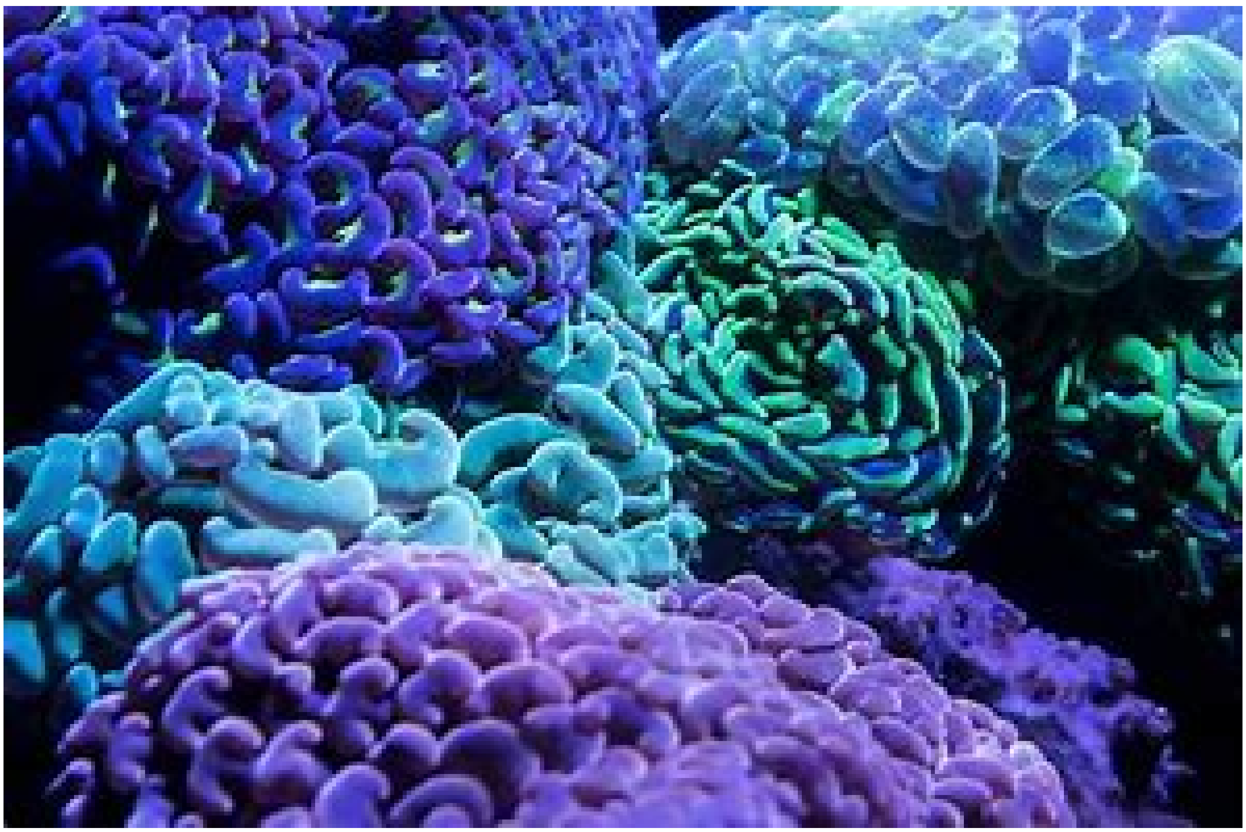
Interim Report

Predicting Algae Bloom spread in the Marlborough Sounds using Bayesian Inference

Luke Pearson



Abstract

Toxic Algae blooms cause shellfish in the affected area to be infected, causing them to carry the toxins that can cause severe illness to humans when they are consumed. When these blooms occur, it causes the shellfish industry in the Marlborough Sounds to shut down collection region-wide out of an abundance of caution, which increases costs. Therefore, they have an interest in predicting Algae blooms, and in particular the size and the spread in an effort to narrow the extent of the shutdown.

For the purposes of this project, the Marlborough sounds have been divided into a set of 386 polygonal approximations of areas of the sounds. The volumes of these polygons have been calculated, and a transport matrix depicting the proportion of particles that move from site i to site j in a particular time interval. (Ross Vennell, 2022)

Our data is recording concentrations of samples retrieved from these sites at discrete time intervals, observed with an error. This data is also observed sparsely.

Due to timing, the actual data is not currently available, so the scope of this project will involve simulating data with randomly generated parameters, and then trying to predict the data by deducing the values of the parameters.

This will be done by constructing a forward in time statistical model of the mass and then using bayesian inference, via a Metropolis-Hastings algorithm to determine the model parameters.

This model will be based upon an exponential growth component during active blooms, a decay component and a transfer component based upon the transport matrix.

In general, the results showed that this method of approximating the model parameters was a successful one, and particularly that the model was able to work with a sparse amount of data, a changing number of events and removing the source sites.

Background

Harmful Algal Blooms, or HABs, occur when there is a large buildup of algae that release materials which are toxic to other species. Particularly, *cyanobacteria* variety algae blooms, when ingested by shellfish become toxic to humans and can cause serious illness if consumed[1].

As a result of this danger, when HABs occur in the Marlborough Sounds, the response is to temporarily halt harvesting in the region until the bloom ends. Considering the Marlborough Sounds is responsible for 80%[2] of domestic shellfish production, this causes an immense cost to the industry; approximately $127 million annually[2].

There are a number of conditions that could cause a HAB to begin, with the main factors being a warmer temperature[3] and increase in nutrients such as phosphorus and nitrogen[4] in the water, however it is hard to predict when exactly one will occur. Climate Change has also had an adverse impact on these conditions causing HABs with more frequency[5].

While the length of the HAB is also variable and hard to predict, there are conditions that will determine the end; primarily, a sudden decrease in temperature that will cause the Algae to go inactive and sink below the surface, stopping the bloom. For the purposes of the Marlborough sounds region, these events happen globally.

The aim of this project is to simulate data of algal blooms over a set time period, and then use Bayesian Inference via a Metropolis-Hastings Algorithm to approximate the parameters of our model. We will also explore other scenarios such as varying amounts of sparseness in the data, and predicting a different number of events than the true amount.

Methods

Bayesian Inference

Bayesian inference is a branch of statistical modelling based upon Bayes' theorem

$$P(\theta|Y) = \frac{P(Y|\theta) \times P(\theta)}{P(Y)}$$

Where:

- $P(\theta)$ is the $prior$ distribution of the parameters
- $P(Y|\theta)$ is the $likelihood$ distribution of the data given the parameters
- $P(\theta|Y)$ is the $posterior$ distribution of the parameters given the data
- $P(Y)$ is the *marginal likelihood* distribution of the data

Where instead of the usual approach where the parameters X are fixed and the data Y is variable to determine the likelihood, the data is used as the fixed variable to determine the distribution of the unknown parameters. Since the data is fixed, $P(Y)$ will be known.

For the purposes of this project, we can use the likelihood for the Concentrations and an educated construction of prior distributions to determine the distribution of the parameters, or $P(Parameters|Concentration)\ \alpha\ P(Concentration|Parameters) \times P(Parameters)$.

However, due to the complex nature of the likelihood in this case $\left(Where\ C_{it} \sim N\left(log\left(\delta + \frac{M_{it}}{V_i}\right), \frac{1}{\sqrt{\tau}}^2\right)\right)$, deriving the posterior distribution mathematically provides a result that is not useful. Therefore, a computational method of approximating these distributions is needed.

Metropolis-Hastings

The Metropolis-Hastings algorithm is a Monte Carlo Markov Chain sampling method that can be used to generate samples from a distribution that is otherwise hard to sample from. This method utilises the Bayesian inference ideas above, using a user defined a $proposal$ distribution to control the chain.

General Method

At iteration s, For each parameter θ, the algorithm generates a candidate value θ^* from a proposal distribution $q\left(\theta^*|\theta_{s-1}\right)$ that is dependent on the previous value for θ. It then calculates a ratio of

$$r = \frac{P\left(Y|\theta^*\right) \times P\left(\theta^*\right) \times q\left(\theta_{s-1}|\theta^*\right)}{P\left(Y|\theta_{s-1}\right) \times P\left(\theta_{s-1}\right) \times q\left(\theta^*|\theta_{s-1}\right)}$$ with:

- $P(Y|\theta)$ being calculated from the $likelihood$ function
- $P(Y)$ being calculated from the $prior$ function
- $P\left(\theta_a|\theta_b\right)$ being calculated from the $likelihood$ function

And then the probability of accepting θ^* as θ_s is $min(1, r)$, with the probability of accepting θ_{s-1} as θ_s is $max(0,\ 1 - r)$

Application

Because our likelihood is different for each site i and time t, and that we have the assumption of independence amongst Y_{it}, the total likelihood $P(Y|\theta)$ is actually $\prod_{i=1}^{n}\left(\prod_{t=1}^{T}\left(P\left(Y_{it}|\theta\right)\right)\right)$. Since our simulation will involve 10 sites and 100 times, this will be the product of 10000 probabilities; to avoid our code rounding the value to zero, the log likelihood $\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\left(log\left(P\left(Y_{it}|\theta\right)\right)\right)\right)$ will be more computationally reliable. Therefore, our total log-ratio is

$$logratio = log\left(P\left(Y|\theta^*\right)\right) + log\left(P\left(\theta^*\right)\right) + log\left(q\left(\theta_{s-1}|\theta^*\right)\right) - \left(log\left(P\left(Y|\theta_{s-1}\right)\right) + log\left(P\left(\theta_{s-1}\right)\right) + log\left(q\left(\theta^*|\theta_{s-1}\right)\right)\right)$$

And then the probability of accepting θ^* as θ_s is $min\left(1, e^{logratio}\right)$, with the probability of accepting θ_{s-1} as θ_s is $max\left(0, 1 - e^{logratio}\right)$.

Transportation Matrix

In preparation for this project, Ross Vennell and Richard Arnold have constructed a transportation matrix, P, to represent the proportion of mass that moves between sites across the Marlborough Sounds for each time step.

This matrix was constructed using OceanTracker, an application designed to efficiently simulate hydrodynamic models via Lagrangian Particle Tracking[6].

He then constructed polygons approximating around the 386 different source sites (Labelled in $Figure\ 1$), and simulated 10000 particles being released in that polygon for a week. The fraction of particles in each polygon after this time period constitute the values for P.

P is an $n \times n$ matrix where P_{ij} is equal to the proportion of particles in site i that travels to site j in a single time step, with two main properties:

- $0 \le P_{ij} \le 1 \; for\ all\ i, j$
- $\sum_{j=1}^{n} P_{ij} = 1 \; for\ all\ i$

$Figure\ 2$ is a heatmap of P. P is a very sparse matrix, with 97% of values being 0, and P is also close to a diagonal matrix with the average value of P_{ii} being 0.935. As a result, we shouldn't expect algae to spread amongst a lot of sites such that a bloom should stay relatively contained to the sites in the near vicinity.

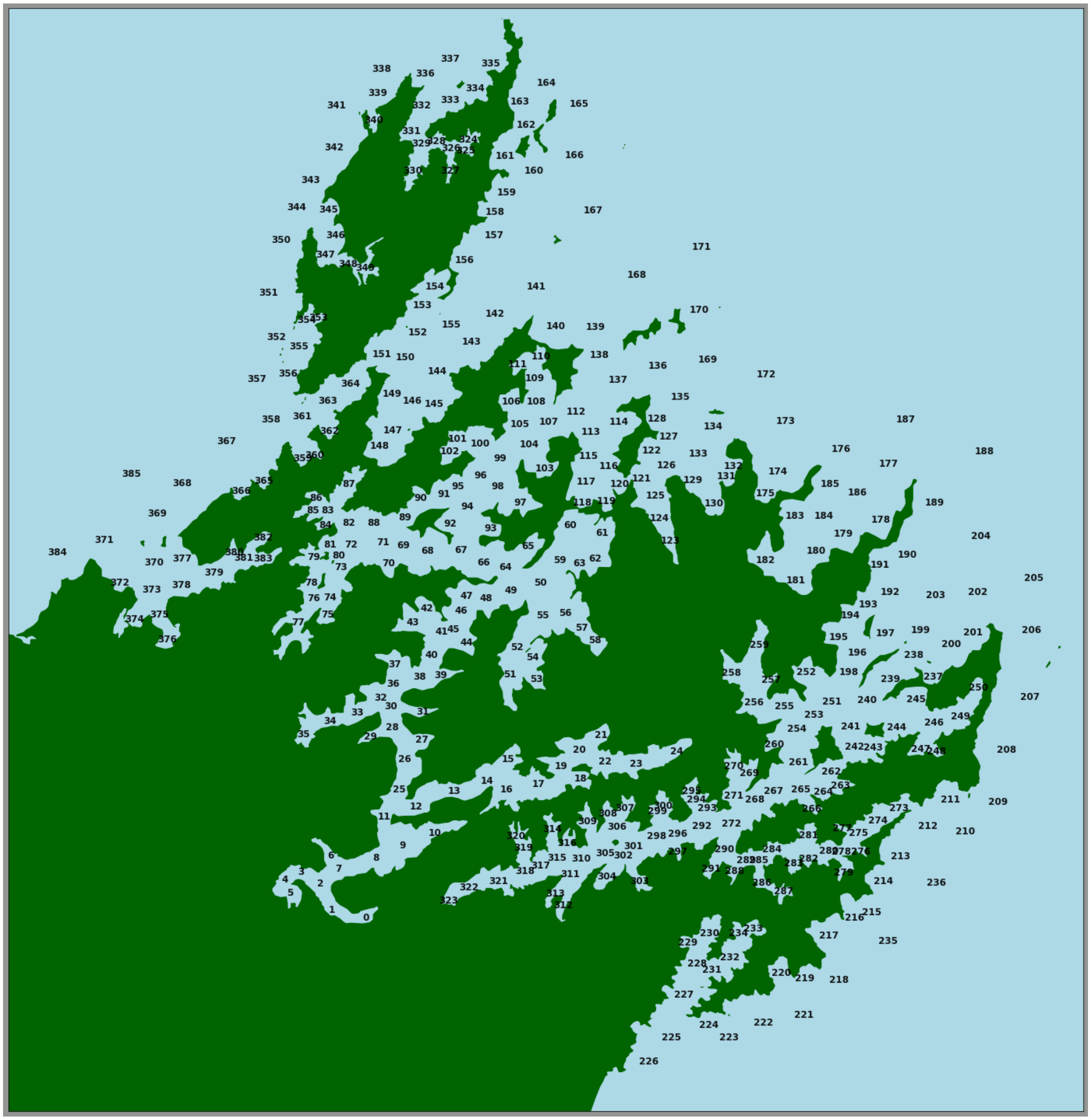
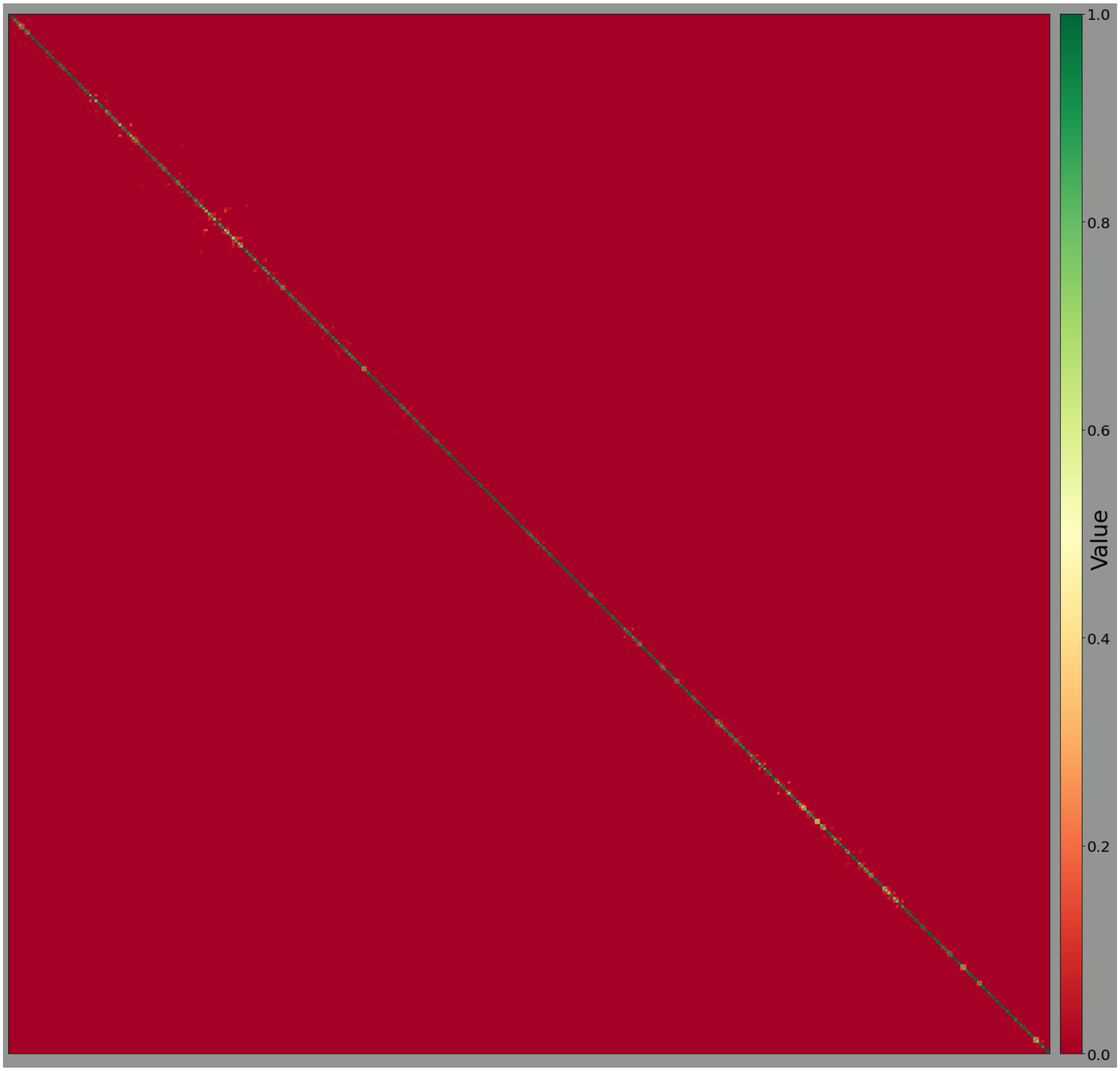


Figure 1: *Map of Marlborough Sounds with locations of sites labelled*



$Figure\ 2\text{:}\ Heatmap\ of\ P$

Metrics

We will be using two metrics to evaluate the comparative performances of our estimation: DIC and the Likelihood. We will be using them over both the full set of data, as well as the observed subset. For each parameter, we will

DIC

DIC, or the Deviance information Criteria, is an extension of Akaike information criterion (AIC) that is particularly useful for performing model selection when they've been approximated via Markov Chain Monte Carlo algorithms like the Metropolis-Hastings algorithm we're using here. This is calculated by using the deviance for a set of parameters $D(\theta)$ as;

$D(\theta) = -2log(P(Y|\theta)) + C$, where C is a constant that will cancel out when comparing models.

Since our likelihood function $log(P(Y|\theta))$ is dependent on i, t and we are assuming independence amongst $Y_{i,t}$ for all i, t, the deviance can be calculated as

$$D(\theta) = -2\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\left(log\left(P\left(Y_{it}|\theta\right)\right)\right)\right)$$

Since $D(\theta)$ fits more easily to models with large sets of parameters, there is a penalty p_D calculated as

$p_D = \overline{D(\theta)} - D\left(\hat{\theta}\right)$, where $\overline{D(\theta)}$ is the posterior mean value of all Deviances and $D\left(\hat{\theta}\right)$ is the Deviance of the posterior mean parameter values. So the total DIC is

$$DIC = \overline{D(\theta)} + p_D$$

Our Model

Mass

Our equation for Mass at site i and time t is:

$$M_{i,t} = \alpha A_{i,t} M_{i,t-1} + \beta\left(\sum_{j=1}^{n} p_{i,j} M_{j,t-1}\right) + B_{i,t}, \text{ where}$$

$$A_{i,t} = Indicator\left(\sum_{k=1}^{N} Indicator\left(I_k = i\right) Indicator\left(T_k \le t \le T_k + W_k\right) > 0\right)$$

and

$$B_{i,t} = \sum_{k=1}^{N} I\left(I_k = i\right) I\left(T_k = t\right) S_k$$

This model is comprised of three main components; start, growth, and transport, where blooms are represented as a fixed number of events N, spanning n sites and over the course of T time periods. For our subscripts, i refers to the site and t refers to the time period.

Start component

The 'Start' component is meant to simulate the starting size of a bloom when it begins at a particular site and time.

If event k starts at site I_k at time T_k, then this component $B_{i,t}$ will be equal to size S_k when $I_k = i$ and $T_k = t$, and all other values will be equal to 0.

In terms of mathematical notation, it is represented as:

$$B_{i,t} = \sum_{k=1}^{n} I\left(I_k = i\right) I\left(T_k = t\right) S_k.$$

Growth Component

The 'Growth' component is meant to simulate how a bloom grows at a particular site over the course of the bloom's length during a single time period.

If event k starts at site I_k and at time T_k with a length of W_k, then this component $A_{i,t}$ will be equal to 1 when $I_k = i$ and $T_k \leq t \leq T_k + W_k$, and all other values will be equal to 0.

In terms of mathematical notation, it is represented as:

$$A_{i,t} = I\left(\sum_{k=1}^{n} I\left(I_k = i\right) I\left(T_k \leq t \leq T_k + W_k\right) > 0\right)$$

If $A_{i,t} = 1$, then the mass at site i will be growing exponentially at a rate of α where α is our growth rate.

Transport Component

The 'transport' component is meant to simulate the mass that transfers from one site to another during a single time period.

For each site i, they will receive $p_{i,j} M_{j,t-1}$ from each site, where $p_{i,j}$ is the proportion of mass that transports from site i to site j, and $M_{j,t-1}$ is the mass at site j at time $t - 1$. All of this mass will be scaled by β, the rate of decay.

In terms of mathematical notation, it is represented as

$$\beta\left(\sum_{j=1}^{n} p_{i,j} M_{j,t-1}\right)$$

Concentration

The observed values for $Y_{i,t}$ at site i and time t is:

$$log(Y_{i,t}) \sim N\left(log\left(\delta + \frac{M_{i,t}}{V_i}\right), \frac{1}{\sqrt{\tau}}^2\right)$$

Concentration $Y_{i,t}$ for a site i at time t is modelled on a logarithmic scale of $\frac{M_{i,t}}{V_i}$, where $M_{i,t}$ is the mass for site i at time t, and V_i is the volume of site i. Due to it being logarithmic and $\frac{M_{i,t}}{V_i}$ having a range of $\{0, \infty^+\}$, an offset δ is added to ensure that $\delta + \frac{M_{i,t}}{V_i} > 0$. This concentration is also presumed to be measured with an error, which is represented as sampling from a normal distribution with mean $log\left(\delta + \frac{M_{i,t}}{V_i}\right)$ and variance $\frac{1}{\tau}^2$, where τ is our precision parameter.

Assumptions

- The growth rate, α. is equal across all sites and times
- The decay rate, β, is equal across all sites and times
- I_k, T_k, W_k, S_k are independent from all other events
- $log(Y_{i,t})$ is independently sampled from $N\left(log\left(\delta + \frac{M_{i,t}}{V_i}\right), \frac{1}{\sqrt{\tau}}^2\right)$
- There are no seasonal components
- The water at each site is well-mixed

Estimation

Parameters

The parameters that are known are:

- Number of sites, n
  - n will be set to 10
- Time period, T
  - T will be set to 100
- Dispersion Matrix, P
  - P is the transportation matrix above
- Number of events, N
  - N will be set to 5 for our sample
- Volumes, V_i
  - V_i will be samples from $NegBinom(1, 1) + 1$
- Concentration offset, δ
  - δ will be set to 10^{-10}

And the parameters that need to be estimated are:

- Growth Parameter, α
- Decay Parameter, β
- Event site, $[I_1, ..., I_N]$
- Event Starting time, $[T_1, ..., T_N]$
- Event time length, $[W_1, ..., W_N]$
- Event size, $[S_1, ..., S_N]$
- Precision, τ

Where the initial values will be derived from the priors below.

Priors

$\alpha \sim Gamma(2, 1)$

Such that $E(\alpha) = 2$, since an $\alpha \leq 1$ would cause it to not grow and blooms would immediately die out.

$\beta \sim Beta(1, 1)$

Which is equivalent to $Uniform(0, 1)$, such that it is an uninformative prior but any values outside the domain of $\{0, 1\}$ have zero probability.

$I_k \sim DiscreteUniform(1, n)$

This is an uninformative prior that considers all sites to be equally likely to have a bloom.

$T_k \sim DiscreteUniform(1, T)$

This is an uninformative prior that considers all time periods to be equally likely.

$W_k \sim NegBinomial(6, 0.75)$

A prior such that $E\left(W_k\right) = 2$, so that event lengths are relatively small but also should give a high enough event length such that α can be approximated.

$S_k \sim Gamma(1, 1)$

An uninformative prior where $E\left(S_k\right) = 1$ and prefers a small cluster of values.

$\tau \sim Gamma(1, 1)$

An uninformative prior where $E(\tau) = 1$ and prefers a small cluster of values.

Simulation

Since we have no real data, we will need to simulate it. We will do this using the RunSimulation function outlined below, where we randomly generate the true values for our parameters and then use the Mass equation outlined above to calculate $M_{i,t}$ for all $i,\ t$.

However, to generate a sample we will use a variation of RunSimulation with a fixed $N\ =\ 5$ such that we can have the most even spread of varying predicted numbers of events, with both undercounting and overcounting blooms.

Due to the fact that we want to compare the effectiveness of the inference amongst different scenarios, we will generate a singular sample and save it along with the true parameters for that sample for repeated use.

Assumptions

- The growth rate, α. is equal across all sites and times
- The decay rate, β, is equal across all sites and times
- I_k, T_k, W_k, S_k are independent from all other events
- There is an equal probability for $I_k = i$ amongst all sites
- There is an equal probability for $T_k = t$ amongst all time periods
- All parameters are independent
- $log\left(Y_{i,t}\right)$ is independently sampled from $N\left(log\left(\delta + \frac{M_{i,t}}{V_i}\right), \frac{1}{\sqrt{\tau}}^2\right)$ for all $i,\ t$
- There are no seasonal components

Sample

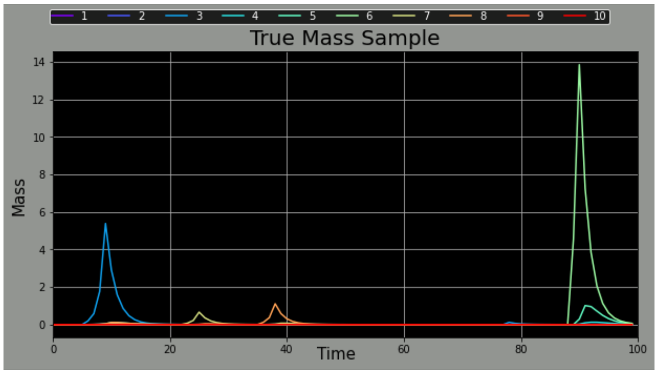
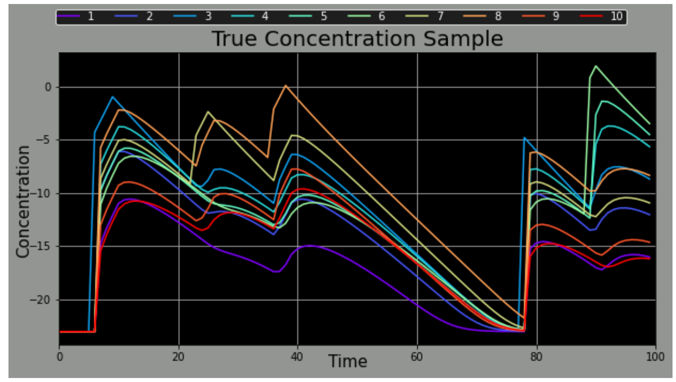
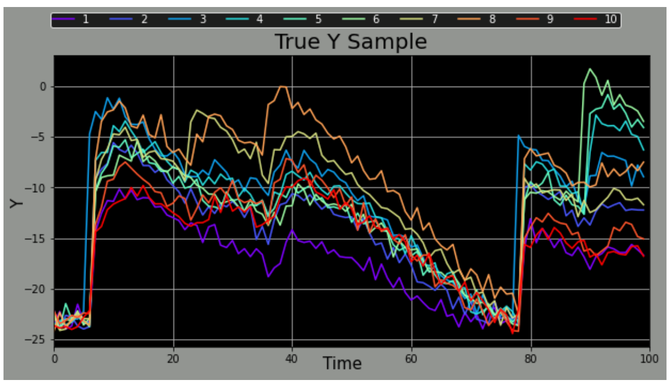
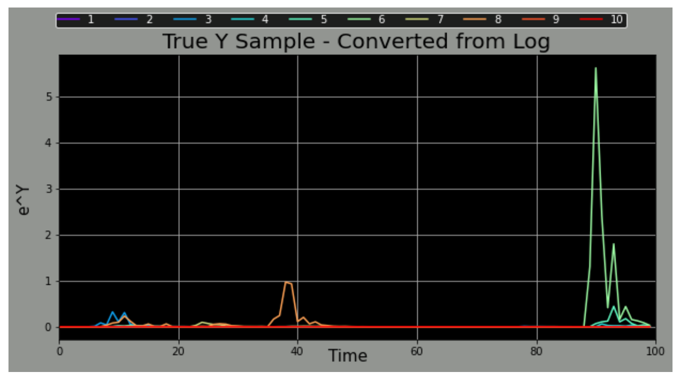
Mass	Concentration
	
	
Y	e^Y

Figure 3: Plots of the Mass, Concentration, Y, and e^Y for the true sample

Parameter	Value(s)
α	2.5086
β	0.5927
I	[7, 2, 2, 5, 6]
T	[36, 78, 6, 89, 23]
W	[3, 1, 4, 2, 3]
S	[0.1164, 0.1162, 0.1900, 4.5746, 00689]
τ	2.2986

Table 1: Table of the true parameters

Pseudocode

$\text{FindM}_{\text{New}}$

The purpose of this function is to calculate and output $M_{i,t}$ from the mass equation outlined above, using $M_{i,t-1}$ and the given parameters. It does this by taking the sum of the separate components which have been calculated. It then returns $M_{i,t}$

Inputs:

- M
- α
- β
- P
- A
- B

$growth = \alpha \times (A \cdot M)$

$transport = \beta \times (P \cdot M)$

$M_{New} = B + growth + transport$

Outputs:

- M_{New}

RunSimulation

The purpose of this function is to generate a random set of parameters and to create the mass data from these parameters. For this, it randomly generates these sets of parameters by taking a sample from predetermined mock distributions, and then uses a for loop over T time periods and continuously feeds $M_{i,t-1}$ and the parameters into **FindM**$_{\textbf{New}}$. It then returns M and all the randomly generated parameters.

Inputs:

- n
- T
- P

$\alpha \sim Gamma(2,\ 1)$

$\beta \sim Beta(4,\ 3)$

$N \sim DiscreteUniform\left(1,\ \frac{T}{10}\right)$

$For\ k\ in\ 1:N:$

$$I_k \sim DiscreteUniform(1,\ n)$$

$$T_k \sim DiscreteUniform(1,\ T)$$

$$W_k \sim NegBinom(6,\ 0.75)\ +\ 1$$

$$S_k \sim Gamma(1,\ 1)$$

$$A_{T_k:min(T_k+W_k,\,T),\,I_k} = 1$$

$$B_{t,i} = S_k$$

$\left[M_{0,0},\ldots,\ M_{n,0}\right] = 0$

$For\ t\ in\ 1:T:$

$$M_{New} = FindM_{New}\left(\left[M_{0,t-1},\ldots,\ M_{n,t-1}\right],\ \alpha,\ \beta,\ P_{i:n,\,i:n},\ \left[A_{t,1},\ldots,A_{t,n}\right],\ \left[B_{t,1},\ldots,B_{t,n}\right]\right)$$

$$\left[M_{0,t},\ldots,\ M_{n,t}\right] = M_{New}$$

Outputs:

- $\left[\left[M_{0,1},\ldots,\ M_{n,1}\right],\ldots,\left[M_{0,T},\ldots,\ M_{n,T}\right]\right]$
- α
- β
- N
- $\left[I_0,\ldots,\ I_N\right]$
- $\left[T_0,\ldots,\ T_N\right]$
- $\left[W_0,\ldots,\ W_N\right]$
- $\left[S_0,\ldots,\ S_N\right]$

GetSample

The purpose of this function is to produce the mass data given a predetermined set of parameters. It does this in a similar fashion to **RunSimulation** but takes the parameters as inputs as opposed to randomly generating themselves within the function. It then returns M.

Inputs:

- n
- T
- P
- α
- β
- N
- $\left[I_1, ..., I_N\right]$
- $\left[t_1, ..., t_N\right]$
- $\left[W_1, ..., W_N\right]$
- $\left[s_1, ..., s_N\right]$

$For\ k\ in\ 1: N$:

$$A_{t_k:min(t_k + W_k, t), I_k} = 1$$

$$B_{t, i} = s_k$$

$$\left[M_{0,0}, ..., M_{n,0}\right] = 0$$

$For\ t\ in\ 1: T$:

$$M_{New} = FindM_{New}\left(\left[M_{0,t-1}, ..., M_{n,t-1}\right], \alpha, \beta, P_{i:n, i:n}, \left[A_{t,1}, ..., A_{t,n}\right], \left[B_{t,1}, ..., B_{t,n}\right]\right)$$

$$\left[M_{0,t}, ..., M_{n,t}\right] = M_{New}$$

Outputs:

- $\left[\left[M_{0,1}, ..., M_{n,1}\right], ..., \left[M_{0,T}, ..., M_{n,T}\right]\right]$

ConvertConcentration

The purpose of this function is to take the mass data returned from either **RunSimulation** or **GetSample**, along with the volumes V and δ, and convert it into $log\left(\delta + \frac{M}{V}\right)$ format that is the mean of $log\left(Y_{i,t}\right)$'s normal distribution. It then returns this array.

Inputs:

- n
- T
- M
- V
- δ

$For\ i\ in\ 1{:}\ n$:

$$C_{i,\,1:T} = \frac{M_{i,\,1:T}}{V_{i}}$$

$$Concentration = log(C + \delta)$$

Outputs:

- $Concentration$

ConvertY

The purpose of this function is to take the array of $log\left(\delta + \frac{M}{V}\right)$ generated from **ConvertConcentration** and sample from the distribution for concentration. It does this by taking this concentration array, along with the precision parameter τ, and sampling from the $N\left(log\left(\delta + \frac{M}{V}\right), \frac{1}{\sqrt{\tau}}^2\right)$ to generate $T \times n$ samples of Y. It then returns the Y_{log} array.

Inputs:

- C
- τ

$$Y_{log} \sim N\left(C, \frac{1}{\sqrt{\tau}}^2\right)$$

Outputs:

- Y_{log}

GenerateObservation

Due to the practical context of taking samples from sites out in the Marlborough sounds, one of the scenarios is dealing with sparse data. Therefore, the purpose of this function is to create an observation array O where $O_{i,t} = 1$ if the data is present and $O_{i,t} = 0$ otherwise.

There are two methods to do this: Remove data randomly based upon a probability $P\left(O_{i,t} = 0\right) = p$ for some value p, where it is independent amongst all i, t; or remove specific sites $\left\{I_1, ..., I_X\right\}$ where $O_{i,t} = 0 \; if \; i \in \left\{I_1, ..., I_X\right\}$ and $O_{i,t} = 1$ otherwise.

Random

Inputs:

- n
- T

- p

$O_{1:n,\,1:T} = 1$

$For\ i\ in\ 1{:}\ n{:}$

$For\ t\ in\ 1{:}\ T{:}$

$u \sim Uniform(0,\ 1)$

$if\ u\ \geq p{:}$

$O_{i,\,t} = 0$

Outputs:

- O

Sites

Inputs:

- n
- T
- $\left[I_1, \ldots I_X\right]$

$O_{1:n,\,1:T} = 1$

$For\ I\ in \left[I_1, \ldots I_X\right]{:}$

$O_{I,\,1:T} = 0$

Outputs:

- O

Metropolis-Hastings Algorithm

The purpose of this code is to execute the Bayesian inference via the Metropolis-Hastings Method previously outlined and approximate the parameters. It does this by first generating a sample with random parameters using **RunSimulation**, and using **ConvertConcentration** and

ConvertY to create the Y sample. It also generates initial values for all parameters θ_0 by sampling from the prior distributions, an Observation Array via **GenerateObservation** and volumes from $NegBinom(1, 1) + 1$. It also generates initial parameter values θ_0 from the priors outlined above, and calculates the initial $M_{Current}$ and $C_{Current}$ arrays. Then, looping over the 10000 iterations, for each parameter to be estimated, it:

- Randomly generates $\theta_{Candidate}$ from the proposal distributions $q(\theta_S|\theta_{S-1})$
- Generate $M_{Candidate}/C_{Candidate}$ using $\theta_{Candidate}$ with **GetSample** and **ConvertConcentration**
- Calculate $ratio = log(P(Y|\theta^*)) + log(P(\theta^*)) + log(q(\theta_{s-1}|\theta^*)) - (log(P(Y|\theta_{s-1})) + log(P(\theta_{s-1})) + log(q(\theta^*|\theta_{s-1})))$
- Accept or reject $\theta_{Candidate}$ with probability e^{ratio}
    - If Accepted, $M_{Current} = M_{Candidate}$, $C_{Current} = C_{Candidate}$
- After all parameters, calculate DIC and $P(Y|\theta)$

Inputs:

- P

$n = 10$

$T = 100$

$\delta = 1^{-10}$

$\tau_{True} \sim Gamma(1, 1)$

$[V_1, ..., V_n] \sim NegBinom(1, 1) + 1$

$[M_{True}, \alpha_{True}, \beta_{True}, N_{True}, I_{True}, T_{True}, W_{True}, S_{True}] = runSimulation(n, T, P)$

$C_{True} = ConvertCalculation(n, T, M_{True}, V, \delta)$

$Y_{True} = ConvertY(C_{True}, \tau_{True})$

$O = generateObservationArray$

$\alpha_0 \sim Gamma(2, 1)$

$\beta_0 \sim Beta(1, 1)$

$N = N_{True}$

$[I_{0,1}, ..., I_{0,N}] \sim DiscreteUniform(1, n)$

$\left[t_{1,1}, \ldots, t_{0,N}\right] \sim DiscreteUniform(1, T)$

$\left[w_{0,1}, \ldots, w_{0,N}\right] \sim NegBinom(6, 0.75)$

$\left[s_{0,1}, \ldots, s_{0,N}\right] \sim Gamma(1, 1)$

$\tau_0 \sim Gamma(1, 1)$

$\phi = \left[\alpha_0, \beta_0, N, \left[I_{0,1}, \ldots, I_{0,N}\right], \left[t_{0,1}, \ldots, t_{0,N}\right], \left[w_{0,1}, \ldots, w_{0,N}\right], \left[s_{0,1}, \ldots, s_{0,N}\right], \tau_0\right]$

$For\ each\ parameter_0\ in\ \phi:$

$Array_{parameter} = parameter_0$

$Accept_{parameter} = 0$

$M_{Current} = GetSample\left(n, T, \phi_\alpha, \phi_\beta, \phi_N, \phi_I, \phi_T, \phi_W, \phi_S\right)$

$C_{Current} = ConvertConcentration\left(n, T, M_{Current}, V, \delta\right)$

$For\ s\ in\ 1: 10000:$

Alpha

$\alpha_{Candidate} \sim N\left(\alpha_{Current}, 8^2\right)$

$M_{Candidate} = GetSample\left(n, T, \alpha_{Candidate}, \phi_\beta, \phi_N, \phi_I, \phi_T, \phi_W, \phi_S\right)$

$C_{Candidate} = ConvertConcentration\left(n, T, M_{Candidate}, V, \delta\right)$

$p_{prior_{Candidate}} = Gamma(2, 1).logPDF\left(\alpha_{Candidate}\right)$

$p_{prior_{Current}} = Gamma(2, 1).logPDF\left(\phi_\alpha\right)$

$$p_{Likelihood_{Candidate}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$p_{Likelihood_{Current}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$p_{Q_{Candidate}} = N\left(\phi_{\alpha}, 8^{2}\right).logPDF\left(\alpha_{Candidate}\right)$$

$$p_{Q_{Current}} = N\left(\alpha_{Candidate}, 8^{2}\right).logPDF\left(\phi_{\alpha}\right)$$

$$ratio_{\alpha} = p_{prior_{Candidate}} + p_{Likelihood_{Candidate}} - p_{prior_{Current}} - p_{Likelihood_{Current}} + p_{Q_{Current}} - p_{Q_{Candidate}}$$

$$u \sim U(0, 1)$$

$$if\ ratio > log(u):$$

$$\phi_{\alpha} = \alpha_{Candidate}$$

$$Accept_{\alpha} = Accept_{\alpha} + 1$$

$$M_{Candidate} = M_{Current}$$

$$C_{Candidate} = C_{Current}$$

$$Likelihood_{\alpha_{Observed_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{\alpha_{Full_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}}, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$$else:$$

$$Likelihood_{\alpha_{Observed_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{\alpha_{Full_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_\tau}}^2\right).logPDF\left(Y_{True_{it}}\right)\right)$$

Beta

$$\beta_{Candidate} \sim Beta(1, 1)$$

$$M_{Candidate} = GetSample\left(n, T, \phi_\alpha, \beta_{Candidate}, \phi_N, \phi_I, \phi_T, \phi_W, \phi_S\right)$$

$$C_{Candidate} = ConvertConcentration\left(n, T, M_{Candidate}, V, \delta\right)$$

$$p_{prior_{Candidate}} = Beta(1, 1).logPDF\left(\beta_{Candidate}\right)$$

$$p_{prior_{Current}} = Beta(1, 1).logPDF\left(\phi_\beta\right)$$

$$p_{Likelihood_{Candidate}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}}, \frac{1}{\sqrt{\phi_\tau}}^2\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$$p_{Likelihood_{Current}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}}, \frac{1}{\sqrt{\phi_\tau}}^2\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$$p_{Q_{Candidate}} = Beta(1, 1).logPDF\left(\beta_{Candidate}\right)$$

$$p_{Q_{Current}} = Beta(1, 1).logPDF\left(\phi_\beta\right)$$

$$ratio_\beta = p_{prior_{Candidate}} + p_{Likelihood_{Candidate}} - p_{prior_{Current}} - p_{Likelihood_{Current}} + p_{Q_{Current}} - p_{Q_{Candidate}}$$

$$u \sim U(0, 1)$$

$if\ ratio > u$:

$$\phi_\beta = \beta_{Candidate}$$

$$Accept_{\beta} = Accept_{\beta} + 1$$

$$M_{Candidate} = M_{Current}$$

$$C_{Candidate} = C_{Current}$$

$$Likelihood_{\beta_{Observed_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{\beta_{Full_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}}, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$else$:

$$Likelihood_{\beta_{Observed_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{\beta_{Full_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$for\ k\ in\ 1{:}\ n$:

$$I_{k_{Candidate}} \sim DiscreteUniform(1, n)$$

$$I_{Candidate} = \phi_{I}$$

$$I_{Candidate_k} = I_{k_{Candidate}}$$

$$M_{Candidate} = GetSample\left(n, T, \phi_{\alpha}, \phi_{\beta}, \phi_{N}, I_{Candidate}, \phi_{T}, \phi_{W}, \phi_{S}\right)$$

$$C_{Candidate} = ConvertConcentration\left(n, T, M_{Candidate}, V, \delta\right)$$

$$p_{prior_{Candidate}} = DiscreteUniform(1, n).logPMF\left(I_{Candidate_k}\right)$$

$$p_{prior_{Current}} = DiscreteUniform(1,\ n).logPMF\left(\phi_{I_k}\right)$$

$$p_{Likelihood_{Candidate}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}},\ \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$$p_{Likelihood_{Current}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}},\ \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$$p_{Q_{Candidate}} = DiscreteUniform(1,\ n).logPMF\left(I_{Candidate_k}\right)$$

$$p_{Q_{Current}}\ DiscreteUniform(1,\ n).logPMF\left(\phi_{I_k}\right)$$

$$ratio_{I_k} = p_{prior_{Candidate}} + p_{Likelihood_{Candidate}} - p_{prior_{Current}} - p_{Likelihood_{Current}} + p_{Q_{Current}} - p_{Q_{Candidate}}$$

$$u \sim U(0,\ 1)$$

$if\ ratio > log(u)$:

$$\phi_{I} = I_{Candidate}$$

$$Accept_{I} = Accept_{I} + \frac{1}{N}$$

$$M_{Candidate} = M_{Current}$$

$$C_{Candidate} = C_{Current}$$

$$Likelihood_{I_{Observed_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}} \cdot O,\ \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{I_{Full_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}},\ \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$else$:

$$Likelihood_{I_{Observed_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right). logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{I_{Full_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right). logPDF\left(Y_{True_{it}}\right)\right)$$

T

$for\ k\ in\ 1{:}\ n{:}$

$$T_{k_{Candidate}} \sim DiscreteUniform(1, T)$$

$$T_{Candidate} = \phi_{T}$$

$$T_{Candidate_{k}} = T_{k_{Candidate}}$$

$$M_{Candidate} = GetSample\left(n, T, \phi_{\alpha}, \phi_{\beta}, \phi_{N}, \phi_{I}, T_{Candidate}, \phi_{W}, \phi_{S}\right)$$

$$C_{Candidate} = ConvertConcentration\left(n, T, M_{Candidate}, V, \delta\right)$$

$$p_{prior_{Candidate}} = DiscreteUniform(1, T). logPMF\left(T_{Candidate_{k}}\right)$$

$$p_{prior_{Current}} = DiscreteUniform(1, T). logPMF\left(\phi_{T_{k}}\right)$$

$$p_{Likelihood_{Candidate}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}}, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right). logPDF\left(Y_{True_{it}}\right)\right)$$

$$p_{Likelihood_{Current}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}}, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right). logPDF\left(Y_{True_{it}}\right)\right)$$

$$p_{Q_{Candidate}} = DiscreteUniform(1,\ T).\, logPMF\left(T_{Candidate_k}\right)$$

$$p_{Q_{Current}} = DiscreteUniform(1,\ T).\, logPMF\left(\phi_{4_k}\right)$$

$$ratio_{T_k} = min\left(p_{prior_{Candidate}} + p_{Likelihood_{Candidate}} - p_{prior_{Current}} - p_{Likelihood_{Current}} + p_{Q_{Current}} - p_{Q_{Candidate}},\ 1\right)$$

$$u \sim U(0,\ 1)$$

$if\ ratio > log(u)$:

$$\phi_T = T_{Candidate}$$

$$Accept_T = Accept_T + \frac{1}{N}$$

$$M_{Candidate} = M_{Current}$$

$$C_{Candidate} = C_{Current}$$

$$Likelihood_{T_{Observed_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}} \cdot O,\ \frac{1}{\sqrt{\phi_\tau}}^2\right).\, logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{T_{Full_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}},\ \frac{1}{\sqrt{\phi_\tau}}^2\right).\, logPDF\left(Y_{True_{it}}\right)\right)$$

$else$:

$$Likelihood_{T_{Observed_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O,\ \frac{1}{\sqrt{\phi_\tau}}^2\right).\, logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{T_{Full_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O,\ \frac{1}{\sqrt{\phi_\tau}}^2\right).\, logPDF\left(Y_{True_{it}}\right)\right)$$

W

$for\ k\ in\ 1{:}\ n{:}$

$$W_{k_{Candidate}} \sim Binom\left(T,\ \frac{1+\phi_{W_k}}{2+T}\right)$$

$$W_{Candidate} = \phi_W$$

$$W_{Candidate_k} = W_{k_{Candidate}}$$

$$M_{Candidate} = GetSample\left(n,\ T,\ \phi_{\alpha},\ \phi_{\beta},\ \phi_{N},\ \phi_{I},\ \phi_{T},\ W_{Candidate},\ \phi_{S}\right)$$

$$C_{Candidate} = ConvertConcentration\left(n,\ T,\ M_{Candidate},\ V,\ \delta\right)$$

$$p_{prior_{Candidate}} = DiscreteUniform(1,\ T).logPMF\left(W_{Candidate_k}\right)$$

$$p_{prior_{Current}} = DiscreteUniform(1,\ T).logPMF\left(\phi_{W_k}\right)$$

$$p_{Likelihood_{Candidate}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}},\ \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$$p_{Likelihood_{Current}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}},\ \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$$p_{Q_{Candidate}} = Binom\left(T,\ \frac{1+\phi_{W_k}}{2+T}\right).logPMF\left(W_{Candidate_k}\right)$$

$$p_{Q_{Current}}\ Binom\left(T,\ \frac{1+W_{Candidate_k}}{2+T}\right).logPMF\left(\phi_{W_k}\right)$$

$$ratio_{W_k} = p_{prior_{Candidate}} + p_{Likelihood_{Candidate}} - p_{prior_{Current}} - p_{Likelihood_{Current}} + p_{Q_{Current}} - p_{Q_{Candidate}}$$

$$u \sim Uniform(0,\ 1)$$

$if\ ratio > log(u)$:

$\phi_W = W_{Candidate}$

$Accept_W = Accept_W + \frac{1}{N}$

$M_{Candidate} = M_{Current}$

$C_{Candidate} = C_{Current}$

$$Likelihood_{W_{Observed_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}} \cdot O, \frac{1}{\sqrt{\phi_\tau}}^2\right). logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{W_{Full_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}}, \frac{1}{\sqrt{\phi_\tau}}^2\right). logPDF\left(Y_{True_{it}}\right)\right)$$

$else$:

$$Likelihood_{W_{Observed_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_\tau}}^2\right). logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{W_{Full_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_\tau}}^2\right). logPDF\left(Y_{True_{it}}\right)\right)$$

S

$for\ k\ in\ 1{:}n$:

$S_{k_{Candidate}} \sim N\left(\phi_{S_k}, 0.1^2\right)$

$S_{Candidate} = \phi_S$

$S_{Candidate_k} = S_{k_{Candidate}}$

$$M_{Candidate} = GetSample\left(n,\ T,\ \phi_{\alpha},\ \phi_{\beta},\ \phi_{N},\ \phi_{I},\ \phi_{T},\ \phi_{W},\ S_{Candidate}\right)$$

$$C_{Candidate} = ConvertConcentration\left(n,\ T,\ M_{Candidate},\ V,\ \delta\right)$$

$$p_{prior_{Candidate}} = Gamma(1,\ 1).logPDF\left(S_{Candidate_k}\right)$$

$$p_{prior_{Current}} = Gamma(1,\ 1).logPDF\left(\phi_{S_k}\right)$$

$$p_{Likelihood_{Candidate}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}},\ \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$$p_{Likelihood_{Current}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}},\ \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$$p_{Q_{Candidate}} = N\left(\phi_{S_k},\ 0.1^{2}\right).logPDF\left(S_{Candidate_k}\right)$$

$$p_{Q_{Current}} N\left(S_{Candidate_k},\ 0.1^{2}\right).logPDF\left(\phi_{S_k}\right)$$

$$ratio_{S_k} = p_{prior_{Candidate}} + p_{Likelihood_{Candidate}} - p_{prior_{Current}} - p_{Likelihood_{Current}} + p_{Q_{Current}} - p_{Q_{Candidate}}$$

$$u \sim U(0,\ 1)$$

$if\ ratio > log(u)$:

$$\phi_{S} = S_{Candidate}$$

$$Accept_{S} = Accept_{S} + \frac{1}{N}$$

$$M_{Candidate} = M_{Current}$$

$$C_{Candidate} = C_{Current}$$

$$Likelihood_{S_{Observed_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{S_{Full_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}}, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

$else$:

$$Likelihood_{S_{Observed_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{S_{Full_{s,k}}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}}\right)\right)$$

Tau

$$\tau_{Candidate} \sim N\left(\phi_{\tau}, \frac{1}{2}^{2}\right)$$

$$p_{prior_{Candidate}} = Gamma(1,\ 1).logPDF\left(\tau_{Candidate}\right)$$

$$p_{prior_{Current}} = Gamma(1,\ 1).logPDF\left(\phi_{\tau}\right)$$

$$p_{Likelihood_{Candidate}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\tau_{Candidate}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$p_{Likelihood_{Current}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right).logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$p_{Q_{Candidate}} = N\left(\phi_{\tau}, \frac{1}{2}^{2}\right).logPDF\left(\tau_{Candidate}\right)$$

$$p_{Q_{Current}} N\left(\tau_{Candidate}, \frac{1}{2}^{2}\right).logPDF\left(\phi_{\tau}\right)$$

$$ratio_{\tau} = p_{prior_{Candidate}} + p_{Likelihood_{Candidate}} - p_{prior_{Current}} - p_{Likelihood_{Current}} + p_{Q_{Current}} - p_{Q_{Candidate}}$$

$$u \sim U(0, 1)$$

$$if\ ratio > log(u):$$

$$\phi_{\tau} = \tau_{Candidate}$$

$$Accept_{\tau} = Accept_{\tau} + 1$$

$$Likelihood_{\tau_{Observed_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right). logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{\tau_{Full_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Candidate_{it}}, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right). logPDF\left(Y_{True_{it}}\right)\right)$$

$$else:$$

$$Likelihood_{\tau_{Observed_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right). logPDF\left(Y_{True_{it}} \cdot O\right)\right)$$

$$Likelihood_{\tau_{Full_s}} = \sum_{t=1}^{T}\left(\sum_{i=1}^{n} N\left(C_{Current_{it}} \cdot O, \frac{1}{\sqrt{\phi_{\tau}}}^{2}\right). logPDF\left(Y_{True_{it}}\right)\right)$$

Proposal Distributions

$\alpha^* \sim N\left(\phi_\alpha, 2^2\right)$

Using a random walk proposal with $E\left(\alpha^*\right) = \phi_\alpha$. Using a standard deviation of $\sigma = 2$ because through testing, α has had a higher variance, especially compared to other parameters with normally distributed proposals like S_k and τ.

$\beta^* \sim Beta(1, 1)$

An independence proposal in order to sample only from β's possible range of values of $[0, 1]$

$I_k^* \sim DiscreteUniform(1, n)$

An independence proposal in order to sample only from I_k's possible range of values of $[0, n]$

$T_k^* \sim DiscreteUniform(1, T)$

An independence proposal in order to sample only from T_k's possible range of values of $[0, T]$

$W_k^* \sim Binom\left(T, \frac{1+\phi_{W_k}}{2+T}\right)$

A random walk sampler over a discrete domain where $E\left(W_k^*\right) \approx \phi_{W_k}$ as possible, with the added requirements that $p\left(W_k^* | \phi_{W_k} = 0\right) \neq 0$ (Hence the $+\ 1$), and $p\left(W_k^* | \phi_{W_k} = T\right) \neq 1$ (Hence the $+\ 2$)

$S_k^* \sim N\left(\phi_{S_k}, 1^2\right)$

Using a random walk proposal with $E\left(S_k^*\right) = \phi_{S_k}$. Using a standard deviation of $\sigma = 1$ because through testing, the default value of 1 allowed an appropriate amount of variance for S_k.

$\tau^* \sim N\left(\phi_\tau, \frac{1}{2}^2\right)$

Using a random walk proposal with $E\left(\tau^*\right) = \phi_\tau$. Using a standard deviation of $\sigma = \frac{1}{2}$ because through testing, τ had a fairly tight distribution compared to α and S_k.

Acceptance Probabilities

Alpha

$$p_{prior_{Candidate}} = log(\alpha_{Candidate}) - \alpha_{Candidate}$$

$$p_{prior_{Current}} = log(\alpha_{Current}) - \alpha_{Current}$$

$$p_{Likelihood_{Candidate}} = \frac{-\tau}{32}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Candidate_{it}}\right)^2\right)$$

$$p_{Likelihood_{Current}} = \frac{-\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Current_{it}}\right)^2\right)$$

$$p_{Q_{Candidate}} = -\frac{1}{8}(\alpha_{Candidate} - \alpha_{Current})^2$$

$$p_{Q_{Candidate}} = -\frac{1}{8}(\alpha_{Current} - \alpha_{Candidate})^2$$

$$ratio_{log} = log\left(\frac{\alpha_{Candidate}}{\alpha_{Current}}\right) - \alpha_{Candidate} + \alpha_{Current} + \frac{\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Current_{it}}\right)^2 - \left(Y_{True_{it}} - C_{Candidate_{it}}\right)^2\right)$$

$$P_{Acceptance}(\alpha) = e^{ratio_{log}}$$

Beta

$$p_{prior_{Candidate}} = 0$$

$$p_{prior_{Current}} = 0$$

$$p_{Likelihood_{Candidate}} = \frac{-\tau}{2} \sum_{t=1}^{T} \left(\sum_{i=1}^{n} \left(Y_{True_{it}} - C_{Candidate_{it}} \right)^2 \right)$$

$$p_{Likelihood_{Current}} = \frac{-\tau}{2} \sum_{t=1}^{T} \left(\sum_{i=1}^{n} \left(Y_{True_{it}} - C_{Current_{it}} \right)^2 \right)$$

$$p_{Q_{Candidate}} = 0$$

$$p_{Q_{Candidate}} = 0$$

$$ratio_{log} = \frac{\tau}{2} \sum_{t=1}^{T} \left(\sum_{i=1}^{n} \left(Y_{True_{it}} - C_{Current_{it}} \right)^2 - \left(Y_{True_{it}} - C_{Candidate_{it}} \right)^2 \right)$$

$$P_{Acceptance}(\beta) = e^{ratio_{log}}$$

I_k

$$p_{prior_{Candidate}} = -\log(n)$$

$$p_{prior_{Current}} = -\log(n)$$

$$p_{Likelihood_{Candidate}} = \frac{-\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Candidate_{it}}\right)^2\right)$$

$$p_{Likelihood_{Current}} = \frac{-\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Current_{it}}\right)^2\right)$$

$$p_{Q_{Candidate}} = -\log(n)$$

$$p_{Q_{Candidate}} = -\log(n)$$

$$ratio_{log} = \frac{\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Current_{it}}\right)^2 - \left(Y_{True_{it}} - C_{Candidate_{it}}\right)^2\right)$$

$$P_{Acceptance}\left(I_k\right) = e^{ratio_{log}}$$

T_k

$$p_{prior_{Candidate}} = -\log(T)$$

$$p_{prior_{Current}} = -\log(T)$$

$$p_{Likelihood_{Candidate}} = \frac{-\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Candidate_{it}}\right)^2\right)$$

$$p_{Likelihood_{Current}} = \frac{-\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Current_{it}}\right)^2\right)$$

$$p_{Q_{Candidate}} = -\log(T)$$

$$p_{Q_{Candidate}} = -\log(T)$$

$$ratio_{log} = \frac{\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Current_{it}}\right)^2 - \left(Y_{True_{it}} - C_{Candidate_{it}}\right)^2\right)$$

$$P_{Acceptance}\left(T_k\right) = e^{ratio_{log}}$$

W_k

$$p_{prior_{Candidate}} = log\left(\left(W_{Candidate_k} + 5\right)C\left(W_{Candidate_k}\right)\right) + 6log\left(\frac{3}{4}\right) - W_{Candidate_k} log(4)$$

$$p_{prior_{Current}} = log\left(\left(W_{Current_k} + 5\right)C\left(W_{Current_k}\right)\right) + 6log\left(\frac{3}{4}\right) - W_{Current_k} log(4)$$

$$p_{Likelihood_{Candidate}} = \frac{-\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Candidate_{it}}\right)^2\right)$$

$$p_{Likelihood_{Current}} = \frac{-\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Current_{it}}\right)^2\right)$$

$$p_{Q_{Candidate}} = log\left((T)C\left(W_{Candidate_k}\right)\right) + W_{Candidate_k} log\left(W_{Current_k} + 1\right) + \left(T - W_{Candidate_k}\right)log\left(T + 1 - W_{Current_k}\right) + Tlog(T + 2)$$

$$p_{Q_{Candidate}} = log\left((T)C\left(W_{Current_k}\right)\right) + W_{Current_k} log\left(W_{Candidate_k} + 1\right) + \left(T - W_{Current_k}\right)log\left(T + 1 - W_{Candidate_k}\right) + Tlog(T + 2)$$

$$\begin{aligned} ratio_{log} = {} & log\left(\left(W_{Candidate_k} + 5\right)C\left(W_{Candidate_k}\right)\right) + \left(W_{Current_k} - W_{Candidate_k}\right)log(4) - log\left(\left(W_{Current_k} + 5\right)C\left(W_{Current_k}\right)\right) \\ & + \frac{\tau}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Current_{it}}\right)^2 - \left(Y_{True_{it}} - C_{Candidate_{it}}\right)^2\right) + log\left((T)C\left(W_{Candidate_k}\right)\right) + W_{Candidate_k} log\left(W_{Current_k} + 1\right) \\ & + \left(T - W_{Candidate_k}\right)log\left(T + 1 - W_{Current_k}\right) - log\left((T)C\left(W_{Current_k}\right)\right) - W_{Current_k} log\left(W_{Candidate_k} + 1\right) - \left(T - W_{Current_k}\right)log\left(T + 1 - W_{Candidate_k}\right) \end{aligned}$$

$$P_{Acceptance}\left(W_k\right) = e^{ratio_{log}}$$

S_k

$$p_{prior_{Candidate}} = -S_{Candidate_k}$$

$$p_{prior_{Current}} = -S_{Current_k}$$

$$p_{Likelihood_{Candidate}} = \frac{-\tau}{2} \sum_{t=1}^{T} \left(\sum_{i=1}^{n} \left(Y_{True_{it}} - C_{Candidate_{it}} \right)^2 \right)$$

$$p_{Likelihood_{Current}} = \frac{-\tau}{2} \sum_{t=1}^{T} \left(\sum_{i=1}^{n} \left(Y_{True_{it}} - C_{Current_{it}} \right)^2 \right)$$

$$p_{Q_{Candidate}} = -50 \left(S_{Candidate_k} - S_{Current_k} \right)^2$$

$$p_{Q_{Candidate}} = -50 \left(S_{Current_k} - S_{Candidate_k} \right)^2$$

$$ratio_{log} = -S_{Candidate_k} + S_{Current_k} + \frac{\tau}{2} \sum_{t=1}^{T} \left(\sum_{i=1}^{n} \left(Y_{True_{it}} - C_{Current_{it}} \right)^2 - \left(Y_{True_{it}} - C_{Candidate_{it}} \right)^2 \right)$$

$$P_{Acceptance}\left(S_k\right) = e^{ratio_{log}}$$

Tau

$$p_{prior_{Candidate}} = log\left(\tau_{Candidate}\right) - \tau_{Candidate}$$

$$p_{prior_{Current}} = log\left(\tau_{Current}\right) - \tau_{Current}$$

$$p_{Likelihood_{Candidate}} = \frac{-\tau_{Candidate}}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Candidate_{it}}\right)^{2}\right)$$

$$p_{Likelihood_{Current}} = \frac{-\tau_{Current}}{2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(Y_{True_{it}} - C_{Current_{it}}\right)^{2}\right)$$

$$p_{Q_{Candidate}} = -2\left(\tau_{Candidate} - \tau_{Current}\right)^{2}$$

$$p_{Q_{Candidate}} = -2\left(\tau_{Current} - \tau_{Candidate}\right)^{2}$$

Since for τ, $C_{Candidate_{it}} = C_{Current_{it}}$ for all i, t

$$ratio_{log} = log\left(\frac{\tau_{Candidate}}{\tau_{Current}}\right) - \tau_{Candidate} + \tau_{Current}$$

$$P_{Acceptance}(\tau) = e^{ratio_{log}}$$

Results

Changing Sparseness

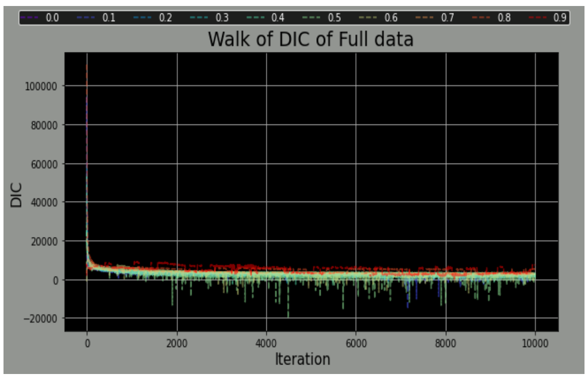
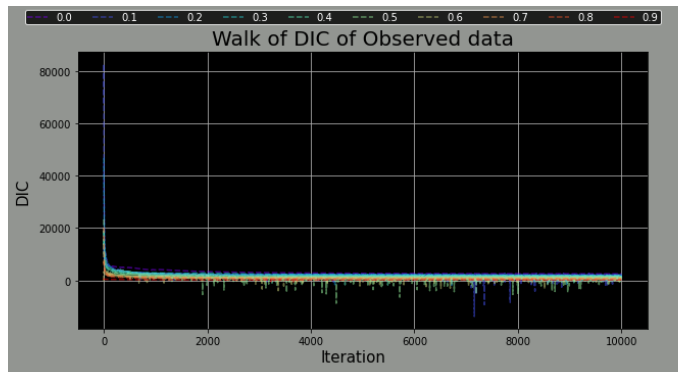
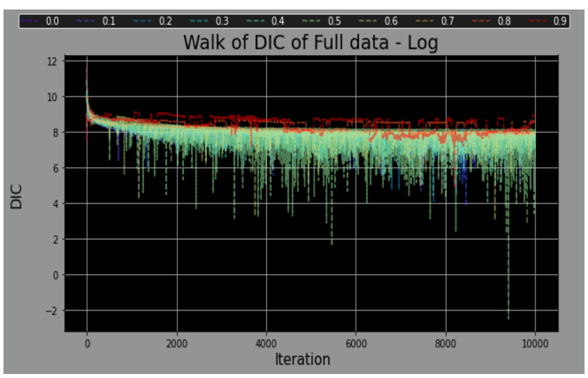
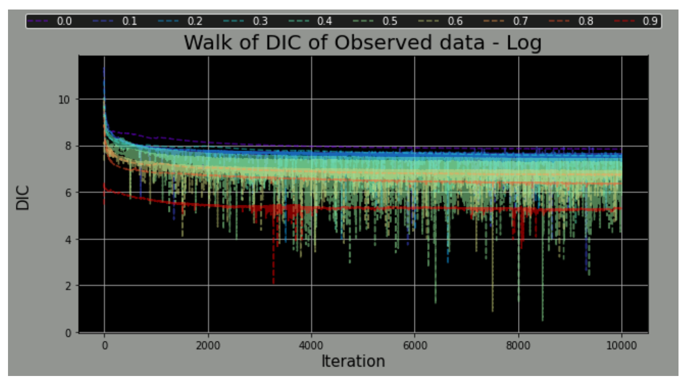
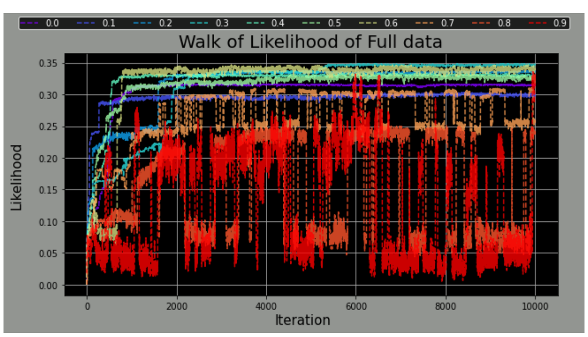
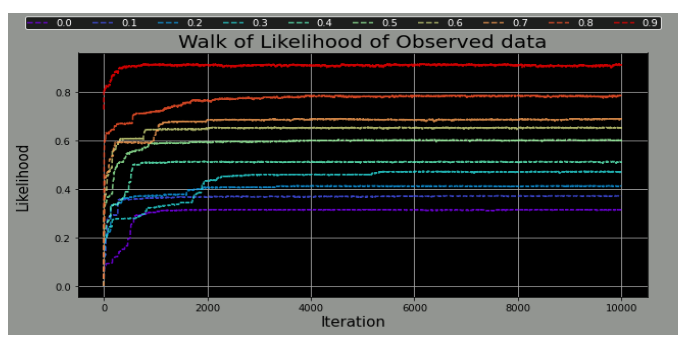
<table>
<tr><th>Model Comparison Metrics</th><th>Full Data</th><th>Observed Data</th></tr>
<tr><td>DIC</td><td>
</td><td>
</td></tr>
<tr><td>DIC - Log</td><td>
</td><td>
</td></tr>
<tr><td>Likelihood</td><td>
</td><td>
</td></tr>
</table>

Figure 4: Table of Model Comparison metrics for models with varying sparseness of data

p = 0.0

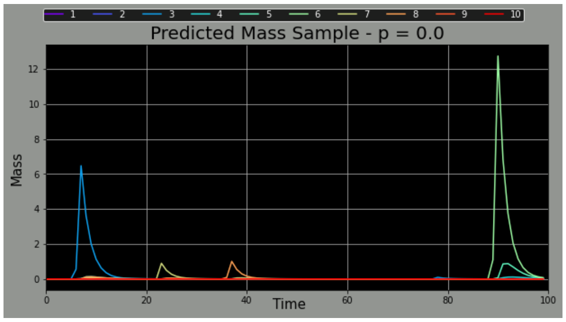
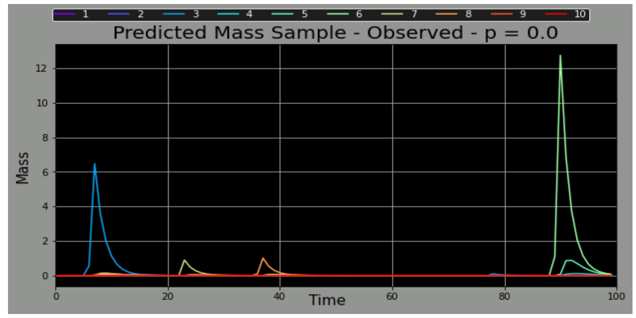
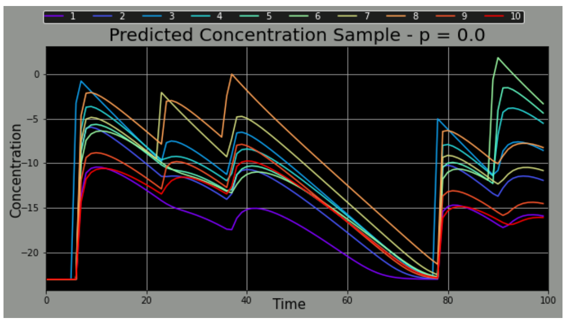
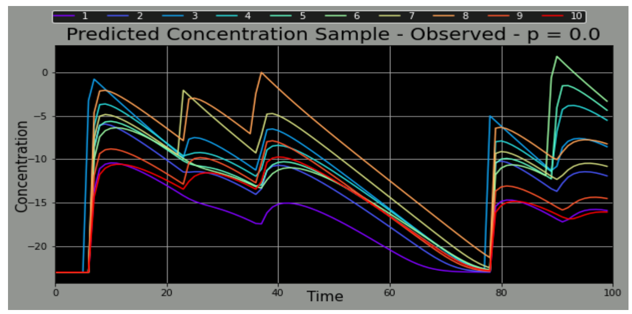
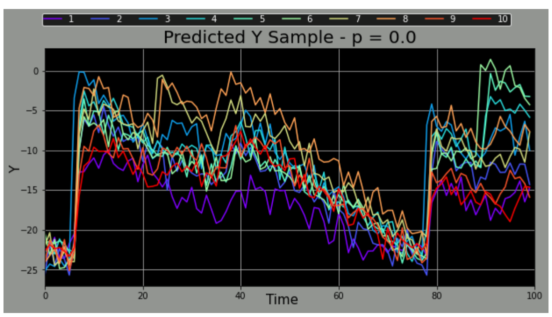
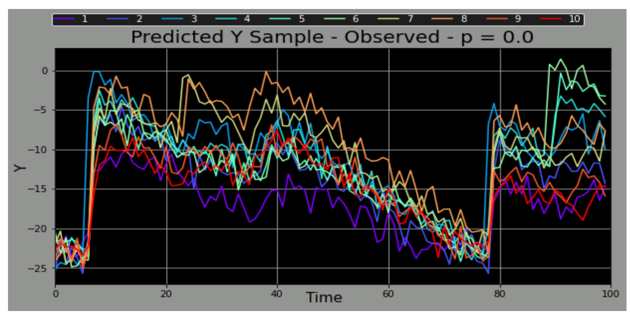
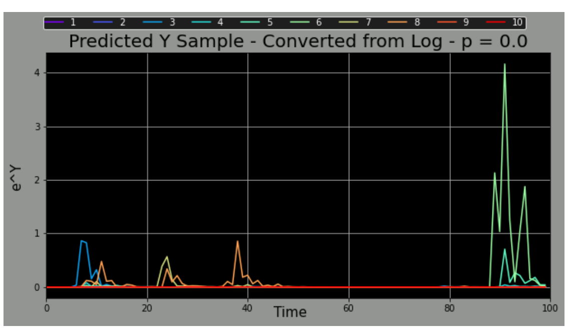
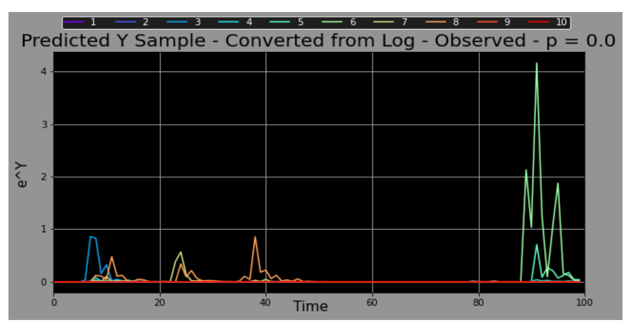
Results	Full	Observed
Mass		
Concentration		
Y		
e^Y		

Figure 5: Table of results for sparse data ($p = 0.0$)

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	11.1733	10.4534	10.5624	0.2137
β	0.6099	0.6185	0.6199	0.0056
I	7, 2, 6, 2, 5	6.9364, 2.0353, 5.9999, 2.0035, 4.9933	6.9518, 2.0491, 6.0003, 2.0077, 4.9981	0.1023
T	36, 6, 23, 78, 89	35.9387, 7.1940, 22.5807, 77.9556,88.8947	36.2395, 7.6202, 22.6791, 78.0048, 88.9957	0.0103
W	2. 2, 1, 2, 2	1.9277, 1.9363, 0.6257, 0.5889, 1.9663	1.9403, 1.9483, 0.6459, 0.6081, 1.9763	0.3346
S	0.0759, 0.5521, 0.8899, 0.0933, 1.0884	0.0832, 0.5851, 0.8549, 0.0894, 1.1534	0.0857, 0.5925, 0.8615, 0.0901, 1.1563	0.3560
τ	1.6762	1.6011	1.6137	0.1819

$Table\ 2\text{:}\ Table\ of\ summary\ statistics\ of\ parameters\ for\ sparse\ data\ (p\ =\ 0.0)$

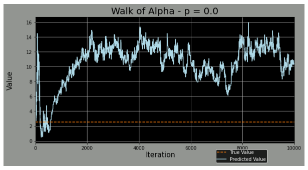
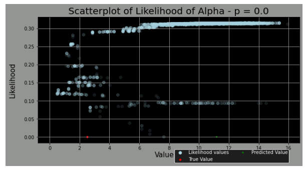
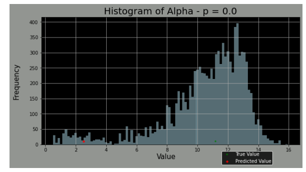
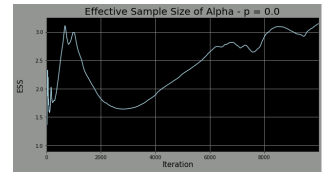
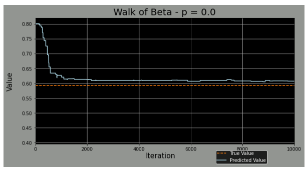
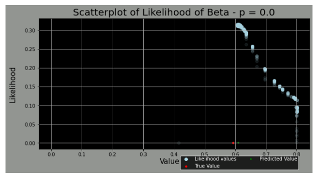
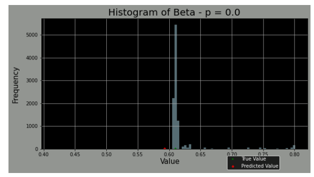
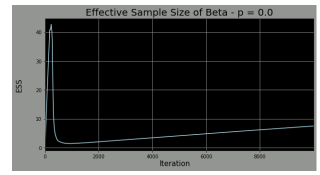
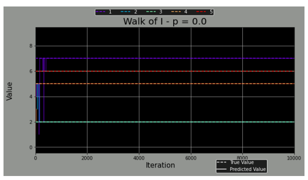
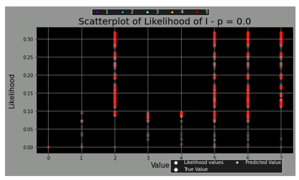
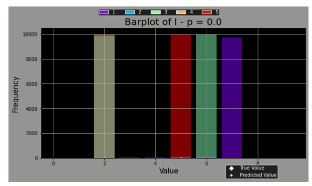
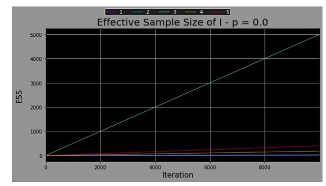
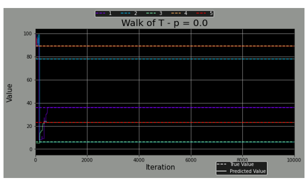
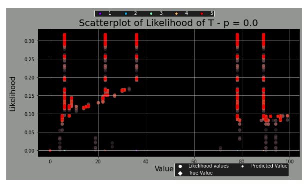
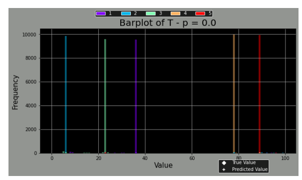
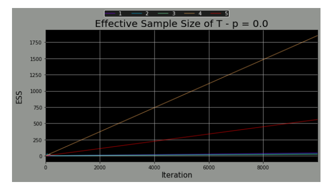
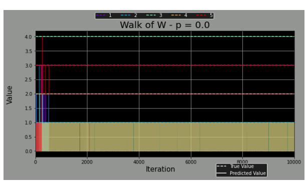
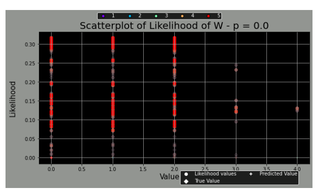
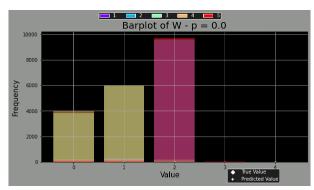
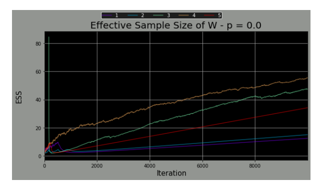
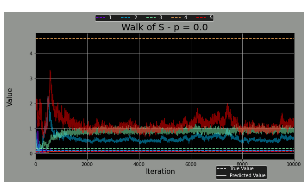
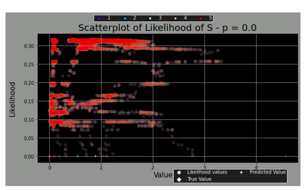
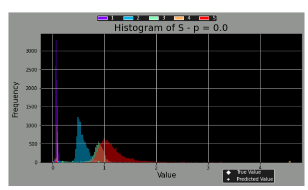
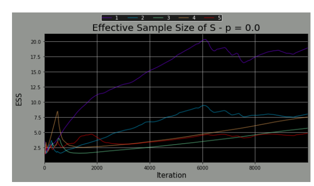
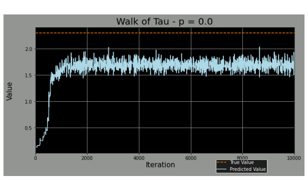
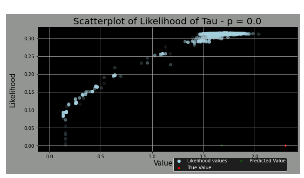
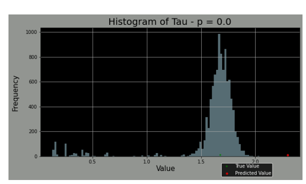
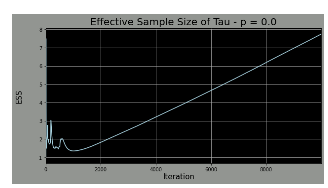
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 6: Table of graphs of parameter metrics for sparse data $(p = 0.0)$

p = 0.1

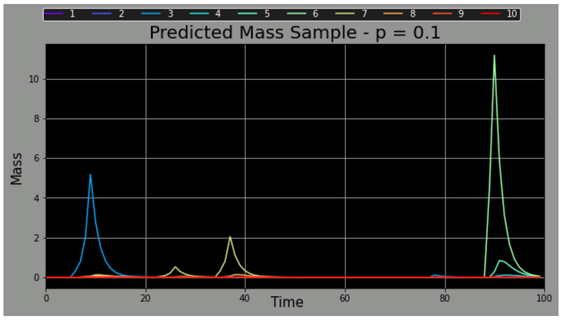
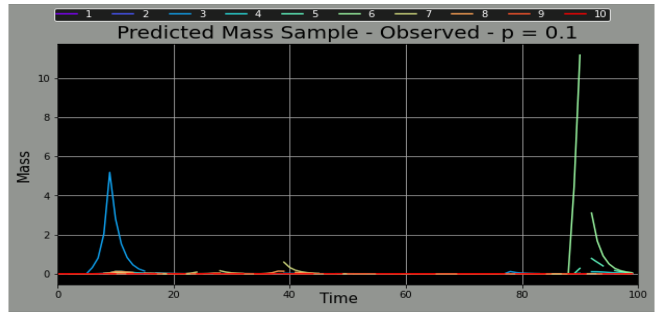
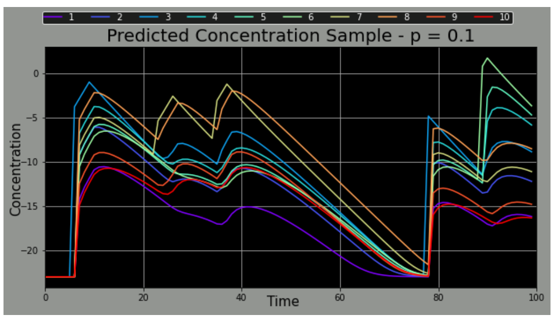
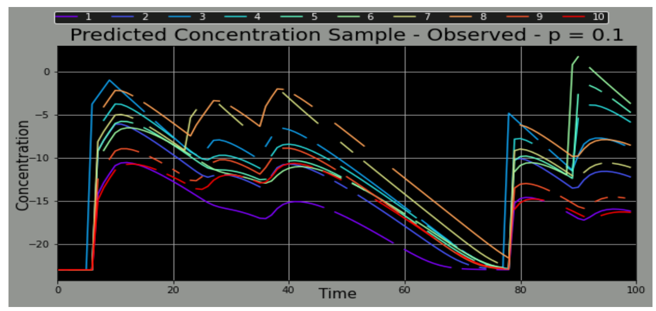
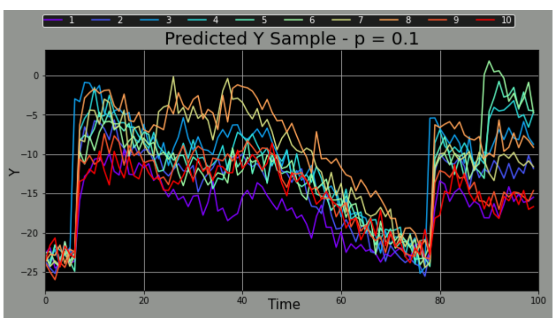
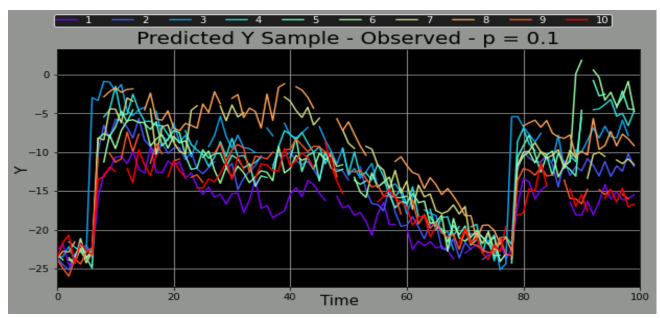
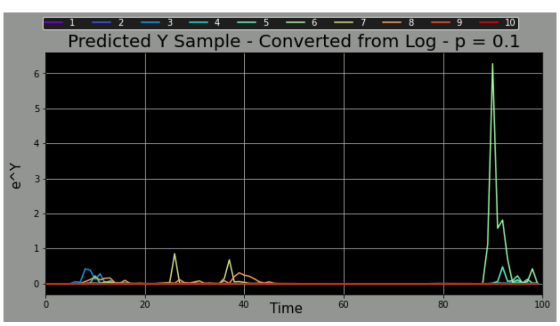
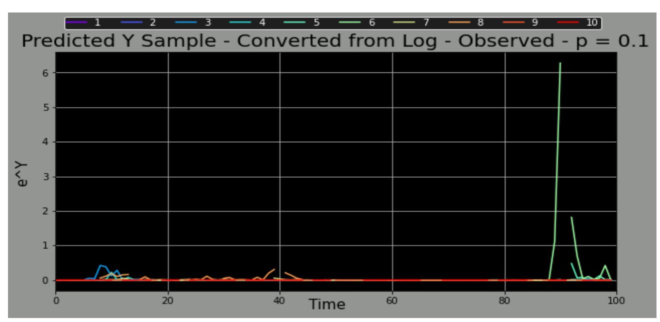
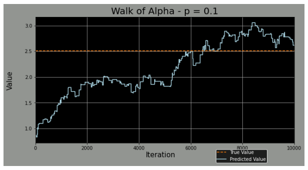
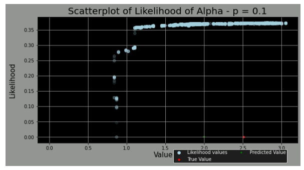
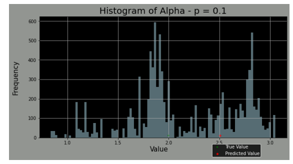
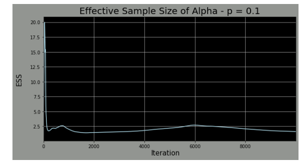
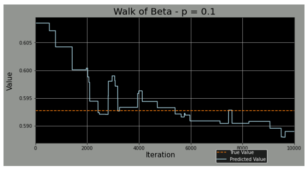
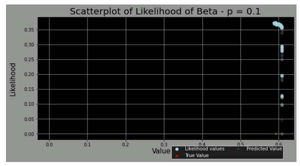
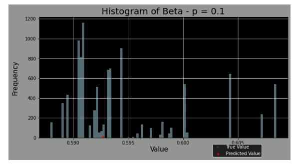
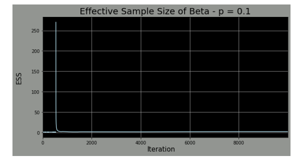
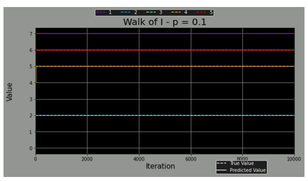
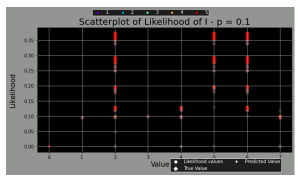
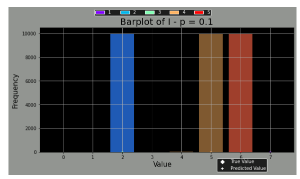
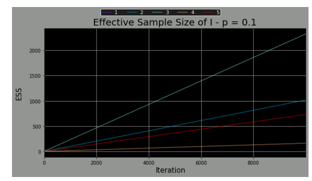
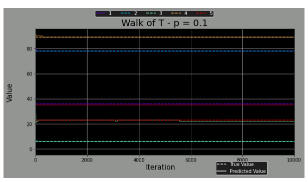
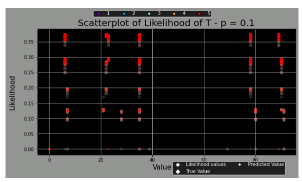
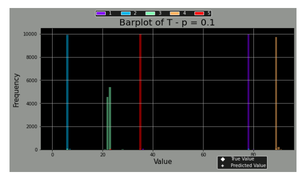
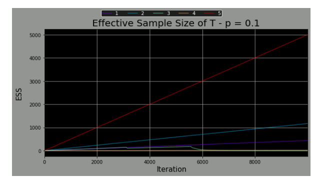
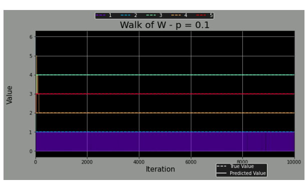
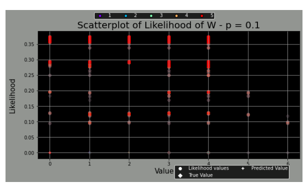
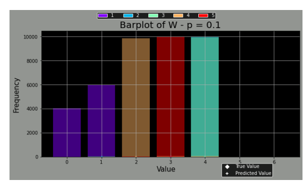
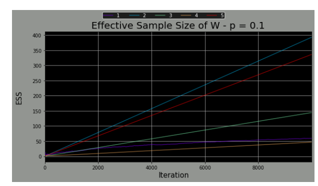
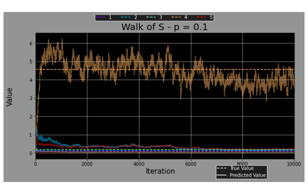
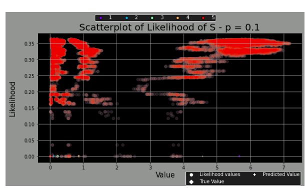
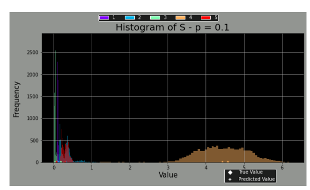
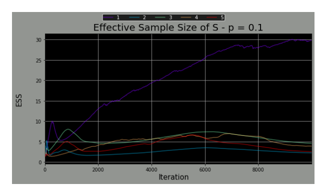
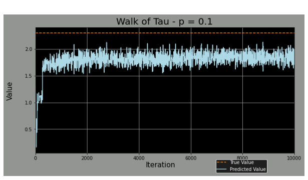
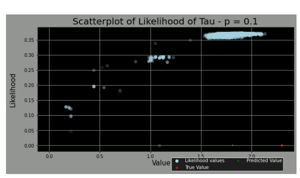
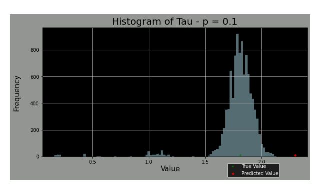
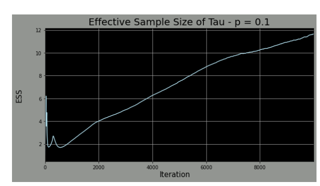
Results	Full	Observed
Mass		
Concentration		
Y		
e^Y		

Figure 7: Table of results for sparse data ($p = 0.1$)

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	1.9981	2.1539	2.1748	0.0231
β	0.5928	0.5945	0.5947	0.0035
I	2, 2, 6, 5, 6	2.0000, 1.9982, 5.9996, 4.9951, 6.0002	2.0000, 2.0002, 6.0016, 4.9977, 6.0016	0.0983
T	78, 6, 23, 89, 35	78.0022, 6.0010, 22.5412, 89.0259, 34.9998	78.0090, 6.0366, 22.5682, 89.0335, 35.0006	0.0104
W	1, 4, 4, 2, 3	0.5871, 4.0008, 3.9872, 2.0162, 2.9974	0.6063, 4.0036, 3.9936, 2.0236, 2.9992	0.2560
S	0.1105, 0.3169, 0.0304, 4.4406, 0.3082	0.1113, 0.3193, 0.0330, 4.4229, 0.3105	0.1120, 0.3260, 0.0342, 4.4496, 0.3140	0.2893
τ	1.8124	1.7884	1.7956	0.2037

$Table\ 3\text{:}\ Table\ of\ summary\ statistics\ of\ parameters\ for\ sparse\ data\ (p\ =\ 0.1)$

Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

$Figure\ 8\!:\ Table\ of\ graphs\ of\ parameter\ metrics\ for\ sparse\ data\ (p\ =\ 0.1)$

p = 0.2

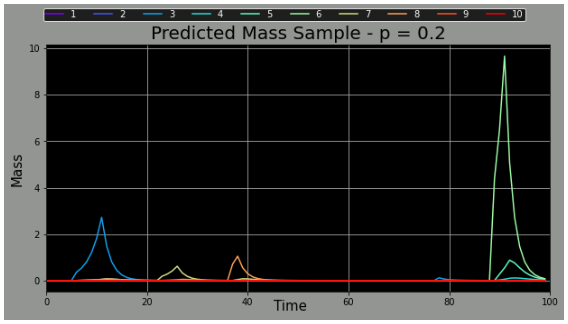
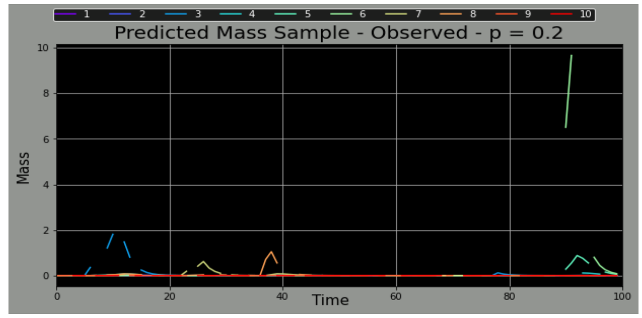
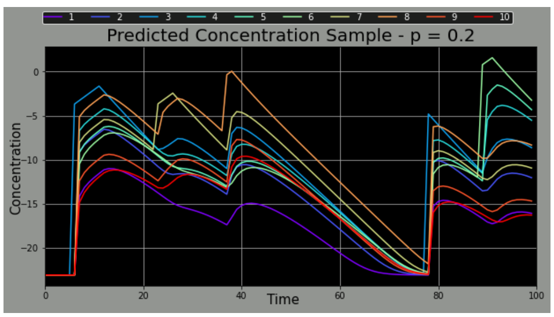
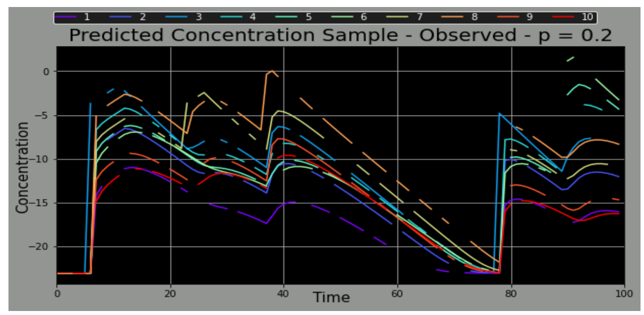
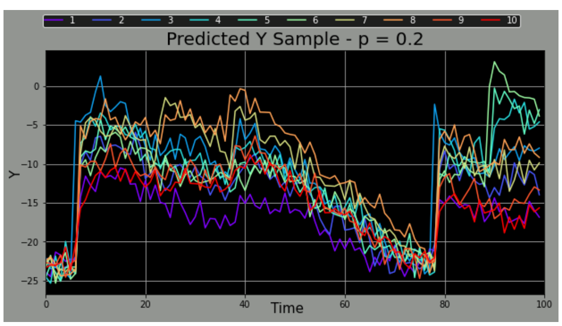
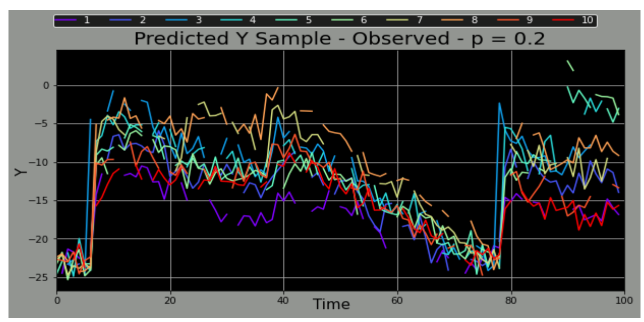
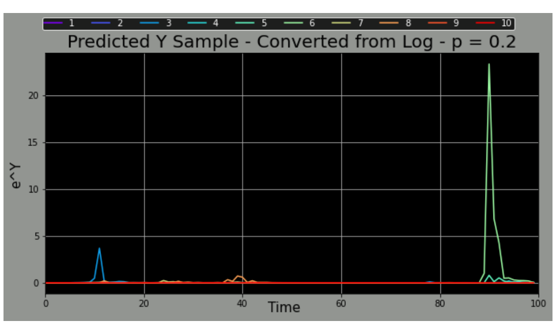
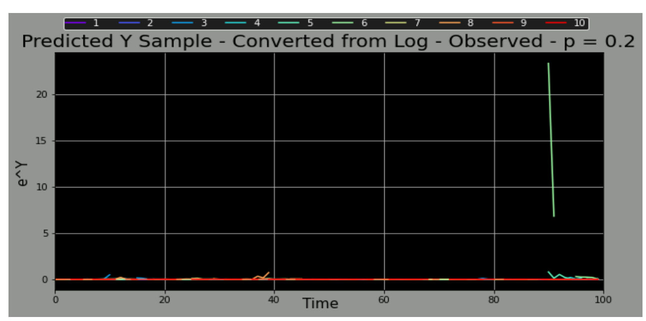
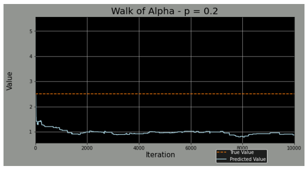
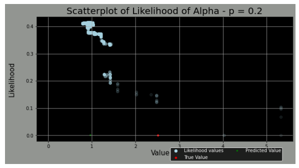
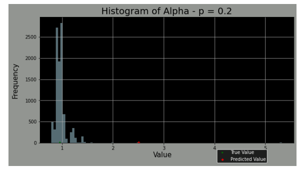
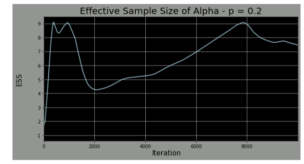
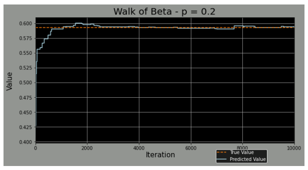
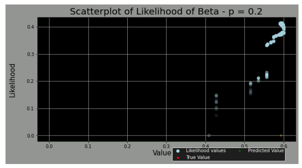
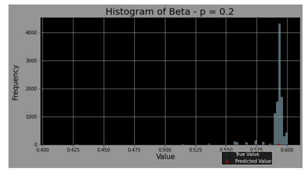
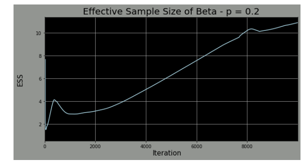
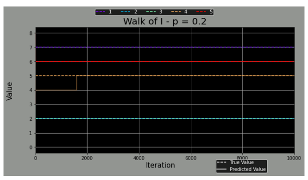
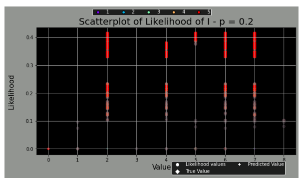
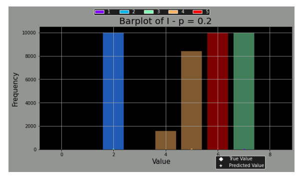
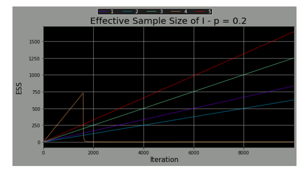
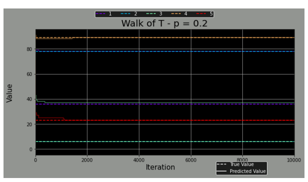
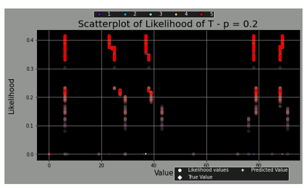
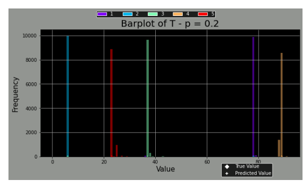
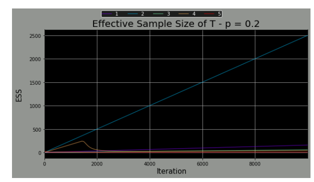
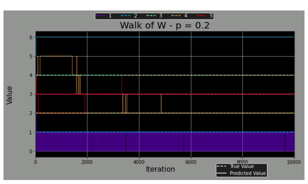
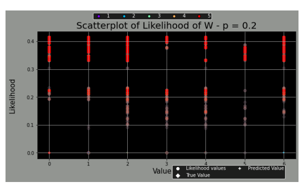
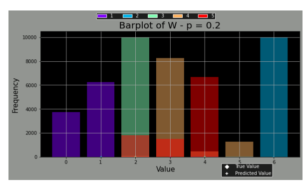
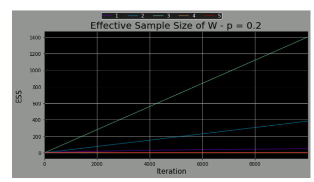
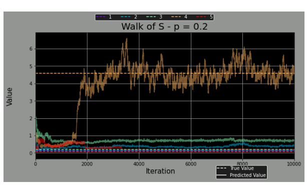
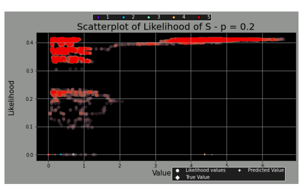
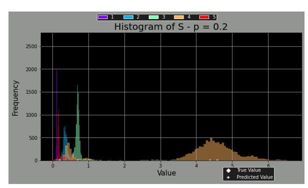
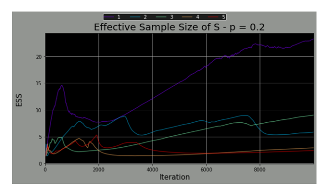
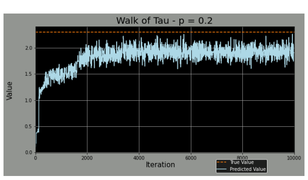
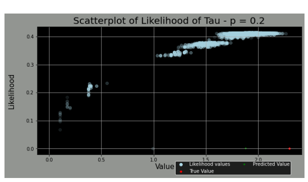
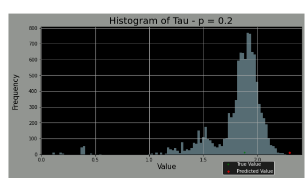
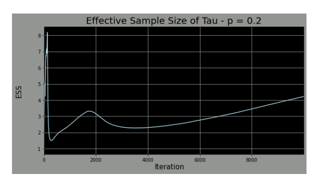
Results	Full	Observed
Mass		
Concentration		
Y		
e^Y		

Figure 9: Table of results for sparse data ($p = 0.2$)

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	0.9612	0.9779	0.9845	0.0116
β	0.5930	0.5914	0.5919	0.0041
I	2, 2, 7, 5, 6	1.9989, 2.0010, 6.9945, 4.8335, 5.9990	1.9999, 2.0054, 6.9999, 4.8479, 6.0016	0.0986
T	78, 6, 37, 89, 23	78.0048, 5.9999, 37.0503, 88.8192, 23.2423	78.0112, 6.0005, 37.0769, 88.8688, 23.2783	0.0114
W	1, 6, 2, 3, 4	0.6229, 5.9957, 1.9990, 3.2738, 3.4724	0.6425, 5.9987, 2.0000, 3.3006, 3.5030	0.2504
S	0.1158, 0.3551, 0.7038, 4.3893, 0.1815	0.1173, 0.3598, 0.7141, 3.8642, 0.2528	0.1180, 0.3628, 0.7184, 3.9254, 0.2586	0.3655
τ	1.8819	1.8138	1.8239	0.2238

$Table\ 4\!:\ Table\ of\ summary\ statistics\ of\ parameters\ for\ sparse\ data\ (p\ =\ 0.2)$

Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

$Figure\ 10:\ Table\ of\ graphs\ of\ parameter\ metrics\ for\ sparse\ data\ (p\ =\ 0.2)$

p = 0.3

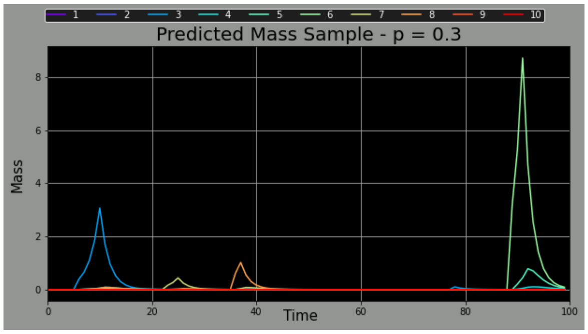
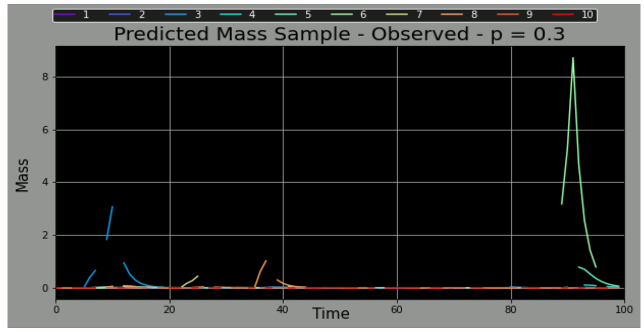
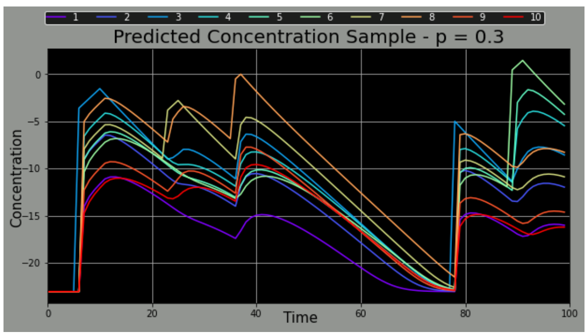
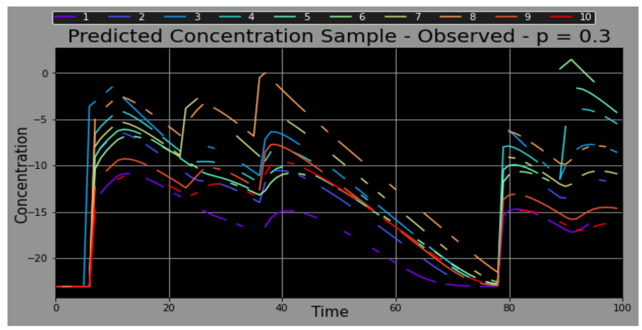
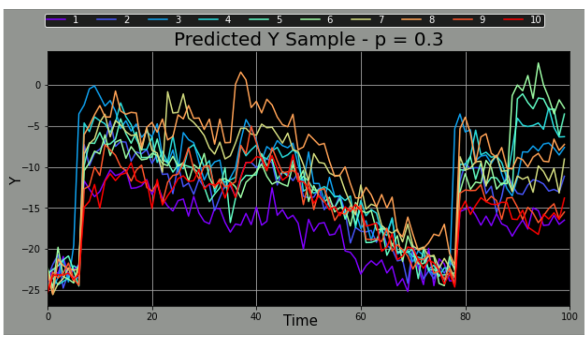
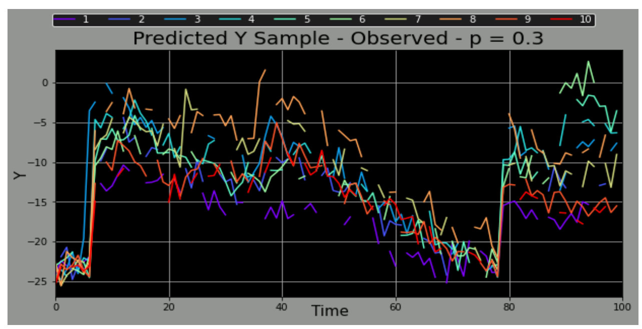
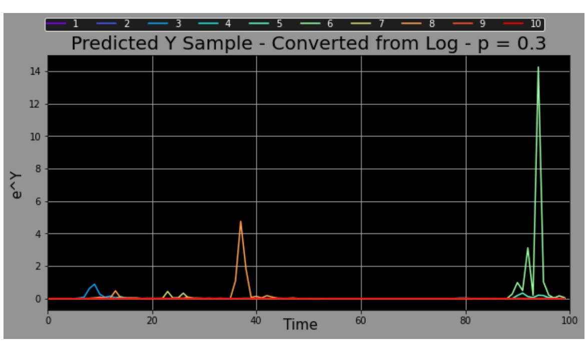
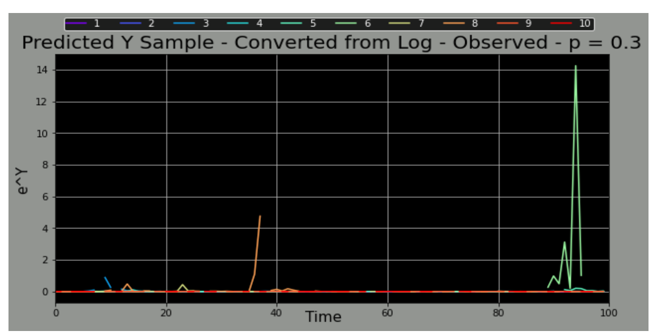
Results	Full	Observed
Mass		
Concentration		
Y		
e^Y		

Figure 11: Table of results for sparse data $(p = 0.3)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	1.1220	1.0334	1.0554	0.0308
β	0.6066	0.6186	0.6197	0.0051
I	7, 2, 2, 6, 5	6.5658, 2.0073, 2.0124, 6.1077, 4.9820	6.6072, 2.0111, 2.0230, 6.1207, 4.9896	0.1020
T	36, 78, 6, 23, 89	45.0593, 77.9936, 6.0006, 23.5989, 88.8422	45.8835, 78.0186, 6.0054, 23.6667, 88.9272	0.0126
W	2, 1, 5, 3, 3	1.9260, 0.6079, 5.1969, 4.2488, 2.8978	1.9526, 0.6277, 5.2509, 4.3974, 2.9188	0.2831
S	0.6126, 0.0963, 0.3890, 0.1553, 3.1741	0.5952, 0.0944, 0.6398, 0.2186, 3.3662	0.6024, 0.0973, 0.6698, 0.2287, 3.4120	0.3986
τ	1.9221	1.7005	1.7215	0.2167

$Table\ 5$: $Table\ of\ summary\ statistics\ of\ parameters\ for\ sparse\ data\ (p\ =\ 0.3)$

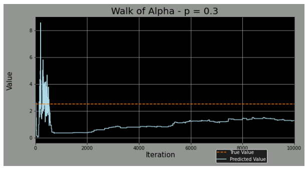
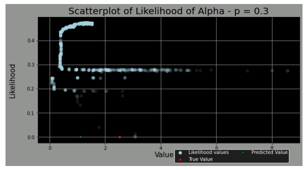
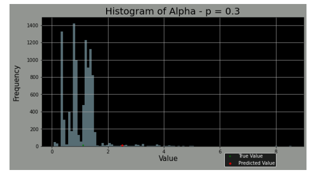
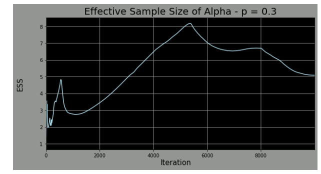
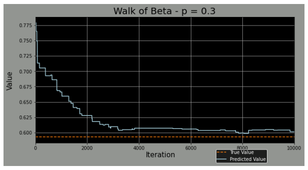
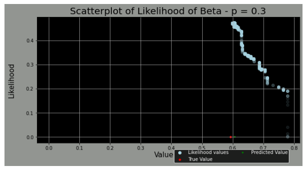
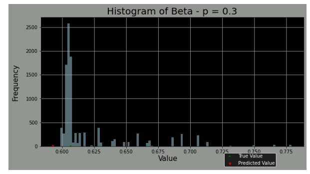
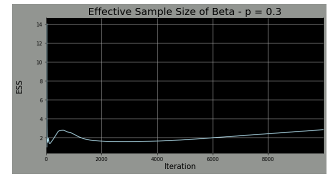
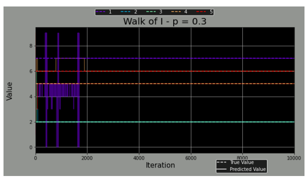
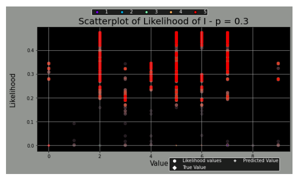
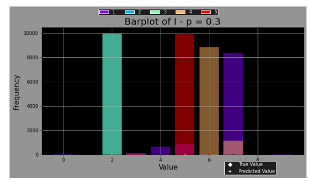
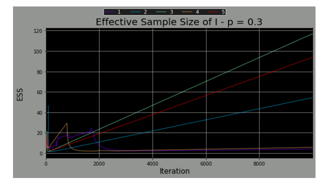
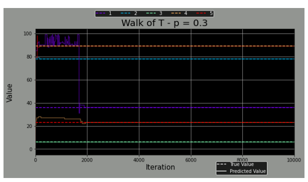
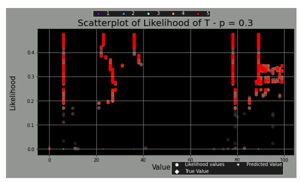
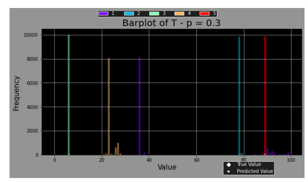
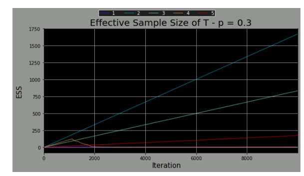
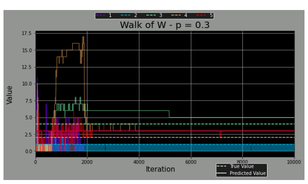
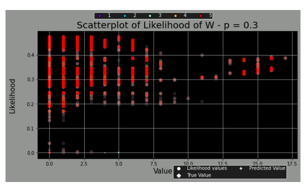
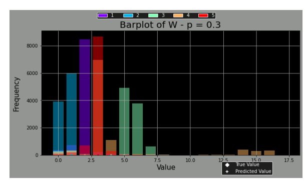
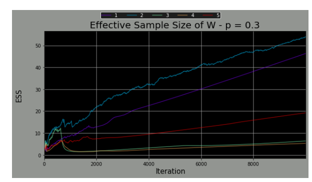
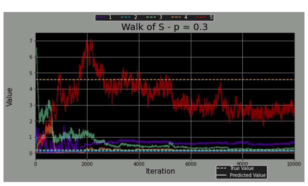
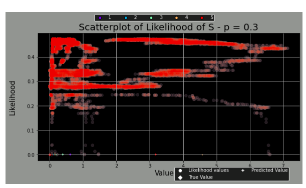
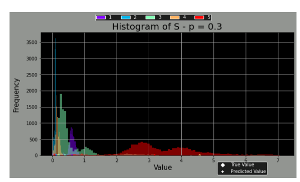
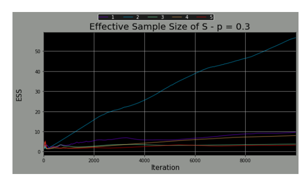
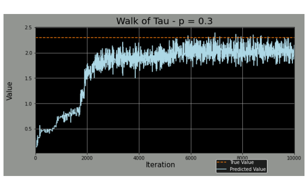
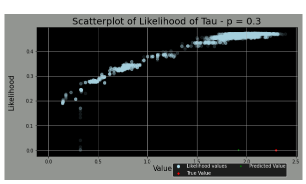
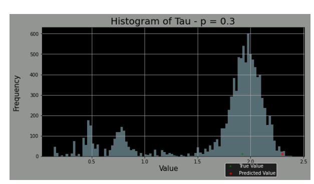
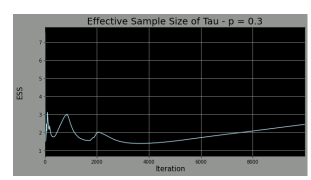
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 12: Table of graphs of parameter metrics for sparse data ($p = 0.3$)

p = 0.4

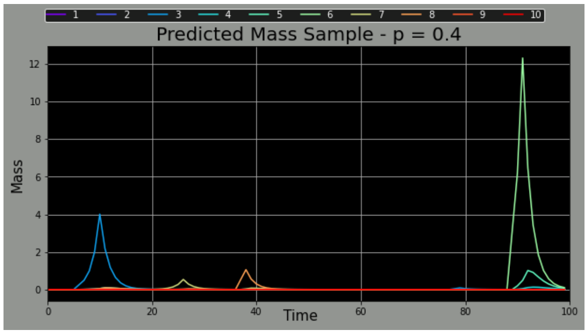
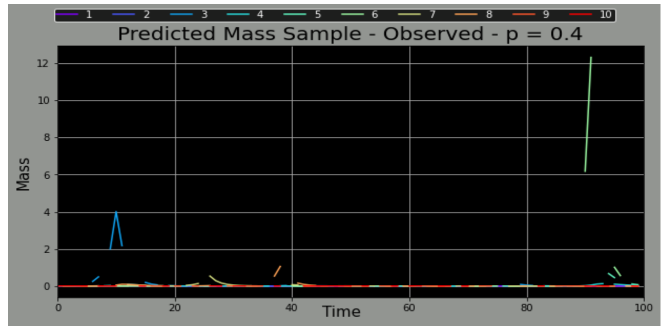
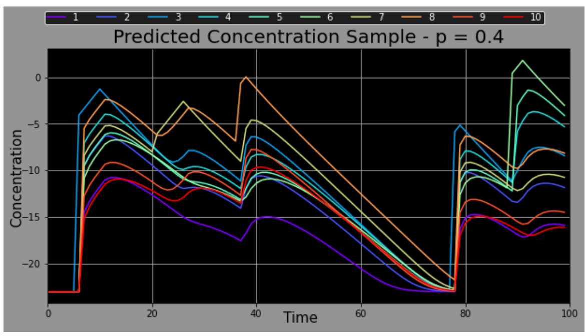
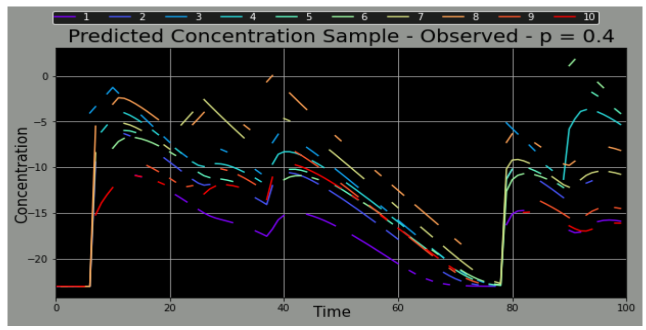
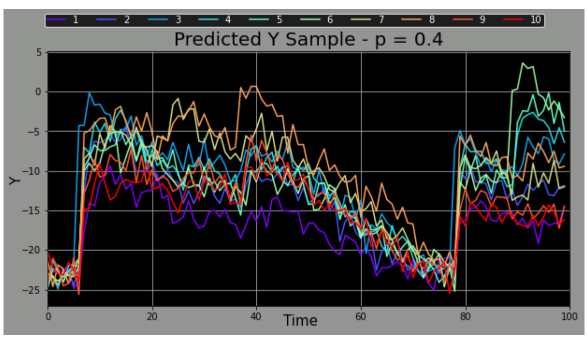
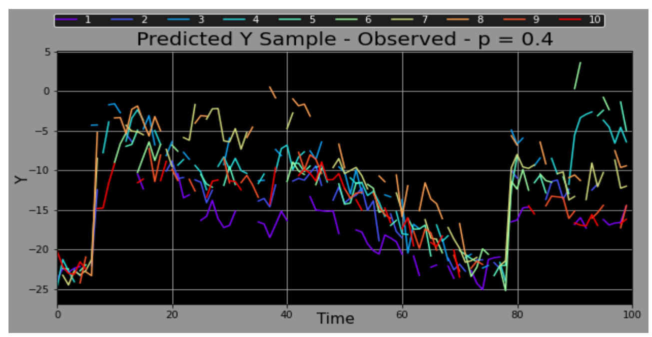
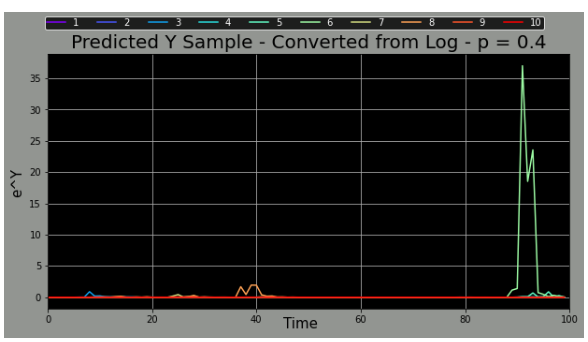
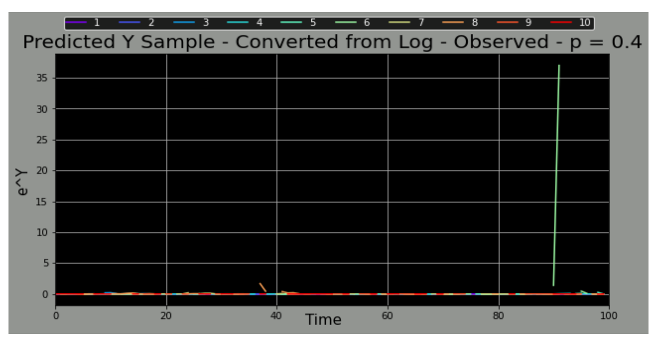
Results	Full	Observed
Mass		
Concentration		
Y		
e^Y		

Figure 13: Table of results for sparse data $(p = 0.4)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	1.4692	1.5017	1.5112	0.0132
β	0.5930	0.5905	0.5910	0.0069
I	5, 2, 2, 7, 6	4.9988, 1.9888, 2.0016, 6.9836, 6.0000	5.0004, 1.9930, 2.0070, 6.9912, 6.0012	0.1010
T	89, 78, 6, 37, 21	89.0067, 77.9842, 5.9970, 37.3660, 21.2626	89.0103, 77.9888, 6.0092, 37.4374, 21.3140	0.0110
W	3, 2, 5, 2, 6	2.5839, 2.0145, 4.9997, 1.9165, 6.0040	2.6031, 2.0197, 5.0009, 1.9301, 6.0118	0.2040
S	3.1055, 0.0412, 0.2455, 0.5287, 0.0115	3.7205, 0.0478, 0.2508, 0.5208, 0.0125	3.7722, 0.0524, 0.2526, 0.5253, 0.0127	0.3015
τ	1.8791	1.8134	1.8264	0.2573

$Table\ 6\text{:}\ Table\ of\ summary\ statistics\ of\ parameters\ for\ sparse\ data\ (p\ =\ 0.4)$

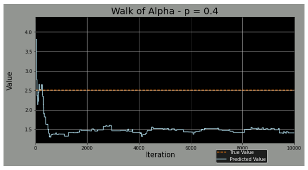
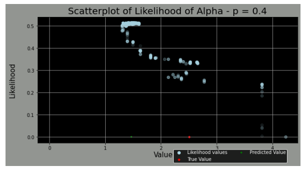
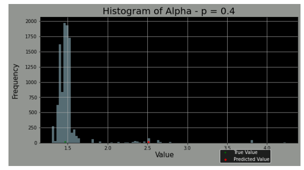
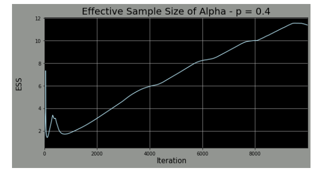
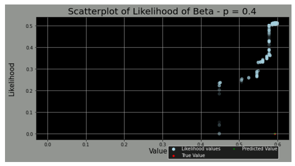
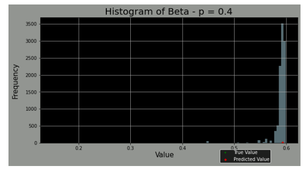
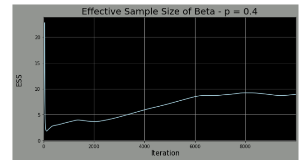
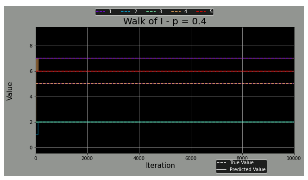
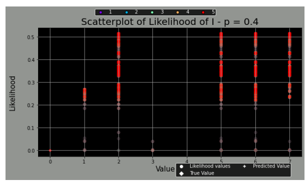
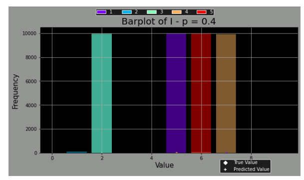
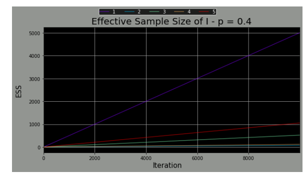
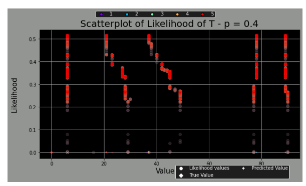
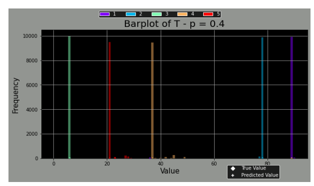
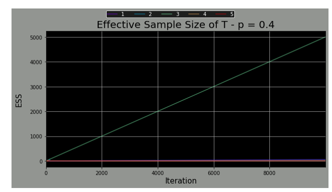
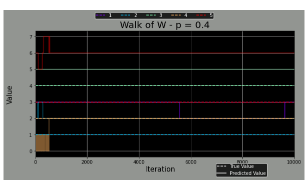
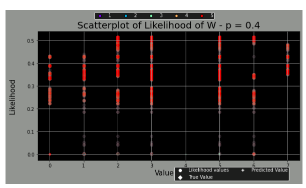
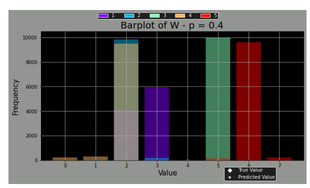
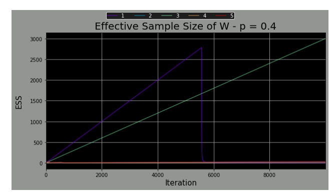
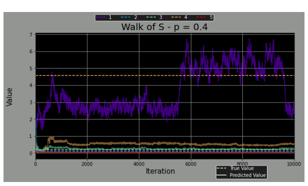
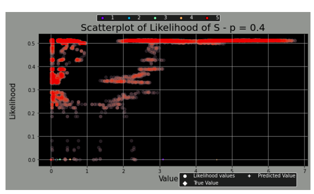
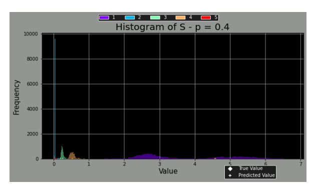
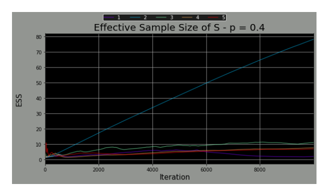
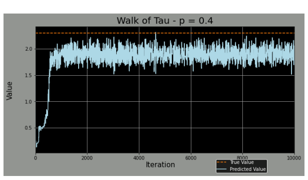
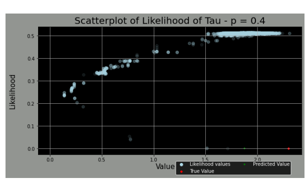
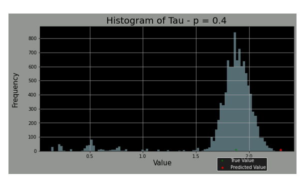
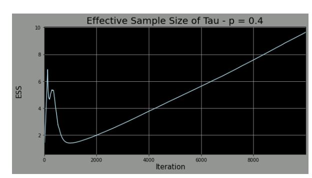
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 14: Table of graphs of parameter metrics for sparse data ($p = 0.4$)

p = 0.5

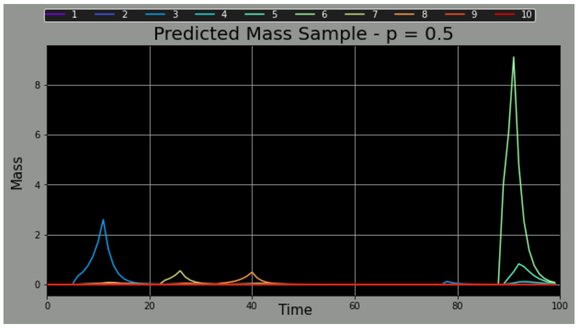
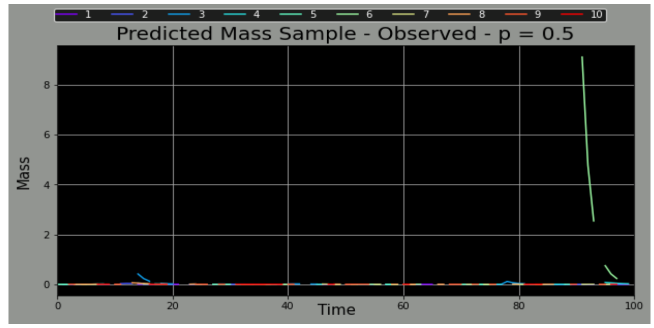
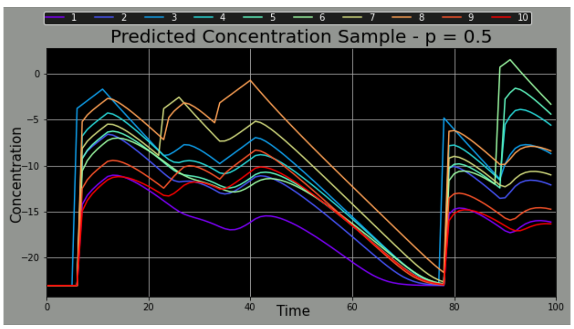
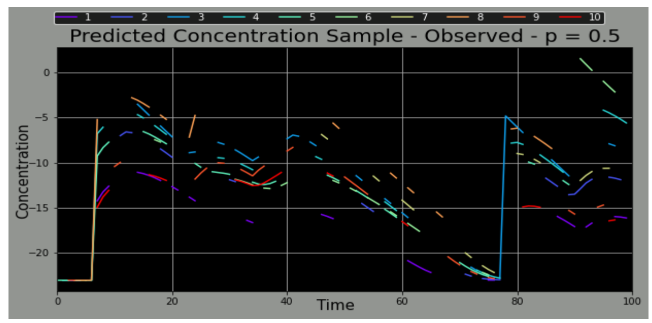
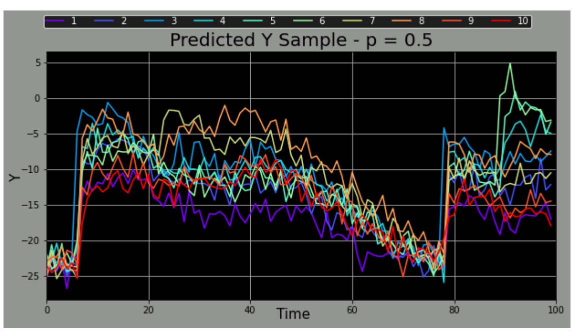
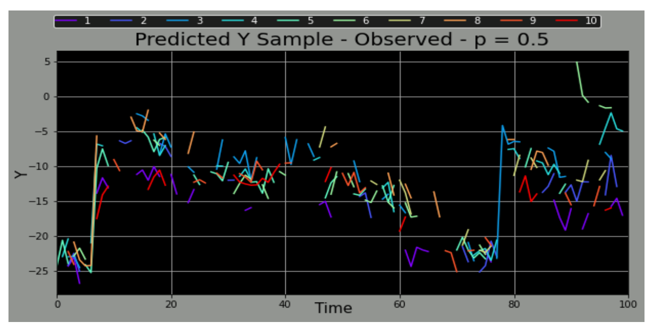
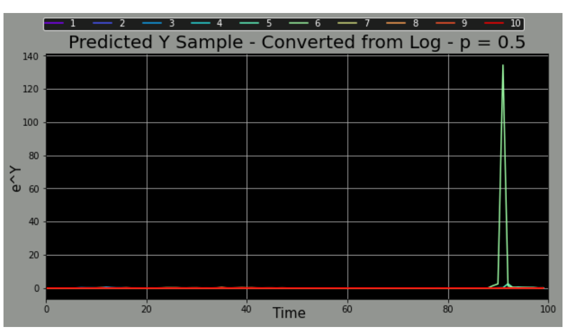
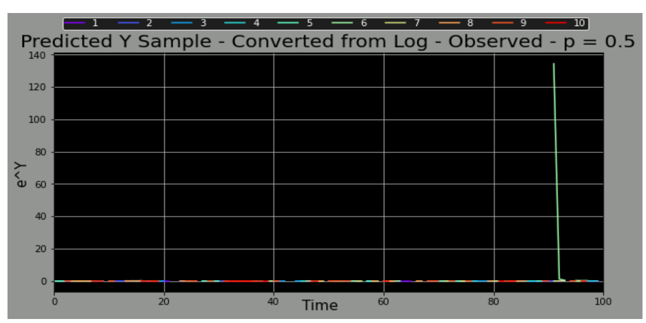
Results	Full	Observed
Mass		
Concentration		
Y		
e^Y		

Figure 15: Table of results for sparse data $(p = 0.5)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	0.9795	0.9105	0.9166	0.0090
β	0.5913	0.5975	0.5982	0.0057
I	5, 2, 6, 2, 7	4.9962, 1.9992, 5.9985, 1.9964, 6.9507	4,9992, 2.0016, 6.0053, 2.0070, 6.9593	0.1009
T	89, 6, 23, 78, 34	88.9992, 6.0445, 23.9136, 76.5947, 33.7967	89.0098, 6.0613, 24.2102, 76.9627, 33.8229	0.0111
W	3, 6, 4, 1, 7	3.0275, 5.8966, 3.6371, 0.6160, 6.9002	3.0401, 5.9188, 3.6655, 0.6354, 6.9186	0.2432
S	4.0505, 0.3204, 0.1551, 0.1131, 0.0342	4.0011, 0.3780, 0.1864, 0.1291, 0.0447	4.0279, 0.3878, 0.1925, 0.1379, 0.0465	0.3342
τ	1,9400	1.87576	1.8882	0.2907

$Table\ 7:\ Table\ of\ summary\ statistics\ of\ parameters\ for\ sparse\ data\ (p\ =\ 0.5)$

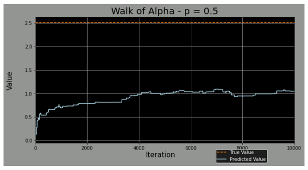
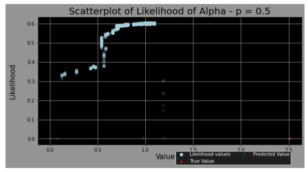
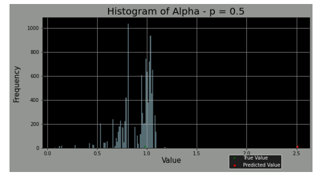
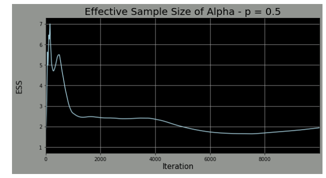
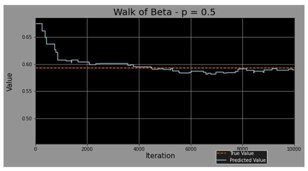
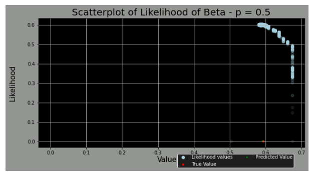
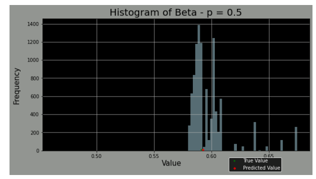
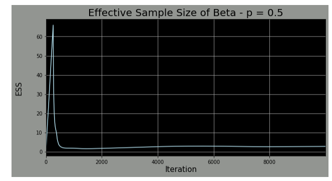
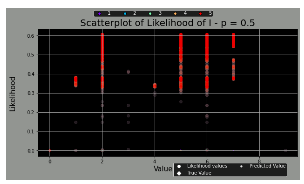
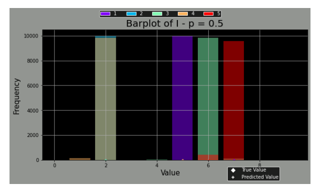
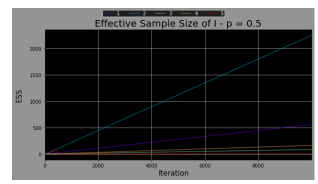
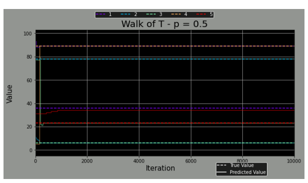
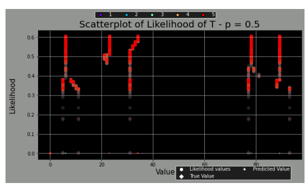
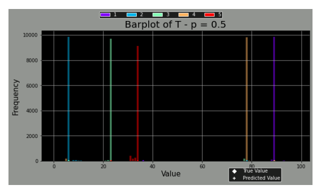
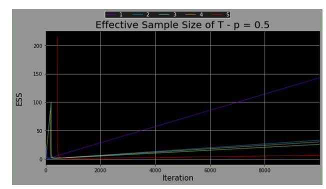
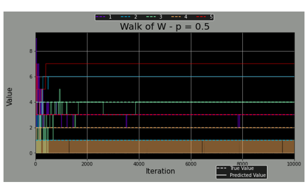
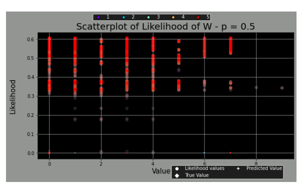
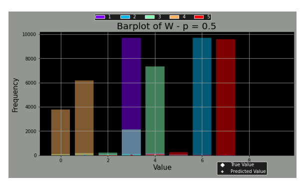
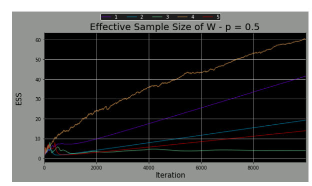
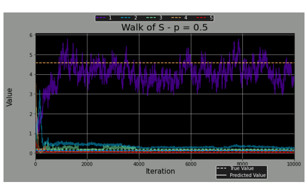
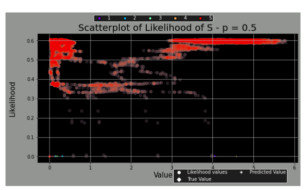
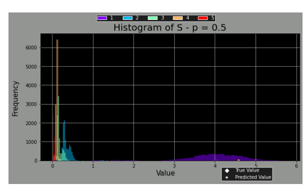
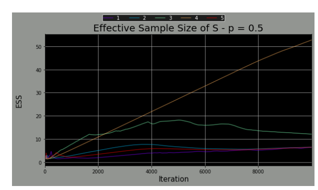
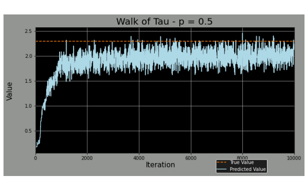
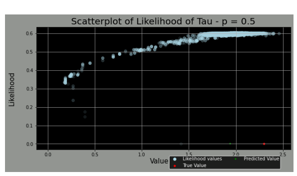
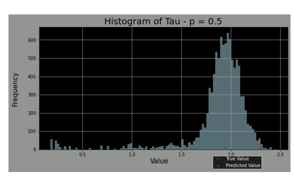
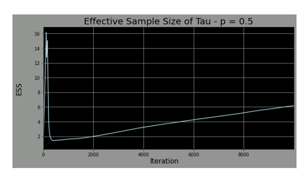
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 16: Table of graphs of parameter metrics for sparse data ($p = 0.5$)

p = 0.6

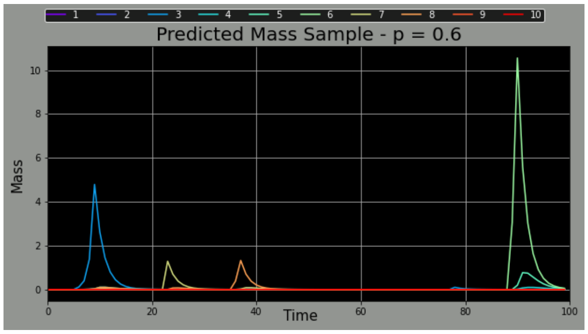
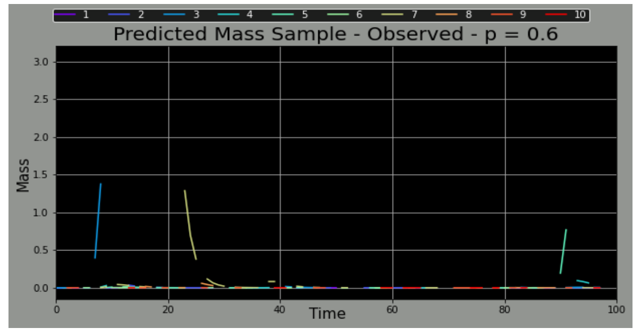
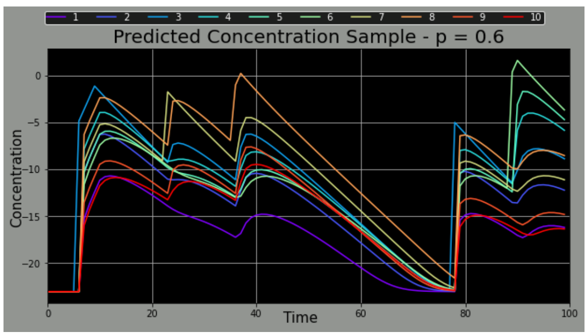
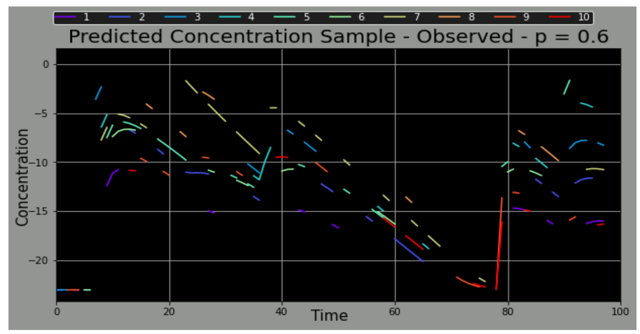
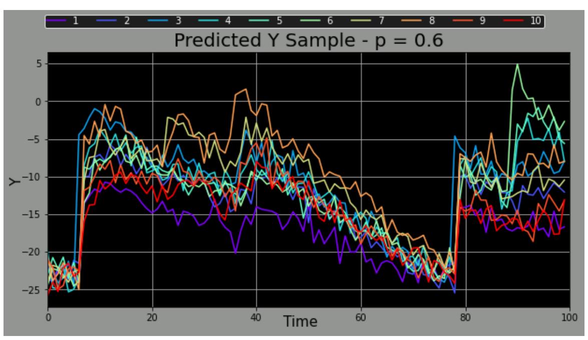
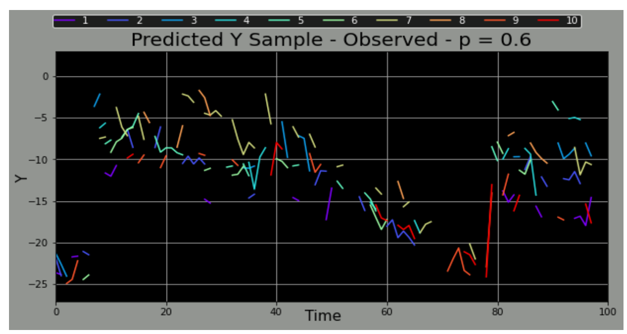
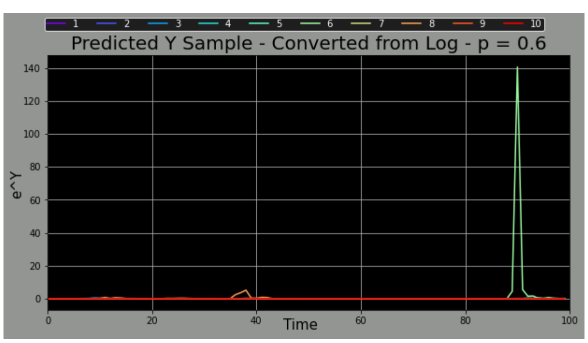
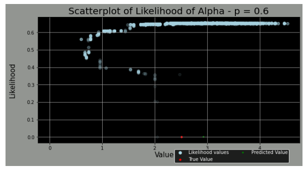
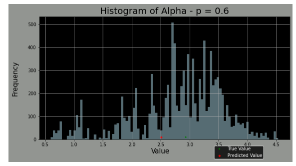
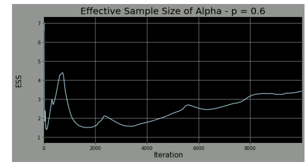
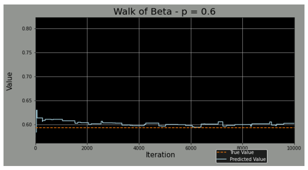
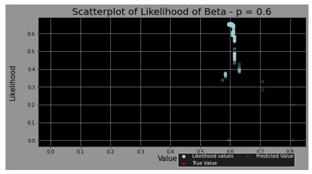
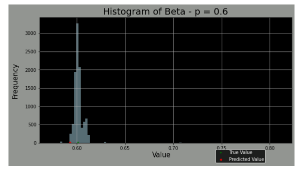
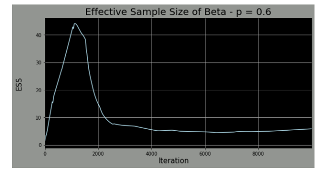
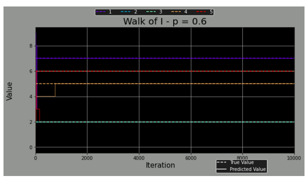
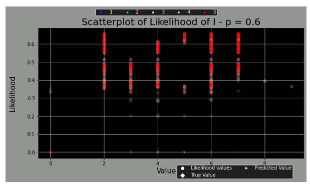
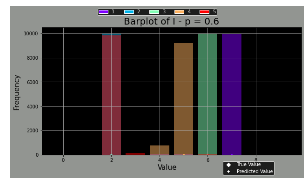
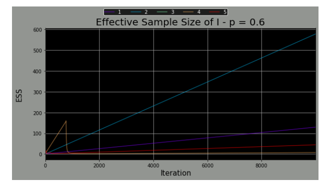
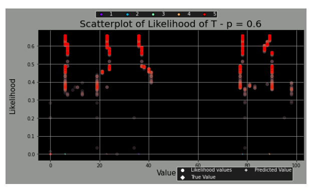
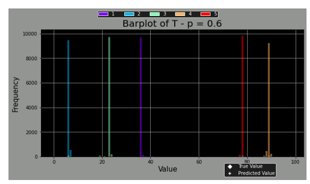
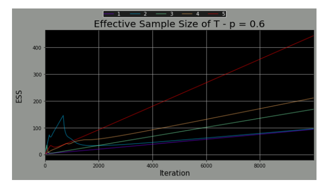
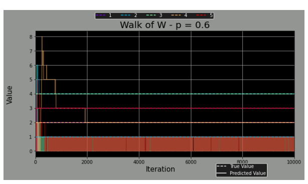
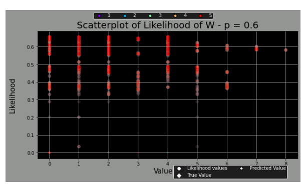
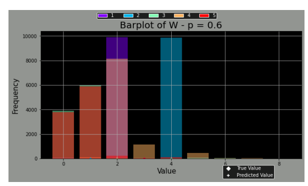
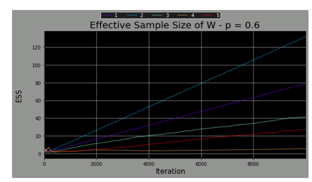
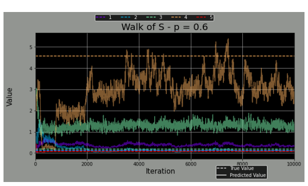
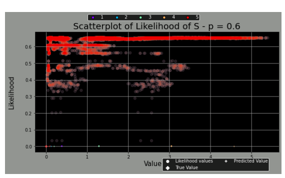
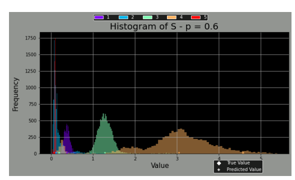
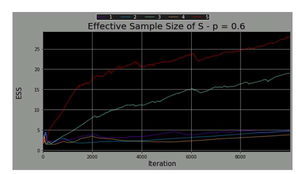
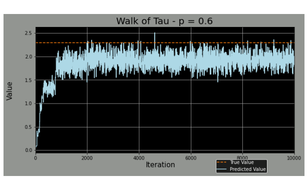
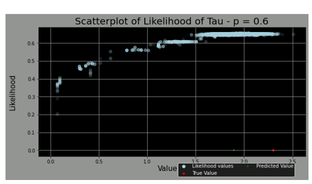
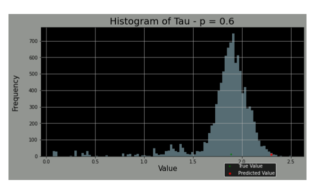
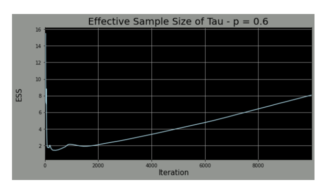
<table>
<tr><th>Results</th><th>Full</th><th>Observed</th></tr>
<tr><td>Mass</td><td>
</td><td>
</td></tr>
<tr><td>Concentration</td><td>
</td><td>
</td></tr>
<tr><td>Y</td><td>
</td><td>
</td></tr>
<tr><td>e^Y</td><td>
</td><td>
</td></tr>
</table>

Figure 17: Table of results for sparse data $(p = 0.6)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	2.9302	2.8023	2.8321	0.0626
β	0.6012	0.6016	0.6018	0.0055
I	7, 2, 6, 5, 2	6.9840, 2.0010, 6.0000, 4.9167, 2.0164	6.9916, 2.0052, 6.0000, 4.9277, 2.0238	0.1023
T	36, 6, 23, 89, 78	36.2009, 6.0428, 22.9796, 88.8610, 75.9589	36.3473, 6.0730, 23.0010, 88.9418, 77.9891	0.0113
W	2, 4, 2, 2, 1	1.9837, 4.0002, 0.6339, 2.2580, 0.6684	1.9901, 4.0060, 0.6591, 0.2900, 0.6932	0.3231
S	0.3788, 0.1798, 1.2839, 3.0497, 0.0984	0.3811, 0.1722, 1.2973, 2.9268, 0.0979	0.3841, 0.1798, 1.3055, 2.9642, 0.0987	0.4555
τ	1.8925	1.8373	1.8488	0.3035

$Table\ 8\text{: } Table\ of\ summary\ statistics\ of\ parameters\ for\ sparse\ data\ (p\ =\ 0.0.6)$

Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

$Figure\ 18\!:\ Table\ of\ graphs\ of\ parameter\ metrics\ for\ sparse\ data\ (p\ =\ 0.6)$

p = 0.7

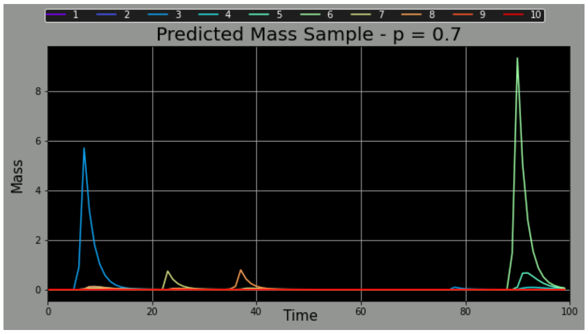
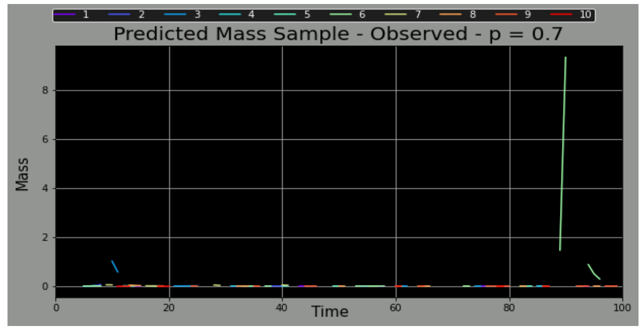
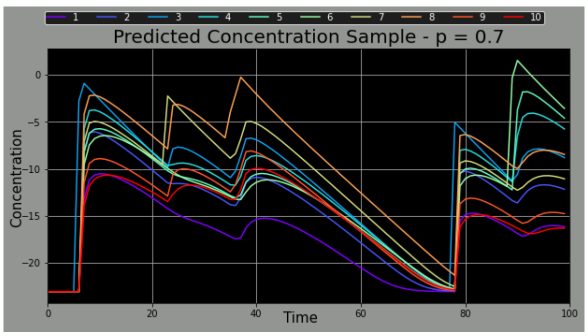
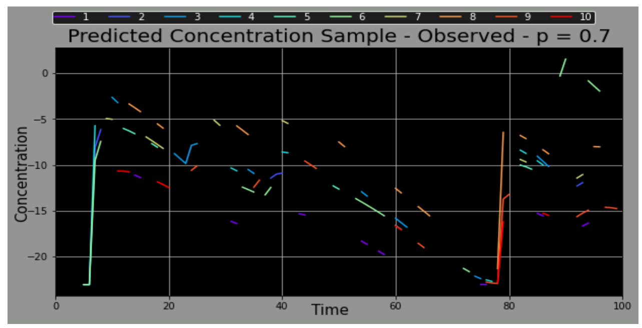
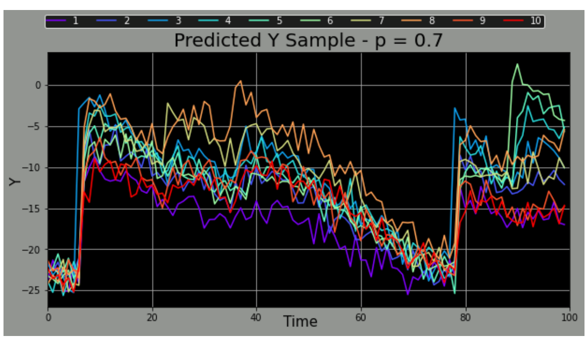
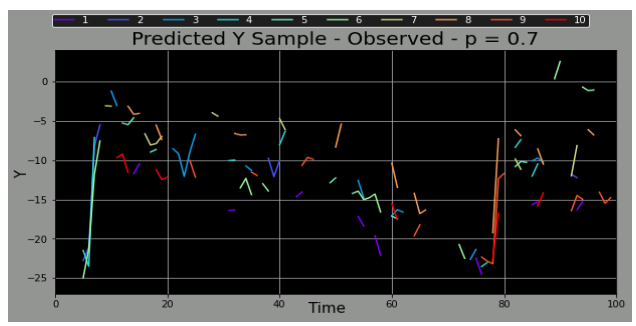
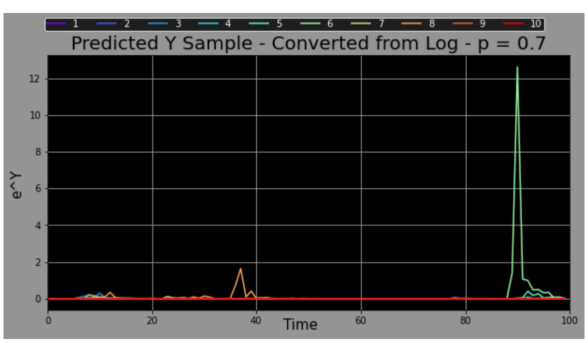
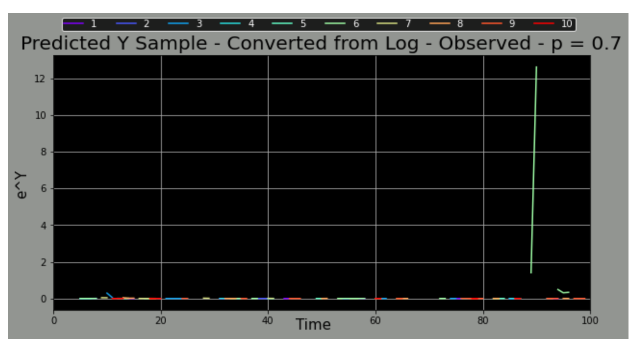
<table>
<tr><th>Results</th><th>Full</th><th>Observed</th></tr>
<tr><td>Mass</td><td>
</td><td>
</td></tr>
<tr><td>Concentration</td><td>
</td><td>
</td></tr>
<tr><td>Y</td><td>
</td><td>
</td></tr>
<tr><td>e^Y</td><td>
</td><td>
</td></tr>
</table>

Figure 19: Table of results for sparse data ($p = 0.7$)

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	5.8331	5.8486	5.9012	0.1356
β	0.6144	0.6195	0.6201	0.0069
I	7, 5, 6, 2, 2	6.8926, 4.9892, 5.6198, 2.1093, 2.0777	6.9048, 4.9938, 5.6646, 2.0289, 2.1025	0.0989
T	35, 89, 23, 78, 6	34.5664, 88.5011, 21.8086, 77.9937, 5.9588	34.6150, 88.5273, 21.9730, 78.0237, 5.9694	0.0114
W	3, 2, 1, 1, 2	2.9990, 2.0041, 0.5962, 0.6119, 1.9331	3.0000, 2.0083, 0.6158, 0.6318, 1.9505	0.3271
S	0.0114, 1.4664, 0.7401, 0.0907, 0.8937	0.0153, 1.4754, 0.9304, 0.0937, 0.8774	0.0161, 1.4892, 0.9598, 0.0966, 0.8859	0.4809
τ	1.4532	1.3673	1.3799	0.2582

Table 9: Table of summary statistics of parameters for sparse data $(p = 0.7)$

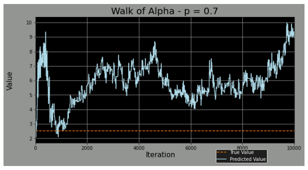
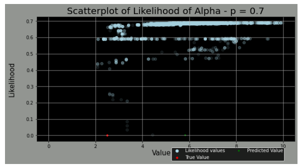
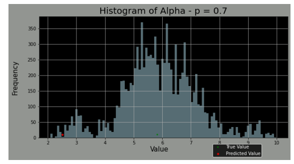
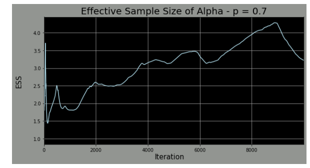
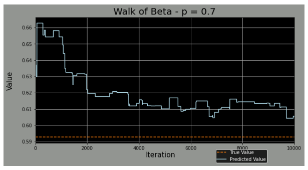
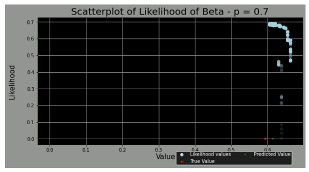
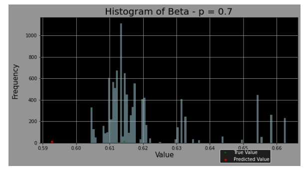
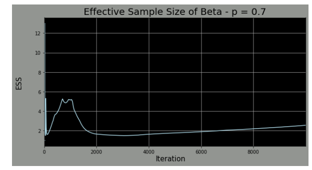
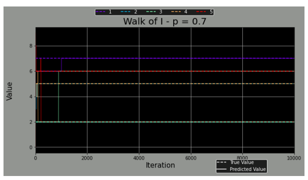
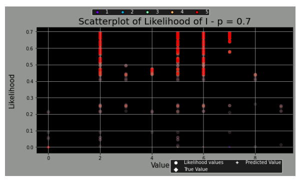
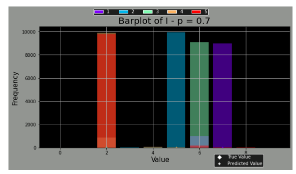
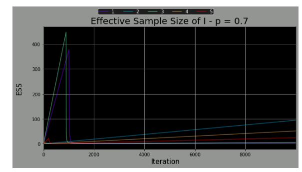
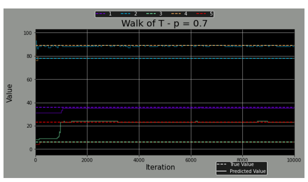
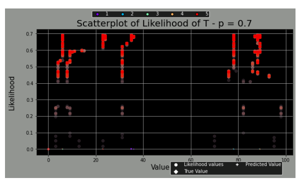
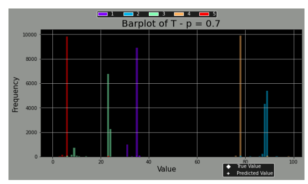
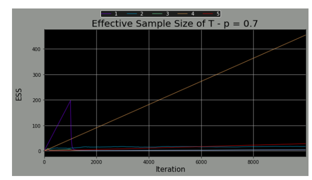
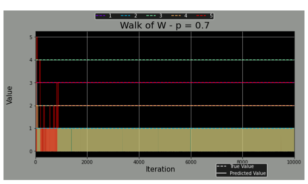
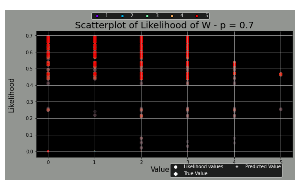
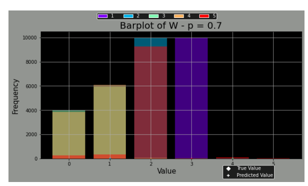
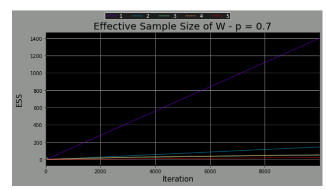
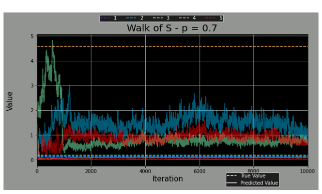
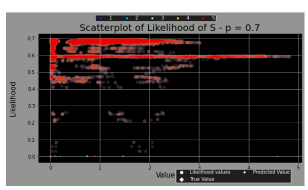
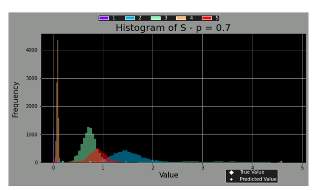
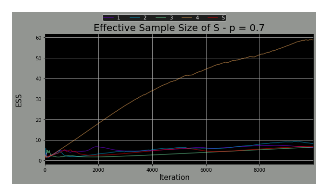
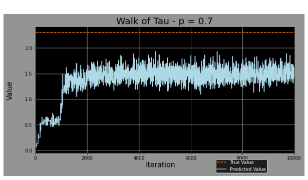
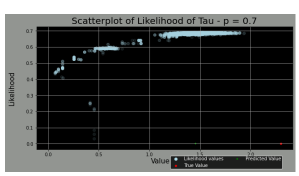
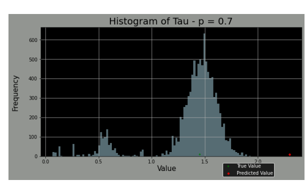
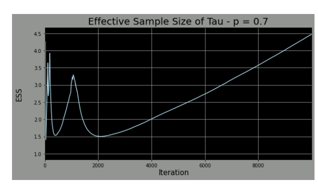
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 20: Table of graphs of parameter metrics for sparse data ($p = 0.7$)

p = 0.8

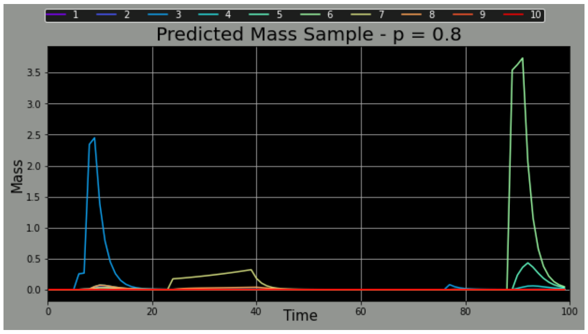
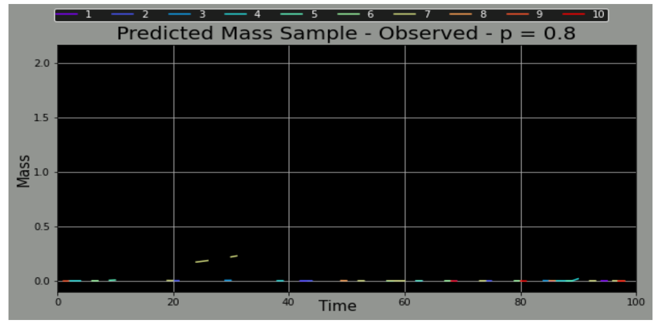
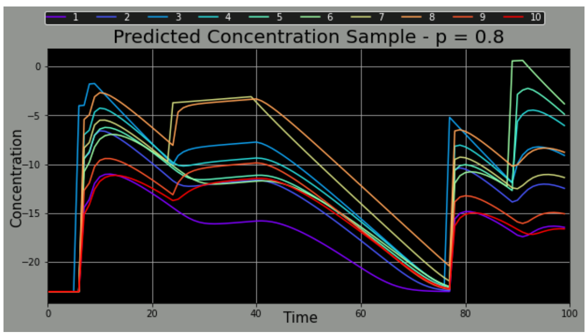
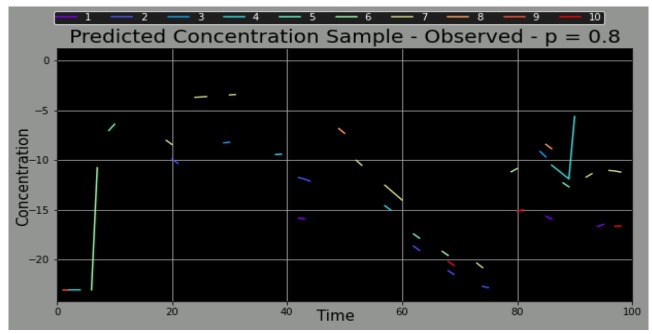
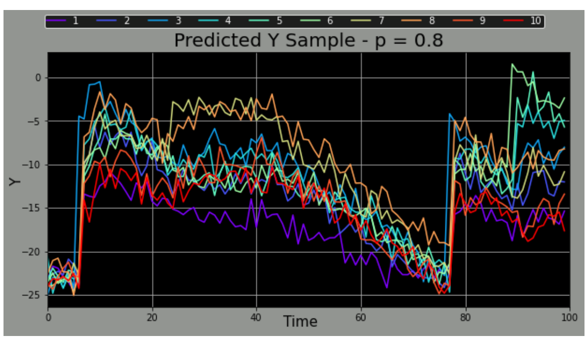
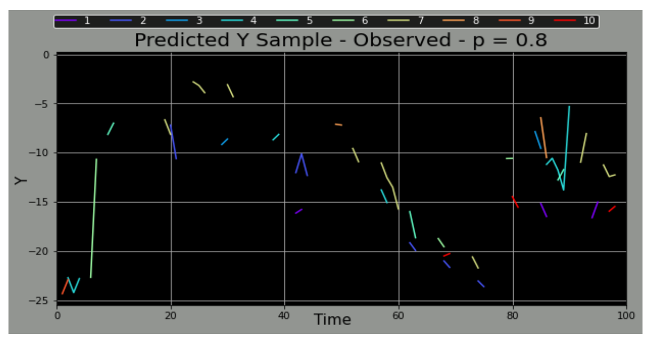
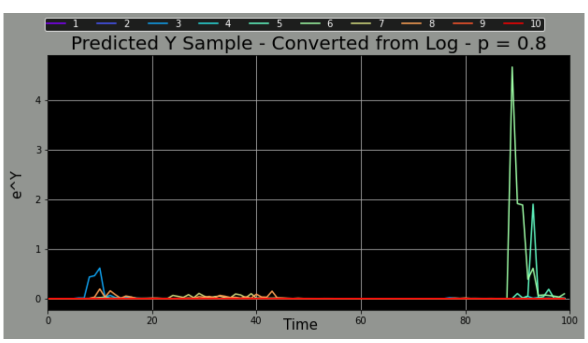
Results	Full	Observed
Mass		
Concentration		
Y		
e^Y		

Figure 21: Table of results for sparse data ($p = 0.8$)

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	0.4803	0.5271	0.5376	0.0120
β	0.6191	0.6404	0.6422	0.0073
I	2, 6, 2, 5, 2	2.3986, 5.7750, 2.0227, 4.9869, 2.0960	2.4994, 5.8100, 2.0321, 4.9989, 2.1180	0.1076
T	8, 24, 77, 89, 6	13.1303, 23.4424, 77.4779, 89.2314, 9.2244	13.9725, 23.7160, 77.5077, 89.2904, 9.8938	0.0143
W	2, 16, 3, 1, 2	2.2571, 13.9046, 1.1200, 2.6004, 2.3640	2.3075, 14.0776, 1.1568, 2.6440, 2.4250	0.3856
S	2.0667, 0.1711, 0.0778, 3.5430, 0.2930	2.1208, 0.1854, 0.0755, 3.4210, 0.2961	2.1607, 0.1890, 0.0770, 3.4731, 0.2997	0.5786
τ	1.3520	1.2440	1.2582	0.2898

$Table\ 10\text{:}\ Table\ of\ summary\ statistics\ of\ parameters\ for\ sparse\ data\ (p\ =\ 0.8)$

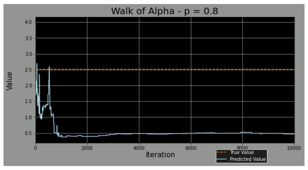
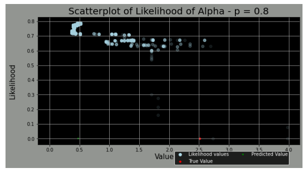
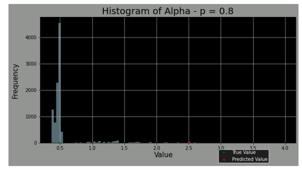
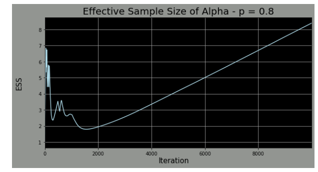
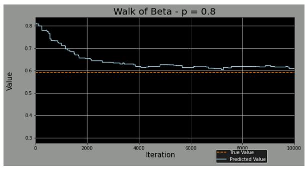
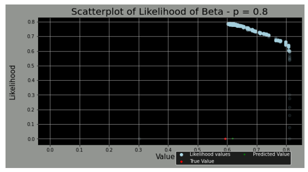
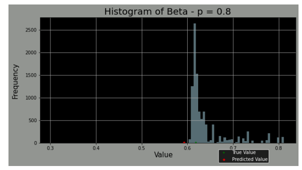
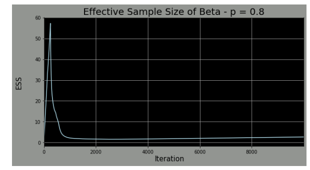
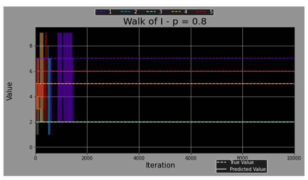
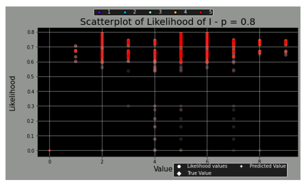
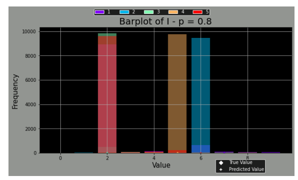
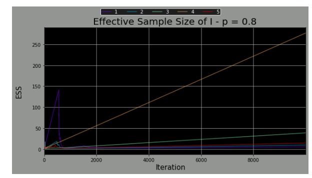
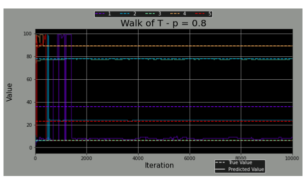
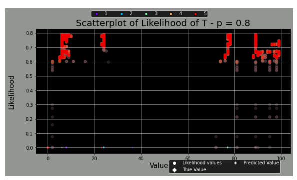
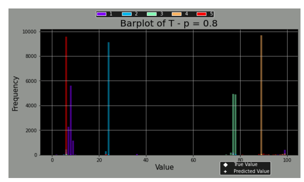
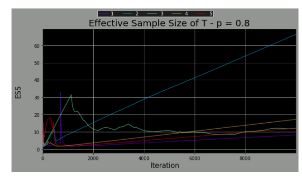
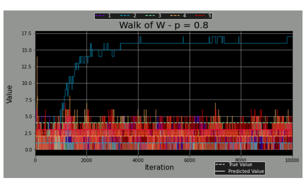
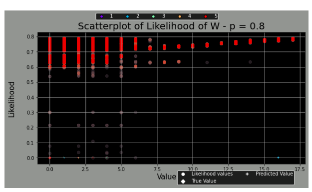
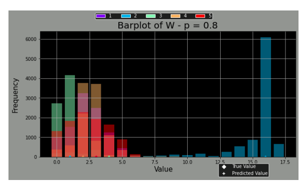
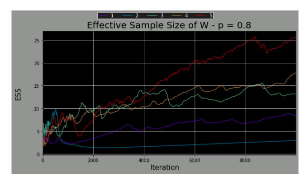
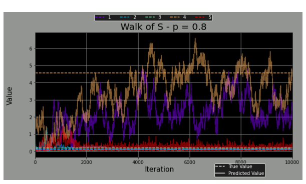
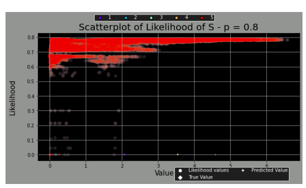
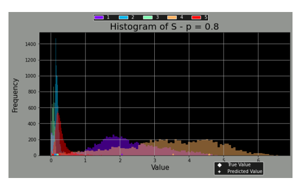
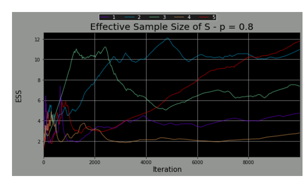
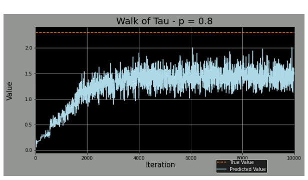
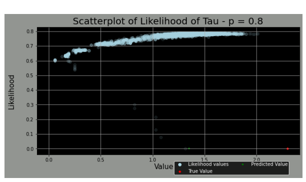
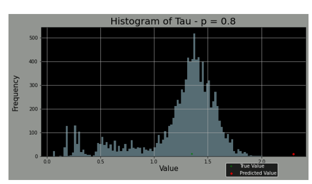
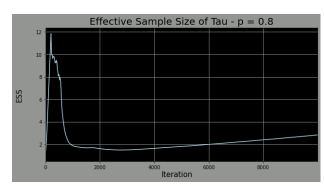
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 22: Table of graphs of parameter metrics for sparse data ($p = 0.8$)

p = 0.9

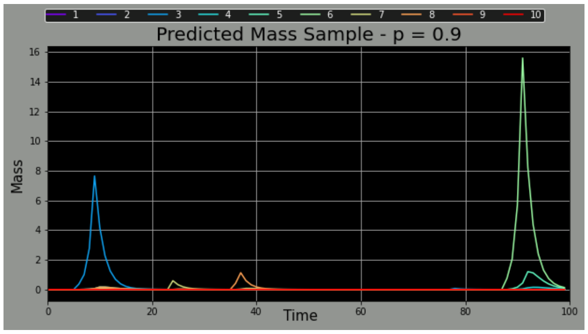
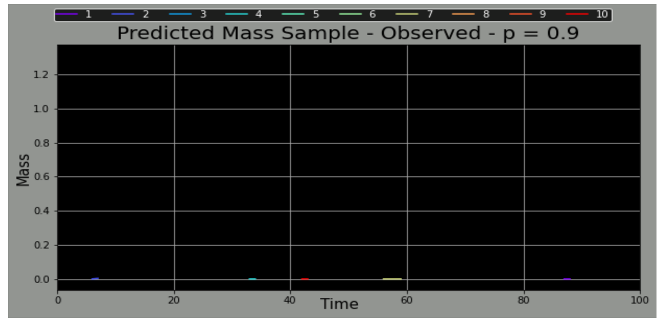
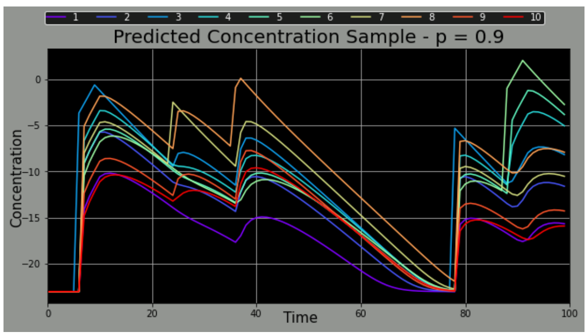
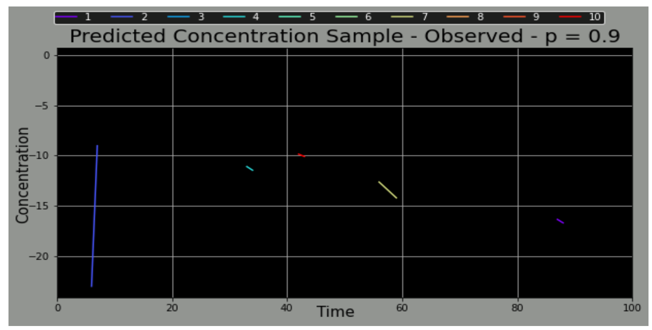
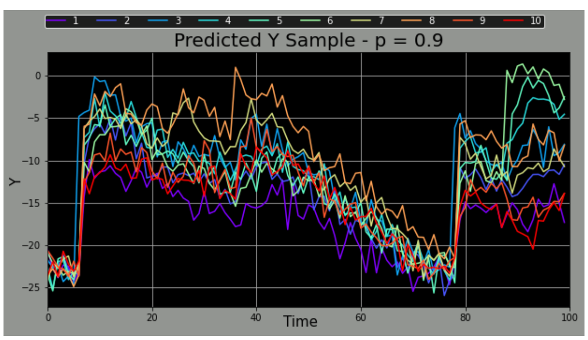
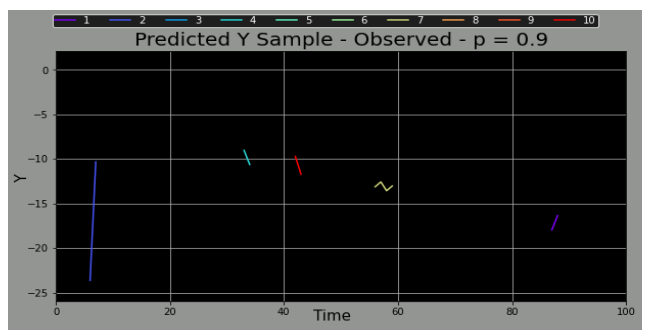
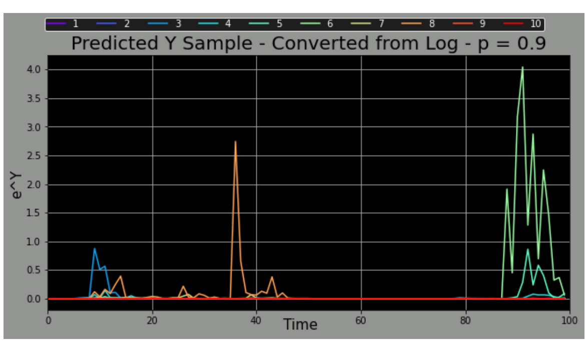
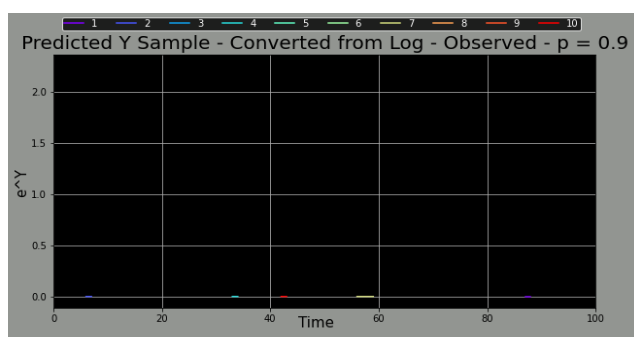
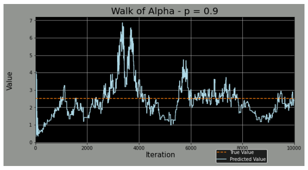
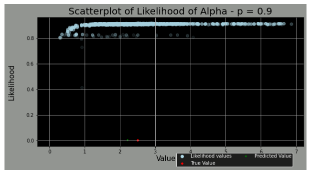
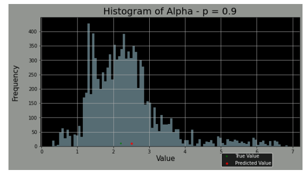
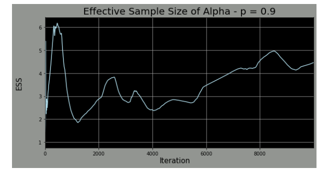
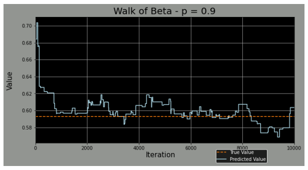
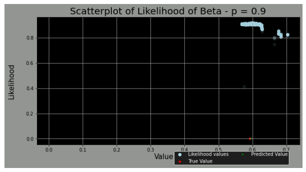
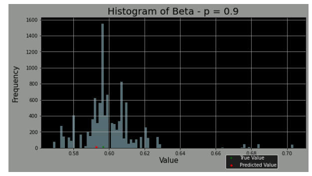
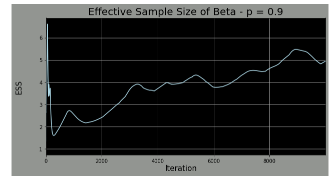
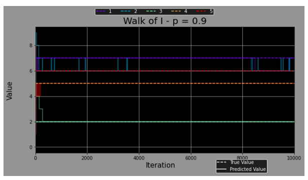
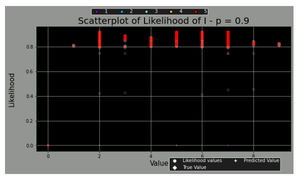
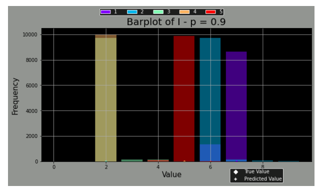
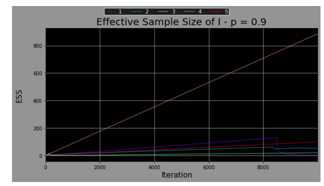
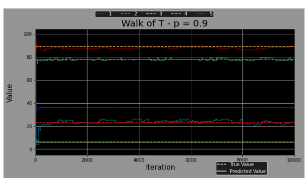
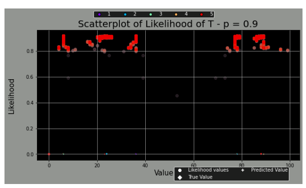
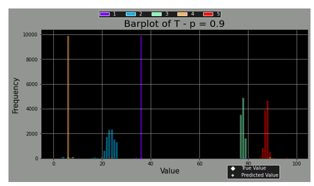
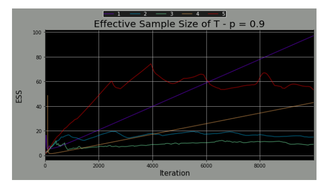
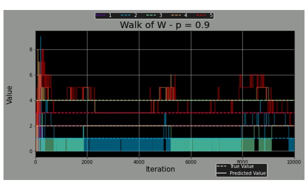
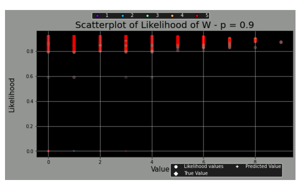
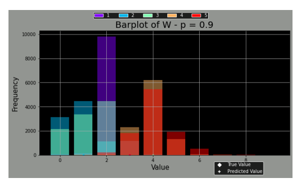
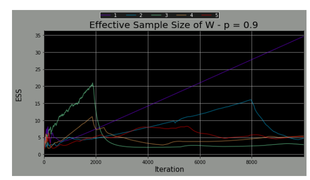
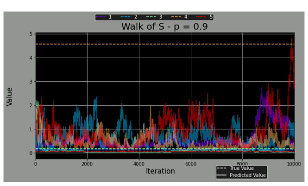
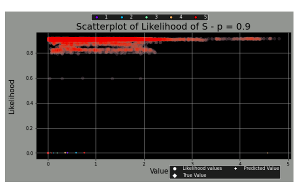
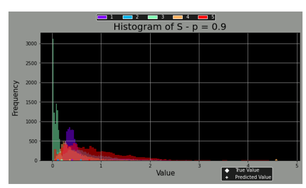
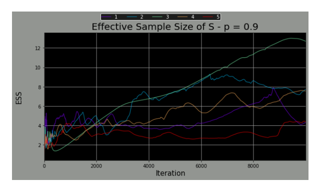
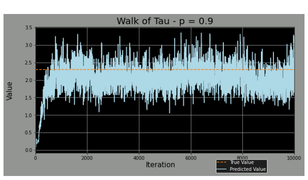
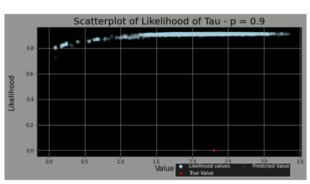
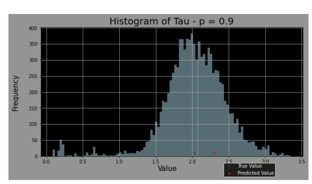
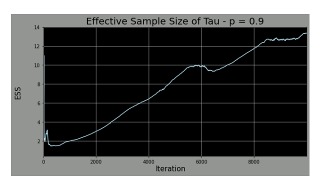
<table>
<tr><th>Results</th><th>Full</th><th>Observed</th></tr>
<tr><td>Mass</td><td>
</td><td>
</td></tr>
<tr><td>Concentration</td><td>
</td><td>
</td></tr>
<tr><td>Y</td><td>
</td><td>
</td></tr>
<tr><td>e^Y</td><td>
</td><td>
</td></tr>
</table>

Figure 23: Table of results for sparse data $(p = 0.9)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	2.2178	2.3364	2.3764	0.0930
β	0.5968	0.5993	0.5999	0.0126
I	7, 6, 2, 2, 5	6.8576, 6.0245, 2.0377, 2.0005, 4.9873	6.8710, 6.0367, 2.0483, 2.0049, 4.9929	0.1035
T	36, 24, 78, 6, 88	35.9667, 23.3786, 77.7880, 6.0179, 87.3527	35.9811, 23.5144, 77.8158, 6.0259, 87.4569	0.0142
W	2, 1, 1, 4, 4	1.9733, 1.0579, 1.2188, 3.8880, 4.0802	1.9807, 1.0987, 1.2496, 3.9142, 4.1146	0.2864
S	0.4093, 0.5891, 0.0696, 0.3627, 0.9171	0.5536, 0.6675, 0.1044, 0.4354, 0.8901	0.5711, 0.6837, 0.1144, 0.4461, 0.9171	0.5632
τ	2.0375	2.0249	2.0419	0.5571

$Table\ 11:\ Table\ of\ summary\ statistics\ of\ parameters\ for\ sparse\ data\ (p\ =\ 0.9)$

Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

$Figure\ 24:\ Table\ of\ graphs\ of\ parameter\ metrics\ for\ sparse\ data\ (p\ =\ 0.9)$

Changing N

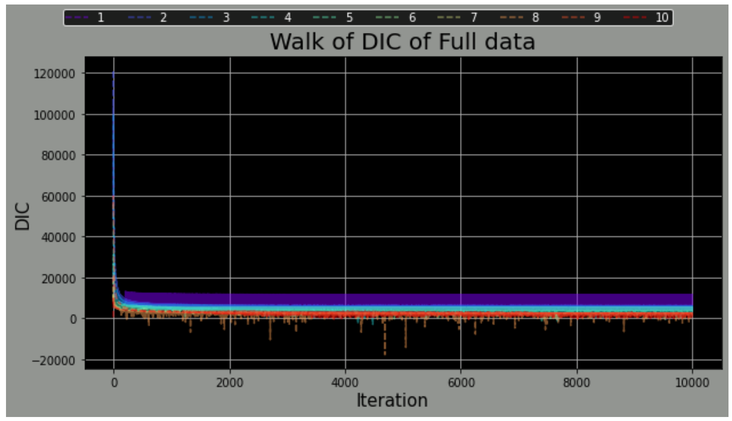
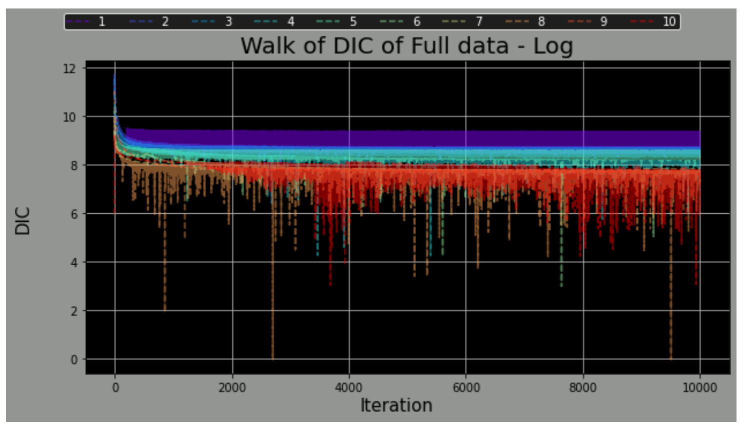
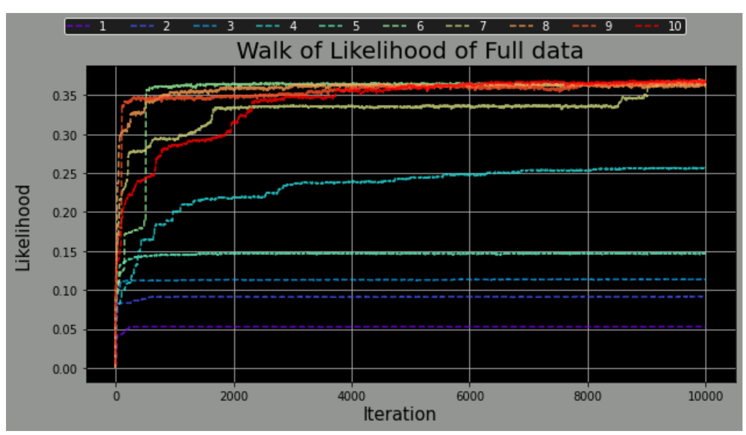
Model Comparison Metrics	Full Data
DIC	
DIC - Log	
Likelihood	

Figure 25: Table of Model Comparison metrics for models with varying numbers of events

N = 1

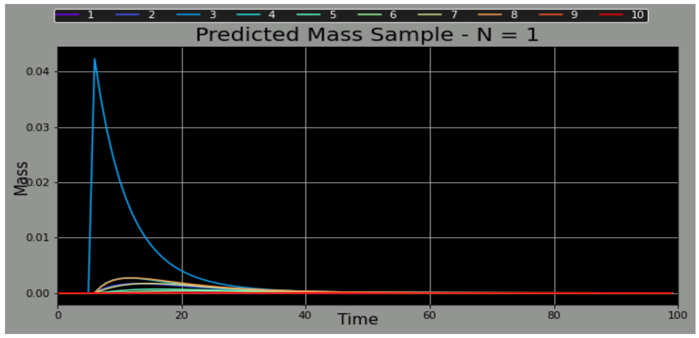
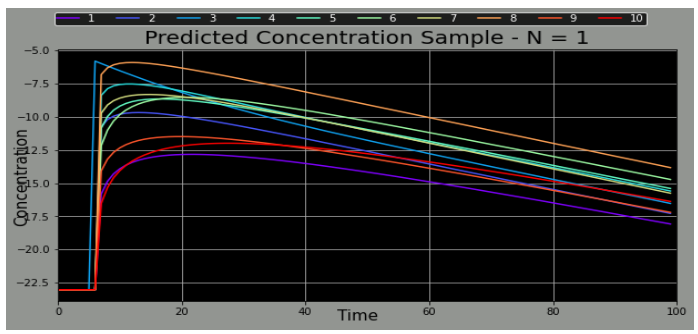
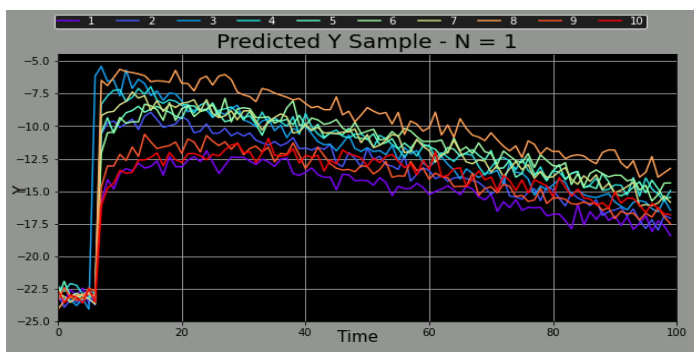
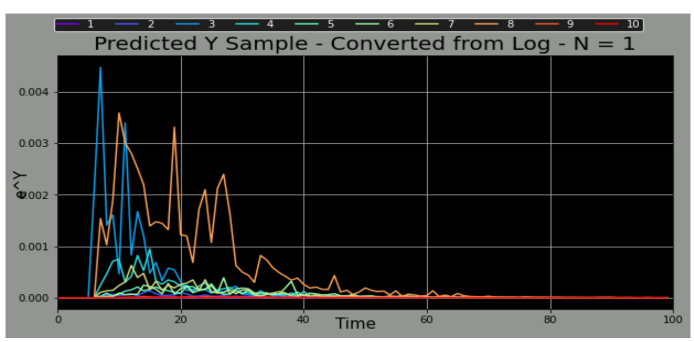
Results	Full Data
Mass	
Concentration	
Y	
e^Y	

Figure 26: Table of results for data (N = 1)

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	1.6679	1.9939	2.0533	0.5140
β	0.9106	0.9077	0.9084	0.0096
I	2	2.0769	2.0999	0.1031
T	6	6.0041	6.0371	0.0105
W	1	0.6516	0.6736	0.4788
S	0.0423	0.0654	0.0722	0.1019
τ	0.0479	0.0477	0.0479	0.0069

$Table\ 12\colon\ Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (N\ =\ 1)$

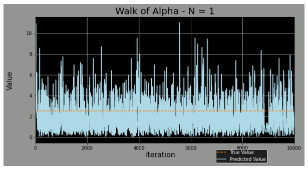
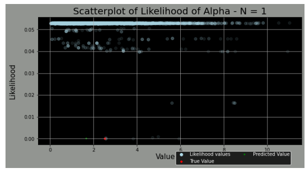
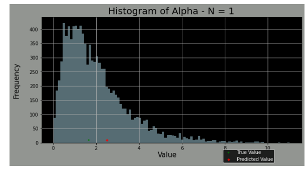
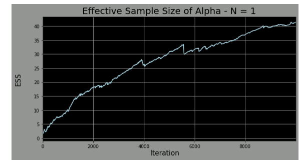
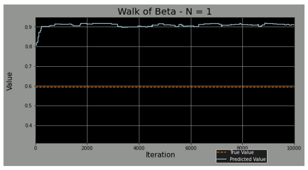
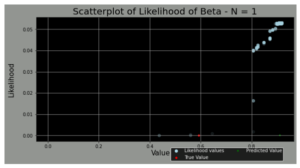
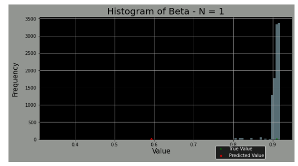
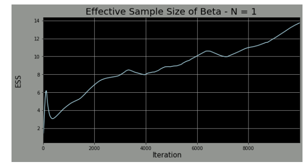
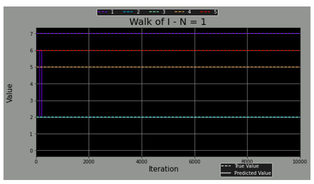
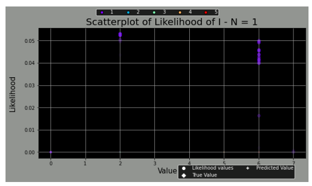
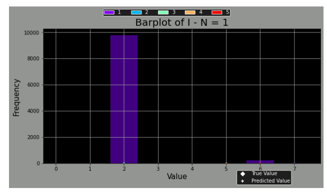
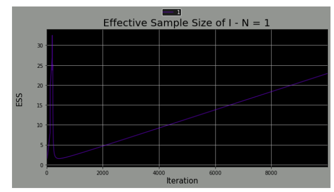
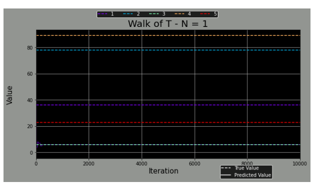
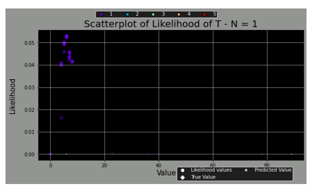
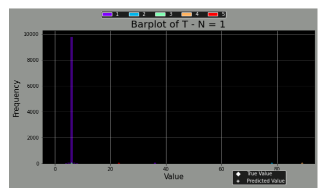
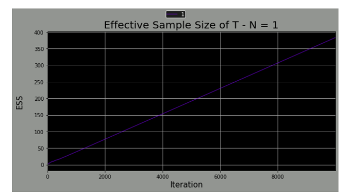
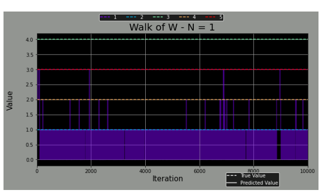
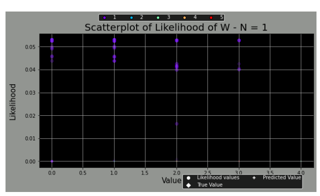
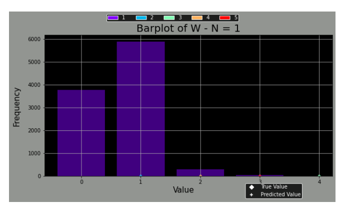
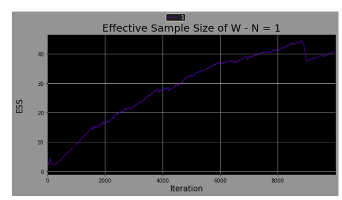
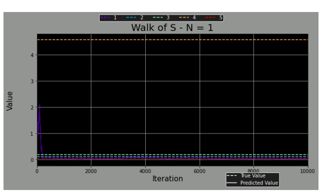
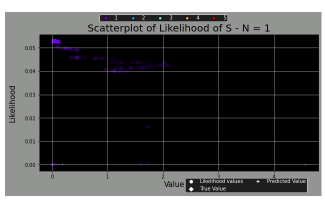
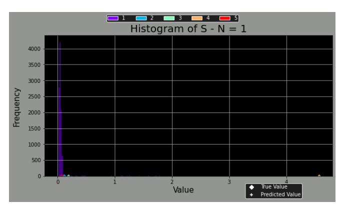
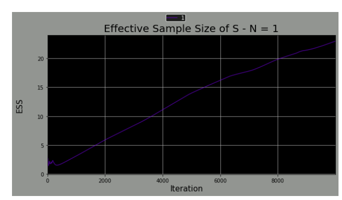
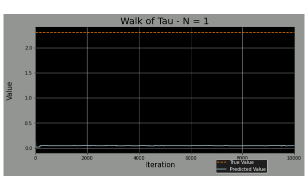
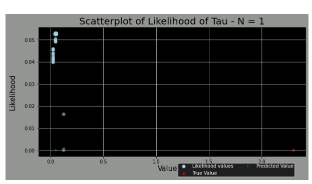
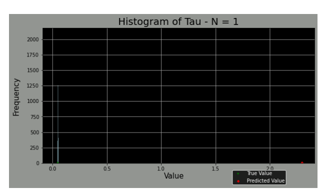
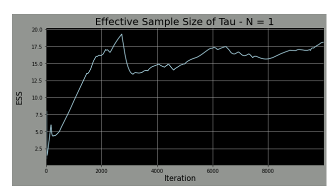
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 27: Table of graphs of parameter metrics for data $(N = 1)$

N = 2

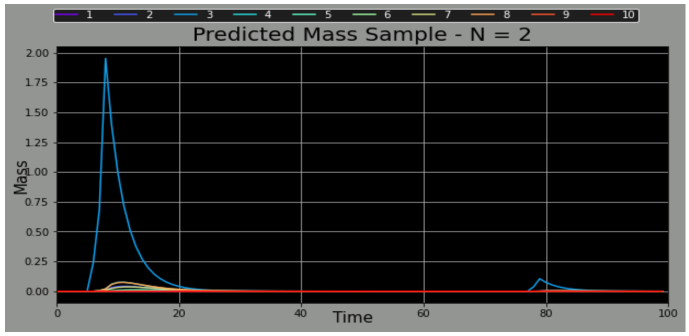
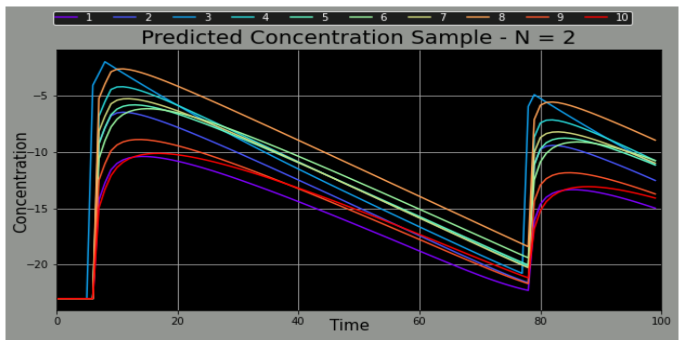
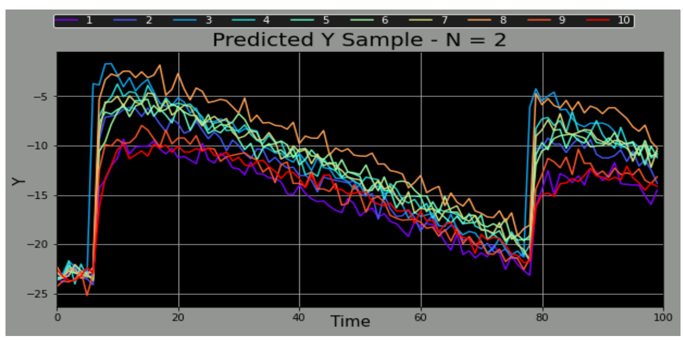
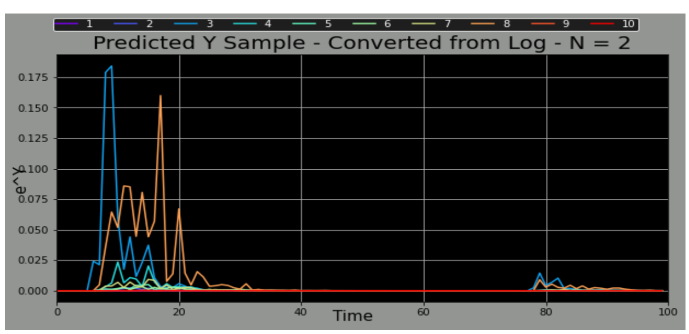
Results	Full Data
Mass	
Concentration	
Y	
e^Y	

Figure 28: Table of results for data $(N = 2)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	2.1388	2.3753	2.4272	0.0973
β	0.7818	0.7786	0.7791	0.0064
I	2, 2	2.1377, 2.0049	2.1677, 2.0105	0.1027
T	6, 78	5.9992, 77.9581	6.0128, 77.9989	0.0101
W	3, 2	3.0583, 1.7395	3.0704, 1.7829	0.2537
S	0.2400, 0.0367	0.2714, 0.0558	0.2775, 0.0574	0.2126
τ	0.1419	0.1413	0.1428	0.2126

$Table\ 13\colon\ Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (N\ =\ 2)$

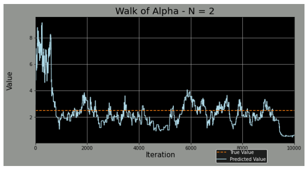
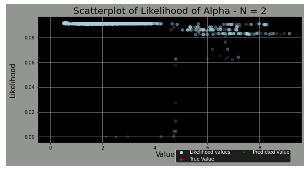
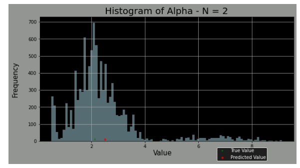
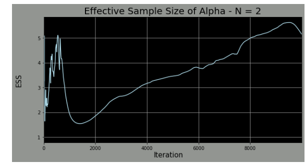
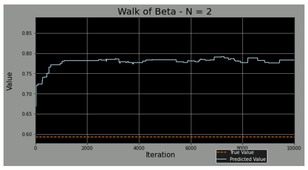
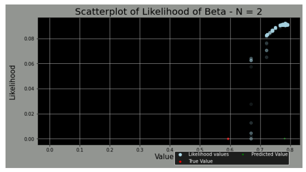
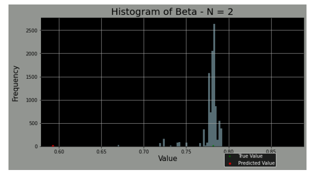
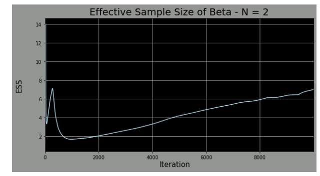
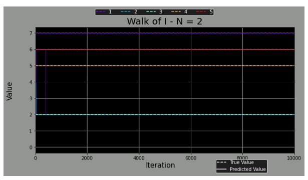
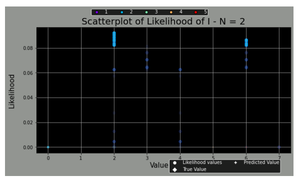
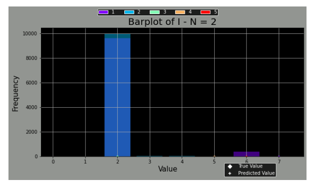
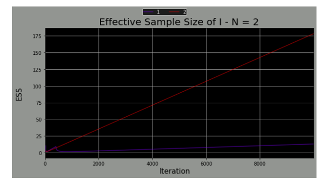
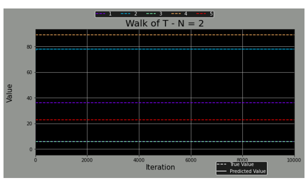
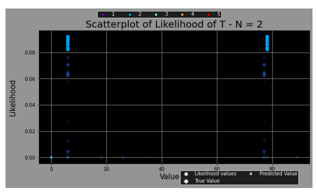
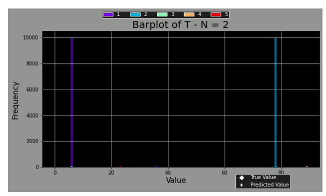
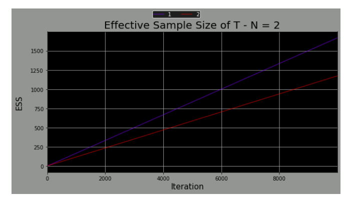
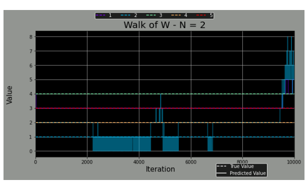
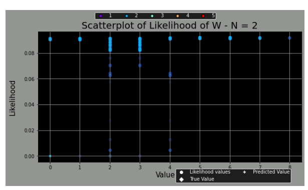
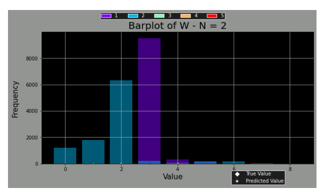
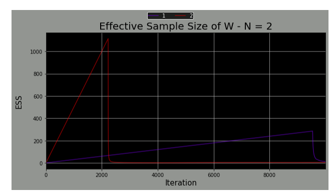
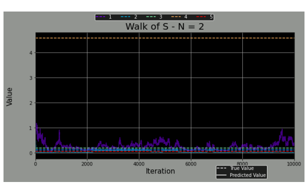
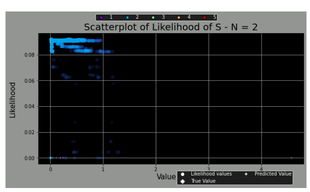
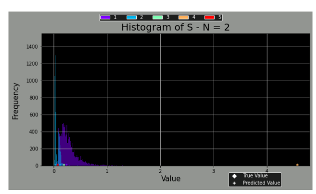
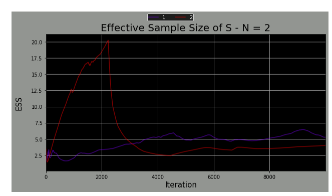
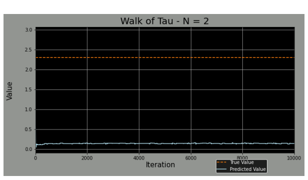
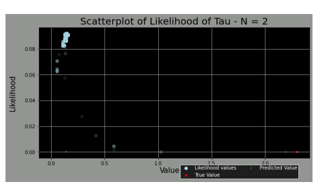
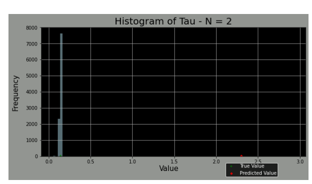
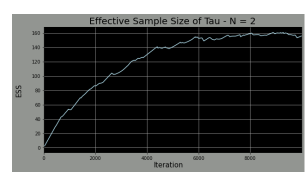
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 29: Table of graphs of parameter metrics for data $(N = 2)$

N = 3

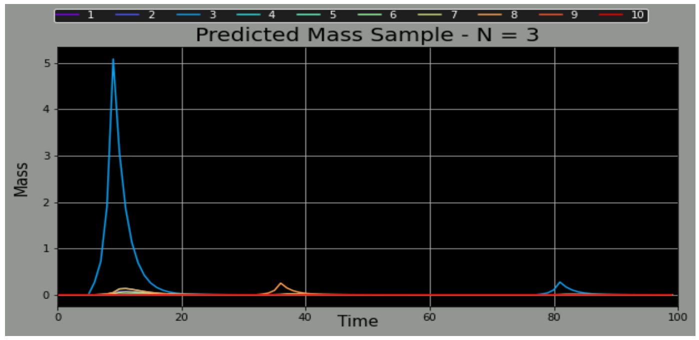
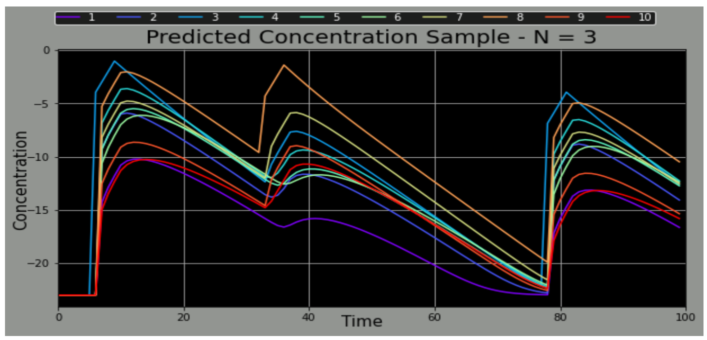
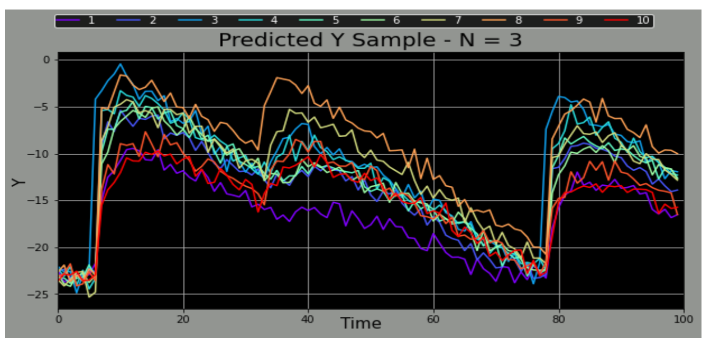
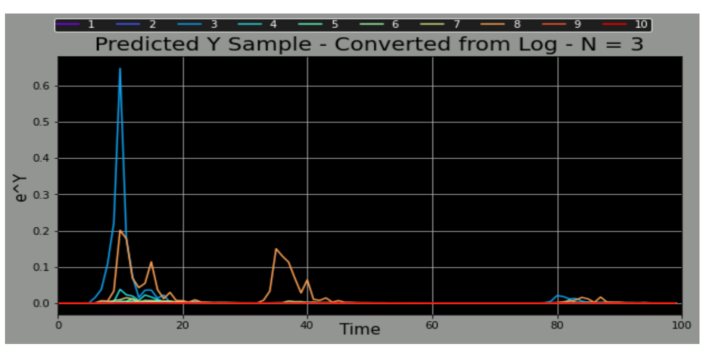
Results	Full Data
Mass	
Concentration	
Y	
e^Y	

Figure 30: Table of results for data $(N = 3)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	2.0542	2.0697	2.0817	0.0434
β	0.6620	0.6613	0.6616	0.0070
I	7, 2, 2	6,9936, 2.0066, 2.0071	6.9990, 2.0152, 2.0117	0.1010
T	33, 6, 78	32.9456, 5.9948, 77.9644	32.9695, 6.0324, 78.0030	0.0114
W	4, 4, 4	3.9968, 4.0037, 4.3092	3.9994, 4.0085, 4.3352	0.1850
S	0.0135, 0.2705, 0.0146	0.0147, 0.2842, 0.0193	0.0150, 0.2882, 0.0204	0.1475
τ	0.2187	0.2175	0.2185	0.0274

$Table\ 14\text{:}\ Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (N\ =\ 3)$

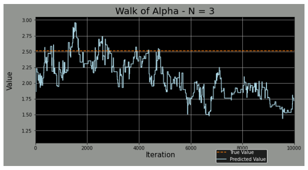
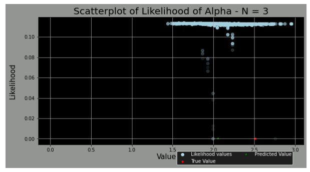
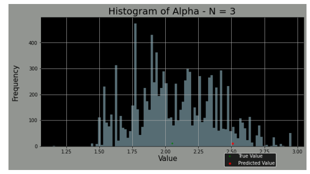
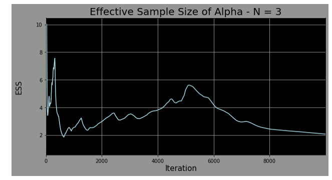
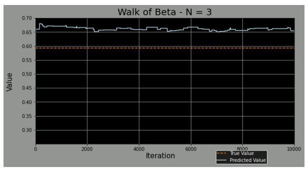
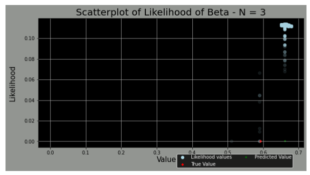
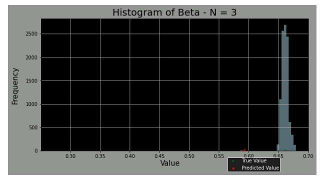
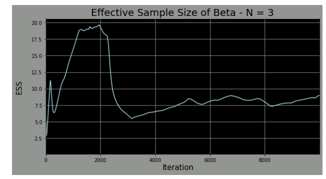
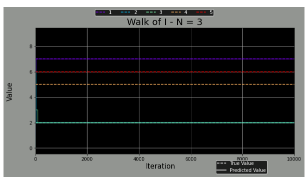
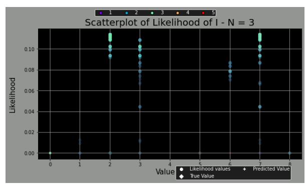
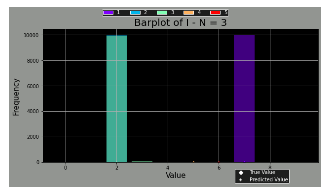
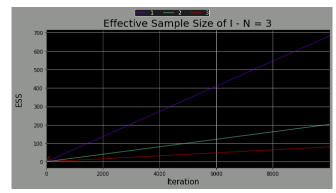
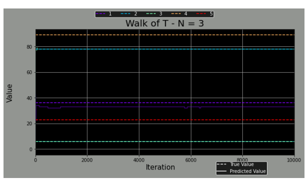
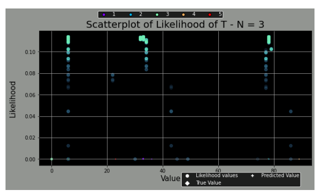
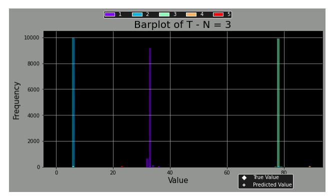
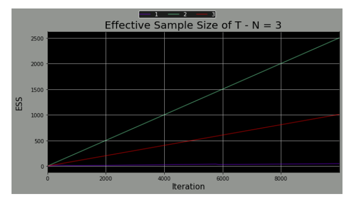
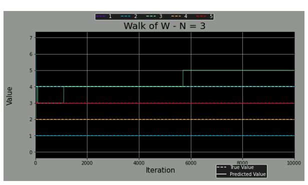
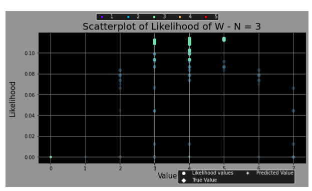
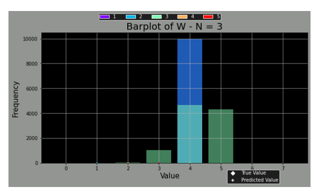
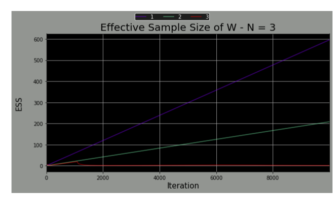
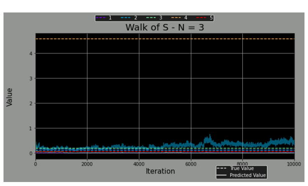
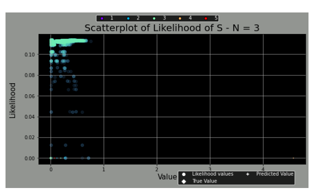
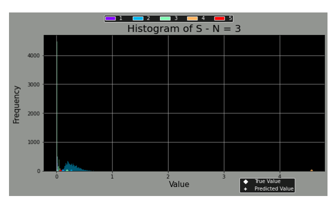
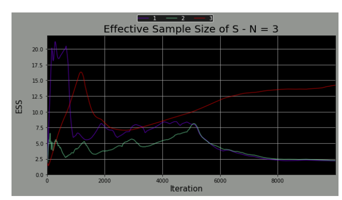
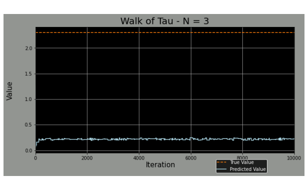
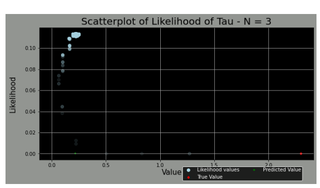
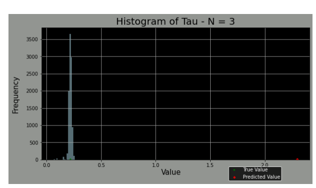
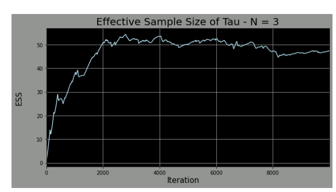
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 31: Table of graphs of parameter metrics for data (N = 3)

N = 4

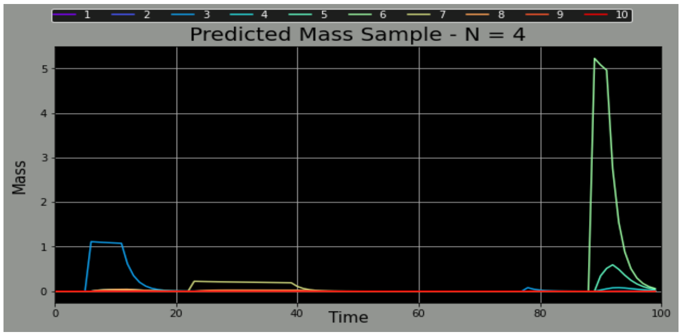
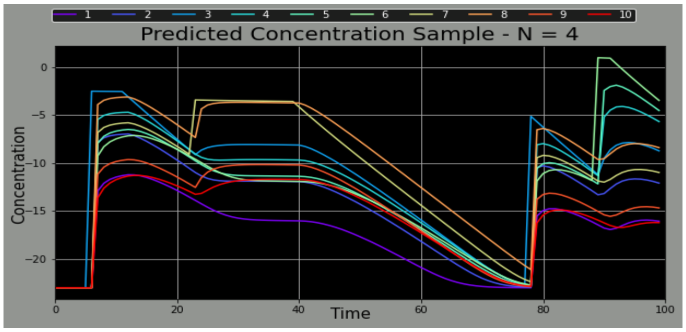
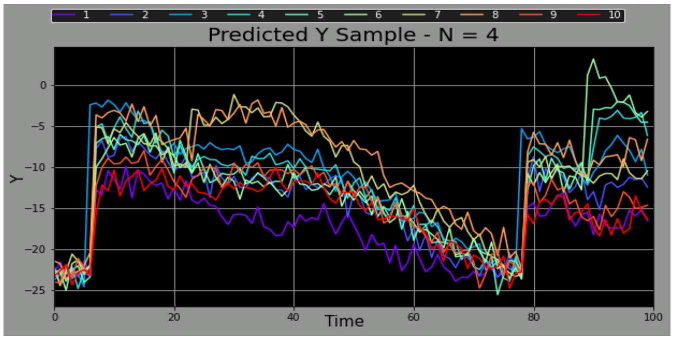
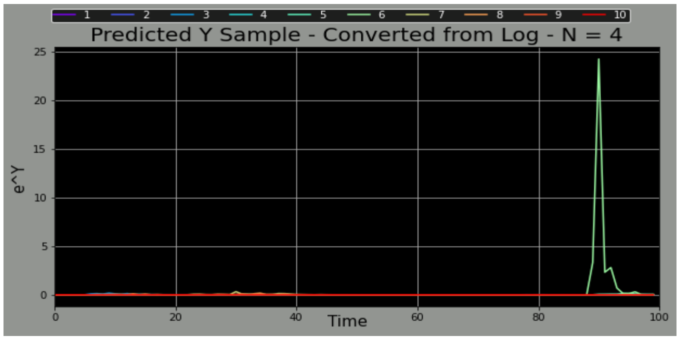
Results	Full Data
Mass	
Concentration	
Y	
e^Y	

Figure 32: Table of results for data $(N = 4)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	0.4246	0.4547	0.4631	0.0084
β	0.6220	0.6363	0.6380	0.0066
I	5, 6, 2, 2	4.9962, 5.9631, 2.0200, 2.0031	4.9992, 5.9719, 2.0328, 2.0077	0.1005
T	89, 23, 6, 78	88.9816, 24.3957, 5.9889, 77.9151	89.0068, 24.8053, 6.0161, 77.9901	0.0110
W	3, 17, 6, 1	3.0306, 16.3456, 5.7511, 0.6681	3.0580, 16.4942, 5.7887, 0.6895	0.2617
S	5.2290, 0.2223, 1.1126, 0.0768	5.0609, 0.2417, 0.1900, 0.0755	5.1135, 0.2478, 1.0953, 0.0768	0.4562
τ	1.1015	0.9421	0.9509	0.1088

$Table\ 15\colon\ Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (N\ =\ 4)$

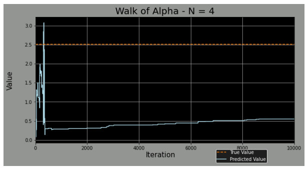
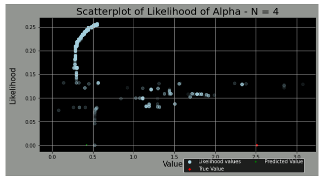
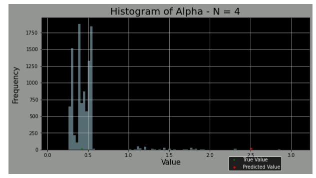
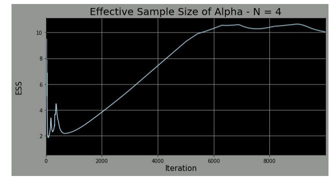
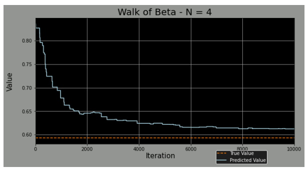
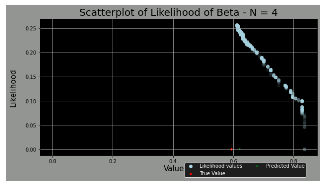
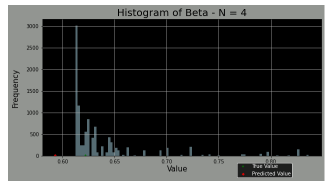
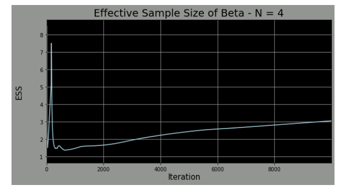
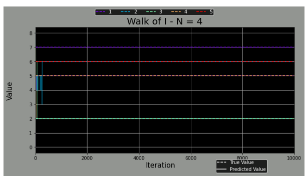
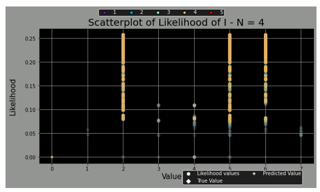
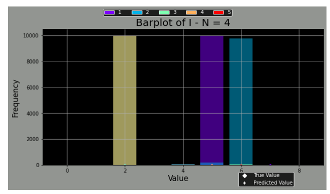
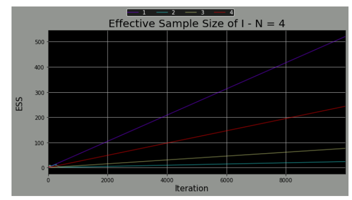
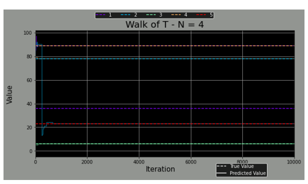
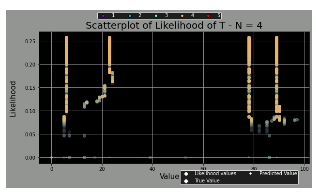
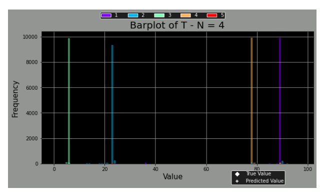
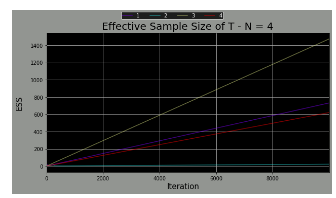
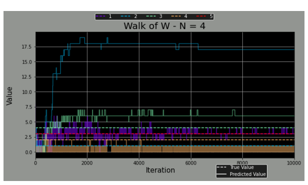
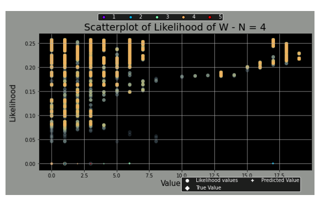
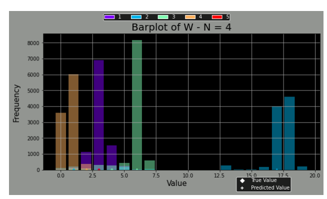
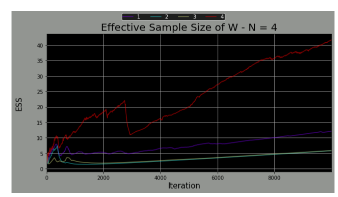
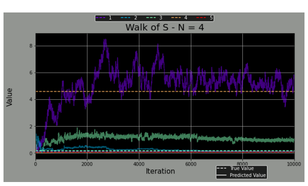
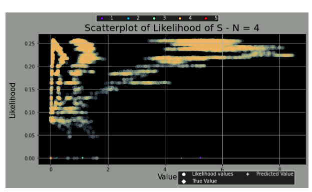
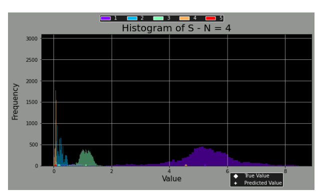
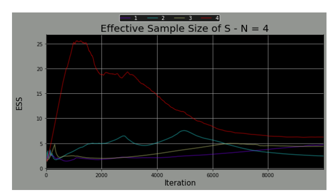
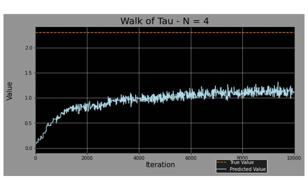
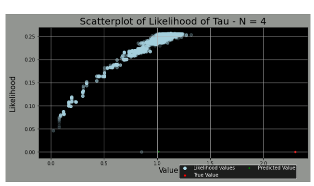
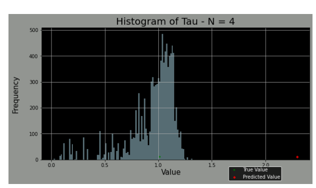
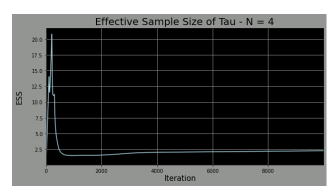
<table>
<tr><th>Parameter</th><th>Walk of value</th><th>Scatter Plot of Likelihood</th><th>Histogram</th><th>Walk of ESS</th></tr>
<tr><td>α</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>β</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>I</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>T</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>W</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>S</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>τ</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
</table>

Figure 33: Table of graphs of parameter metrics for data $(N = 4)$

N = 5

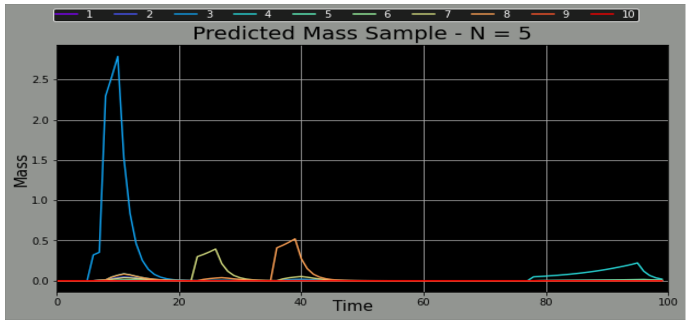
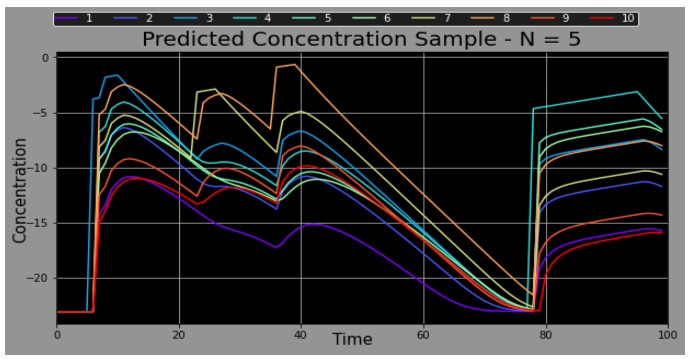
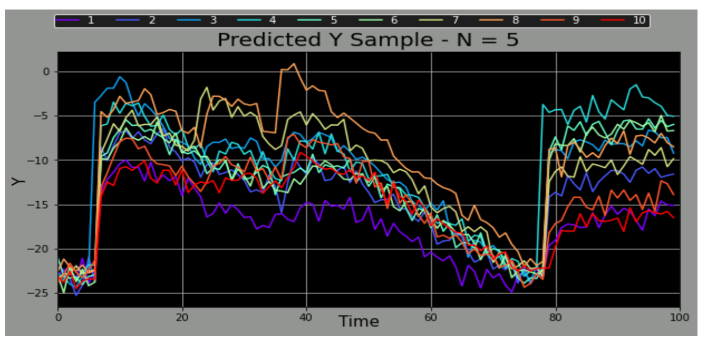
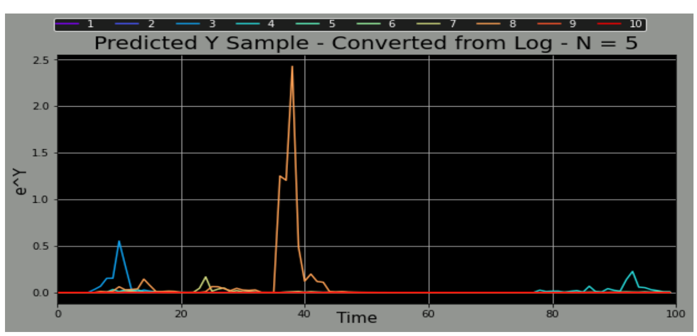
Results	Full Data
Mass	
Concentration	
Y	
e^Y	

Figure 34: Table of results for data $(N = 5)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	0.5548	0.5707	0.5743	0.0091
β	0.5994	0.5990	0.5993	0.0078
I	6, 7, 2, 2, 3	6.0000, 6.9972, 2.0004, 1.9992, 3.0064	6.0018, 7.0000, 2.0048, 2.0006, 3.0114	0.1004
T	23, 36, 6, 8, 78	22.9822, 36.1541, 5.9970, 7.8128, 77.9918	22.9924, 36.1876, 6.0186, 7.8518, 77.9956	0.0111
W	4, 4, 3, 3, 18	3.5851, 3.4623, 2.9544, 2.6298, 17.8520	3.6201, 3.5087, 3.0220, 2.6822, 17.9038	0.3027
S	0.2981, 0.4059, 0.3213, 1.9060, 0.0480	0.3255, 0.5565, 0.3370, 1.9980, 0.0520	0.3311, 0.5750, 0.3431, 2.0263, 0.0562	0.5236
τ	0.3665	0.3651	0.3662	0.0411

$Table\ 16\text{:}\ Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (N\ =\ 5)$

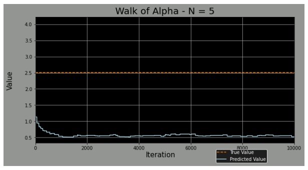
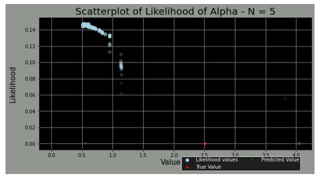
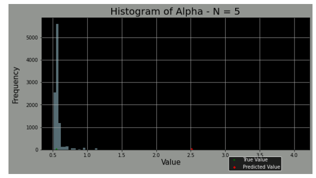
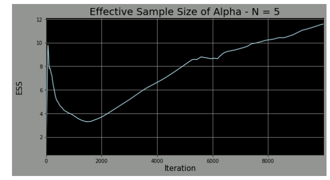
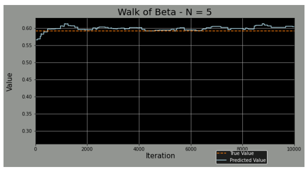
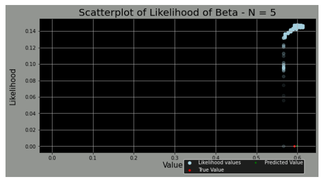
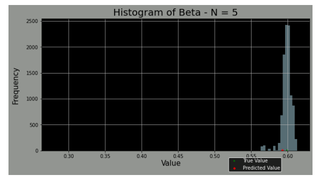
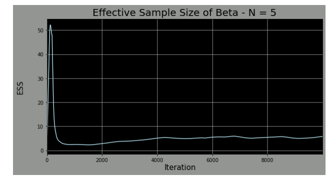
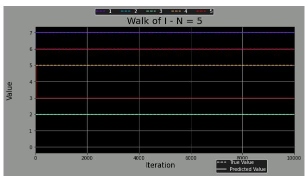
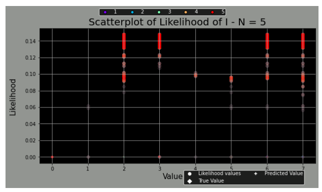
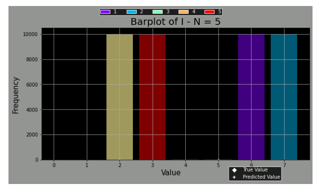
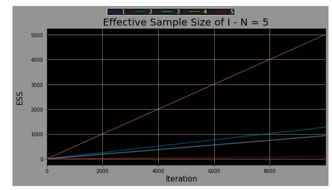
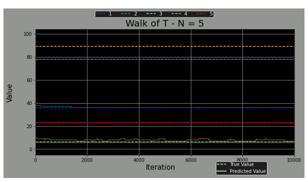
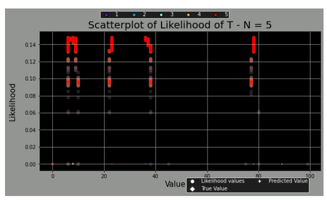
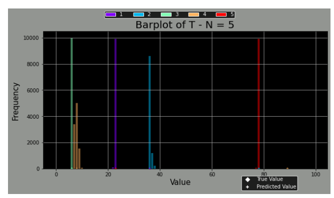
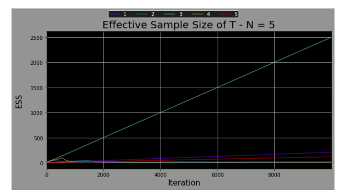
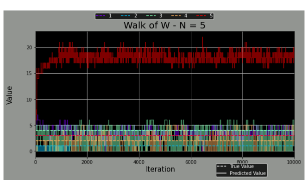
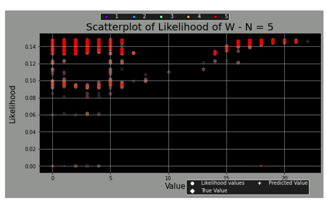
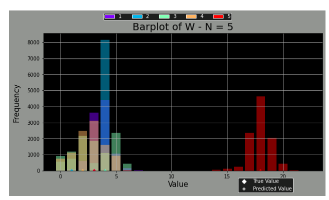
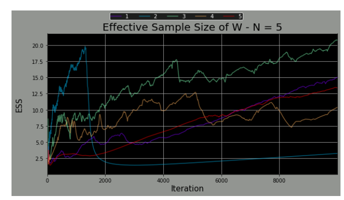
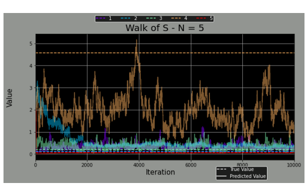
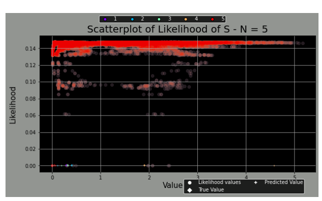
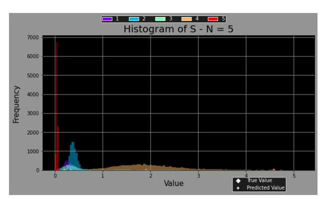
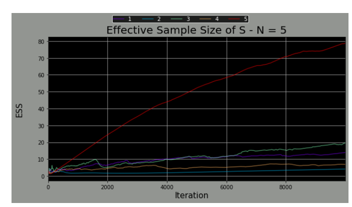
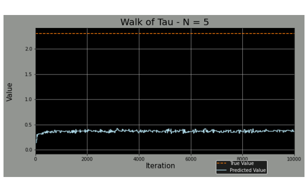
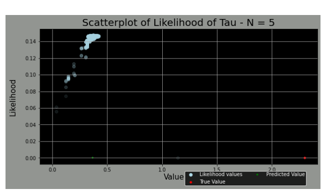
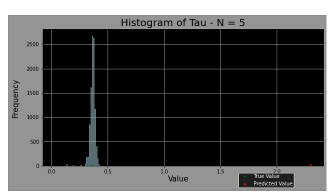
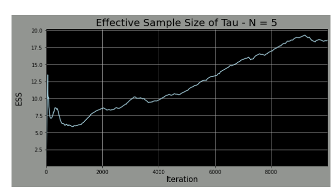
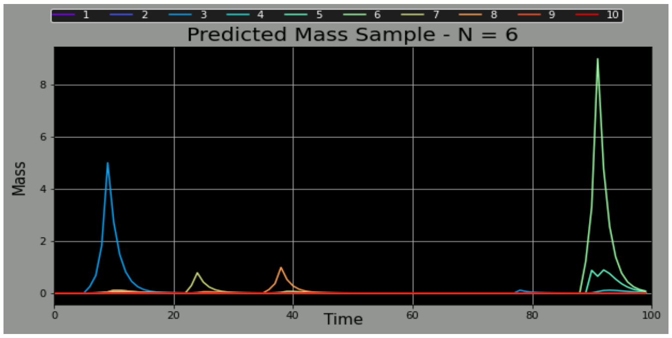
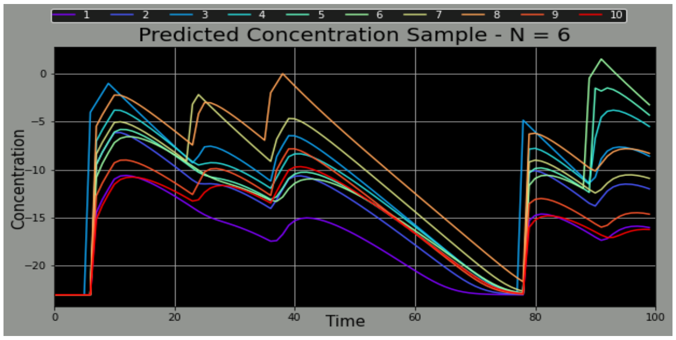
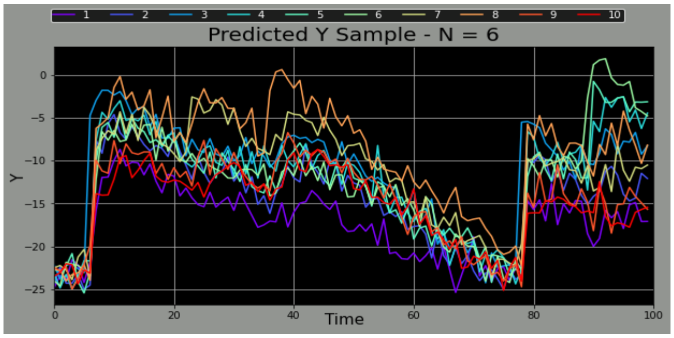
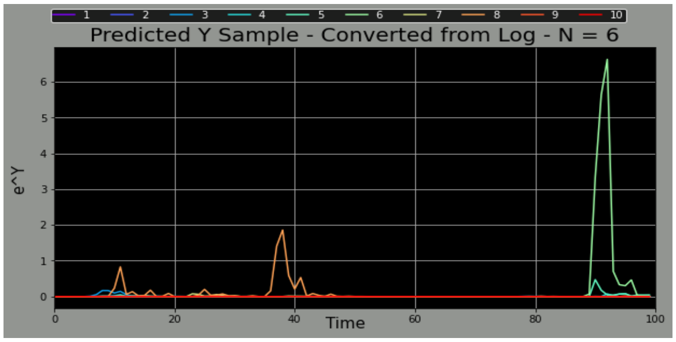
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 35: Table of graphs of parameter metrics for data $(N = 5)$

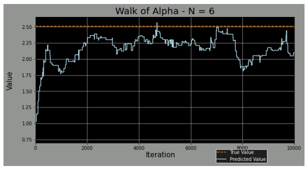
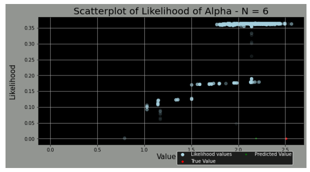
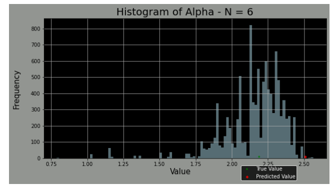
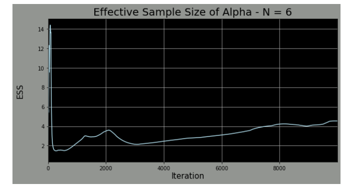
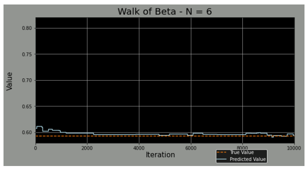
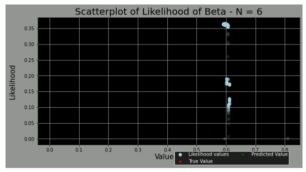
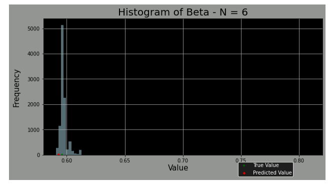
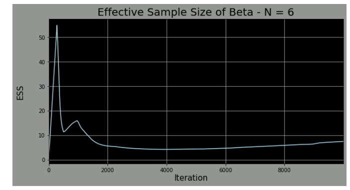
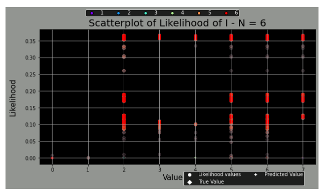
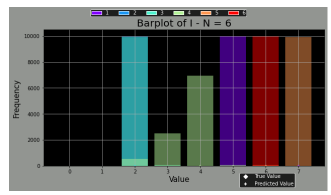
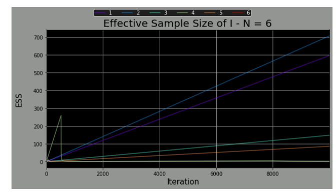
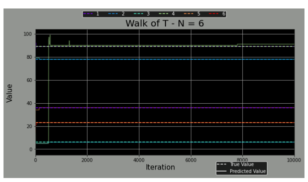
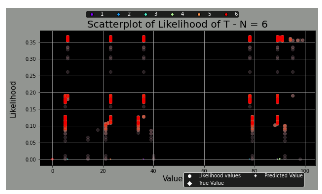
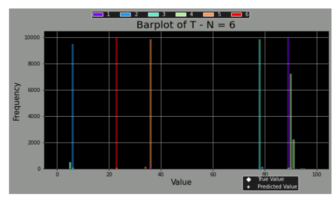
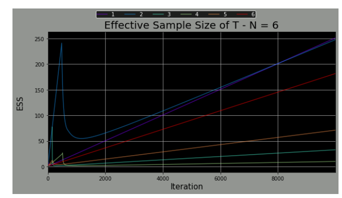
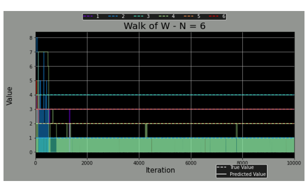
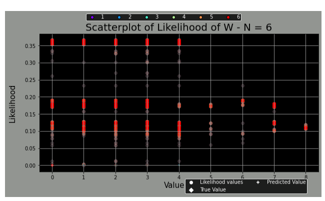
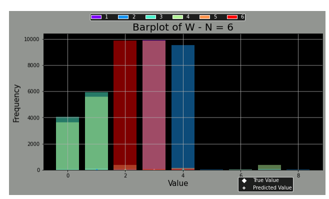
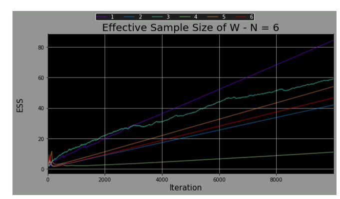
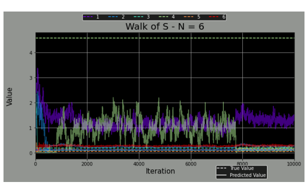
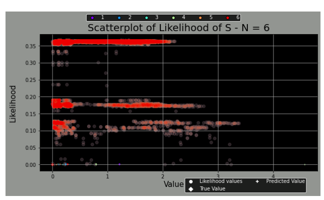
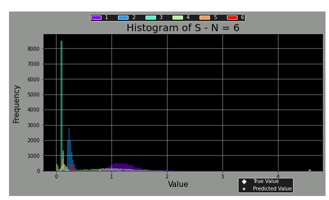
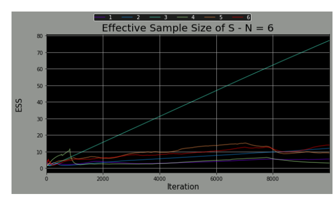
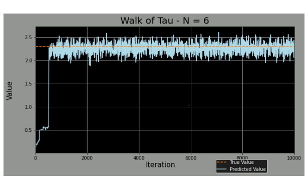
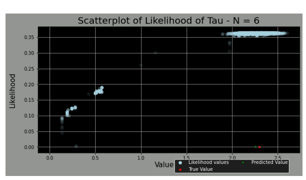
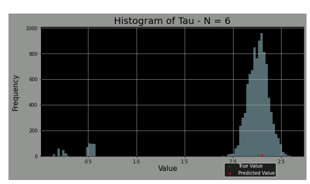
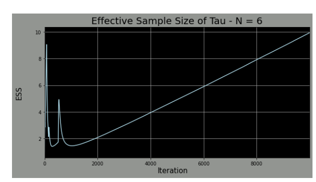
N = 6

Results	Full Data
Mass	
Concentration	
Y	
e^Y	

$Figure\ 36\colon\ Table\ of\ results\ for\ data\ (N\ =\ 6)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	2.1897	2.1545	2.1626	0.0411
β	0.5961	0.5972	0.5974	0.0037
I	5, 2, 2, 4, 7, 6	4.9968, 2.0005, 2.0034, 3.6431, 6.9888, 6.0000	4.9994, 2.0033, 2.0070, 3.6661, 6.9934, 6.0000	0.1001
T	89, 6, 78, 90, 36, 23	89.0048, 5.9432, 78.0129, 85.5607, 35.9545, 22.9819	89.0126, 5.9776, 78.0177, 86.2927, 35.9716, 22.9917	0.0101
W	3, 4, 1, 1, 3, 2	2.9866, 3.9342, 0.5876, 0.8967, 3.0099, 2.0163	2.9920, 3.9592, 0.6072, 0.9499, 3.0149, 2.0239	0.2960
S	1.2125, 0.2446, 0.1080, 0.7943, 0.1305, 0.2825	1.2504, 0.2887, 0.1177, 0.7355, 0.1322, 0.2802	1.2621, 0.2982, 0.1240, 0.7561, 0.1330, 0.2814	0.3371
τ	2.2587	2.1605	2.1768	0.2390

$Table\ 17\!:\ Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (N\ =\ 6)$

Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

$$Figure\ 37:\ Table\ of\ graphs\ of\ parameter\ metrics\ for\ data\ (N\ =\ 6)$$

N = 7

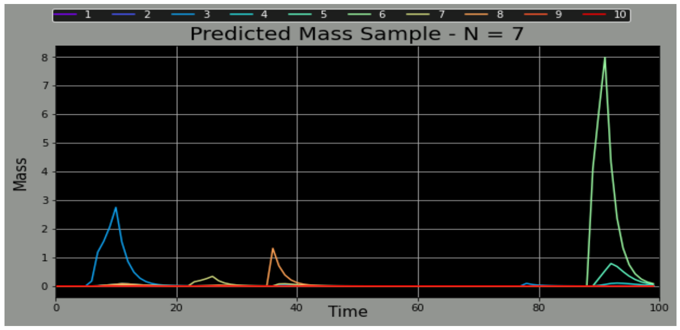
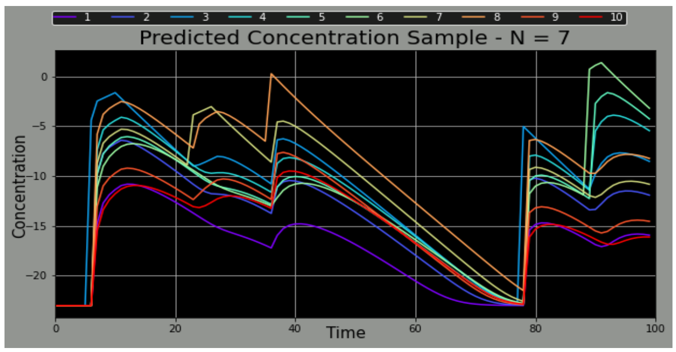
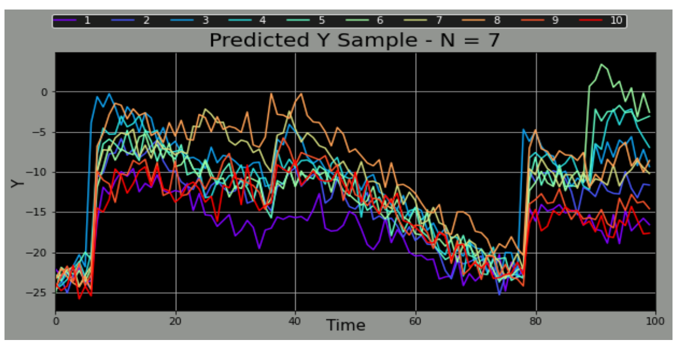
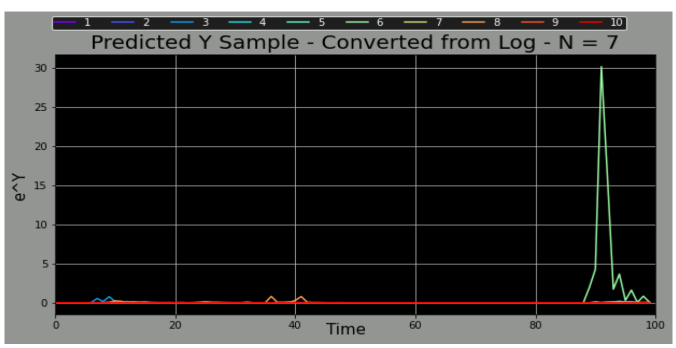
<table>
<tr><th>Results</th><th>Full Data</th></tr>
<tr><td>Mass</td><td>
</td></tr>
<tr><td>Concentration</td><td>
</td></tr>
<tr><td>Y</td><td>
</td></tr>
<tr><td>e^Y</td><td>
</td></tr>
</table>

Figure 38: Table of results for data (N = 7)

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	0.7676	0.7280	0.7342	0.0133
β	0.6103	0.6127	0.6131	0.0033
I	2, 2, 6, 5, 2, 7, 5	1.9997, 2.0486, 5.9985, 4.5941, 1.9992, 6.9974, 4,9964	2.0041, 2.0672, 6.0005, 4.6455, 2.0010, 7.0000, 4.9994	0.1031
T	6, 7, 23, 90, 78, 36, 89	5,9958, 6.9973, 21.1177, 82.4545, 78.0000, 36.0160, 88.9872	6.0148, 7.0365, 21.2919, 83.1937, 78.0000, 36.0476, 88.9928	0.0112
W	3, 4, 4, 1, 1, 1, 3	2.7228, 3.2627, 5.9804, 0.9212, 0.5941, 0.7071, 2.3628	2.7964, 3.3125, 6.1973, 0.9552, 0.6157, 0.7313, 2.4016	0.3766
S	0.1725, 0.9525, 0.1457, 0.7545, 0.0918, 1.3144, 4.1075	0.1866, 1.0007, 0.1447, 1.2188, 0.0907, 1.1219, 4.0826	0.1905, 1.0124, 0.1502, 1.2724, 0.0911, 1.1381, 4.1156	0.4567
τ	1.9146	1.8676	1.8787	0.1990

$Table\ 18\text{:}\ Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (N\ =\ 7)$

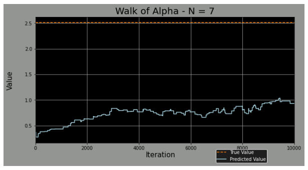
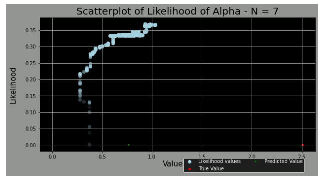
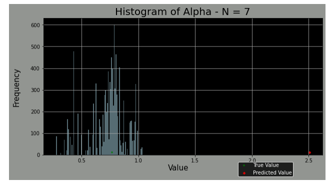
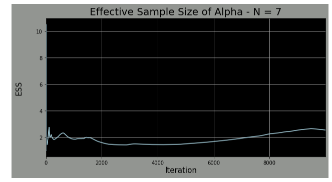
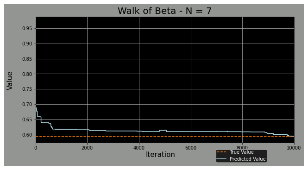
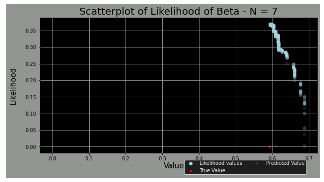
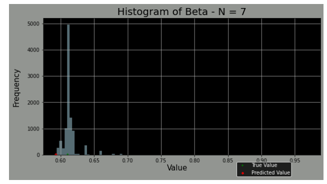
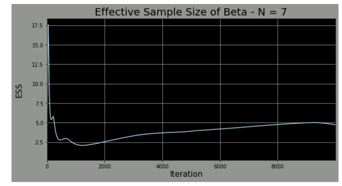
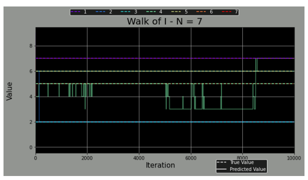
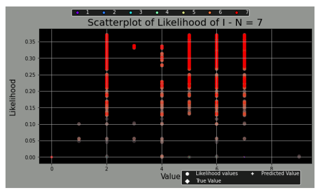
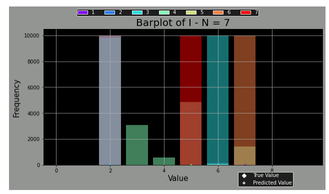
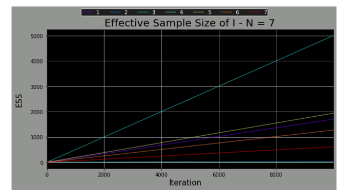
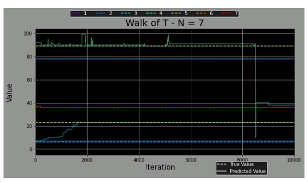
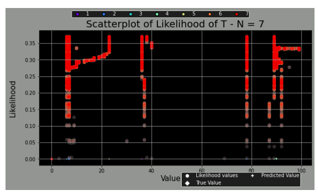
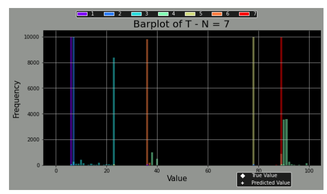
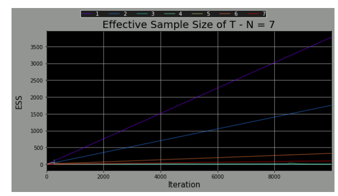
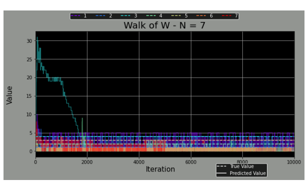
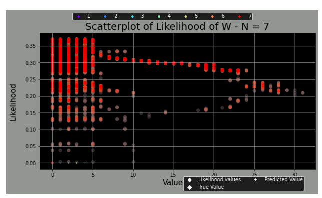
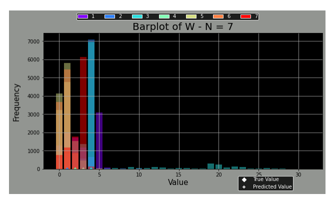
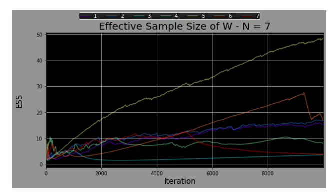
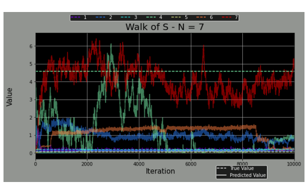
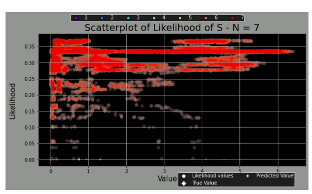
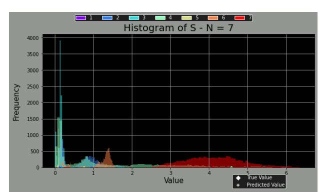
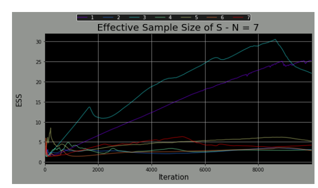
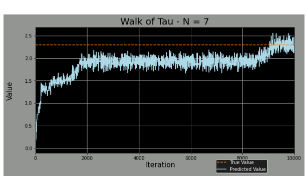
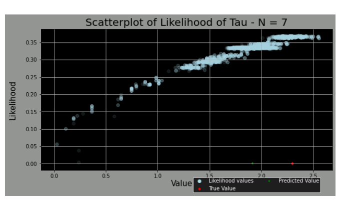
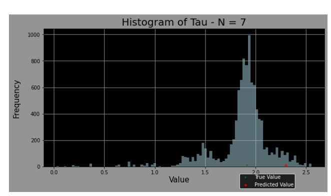
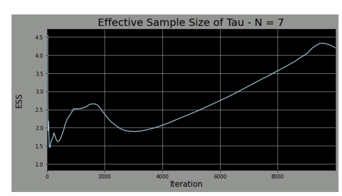
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 39: Table of graphs of parameter metrics for data $(N = 7)$

N = 8

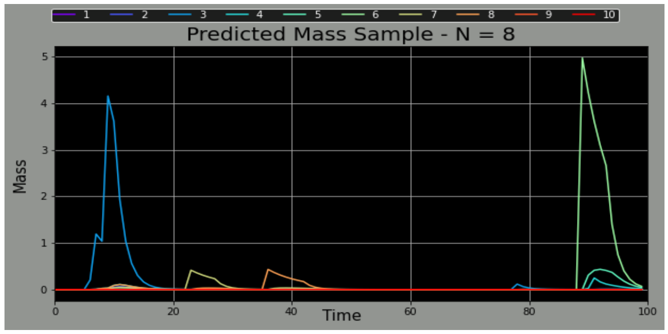
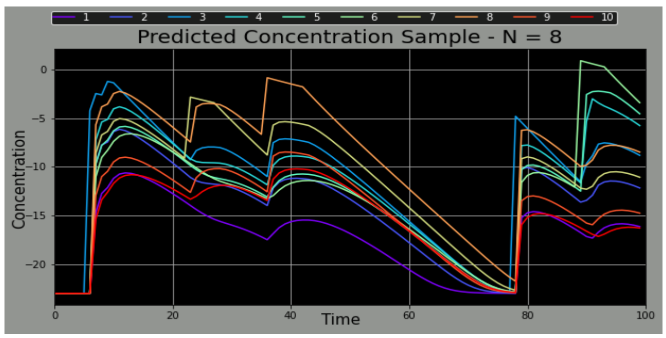
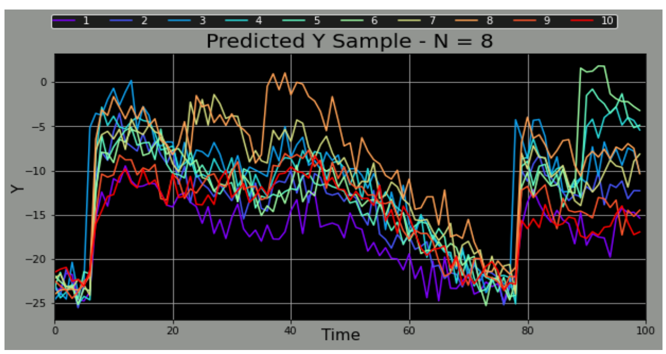
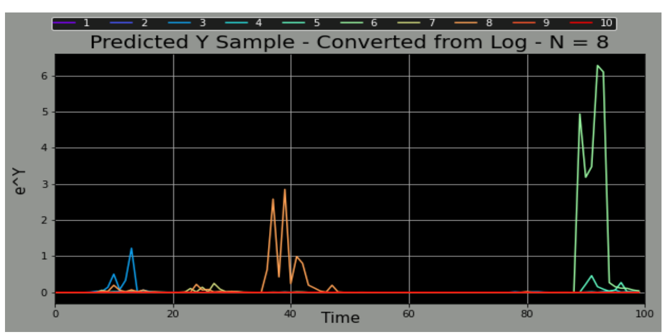
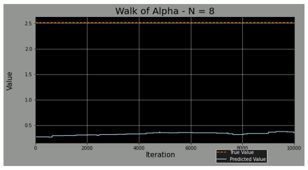
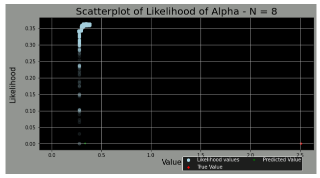
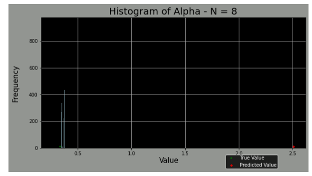
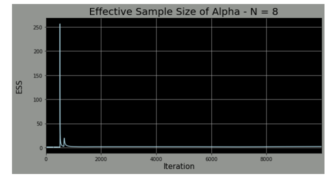
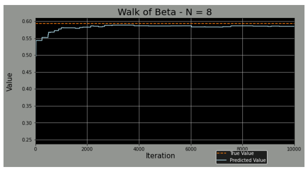
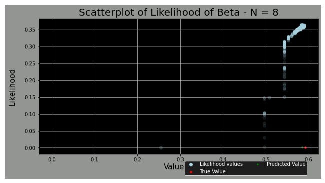
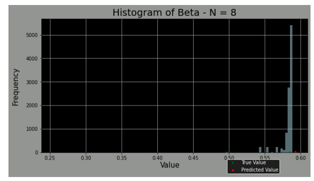
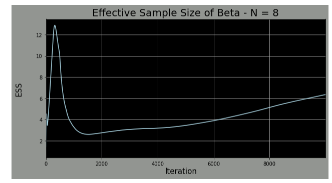
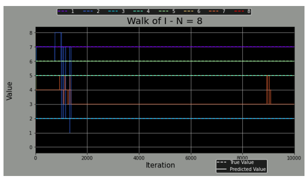
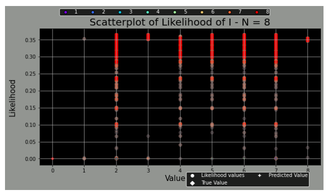
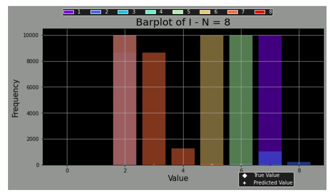
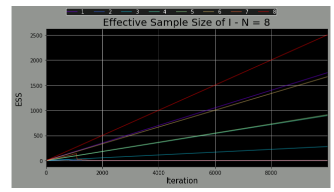
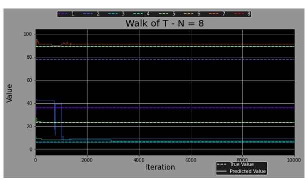
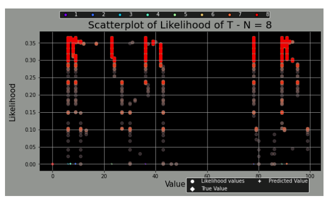
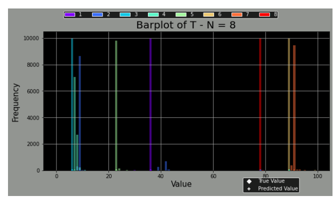
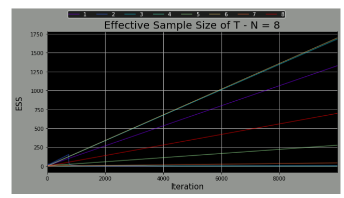
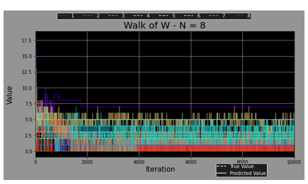
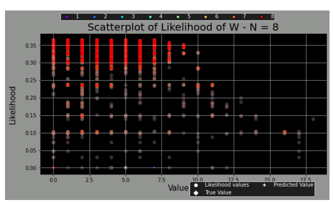
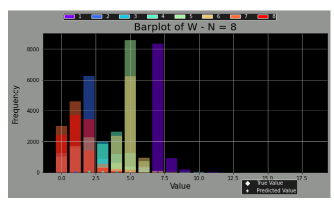
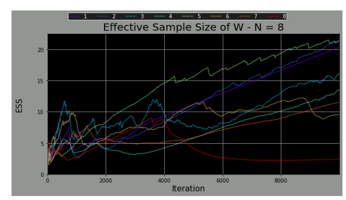
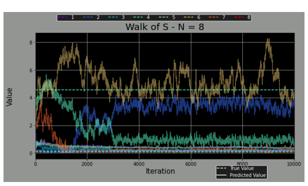
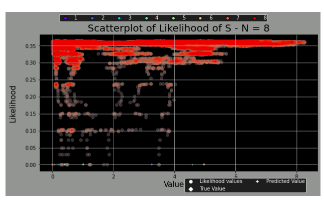
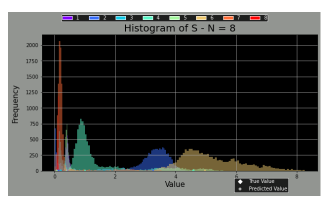
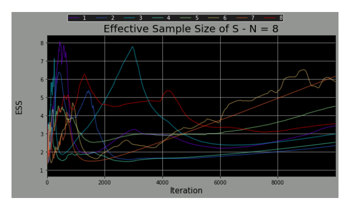
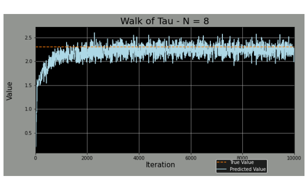
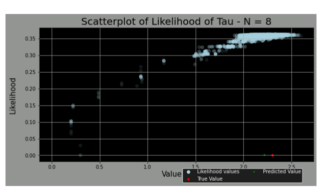
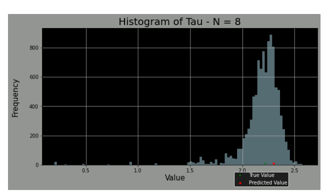
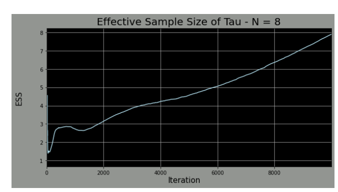
<table>
<tr><th>Results</th><th>Full Data</th></tr>
<tr><td>Mass</td><td>
</td></tr>
<tr><td>Concentration</td><td>
</td></tr>
<tr><td>Y</td><td>
</td></tr>
<tr><td>e^Y</td><td>
</td></tr>
</table>

Figure 40: Table of results for sparse data $(N = 8)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	0.3346	0.3327	0.3338	0.0038
β	0.5855	0.5821	0.5825	0.0041
I	7, 2, 2, 2, 6, 5, 3, 2	6.9963, 2.6556, 2.0047, 2.0003, 5.9955, 4.9973, 3.1329, 1.9990	7.0003, 2.7244, 2.0129, 2.0033, 5.9997, 5.0005, 3.1473, 2.0006	0.1002
T	36, 9, 6, 7, 23, 89, 91, 78	35.9760, 12.0173, 5.9982, 7.3105, 23.0328, 88.9807, 90.9813, 78.0016	36.0020, 12.3969, 6.0286, 7.3315, 23.0892, 89.0159, 90.9967, 78.0102	0.0101
W	7, 2, 2, 3, 5, 5, 1, 1	7.0800, 1.7483, 2.4983, 2.5014, 4.9897, 4.8012, 1.0936, 1.1769	7.1051, 1.7813, 2.5633, 2.5638, 5.0077, 4.8302, 1.1386, 1.2113	0.3481
S	0.4298, 3.2537, 0.2035, 1.0149, 0.4125, 4.9694, 01947, 0.1153	0.4446, 2.7444, 0.2339, 1.5086, 0.4307, 5.1501, 0.3485, 0.1102	0.4479, 2.7932, 0.2379, 1.5542, 0.4340, 5.1895, 0.3713, 0.1111	0.5092
τ	2.2185	2.1853	2.1931	0.2409

$Table\ 19\text{:}\ Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (N\ =\ 8)$

Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

$Figure\ 41:\ Table\ of\ graphs\ of\ parameter\ metrics\ for\ data\ (N\ =\ 8)$

N = 9

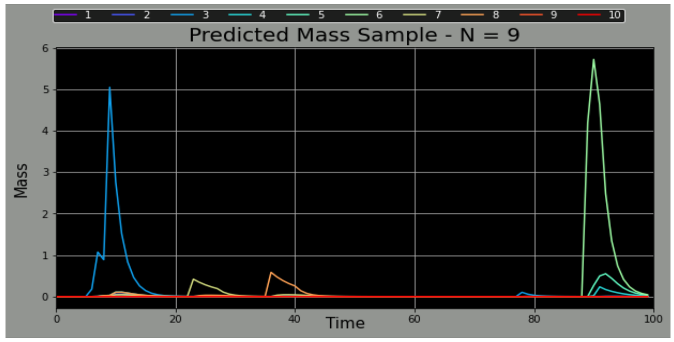
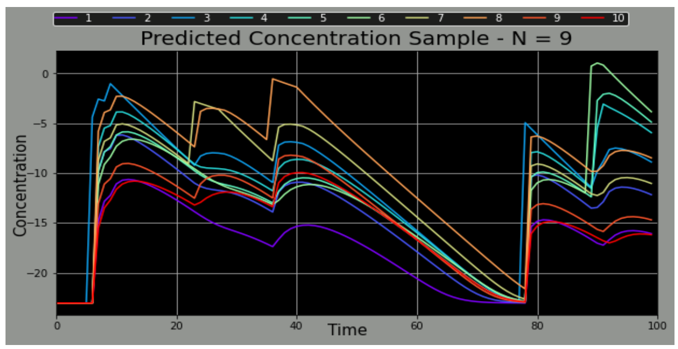
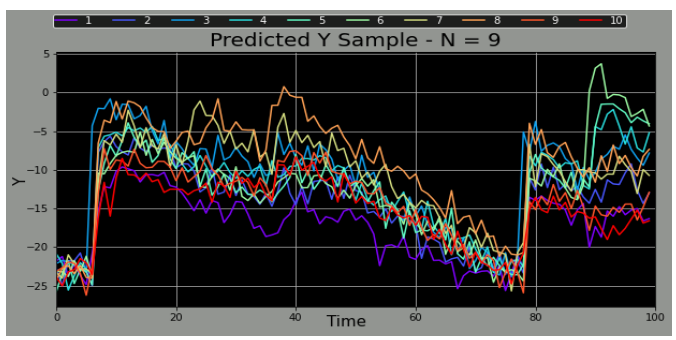
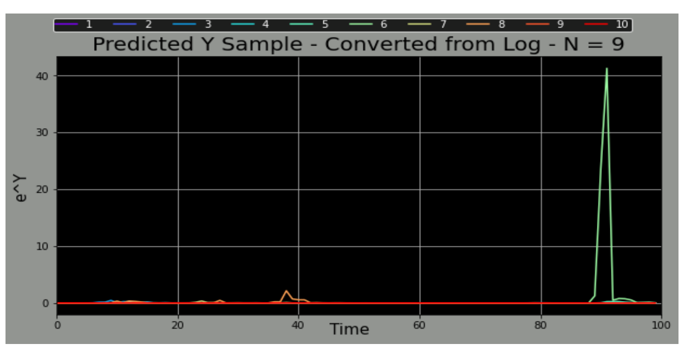
Results	Full Data
Mass	
Concentration	
Y	
e^Y	

Figure 42: Table of results for data (N = 9)

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	0.2790	0.2318	0.2357	0.0045
β	0.5999	0.6006	0.6008	0.0033
I	2, 6, 7, 2, 2, 2, 5, 3, 5	1.9990, 6.0000, 6.9990, 2.0038, 2.0001, 2.0003, 4.9974, 3.8108, 4.8138	2.0030, 6.0000, 7.0002, 2.0112, 2.0037, 2.0041, 4.9990, 3.8604, 4.8290	0.1019
T	9, 23, 36, 78, 7, 6, 89, 91, 90	8.9666, 22.9945, 35.9784, 77.6184, 7.5842, 6.0382, 88.9931, 74.8080, 90.7381	8.9922, 23.0063, 35.9984, 77.7984, 7.8626, 6.0382, 88.9971, 75.9220, 90.8145	0.0113
W	1, 5, 5, 1, 2, 2, 3, 1, 2	0.9211, 4.9976, 6.1584, 0.8399, 1.7098, 1.9261, 2.9857, 1.4442, 2.4470	0.9593, 5.0762, 6.2148, 0.8691, 1.7566, 1.9793, 3.0379, 1.4932, 2.1530	0.4487
S	4,3175, 0.4177, 0.6689, 0.1006, 0.9261, 0.1777, 4.1973, 0.1772, 2.3397	4.2841, 0.4424, 0.6621, 0.0987, 0.9227, 0.1847, 4.1714, 0.2087, 2.8449	4.3032, 0.4292, 0.6689, 0.0991, 0.9314, 0.1870, 4.2161, 0.2142, 2.9274	0.5679
τ	2.1676	2.1466	2.1548	0.2306

$Table\ 20\text{: } Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (N\ =\ 9)$

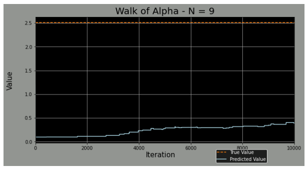
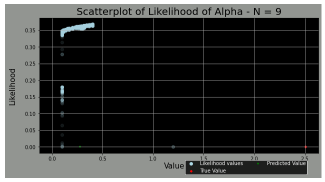
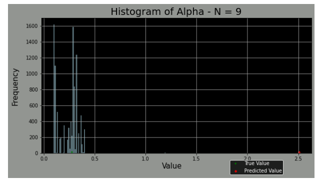
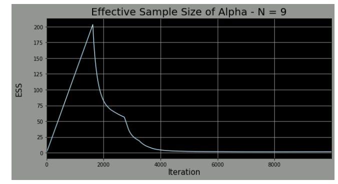
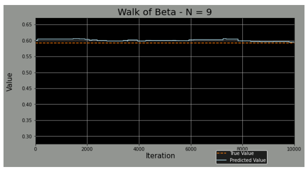
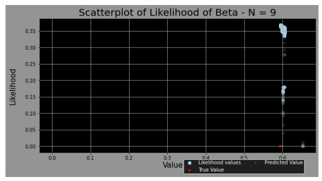
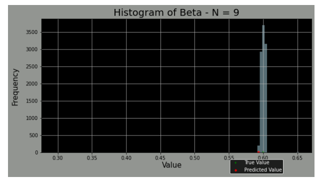
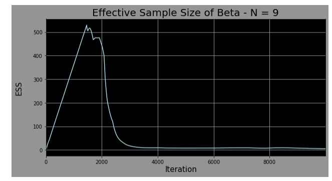
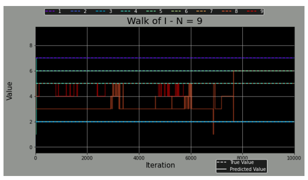
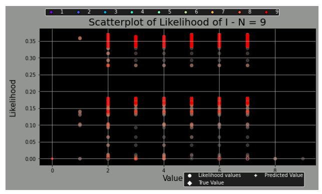
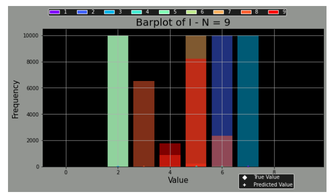
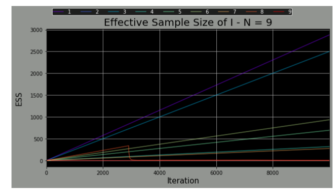
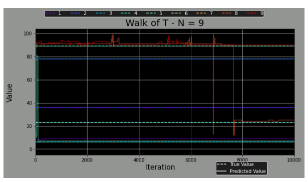
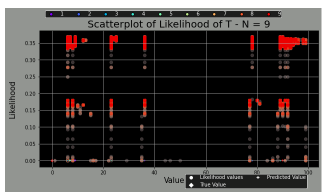
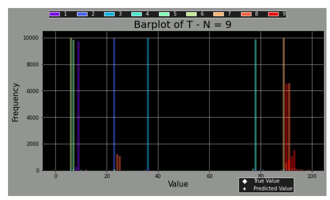
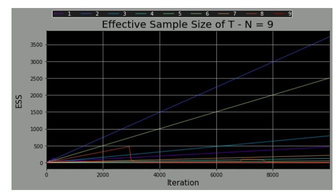
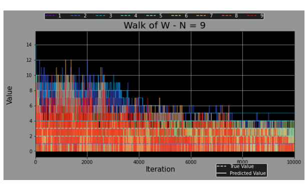
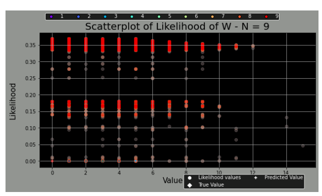
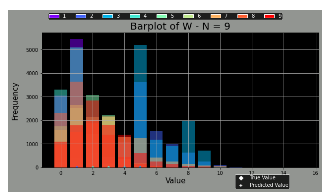
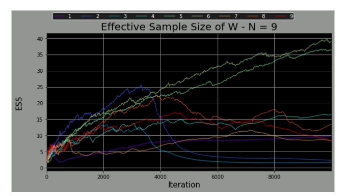
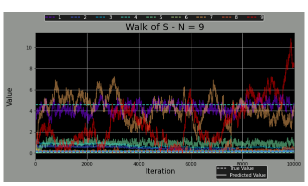
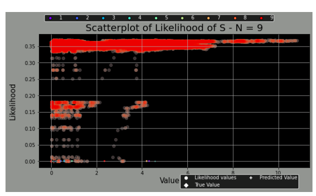
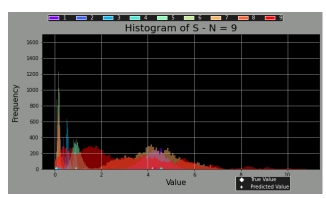
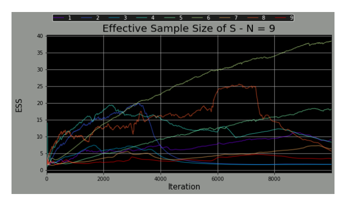
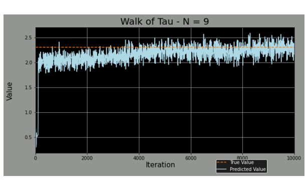
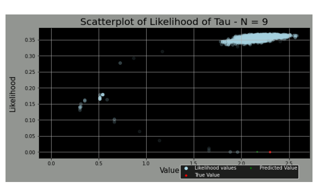
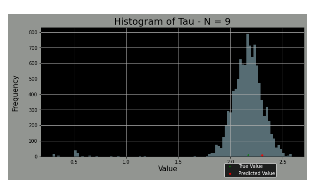
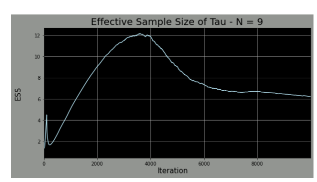
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 43: Table of graphs of parameter metrics for data $(N = 9)$

N = 10

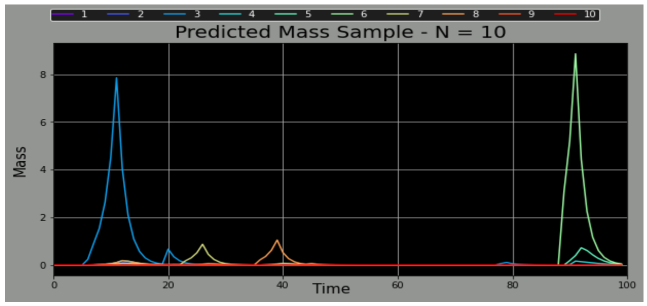
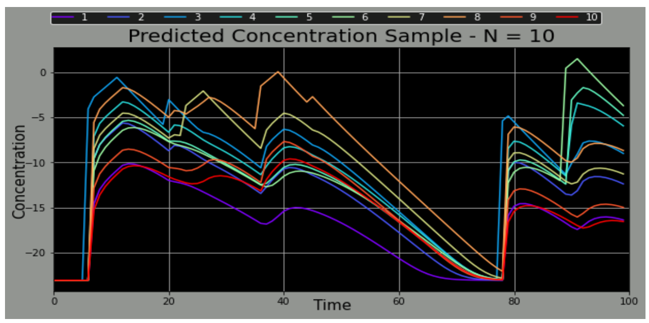
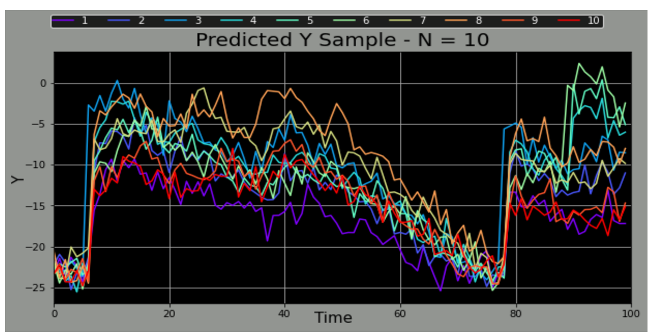
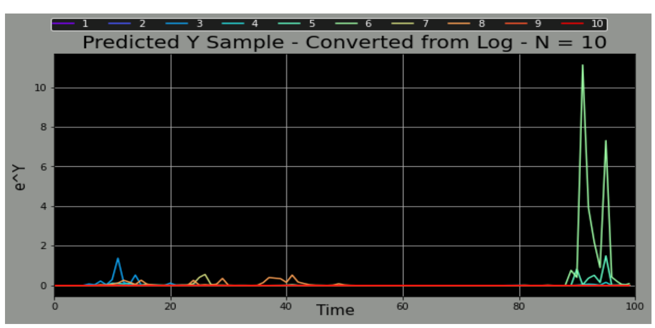
<table>
<tr><th>Results</th><th>Full Data</th></tr>
<tr><td>Mass</td><td>
</td></tr>
<tr><td>Concentration</td><td>
</td></tr>
<tr><td>Y</td><td>
</td></tr>
<tr><td>e^Y</td><td>
</td></tr>
</table>

Figure 44: Table of results for data $(N = 10)$

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	1.1975	1.1738	1.1864	0.0128
β	0.5660	0.5487	0.5502	0.0063
I	2, 2, 2, 7, 6, 3, 2, 2, 5, 7	2.4504, 2.0000, 3.0613, 6.9950, 5.9984, 4.7778, 2.0000, 3.2991, 4.9258, 7.0903	2.4732, 2.0000, 3.1409, 7.0000, 6.0008, 4.8538, 2.0000, 3.3537, 4.9358, 7.1421	0.1031
T	7, 78, 11, 36, 23, 91, 6, 20, 89, 45	42.4942, 77.9880, 24.1065, 36.0054, 23.0545, 70.5421, 5.9991, 51.1260, 89.4992, 41.1573	44.1200, 77.9922, 24.8001, 36.0116, 23.0683, 71.4045, 6.0027, 52.6342, 89.5561, 41.5153	0.0136
W	1, 2, 1, 4, 4, 1, 5, 1, 3, 1	1.7355, 2.1430, 1.2912, 3.9217, 3.8189, 1.5532, 3.5158, 1.1166, 2.7046, 0.8600	1.7927, 2.1666, 1.3352, 3.9381, 3.8397, 1.6070, 3.5836, 1.1570, 2.7482, 0.8814	0.3099
S	0.4885, 0.0643, 0.1065, 0.2113, 0.1618, 0.1092, 0.2380, 0.5931, 3.0782, 0.0030	0.4923, 0.0680, 0.4568, 0.2159, 0.1720, 0.1346, 0.3284, 1.0591, 2.8432, 0.0339	0.5034, 0.0690, 0.4790, 0.2178, 0.1747, 0.1442, 0.3359, 1.1055, 2.8870, 0.0392	0.3766
τ	2.1647	1.9959	2.0131	0.2242

Table 21: Table of summary statistics of parameters for data $(N = 10)$

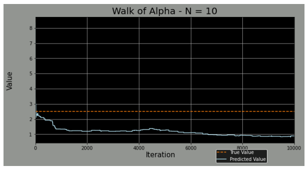
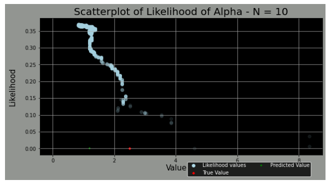
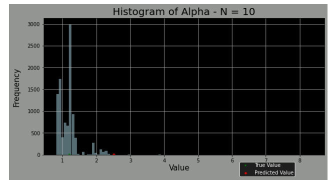
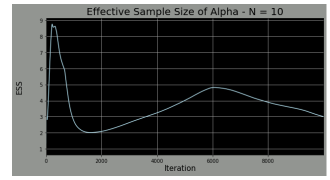
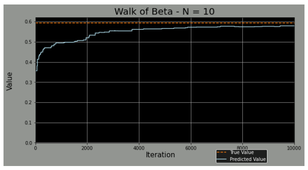
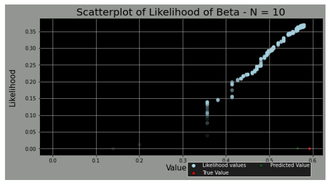
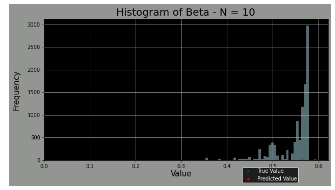
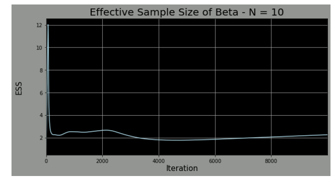
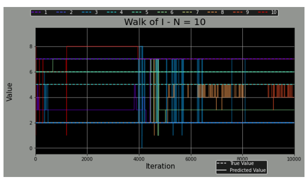
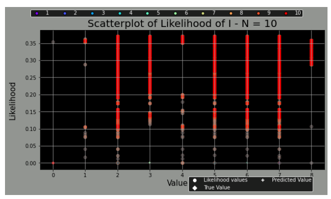
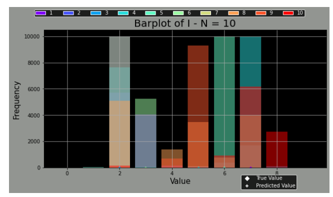
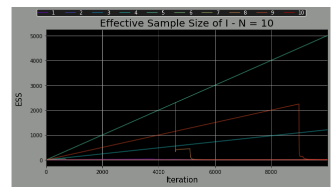
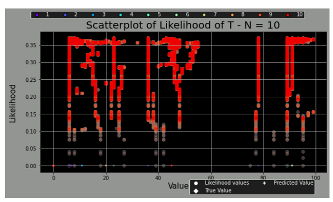
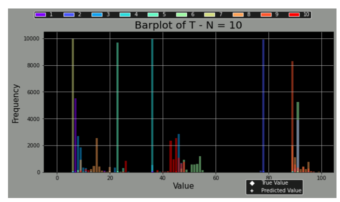
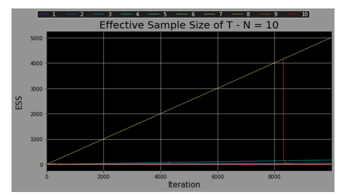
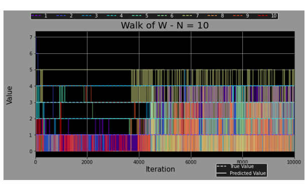
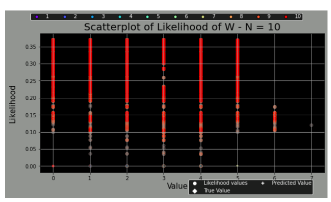
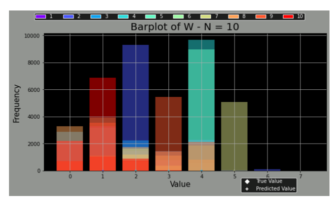
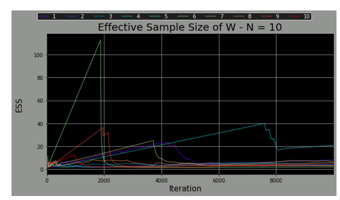
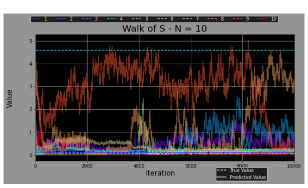
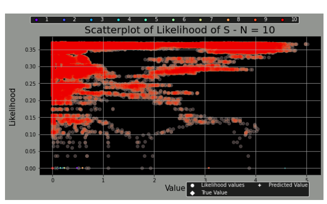
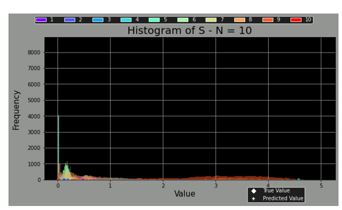
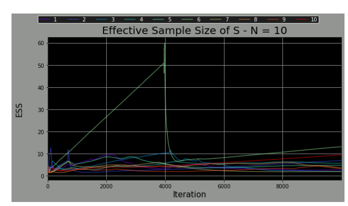
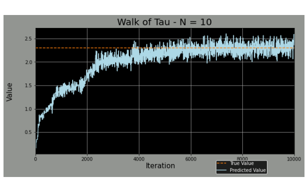
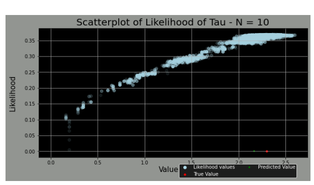
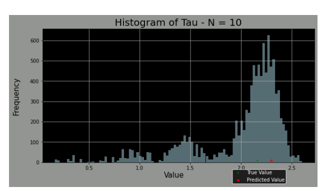
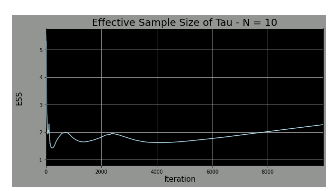
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 45: Table of graphs of parameter metrics for sparse data ($N = 10$)

Removing Source Sites

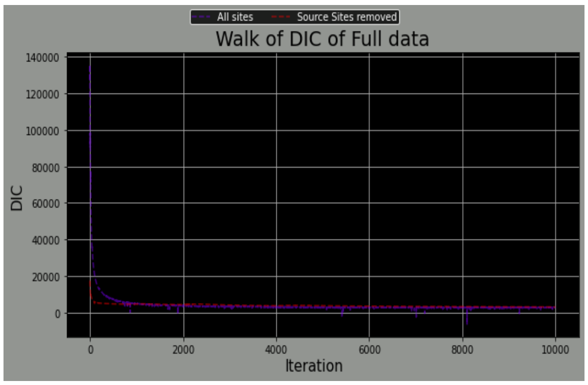
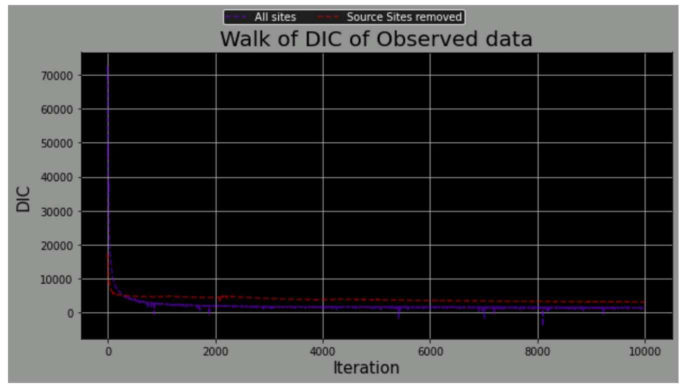
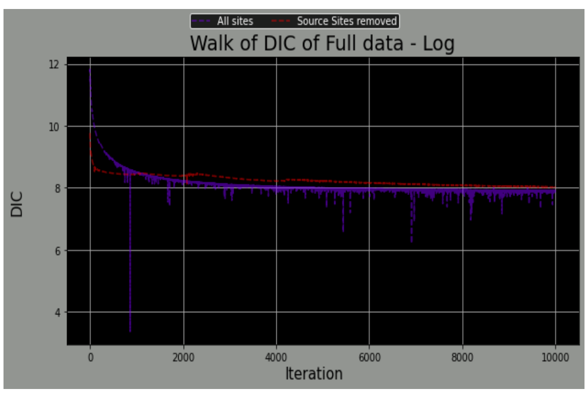
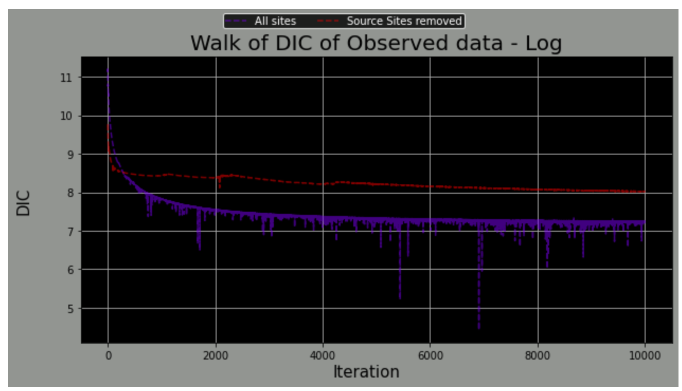
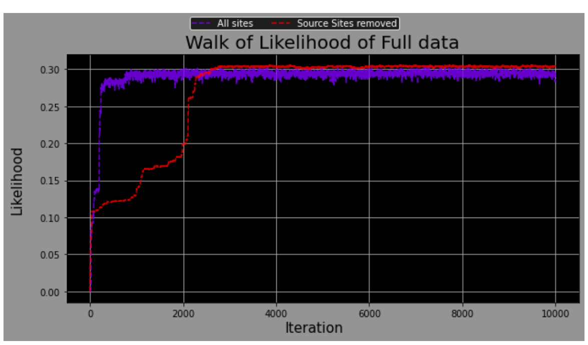
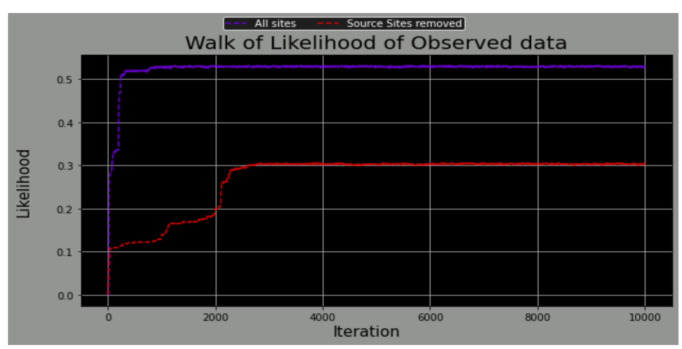
Model Comparison Metrics	Full Data	Observed Data
DIC		
DIC - Log		
Likelihood		

Figure 46: Table of Model Comparison metrics for models with different sites observed

All Sites

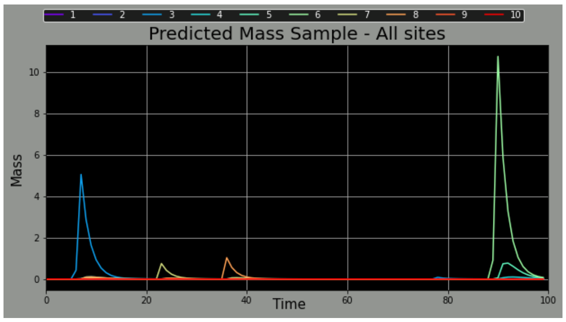
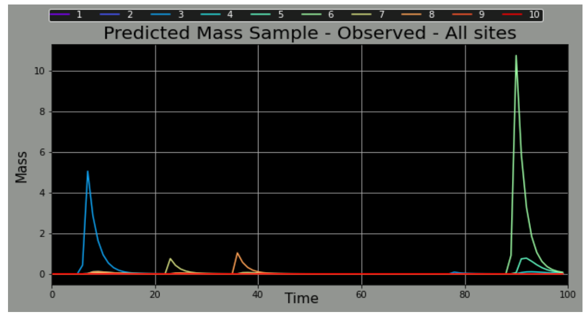
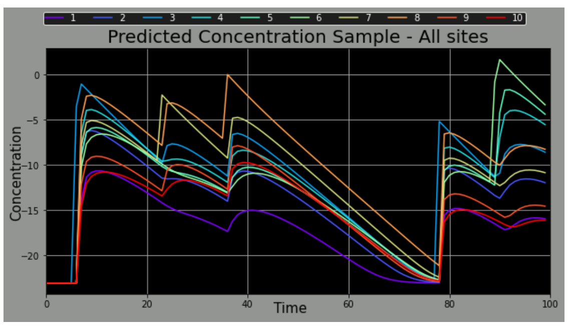
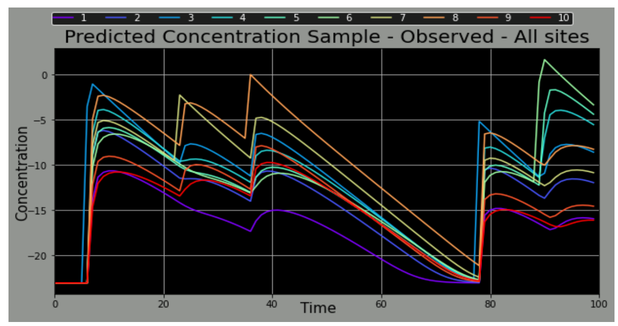
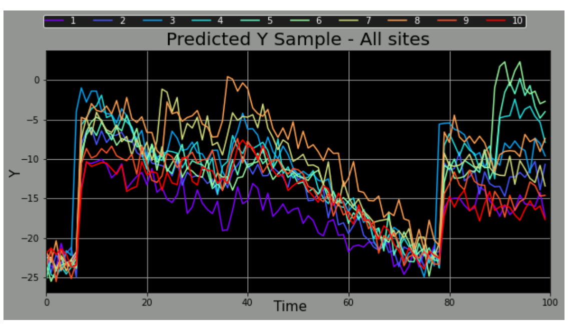
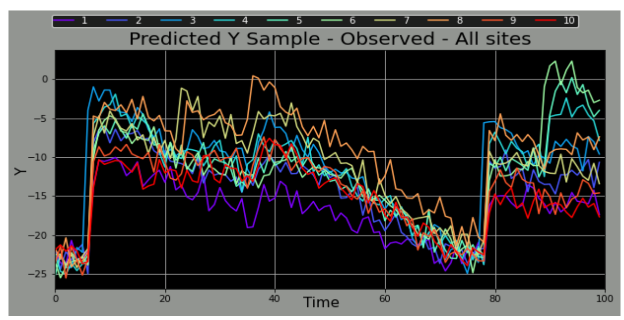
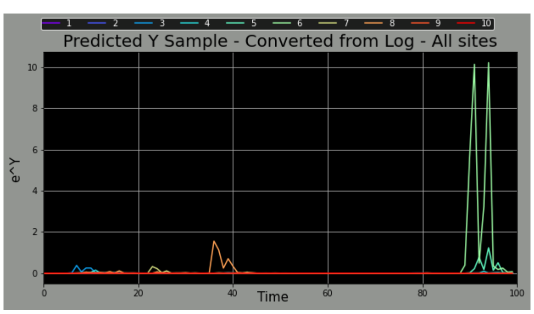
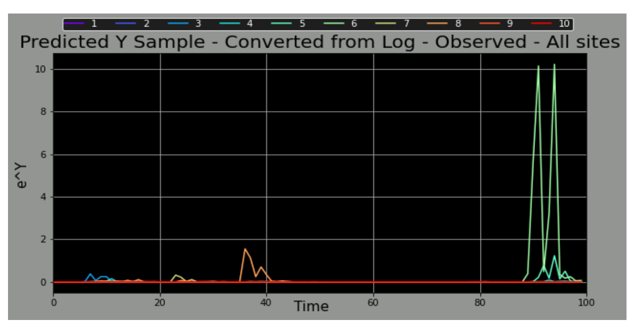
Results	Full	Observed
Mass		
Concentration		
Y		
e^Y		

Figure 47: Table of results for data (All Sites)

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	11.3993	10.3854	10.5535	0.3539
β	0.6213	0.6469	0.6491	0.0070
I	2, 6, 5, 2, 7	2.0009, 5.7385, 4.9984, 2.0004, 6.7876	2.0040, 5.7639, 5.0000, 2.0040, 6.8038	0.1011
T	6, 23, 89, 78, 36	6.0493, 34.2564, 88.9611, 77.9956, 37.1762	6.1631, 35.2894, 89.0059, 78.0020, 32.4912	0.0114
W	2, 1, 2, 1, 1	1.9126, 0.6523, 1.9583, 0.6040, 0.5810	1.9264, 0.6799, 1.9741, 0.6232, 0.6006	0.3882
S	0.4232, 0.7434, 0.9016, 0.0813, 1.0226	0.4101, 0.6891, 0.9182, 0.0712, 1.0098	0.4142, 0.6986, 0.9264, 0.0722, 1.0206	0.4333
τ	1.5388	1.3000	1.3194	0.1470

$Table\ 22\text{:}\ Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (All\ Sites)$

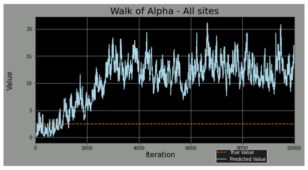
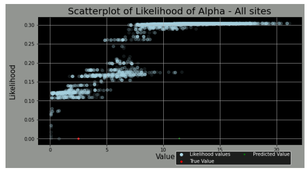
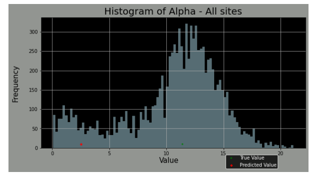
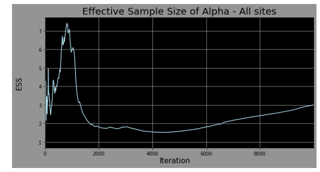
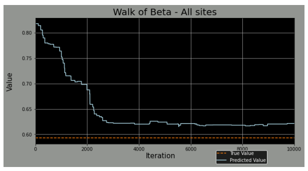
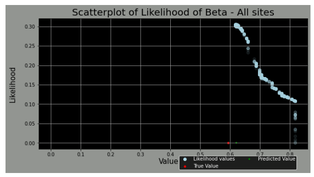
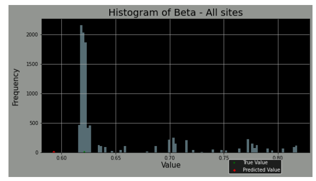
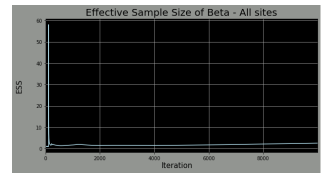
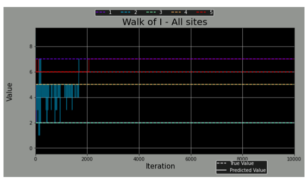
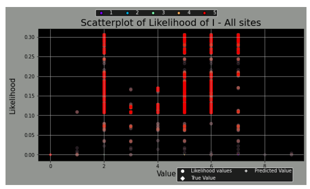
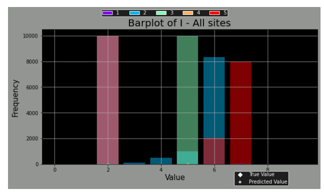
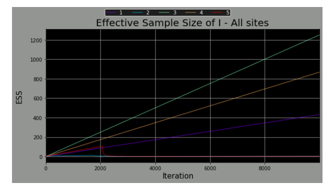
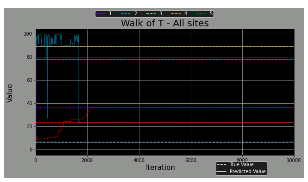
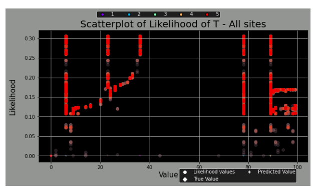
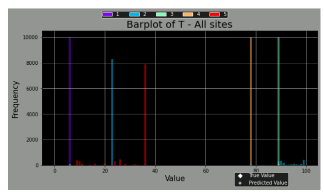
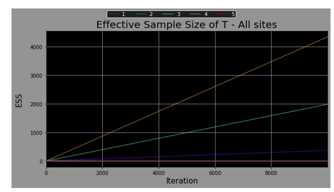
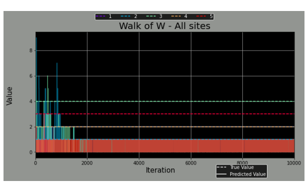
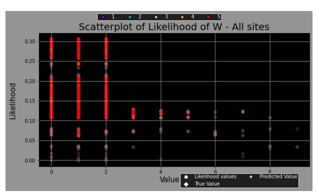
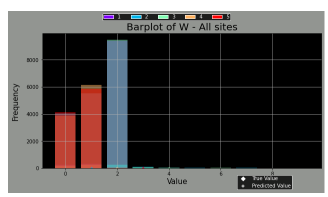
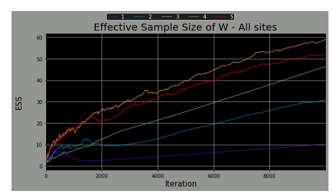
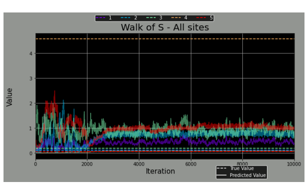
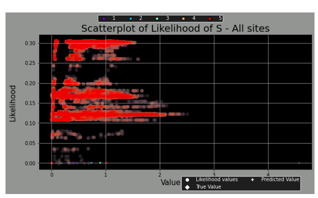
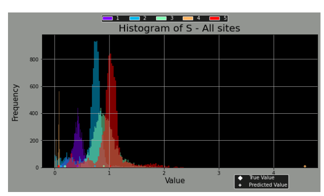
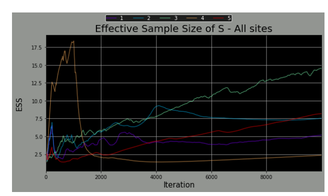
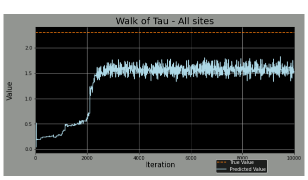
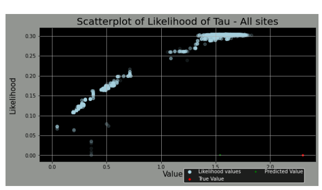
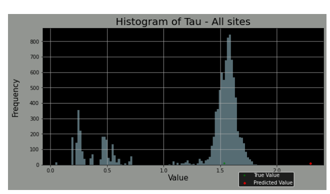
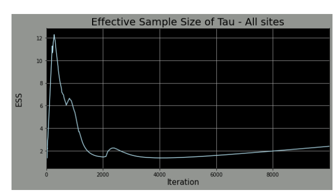
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 48: Table of graphs of parameter metrics for data (All Sites)

Source Sites Removed

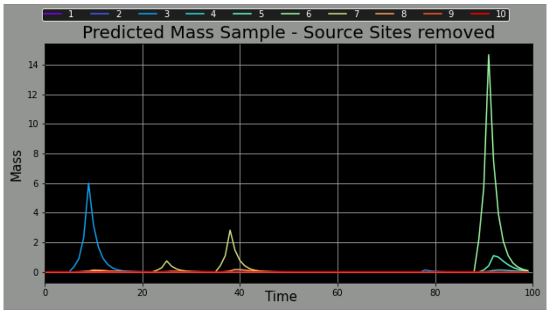
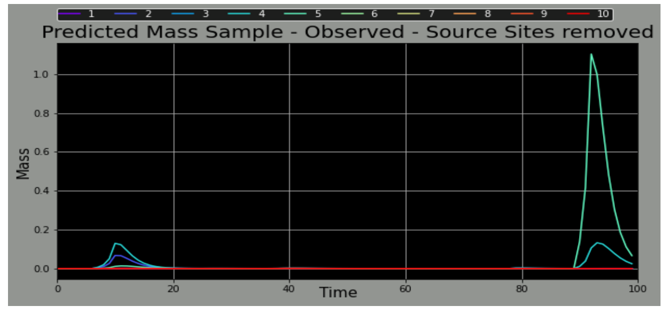
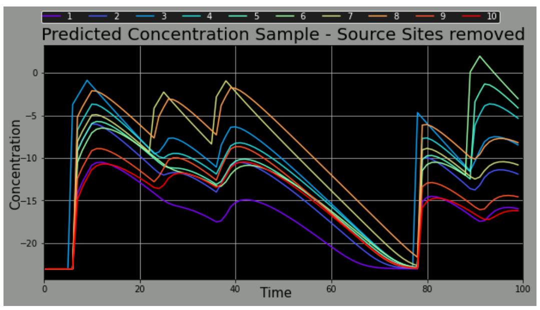
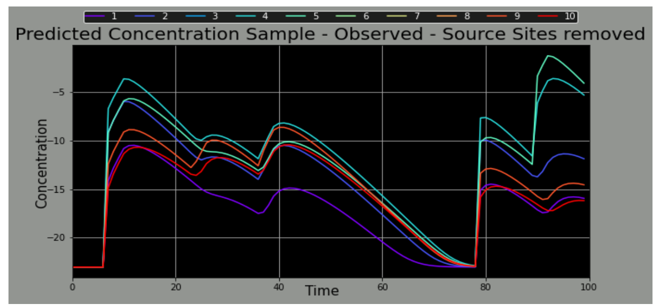
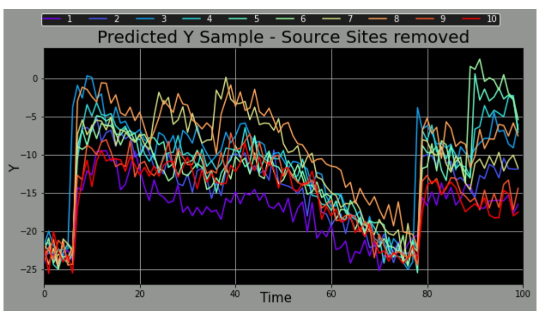
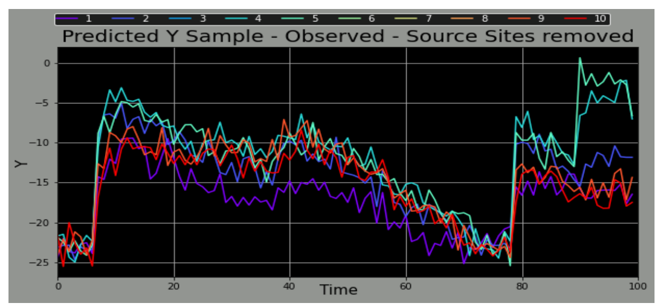
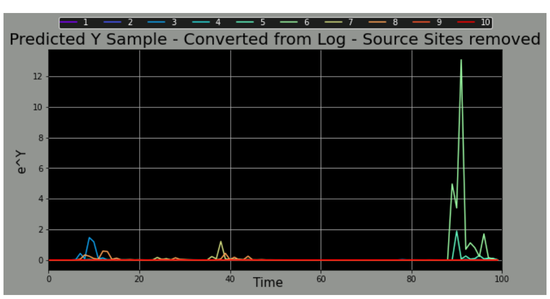
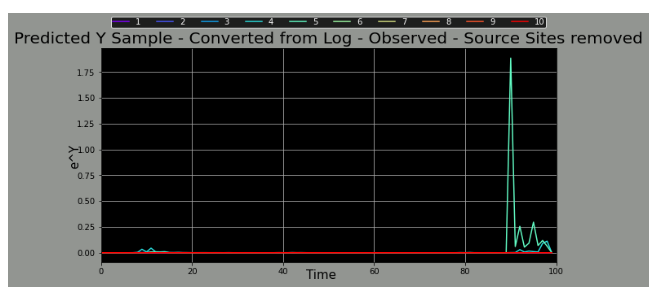
Results	Full	Observed
Mass		
Concentration		
Y		
e^Y		

Figure 49: Table of results for data (Source Sites Removed)

Parameter	Median	95% CI - Lower Bound	95% CI - Upper Bound	P(Acceptance)
α	2.0512	2.0678	2.0750	0.0275
β	0.5804	0.5773	0.5779	0.0053
I	6, 2, 5, 6, 2	6.0000, 2.0004, 4.9950, 5.9981, 2.0225	6.0006, 2.0016, 4.9976, 6.0003, 2.0359	0.1003
T	36, 78, 89, 23, 6	36.0359, 78.0181, 88.9866, 23.0133, 5.9969,	36.0477, 78.0414, 88.9978, 23.0463, 6.0011	0.0103
W	3, 1, 3, 3, 4	3.0019, 0.5810, 3.0009, 2.9925, 4.0091	3.0041, 0.6006, 3.0140, 2.9959, 4.0161	0.2553
S	0.4213, 0.1333, 2.2327, 0.1112, 0.3485	0.4437, 0.1389, 2.2331, 0.1209, 0.3527	0.4513, 0.1405, 2.2450, 0.1247, 0.3553	0.3396
τ	2.029	1.996	2.006	0.2794

$Table\ 23\text{:}\ Table\ of\ summary\ statistics\ of\ parameters\ for\ data\ (Source\ Sites\ Removed)$

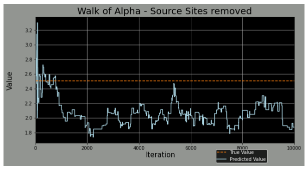
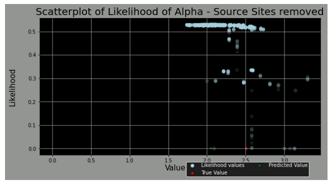
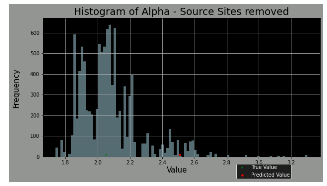
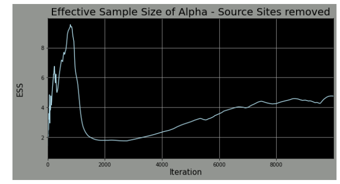
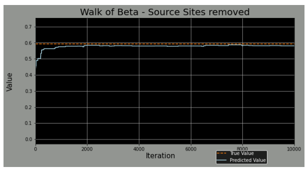
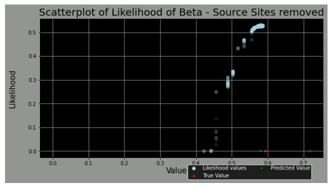
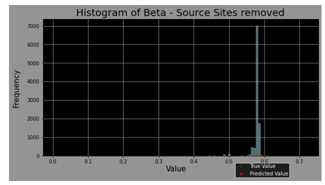
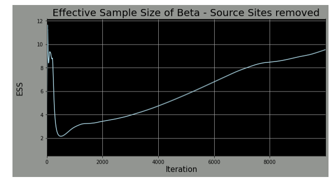
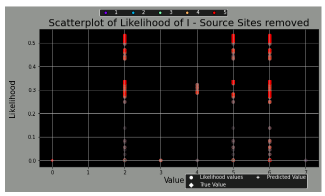
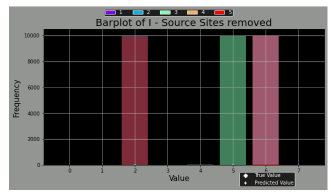
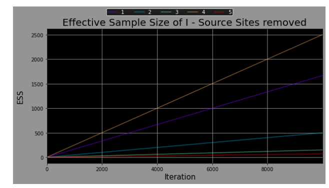
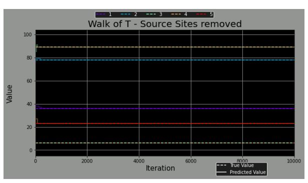
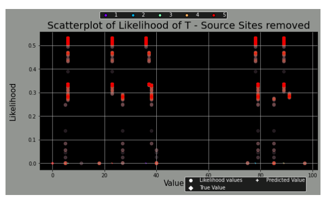
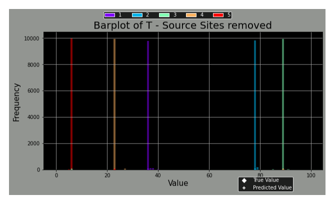
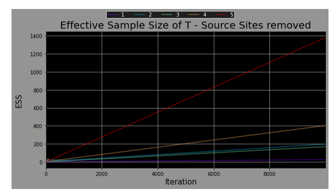
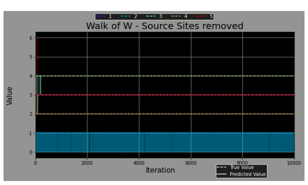
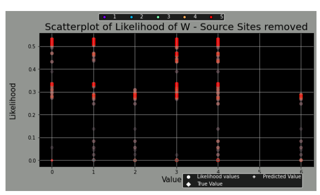
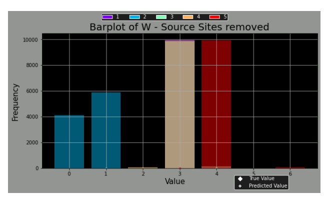
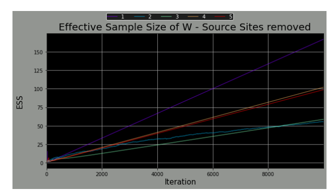
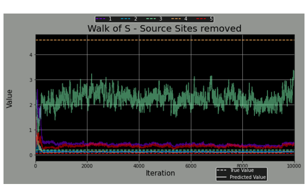
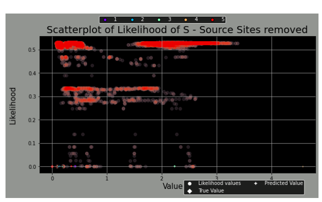
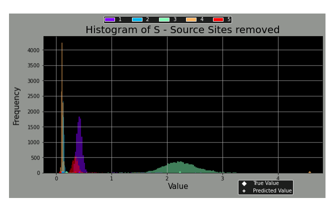
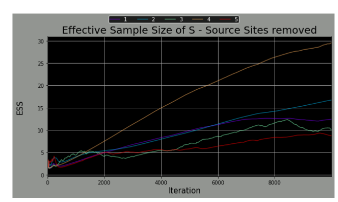
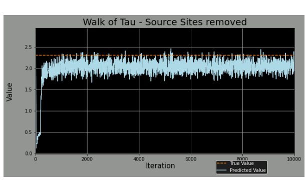
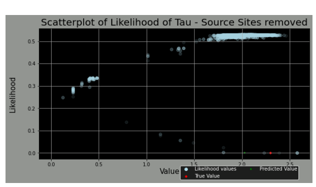
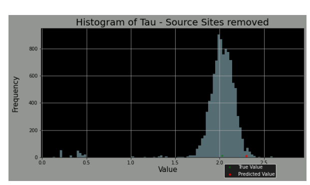
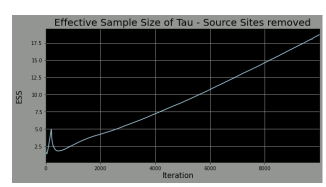
Parameter	Walk of value	Scatter Plot of Likelihood	Histogram	Walk of ESS
α				
β				
I				
T				
W				
S				
τ				

Figure 50: Table of graphs of parameter metrics for data (Source Sites Removed)

Discussion

These results provide a number of interesting aspects regarding the implementation of Metropolis-Hastings to this model, but first and foremost they show that the concept does indeed work for this problem - the ‘base’ models (Where $p = 0$, $N = 5$ and $All\ sites$ included) are able to identify the starting times and sites of the events, with only $N = 5$ misidentifying a single event. However, considering the inherent randomness of this method with sampling possible values for parameters, each model had some noise in regards to how well it performed that doesn’t correctly estimate the true values of the parameters, but estimated the true data correctly.

Generally, the models were best at identifying I and T and could all do so with a relatively high accuracy. This makes sense considering that the features associated with these parameters are pretty obvious to spot; for a particular event k, I_k is the site that will experience a sharp increase at time T_k. β was also predicted well, which also makes sense since it can be seen as the rate of decrease once the values actually start to decrease after an event has finished, and the only other parameter interacting with it is the error, τ.

α, W, S had less luck with being able to determine the true values. Part of this is simply due to the nature of the data ($Table\ 1$)- the error around event 3 tends to reflect a longer event with a smaller growth rate than what is actually true. Additionally, these parameters all interact with each other; for instance, during an event k, the model might predict a lower α but compensate with a larger S_k, and vice versa. However, it still managed to model the correct outputs.

τ tended to perform better than α, W, S but worse than β, I, T. Interestingly, the models that overfitted to a subset of the data (Particularly models with high p values) predicted a lower τ compared to other models with less overfitting. This is likely because the model will fit the other parameters to best reflect the true data with the error, which necessarily minimises the error of this predicted data and so τ is minimised; and for overfitted models, this difference between the predicted Concentration and the true Y will be significantly smaller.

In regards to the DIC metric, changing the sparseness ($Figure\ 4$) or number of events ($Figure\ 25$) didn't necessarily result in a *significant* fluctuation in DIC.

For changing sparseness, the general trend was that p increased, the amount of overfitting to the observed set of data increased. However, the fit on the full data surprisingly didn't decrease in a similar fashion like one might expect with full data. Instead, $p \leq 0.6$ all reached a similar DIC and likelihood on the full data, and only once $p \geq 0.7$ did we see worse results; particularly, both DIC and likelihood on the full datasets were both worse and subject to a higher variance.

This may be due to the method that we removed data; for each separate model, each $Y_{i,t}$ had a $100p\,\%$ chance of being removed, and this was done randomly both over i, t but also the separate p values - A higher p value may have had data that the lower ones didn't necessarily have. Additionally, not all data points are created equal; a subset that only included all sites at all times T_k, $T_k + W_k$ should be able to identify the parameters with relatively good accuracy, which would only be 10% of our dataset.

In regards to changing the number of events, the general trend was that as N increased, the models performed better. Particularly, as $N \geq N_{True}$ (5) the models were able to identify all the events. This makes sense, as a model where $N < 5$ is physically unable to identify all 5 events, and a model where $N > 5$ can identify events with either a minimal S_k such that it has no effect, or fit it to some of the noise in the data; for this data in particular, models with a higher N identified events around $T = 90$, because the error tends to suggest a sharper increase than would be expected simply due to the transfer.

As a result, for fitting with unknown N one approach might simply be to run this with a high N and pare back until a significant dropoff is detected. However, this would be a computationally expensive approach; for each additional event, this model took roughly 20 minutes longer to complete.

In regards to removing the source sites, this model was also able to identify the parameters with a similar accuracy to the model with the full dataset, correctly identifying $\frac{4}{5}$ of I_k ($Table\ 23$). However, it took *longer* for the model to reach the same likelihood on the full dataset, taking roughly 2000 more iterations to reach a likelihood of 0.3 than the baseline model ($Figure\ 46$)

Conclusion

To conclude, these results have shown that this method of approximating the model for our algae blooms is a successful one. Additionally, this method was able to work with a significantly smaller observed subset (Both with random data removed, and with removing the source sites for blooms), and was also able to generally model the true number of events. However, as useful as this proof of concept may be, there are some possible improvements that could be implemented with this algorithm.

Firstly, for all parameters the Effective Sample Size tended to be quite low across all models, reflecting a large amount of autocorrelation between the parameter iterations, θ_s. Therefore, the next step could be to implement burn-in[7] (Removing the first X iterations such that the chain is less dependent on the initial value θ_0) or thinning[7] (Only selecting every k-th value for θ such that the correlation between θ_k and θ_{k-1} is reduced).

Additionally, this project implemented a rather rough way of trying to determine the true number of events, by simply repeating the estimation with an increasing N value and comparing the results. Since having N being an estimated parameter would entail having an *unknown* number of parameters to actually estimate, this would be a significantly more complicated algorithm to achieve. One method of doing this may be to implement a reversible jump simulator[8].

References

1. Rhodes, L., Munday, R. (2016). Harmful Algae and Their Commercial Implications. In: Bux, F., Chisti, Y. (eds) Algae Biotechnology. Green Energy and Technology. 301-315. https://doi.org/10.1007/978-3-319-12334-9_15
2. Beston, A. (2000). Toxin puts mussel industry in danger as algal bloom spreads south. *New Zealand Herald*. Toxin puts mussel industry in danger as algal bloom spreads south - NZ Herald
3. Paerl, H. W., Huisman, J. (2008). Blooms Like it Hot. *Science 320*(5282) 57-58. https://doi.org/10.1126/science.1155398
4. Paerl, H. W., Hall, N. S., Calandrino, E. S. (2011). Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. *Science of the Total Environment 409*(10) 1739-1745. https://doi.org/10.1016/j.scitotenv.2011.02.001
5. P.M. Glibert et al. (eds.), Global Ecology and Oceanography of Harmful Algal Blooms, Ecological Studies 232, https://doi.org/10.1007/978-3-319-70069-4_5
6. Vennell, R., Scheel, M., Weppe, S. *et al.* (2021) Fast lagrangian particle tracking in unstructured ocean model grids. *Ocean Dynamics 71.* 423–437. https://doi.org/10.1007/s10236-020-01436-7
7. Sibanda, N. (n.d.) Computing and MCMC Diagnostics. [PowerPoint slides]
8. Green, P., Hastie, D. (2009). Reversible jump MCMC. *Genetics 155*.