The School of Mathematics

Identifying Local Weather Conditions Affecting Raptor Migration at Allegheny Front, USA

by

Yinying Xu, s1928699

July 2020

Supervised by
Dr. Nicole Augustin
Dr. Victor Elvira

Contents

Own Work Declaration

I confirm that the work of this MSc dissertation project is my own expect where otherwise declared.

Yinying Xu.

Word count

4796 words (only applies to the main body of the text)

Executive summary

As migrating birds can be affected under global warming to change migration timing(Jenni & Kry 2003), we are interested in exploring local weather factors affecting abundance of migratory raptors. We used generalized linear mixed model considering the observers and counters involved with a wide range of ability levels. We found that there is a significant positive relationship between some local environment factors, i.e. wind speed, temperature, humidity,barometric pressure, visibility, and overall raptor abundance. However, precipitation can affect raptor migration negatively. Among groups of species we are studying, hawks can be most significantly influenced by changes in local weather patterns, while buzzards are the opposite.

1 Introduction

Migration is one of adaptation strategies for those animals living at high-latitude breeding area to get through severer local climate on a seasonal basis. Most birds are suitable for relatively long-distance movement along a flyway. In the northern hemisphere, the general pattern of bird migration involves flying north in the spring during reproductive period and heading south in the fall to the wintering grounds in warmer regions (Newton & Brockie 2008). There is evidence that raptors use thermals as their migration strategies, which means that they reply on good weather conditions to support their movement behavior (Spaar & Bruderer 1996).

It has been proved that migrating birds can be influenced under global warming to change migration timing. This can result in migrants starting their seasonal journey earlier or later than usual (Jenni & Kry 2003). However, few researches are available about the effect of local weather patterns on migration behavior. This report aims to contribute to this part by examining how local weather conditions influence raptor migration. To be more specific, the first aim is to quantify the relationship between various weather conditions affecting raptor's flight conditions and the total abundance of migrating raptors, and the other one is to examine the influence of weather conditions on the individual counts of groups of raptor species observed (eagles, falcons, hawks, buzzards). We hope that a better understanding of the effect of changes in local weather linked to climate change on raptor migration can help researchers to recognize species most exposed to such changes.

We will analyze a sample of 1319 observations including two migration period in 2018 and one in 2019 respectively from a major migration observation stations, Allegheny Front, Pennsylvania, USA, for monitoring and counting migratory raptors including eagles, falcons, hawks and buzzards. Our goal is to identify the most relevant factors for the explanation of migrating raptor behavior among the following ones: migration period, wind speed, wind direction, temperature, humidity, barometric pressure, cloud cover, visibility, precipitation levels.

2 Data analysis and manipulation

2.1 Data source

Observation records from Spring migration 2018 to end of Spring migration 2019 were obtained by volunteers who monitored migratory raptors passing over the site during daylight hours at an important southeast escarpment in the Allegheny Front, located in southern Pennsylvania, USA. These records provide information about local weather, the counts of observed migrants up to 16 known species (mostly can be grouped in eagles, falcons, hawks and buzzards), birds' flight behavior, the IDs of the counter and observers present, and the duration at a single observation.

2.2 Characteristics of local weather conditions

In Figure 1, we can see that total counts of migratory raptors are unstable over the migrating period and there are a great number of zero counts available during the whole observation period. The large number of raptors are more likely to be observed during migrating journey in the autumn than spring. Thus we will consider migration period, contrasting Spring from mid-February to mid-May versus Fall from mid-August to November in the model.

Temperature and barometric pressure can be included in the model, considering that some species, e.g. broad-winged hawks, eagles and vultures, tend to migrate using thermals during the day (Wikipedia contributors 2020). They are reliant on hot air to assist with their migration by using it for lift and soar. Figure 2 shows that more frequent and larger counts are presented when it is warm enough, around 20 degrees Celsius. Most of barometric pressure is ranging from 26 to 30 inches per Mercury(in Hg) but several points at nearly zero and at around 40 in HG are implausible. In that case, we looked further at the HawkCount databases for the exact day with such outliers and then imputed a reasonable value to consistent with other observations on the same date. In addition, 57 points at barometric pressure are missing, accounting for 4.32%

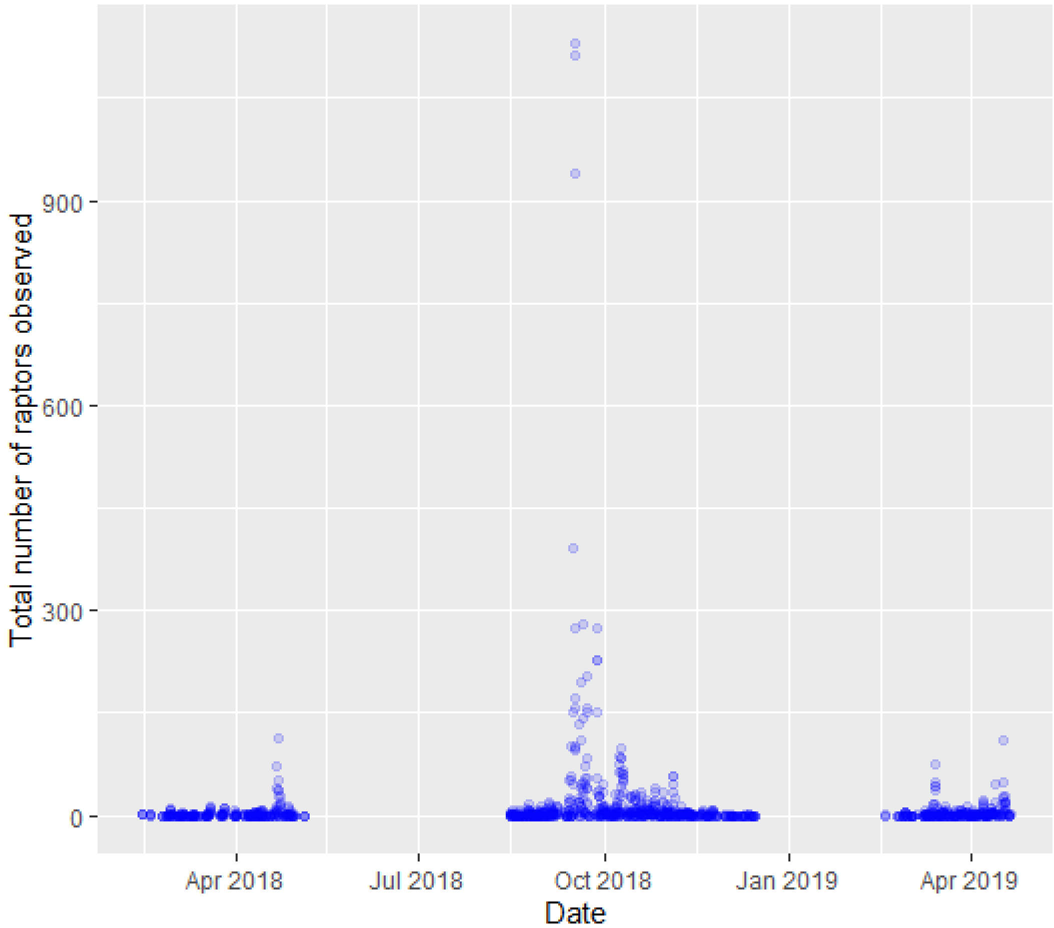

Figure 1: Distribution of total number of raptors observed from Spring migration 2018 to end of Spring migration 2019

of total records. The standard error of that variables (0.0114) is so low that a single imputation strategy could be used (Enders 2010). Thus, each missing value of air pressure can be filled by the overall mean of the observed values for that variable. The only opposite effect of this mean imputation is that it will reduce the standard deviation of that predictor variable, attenuating the variability of the data.

We will also consider humidity measured in percent as numeric predictor and precipitation levels as factor one. In order to read data easily, We plotted precipitation levels against squared root of total counts as a few points are much larger than the bulk of the data.(See Figure 3) we can see that there are a fewer cases with high level of dispersion when encountering drizzle, thunderstorm, snow and wind-driven dust/sand/snow. Thus we grouped all of them as a new level. It also shows that the highest average number of raptors observed is in the precipitation categories of None, and Haze or Fog. The distribution of total counts in None is also similar to that in Haze or Fog, but with more large values of outliers. This implies that raptors might be more active on the day with lower precipitation. It is true that their plumages are more likely to be saturated by torrential rain, which increases wind loading and consumes body heat (Newton 2007). However, in Figure 4, we found that more large values of total counts with high humidity were recorded than with low humidity. The issue for analysis is that there are 23 missing values of precipitation levels available, accounting for 1.75% of total records and 45 of humidity for 3.41%. After looking further in the detailed description of weather condition on the exact day with those missing values, we found that most of them were nearly sunny the whole day, except two records on March 16th, 2019 with some uncertain snow flurries. Thus we deleted two unsure records and imputed none for the rest. In term of missing values of humidity, we used MICE package, taking account of uncertainty in that missing values. We assumed that the missing data in humidity were Missing at Random (MAR). This means that the probability that a value is missing depends only on observed value and can be predicted using them (Enders 2010). Those missing values of humidity were imputed by the predicted values according to observed variables that might be related with humidity, i.e. temperature, precipitation, barometric pressure.

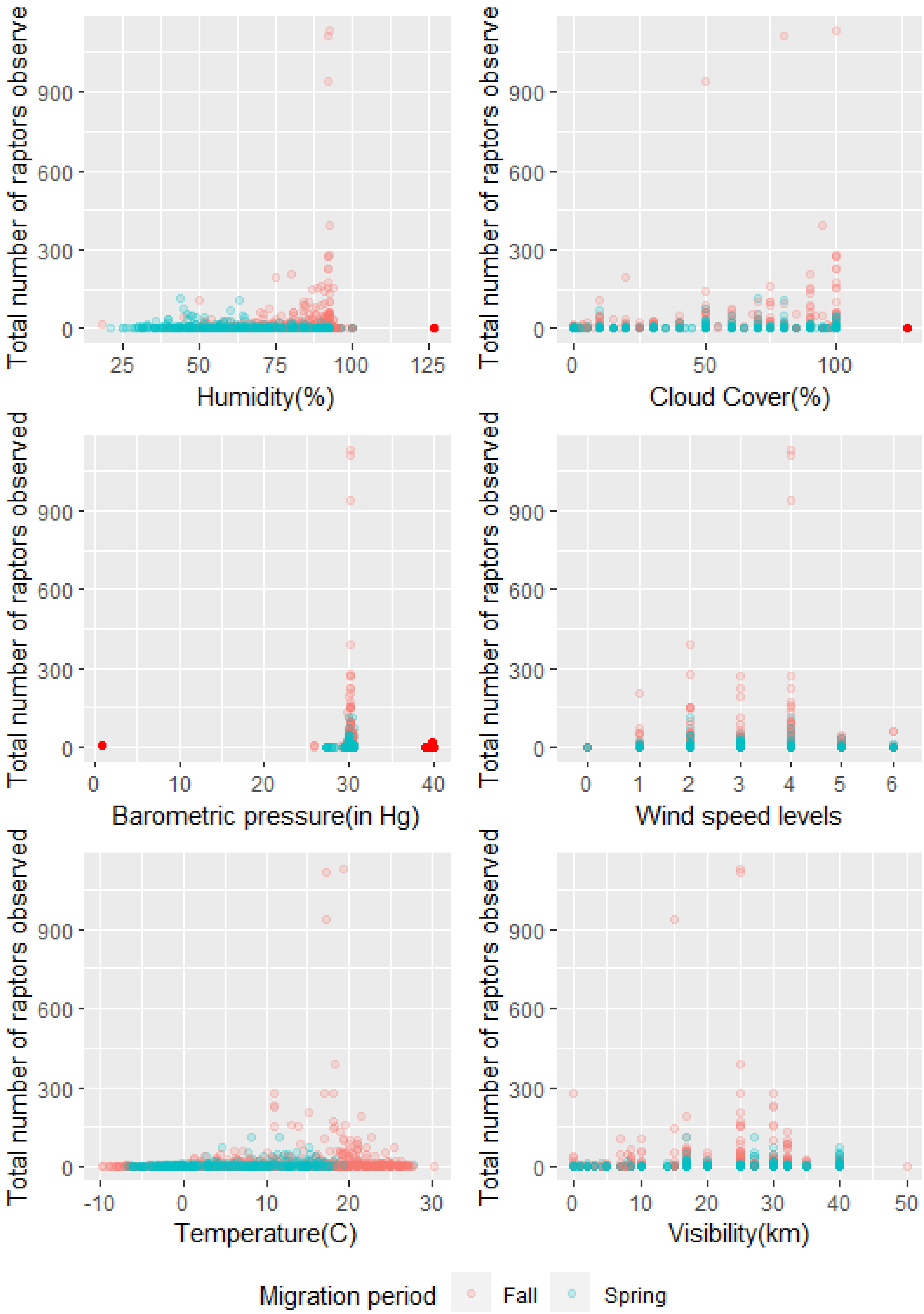

Figure 2: Plots of numerical variables (Humidity, Cloud cover, Barometric pressure, Wind speed, Temperature) against total number of raptors observed. (Outliers shown as red points)

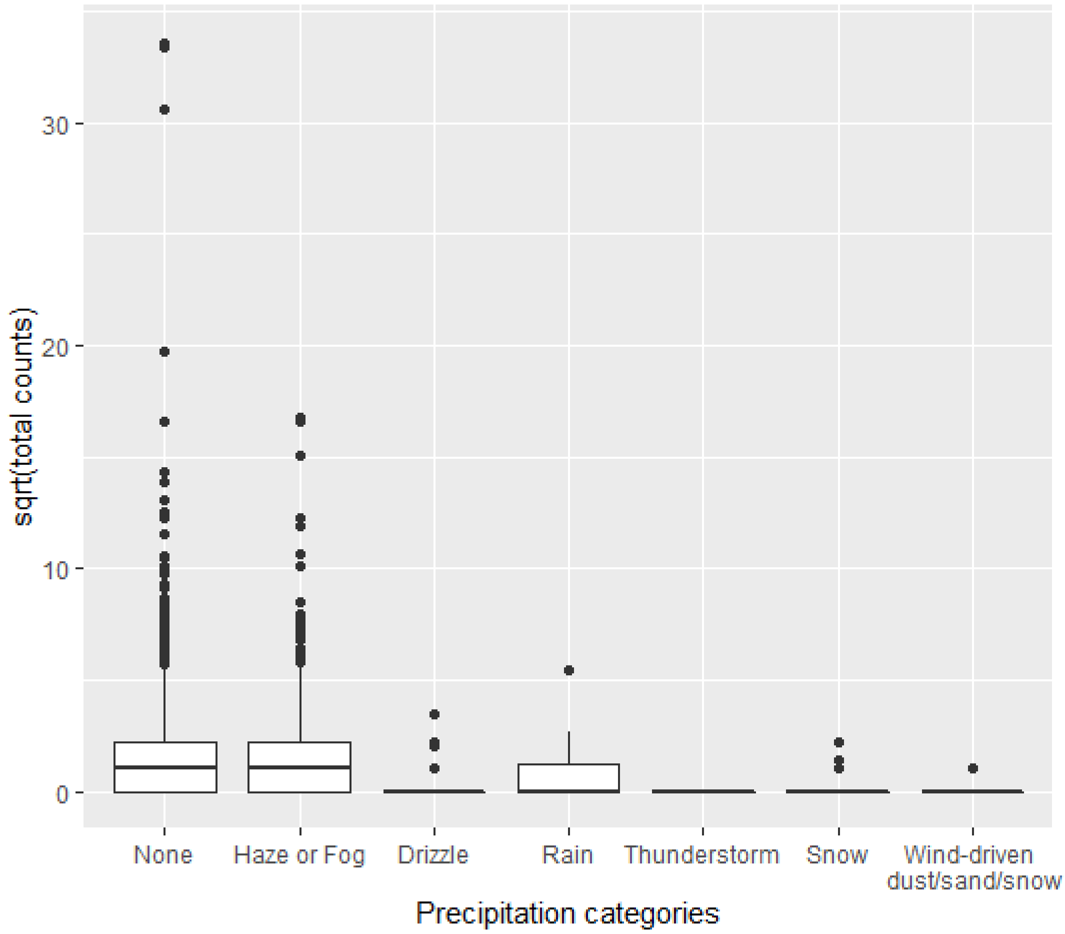

Figure 3: Distribution of precipitation levels against squared root of the total counts

Cloud cover and visibility, measured in percent and kilometers (km) restively, are also worth well to be considered as these may impact on ability to spot high-flying raptors for volunteers. It is interesting to note that more raptors are observed with a thick cloud cover than a thin one. (See Figure 2) We treated both predictors as numeric but two records on November 17th, 2018 and March, 16th, 2019 respectively with unknown visibility were delete for further modeling.

It is well known that migrants usually utilize the high wind speeds to save their energy during migration. They would adjust their flight heights to meet the most beneficial wind speeds and directions and may start their migrating journey at all until favorable conditions occur. By contrast, the birds would be forced by headwinds to fly for longer, exhausting themselves (Newton 2007). In Figure 2, we can see that the large total counts are most likely to occur when the wind speed is relatively high, between level 2 and 4. Although we can't see a clear relationship between wind direction and the squared root of total counts from Figure 4), we included this predictor variables in the full model to find out whether there is actually an association with the counts. However, flight height and flight direction will not be considered, mainly because most of them were not recorded when there was no migrating raptor observed. There are 7 levels of wind strengths, ranging from less than 1km/h to 39-49km/h, which are recorded as numeric codes from 0 to 6. We treated it as numeric because the value of those numeric codes implies wind strength. As for wind direction, there were 16 various levels available, which seems rather a lot if specified as a fixed effect. Based on the fact that migrants are going to north in the fall and heading to south in the spring, we grouped multiple levels of wind direction as three factor levels: North, South and Other (including west, east and variable). There are 10 points at wind direction that were not recorded but we imputed as "Other" by information provided in the detailed description on that observation period with missing values from the HawkCount databases.

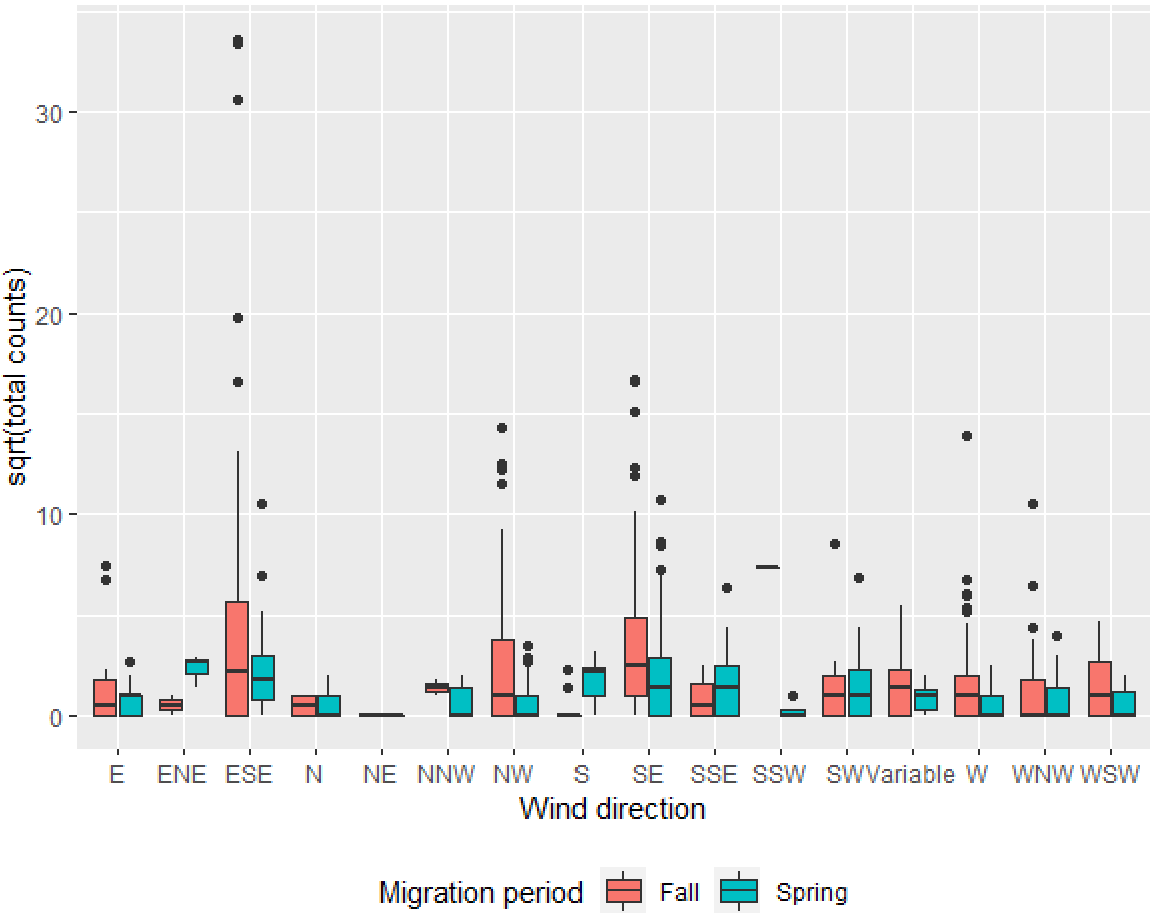

Figure 4: Distribution of wind direction against squared root of the total counts

In Figure 2, We found that the outliers are available in the humidity and cloud cover, both at point of 127% on August 31st, 2018, which is at least physically impossible. Although humidity could go beyond 100 percent in the environment, known as supersaturation, two same numerical values in two different variables seems not sensible. Thus we deleted those two relating records. Additionally, we also did not include one record on August 18th, 2018 when zero bird was observed with counter and observer absent over a period lasting zero minutes. Finally, we have a total of 1312 distinct observation period include in the model. More detailed data manipulation could be found in Appendix.

2.3 Characteristics of the counter and observers

For each observation, there was at least one counter present, and observers might be absent but up to four of them could join the watch. The data for them was anonymous, i.e. they were recorded as numeric codes, and they might vary in the ability to spot birds. This would cause observer and count effects as there were different skill levels of observers and counter, and inconsistent numbers of observers during disparate observation period. It is evidence that there were significant observer differences found which impacted on the estimates of overall abundance (Meer & Camphuysen 1996). Thus observer and counter effects should be specified as crossed random effects in the model. This tested how the counts recorded varied among different counter and observer(s) when at least one of the levels of count effect was appeared in multiple levels of observer effect (Faraway 2016). However, the problem for measuring observer and counter effect is that counter and observer(s) sometimes were not recorded, or a same ID was recorded for both the counter and Observer by mistake. For instance, we had a duration of 60 minutes and a total observation period of 120 minutes. It should have recorded two IDs for volunteers involved but only one counter's ID was written down. We supposed that the ID for the observer was forgotten to be recorded. In that case, we created a unique ID manually for each missing value found when the recorded number of volunteers involved was inconsistent with the total duration of all watches divided by the duration of watch. On the occasion where the calculated values were smaller than the recorded ones, we supposed that one individual watched for a different duration of time to another and left earlier than we expect. In terms of mistake in writing wrong ID for observers, we also gave them artificial IDs to replace the error ones, avoiding two same ID presented at a single watch. Another thing to consider when measuring observer effect is the instability in the number of observer present. To tackle this problem, we sorted out the relatively most skilled observer as a representative among more than one observers present for analysis by counting their frequency in the overall sample period. (See further Appendix)

plot/replaced.png

Figure 5: The boxplot of indication for the replacement of lead pipes against the number of samples detected lead

plot/builtafter1970.png

Figure 6: Plot of the proportion of houses built after 1970 against the average of lead concentration

3 Models and results: overall abundance

3.1 Rate model

When the response is a count of some events, the number of events may rely on a size variable that controls the probability of events happening (Faraway 2016). Time spent per watch on counting the number of different raptor species was varied, with common 30-minute or an hour duration. Obviously, the duration of watch made difference on the counts of different species. The longer volunteers were there, the more birds can be seen. Sometimes, they did not detect any birds simply because there was no migratory raptor passing over the site on that day. Therefore, we took account of the differing length of observation period of each distinct observation and assumed that the counts might be directly proportional to the amount of time spent. We did so by modeling the rate of counts (TOTAL) divided by the duration of watch (Duration).

3.2 Generalized linear mixed model

We applied generalized linear mixed model (GLMMs) to estimate the parameters of weather conditions in response to the overall counts of raptor migrants, and also to groups of different species, i.e. eagles, falcons, hawks, buzzards (we will talk about it later). GLMMs is commonly used in ecology and evolution, which integrates the characteristics of linear mixed model including random effects, with generalized linear models analyzing nonmoral data by using link functions and exponential family distribution (Bolker et al. 2009). We specifies that the total number of raptors observed (TOTAL) and the responses to different weather conditions could vary randomly across different observer(s) and the counter. To be more specific, in our model, the fixed -effect parameters are the effects of following predictor variables:

1. Numeric variables:

    - Wind speed(Wind.Spd): numeric code from 0 to 6 indicating wind strength (increasing scale);
    - Temperature (Temp): measured in degree Celsius;
    - Humidity: measured in percent;
    - Barometric pressure (BARO): measured in Inch of mercury;
    - Cloud cover (Cloud.Cov): measured in percent;

- Visibility: measured in kilometers.

2. Categorical variables:

   - Period: contrasting Spring versus Fall;
   - Wind direction (Wind,Dir): three levels, i.e. N (denotes the main direction of north), S (denotes the main direction of south), other (denotes the main direction of west and east, and variable);
   - Precipitation: four levels, i.e. None, 1 (denotes haze or fog), 2/4/5/6 (denotes drizzle, thunderstorm, snow, and wind-driven dust/sand/snow), 3 (denotes rain).

and the random-effect parameters are the standard deviations of skilled observers and counter effects.

We considered the poison distribution and the negative binomial distribution as the raptor observation data are count data and chose log link for both distribution (Faraway 2016) The method of Restricted maximum likelihood (REML), a variant that averages over some of the uncertainty in the fixed-effect parameters, was used to estimate the random-effect parameters (i.e. standard deviation), providing less biased result than maximum likelihood (Bolker et al. 2009).

In this report, we used "gam" function from "mgcv" package in R to fit our model. We started with a Poisson model with a logarithmic link including all predictor variables and random effects mentioned above but we found that the data was ovserdispered as the variability of residuals increases with fitted values. In other words, its variance is not equivalent to mean, which implies that it violates the mean variance relationship assumed to some extent. Apart from that, it does not seem to be in the band of residuals simulated from the fitted model. We also had zero-inflation which means that there were too many zeros given the specified distribution. Thus we tried an initial negative binomial "full" model for the rate $TOTAL/Duration$ next. The distribution of the rate is modeled as

$$TOTAL/Duration \sim \mathcal{NB}(\theta,\, p)\,,$$

Where θ is the dispersion parameter, p is the probability of success. We used log link in our model

$$\log(E(TOTAL)/Duration) = X\beta,$$

which can be rearranged as

$$\log(E(TOTAL)) = \log(Duration) + X\beta.$$

Therefore, our "full" model can be shown in detail by the following set of equations

$$\mu = E(TOTAL|u_{observer}, u_{counter}) = exp(\beta_0 + \beta_{\log(Duration)} + \beta_{periodSpring} + \beta_{Wind.Spd} + \beta_{Temp} + \beta_{Humidity} + \beta_{BARO} + \beta_{Cloud.Cov} + \beta_{Visibility} + \beta_{Precipitation1} + \beta_{Precipitation2/4/5/6} + \beta_{Precipitation3} + \beta_{Wind.DirOther} + \beta_{Wind.DirS} + u),$$

$$\textbf{Decomposition.rate} = 13.6\textbf{Growth.rate.f2} - 5.92\textbf{Growth.rate.f4} - 5.83\textbf{Weighted.tolerance\_High} - 38.98\textbf{interaction23} + 23.84\textbf{interaction25} + 13.54\textbf{interaction45}34.54$$

$$u_{observer} \sim \mathcal{N}(0,\, \sigma^2_{u_{observer}})\,,$$

$$u_{counter} \sim \mathcal{N}(0,\, \sigma^2_{u_{counter}})\,,$$

$$\sigma^2 = Var(TOTAL|u_{observer}, u_{counter}) = \mu + \mu^2/\theta,$$

Where $u_{observer}$ and $u_{counter}$ is a observer and a counter specific random effect respectively, μ is the average number of the total counts of raptors observed and β's are regression coefficients with subscript denoting predictor/level (with 0 denoting intercept).

3.3 Model checking

The full detailed process of checking model is included in the Appendix. Here we give the brief description of the result of model checking. On the one hand, we looked at the diagnostic of the residuals versus fitted values, which shows that the mean variance relationship seems to be assumed well, i.e. $\sigma^2 = \mu(1 + \mu/\theta)$. Looking on the diagnostic of deviance residuals against theoretical quantiles, it is also acceptable that only a few points in the tail run out of the band of residuals simulated from the fitted model. Overall, by comparing the residual plots from this model with the ones form a model where the negative binomial distribution was replaced with a Poisson, we can see that the data can be fitted better with the negative binomial distribution. On the other hand, zero counts can be predicted more precisely with the negative binomial distribution than a Poisson distribution and the true number of zeros (576) lie in the range of the distribution of simulated zeros. The coefficient for $\log(Duration)$ is far from 1. This indicates that the total number of birds observed may be not directly proportional to the duration of watch. Thus we simply included $\log(Duration)$ as a variable in the model, instead of using offset.

3.4 Model comparison and interpretation of results

Then we decided to simplify our full model by backward selection based on the Akaike information criterion(AIC), an information-theoretic tool that uses deviance as a measure of fit but also incorporate a penalty for the number of parameters (Bolker et al. 2009). Generally, the lower the AIC the better the fit. In this case, we begin with a full model with the AIC value at 6082.81 including all the predictor variables and random effect, and drop one predictor variable at each candidate model until none variables could be dropped. We repeated the above steps twice until none model with a lower AIC can be found. Finally, we chose one simplifies candidate model that removed the covariates of cloud cover and wind director, with the lowest AIC value at 6078.873, as our final model. We also found that temperature variables is an significant independent variables as when we drop it the AIC arise considerably.

Figure 7 shows the obtained coefficients and corresponding confidence intervals for two models described above, i.e. full model and final model (without Cloud.Cov and Wind.Dir). Generally, in both model, the variability of confidence intervals is the highest in the precipitation levels of rain, while much lower in visibility, temperature and humidity variables.

Table 1:

	Dependent variable:				
	TOTAL	hawk	buzzard	falcon	eagle
	link = log	*link = log*	*link = log*	*link = log*	*link = log*
	(1)	(2)	(3)	(4)	(5)
log(Duration)	3.212***	3.061***		2.880*	1.945***
	(0.406)	(0.452)		(1.545)	(0.688)
periodSpring		−0.270	−0.579*	−0.319	−0.588***
		(0.168)	(0.304)	(0.340)	(0.218)
Wind.DirHead		−0.053	0.057	0.726***	0.036
		(0.150)	(0.268)	(0.249)	(0.185)
Wind.DirTail		0.211	0.833***	0.331	0.112
		(0.155)	(0.277)	(0.305)	(0.208)
Wind.Spd	0.152***	0.134**	0.056	0.502***	0.362***
	(0.047)	(0.052)	(0.096)	(0.099)	(0.067)
Humidity	0.020***	0.021***	−0.003	0.011	0.012**
	(0.004)	(0.005)	(0.008)	(0.008)	(0.006)
Temp	0.069***	0.085***	−0.001	0.067***	0.003
	(0.007)	(0.008)	(0.014)	(0.015)	(0.010)
BARO	0.370**	0.406**	0.097	1.261**	−0.206
	(0.153)	(0.179)	(0.280)	(0.629)	(0.185)
Precipitation1	−0.536***	−0.595***	−0.201		
	(0.147)	(0.165)	(0.292)		
Precipitation2/4/5/6	−1.519***	−1.320***	−1.730**		
	(0.323)	(0.361)	(0.808)		
Precipitation3	−0.924*	−1.847***	1.622*		
	(0.554)	(0.706)	(0.936)		
Visibility	0.028***	0.031***	0.045***	0.031***	0.026***
	(0.007)	(0.007)	(0.013)	(0.012)	(0.008)
Constant	−26.347***	−27.533***	−4.401	−56.738***	−5.991
	(4.900)	(5.731)	(8.490)	(19.975)	(6.206)
Observations	1,312	1,312	1,312	1,312	1,312
Adjusted R^2	0.067	0.052	0.081	0.181	0.124
Log Likelihood	−3,036.114	−2,664.865	−934.593	−394.296	−906.754
UBRE	3,065.235	2,692.462	939.715	401.489	922.684
Note:					*p<0.1; **p<0.05; ***p<0.01

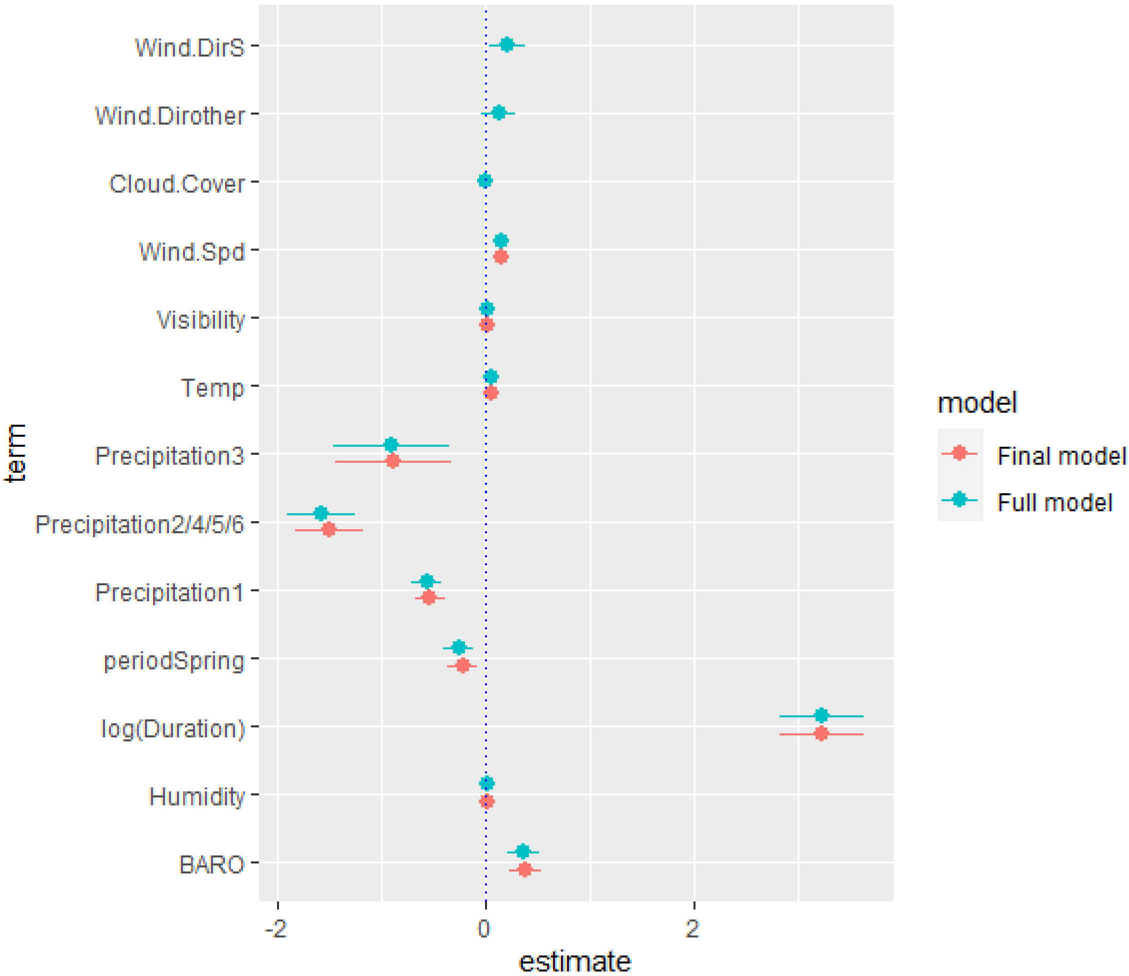

Figure 7: Regression estimates of full model and final model

Table 2: Coefficients of full model and final model

	Dependent variable:	
	TOTAL	
	full model	final model
log(Duration)	3.228***	3.234***
	(0.404)	(0.406)
Wind.DirOther	0.129	
	(0.163)	
Wind.DirS	0.213	
	(0.168)	
periodSpring	−0.252*	−0.218
	(0.145)	(0.145)
Wind.Spd	0.145***	0.148***
	(0.047)	(0.047)
Cloud.Cover	0.002	
	(0.002)	
Humidity	0.016***	0.017***
	(0.004)	(0.004)
Temp	0.063***	0.066***
	(0.007)	(0.007)
BARO	0.364**	0.381**
	(0.154)	(0.153)
Precipitation1	−0.562***	−0.532***
	(0.147)	(0.147)
Precipitation2/4/5/6	−1.571***	−1.502***
	(0.325)	(0.323)
Precipitation3	−0.911	−0.885
	(0.556)	(0.556)
Visibility	0.028***	0.027***
	(0.007)	(0.007)
Constant	−25.942***	−26.366***
	(4.956)	(4.923)
Observations	1,312	1,312
AIC	6082.81	6078.873
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 3: Coefficients of final model and groups of species model

	Dependent variable:				
	TOTAL	hawk	buzzard	falcon	eagle
log(Duration)	3.234***	3.042***		2.875*	1.933***
	(0.406)	(0.452)		(1.538)	(0.686)
periodSpring	−0.218	−0.201	−0.437	−0.489	−0.552***
	(0.145)	(0.162)	(0.297)	(0.317)	(0.204)
Wind.Spd	0.148***	0.131**	0.043	0.482***	0.359***
	(0.047)	(0.052)	(0.097)	(0.099)	(0.067)
Humidity	0.017***	0.020***	−0.008	0.010	0.012**
	(0.004)	(0.005)	(0.008)	(0.008)	(0.006)
Temp	0.066***	0.086***	−0.003	0.062***	0.004
	(0.007)	(0.008)	(0.014)	(0.015)	(0.010)
BARO	0.381**	0.422**	0.214	1.305**	−0.199
	(0.153)	(0.178)	(0.285)	(0.634)	(0.183)
Precipitation1	−0.532***	−0.566***	−0.130		
	(0.147)	(0.165)	(0.292)		
Precipitation2/4/5/6	−1.502***	−1.310***	−1.720**		
	(0.323)	(0.361)	(0.818)		
Precipitation3	−0.885	−1.844***	1.565		
	(0.556)	(0.707)	(0.953)		
Visibility	0.027***	0.032***	0.045***	0.034***	0.026***
	(0.007)	(0.007)	(0.013)	(0.012)	(0.008)
Constant	−26.366***	−27.935***	−7.327	−57.501***	−6.113
	(4.923)	(5.704)	(8.646)	(20.127)	(6.169)
Observations	1,312	1,312	1,312	1,312	1,312
Adjusted R^2	0.067	0.052	0.070	0.204	0.123
Note:					*p<0.1; **p<0.05; ***p<0.01

Before analyzing the effect on weather conditions on the total counts, we used the exponential function to translate back the regression coefficients as we applied log transformation to our dependent variable. Table 4 presents the output of full model and final model with exponential coefficients. In Table 2, we note that spring indicator is only marginally significant (at the 10% level) in the full model, suggesting raptors are less likely to be observed in spring than in fall. Full model also shows that wind direction and cloud cover variables do not have a significant impact on the average total counts. However, dropping cloud cover and wind direction variables in the final model renders the original spring indicator non-significant. Both model show that rain indicator in precipitation variable appears to be irrelevant in predicting the average total number of raptors observed. By contrast, barometric pressure variables significantly affect the average total counts (at the 5% level), while the response is highly more likely to be influenced by the log of duration, wind speed, humidity, temperature, precipitation (expect rain indicator) and visibility variables (at the 1% level).

In Table 4 , we can see that There are four variables, i.e. Wind speed, humidity, temperature and visibility that have positive effects on the rate of birds observed on average. To be more specific, an extra level of wind speed increases the total counts by 16% , while an extra unit of temperature or visibility increases the total counts by 7% or 3% respectively. It is noticeable that barometric pressure, among all the weather conditions considered in our final model, has the most dramatic effect on the rate of migratory raptors observed. When the barometric pressure raises 1 unit, it could result in a increase of the total number of raptors observed by 46%. By contrast, the total counts can be most slightly influenced by humidity: a increase of one unit raises the total counts by 2%. In addition, for every unit of the log of duration increase, total number of raptors can be be multiplied by 25.39 times.

Table 4: Exponential Transformed coefficients

	TOTAL	hawk	buzzard	falcon	eagle
log(Duration)	25.387	20.939		17.734	6.911
periodSpring	0.804	0.818	0.646	0.613	0.576
Wind.Spd	1.159	1.140	1.044	1.620	1.432
Humidity	1.017	1.020	0.992	1.010	1.012
Temp	1.068	1.090	0.997	1.064	1.004
BARO	1.464	1.526	1.239	3.688	0.820
Precipitation1	0.587	0.568	0.878		
Precipitation2/4/5/6	0.223	0.270	0.179		
Precipitation3	0.413	0.158	4.784		
Visibility	1.027	1.032	1.046	1.034	1.026
Constant	0.000	0.000	0.001	0.000	0.002

3.5 The significance of random effects

In Table 5, we can be 95 percent confident that the parameters of the distribution of the counter effect (i.e. the standard deviation) is between 0.64 and 1.81, while that of the observer effect is between 0.55 and 1.16. The estimates of the standard deviation of both random effects are very close to 1, which indicates considerable heterogeneity across the counter and observers. In other words, the level of between-group variability is considerable to warrant including random effects in our model.

In order to testing significance of random effects further, we also create three different models:

- Model1: remove counter effect based on the final model;
- Model2: remover observer effect based on the final model;
- Model 3: remove both random effects based on the final model.

Figure 8 shows that the result of comparison of the AIC values of the final model with the random effects of the counter and observers with the three new model that don't take both random effects in account. It shows that counter and observers effects have significant effect on the goodness of the fit of the model as when we remove either random effect the AIC values goes up. Thus we contained both random effects in our model.

Table 5: Standard deviations and 0.95 confidence intervals of random effects

	std.dev	lower	upper
Counter	1.0752	0.6383	1.8114
skilled observer	0.8052	0.5543	1.1696

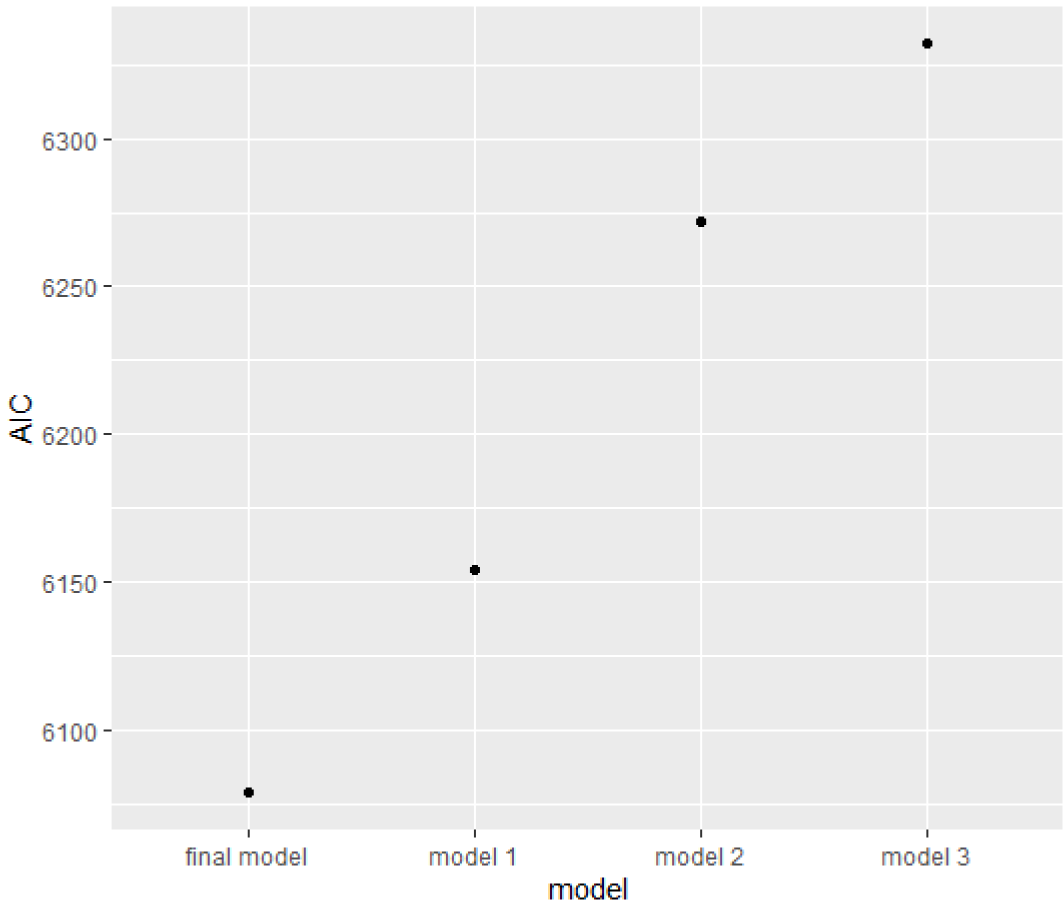

Figure 8: Test of the significance of random effect

4 Models and results: groups of specials

4.1 Model for groups of specials

We examined the effect of weather condition on groups of species (hawks, buzzards, eagles and falcons). We created four new models based on the final model and replace the original response to the individual count rate for each group of species. The duration was not considered in the model of buzzards as the duration for each watch detecting buzzards is lasting 60 minutes which means that the average number of buzzards is not affected by the time spent on that observation. In the models of falcon and eagle, as there are not sufficient samples of observations spotting those species and some levels of precipitation with the p-values at nearly 1 seems to be irrelevant to estimates the effect of weather condition on the total counts of both species, we removed precipitation variable in the models that specify the number of falcon and eagle as response variables.

4.2 Model checking

The models for all groups of species can be fitted well with the negative binomial distribution with log link based on the diagnostic of residual. And also it can predict zeros more precisely than the final model to some extent. Detailed model checking can be seen in the Appendix.

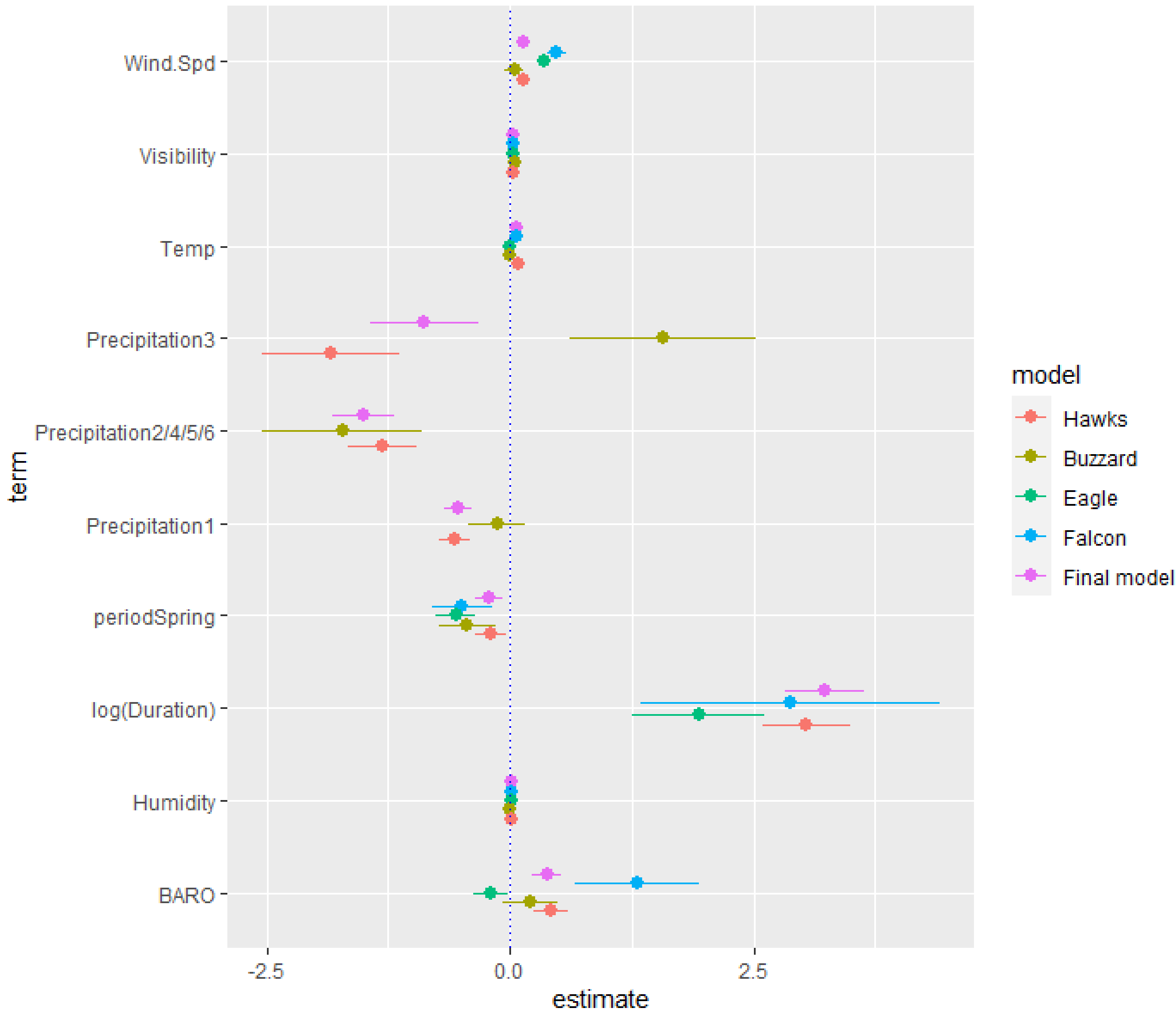

Figure 9: Regression estimates of final model and models of groups of species

4.3 Model interpretation of results

Figure 9 shows the estimated coefficients along with confidence intervals at 95 percent level for the models describing groups of species mentioned above. It notes that the confidence interval of the model for the group of hawks is almost the same as the final one, which implies that the data of the number of hawks observed is dominated in the final model. We also found temperature, wind speed and barometric pressure variables, and haze or fog indicator have uncertainty effects on the migration of buzzards as the confidence intervals of those predictor variables, in the model responding the average number of buzzards observed, contain zero. Similarly, eagles does not seem to be sensible to temperature changes.

In Table 3, it is noticeable that the precipitation level of rain can positively affect the counts of migrating buzzards, while we can have a lower probability to observe hawks on the rainy day than usual. However, this effect of weather condition for buzzards with a wide confidence interval is not significant at standard levels. We can also see that eagle is the only one group of species studying that can be negatively affected by the barometric pressure, but this is also not significant at standard levels. Moreover, buzzards are more likely the group of species who is the least influenced by the weather conditions during migration among the groups of species that we are studying, while hawks tend to be the most affected one. It is evidence that humidity, temperature, precipitation, visibility highly have impact on the number of hawks spotted (at the 1% level), while wind speed and barometric pressure less significantly influence that (at the 5% level) . By contrast, visibility (at the 1% level) ,and the precipitation levels of drizzle, thunderstorm, snow, wind-driven dust or sand or snow (at the 5% level) are only two weather conditions that have significant effect on the counts of buzzards. In term of the group of falcons, under excluding the precipitation variable, migration period of spring and humidity are less likely to influence the counts of it significantly. When refer to eagles, without examining the effect of precipitation, temperature and barometric pressure cannot have significant impact on the counts of them.

Table 4 shows that barometric pressure has the greatest influence on the number of migrating falcons observed among that of other groups of species. For every Inch of mercury of increase in barometric pressure, the counts of falcons passing over the watch site can increase by 3.69 times. Meanwhile, falcons may be the group of species that most reply on the wind speed, followed by eagle. An extra level of wind speed increases the counts of the group of falcons or that of eagles by 62% or 43% respectively. Visibility have a similar effect on the individual count rate of each group of species. An extra kilometer of visibility increases the counts of groups of species detected by between 3% and 5%. We supposed that it may be caused by influencing the probability to spot high-flying birds for volunteers involved in the observation: the greater visibility, the more number of migrants could be counted. All groups of species excluding buzzard are slightly positively affect by temperature and humidity. Both an extra unit of temperature and that of humidity increase the number of groups of species by less than 10%.

5 Conclusions

We analyzed a sample of observation spanning 3 migration period from spring in 2018 to fall in 2019 at Allegheny Front, USA with the goal of exploring which local environmental factors-among wind speed, temperature, wind direction, precipitation, humidity, barometric pressure, cloud cover, visibility are significant to influence raptor migration.

We used GLMMs fitted in negative binomial distribution with log link and choose the best fit model by comparing AIC values of different model. This analysis has shown a significant positive relationship between some weathers conditions, i.e. wind speed, temperature, humidity, barometric pressure, visibility, and overall raptor abundance. However, there was a significant negative association with the precipitation levels of haze or fog drizzle, thunderstorm, snow, and wind-driven dust or sand or snow. The count rate of the group of hawks achieved similar conclusion regarding the direction of the effects of local weather conditions with the total number of raptors observed: the positive associated factors mentioned above can assist with hawks' migration, while, in addition to the negative associated factors discussed above, rain with a significant effect is also not favorable weather patterns. Other groups of species are impacted on changes in local weather conditions with different levels of significance. Buzzards is less likely to be affected by local weather patterns during migration. Falcons are most reliant on the barometric pressure to support migration, while eagles depend on wind speed. The observer and counter effect also have significant effect on the number of birds observed, which suggest that it will increase the variability of count rates.

A first limitation of the analysis is that our model may not be suitable for prediction. This is because there is no sufficient data available, only 1320 observations with some missing value of different predictor variables that we are interested in. Besides, some observers and counters were not recorded so we created a unique artificial IDs for each of them but they may be the person present as a known ID, which would increase the variation of the fixed effect across observers and the counter. The points mentioned above would cause biased estimates.

A second limitation relates to the improvement of fit of the model. It is sensible that some local weather factors could have interactions between each other. However, sample size is not so sufficient that we could examine whether there is interaction effect between two variables we are interested in. if we could have enough samples to identify interaction effect, we would find a better fit by considering appropriate interactions.

References

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H. & White, J.-S. S. (2009), 'Generalized linear mixed models: a practical guide for ecology and evolution.', *Trends in Ecology & Evolution* **24**(3), 127–135.

Enders, C. K. (2010), *Applied missing data analysis*, Methodology in the social sciences, The Guilford Press, New York.

Faraway, J. J. (2016), *Extending the Linear Model with R*, Chapman Hall/CRC Texts in Statistical Science, 2 edn, CRC Press.

Jenni, L. & Kry, M. (2003), 'Timing of autumn bird migration under climate change: advances in longdistance migrants, delays in shortdistance migrants', **270**(1523), 1467–1471.

Meer, J. V. D. & Camphuysen, C. J. (1996), 'Effect of observer differences on abundance estimates of seabirds from ship-based strip transect surveys', *Ibis* **138**(3), 433–437.

Newton, I. (2007), 'Weather-related mass-mortality events in migrants', *Ibis* **149**(3), 453–467.

Newton, I. & Brockie, K. (2008), *The migration ecology of birds*, first edition edn, ScienceDirect, Amsterdam, Netherlands].

Spaar, R. & Bruderer, B. (1996), 'Soaring migration of steppe eagles aquila nipalensis in southern israel: Flight behaviour under various wind and thermal conditions', *Journal of Avian Biology* **27**(4), 289–301.

Wikipedia contributors (2020), 'Bird migration — Wikipedia, the free encyclopedia'. [Online; accessed 3-July-2020].

Appendix

Detailed R code can be seen from the following link to a Github repository:`https://github.com/xu-echo/migration_raptor_pro1.git`