

A Bayesian Analysis of Global Warming

Lewis Fishgold

May 7, 2014

1 Introduction

Global temperature has risen 0.8°C since the early 20th century [2]. The Intergovernmental Panel on Climate Change (IPCC) has stated that scientists are 90% certain that the bulk of global warming is being caused by human activities including fossil fuel combustion, cement production, and deforestation. The geological effects of global warming include the rise of sea level, changes in patterns of precipitation, and expansion of subtropical desert, and are expected to have dire consequences for human wellbeing. The goal of this paper is to perform a Bayesian analysis of the rate of temperature change in different regions of the world.

2 Data Set

I am using the GISTEMP data set from NASA [1], which contains monthly temperature anomalies for the past 100 years for $2^{\circ} \times 2^{\circ}$ regions around the globe. Temperature anomalies measure how much warmer or colder it was in a particular region and month than the average for that region and month, where the average is computed over the period of 1951-1980.

For the sake of simplicity, I am only using data from the past 50 years from 36 regions lying on a grid over Europe, Asia, Africa, and Australia. I have also made the data more temporally coarse-grained, by averaging over the months for each year. To get some intuition for what this data is like, Figure 1 shows the anomaly curve for each of the 36 regions. Although there is much noise, an upward trend is apparent.

3 Models

To model the data, I used three different linear models. Each model attempts to capture the slope in the anomaly curve for each region. The first model, in Figure 2, is a linear model with pooling, where each region is modelled by the same coefficient. The second model, in Figure 3, is a linear model without pooling, where each region has its own coefficient. I expect that this is a better

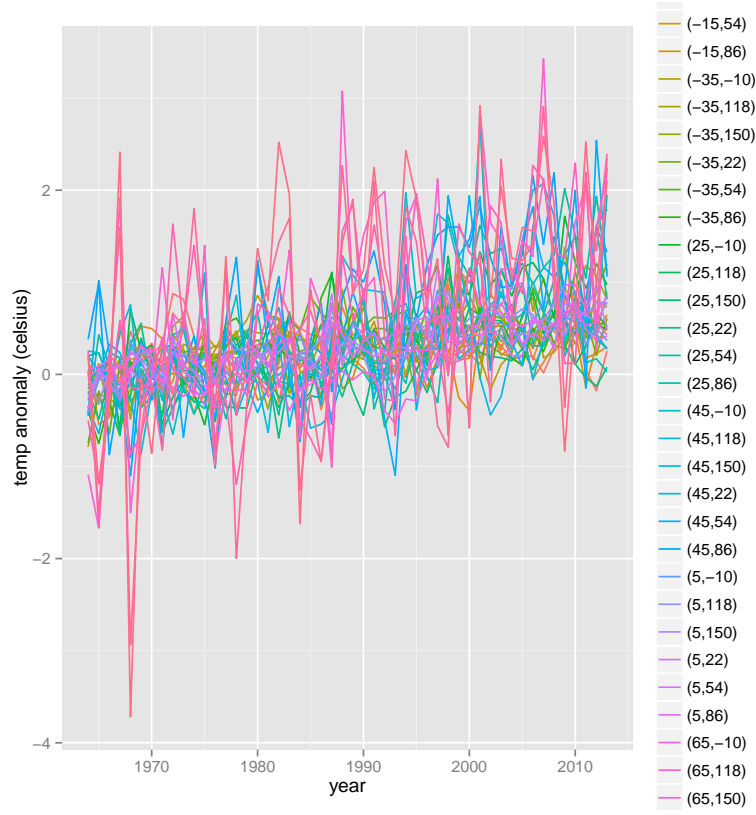


Figure 1: The anomaly curve for each of the 36 regions over the past 50 years, labelled by latitude and longitude.

model than the former, as global warming varies by region. The third model, in Figure 4, is a hierarchical model where each region has its own coefficient. I expect that this is a better model than the former, as there is a global process influencing regional temperature change, and the hierarchical aspect of the model will help reign in outliers.

In each of the models, the variance parameter for each anomaly curve is shared across regions. I have centered the data from each region, so that no bias term is necessary in the models. The prior knowledge I have about global warming is indirectly based on the GISTEMP data set, so using informative priors would amount to using the same data twice. Therefore, I have used uninformative priors.

```

1 model {
2
3   for(r in 1:R) {
4     for(y in 1:Y) {
5       temps.hat[y,r] <- b * times[y]
6       temps[y,r] ~ dnorm(temps.hat[y,r], tau.temps)
7     }
8   }
9
10  b ~ dnorm(0, 0.0001)
11  sigma.temps ~ dunif(0, 100)
12  tau.temps <- pow(sigma.temps, -2)
13
14 }

```

Figure 2: Bugs code for the linear model with pooling.

4 Results

I used a Gibbs sampler to draw 10,000 samples (using 5 chains, 2,000 burnin samples per chain, and no thinning) from the posterior distribution for each of the three models. To assess convergence, I noted that the \hat{R} values from the BRG diagnostic are close to one for all parameters. In addition, the number of effective samples for all parameters is greater than 2000. I used DIC (Deviance Information Criterion), an estimate of expected predictive error, to compare models, and the results are in Table 1. The DIC scores for the non-pooled and hierarchical models are tied, but are better than the score of the pooled model. I suspect that the hierarchical model did not provide an advantage over the non-pooled model because there was enough data to model each slope parameter individually without sharing statistical strength between regions.

Model	DIC (lower better)
Pooled	2508
Not Pooled	2363
Hierarchical	2364

Table 1:

The following results are based on the hierarchical model. The 95% credible intervals for the marginal posteriors of all but one region are above zero, indicating that there is warming with a high degree of certainty. The 80% credible intervals are shown in Figure 5. To gain better insight into the variation across regions, I have plotted the expected value of the slope for each region on a map in Figure 6. From the plot, it can be seen that the rate varies widely from

```

1 model {
2
3   for(r in 1:R) {
4     for(y in 1:Y) {
5       temps.hat[y,r] <- b[r] * times[y]
6       temps[y,r] ~ dnorm(temps.hat[y,r], tau.temps)
7     }
8   }
9
10  for(r in 1:R) {
11    b[r] ~ dnorm(0, 0.0001)
12  }
13  sigma.temps ~ dunif(0, 100)
14  tau.temps <- pow(sigma.temps, -2)
15
16 }

```

Figure 3: Bugs code for the linear model without pooling.

around $1 - 4^\circ$ per century, and that global warming tends to be greater at more northern latitudes.

References

- [1] James Hansen, Reto Ruedy, Makiko Sato, and Ken Lo. Global surface temperature change. *Reviews of Geophysics*, 48(4), 2010.
- [2] Wikipedia. Global warming — wikipedia, the free encyclopedia, 2014. [Online; accessed 7-May-2014].

```

1 model {
2
3   for(r in 1:R) {
4     for(y in 1:Y) {
5       temps.hat[y,r] <- b[r] * times[y]
6       temps[y,r] ~ dnorm(temps.hat[y,r], tau.temps)
7     }
8   }
9
10  for(r in 1:R) {
11    b[r] ~ dnorm(mu.b, tau.b)
12  }
13
14  mu.b ~ dnorm(0,0.0001)
15  sigma.b ~ dunif(0,100)
16  tau.b <- pow(sigma.b, -2)
17
18  sigma.temps ~ dunif(0, 100)
19  tau.temps <- pow(sigma.temps, -2)
20
21 }

```

Figure 4: Bugs code for the hierarchical linear model.

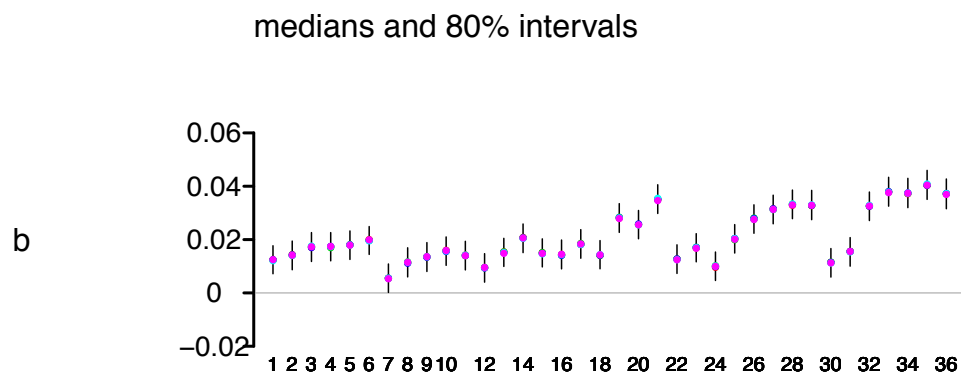


Figure 5: 80% credible intervals for posteriors over hierarchical model.

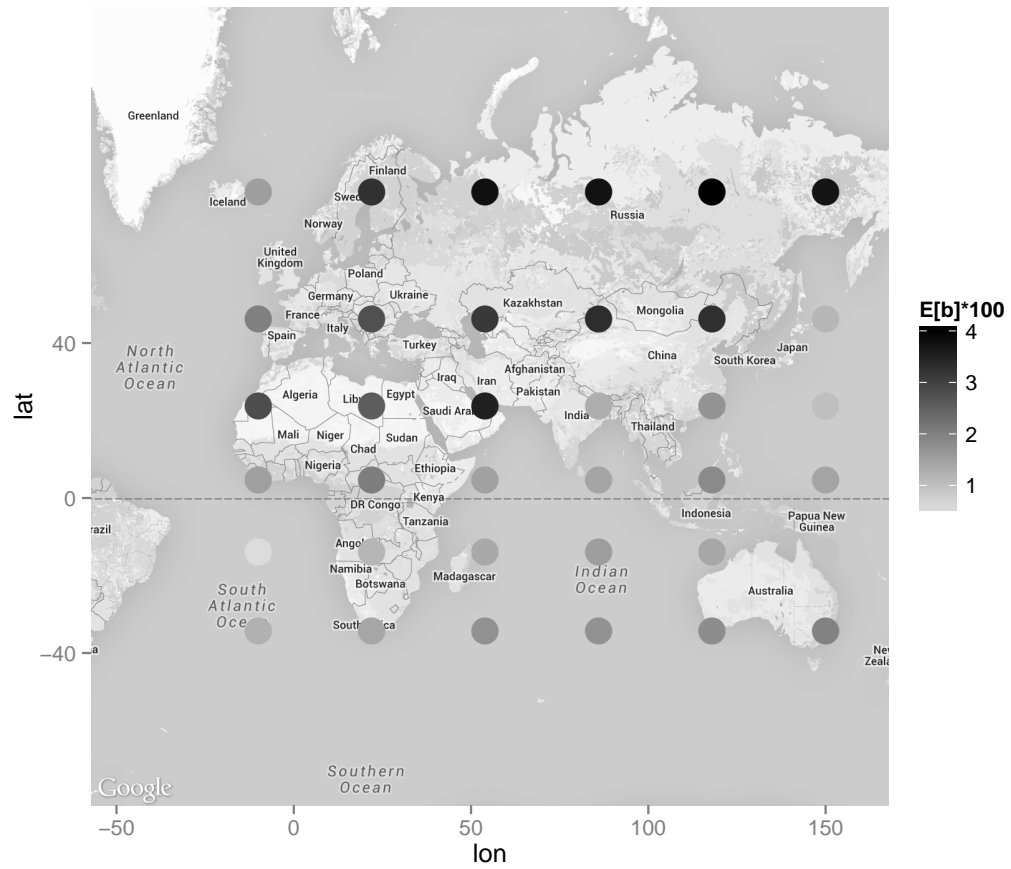


Figure 6: The expected rate of change in temperature anomaly per century.