1 Executive Summary

In this paper we present a time series analysis on the number of worldwide earthquakes with magnitude from the year 1918 to 2020. We are interested on analyzing how this number varies, as well as the forecasting pattern for the number of earthquakes of such magnitude in the near future .We concluded that the AR(3) model is better than the mixture of AR(3) models. We also concluded that for the next four years, world wide earthquakes with magnitude greater than 7 will be 12,10,10,11.

2 Introduction

We will use a stand alone AR(3) model, as well as a mixture of AR(3) model for this analysis. The Dataset used is as mentioned previously.

3 Data

The dataset we used is the annual number of worldwide earthquakes with magnitude greater than 7 on the Richter scale for T = 100 years, from 1918 to 2020. The original dataset can be obtained from USGS Website.

4 AR Model

4.1 Model Order Selection

To detect the model order for our earthquake dataset, we used the Partial Auto correlation Function (PACF). In the sample PACF plot shown in figure 1, we can see that the last significant value is at lag of 3. This plot favours model of order 3. Additionally, we also used the Akaike Information Criterion (AIC) and Bayesian Informatino Criterion BIC metrics for deciding the model order. These metrics are plotted against the model order in the figure 2, where a indicates the AIC value, while b indicates the BIC value. Visibly, both the criterions/metrics have their lowest value when the AR order equals 3. Based on this plot, as well as the PACF plot, we concluded that appropriate model order for the earthquake dataset should be 3.

4.2 Model Definition

Following is the model definition for our AR(3) model:

$$y_t = \sum_{j=1}^{3} \Phi_j y_{t-j} + \epsilon_t; \;\; \epsilon_t \sim^{iid} N(0, \nu)$$

We used the following conditional conjugate prior:

$$p(\Phi, \nu) = p(\Phi|\nu)p(\nu)$$

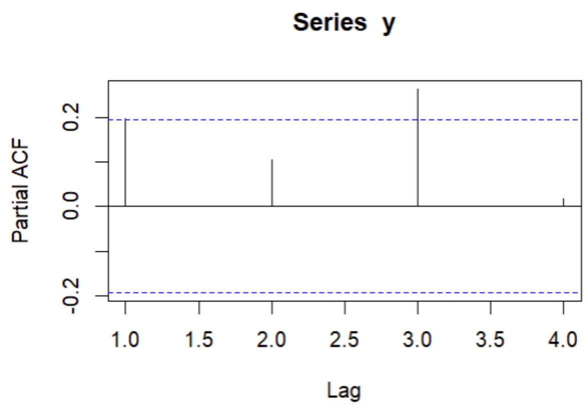

Figure 1: Sample PACF

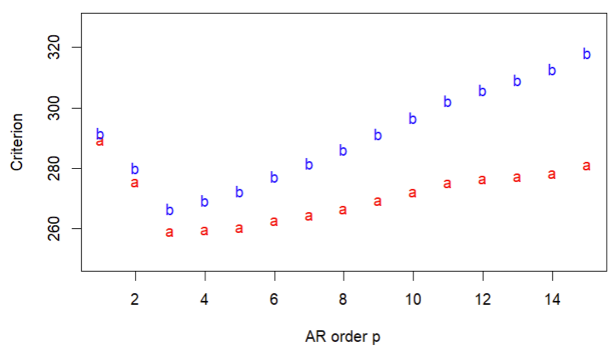

Figure 2: AIC and BIC plot against model orders. Here *a* indicates AIC value, and *b* indicates BIC value. Both metrics/criterion have their lowest value at order value *3*.

$$p(\Phi|\nu) \sim N(\Phi|m_0, \nu C_0)$$
$$p(\nu) \sim IG(n_0/2, d_0/2)$$

The likelihood for our model is:

$$\mathbf{y} \sim N(F^T\Phi|\nu I_n)$$

4.3 Prior Sensitivity Analysis

To decide on how sensitive is our analysis to the choice of prior, we computed posterior distribution for $\Phi = (\phi_1, \phi_2, \phi_2)$ for three different choice of prior hyper-parameters. Recall, are prior family of distribution is as mentioned in the section above. Here, we only varied the hyper-parameters n_0, d_0, C_0, The choice of hyper-parameters for the prior,the corresponding posterior distribution of the Φ in terms of its mean $\mathbf{m}$, and the variance-covariance matrix $\nu\mathbf{C}$, is presented below.

The choice, $n_0 = d_0 = 2, C_0 = diag(3), m_0 = (0, 0, 0)^T$, resulted in:

$$\mathbf{m} = (0.3332666, 0.2060106, 0.4427255)^T$$

$$\mathbf{C} = \begin{bmatrix} 0.0005228847 & -0.0002742564 & -0.0002326578 \\ -0.0002742564 & 0.0005629060 & -0.0002698535 \\ -0.0002326578 & -0.0002698535 & 0.0005278381 \end{bmatrix}$$

The choice, $n_0 = 6, d_0 = 1, C_0 = diag(3), m_0 = (0, 0, 0)^T$, resulted in:

$$\mathbf{m} = (0.3332666, 0.2060106, 0.4427255)^T$$

$$\mathbf{C} = \begin{bmatrix} 0.0005228847 & -0.0002742564 & -0.0002326578 \\ -0.0002742564 & 0.0005629060 & -0.0002698535 \\ -0.0002326578 & -0.0002698535 & 0.0005278381 \end{bmatrix}$$

The choice, $n_0 = 6, d_0 = 1, C_0 = diag(3), m_0 = (-0.5, -0.5, -0.5)^T$, resulted in:

$$\mathbf{m} = (0.3332586, 0.2060012, 0.4427129)^T$$

$$\mathbf{C} = \begin{bmatrix} 0.0005228847 & -0.0002742564 & -0.0002326578 \\ -0.0002742564 & 0.0005629060 & -0.0002698535 \\ -0.0002326578 & -0.0002698535 & 0.0005278381 \end{bmatrix}$$

The mean values $\mathbf{m}$ and variance-covariance matrix $\nu\mathbf{C}$ (we have showed $\mathbf{C}$ matrix above) of the Φ parameter resulting from these choices of prior-hyper-parameters are almost exact, indicating that the posterior distribution is insensitive to the choice of prior. So we will just pick the prior with $n_0 = 0.02, d_0 = 0.02, C_0 = 10 * diag(3), m_0 = (0, 0, 0)^T$ The model predictions for the next four years are 12,10,10,11.

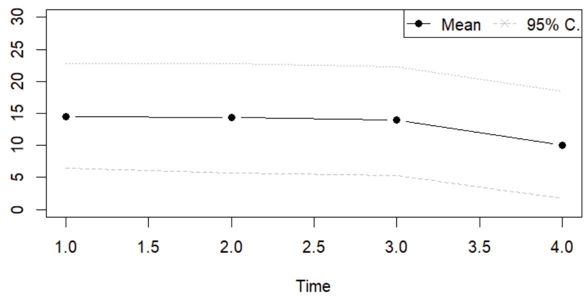

Figure 3: Predictions from a mixture of two AR(3) model

5 Mixture of AR Model

Here, we will now fit a mixture of AR(3).

5.1 Deciding the number of components

For deciding the number of components, we will fit a mixture of two, three, four, and five models, and then compare there fit and complexity using the criterion Deviance Information Criterion (DIC). The mixture model with lowest DIC will be the one we will choose. Follwing is the table summarizing the results:

Mixture components	DIC
2	575
3	576
4	579
5	580

Consequently, we will choose the mixture model with two components.

5.2 Chosen Model

The chosen mixture model was the one with two components. The prediction for the next four years can be seen in the figure 3.

6 Model Comparison

Here we will compare our standalone AR(3) model with our mixture of AR(3) models using the Deviance Information Criterion (DIC). The standalone AR model has a DIC of 572.4, whereas the mixture of two model had a DIC of 575. Since the stan alone model has a lower DIC, we will go ahead with that model for predictions.

7 Conclusions

We concluded that the AR(3) model is better than the mixture of AR(3) models. We also concluded that for the next four years, world wide earthquakes with magnitude greater than 7 will be 12,10,10,11.