate for multiparameter 20-40%, but still there is bad mixing. This problem may be due to correlation in parameter space. It can be solved by discarding the burn-in part and reduce short term autocorrelation by taking every kth draw from the chain and throw away most of the samples. Also, autocorrelation can be reduced by increasing the variance of the proposal distribution which will increase the rejection of the proposal distribution and lead to bad mixing. Because of that, a good posterior estimation will need a very large sample.

Another way to assess convergence is to check the autocorrelations between the draws of our Markov chain. It can be noticed from the plots of sigma for different chains, the kth lag autocorrelation is slowly decreased as k increases as an indicator of poor mixing (see Figure 9).

By taking every 2 draws from chains we decrease the autocorrelation as shown in Figure 11and Figure 12 ,but we still need to increase the sample size to have a good posterior estimation.

Also, the PSRF changes through the iterations till it reaches almost 1 so there is no significant difference between the variance within the three chains and the variance between them (see Figure 10). Finally, the visualisation of the posterior on the map is shown below (see Figure 13). Our location is outside the triangle and closer to the middle of the nearest road. In this case the posterior is mainly affected by the prior and the data has less effect on it that's why the estimated location is outside the triangle and near the middle of the road.

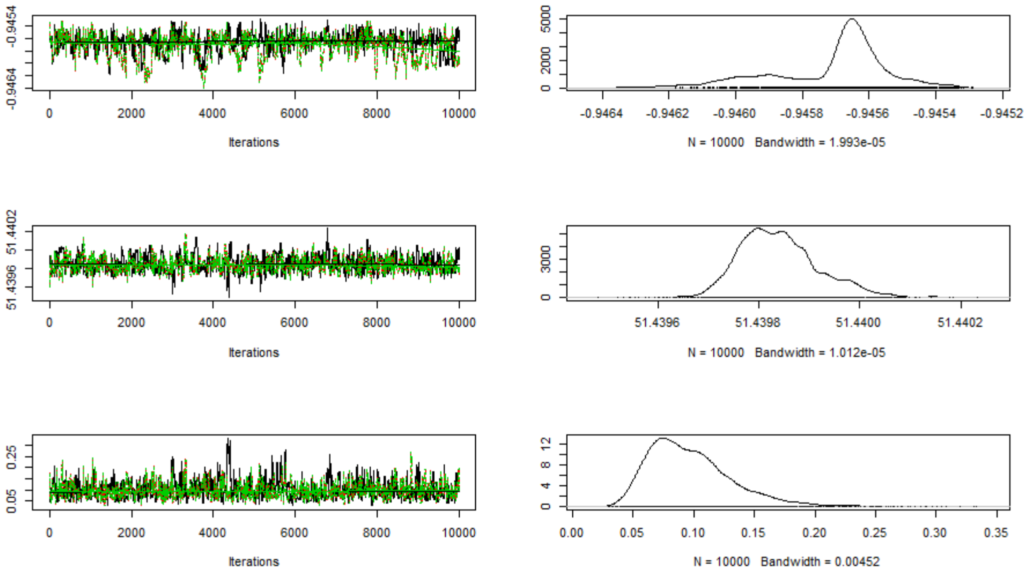

Figure 8:Traceplot and density plot for Lambda, Phi and Sigma

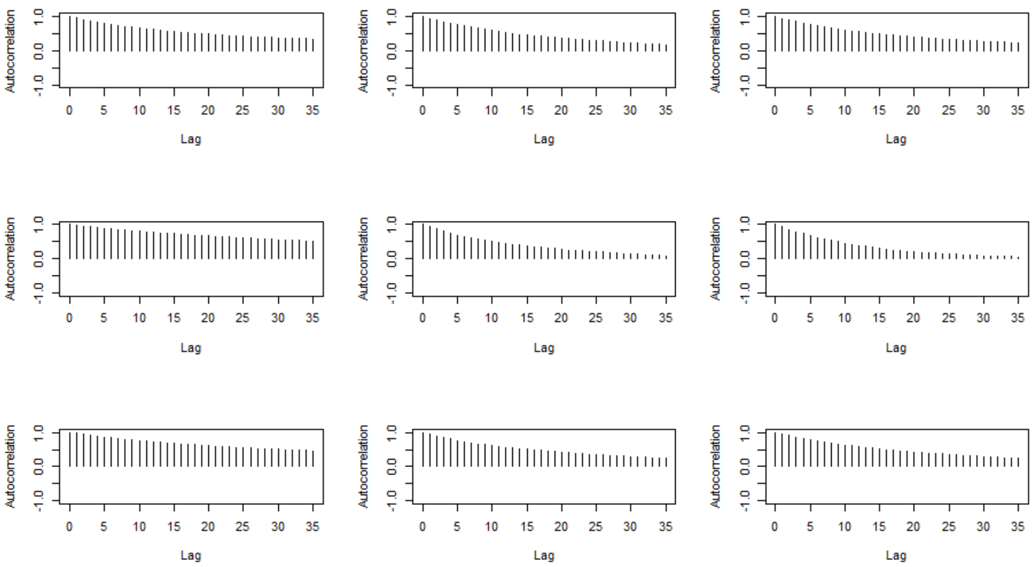

Figure 9:Autocorrelation plots for the three chains for Each Variable (λ, φ and σ)

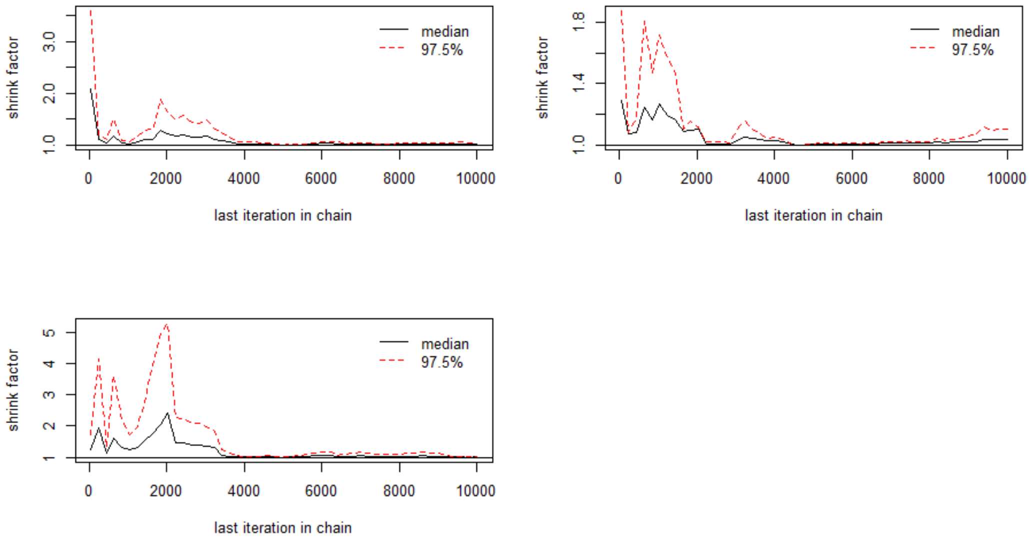

Figure 10: Potential Scale Reduction factor(PSRF) changes through the iterations

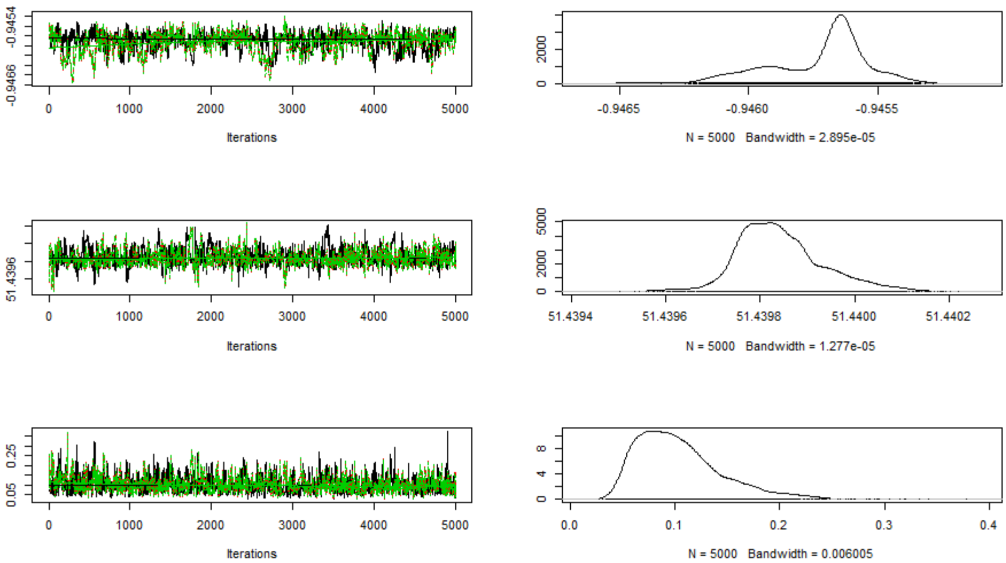

Figure 11: Traceplot and density plot for Lambda, Phi and Sigma after Thinning

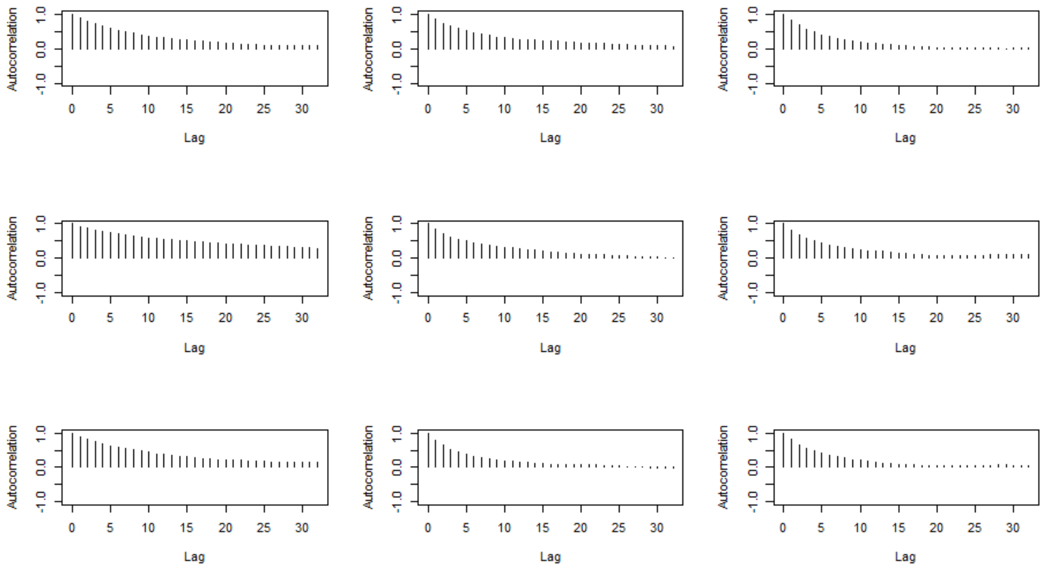

Figure 12: Autocorrelation plots for the three chains for Each Variable (λ, φ and σ) after Thinning

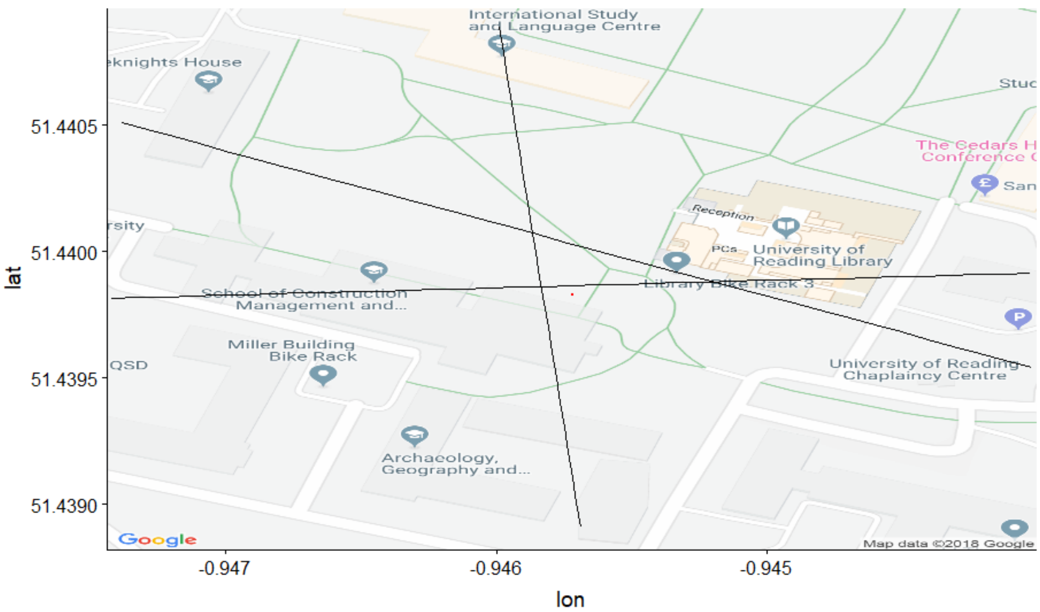

Figure 13: The posterior of the position assuming proximity to a road.

Conclusion:

The proposal distribution has an important role in convergence rate of Markov chains. So, convergence rate can be longer or shorter depends on the proposal distribution. Also, chain reaches convergence doesn't mean that an optimal sample has been produced from the posterior distribution. High autocorrelation between consecutive samples mean the given posterior distribution information is small and we need higher sample size to cover the parameter space. Model identification problems affect the convergence in MCMC. The choice of priors one of the reasons that affect model identification and convergence in MCMC. Our estimated location (posterior) is mainly affected by the choice of the prior. Metropolis-Hastings is very slow for multiparameter models, so an adaptive Metropolis-Hastings would perform better.