Modeling Stopover Sites of Migratory Birds' Routes for Conservation of Population and Prevention of Disease

Ulya Bayram[1,*], Ruichen Sun[2], and William Lee[3]

[1]PhD student, Computer Science at University of Cincinnati, Cincinnati-Ohio, 45219, USA
[2]PhD student, University of California, San Diego, 92093, USA
[3]MITRE, Bedford, 01824, USA
[*]bayramua@mail.uc.edu

ABSTRACT

Modeling migratory paths of birds is an emerging area of complexity science in which attracts interdisciplinary collaborations from ornithologists, computer scientists, epidemiologist, and policy makers. However, due to the unpredictability of climate and habitat changes, our understanding of bird migratory paths is still limited. Not accounting for environmental changes when modeling bird migration will mislead our conservation efforts. Hence, in this paper, we are using validated training data to predict future stopover sites of migratory bird species. We used white-fronted geese migratory paths as training data and modeled stopover sites with Markov Chains to predict future changes on the stopover sites.This prediction will allow researchers to realize how to better conserve the habitat locations at and around stopover sites in the light of current climate changes. Since birds often stop and interact with the nearby environment, our work will also allow researchers to predict potential dangers emerged from long-range migrations, such as new avian viruses along these routes.

Introduction

Migration is arguably the most fascinating behavior of birds. Aristotle thought birds hibernated between Fall and Spring since after Fall they were nowhere to be seen and only reappeared after Spring[1]. This myth has survived surprisingly for a long time since passive observational records were the only available materials back then. However, the Aristotle myth became obsolete when Frederick II, the Holy Roman Emperor, documented bird migration for the first time after the beginning of the 20^{th} century by tagging birds with specially made rings that became standard practice in birds tracking[2]. Over the past century, technological advancement has enabled researchers to study bird migration with much greater spatial and temporal details. Particularly, satellite tracking provides the most detailed picture of how birds migrate as this method can record where a tagged bird has been over a long period of time.

Tracking bird populations in space and time is a central challenge for many scientific and environmental aspects, including conservation science, particularly in the light of the potential changes on our planet's climate[3]. There are two main reasons for tracking and studying bird migration. One reason is to understand and identify any expected and/or unexpected changes in the number and timing of birds using the affected areas/habitat due to man-made modifications/changes. Another reason is to see what might happen if all habitats are simultaneously affected by global climate change[4] since such drastic change could eventually leads to extinction. In fact, Dodo bird is the first bird specie whose extinction was conceded, and admitted in writing, to have been caused by humans[5]. A study[6] states that we are now on a trajectory to lose between 1/3 to 2/3 of all currently living species due to human alteration of the environment. Another study[7] outlines that today 1 in 8 known bird species are threatened by global extinction. So clearly the threat of extinction is real, and there is a need for urgent intervention. In his book[8], Van Dooren states that a more interdisciplinary approach to research birds' behavior can shed light on the nature of the birds and their interactions with the environment. Yet, existing findings seem limited due to the potential limitations introduced by traditional modeling approaches and disciplinary boundaries.

In addition to extinction factor, studies on bird migration are important for disease prevention as well. Migratory birds are potential agents that could spread West Nile virus to distant places from where it orignated[9]. A study[10] identifies the relationship between timing of outbreaks and migration of birds from a correlation analysis of the spread of H7N9 bird flu and the migratory flyways of different bird species. Another study states that there is evidence suggesting the H5N1 outbreaks are possibly spread and exacerbated as the birds stopover to rest and spread the viruses to nearby species and environment[11].

It's a common practice in migration studies to focus on optimal migratory path modeling. For example Vrugt[12] et al. created a model to simulate the spatial location, muscle, and fat amounts of an individual bird using a given set of behavioral rules to detect the optimal path of their migration. However, navigational skills of the birds during migration do not always seek the shortest and the most optimal path. Instead, there are plenty of factors involved in the migration process including both environmental and metabolism[13]. As a matter of fact, in real-world environments, birds often don't match any single predicted optimum. For example birds unconsciously decide on migration timing, amount of fuel to store, choice of stopover sites, and breeding locations based on on environmental and their physical abilities[4]. Considering this, it is clear that attempting to embed environmental information such as habitat locations, bird's physical information, and weather/wind data into migration modeling problems are not likely to bring any better predictive abilities to the task. An example of this situation is explained by[13]. The authors stated that birds cannot always foresee the future situations. And, for that reason, they attempt to carry extra fuel for potential unexpected situations. This problem is also clear when we consider the chaotic behavior of the weather. Although there are some areas with predictable weather conditions[14], in general it is chaotic which makes it highly dependent on the initial conditions. Hence, making it likely to be impossible to predict more than 10 days ahead of the weather[15,16]. By embedding weather, environment and other factors to predict migratory behavior is inappropriate since weather and environment are also unpredictable and unstable factors.

Considering all these stated factors that are effective on migratory behavior, timing, route selection, survival etc., we chose to focus on modeling and prediction of stopover sites on migratory routes. Main reasons behind our selection/decision are as the following:

- Birds decide to stop at some sites during their migratory flights for purposes of feeding, drinking, resting etc. so they can continue to their migration. Due to factors explained before, this decision is dependent on the weather conditions which is not possible to predict a year ahead. Another factor is habitat at the location so bird can decide to stop there which is also depending on the weather, climate, and human factors.

- Modeling solely by the stopover sites used by the same specie in previous years could provide generalizable information as long as we keep the characteristics of the habitat and the birds' preferred feeding locations for feeding constant. So modeling only using the stopover locations allows anybody to apply our model/method to their specie of interest so long as they have latitude and longitude of stopover locations during the migration of the specie for couple of years for training the model.

- Modeling the stopover locations and predicting the likely stopover locations for the following years will allow us to conserve the habitat on that predicted locations, to keep the bird specie from extinction, as well as to plan ahead and control potential disease spreads (e.g., bird-flu like virus). Knowing where the migratory birds that are likely to carry that virus will minimize the interaction with the environment at that location and will allow people to take precautions - hence very likely saving lives.

In the following sections, we explain our method on modeling and predicting the stopover sites, results of experiments we conducted on a sample datasets, implications of the results, and potential future directions of our research project.

Methods

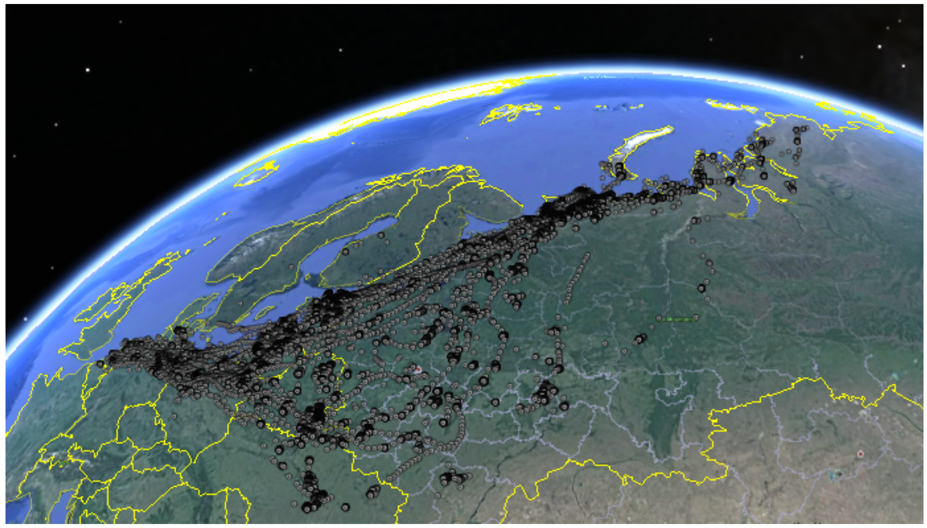
0.1 Background

Numerous studies exist in modeling birds' migratory paths[17] including reconstructing of migration velocities[18] and prediction of migration timings by utilizing acoustic feature extraction techniques[3]. However, most of the results of these studies are limited. For example, some studies embed the temperature and wind data as features in their methods[18]; however, due to chaotic behavior of the weather,[15] embedding temperature data and making decisions using the past information may not return the most reliable results. Generalization is also another issue, some studies do not model migration as a round trip (i.e., only by autumn or only by spring)[12], but not using the same method they propose on both spring and fall migrations. And a model proposed on a specific bird breed may not fit another breed, so modeling by specific morphological features would not work well for other species. Hence, this is not a generalizable way of modeling either.

0.2 Data and Initial Exploratory Analysis

In the task of bird migratory path modeling, there are plenty of available datasets collected by bird watchers[19] as well as data collected by biologists[20]. After searching for suitable datasets available publicly in movebank[20], we decided to work with white-fronted geese data tracked by GPS. Tracking began in 2006 with over 65 geese tagged with GPS device, the dataset contains 6 years of migratory route information collected from satellite. White-fronted geese, known in latin as Anser albifrons, is shown in Figure 1a. We transformed these data into KML format and mapped them onto Google Earth for exploratory analysis (Figure 1b). The Google Earth visualization shows both spring and fall migratory routes taken by 65 birds in 6 years. As seen in the figure, these birds migrate between western Europe and the Russian Arctic during spring and fall.

(a) Anser albifrons[21],[22]



(b) Migration route

Figure 1. On the left, we see the white-fronted geese, and on the right side we see the path they use during Fall and Spring in 6 years of data during migration.

When we analyzed the dataset, we found that there are ten different actions performed by white-fronted geese on their migration routes in 6 years: breeding, flying, stopping over, moulting and breeding, prospecting, getting shot, getting lost, spending the winter, and moulting and prospecting. The frequency of these actions in the dataset are shown in the histogram in Figure 2.

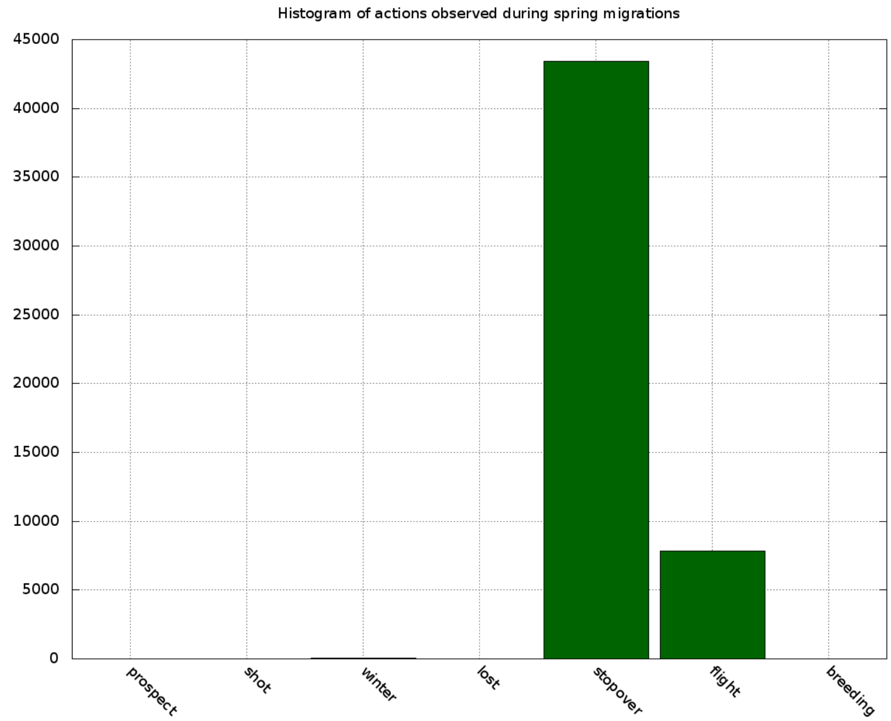
In Figure 2, we observed that the most frequent action during the Fall and Spring migrations is stopping. The next frequent action is flying in both seasonal migrations. These observations are in accordance with the literature findings[13] since during migration, birds need to stop frequently to rest, and refuel so they can prepare for potential unexpected surprises (e.g., unable to locate usable habitats or severe weather). This further indicates our appropriate choice of selecting stopover sites as the main focus in our migratory modeling task. An interesting discovery made by the researchers who collected this dataset[23] is in contradiction with previous theories on migratory birds' behavior. The researchers in fact found that there is no evidence to suggest any competition within specie to reach the breeding grounds as quickly as possible to catch the best area to settle, in terms of proximity to food and water resources and security from predators[4]. They also found that birds are using different strategies in the two seasons. In Spring, they spread out wider to acquire more energy in many stopover sites to store more fuel for breeding. In Autumn, they store supplies and food around their breeding grounds in order to allow them to return to their Winter grounds more quickly.

In Figure 4, we learned how stopover sites are connected to each other, as well as where birds start and stop their migrations. White-fronted geese prefer spending their winters in southern regions of Europe as visible in Figure 4, and they fly to north-east to spend their spring and breed, raise their chicks, and at the end of the Spring, they return to their wintering grounds during Fall.

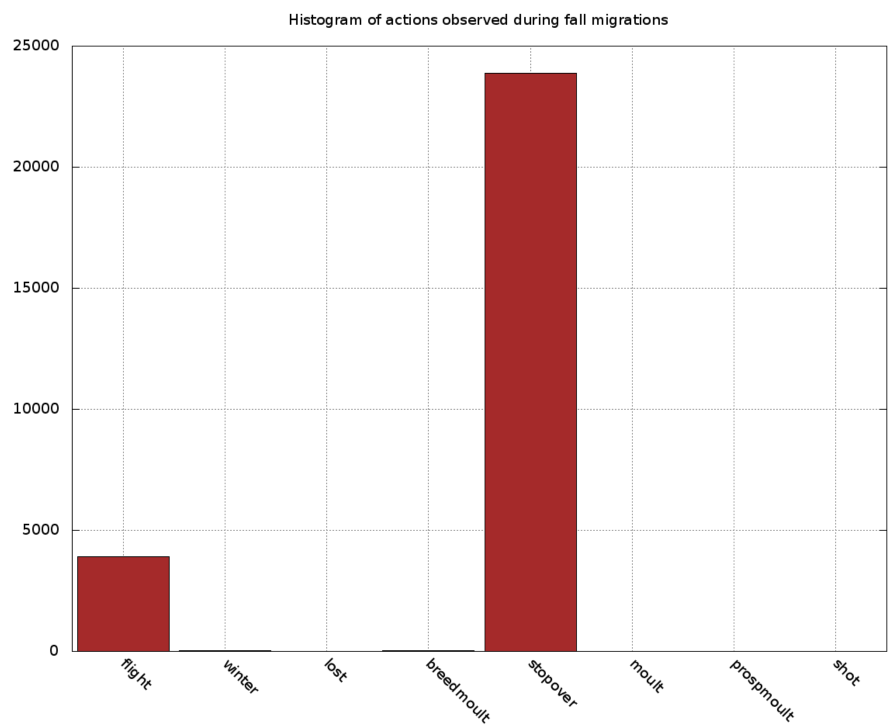
0.3 Algorithm

To start the modeling of the migratory route stopover locations, we started with dividing the latitude and longitude locations of the world map into grids. This grid mapping is a standard procedure in many migratory modeling studies for the purpose of reducing prediction sensitivity[24]. Kranstauber et al. used hexagonal grids with 8^{th} resolution and aperture 3, to utilize the advantage of equally spaced points and uniformly distributed grid cells. Another study discretized the map into 2D rectangular equidistant grid cells while choosing 0.5^0 as the resolution parameter[12]. A bird's migratory path modeling study that focused on predicting disease spreading task used 0.1^0 as the grid resolution in horizontal and vertical directions[9]. In this study, we selected 0.5^0 as the cell dimensions of the grids since as we wanted to have enough sensitivity to capture the stopping behavior of birds during migration. We also did not want to have too many cells that would be redundant in modeling and won't return any important behavioral information. Therefore, we chose the same grid mapping method as in[12]. Grid-mapped locations can be seen in Figure 3.

Next, we divided the dataset into two temporal segments. For example, we used 3 years of the dataset for training the Spring and Fall migrations. Then, the rest of the other 3 years was for predictive purposes. Training data for both Spring and



(a)



(b)

Figure 2. Actions observed during 6 years of (a) Spring and (b) Fall migration dataset of white-fronted geese.

Fall migrations are seen in Figure 4. Migration in Spring typically starts from the wintering grounds in southern regions, and during this migration, they fly towards the potential breeding grounds that are habitable thanks to the Spring warm season.

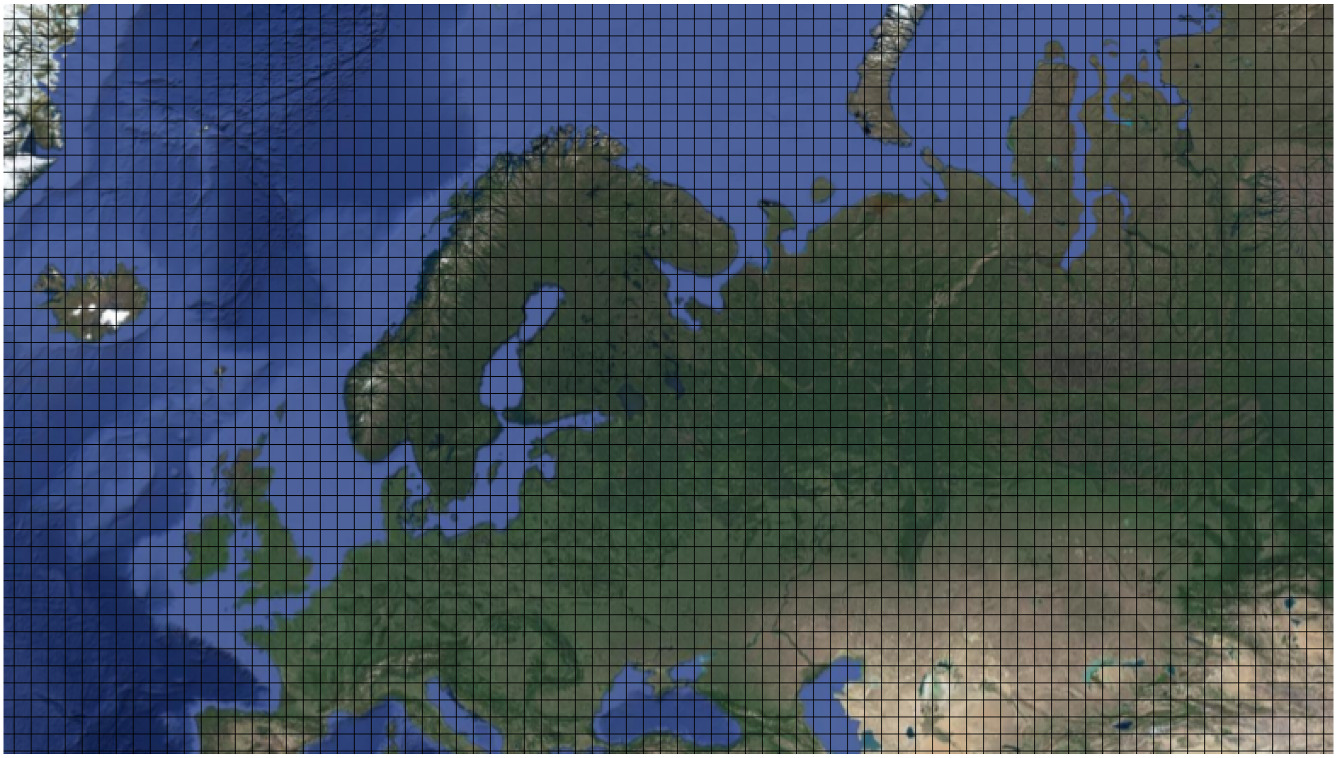


Figure 3. Map divided into 0.5^0 grids where each cell is to be treated as a pixel.

During Fall, migration starts from the breeding grounds and ends with the settling at wintering grounds where geese settle to spend their cold winter months in southern regions because the weather is not as tough as the north arctic regions. In Figure 4 we can see the frequency of usage of stopover sites by the birds through the heat-map points. Frequency numbers corresponding to the colors are seen on the legends on the right of the plots.

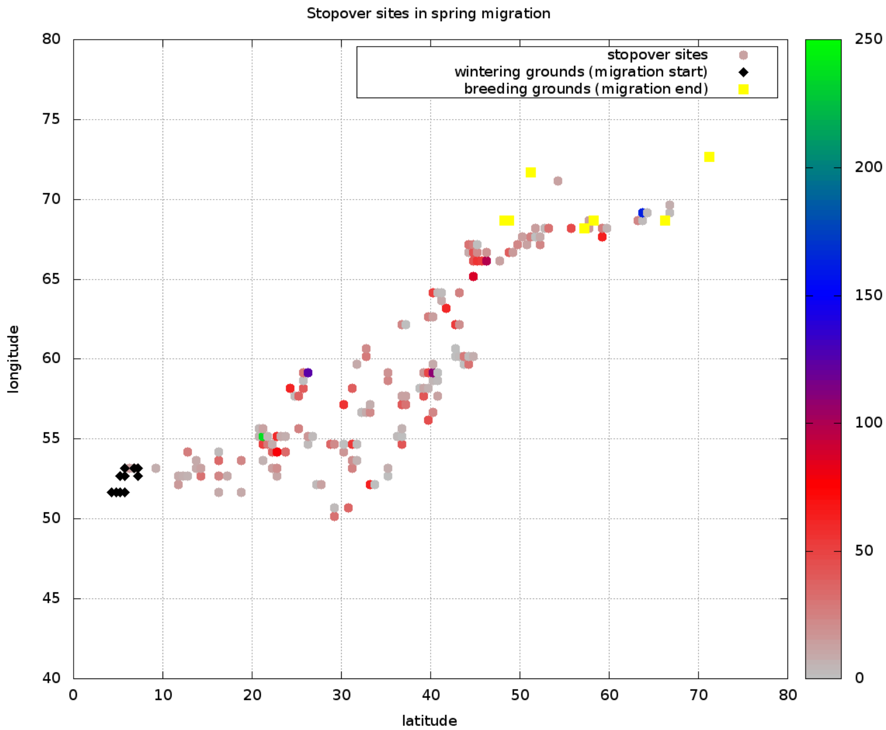
The locations in Figure 4 show where the stopover locations are in the box of interest where white-fronted geese move and live within their lifetimes. These locations span a great portion of Europe and part of Russia. And since we already divided the map into grids as described earlier, we decided to treat the grids and the regions of stopover as pixels of an image. So we converted the grids with number of stops in stopover sites for Fall and Spring migrations separately. Frequency of stopping at those grid cells for each bird shows the importance of that habitat regions. In our project, since our goals were to secure the habitats for specie conservation as well as to restrict the search locations for predicting and preventing disease spreading, we applied convolution to the image with a kernel size of [5x5] using 1 for all elements (i.e., our dilation kernel). This would allow us to capture proper locations and also to have buffer regions in case the grids are not able to cover enough portion of the habitat. We adopted the convolution method described in[25]. Dilated habitat regions that are frequently used stopover sites during Spring and Fall are shown in Figure 5.

In addition to the dilation approach, we also decided to investigate the stopover routes used during Spring and Fall migrations. The most appropriate method for this purpose would be Markov Chains. Every step taken by the birds, which would be flying from stopover site to another stopover site until the migration ends, is treated as a time step. The dataset provides us the snapshot timing from the GPS data. We treated each time step as a unit of movement instead of embedding the hour, minute, second time difference on stopover site to stopover site jumps.

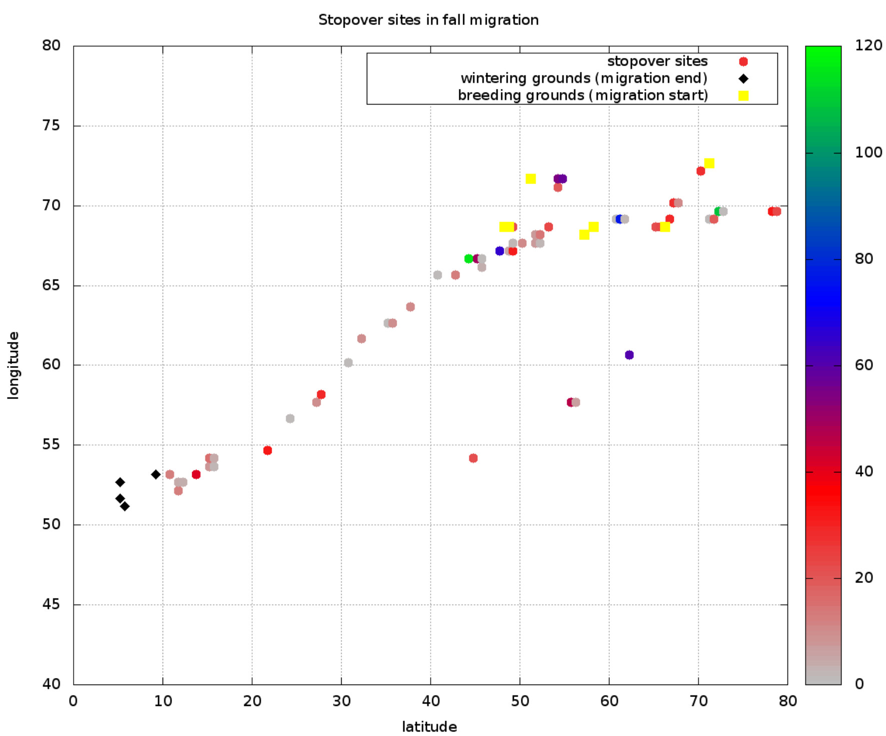
In our method, we modeled each stopover site, which is the grid cell of interest, during a migration season. Each stopover cell was modeled as a Markov Chain node in the network. Each jump of a bird from a stopover site node to another stopover site node would increase the transition probability. This modeling technique is to allow researchers to find the most likely paths connecting stopover sites and to identify the likelihood of a bird staying within the same stopover site during snapshot times/time steps. Results of our methods are illustrated in the following section.

Results

First thing we did was to perform dilation-based stop-over site prediction. Using the dilated stopover sites over grid-image generated from the training data, we performed predictions to find out if we can predict the stopover sites the geese might use



(a)



(b)

Figure 4. Wintering grounds, breeding grounds and stopover sites white-fronted geese frequently use during (a) Spring and (b) Fall migrations. Frequency of usage of stopover sites they used until 2009 are seen from the legend of colors in right side of the plots.

	Predicted	
Actual	TP=48	FN=109
	FP=201	TN = 10658

	Predicted	
Actual	TP=21	FN=41
	FP=78	TN = 10876

Table 1. Actual and predicted stopover locations on (left) Spring and (right) Fall migrations.

in test data years ranging from 2009 to the end of the data collection years. When we performed predictions, prediction results were observed in Table 1. Table on the left shows the prediction performance for the Spring routes while table on the right shows the Fall predictions on test data.

These results show the accuracy of true predictions of actual stopover sites (true positive) and actual sites that geese prefer not to stop at due to a lack of habitability and many other reasons (true negative). Compared to these true predictions, false negative and false positives are very low on spring and fall migrations. This alone shows that we can use dilated images where each pixel location indicates a stopover site has great predictive ability.

In addition to predicting the future stopover sites, we also used Markov Chains to model stopover site routes that birds prefer every year for Spring and Fall migrations separately. To discover this, we modeled grid cells and the preferences of birds jumping from one cell to another. We then converted that occurrences into transitional probabilities. Spring migration stopover site Markov Chain network is seen in Figure 6. Markov Chain network for Fall migration route is also seen in Figure 7. When we examined both networks, we saw that there were few independent routes that geese preferred during their migrations. This might be due to their strategies for finding the best habitats during their migration stopovers. If we consider the entire flock, it would be the best choice for birds to migrate to the same destination by stopping over at different stopover locations. This is likely due to the competitiveness theory proposed in[13]. Klaassen et al. found that birds compete with each other within a flock in order to find the best sites to settle, to breed, and to rest in which best location is the one with closest proximity to habitat, and most secure place against potential predators that are present at the location. Hence, our finding is in accordance with the theory.

When we reviewed the Markov Chain networks in Figures 6 and 7, we identified sub-networks also existed within the complete network. For Spring migrations, there seemed to be more connected sub-networks on stopover sites. However, for Fall migration, the birds seemed to deviate less from a directed path. Hence, the networks are not interconnected to each other as much as Spring.

Markov Chains also showed us self-loops. These self-loops told us how many times a bird preferred to stay at stopover locations. Self-loops may also be indicating to us how habitable the habitat in those stopover locations, so that birds chose to spend more than few once at that location, to rest, to refuel, and perhaps for other purposes.

Markov Chains can also be used by the researchers to perform appropriate predictions. For example, if a researcher wants to know which stopover route the bird will prefer during their migration in a given season, all they need to do is to construct a vector of zeros where the migration starting location is. A 1 within that vector would indicate the first stopover location the bird used. Multiplication of that vector with the transitional probabilities of the Markov Chains we constructed will provide the researcher a vector of probabilities. The result would show the most likely stopover sites and least likely stopover sites that are used by the birds during migration of a particular season.

Such predictive power is very important especially for the case of predicting the disease spreads for bird flu or similar disease. If researchers wants to know if a bird stops at a certain location during migration, researchers can predict where the bird is most likely to stop during migration to rest, so researchers can let the officials know which locations are likely to have the bird stopover, to take precautions, and to quarantine suspected birds[10].

Discussion

Our results showed that conversion of data into images and performing image processing on location data is able to provide suitable predictions.

Also, our results indicated that it is very useful to construct Markov Chain networks of stopover locations, which is novel, never be done before in the literature in migratory modeling studies. Researchers can perform predictions and control to find out how habitats are changing and how diseases are likely to spread so they can take precautions.

Acknowledgements (not compulsory)

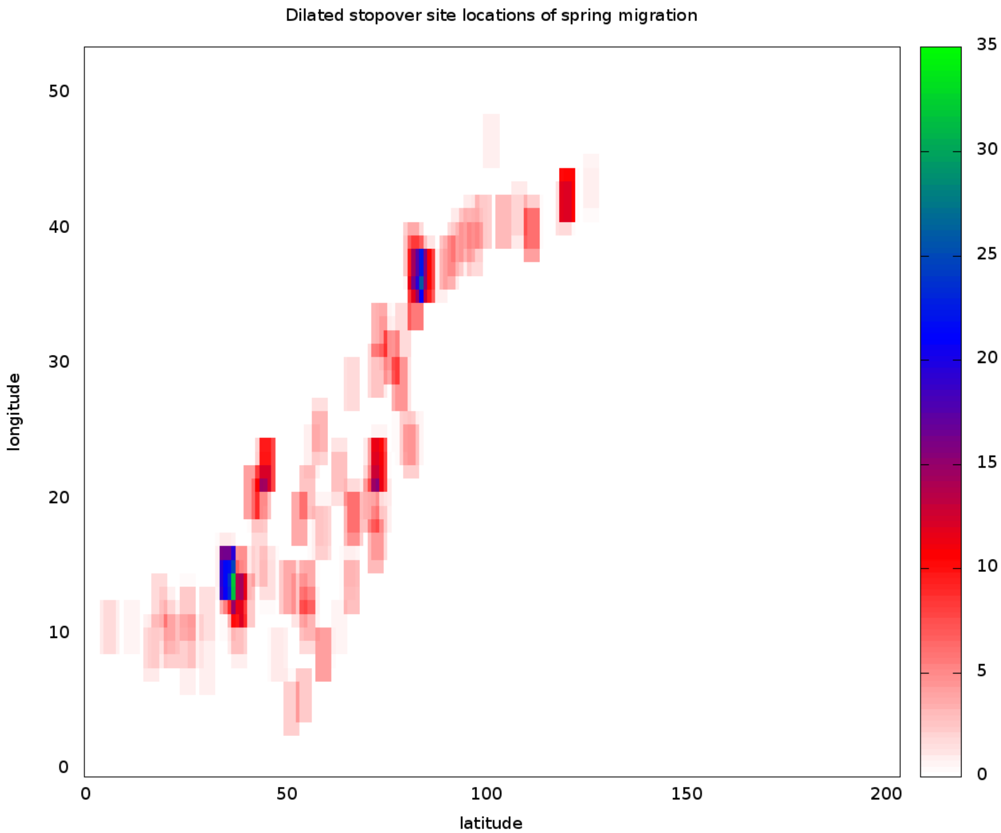
Thanks to Santa Fe Institute and Evelyn Strombom.

Author contributions statement

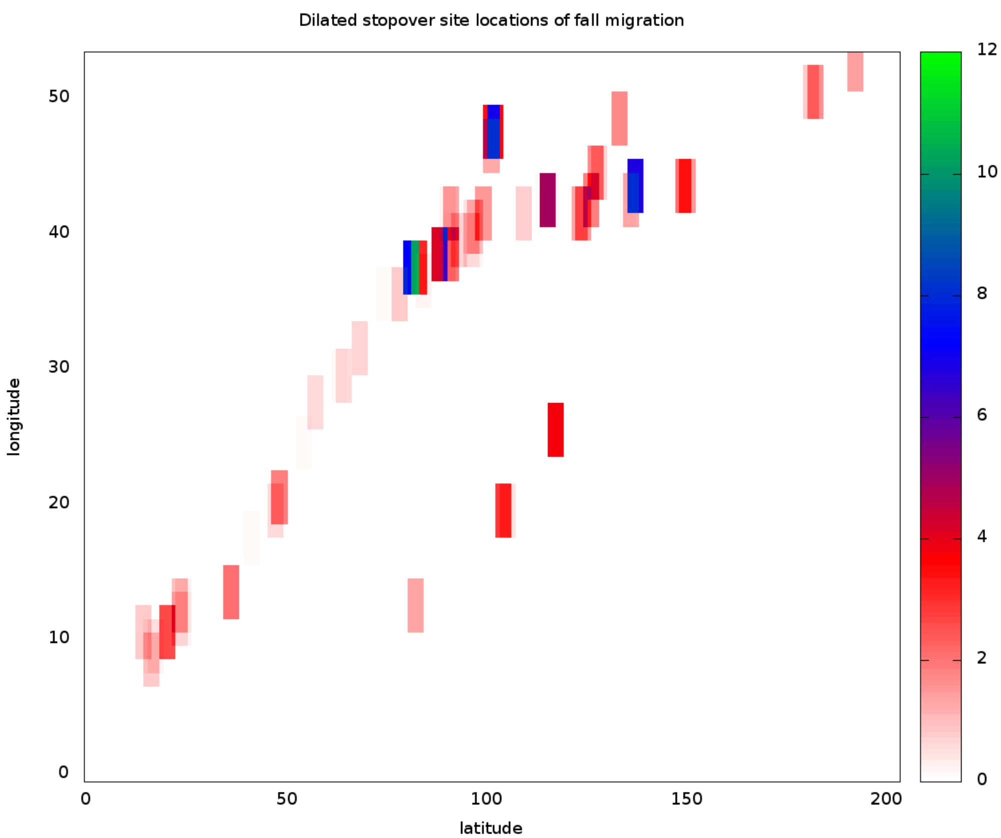
U.B. conceived and conducted the experiments, U.B. and W.L. built the algorithm, U.B. implemented the algorithm, R.S. analyzed the results, All authors reviewed the manuscript.

References

1. Berthold, P. *Bird migration: a general survey* (Oxford University Press on Demand, 2001).
2. Campbell, B. & Lack, E. *A dictionary of birds*, vol. 108 (A&C Black, 2011).
3. Damoulas, T., Henry, S., Farnsworth, A., Lanzone, M. & Gomes, C. Bayesian classification of flight calls with a novel dynamic time warping kernel. In *Machine Learning and Applications (ICMLA), 2010 Ninth International Conference on*, 424–429 (IEEE, 2010).
4. Ens, B. J., Piersma, T. & Tinbergen, J. M. *Towards predictive models of bird migration schedules: theoretical and empirical bottlenecks* (Nederlands Instituut voor Onderzoek der Zee, 1994).
5. David, Q. The song of the dodo: Island biogeography in an age of extinctions (1996).
6. Myers, N. & Knoll, A. H. The biotic crisis and the future of evolution. *Proceedings of the National Academy of Sciences* **98**, 5389–5392 (2001).
7. International, B. L. State of the world´ s birds: Indicators for our changing world. *Birdlife International.* (2008).
8. Van Dooren, T. *Flight ways: life and loss at the edge of extinction* (Columbia University Press, 2014).
9. Peterson, A. T., Vieglais, D. A. & Andreasen, J. K. Migratory birds modeled as critical transport agents for west nile virus in north america. *Vector-Borne and Zoonotic Diseases* **3**, 27–37 (2003).
10. Wiwanitkit, V. *et al.* Research priorities in modeling the transmission risks of h7n9 bird flu. *Infectious diseases of poverty* **2**, 1 (2013).
11. Liang, L. *et al.* Combining spatial-temporal and phylogenetic analysis approaches for improved understanding on global h5n1 transmission. *PLoS One* **5**, e13575 (2010).
12. Vrugt, J. A., Van Belle, J. & Bouten, W. Pareto front analysis of flight time and energy use in long-distance bird migration. *Journal of Avian Biology* **38**, 432–442 (2007).
13. Klaassen, M. Metabolic constraints on long-distance migration in birds. *Journal of Experimental Biology* **199**, 57–64 (1996).
14. Shukla, J. Predictability in the midst of chaos: A scientific basis for climate forecasting. *science* **282**, 728–731 (1998).
15. Slingo, J. & Palmer, T. Uncertainty in weather and climate prediction. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **369**, 4751–4767 (2011).
16. Hilborn, R. C. Sea gulls, butterflies, and grasshoppers: A brief history of the butterfly effect in nonlinear dynamics. *American Journal of Physics* **72**, 425–427 (2004).
17. Caruana, R. *et al.* Mining citizen science data to predict orevalence of wild bird species. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, 909–915 (ACM, 2006).
18. Sheldon, D. R. *et al.* Approximate bayesian inference for reconstructing velocities of migrating birds from weather radar. In *AAAI* (2013).
19. Sullivan, B. L. *et al.* ebird: A citizen-based bird observation network in the biological sciences. *Biological Conservation* **142**, 2282–2292 (2009).
20. Wikelski, M. & Kays, R. Movebank: archive, analysis and sharing of animal movement data. *World Wide Web electronic publication* (2014).
21. White-fronted geese. http://deltafarmland.ca/admin/userfiles/image/white-front.jpg.
22. White-fronted geese. http://www.crb-photoguide.com/oche/ansal_ned_matthias_haupt.jpg.
23. Kölzsch, A. *et al.* Towards a new understanding of migration timing: slower spring than autumn migration in geese reflects different decision rules for stopover use and departure. *Oikos* (2016).
24. Kranstauber, B., Weinzierl, R., Wikelski, M. & Safi, K. Global aerial flyways allow efficient travelling. *Ecology letters* **18**, 1338–1345 (2015).
25. Gonzalez, R. C., Woods, R. E. *et al.* Digital image processing (2002).



(a)



(b)

Figure 5. Actions observed during 6 years of (a) Spring and (b) Fall migration dataset of white-fronted geese.

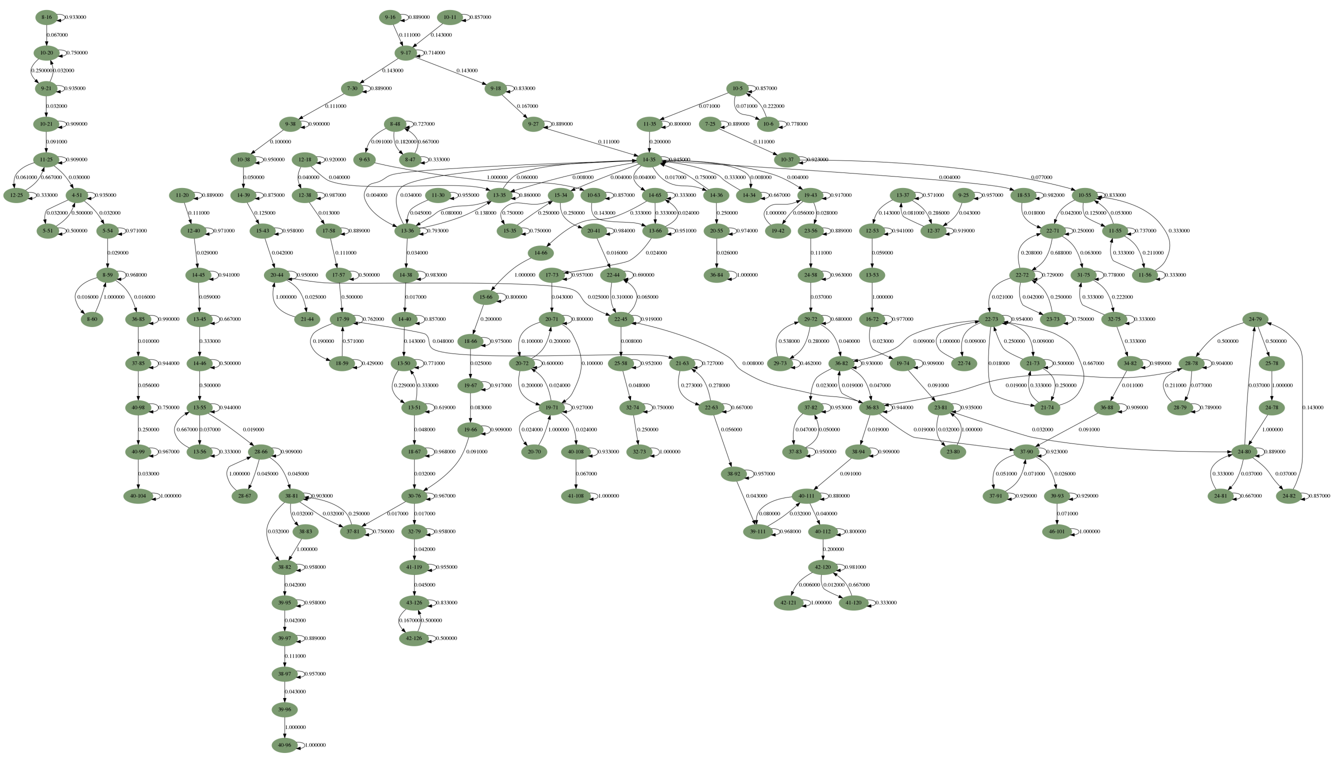


Figure 6. Markov chains network between stopover sites on Spring migration routes.

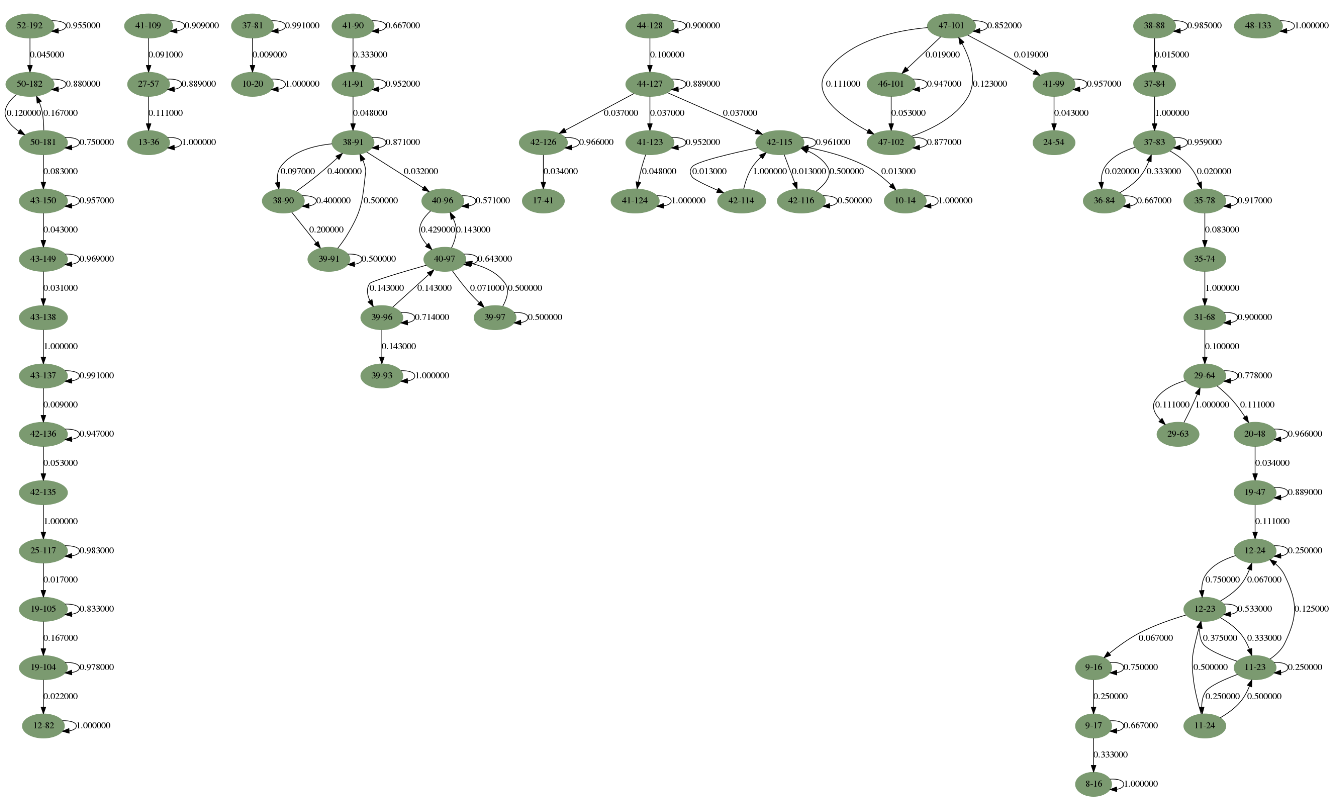


Figure 7. Markov chains network between stopover sites on (a) Spring and (b) Fall migration routes.