*Working title 1: Adapting hidden Markov models to data from small GPS transmitters*

*Working title 2: Adapting hidden Markov models to data from small GPS transmitters reveals variation in American Woodcock migration patterns*

**2 Methods**

*2.1 Collecting data via GPS transmitters*

We delineated American Woodcock movements throughout the full annual cycle using GPS-tracking data from the Eastern Woodcock Migration Research Cooperative, a collaboration of 42 federal, state, provincial, non-profit, and university partners throughout the United States and Canada (www.woodcockmigration.org). We captured woodcock at 42 sites in Quebec, Ontario, Nova Scotia, Maine, Vermont, New York, Rhode Island, Pennsylvania, Maryland, West Virginia, Virginia, North Carolina, South Carolina, Georgia, Alabama, and Florida using mist nets during morning and evening flights (Sheldon 1960), and on night roosts using spotlights and dip nets (Rieffenberger and Kletzly 1966, McAuley et al. 1993). We attached 4g, 5g, and 6.3g PinPoint GPS Argos transmitters (Lotek Wireless Inc., Newmarket, Ontario, CA) to captured woodcock. Transmitters, bands, and attachment materials never exceeded 4% of a bird’s body weight, and all capture and handling were conducted with methods approved by the University of Maine Institutional Animal Care and Use Committee (Protocol # A2020-07-01).

GPS transmitters were programmed with one of several schedules, each of which collected data at a slightly different pace to optimize battery life. Transmitters with a frequent schedule collected locations daily, and typically lasted throughout the duration of a single migration. Transmitters with an infrequent schedule collected locations every 3–5 days, with the aim of collecting data from several migrations. Hybrid schedules combined characteristics of frequent and infrequent schedules, collecting locations every 3–7 days outside of the migratory season, and collecting data every 1–2 days during the migratory season. Gap schedules used a similar timing to hybrid schedules, but additionally included a period of 1–3 months in which the transmitter did not collect locations to extend battery life over the course of multiple seasons. Each transmitter collected tk–tk GPS locations, depending on transmitter size, at 12–60m accuracy. Transmitters relayed GPS locations back to the ARGOS satellite network after every 3rd location, allowing recovery of data without recapture of the bird. However, transmissions back to satellites are energetically demanding and transmitters occasionally fail to relay data as their battery life wanes, creating issues with sporadic missing data near the end of migration.

*2.2 Delineating spring and fall migration*

*2.2.1 Using HMMs to delineate woodcock migration*

For woodcock with a full migratory track (Fall: tk–tk, Spring: tk–tk), we define woodcock migration as beginning when woodcock make their first 30.2 km movement in a season and ending after woodcock make their final 30.2 km movement in a season. However, tagged woodcock in our study frequently had incomplete migratory tracks, either due to waning battery life or bird mortality before the end of the season. We therefore used hidden Markov Models (HMMs) trained on those birds with complete migratory tracks to estimate the final migratory state of birds with incomplete migratory tracks. We made several adjustments to these HMMs to allow them to better function with infrequent locations and missing data and provided additional data streams to allow the models to better differentiate between stopovers and post-migratory locations.

We accommodated for infrequent locations and missing data using a correlated random walk model to interpolate daily locations at stopover, pre-migratory, and post-migratory sites using the R package tk (cite tk). This allowed the HMM to detect recursive movements more accurately during stationary periods using small numbers of infrequent locations. We only used the correlated random walk model to interpolate locations between points that were <30.2 km apart (i.e. when the bird was either at a stopover or not migrating), as the correlated random walk model tended to break long, single night migratory flights into a series of short movements over multiple days, making the HMM less likely to identify these movements as migratory. Due to the directional consistency inherent in correlated random walks, our models occasionally produced erroneous loops of interpolated points to explain directional changes when birds made recursive movements during non-migratory periods. These interpolated loops were often lengthy and could artificially create step lengths >30.2 km. To ensure that HMMs didn’t incorrectly interpret these loops as migratory movements, we removed all loops of interpolated points for which the total length of the loop was more than 10 times the distance between observed points. We replaced these loops with sets of predicted locations spaced evenly along a line between the observed points.

While HMMs generally differentiate between different movement states using step lengths and turn angles, we encountered difficulties inferring the final movement state of incomplete tracks using these data streams alone. To overcome this issue, we added additional data streams that would allow the models to better differentiate between stopovers and post-migratory locations. Additional data streams measured recursiveness, residence time, day of season, and location of a given point (Table tk). We measured recursive movements using the mean distance to the nearest 7 points, transformed using a natural logarithm. This measured whether bird locations over the period of a week reflected intensive use of the same area (presumably resource utilization) or spread-out movement throughout the area (reflecting exploration). We measured residence time using the time difference between the first and last day that the bird was within a 10km radius of the point. This reflected the difference between the amount of time that woodcock spent occupying stopover sites as opposed to their post-migratory sites. We measured day of season using an ordinal day variable, reflecting woodcocks’ tendency to migrate at similar times each year. We measured location using latitude and two binomial variables. The first binomial location variable determined whether a bird had or had not moved tk km from its position at the beginning of the season and had thus departed its initial site to begin migration. The second binomial location variable reflected whether the bird was currently within the woodcock breeding range, as delineated using the eBird 2021 Status and Trends abundance maps (Fink et al. 2022).

Table tk. Covariates for all spring/fall migratory models. Point-specific attributes (Latitude, Ordinal day, Distance from start, breeding range, log distance to nearest points, residence time) are based on the woodcock’s location at the beginning of the step.

|  |  |  |
| --- | --- | --- |
| **Covariate** | **Distribution** | **Description** |
| Step length | Gamma | Length of the current step |
| Turn angle | Wrapped Cauchy | Angle between the current and previous step |
| Latitude | Normal | Latitude at the beginning of the step |
| Ordinal day | Normal | Days since the beginning of the migratory season |
| Distance from start threshold | Bernoulli | Has the bird moved >tk km from its location at the beginning of the migratory season? |
| Step length threshold | Bernoulli | Is the current step length > 30.2km? Implemented with a fixed distribution so that all steps >30.2km are migratory. |
| Breeding range | Bernoulli | Does the step begin in the American Woodcock breeding range (Fink et al. 2022)? |
| Log(distance to nearest points) | Normal | Natural logarithm of the mean distance to the nearest 7 points. |
| Residence time | Normal | Number of days that the bird has spent/will spend within a 10km radius. |

*2.2.2 Model design*

We implemented HMMs in the R package momentuHMM (McClintock and Michelot 2018) due to its accommodation of multiple data streams and customized model structures. We designed separate models for the fall (tk–tk) and spring (tk–tk) migratory seasons. We used separate HMMs for males and females in spring due to differential breeding movements, and together during the fall as movements were similar between sexes. We constructed a multi-state model for each HMM to distinguish between pre-migratory, migratory, and post-migratory movements (Fig. tk). Spring females and fall birds had 4 states in their model: pre-migration, migration, stopover, and post-migration. Pre-migration was the initial state that birds entered at the beginning of the season and continued as long as the bird made no movements >30.2km. When the bird made its first movement >30.2km it moved into a migration state, which was the only state in which movements >30.2km were permitted. From the migration state, birds could enter either a stopover state or a post-migration state. The stopover state was characterized by a lengthy period of recursive movements of <30.2 km, where the only possible state transition was back into the migration state. The post-migration state could only be reached from the migration state and did not allow for any further state transitions. This state characterized when birds had completed movements >30.2km for the rest of the season.

The spring male model used one additional state, settling, to reflect an additional movement state that males entered in spring at the conclusion of migration. The settling state was likely due to male breeding behavior causing differential step lengths immediately after migration that faded as breeding displays concluded near the end of the spring. Birds were able to transition freely between the settling and post-migratory states, but birds were not able to reenter migration after entering either of these states.

Most tagged birds were assigned a fixed pre-migration state for their initial step, as they were captured during either the breeding or the wintering season. However, birds captured in Virginia, New Jersey, and Maryland were sometimes captured either late in the fall migratory season (Aug. 1 – Feb. 25) or early in the spring migratory season (Jan. 1 – date tk). As we could not be certain whether a Virginia or Maryland-caught bird entered the HMM in a pre-migration or a migration state, we allowed the HMM to estimate the initial state of the model instead of using a fixed pre-migration state for the initial step.

During our study, we observed an overlap in a small proportion of birds between the dates on which the final fall migrants settled into their wintering ranges and when the first spring migrants began their spring migrations. We addressed this by using an individual-specific date range for each bird when designating their fall migrations. We delineated the spring migrations before the fall migrations. If a given bird had a fall migration in the season immediately preceding a delineated spring migration, we shortened the end of the default fall migration timeframe (August 1 – Feb 25) to a day before spring migration was set to start for that bird. By doing this, we were able to delineate fall migration using custom timeframes for each bird to overcome the fact that there is no universal cutoff date between spring and fall migration.

*2.3 Assessment of model accuracy*

In all models, transition to post-migratory states was incumbent on no further steps >30.2 km occurring after the transition. As we had a high degree of certainty that no further >30.2 km steps would occur after certain dates (tk in fall, tk in spring) we considered the date of post-migratory transition (the date of the bird’s final >30.2 km movement) to be known for birds that survived and continued to transmit past the cutoff date. For birds which stopped sending data before the cutoff date, we relied on HMMs to determine whether that bird had reached its final >30.2 km movement of the season or was still migrating. We assessed the accuracy of our final state assignments of incomplete tracks using a leave-one-out simulation, in which we artificially removed the ends of tracks of birds with known post-migratory transition dates to see if the model would still correctly determine the movement state that the bird was in at the end of its artificially constrained track. We iterated through 50 individuals for the simulation, testing the accuracy of final state assignment when tracks were cut off to dates from tk–tk in fall and tk–tk in spring. We measured using two metrics modified from statistical hypothesis testing, Type I and Type II error. In our case, Type I errors occurred when an ending state was falsely classified as migratory, when the bird had in fact stopped migrating. Type II errors occurred when an ending state as falsely classified as settling or post-migratory, when the bird was in fact still migrating. We conducted this assessment on the full model for each HMM and a parallel model for each season which used only step length and angle to train the HMM. Running both full and reduced models allowed us to examine how both Type I and Type II errors changed with the inclusion of additional variables in the HMM and allowed us to evaluate whether the inclusion of these variables improved the predictive capacity of the HMM.

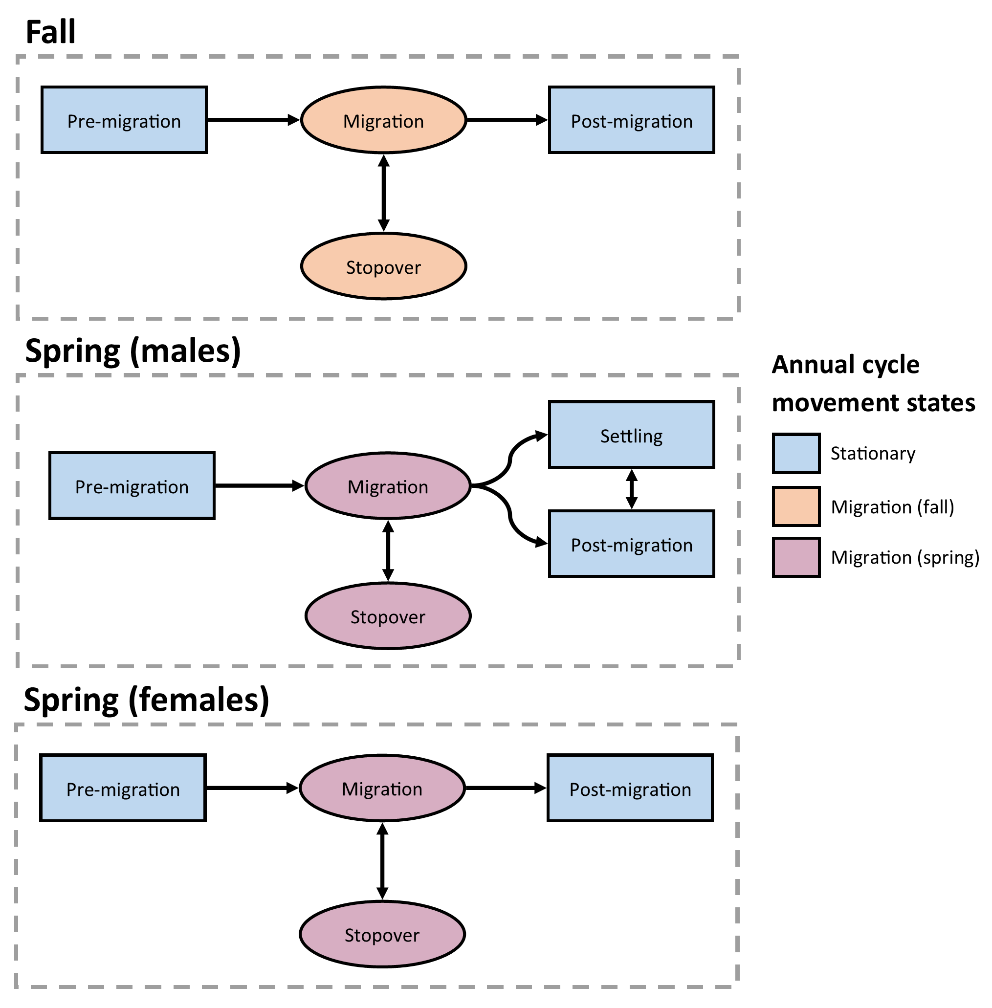


Figure tk.

*2.4 Classifying movements outside of spring and fall migration*

In addition to the timing of fall and spring migration, we were also interested in movements before or after migration had concluded. To examine these movements, we first matched the classifications from the fall and spring HMMs with the locations from each woodcocks’ full trajectory, which included locations in the summer which were not considered in either HMM. Before we matched the classifications, we simplified the classification states to 3 classes, stationary, migration (fall), and migration (spring; Fig. tk). We then manually classified movements from each bird which fell outside the migratory period. We added 3 additional movement states in this manual classification: summer migrations, foray loops, and dispersal movements (Table tk). Summer migrations initiated after the conclusion of breeding activity but prior to the first date of consideration for fall migration (August 1). These movements resulted in southerly displacement which was considerable enough to preclude further migration in the fall. Foray loops were recursive movements which included step lengths >30.2 km but for which the resulting displacement was <30.2 km, reflecting exploratory movements which would eventually return to the original breeding or wintering site. Dispersal movements were one-way movements which included step lengths >30.2 km and resulted in >30.2 km of displacement, reflecting one-way movements to a new breeding or wintering site. All 3 additional movement states were further constrained to begin and end in a stationary state. For example, a bird could make a dispersal movement after its migration had concluded, but it could not transition directly from a migration state into a dispersal event, as this would be classified as a continued migration.

Table tk. Definitions of all full annual cycle movement states delineated for American Woodcock.

|  |  |
| --- | --- |
| **State** | **Definition** |
| Stationary | Recursive movements with steps < 30.2km that are associated with non-migratory behavior. |
| Migratory (fall/spring) | Directional movements along a latitudinal gradient that bring the bird from the breeding to the wintering range, or visa-versa. Steps can be >, =, or < 30.2 km. Fall migratory movements initiate between Aug 1 and Feb 1, while spring migratory movements initiate between Jan 1 and June 1. |
| Migratory (summer) | Post-breeding, southerly movements that initiate before Aug 1 and would preclude further migratory movements in the fall. |
| Foray loops | Movements that include step lengths > 30.2 km but result in < 30.2 km of net displacement between the first and last point. Can occur in any season, provided the starting and ending state is stationary. |
| Dispersal events | Movements that include step lengths > 30.2 km and result in > 30.2 km of net displacement between the first and last point, and do not preclude a fall or spring migration. Can occur in any season, provided the starting and ending state is stationary. |

*2.5 Exceptions to the rules and mortalities*

Because of the incredible variation that we observed in the data, there were several occasions in which we needed to make individual exceptions to the rules to allow a bird’s HMM to fit correctly. Our analysis was also stymied by mortality issues, as birds which died during migration occasionally continued to transmit from the mortality location, causing the HMM to think that they were still alive and had ceased migrating. We have detailed these individual bug fixes and exceptions, as well as methods used in delineating GPS mortalities and removing them from the dataset, in Appendix tk.

Remove birds with no individual step lengths >30.2km (20 miles). In practice this removes birds that never initiate a substantial migratory movement but doesn't penalize birds that DO migrate, as they always make at least one substantial movement.

**3 Results**

*3.1 Full annual cycle phenology*

Basic migration phenological stats

Fall migrations lasted tk ± tk days and were tk ± tk km long. Spring migrations lasted tk ± tk days and were tk ± tk km long for males, and lasted tk ± tk days and were tk ± tk km long for females.

The median date of departure for fall migration was Nov. 8th (Aug. 29th–Jan. 11th), and the median date of migratory termination was Dec. 1st (Oct. 25th–Feb. 4th). The median date of departure for males during spring migration was Mar. 1st (Jan. 6th–Apr. 14th), and the median date of migratory termination was Apr. 4th (Jan. 11th–Jun. 4th). The median date of departure for females during spring migration was Mar. 6th (Jan 19th–Apr. 26th), and the median date of migratory termination was Apr. 21st (Feb. 20th–Jul. 28th).

Fall migration peaked on the week of tk–tk, with X% of locations in that week being migratory. Spring migration peaked on the week of tk–tk for males and tk–tk for females, with tk% and tk% of locations in that week being migratory, respectively.

Incidence and statistics for each behavior (non-migratory, auxiliary)

X% of birds tracked from tk–tk skipped their fall migration, and X% of birds tracked from tk–tk skipped their spring migration. These non-movements included tk birds overwintering in tk, and the birds spending the breeding season in tk.

X% of birds tracked from tk–tk underwent a summer migration. Summer migratory movements initiated between tk and tk and terminated between tk and tk. These movements lasted tk–tk days and were tk–tk km long.

X% of birds tracked for at least tk months underwent a dispersal movement. Dispersal movements were observed in December, February, June, and July, and were restricted to the northeastern United States. They lasted tk–tk days and were tk–tk km long.

X% of birds tracked for at least tk months underwent a foray loop. Foray loops were observed in all months except for March and April during the peak of spring migration. They lasted tk–tk days and were tk–tk km long.

*3.2 Model evaluation*

How many birds did weird things that required me to make additional rules for them?

The full fall model had a comparable Type I error rate (0–28.6%) to the reduced model (0–28.2%), but the full model had a Type II error rate (0–4.0%) which was lower than that of the reduced model (0–26.1%).

The spring male full model exhibited lower rates of Type I (0–13.0%) and Type II (0–8.9%) error during all time periods than the reduced model (Type I: 0–55.6%, Type II: 0–19.6%).

The spring female full model exhibited higher Type I errors (0–20.8%) during some time periods than the reduced model (2.0–17.9%), but the full model also experienced substantially reduced Type II error rates (2.0%–16.7%) than the reduced model (7.7–70.0%).

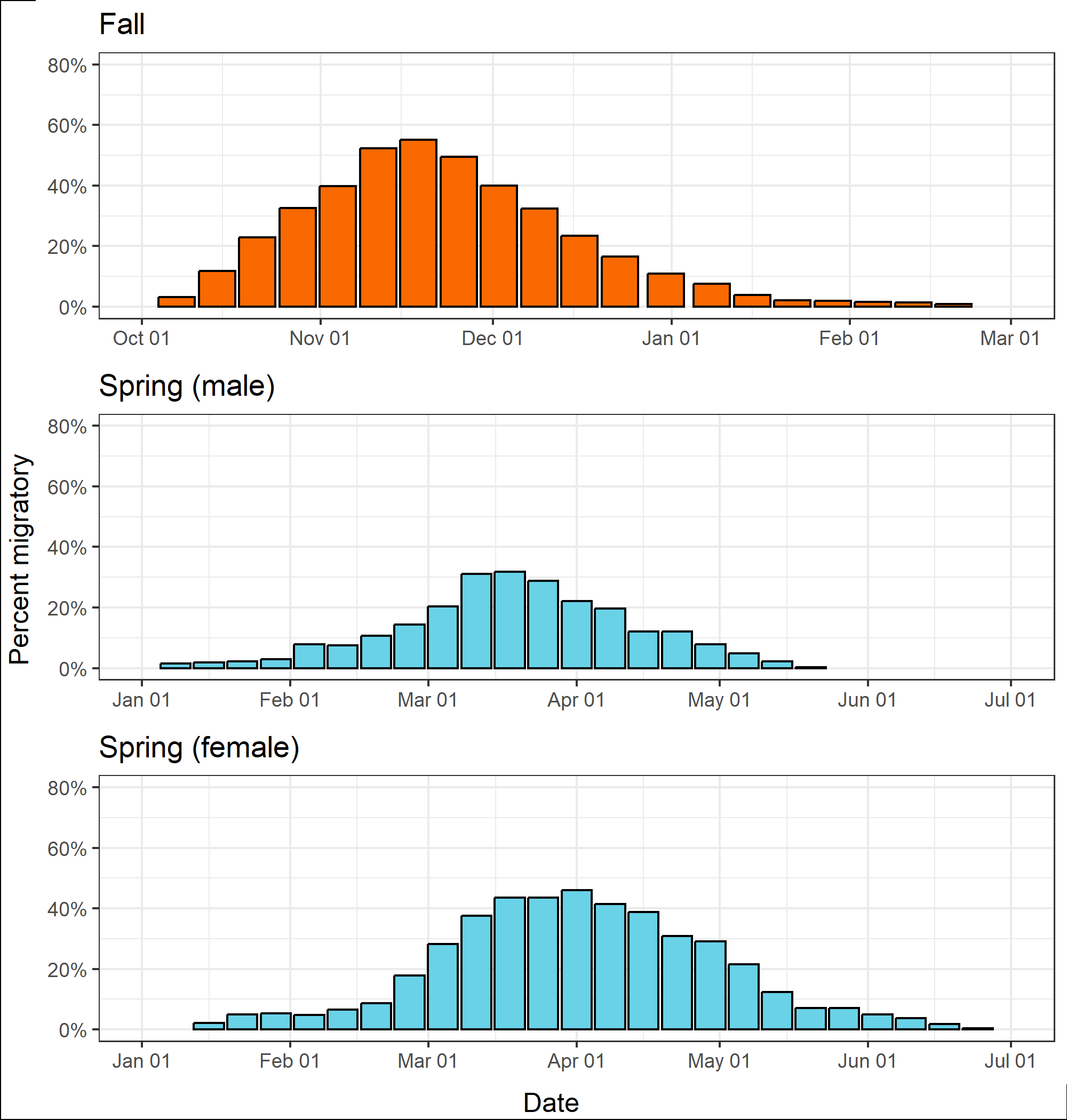


Figure tk. The Y axis is the percent of all locations in a week that belong to a give migratory state.

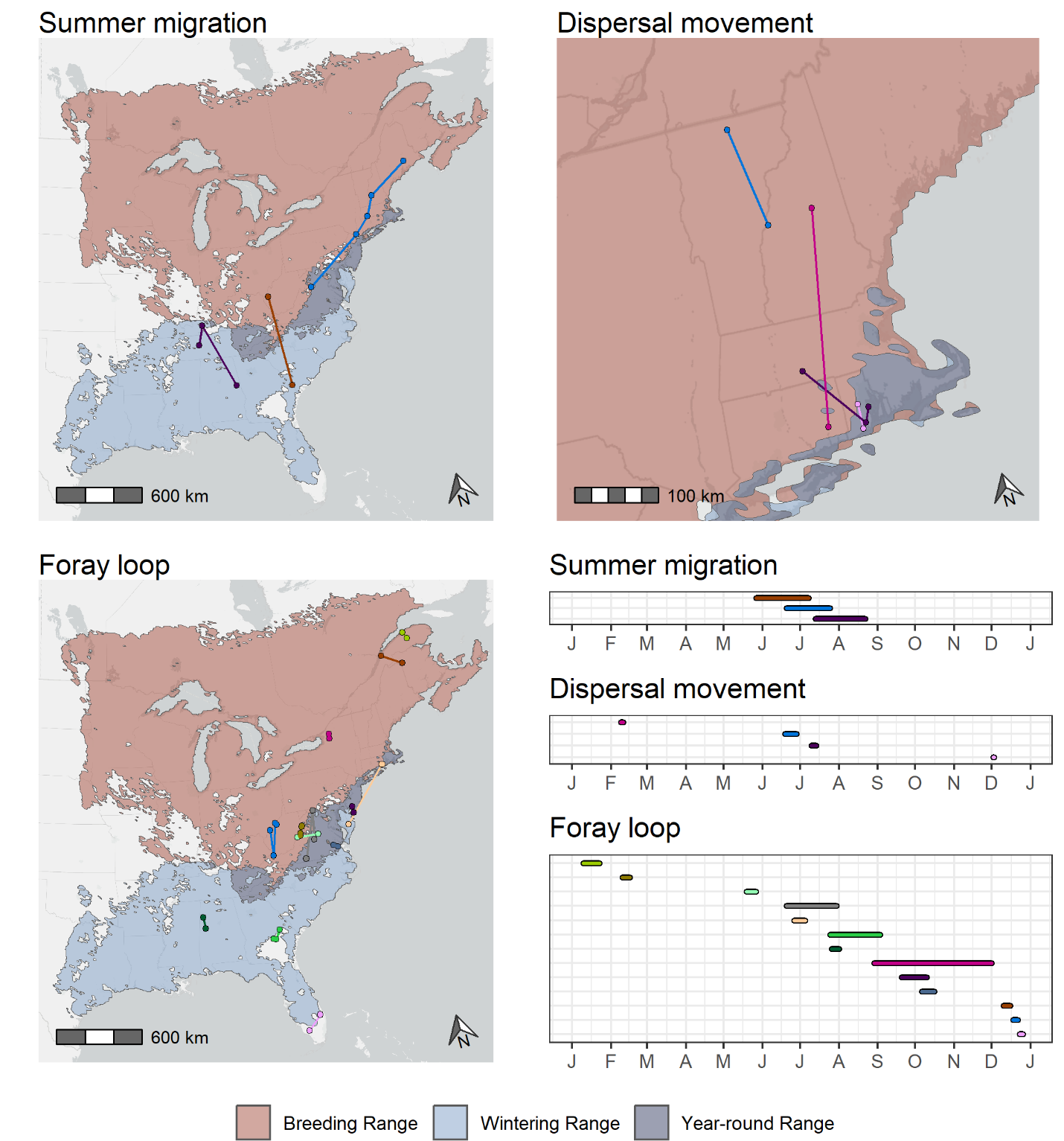


Figure tk

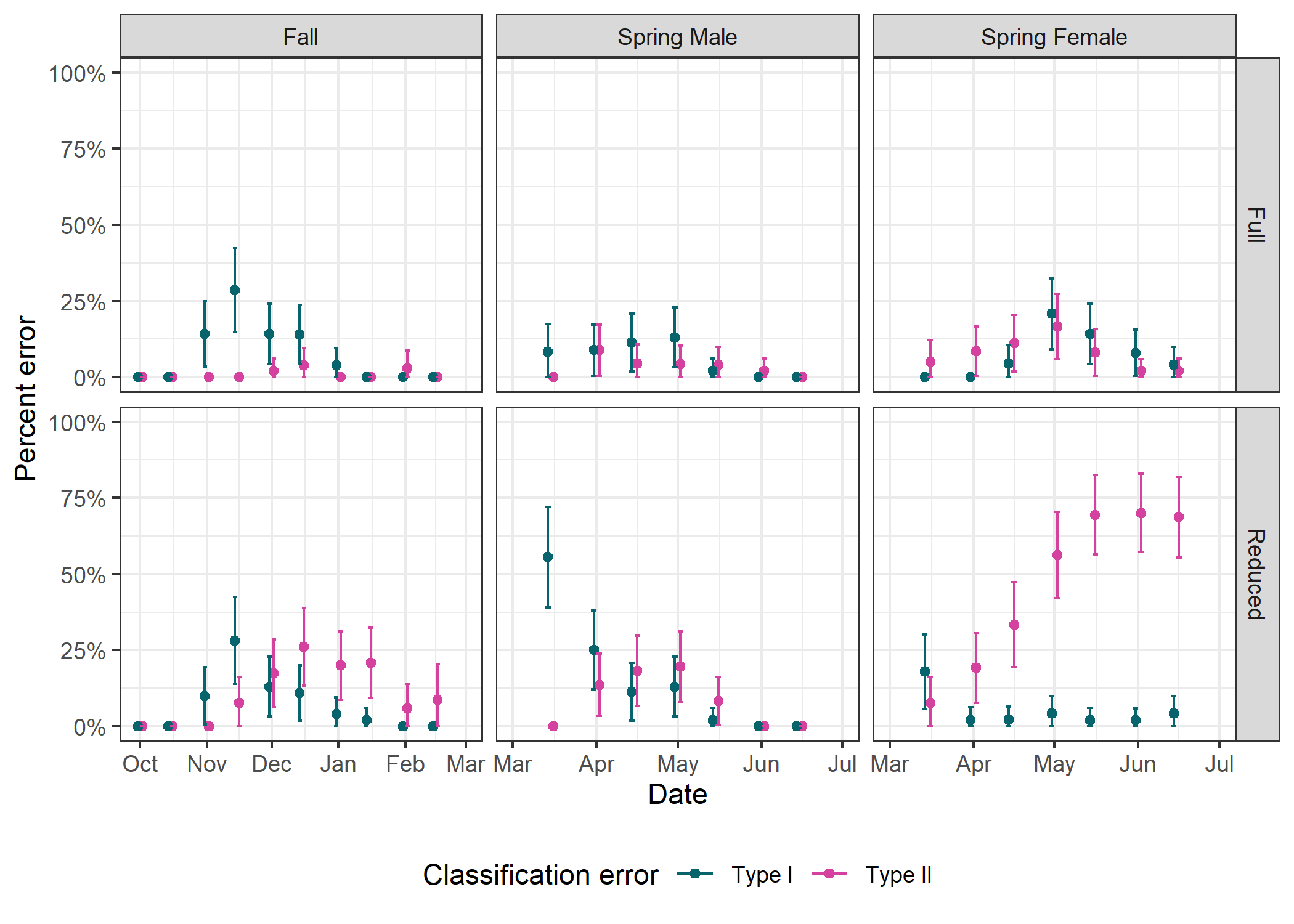


Figure tk. Type 1 errors are falsely classified as post-migratory while the true state is migratory. Type 2 errors are falsely classified as migratory while the true state is post-migratory. Bars reflect the 95% confidence interval of the mean. The full model was an HMM including all variables, while the reduced model included only step lengths and turn angles.

**References**

Fink, D., T. Auer, A. Johnston, M. Strimas-Mackey, S. Ligocki, O. Robinson, W. Hochachka, L. Jaromczyk, A. Rodewald, C. Wood, I. Davies, and A. Spencer. 2022. eBird Status and Trends. Cornell Lab of Ornithology, Ithaca, New York.

McAuley, D. G., J. R. Longcore, and G. F. Sepik. 1993. Techniques for Research into Woodcocks: Experiences and Recommendations. Page 5 *in*. Proceedings of the eighth American woodcock symposium. Volume 16. U.S. Fish and Wildlife Service.

McClintock, B. T., and T. Michelot. 2018. momentuHMM: R package for generalized hidden Markov models of animal movement. Methods in Ecology and Evolution 9:1518–1530.

Rieffenberger, J. C., and R. C. Kletzly. 1966. Woodcock night-lighting techniques and equipment. WH Goudy, compiler. Woodcock research and management 33–35.

Sheldon, W. G. 1960. A method of mist netting woodcocks in summer. Bird-banding 31:130–135.