**Gapeworm Dynamics in First-year European Starlings: Statistical Analysis in R**

**Abstract**

Host behavior and interaction in the community level can have a significant effect on parasites with complex life cycles. As an invasive species, European starling serves as a good reservoir for several parasite species. Gapeworm is a parasitic nematode, which can infect birds and poultry animals by an intermediate host and cause economic loss to small open backyard poultry producers. Here, I investigated the seasonal gapeworm dynamics in the first-year European starlings and how did rainfall level difference, bursa and spleen weight change the gapeworm dynamics in the starlings. Statistical analysis had been used to analyze the data in R and I compared different models to better describe the gapeworm data. A better understanding of gapeworm dynamics in each year could help poultry producers or managers to have an insight into when and where to apply the gapeworm treatment.

**Introduction**

After a historical ignorance of parasites in community ecology, epidemiologists and ecologists have melded these two fields and addressed the importance of parasites in different host interactions (Hatcher and Dunn, 2011). A large body of literature has been focusing on the changes in infected host behavior and how could these changes contribute to parasite infections (Dobson, 1988). But in parasite ecology, how could host behavior and interaction at community level influence the parasite infection is still not well understood. In most cases, host behavior is difficult to measure and is assumed to be homogenous in models of disease spread (Grear *et al*., 2013). The lack of study of host behavior and interaction on complex life cycle parasite dynamics might be a missing step in parasite regulation and causes epizootic problems. In 1998, Hudson *et al.* addressed how parasite removal can prevent the red grouse population collapse. It is important to understand how to regulate parasite population and to prevent unexpected results. For poultry producers, a better understanding of gapeworm dynamics is also important for preventing bad helminthic disease in chickens and turkeys (Akand *et al.*, 2020). However, for parasites with complex life cycles, the regulation and treatment might be more complicated due to the presence of the intermediate hosts. The population dynamics of parasites in avian hosts have been the least investigated, primarily because of the difficulties in getting collection permits. The current study has been made possible because invasive species like European Starling (*Sturnus vulgaris*) are routinely deprecated at local airports (Newark Liberty International Airport, John F. Kennedy International Airport, LaGuardia Airport, and Teterboro Airport) in an effort to reduce bird-aircraft collisions. The carcasses are available to museums and universities, which offered the opportunity to gather a wealth of parasite data. Also, all the starlings we got from airports are yearlings, which means we could focus on the parasite infection in one year.

European starlings have been an invasive species in North America since the 1890s. They have a broad diet range and prefer invertebrates and insect larvae during nesting season (Linz *et al*., 2018). Gapeworm (*Syngamus trachea*) is a parasitic nematode with a complex life cycle that infects the birds’ trachea. Starlings get infected from ingesting the encysted parasite larvae in earthworms (Clapham, 1934). As an invasive species, starlings can be a good reservoir of the parasites and transmit gapeworms to local birds or poultry animals (Valente *et al*., 2014). Our preliminary results suggest that the mean gapeworm abundance in 2019 was higher than in 2018 (Fig. 1). The higher monthly rainfall levels in 2019, compared to 2018, could be possible indirect causation of the high parasite infection in 2019 (Fig. 2). Starlings often become infected by eating infected earthworms (*Lumbricus terrestris*). Rain affects earthworm behavior by bringing them to the surface where they are more easily predated by the birds (Onrust *et al*., 2019), and this might influence the encounter rate between gapeworm larvae and the starlings. Thus, rain seems to be an important abiotic parameter in gapeworm biology.

In this paper, I will focus on the statistical methods I used to analyze the gapeworm dynamics. A generalized additive model (GAM) had been used to explore the correlation between seasonal dynamics and gapeworm infection. I also used a generalized linear mixed-effects model (GLMM) to investigate the relationship between host immune system response and the gapeworm population.

**Material and Methods**

Data Collection:

The European Starling carcasses were transported from four airports from May to September. All the starlings will be kept in the freezer under -20℃ for further necropsy. Each frozen bird will be weighted (top-loading scale), and then be defrosted under running hot water for 20 minutes. During necropsy, the thoracic cavity was opened by cutting through the sternum, up to the throat. To recover the digestive system, the belly skin was carefully cut down to the vent. When the bird was opened, the spleen and bursa were removed and weighed. Then beginning with a cut at the trachea and the esophagus, and another at the end of the intestine, the entire intestinal tract was extracted. Each part of the intestinal tract was placed in a 100mm x 15 mm petri dish with about 5 ml of tap water and the section was slit along its length with a small nose hair scissors. These scissors have rounded tips to prevent the cutting of parasites. All the macroparasites find in the intestinal tract will be recorded. *Syngamus trachea* would be found, specifically in the trachea.

Rainfall level data was gathered from four airports’ weather station. For the model analysis, I only considered the previous month rainfall for each collected bird. Here, I assume one month before the collecting date is the time interval that those starlings could get the infection.

Statistical Analysis:

Part 1: Seasonal Variation and Rainfall effects

Based on the preliminary result in Figure 1, the gapeworm infection and time did not follow the linear relationship, so more functions might be needed to describe the model. I used the generalized additive model to explore how seasonal variation and rainfall affect the gapeworm infection in first-year starlings. I used 2019 data to generate the model and 2018 data to test the model prediction. The overview of gapeworm count data showed that the variance of the data is much larger than the mean in both years, which indicated the overdispersion problem. Usually, the Poisson distribution will be used to fit the model with count data. But the Poisson distribution assumes that the data variance equals the mean. In this case, I need to use a different distribution for my model fitting.

Historically, parasite burden in the host population tends to be aggregated and has been described as a negative binomial distribution. A small portion of hosts will have a high parasite burden and most hosts will have no or few parasite infections (Wilson et al., 1996). This suggests the negative binomial distribution can be a good candidate for my data analysis. Different from Poisson distribution, the negative binomial treats the variance as a quadratic function of the mean. However, zero-inflated Poisson distribution here might be another good candidate. The zero-inflated Poisson distribution suggests the zero counts in the data are generated by another process. For parasite infection, we could expect the host might not encounter the parasite at all or their immune system win the fight, and this could be the main reason for zero counts. To compare these two different distributions, I used AIC values as the criteria. The lower AIC value gives me a better model to describe the relationship.

I also compared two models with different ways of describing rainfall terms. One model described the rainfall by simply adding the term, and another model multiplies the rainfall term to the smooth function. To select the best model I want, I used 2018 data as the input value to predict the model results and compare it to the real 2018 data to see which one give the better prediction.

Part 2: Bursa and Spleen weight

Bursa and spleen are two important organs that responsible for the bird immune system. Here, I fitted the generalized linear mixed-effects models with negative binomial distribution for both bursa and spleen weight and set date difference and rainfall level as the random effects. To avoid the birds weight effect on bursa and spleen size, I used bursa (spleen) and bird weight ratio to run the model.

**Results**

For the overdispersion test, I fitted a generalized linear model with gapeworm infection as the response variable, and date difference with rainfall as explanatory variables. To check if the overdispersion happened in my data, I looked at the residual deviance and the degree of freedom. The residual deviance for this model is 921.33, which is much larger than the degree of freedom (400). This indicated the overdispersion problem in my dataset.

I ran four models with different distribution families and ways of interpreting rainfall. After fitting the 2019 data into 4 different model combinations, the results of AIC values are shown in Table 1. The negative binomial distribution model showed a lower AIC value (1381.10 and 1381.76), which means the model with this distribution can be a better fit for my data. However, models with two different ways of interpreting rainfall levels did not show a big difference in AIC values. To decide which condition is more suitable for my data, I used 2018 data to predict the gapeworm infection and compare the predicted values to real data. Figure 4 shows that the model has rainfall level as a multiply term might be better in this case.

**Table 1.** Model Comparison for Different Distribution Family and Ways of Interpreting Rainfall Term.

|  |  |  |
| --- | --- | --- |
|  | Df | AIC |
| Negative Binomial (add) | 7.768459 | 1381.100 |
| Negative Binomial (multiply) | 8.014298 | 1381.766 |
| Zero-inflated Poisson (add) | 9.575260 | 1419.476 |
| Zero-inflated Poisson (multiply) | 9.626372 | 1420.259 |

For the bursa and spleen data, I ran two linear models to check if the birds weight affects the bursa or spleen size. Two models were all significant (p-value: bursa: < 2.2e-16; spleen: < 2.2e-16). This indicated the possible effects of bird weight. To get rid of the effects, I used bursa (spleen) and weight ratio to run the gapeworm infection model. For the two GLMM models, the bursa model showed a significant negative correlation between bursa weight and gapeworm infection (Intercept: -12.03088; p-value: 0.0175). But no correlation between the gapeworm infection and spleen weight was detected (p-value: 0.134). The model prediction was shown in Figure 5.

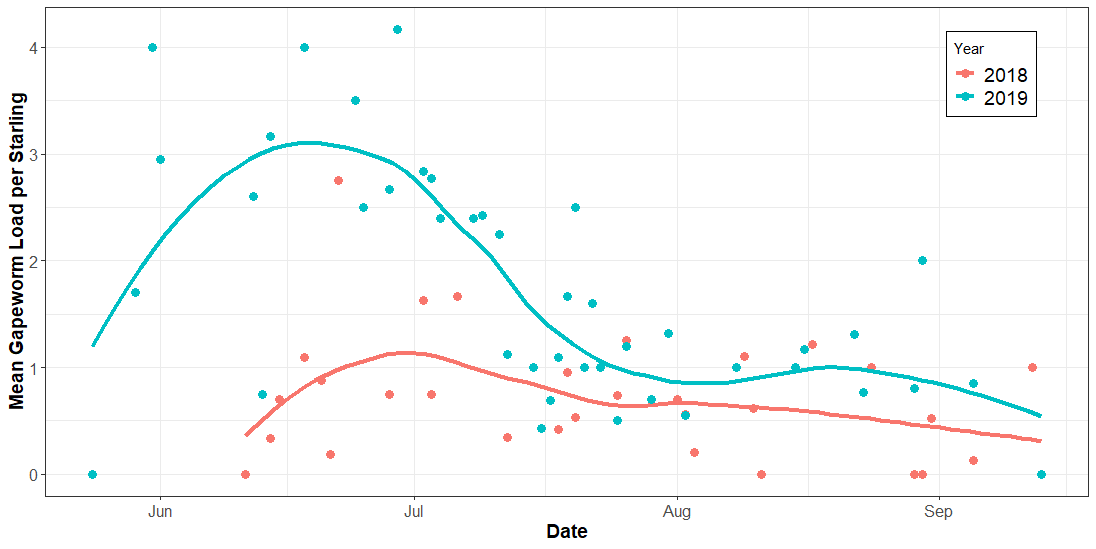
**Discussion**

The model prediction for seasonal variation and rainfall level gave us an insight into how gapeworm infection changed during the year. The high peak of infection in June and July may be corresponding to the parental feeding in April and May. Also, the rainfall level in April and May caused a big difference in model prediction. This can be a reminder that the precipitation in nesting season can be a predictor for the intensity of gapeworm infection. The only issue here is the birds age was not clear for each starling, which would make the rainfall effect in the model less accurate. In this study, I assumed the rainfall level one month before the birds collecting date determined the gapeworm infection. If the birds age can be estimated in some ways, the time interval for rainfall level can be more accurate. The model for bursa weight suggested that the starling with larger bursa tend to have fewer gapeworms. This could be explained by the size of bursa reflects how strong the immune response in that bird. In Møller and Erritzøe (1996) research, they found the bird with higher chance encounter parasites tend to have a larger bursa for immune defense.

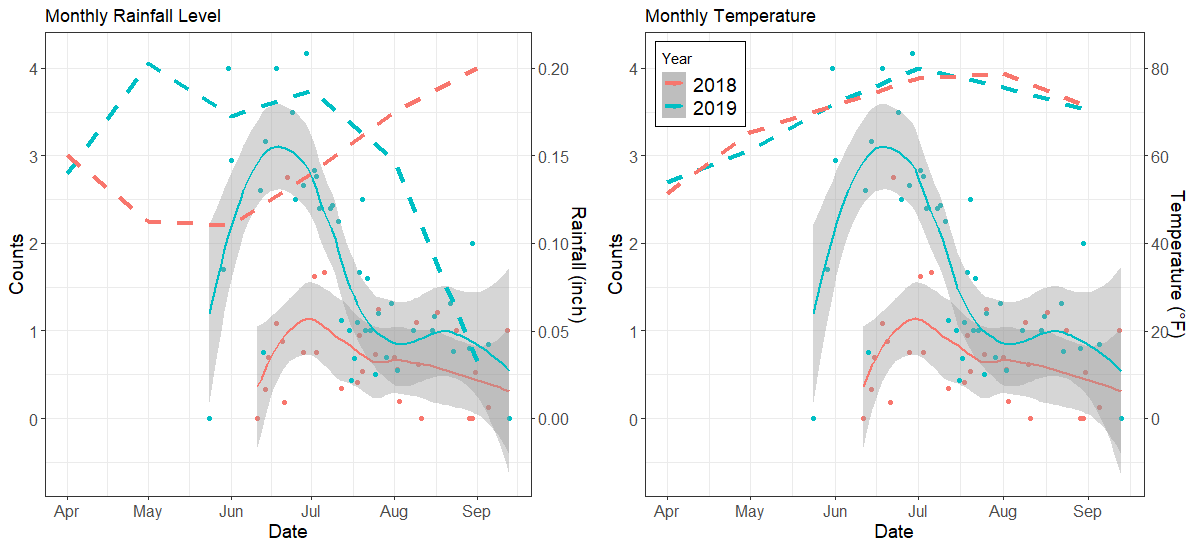
Negative binomial and zero-inflated Poisson distribution might both work well for parasite data modeling. In my case, the gapeworm infection seems to be common in European starlings. This might give the negative binomial distribution a better place to describe the data. The encounter rate for gapeworm larvae and starlings can be too high that most starlings might have s chance to eat gapeworm larvae. Zero-inflated Poisson distribution might be better to use if the parasite prevalence is relatively low that the hosts have a lower chance to encounter the parasites. In this case, zero values are generated by a lower encounter rate.

**Appendix**

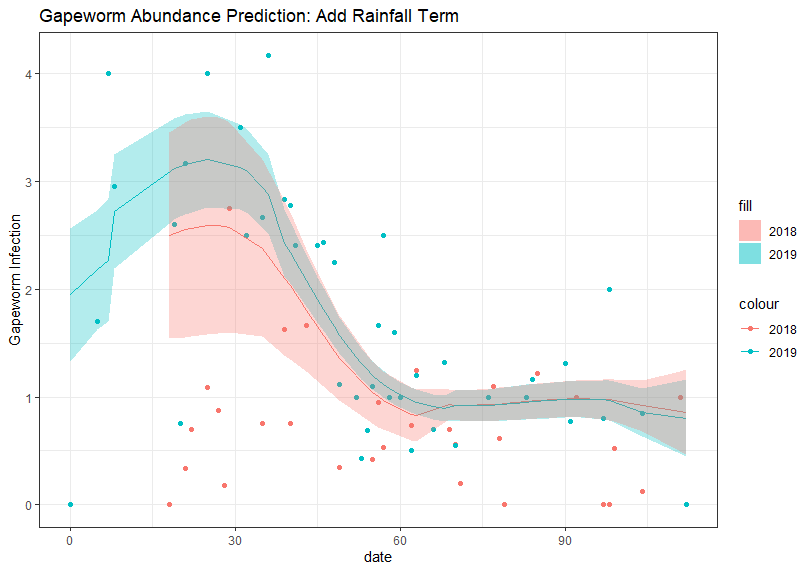
**Figure 1.** Mean Gapeworm Load (Abundance) for Starlings at Different Sampling Dates in 2018 and 2019. Parasite abundance was much higher in 2019 than 2018. The trendlines are generated by R using loess regression. In 2018, the mean intensity of gapeworms was 1.816, and the mean prevalence was 0.427. In 2019, the mean intensity of gapeworm was 2.861, and the mean prevalence was 0.591.



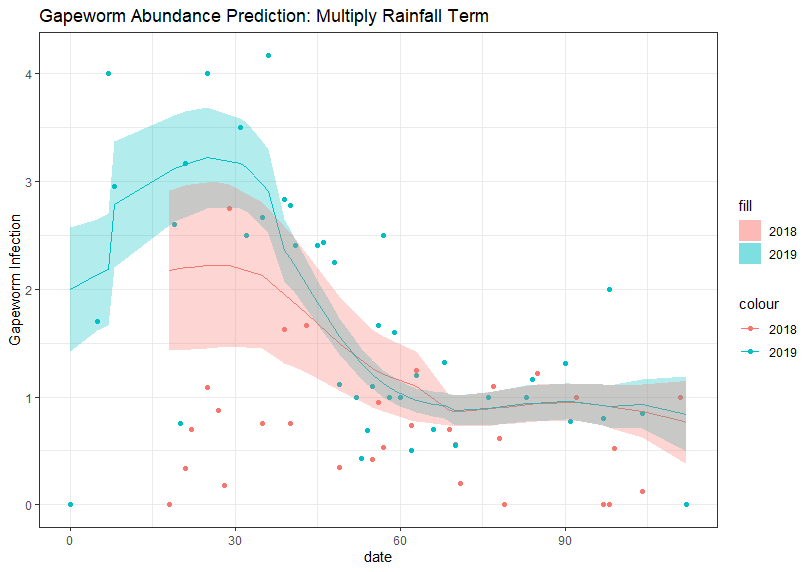
**Figure 2.** Direct Comparison between Abiotic Factors and Gapeworm Dynamics in 2018 and 2019. Mean monthly rainfall level and temperature for 4 airports are shown in dashed lines. The solid trendlines are gapeworm dynamics generated by R using loess regression with grey range of standard error. Rainfall and temperature data are derived from airport weather stations. These data support the hypothesis that rainfall is more important than temperature in gapeworm population dynamics.



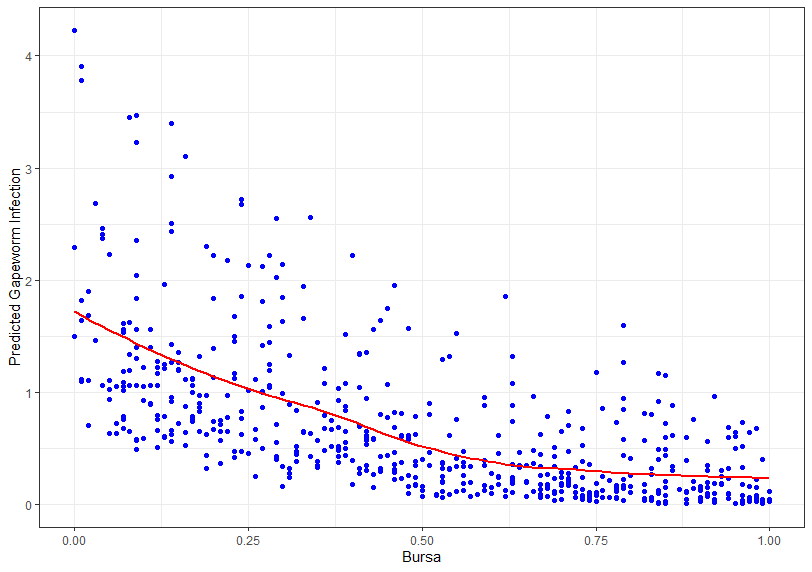
**Figure 3.** Generalized Additive Model Prediction with Adding Rainfall Level Term. The model with negative binomial distribution was generated using 2019 data and predicted by 2018 data. The curve for 2018 prediction does not show a good fit to describe the infection. Shaded area is the standard error for the model prediction.



**Figure 4.** Generalized Additive Model Prediction with Multiply Rainfall Level Term. The model with negative binomial distribution was generated using 2019 data and predicted by 2018 data. The curve for 2018 prediction shows a better fit compare to the model simply add the rainfall term. Shaded area is the standard error for the model prediction.



**Figure 5.** Prediction from Generalized Linear Mixed-Effects Model for Starling Bursa Weight. The model was fitted by using combined 2019 and 2018 data. The model outcome showed the significant negative correlation between bursa weight and gapeworm infection. Bursa weight data was generated by R randomly and used to predict the infection. The trendline was by R using loess regression.



**Literature Cited**

1. Akand, A.H., Bulbul, K.H., Hasin, D., Parbin S., Hussain, J., and Sheikh, I.U. (2020). Perspective of gapeworm infection in birds. *International Journal of Veterinary Sciences and Animal Husbandry,* 5(3), 68-71.
2. Clapham, P.A. (1934). Experimental studies on the transmission of gapeworm (Syngamustrachea) by earthworms. *Proc. R. Soc. Lond.* B.115, 18–29.
3. Dobson, A.P. (1988). The Population Biology of Parasite-Induced Changes in Host Behavior. *The Quarterly Review of Biology,* 63(2), 139-165.
4. Grear, D.A., Luong, L.T. and Hudson, P.J. (2013). Network transmission inference: Host behavior and parasite life cycle make social networks meaningful in disease ecology. *Ecological Applications*, 23, 1906-1914. https://doi.org/10.1890/13-0907.1
5. Hatcher, M. J., & Dunn, A. M. (2011). Parasites in ecological communities: From interactions to ecosystems. *ProQuest Ebook Central.* <https://ebookcentral-proquest-com.proxy.libraries.rutgers.edu>.
6. Hudson, P. J., A. P. Dobson, and D. Newborn. 1998. Prevention of Population Cycles by Parasite Removal. *Science*,282:2256-2258.
7. Møller, A.P. and Erritzøe, J. (1996) Parasite Virulence and Host Immune Defense: Host Immune Response is Related to Nest Reuse in Birds*. Evolution*, 50(5): 2066-2072.
8. Linz, G.M., Johnson, R.J., and Thiele, J. (2018). European Starlings. USDA National Wildlife Research Center –Staff Publications. 2027.
9. Onrust, J., Wymenga, E., Piersma, T. and Olff, H. (2019). Earthworm activity and availability for meadow birds is restricted in intensively managed grasslands. *Journal of Applied Ecology*, 56, 1333-1342.
10. Valente, R., Ibañez, L.M., Lorenti, E., Fiorini, V.D., Montalti, D., Diaz, J.I. (2014). Helminth parasites of the European starling (Sturnus vulgaris) (Aves, Sturnidae), an invasive bird in Argentina. *Parasitology Research,* 113(7), 2719–2724.
11. Wilson, K., Grenfell, B., & Shaw, D. (1996). Analysis of Aggregated Parasite Distributions: A Comparison of Methods. Functional Ecology, 10(5), 592-601.