How has bird body size responded to temperature changes across latitude?

```
library(knitr)
opts_chunk$set(tidy.opts=list(width.cutoff=60),tidy=TRUE)
```

Overview and approach

In this study, we have three main questions: 1) Have bird body sizes across a diverse range of species and locations changed over the past few decades, and if so, how? 2) Are these changes related to realized changes in temperature and / or precipitation due to anthropogenic global warming? 3) Does latitude predict the degree to which body size responds to temperature? We hypothesized we would indeed see a response in body size to recent climate change, that body size would be negatively correlated to temperature, and that body size changes would be greatest at low latitudes. I've put all code and some early figures up on GitHub (https://github.com/elinck/bird_body_size) if you're interested in digging in deeper, but this document outlines the approach we've taken so far to cleaning and filtering data, checking assumptions and exploring relationships, and statistical analysis.

Data cleaning and filtering

To address these questions, we're taking advantage of 1) longitudinal climate data from the University of Delaware and 2) seven longitudinal bird banding datasets from across the Americas, with sites in Brazil, Panama, Puerto Rico, California, Wyoming, Pennsylvania, and Illinois. The climate data was extracted and assembled by Xingli; you can see the code in the script `01_extract_climate.R`. The bird banding data were previously collated and formatted by hand by Matt McGee—thanks, Matt!.

Here are the number of data points and unique species from these files prior to any filtering:

```
raw_data
```

```
##         site observations species
## 1      brazil        24104     143
## 2      panama        11679     189
## 3     guanica         8061      87
## 4   powdermill       719781     196
## 5    palomarin       158828     179
## 6       teton        11776      79
## 7    waterfall         2393      68
```

As you can see, a good amount of data! Depending on the site, we have anywhere from 68 species and 2393 records (at Waterfall Glen) to 189 species (in Panama) to 719781 records (at Powdermill). But of course, our project requires filtering these data down to records that will allow us to make comparisons among sites with as few assumptions as possible. Ultimately, we're hoping to analyze three different datasets, to see whether patterns remain the same depending on which species and records we look at: a dataset that includes records that meet our criteria regardless of sex, a dataset that includes only male birds meeting our criteria, and a dataset that includes only female birds meeting our criteria.

For our sex-blind dataset, we apply the following filters:

1) We exclude all juvenile birds, to ensure body size data represented mature individuals. However, we retain records that had either NA or "Unknown" in the "Age" column, as this would dramatically reducer the numbe of observations in certain datasets;

2) We exclude all records without mass data, or records with a mass of 0;

1

3) We drop all records with a mass greater or less than 4 standard deviations from the mean for each species, as these are likely errors;

4) We include only the first recorded data point for each individual bird, based on date of capture and band number (when available);

5) To ensure a sufficient number of points for estimating the slope of change in mass over time, we include only species with at least five records in each of two time periods, which were define as either less than or equal to the median year of data collection for a given locality, or greater than the median year of data collection. E.g., for the Brazil dataset, we exclude all species without at least 5 records before 1984, or five records from 1984 to 2013.

6) For temperate sites (Palomarin, Powdermill, Teton Science School, Waterfall Glen), we include only records from the breeding season, which we define as the months of June and July. In contrast, we do not apply any temporal filter on data from tropical sites. While comparing body size of birds in breeding condition across latitude would obviously be ideal, year-round breeding in the tropics— or breeding activity correlated with locally-specific dry and wet seasons—make selecting a particular set of months as comparable to the temperate breeding season difficult to impossible. Further, we are constrained both by the limited number of total records from tropical sites, and the temporal distribution of banding effort.

For our sex-specific datasets, we apply the same filters as above, but only to the individuals of a single known sex.

We implement these filters using our custom `clean_data()` function from the `00_functions.R` script.

Here's the output from running this function on each dataset, which exports information on the number of species, observations, and years included:

```r
source("~/Dropbox/Bird_body_size-analysis/bird_body_size/scripts/00_functions.R")

# run function on Brazil data
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/brazil_raw.csv",
           sex="ALL", climate = "TROPICAL", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/brazil_filtered.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/brazil_raw.csv", : NAs introduced by coercion

## [1] "here are the unique years in this dataset:"
##  [1] 1980 1981 1982 1983 1984 1985 2008 2010 2011 2012 2013
## [1] "there are  15609  records"
## [1] "there are  79  species"
```

```r
# run function on Panama data
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/panama_raw.csv",
           sex="ALL", climate = "TROPICAL", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/panama_filtered.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/panama_raw.csv", : NAs introduced by coercion

## [1] "here are the unique years in this dataset:"
##  [1] 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1990 1992
## [15] 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006
## [29] 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016
## [1] "there are  8492  records"
```

```
## [1] "there are  57  species"
```
```
# run function on Guanica data
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/guanica_raw.csv",
           sex="ALL", climate = "TROPICAL", drop_juveniles = TRUE,
           output_path =
              "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/guanica_filtered.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984
## [15] 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998
## [29] 1999 2000 2001 2002 2003 2004 2005 2006 2008 2009 2010 2011 2012 2013
## [43] 2014 2015 2016 2017
## [1] "there are  5893  records"
## [1] "there are  27  species"
```
```
# run function on Powdermill data
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/powdermill_raw.csv",
           sex="ALL", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
              "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/powdermill_filtered.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/powdermill_raw.csv", : NAs introduced by coercion
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/powdermill_raw.csv", : NAs introduced by coercion
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/powdermill_raw.csv", : NAs introduced by coercion
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974
## [15] 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988
## [29] 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002
## [43] 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
## [1] "there are  20952  records"
## [1] "there are  74  species"
```
```
# run function on Palomarin data
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/palo_raw.csv",
           sex="ALL", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
              "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/palo_filtered.csv")
```

```
## [1] "here are the unique years in this dataset:"
```

```
##  [1] 1971 1972 1973 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985
## [15] 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999
## [29] 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013
## [43] 2014 2015 2016
## [1] "there are  9478  records"
## [1] "there are  38  species"
```

```r
# run function on Teton Science School data
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/tss_raw.csv",
           sex="ALL", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/tss_filtered.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/tss_raw.csv", : NAs introduced by coercion
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008
## [15] 2009 2010 2011 2012 2013 2014 2015 2016 2017
## [1] "there are  4506  records"
## [1] "there are  36  species"
```

```r
# run function on Waterfall Glen data
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/wate_raw.csv",
           sex="ALL", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/wate_filtered.csv")
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
## [1] "there are  268  records"
## [1] "there are  7  species"
```

```r
# run function on Brazil data, males only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/brazil_raw.csv",
           sex="MALES", climate = "TROPICAL", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/brazil_filtered_males.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/brazil_raw.csv", : NAs introduced by coercion
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1980 1981 1982 1983 1984 1985 2008 2010 2011 2012 2013
## [1] "there are  4021  records"
## [1] "there are  34  species"
```

```r
# run function on Panama data, males only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/panama_raw.csv",
           sex="MALES", climate = "TROPICAL", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/panama_filtered_males.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/panama_raw.csv", : NAs introduced by coercion
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1990 1992
## [15] 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006
```

```
## [29] 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016
## [1] "there are  1215  records"
## [1] "there are  15  species"
```

```r
# run function on Guanica data, males only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/guanica_raw.csv",
           sex="MALES", climate = "TROPICAL", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/guanica_filtered_males.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion

## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion

## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion

## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion

## [1] "here are the unique years in this dataset:"
##  [1] 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985
## [15] 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999
## [29] 2000 2001 2002 2003 2004 2005 2006 2009 2010 2012 2013 2014 2015 2016
## [43] 2017
## [1] "there are  1300  records"
## [1] "there are  18  species"
```

```r
# run function on Powdermill data, males only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/powdermill_raw.csv",
           sex="MALES", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/powdermill_filtered_males.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/powdermill_raw.csv", : NAs introduced by coercion

## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/powdermill_raw.csv", : NAs introduced by coercion

## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/powdermill_raw.csv", : NAs introduced by coercion

## [1] "here are the unique years in this dataset:"
##  [1] 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974
## [15] 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988
## [29] 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002
## [43] 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
## [1] "there are  9513  records"
## [1] "there are  50  species"
```

```r
# run function on Palomarin data, males only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/palo_raw.csv",
           sex="MALES", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/palo_filtered_males.csv")
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1971 1972 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986
## [15] 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000
## [29] 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014
## [43] 2015 2016
## [1] "there are  3618  records"
## [1] "there are  26  species"
```

```r
# run function on Teton Science School data, males only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/tss_raw.csv",
           sex="MALES", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/tss_filtered_males.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/tss_raw.csv", : NAs introduced by coercion
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008
## [15] 2009 2010 2011 2012 2013 2014 2015 2016 2017
## [1] "there are  1825  records"
## [1] "there are  24  species"
```

```r
# run function on Waterfall Glen data, males
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/wate_raw.csv",
           sex="MALES", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/wate_filtered_males.csv")
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
## [1] "there are  65  records"
## [1] "there are  3  species"
```

```r
# run function on Brazil data, females only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/brazil_raw.csv",
           sex="FEMALES", climate = "TROPICAL", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/brazil_filtered_females.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/brazil_raw.csv", : NAs introduced by coercion
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1980 1981 1982 1983 1984 1985 2008 2010 2011 2012 2013
## [1] "there are  3502  records"
## [1] "there are  34  species"
```

```r
# run function on Panama data, females only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/panama_raw.csv",
           sex="FEMALES", climate = "TROPICAL", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/panama_filtered_females.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/panama_raw.csv", : NAs introduced by coercion
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1990 1992
```

```
## [15] 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006
## [29] 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016
## [1] "there are  1147  records"
## [1] "there are  19  species"
```

```r
# run function on Guanica data, females only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/guanica_raw.csv",
           sex="FEMALES", climate = "TROPICAL", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/guanica_filtered_females.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion

## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion

## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion

## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/guanica_raw.csv", : NAs introduced by coercion

## [1] "here are the unique years in this dataset:"
##  [1] 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984
## [15] 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998
## [29] 1999 2000 2001 2002 2003 2004 2005 2006 2008 2009 2010 2011 2012 2013
## [43] 2014 2015 2016 2017
## [1] "there are  1465  records"
## [1] "there are  17  species"
```

```r
# run function on Powdermill data, females only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/powdermill_raw.csv",
           sex="FEMALES", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
             "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/powdermill_filtered_females.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/powdermill_raw.csv", : NAs introduced by coercion

## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/powdermill_raw.csv", : NAs introduced by coercion

## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/powdermill_raw.csv", : NAs introduced by coercion

## [1] "here are the unique years in this dataset:"
##  [1] 1961 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975
## [15] 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989
## [29] 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003
## [43] 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
## [1] "there are  8112  records"
## [1] "there are  52  species"
```

```r
# run function on Palomarin data, females only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/palo_raw.csv",
           sex="FEMALES", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
```

```
                    "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/palo_filtered_females.csv")
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1971 1972 1973 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985
## [15] 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999
## [29] 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013
## [43] 2014 2015 2016
## [1] "there are  3857  records"
## [1] "there are  34  species"
```

```
# run function on Teton Science School data, females only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/tss_raw.csv",
           sex="FEMALES", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
               "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/tss_filtered_females.csv")
```

```
## Warning in clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/
## bird_body_size/data/tss_raw.csv", : NAs introduced by coercion
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008
## [15] 2009 2010 2011 2012 2013 2014 2015 2016 2017
## [1] "there are  1692  records"
## [1] "there are  28  species"
```

```
# run function on Waterfall Glen data, females only
clean_data(file_path = "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/wate_raw.csv",
           sex="FEMALES", climate = "TEMPERATE", drop_juveniles = TRUE,
           output_path =
               "~/Dropbox/Bird_body_size-analysis/bird_body_size/data/wate_filtered_females.csv")
```

```
## [1] "here are the unique years in this dataset:"
##  [1] 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
## [1] "there are  116  records"
## [1] "there are  3  species"
```

Even with those fairly loose filters, we've lost a lot of data. So it goes.

(I've generated plots body size change over time for all species. While there are too many / they are too big to include in this document, you can view them here: https://github.com/elinck/bird_body_size/tree/master/plots).

Checking assumptions

Before analyzing these data, we should evaluate a few basic assumptions we're making. Because mass (in grams) is the only metric of body size recorded for the majority of observations across datasets, we'll be using it as the response variable in our model. But mass can be an imperfect proxy for body size for a variety of reasons. As a gut check, we'll see if the slope of change in mass over time is correlated with the slope of our two other body size measures—wing length and tarsus length—for the subset of all records that have these data. Let's visualize this:
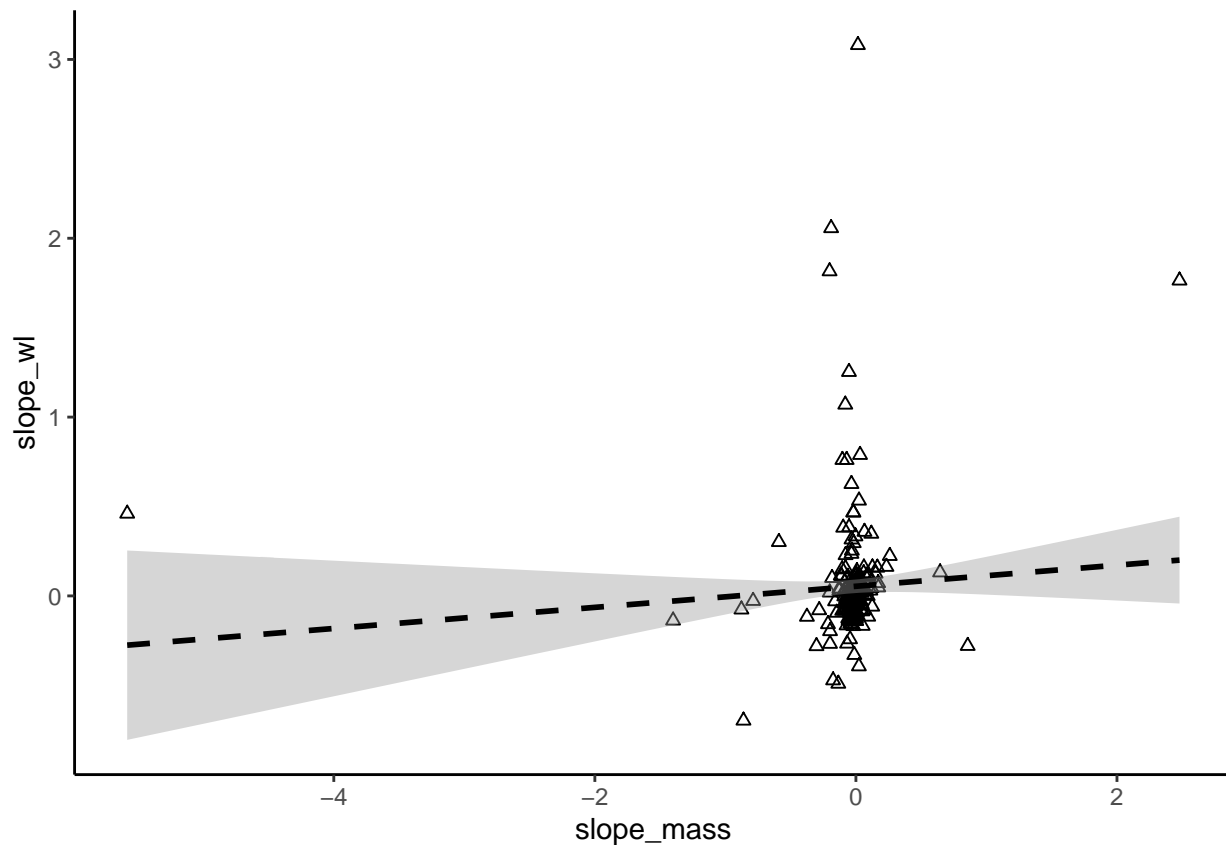
```r
mod.wl <- lm(slope_wl ~ slope_mass, wl.df) # summarize
mod.wl.sum <- summary(mod.wl) # summarize model
mod.wl.sum$adj.r.squared # r^2 value
```

```
## [1] 0.001714323
```

```r
mod.wl.sum$coefficients[2,4] # p-value
```

```
## [1] 0.2238589
```

Well, not a very good correlation, is it! It's possible this is being driven by a handful of outliers, but it *does* make some degree of biological sense; wing length is going to vary quite a bit across molt cycle, etc. Let's check out tarsus length, though we have fewer data.
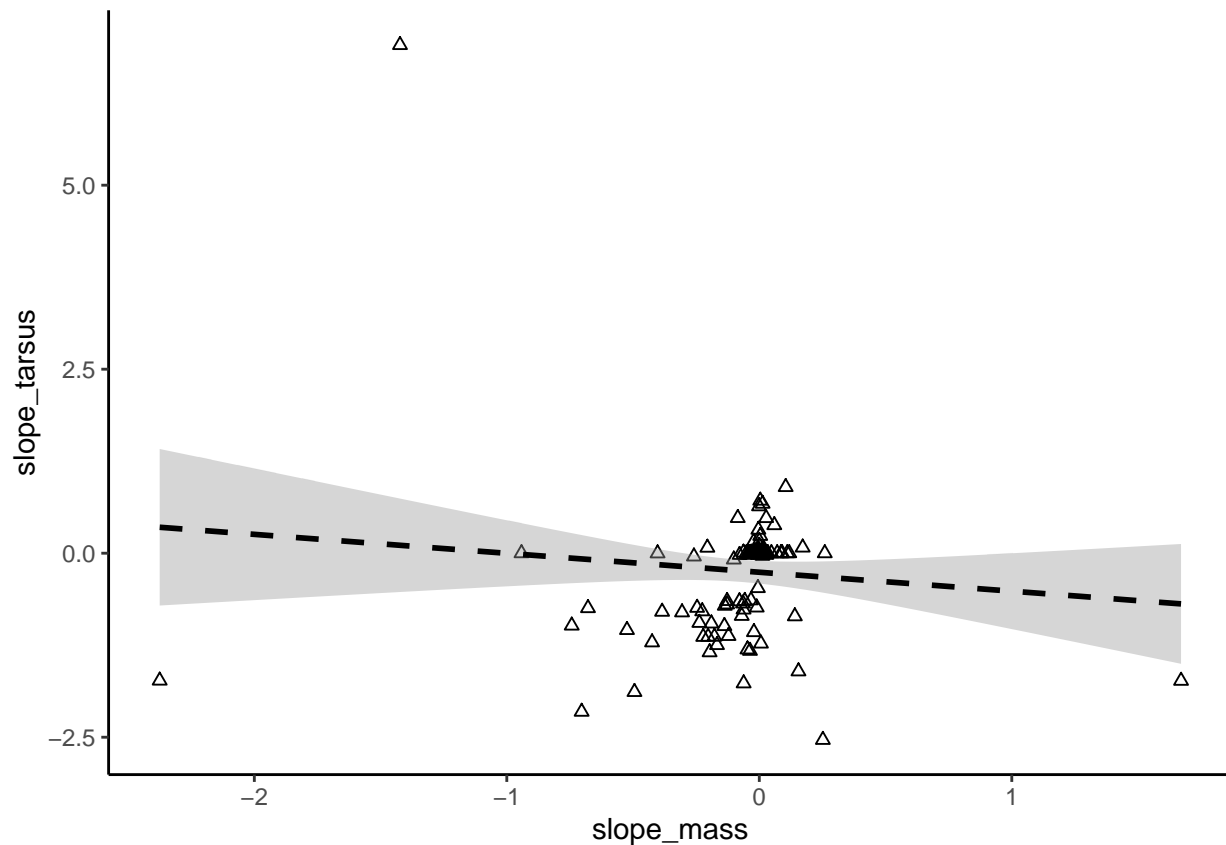
```
mod.tarsus <- lm(slope_tarsus ~ slope_mass, tarsus.df) # summarize
mod.tarsus.sum <- summary(mod.tarsus) # summarize model
mod.tarsus.sum$adj.r.squared # r^2 value
```
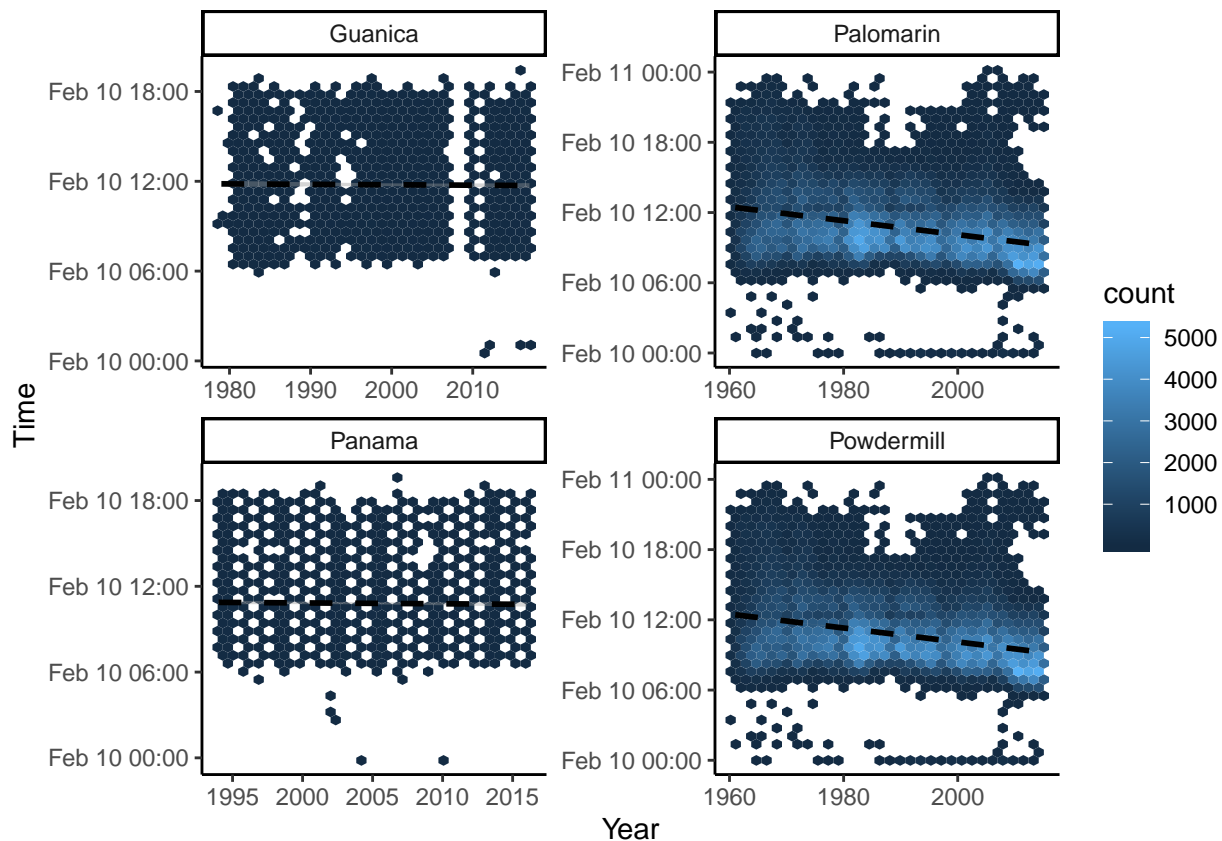
```
## [1] 0.00179748
```

```
mod.tarsus.sum$coefficients[2,4] # p-value
```

```
## [1] 0.2683156
```

Still bad! Well, I hope mass is a good metric...

Lastly, let's check to make sure there's not a correlation between time of capture and year, to make sure we aren't introducing a systematic bias into our data related to changing banding practices.

Seems like there are some trends in the data. Let's check with some simple linear models.

```
# all data
master.time$yr.num <- as.numeric(master.time$Year)
master.time$time.num <- as.numeric(master.time$Time)
mod.time <- lm(time.num ~ yr.num, master.time) # summarize
mod.time.sum <- summary(mod.time) # summarize model
print(mod.time.sum)
```

```
##
## Call:
## lm(formula = time.num ~ yr.num, data = master.time)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -43871  -5975  -1398   4790  51782
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.582e+09  9.292e+02 1702237   <2e-16 ***
## yr.num      -2.115e+02  4.669e-01    -453   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8552 on 1439571 degrees of freedom
## Multiple R-squared:  0.1248, Adjusted R-squared:  0.1248
## F-statistic: 2.052e+05 on 1 and 1439571 DF,  p-value: < 2.2e-16
```

11

```
# guanica
guanica.full$yr.num <- as.numeric(guanica.full$Year)
guanica.full$time.num <- as.numeric(guanica.full$Time)
mod2.time <- lm(time.num ~ yr.num, guanica.full) # summarize
mod2.time.sum <- summary(mod2.time) # summarize model
print(mod2.time.sum)
```

```
##
## Call:
## lm(formula = time.num ~ yr.num, data = guanica.full)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -40260  -11640   -2817   11610   27102
##
## Coefficients:
##                Estimate Std. Error    t value Pr(>|t|)
## (Intercept)   1.581e+09  3.036e+04  52084.530   <2e-16 ***
## yr.num       -1.042e+01  1.517e+01     -0.687    0.492
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12910 on 5548 degrees of freedom
## Multiple R-squared:  8.512e-05,  Adjusted R-squared:  -9.511e-05
## F-statistic: 0.4723 on 1 and 5548 DF,  p-value: 0.492
```

```
# panama
panama.full$yr.num <- as.numeric(panama.full$Year)
panama.full$time.num <- as.numeric(panama.full$Time)
mod3.time <- lm(time.num ~ yr.num, panama.full) # summarize
mod3.time.sum <- summary(mod3.time) # summarize model
print(mod3.time.sum)
```

```
##
## Call:
## lm(formula = time.num ~ yr.num, data = panama.full)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -38931   -9426   -2886    7969   31636
##
## Coefficients:
##                Estimate Std. Error    t value Pr(>|t|)
## (Intercept)   1.581e+09  4.380e+04  36107.901   <2e-16 ***
## yr.num       -2.224e+01  2.186e+01     -1.018    0.309
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11400 on 7335 degrees of freedom
## Multiple R-squared:  0.0001412,  Adjusted R-squared:  4.846e-06
## F-statistic: 1.036 on 1 and 7335 DF,  p-value: 0.3089
```

```
# palomarin
palo.full$yr.num <- as.numeric(palo.full$Year)
palo.full$time.num <- as.numeric(palo.full$Time)
```

```r
mod4.time <- lm(time.num ~ yr.num, palo.full) # summarize
mod4.time.sum <- summary(mod4.time) # summarize model
print(mod4.time.sum)
```

```
##
## Call:
## lm(formula = time.num ~ yr.num, data = palo.full)
##
## Residuals:
##     Min     1Q Median     3Q    Max
## -43901  -5927  -1344   4787  51889
##
## Coefficients:
##               Estimate Std. Error   t value Pr(>|t|)
## (Intercept)  1.582e+09  1.312e+03 1205412.6   <2e-16 ***
## yr.num      -2.144e+02  6.594e-01    -325.1   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8500 on 713341 degrees of freedom
## Multiple R-squared:  0.129,  Adjusted R-squared:  0.129
## F-statistic: 1.057e+05 on 1 and 713341 DF,  p-value: < 2.2e-16
```

```r
# powdermill
powdermill.full$yr.num <- as.numeric(powdermill.full$Year)
powdermill.full$time.num <- as.numeric(powdermill.full$Time)
mod5.time <- lm(time.num ~ yr.num, powdermill.full) # summarize
mod5.time.sum <- summary(mod5.time) # summarize model
print(mod5.time.sum)
```

```
##
## Call:
## lm(formula = time.num ~ yr.num, data = powdermill.full)
##
## Residuals:
##     Min     1Q Median     3Q    Max
## -43901  -5927  -1344   4787  51889
##
## Coefficients:
##               Estimate Std. Error   t value Pr(>|t|)
## (Intercept)  1.582e+09  1.312e+03 1205412.6   <2e-16 ***
## yr.num      -2.144e+02  6.594e-01    -325.1   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8500 on 713341 degrees of freedom
## Multiple R-squared:  0.129,  Adjusted R-squared:  0.129
## F-statistic: 1.057e+05 on 1 and 713341 DF,  p-value: < 2.2e-16
```

So, a definite trend towards earlier captures in later years for the Powdermill and Palomarin data. We should discuss how to deal with this.

Assembling the data frame for statistical analyses

We next take our raw data and extract the information we need from each species using the `make_stats_df()` function from our `00_functions.R` script. Here's what's under the hood:

```r
make_stats_df <- function(file_path, site_name, output_path){

  # load data
  data.all <- read.csv(file = file_path, fileEncoding="latin1")[-1]

  # subset temp data, merge with df
  data.temps <- cbind.data.frame(temp.df$year, temp.df[,colnames(temp.df)==site_name])
  data.precip <- cbind.data.frame(precip.df$year, precip.df[,colnames(precip.df)==site_name])
  colnames(data.temps) <- c("year", "MAT")
  colnames(data.precip) <- c("year", "MAP")
  data.temps$MAT <- as.numeric(as.character(data.temps$MAT))
  data.precip$MAP <- as.numeric(as.character(data.precip$MAP))
  data.all <- merge(data.all, data.temps, by.x="year", by.y="year", all.y=FALSE)
  data.all <- merge(data.all, data.precip, by.x="year", by.y="year", all.y=FALSE)

  # loop to get mean mass change per species
  data.delta <- list()
  for(i in unique(data.all$species)){
    tmp <- data.all[data.all$species==i,]
    n <- nrow(tmp) # get sample size
    n.years <- length(unique(tmp$year)) # get number of years
    min.year <- min(tmp$year)
    mass.start <- mean(tmp[tmp$year==min.year,]$mass) # average mass for starting year
    mod.mass <- lm(mass ~ year, tmp) # simple regression for mass change
    mod.mass.sum <- summary(mod.mass)
    slope.mass <- mod.mass.sum$coefficients[2,1] # get mass slope
    r2.mass <- mod.mass.sum$r.squared # get mass r-squared
    se.mass <- mod.mass.sum$coefficients[2,2] # get mass SE
    mod.temp <- lm(MAT ~ year, tmp) # simple regression for temp change
    mod.temp.sum <- summary(mod.temp)
    slope.temp <- mod.temp.sum$coefficients[2,1] # get temp slope
    r2.temp <- mod.temp.sum$r.squared # get temp r-squared
    mod.precip <- lm(MAP ~ year, tmp) # simple regression for temp change
    mod.precip.sum <- summary(mod.precip)
    slope.precip <- mod.precip.sum$coefficients[2,1] # get temp slope
    r2.precip <- mod.precip.sum$r.squared # get temp r-squared
    data.delta[[i]] <- cbind.data.frame(i,n,n.years,slope.mass,r2.mass,se.mass,
                                        slope.temp,r2.temp,slope.precip,r2.precip,mass.start)
  }

  # assemble dataframe
  data.mass.df <- do.call(rbind, data.delta)
  rownames(data.mass.df) <- NULL
  colnames(data.mass.df) <- c("species","sample_size","no_years","slope_mass",
                              "variance_mass","se_mass","slope_temp", "variance_temp",
                              "slope_precip", "variance_precip","starting_mass")
  data.mass.df$lat <- rep(latlong.df[latlong.df$site==site_name,]$lat, nrow(data.mass.df))
  data.mass.df$long <- rep(latlong.df[latlong.df$site==site_name,]$long, nrow(data.mass.df))

  write.csv(data.mass.df, file=output_path)
```

```
}
```

What we're doing here is looping over species in a particular site and 1) getting the number of years of data for that species; 2) getting the mean starting mass of the species in the first year we gave data for it; 3) getting the slope, R^2, and standard error of the change in mass of the speciesby regressing mass against year; 4) getting the slope, R^2, and standard error of the change in temperature at the site that species is present at by regressing temperature against year; and 5) getting the slope, R^2, and standard error of the change in precipitation at the site that species at present in by regressing precipitation against year. We then merge all the site-specific dataframes and drop each species that occurs more than once in any site subsequent to its first appearance in the dataframe (in the order Brazil, Panama, Guanica, Powdermill, Palomarin, Teton Science School, Waterfall Glen) using the following line of code:

```r
master.df <- master.df %>% group_by(species) %>% filter(row_number() == 1)
```

Statistical analysis

Because we're interested in body size changes in species, and species are evolutionarily non-independent units, we're going to be analysing our data using phylogenetic multilevel models. This framework will allow us to include a varying intercept for each species, but dictate that these intercepts are dependent on an underlying phylogenetic distance matrix. We'll include the slope of temperature change, initial mass of each species, and latitude as predictors in the model, and weight by the number of years of data we have for each species. We'll be using the R package phangorn—specifically, its pgls.SEy() function, which allows us to account for sampling variance.

First, we'll scale our variables to the same order of magnitude:

```r
# scale variables
master.df$starting_mass <- master.df$starting_mass/1000
master.df$lat <- master.df$lat/1000

male.df$starting_mass <- male.df$starting_mass/1000
male.df$lat <- male.df$lat/1000

female.df$starting_mass <- female.df$starting_mass/1000
female.df$lat <- female.df$lat/1000

powdermill.df$starting_mass <- powdermill.df$starting_mass/1000
powdermill.df$lat <- powdermill.df$lat/1000

palo.df$starting_mass <- palo.df$starting_mass/1000
palo.df$lat <- palo.df$lat/1000
```

Next, we'll assemble our variables of interest into a dataframe for our full / sex-blind dataset, and run a PGLS model with slope of change in mass as our response variable, slope of temperature change, slope of change in precipitation, starting mass, and latitude as our predictor variables, standard error of change mass as our sampling variance measure, all while controlling for phylogeny.

```r
### pgls approach

# create data frame, all birds
X1 <- cbind.data.frame(master.df$slope_mass, master.df$slope_temp,
                       master.df$slope_precip, master.df$starting_mass,
                       master.df$lat)
colnames(X1) <- c("slope_mass","slope_temp", "slope_precip","starting_mass","lat")
species1 <- as.vector(as.character(master.df$species))
rownames(X1) <- species1
```

```
# get SE of sample size
SE1 <-setNames(master.df$se_mass, master.tree$tip.label)

# fit model, all birds
fit1<-pgls.SEy(slope_mass ~ slope_temp + slope_precip + starting_mass + lat,
               data=X1, se=SE1, tree=master.tree, method="ML")
summary(fit1)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: slope_mass ~ slope_temp + slope_precip + starting_mass + lat
##   Data: X1
##         AIC       BIC   logLik
##   -499.4816 -478.3528 255.7408
##
## Correlation Structure: corBrownian
##   Formula: ~1
##   Parameter estimate(s):
## numeric(0)
## Variance function:
##   Structure: fixed weights
##   Formula: ~vf
##
## Coefficients:
##                   Value Std.Error    t-value p-value
## (Intercept)   -0.0362624 0.0551928 -0.6570133  0.5118
## slope_temp    -0.2444298 0.2907133 -0.8407931  0.4013
## slope_precip  -0.0167951 0.0729105 -0.2303515  0.8180
## starting_mass -0.2555701 0.2781498 -0.9188216  0.3591
## lat            0.6590960 0.5466186  1.2057694  0.2291
##
##   Correlation:
##               (Intr) slp_tm slp_pr strtn_
## slope_temp    -0.141
## slope_precip  -0.292  0.313
## starting_mass -0.408  0.000  0.063
## lat           -0.283  0.082  0.628  0.057
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -5.77141978  0.05171455  0.14121161  0.22988543  2.51821424
##
## Residual standard error: 0.7452966
## Degrees of freedom: 250 total; 245 residual
```

So, while the relationships look about what we would expect, there are no significant predictors.

Let's repeat this for our male-only data…

```
# create data frame, males
X2 <- cbind.data.frame(male.df$slope_mass, male.df$slope_temp,
                       male.df$slope_precip, male.df$starting_mass, male.df$lat)
colnames(X2) <- c("slope_mass","slope_temp", "slope_precip",
                  "starting_mass","lat")
species2 <- as.vector(as.character(male.df$species))
rownames(X2) <- species2
```

```r
# get SE of sample size
SE2 <-setNames(male.df$se_mass, male.tree$tip.label)

# fit model, males
fit2<-pgls.SEy(slope_mass ~ slope_temp + slope_precip +
                starting_mass + lat,
            data=X2, se=SE2, tree=male.tree, method="ML")
summary(fit2)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: slope_mass ~ slope_temp + slope_precip + starting_mass + lat
##   Data: X2
##         AIC       BIC    logLik
##    -385.8786 -368.5365 198.9393
##
## Correlation Structure: corBrownian
##  Formula: ~1
##  Parameter estimate(s):
## numeric(0)
## Variance function:
##  Structure: fixed weights
##  Formula: ~vf
##
## Coefficients:
##                   Value Std.Error   t-value p-value
## (Intercept)    0.0540654 0.0454892  1.188533  0.2368
## slope_temp    -0.3437739 0.1980892 -1.735450  0.0851
## slope_precip  -0.2888597 0.0534845 -5.400811  0.0000
## starting_mass -0.4568879 0.2624652 -1.740756  0.0841
## lat           -1.0068366 0.4613051 -2.182583  0.0309
##
##   Correlation:
##               (Intr) slp_tm slp_pr strtn_
## slope_temp    -0.142
## slope_precip  -0.230  0.520
## starting_mass -0.208  0.000 -0.051
## lat           -0.257  0.213  0.488  0.009
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.18460199 -0.18770429 -0.03825834  0.07976642  1.77631956
##
## Residual standard error: 0.0004125338
## Degrees of freedom: 133 total; 128 residual
```

...our female-only data...

```r
# create data frame, females
X3 <- cbind.data.frame(female.df$slope_mass, female.df$slope_temp,
                    female.df$slope_precip,
                    female.df$starting_mass, female.df$lat)
colnames(X3) <- c("slope_mass","slope_temp","slope_precip",
                "starting_mass","lat")
species3 <- as.vector(as.character(female.df$species))
```

```
rownames(X3) <- species3

# get SE of sample size
SE3 <-setNames(female.df$se_mass, female.tree$tip.label)

# fit model, females
fit3<-pgls.SEy(slope_mass ~ slope_temp + slope_precip + starting_mass + lat,
             data=X3, se=SE3, tree=female.tree, method="ML")
summary(fit3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: slope_mass ~ slope_temp + slope_precip + starting_mass + lat
##   Data: X3
##         AIC       BIC    logLik
##   -383.5727 -365.5089 197.7864
##
## Correlation Structure: corBrownian
##  Formula: ~1
##  Parameter estimate(s):
## numeric(0)
## Variance function:
##  Structure: fixed weights
##  Formula: ~vf
##
## Coefficients:
##                    Value Std.Error    t-value p-value
## (Intercept)   -0.0381266 0.0528912 -0.7208507  0.4722
## slope_temp     0.5962575 0.1927291  3.0937598  0.0024
## slope_precip   0.1745832 0.0628448  2.7780063  0.0062
## starting_mass  0.9859784 0.3458430  2.8509425  0.0050
## lat            0.4614300 0.5686409  0.8114612  0.4184
##
##  Correlation:
##              (Intr) slp_tm slp_pr strtn_
## slope_temp   -0.064
## slope_precip -0.275  0.089
## starting_mass -0.254  0.130  0.099
## lat          -0.276  0.021  0.563  0.081
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -0.96713861 -0.26240175 -0.11621530  0.05486568  1.65760595
##
## Residual standard error: 0.1146413
## Degrees of freedom: 150 total; 145 residual
```

...data from Palomarin, where previous researchers have NOT seen coordinated responses to climate change...

```
# create data frame, palomarin
X4 <- cbind.data.frame(palo.df$slope_mass, palo.df$slope_temp,
                    palo.df$slope_precip,
                    palo.df$starting_mass, palo.df$lat)
colnames(X4) <- c("slope_mass","slope_temp", "slope_precip",
                "starting_mass","lat")
species4 <- as.vector(as.character(palo.df$species))
```

```r
rownames(X4) <- species4

# get SE of sample size
SE4 <-setNames(palo.df$se_mass, palo.tree$tip.label)

# fit model, palomarin
fit4<-pgls.SEy(slope_mass ~ slope_temp + slope_precip + starting_mass,
               data=X4, se=SE4, tree=palo.tree, method="ML")
summary(fit4)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: slope_mass ~ slope_temp + slope_precip + starting_mass
##   Data: X4
##        AIC      BIC    logLik
##   -94.18154 -86.12695 52.09077
##
## Correlation Structure: corBrownian
##  Formula: ~1
##  Parameter estimate(s):
## numeric(0)
## Variance function:
##  Structure: fixed weights
##  Formula: ~vf
##
## Coefficients:
##                    Value Std.Error   t-value p-value
## (Intercept)   -0.0132983 0.0505703 -0.262967  0.7942
## slope_temp    -0.1802975 1.4415729 -0.125070  0.9012
## slope_precip  -0.2880078 0.2365720 -1.217421  0.2321
## starting_mass  1.5741064 0.4106501  3.833206  0.0005
##
##  Correlation:
##               (Intr) slp_tm slp_pr
## slope_temp    -0.117
## slope_precip   0.098 -0.105
## starting_mass -0.286  0.103 -0.305
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -1.98570829 -0.21597345 -0.07033377  0.04917365  1.15001545
##
## Residual standard error: 0.0003527163
## Degrees of freedom: 37 total; 33 residual
```

...and Powdermill, where they have, albeit while not controlling for phylogeny.

```r
# create data frame, powdermill
X5 <- cbind.data.frame(powdermill.df$slope_mass, powdermill.df$slope_temp,
                       powdermill.df$slope_precip,
                       powdermill.df$starting_mass, powdermill.df$lat)
colnames(X5) <- c("slope_mass","slope_temp","slope_precip",
                  "starting_mass","lat")
species5 <- as.vector(as.character(powdermill.df$species))
rownames(X5) <- species5
```

```
# get SE of sample size
SE5 <-setNames(powdermill.df$se_mass, powdermill.tree$tip.label)

# fit model, powdermill
fit5<-pgls.SEy(slope_mass ~ slope_temp + slope_precip + starting_mass,
               data=X5, se=SE5, tree=powdermill.tree, method="ML")

fit5.sum <- summary(fit5)
```

So, no effect! Let's chat about this, but in some ways it isn't suprising.

Finally, let's see if controlling for phylogeny is having a significant effect on our results:

```
# simple linear models to see if phylogeny matters
mod1 <- lm(slope_mass ~ slope_temp + slope_precip + starting_mass, master.df)
summary(mod1)
```

```
##
## Call:
## lm(formula = slope_mass ~ slope_temp + slope_precip + starting_mass,
##     data = master.df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.94295 -0.01283 -0.00208  0.01262  0.44633
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.01012    0.01035   0.978 0.329167
## slope_temp    -0.33665    0.36554  -0.921 0.357959
## slope_precip  -0.02693    0.05988  -0.450 0.653305
## starting_mass -0.60343    0.16246  -3.714 0.000252 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08954 on 246 degrees of freedom
## Multiple R-squared:  0.06036,    Adjusted R-squared:  0.0489
## F-statistic: 5.267 on 3 and 246 DF,  p-value: 0.001544
```

For our full dataset, yes——and in the expected way, as starting mass becomes a significant predictor.

```
mod2 <- lm(slope_mass ~ slope_temp + slope_precip + starting_mass, male.df)
summary(mod2)
```

```
##
## Call:
## lm(formula = slope_mass ~ slope_temp + slope_precip + starting_mass,
##     data = male.df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.260319 -0.017606 -0.003012  0.015403  0.197510
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.011258   0.007839   1.436   0.1534
```

```
## slope_temp     0.053998    0.222248    0.243    0.8084
## slope_precip  -0.169387    0.041887   -4.044 8.99e-05 ***
## starting_mass -0.370906    0.161964   -2.290    0.0236 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04902 on 129 degrees of freedom
## Multiple R-squared:  0.2238, Adjusted R-squared:  0.2057
## F-statistic:  12.4 on 3 and 129 DF,  p-value: 3.545e-07
```

Curiously, precipitation becomes significant when you fail to control for phylogeny using our male-only data. But the rest of the models (female-only, Powdermill-only, Palomarin-only) match the full dataset:

```
mod3 <- lm(slope_mass ~ slope_temp + slope_precip + starting_mass, female.df)
summary(mod3)
```

```
##
## Call:
## lm(formula = slope_mass ~ slope_temp + slope_precip + starting_mass,
##     data = female.df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.123817 -0.016380 -0.000182  0.015413  0.253503
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.022327   0.006964  -3.206  0.00165 **
## slope_temp    0.253429   0.198084   1.279  0.20278
## slope_precip  0.040895   0.037308   1.096  0.27482
## starting_mass 0.821718   0.180890   4.543 1.15e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04897 on 146 degrees of freedom
## Multiple R-squared:  0.1332, Adjusted R-squared:  0.1154
## F-statistic: 7.481 on 3 and 146 DF,  p-value: 0.000108
```

```
mod4 <- lm(slope_mass ~ slope_temp + slope_precip + starting_mass, palo.df)
summary(mod4)
```

```
##
## Call:
## lm(formula = slope_mass ~ slope_temp + slope_precip + starting_mass,
##     data = palo.df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.164870 -0.007544  0.005374  0.017532  0.147184
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.02125    0.01306  -1.627    0.113
## slope_temp   -0.39564    1.42748  -0.277    0.783
## slope_precip -0.18146    0.24461  -0.742    0.463
## starting_mass 1.34727    0.30056   4.483  8.4e-05 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0514 on 33 degrees of freedom
## Multiple R-squared:  0.3836, Adjusted R-squared:  0.3275
## F-statistic: 6.845 on 3 and 33 DF,  p-value: 0.001034
```

```
mod5 <- lm(slope_mass ~ slope_temp + slope_precip + starting_mass, powdermill.df)
summary(mod5)
```

```
##
## Call:
## lm(formula = slope_mass ~ slope_temp + slope_precip + starting_mass,
##     data = powdermill.df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.142738 -0.009874  0.000346  0.010862  0.251592
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.02466    0.01421   1.736    0.087 .
## slope_temp    -0.43776    0.46695  -0.937    0.352
## slope_precip  -0.30482    0.32641  -0.934    0.354
## starting_mass -1.00075    0.17396  -5.753 2.13e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0486 on 70 degrees of freedom
## Multiple R-squared:  0.3312, Adjusted R-squared:  0.3025
## F-statistic: 11.55 on 3 and 70 DF,  p-value: 3.067e-06
```