Robust Inference for Generalized Partially Linear Mixed Models subject to censored responses and missing covariates—an Arctic Data Analysis

Kalyan Das
Department of Statistics, University of Calcutta, Kolkata, INDIA

Angshuman Sarkar
Novartis India Ltd., Hyderabad, INDIA
Dependence of the Growth of Phytoplankton on Water Salinity
Background

- **Phytoplanktons** photosynthesizing microscopic organisms
- Responsible for **half of the total amount of oxygen** produced by all plants present in the Earth's atmosphere.
- Serve as the foundation of the aquatic food web, feeding everything from microscopic, animal-like zooplankton to multi-ton whales.
- **Phytoplankton** can be the harbingers of death or disease. Certain species of phytoplankton produce powerful biotoxins called “red tides,” or harmful algal blooms. These toxic blooms can kill marine life and people who eat contaminated seafood.
Dead fish washed onto a beach at Padre Island, Texas, in October 2009, following a red tide (harmful algal bloom).

In spring, they multiply rapidly, in what is referred to as a bloom, to take advantage of the warmer, more favorable conditions. In fall, their numbers drop as the water temperature decreases.
A picturesque phytoplankton bloom near Ireland observed by NASA's Terra satellite
Objective

Now a days, due to the overwhelming presence of various types of pollutants in the water bodies, its salinity level may increase over time. **Objective is to study the extent to which water salinity influences the cell concentration of the phytoplankton named *Thalassiosira Nordenskioeldii***
**Thalassiosira nordenskioeldii**

**Length:** 0.01-0.05 mm, **Habitat:** Cold Water,
**Diet:** phytosynthetic **Morphology:** yellow brown in color, not green
**Reproduction:** general for group: both sexual and asexual
Road Map

- An Arctic Data
- GPLMCM
- Methods for Inference
  i) Exact MCEM based approach
  ii) Approximate approach
- Simulation Study
- Data Analysis
An Arctic Data

Phytoplankton cell concentrations (cells/ l) and relative abundance (% total phytoplankton density) in water samples collected from different stations in Svalbard, Norway during blooming time early spring 2010.
Most phytoplankton are too small to be individually seen with the unaided eye. While observing the count of Phytoplankton, observations less than a certain value, are approximated. Thus the dataset is left-censored.
Experiment and Objective

- **Experiment**: 100 ml. water collected from each sample taken and the phytoplankton named *Thalassiosira Nordenskioeldii* detected. Cell densities, relative abundance, temperature, salinity etc. were observed. Across 37 stations, the cell count of *T. Nordenskioeldii* varied considerably due to ice cover in the surface water column. The maximum cell count was 256 and minimum had been censored to 25.

- **Main focus**: To what extent the salinity influences the cell concentration of *Thalassiosira Nordenskioeldii*?
## Data in hand

<table>
<thead>
<tr>
<th>Count</th>
<th>Depth</th>
<th>Temperature</th>
<th>Salinity</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>139</td>
<td>2.2 m</td>
<td>1.3 deg c</td>
<td>34.54</td>
<td>2:38:40</td>
</tr>
</tbody>
</table>

Count are subject to left censoring
Boxplot showing count of Thalassiosira nordenskioeldii in 8 different stations.
Count of Thalassiosira Nordenskioeldii at 3.2 Mitter

Count of Thalassiosira Nordenskioeldii vs Depth

Depth in 0.2 Meters
How to analyse such data?

Imp. Issues to take into account

i) Count varies with temp. through some unknown function
ii) Count below 25 / 100ml censored
iii) A few counts tremendously huge (outliers)
iv) Some covariates missing
Analyse through GPLMCM

What is GPLMCM?

Outcome $y_{ij}$ measured on $i$th subject at time point $t_{ij}$.
May be Discrete or Continuous

- **GLM:** $\mathbb{E}(y_{ij} | x_{ij}) = \mu_{ij}$

  We consider a covariate vector $x_{ij} \in \mathbb{R}^p$ observed for the $i$th subject at time $t_{ij}$.

  $\eta_{ij} = g(\mu_{ij}) = x'_{ij} \beta$ - A given link function

- **GPLM:** $g(\mu_{ij}) = x'_{ij} \beta + f_0(t_{ij})$

  $-f_0$ a smooth function (parameter free)
We approximate $f_0(t)$ by $\pi(t) \alpha$

where $\pi(t) = (B_1(t), \ldots B_N(t))'$ vector of $B$-spline basis functions.

GPLLMM:-

$$g(\mu_{ij}^{b_i}) = x'_{ij} \beta + z'_{ij} b_i + \pi'_{ij} \alpha$$

$b_i$ subject specific effect (random)$\sim N_q(0, D(\nu))$

GPLMCM:-

$$y_{ij} = y^*_{ij} \text{ if } c_{ij}^* = 0 \text{ i.e. } y^*_{ij} > d$$
$$= d \text{ if } c_{ij}^* = 1 \text{ i.e. } y^*_{ij} \leq d$$

Left censored - $d$ detection limit
GPLMCM with missing covariates

We write $x_i = (x_{i1}, \ldots x_{in_i})' = (x_{mi}, x_{oi})$
where $x_{mi}$ corresponds to the missing components
$x_{oi}$ contains the observed components of $x_i$.

**Purpose:** Carry out Robust analysis to guard against outliers.

Data: $\{y_{ij}^*, c_{ij}^*, x_{ij}, z_{ij}, i = 1, \ldots, m, j = 1, \ldots, n_i\}$
Likelihood (no missing covariates)

\[ L = L(\theta | y_{i}^{*}, c_{i}^{*}, x_{i}, z_{i}) \]

\[ = \sum_{i=1}^{m} \left[ \int \prod_{j=1}^{n_{i}} \left\{ f(y_{ij}^{*} | x_{0ij}, b_{i}) \right\}^{1-c_{ij}^{*}} \left( F(d | b_{i}, x_{ij}) \right)^{c_{ij}^{*}} \right] g(b_{i}) db_{i} \]

where \( \theta = (\beta', \alpha', \nu')' \)

\[ y_{i}^{*} = (y_{i1}^{*} \ldots y_{in_{i}}^{*})', \quad c_{i}^{*} = (c_{i1}^{*} \ldots c_{in_{i}}^{*})' \]

\[ f(y_{ij} | x_{ij}, z_{ij}, b_{i}) = \exp \{ a^{-1}(\phi) \left( y_{ij} \delta_{ij} - h(\delta_{ij}) \right) + c(y_{ij}, \phi) \} \]

\[ \delta_{ij} = \delta(\eta_{ij}), \phi = 1, F(d | x_{ij}, b_{i}) = P(y_{ij} \leq d | x_{ij}, z_{ij}, b_{i}) \]
ML Estimation- MCEM approach

**EM**: Complete Data \((y_i, x_i, b_i, i = 1, \ldots, m)\) where missing data

- \(y_{ci}\) non-detectable components of \(y_i\).
- \(x_{mi}\) missing components of \(x_i\).
- \(b_i\) unobserved random components.
The complete data Likelihood

The complete data likelihood is

\[ L_C = \prod_{i=1}^{m} \{ f(y_i \mid x_{oi}, x_{mi}, b_i) f(x_{mi} \mid \gamma) f(b_i \mid \nu) \} \]

Then at \((t + 1)\)st EM iteration, the \(E\)-step for individual \(i\) is

\[ Q_i(\theta \mid \theta^{(t)}) = E_{y_{ci}, x_{mi}, b_i} \left[ l_c(\theta, y_i, x_i, b_i) \mid y^*_i, c^*_i, x_{oi}, \theta^{(t)} \right] \]

where

\[ \sum_{i=1}^{m} l_c(\theta; y_i, x_i, b_i) = \log L_C \]
Monte Carlo EM

Gibbs sampler to generate samples from $[y_{ci}, x_{mi}, b_i | z_{oi}, y_i^*, c_i^*, \theta^{(t)}]$

$$f(y_{ci} | x_i, b_i, y_i^*, c_i^*, \theta^{(t)}) \propto f(y_{ci} | x_i, b_i, \theta^{(t)}) f(y_i^*, c_i^* | x_i, y_{ci}, b_i, \theta^{(t)})$$

$$f(x_{mi} | x_{oi}, y_i, b_i, y_i^*, c_i^*, \theta^{(t)}) \propto f(x_{mi} | \theta^{(t)}) f(y_i^*, c_i^* | x_i, y_{ci}, b_i, \theta^{(t)})$$

$$f(b_i | x_i, y_i, y_i^*, c_i^*, \theta^{(t)}) \propto f(b_i | \theta^{(t)}) f(y_i^*, c_i^* | x_i, y_{ci}, b_i, \theta^{(t)})$$

$$Q(\theta | \theta^{(t)}) = \sum_{i=1}^{m} Q_i(\theta | \theta^{(t)})$$

$$= \sum_{i=1}^{m} \left\{ \frac{1}{M_i} \sum_{l=1}^{M_i} l_c(\theta; y_i^*, c_i^*, x_{oi}, y_{ci}^{(l)}, x_{mi}^{(l)}, b_i^{(l)}) \right\}$$

The $M$-step of the Monte Carlo EM (MCEM) algorithm maximizes $Q(\theta | \theta^{(t)})$ to produce an updated estimate $\theta^{(t+1)}$. 
Robust Method
Robust Method

Why? MLE Sensitive to outliers
small fraction of responses from an arbitrary heavy tailed distribution.
Maximize
\[ Q^R(\gamma|\theta^{(t-1)}) = \sum_{i=1}^{m} E \left\{ l^R_c(\gamma|(y_i, x_{0i}, b_i, \theta_0^{(t-1)})) \right\} \]
\[ = \sum_{i=1}^{m} \sum_{i=1}^{n} \rho \left( \frac{y_{ij}^*-x'_{ij} \beta - \pi'_{ij} \alpha - z'_{ij} b_i}{\sigma} \right) \]
\( \rho \) corresponds to the Huber’s least favourable distribution for location:
\[ \rho(r) = \begin{cases} 
\frac{1}{2} r^2 & \text{if } |r| \leq a \\
 a |r| - \frac{1}{2} a^2 & \text{otherwise.} 
\end{cases} \]
Note that the derivative of the \( \rho \) function leads to the
Huber’s \( \psi \)-function \( \psi_a(r) = \max(-a, \min(r, a)) \).
Robust version of the score function

\[ E_{y_{ci}, x_{mi}, b_i} \left[ \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left\{ \psi_1 \left( \frac{y_{ij} - \mu_{ij}^b}{\sigma_{ij}^b} \right) \sigma_{ij}^b w(x_{ij}) x_{ij} - q_{ij}^{(\theta, b_i)} \right\} \right| y_{i}^*, x_{0i}, c_i \right] = 0 \]

where

\[ \mu_{ij}^b = g^{-1}(x_{ij}^l \beta + z_{ij}^l b_i + f_0(t_{ij})) \]

\[ q_{ij}^{(\theta, b_i)} = \frac{1}{n} \left[ \psi_2 \left( \frac{y_{ij} - \mu_{ij}^b}{\sigma_{ij}^b} \right) \right| y_{ci}, x_{mi}, b_i \left] \sigma_{ij}^b w(x_{ij}, \pi_{ij}) x_{ij} \right. \]
More generally, Robust version of the score function leads to

\[ \sum_{i=1}^{m} E_{y_{ci}, x_{mi}, b_i} \left[ G_i' W_i(\theta, b_i) h_i(\mu_i^{b_i}, \theta) | y_i^*, c_i^* \right] = 0 \]

where \( G_i = (x_i', \pi_i') \),

\( j \)th diagonal element of \( W_i \) is \( \sigma_{ij}(\theta, b_i) w(x_{ij}, \pi_{ij}) \),

\( j \)th diagonal element of \( \cdot h_i(\mu_i^{b_i}, \theta) \) is \( \psi_{1c_0} \left( \frac{y_{ij} - \mu_i^{b_i}}{\sigma_{ij}} \right) - v(\theta; b_i, x_{mi}, y_{ci}) \)

\( \mu_i^{b_i} = E_{y_{i}|x_{i}, b_i} \) and \( \sigma_{ij} = \sigma_{i,j}^{b_i} = \text{var}(y_{ij}|x_i, b_i) \).

\[ w(x) = \min \left\{ 1, \left( \frac{p_0}{(x - m_x)' S_x^{-1} (x - m_x)} \right)^{\nu_0/2} \right\} \]

\( p_0 \) percentile of \( \chi^2_{\dim(x)} \)
Here \( \psi_{1c} = \max(-c, \min(x, c)) \), \( c \) is typically chosen to be 1.5 to ensure asympt. efficiency at a certain level and

\[
v_{ij}(\theta, b_i, x_m^i, y_{ci}) = \text{E}_{y_i, c_i}^* \left\{ \psi_{1c} \left( \frac{y_{ij} - \mu_{ij}^b}{\sigma_{ij}^b} \right) \right\} | y_{ci}, x_m^i, b_i
\]

\( m, a, n, d, S \) are robust estimates of location and scale of the distribution of \( x \). (Rousseeuw and Van Zomeren (1990, Jasa))
Algorithm

1. Start with $t = 0$. Choose initial values values of $\theta^{(0)}$. These initial estimates can be chosen as the ordinary Monte Carlo based Newton Raphson approach.

2. Using multivariate rejection sampling method or importance sampling method (depending on the dimensions of missing responses, covariates and random effects) generate $S_i$ observations $\{y^{(s)}_{ci}, x^{(s)}_{mi}, b^{(s)}_i, s = 1, \ldots, S_i\}$ from $f(y_{ci}, x_{mi}, b_i|x_{oi}, y^*_i, c^*_i; \theta^{(t)})$ and compute $\xi^{(t+1)} = (\beta^{(t+1)}, \alpha^{(t+1)})$ from the expression

$$
\xi^{(t+1)} = \xi^{(t)} + \left[ \sum_{i=1}^{m} \left\{ \frac{1}{S_i} \sum_{s=1}^{S_i} G'_i \cdot W_i (\theta^{(t)}, y^{(s)}_{ci}, b^{(s)}_i, x^{(s)}_{mi}) \right. \right. \\
\left. \left. H_i (\theta^{(t)}, y^{(s)}_{ci}, b^{(s)}_i, x^{(s)}_{mi} G'_i) \right\} \right]^{-1} \\
\left[ \sum_{i=1}^{m} \left\{ \frac{1}{S_i} \sum_{s=1}^{S_i} G'_i W_i (\theta^{(t)}; y^{(s)}_{ci}, b^{(s)}_i, x^{(s)}_{mi}) \right. \right. \\
\left. \left. h_i (\theta^{(t)}; y^{(s)}_{ci}, b^{(s)}_i, x^{(s)}_{mi}) \right\} \right]
$$
Where

\[ H_i = Diag.(\ldots, \frac{\partial h_{ij}}{\partial \eta_{ij}^b}, \ldots) \]

\[ W_i = Diag.(\ldots, \sigma_{ij}^b w(x_{ij}, \pi_{ij}), \ldots) \]
Algorithm Contd.

3. $\gamma^{(t+1)}$ and $\nu^{(t+1)}$ are updated by maximizing

$$\sum_{i=1}^{m} \frac{1}{S_i} \sum_{s=1}^{S_i} \log f(x_{oi}, x_{mi}^{(s)} | \gamma)$$

and

$$\sum_{m=1}^{m} \frac{1}{S_i} \sum_{s=1}^{S_i} \log f(b_{i}^{(s)} | \nu)$$

respectively.

4. Continue Step - 2 until a convergence is achieved.
Approximate Approach

GPLMM can be approximated by IPLMM (iteratively PLMM)

\[ u_i = x_i \beta + \pi_i \alpha + z_i b_i + e_i, \quad i = 1, \ldots, m \]

\[ e_i \mid b_i \sim N(0, \sigma^2 I), \quad b_i \sim N(0, D), \quad D = D(\nu) \]

\[ x_i = (x_{i1}, \ldots, x_{in_i})', \quad \pi_i = (\pi(t_{i1}) \ldots \pi(t_{in_i}))' \]

\[ u_i = \left( \frac{\partial \eta_i}{\partial \mu_i} \right)_{\hat{b}_i, \hat{\theta}} (\hat{y}_i - \hat{\mu}_i) + X_i \hat{\beta} + \pi_i \hat{\alpha} + Z_i \hat{b}_i \]

\[(\hat{\theta}, \hat{b}_i)' \text{ is the current estimate} \]

\[(b_i \mid u_i, x_i, \hat{\theta}) \sim N(\tilde{b}_i, \tilde{\Sigma}) \text{ where } \tilde{b}_i = \tilde{\Sigma}_i z_i' (u_i - X_i \hat{\beta} - \pi_i \hat{\alpha}) / \tilde{\sigma}^2 \]

\[ \tilde{\Sigma}_i = (\tilde{\sigma}^{-2} z_i' z_i + \hat{D}^{-1})^{-1}. \]
This approximation could be done in view of the linearization
\[ u = g(y) = g(\mu) + (y - \mu) g'(\mu) \]
\[ = \eta + (y - \mu) \frac{d\eta}{d\mu} \]
\[ = x\beta + \pi\alpha + zb + (y - \mu) \frac{d\eta}{d\mu} \]
Asymptotics

**Theorem 1:** Under Suitable assumptions, when the number of Knots \(K_n \sim n^{1/(2r+1)}\), then

\[
\frac{1}{n} \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left( \hat{f}(t_{ij}) - f_0(t)_{ij} \right)^2 = O_p \left(n^{-\frac{2r}{2r+1}}\right)
\]

and

\[
\sqrt{n}(\hat{\beta} - \beta_0) \xrightarrow{L} N(0, B^{-1}SB^{-1})
\]
Where

\[ B = \lim_{n \to \infty} B_n, B_n = \sum_{i=1}^{n} E \left[ X_i^* \Phi_i X_i^* \mid y_i^*, c_i^* \right] \]

\[ S = \lim_{n \to \infty} S_n, S_n = \sum_{i=1}^{n} E \left[ X_i^* \Phi_i^1 X_i^* \mid y_i^*, c_i^* \right] \]

\[ \Phi_i = W_i(\theta_0, b_i) H_i(\theta_0, b_i, x_{mi}) \]
Simulation

Here we have considered the model

\[ y_{ij}^* = \beta_0 + \beta_1 x_{ij} + \sin(e_{ij}) + b_i + e_{ij} \quad i = 1, 2, ..., 100; j = 1, 2, 3. \] (1)

where \( x_{ij} \) are generated from a normal distribution with mean 2 and variance 1 and \( t_{ij} \) are independently generated from a uniform distribution \([0, 1]\). We have considered the true values of \( \beta_0 = 0.5 \) and \( \beta_1 = 1 \). \( b_i \)'s are generated from \( N(0, \sigma_b^2 = 2) \) and \( e_{ij} \)'s from \( N(0, \sigma_e^2 = 1) \).
Generation of missing

For generating missing covariates we have generated a binary column of order $300 \times 1$ and considered where there is 1 those covariates are missing. The above vector is generated by drawing random samples from binomial distribution with parameters 1 and 0.2. That is, a particular observation of the covariates will be missing with probability 0.2.
Introduction of Outliers

introduce the 20 percent outliers in the data we generate 80 percent of random errors from $N(0, 1)$ and the remaining 20 percent from an arbitrary $N(0, 5)$ distribution.
Table 1: Comparison of estimates of $\beta_0$, $\beta_1$, $\sigma^2$ and $\sigma_b^2$ in terms of MSE in two methods of estimation when there is no outliers in the data

<table>
<thead>
<tr>
<th>Estimation Procedure</th>
<th>True Value</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$\sigma^2$</th>
<th>$\sigma_b^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Without Robust Estimation</strong></td>
<td>Estimates</td>
<td>0.4991</td>
<td>1.0046</td>
<td>1.0026</td>
<td>1.9968</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0101</td>
<td>0.0106</td>
<td>0.0398</td>
<td>0.0425</td>
</tr>
<tr>
<td></td>
<td>IMMSE</td>
<td>0.1273</td>
<td>0.1149</td>
<td>0.1329</td>
<td>0.1345</td>
</tr>
<tr>
<td></td>
<td>Coverage Probability</td>
<td>0.9559</td>
<td>0.9404</td>
<td>0.9168</td>
<td>0.9364</td>
</tr>
<tr>
<td><strong>Using Robust Estimation</strong></td>
<td>Estimates</td>
<td>0.4708</td>
<td>1.0035</td>
<td>0.9994</td>
<td>1.9952</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0216</td>
<td>0.0120</td>
<td>0.0421</td>
<td>0.0616</td>
</tr>
<tr>
<td></td>
<td>IMMSE</td>
<td>0.0673</td>
<td>0.0907</td>
<td>0.0907</td>
<td>0.0784</td>
</tr>
<tr>
<td></td>
<td>Coverage Probability</td>
<td>0.9300</td>
<td>0.9436</td>
<td>0.9143</td>
<td>0.9231</td>
</tr>
<tr>
<td><strong>Approximate EM Method (Classical)</strong></td>
<td>Estimates</td>
<td>0.5399</td>
<td>1.0402</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0146</td>
<td>0.0140</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>IMMSE</td>
<td>0.1168</td>
<td>0.1148</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Coverage Probability</td>
<td>0.9557</td>
<td>0.9388</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Approximate EM Method (Robust)</strong></td>
<td>Estimates</td>
<td>0.4560</td>
<td>1.0012</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0228</td>
<td>0.0151</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>IMMSE</td>
<td>0.0817</td>
<td>0.0872</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Coverage Probability</td>
<td>0.9362</td>
<td>0.9416</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 2: Comparison of estimates of $\beta_0$, $\beta_1$, $\sigma^2$ and $\sigma_b^2$ in terms of MSE in two methods of estimation when there is 20 percent outliers in the data

<table>
<thead>
<tr>
<th>Estimation Procedure</th>
<th>Estimates</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$\sigma^2$</th>
<th>$\sigma_b^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True Value</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>With out Robust</td>
<td>MSE</td>
<td>0.4556</td>
<td>0.4711</td>
<td>0.4690</td>
<td>0.5025</td>
</tr>
<tr>
<td>Estimation</td>
<td>IMSE</td>
<td>0.1156</td>
<td>0.1342</td>
<td>0.1210</td>
<td>0.1458</td>
</tr>
<tr>
<td></td>
<td>Coverage Probability</td>
<td>0.9019</td>
<td>0.8953</td>
<td>0.8922</td>
<td>0.8808</td>
</tr>
<tr>
<td>Using Robust</td>
<td>MSE</td>
<td>0.0102</td>
<td>0.0100</td>
<td>0.0382</td>
<td>0.0406</td>
</tr>
<tr>
<td>Estimation</td>
<td>IMSE</td>
<td>0.0571</td>
<td>0.0782</td>
<td>0.0687</td>
<td>0.0849</td>
</tr>
<tr>
<td></td>
<td>Coverage Probability</td>
<td>0.9451</td>
<td>0.9569</td>
<td>0.9359</td>
<td>0.9336</td>
</tr>
<tr>
<td>Approximate EM</td>
<td>MSE</td>
<td>0.4496</td>
<td>0.4388</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Method</td>
<td>IMSE</td>
<td>0.1025</td>
<td>0.1086</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Coverage Probability</td>
<td>0.9023</td>
<td>0.9104</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Approximate EM</td>
<td>MSE</td>
<td>0.0219</td>
<td>0.0145</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Method (Robust)</td>
<td>IMSE</td>
<td>0.0470</td>
<td>0.0708</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Coverage Probability</td>
<td>0.9559</td>
<td>0.9556</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Convergence of $\beta_1$ when there is 20% outliers in the data

Convergence of $\beta_1$ when there is no outliers in the data
Table 3: Analysis of the Arctic data: Effect of Temperature and Salinity on the phytoplankton (Thalassiosira nordenskioeldii) cell count has been studied by the three methods of estimation.

<table>
<thead>
<tr>
<th>Estimation Procedure</th>
<th></th>
<th>Constant</th>
<th>Temp</th>
<th>Salinity</th>
<th>$\sigma_b^2$</th>
<th>$\sigma^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>With out Robust Estimation</td>
<td>Estimates</td>
<td>1.0463</td>
<td>1.0953</td>
<td>-0.9836</td>
<td>15.8330</td>
<td>6.3492</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0803</td>
<td>0.0927</td>
<td>0.0781</td>
<td>1.8251</td>
<td>1.3333</td>
</tr>
<tr>
<td>Using Robust Estimation</td>
<td>Estimates</td>
<td>1.0740</td>
<td>1.0536</td>
<td>-1.0066</td>
<td>19.3642</td>
<td>6.9477</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0712</td>
<td>0.0605</td>
<td>0.0519</td>
<td>1.3346</td>
<td>0.9972</td>
</tr>
<tr>
<td>Approximate EM (Classical)</td>
<td>Estimates</td>
<td>1.0371</td>
<td>1.0962</td>
<td>-0.9900</td>
<td>15.1583</td>
<td>6.3021</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0875</td>
<td>0.0971</td>
<td>0.0717</td>
<td>1.8841</td>
<td>1.5732</td>
</tr>
<tr>
<td>Approximate EM (Robust)</td>
<td>Estimates</td>
<td>1.0777</td>
<td>1.0038</td>
<td>-1.0053</td>
<td>19.0133</td>
<td>6.8447</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0819</td>
<td>0.0662</td>
<td>0.0532</td>
<td>1.4738</td>
<td>1.0361</td>
</tr>
</tbody>
</table>