**Supplementary material: Enhancing hidden Markov models using additional data streams can better track a migratory bird with incomplete or missing GPS data**

Liam Berigan, Rachel Darling, Sarah Clements, Alex Fish, Amber Roth, Greg Balkcom, Bobbi Carpenter, Gary Costanzo, Jeffrey Duguay, Kayleigh Filkins, Clayton Graham, William Harvey, Michael Hook, Douglas Howell, Seth Maddox, Scott McWilliams, Shawn Meyer, Theodore Nichols, J. Bruce Pollard, Christian Roy, David Sausville, Colby Slezak, Josh Stiller, Jacob Straub, Mathieu Tetreault, Dawn Washington, Lisa Williams, Erik Blomberg.

**S1 – Individual exceptions**

Throughout the study, we encountered 45 of 522 (9%) classified migratory tracks that failed to fit the parameters of the full model due to erratic behavior. These exceptions were broadly categorized as misclassification errors due to the presence of foray loops (n = 19 misclassification errors), directional movements (3), and summer migrations (2), as well as errors caused by continued migration after the end of the HMM’s consideration period (14) and transmissions beginning after the start of migration (3). We additionally edited the known state classification for 4 birds, two of which were captured on their nests late in spring migration, one which was recaptured at a suspected post-migratory site several months after its transmitter had prematurely died, and one which continued regular ≥16.1 km movements between sections of its wintering home range. We modified the HMM classification process for these individuals by removing locations or setting known movement states to ensure that these tracks were classified correctly by our HMMs (Table S1).

The base model required all the same individual exceptions as the full model, in addition to several additional fixes to pre-migratory state delineation (n = 11). All eleven exceptions were made to individuals captured in New Jersey or Virginia for whom the initial state was erroneously estimated to be migratory, despite pre-existing knowledge that the bird was in a pre-migratory state (either due to capture at that site prior to the migratory season or migratory data from previous seasons). For all 11 exceptions we addressed the issue by setting a known pre-migration state for the first location in that individual’s seasonal HMM.

Table S1. Individual exceptions to the full model ruleset made in hidden Markov Model delineation to improve seasonal model fits for individual birds. 1 Year in which the issue occurred. 2 Seasonal HMM that encountered the issue. 3 Short description of why the edit was necessary. 4 Steps that were taken to alleviate the corresponding issue.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Bird ID** | **Year** | **Season2** | **Issue3** | **Edit4** |
| RI-2020-31 | 2020 | Fall | Foray loop caused early initiation of migration | Removed points on Aug. 29th and Nov. 21st from the HMM |
| NY-2018-04 | 2018 | Fall | Foray loop caused late termination of migration | Removed points on Nov. 21st–23rd from the HMM |
| NJ-2018-03 | 2018 | Fall | Foray loop caused an apparent migration | Removed points on Jan. 10th and 18th from the HMM |
| PA-2018-01 | 2018 | Fall | Foray loop caused late termination of migration | Removed points on Dec. 18th–20th from the HMM |
| ME-2018-08 | 2018 | Fall | Foray loop caused late termination of migration | Removed points on Dec. 23rd–25th from the HMM |
| PA-2019-15 | 2019 | Fall | Foray loop caused late termination of migration | Removed points on Dec. 10th–13th from the HMM |
| VA-2021-92 | 2021 | Fall | Foray loop caused early initiation of migration | Removed points on Sep. 20th–October 6th from the HMM |
| RI-2021-46 | 2021 | Fall | Foray loop caused late termination of migration | Removed points on Feb. 10th–14th from the HMM |
| RI-2021-59 | 2021 | Fall | No locations prior to the start of migration due to transmitter glitch | Set a known stopover state for the first location in the HMM |
| VA-2019-48 | 2019 | Fall | Bird settles after Feb. 25th | Set a known post-migration state for the final location in the HMM |
| RI-2021-58 | 2021 | Fall | Bird settles after Feb. 25th | Set a known post-migration state for the final location in the HMM |
| RI-2019-29 | 2019 | Fall | Series of directional movements caused an apparent migration | Removed bird from HMM classification |
| SC-2020-13 | 2020 | Fall | Summer migration erroneously classified as a fall migration | Removed from HMM classification |
| GA-2021-18 | 2021 | Fall | Foray loop caused an apparent migration | Removed from HMM classification |
| NY-2018-06 | 2018 | Fall | Foray loops caused late termination of migration | Removed points on Nov. 17th–18th and Dec. 25th–27th from the HMM |
| PA-2021-37 | 2021 | Fall | Foray loops caused late termination of migration | Removed points on Dec. 8th from the HMM |
| NY-2018-03 | 2018 | Fall | Foray loops caused early initiation of migration | Removed points on Oct. 6th–15th from the model |
| PA-2021-34 | 2021 | Fall | Regular movements ≥16.1 km between two sections of the wintering home range caused late termination of migration. | Set a known post-migratory state on Nov. 7th. |
| VA-2020-52 | 2020 | Spring (male) | Directional movement caused late termination of migration | Removed points on Jun. 18th–28th from the HMM |
| NJ-2018-03 | 2019 | Spring (male) | Foray loop caused early initiation of migration | Removed points on Jan. 10th–17th from the HMM |
| RI-2019-21 | 2020 | Spring (male) | Late termination of fall migration, never initiated spring migration | Removed from HMM classification |
| RI-2019-29 | 2020 | Spring (male) | Directional movement caused an apparent migration | Removed from HMM classification |
| VA-2018-03 | 2018 | Spring (male) | Foray loop caused an apparent migration | Removed from HMM classification |
| FL-2021-01 | 2021 | Spring (male) | Foray loop caused an apparent migration | Removed from HMM classification |
| NJ-2018-08 | 2019 | Spring (female) | Late termination of fall migration | Removed points prior to Jan. 12th from the HMM |
| NJ-2018-15 | 2019 | Spring (female) | Late termination of fall migration | Removed points prior to Jan. 14th from the HMM |
| NJ-2018-13 | 2019 | Spring (female) | Late termination of fall migration | Removed points prior to Feb. 5th from the HMM |
| RI-2020-31 | 2021 | Spring (female) | Late termination of fall migration | Removed points prior to Mar. 2nd from the HMM |
| ME-2018-13 | 2019 | Spring (female) | Late termination of fall migration | Removed points prior to Jan. 8th from the HMM |
| RI-2021-46 | 2022 | Spring (female) | Foray loop caused early initiation of migration | Removed points on Feb. 10th–14th from the HMM |
| NY-2018-07 | 2019 | Spring (female) | Foray loop caused early initiation of migration | Removed points on Dec. 2nd–Dec. 9th from the HMM |
| RI-2020-44 | 2021 | Spring (female) | Foray loop caused early initiation of migration | Removed points on Dec. 4th–Dec. 18th from the HMM |
| NJ-2018-13 | 2019 | Spring (female) | Locations began later in the season than other New Jersey transmitters, creating issues with initial state estimation | Excluded from initial state estimation |
| NJ-2018-15 | 2019 | Spring (female) | Locations began later in the season than other New Jersey transmitters, creating issues with initial state estimation | Excluded from initial state estimation |
| NY-2022-40 | 2022 | Spring (female) | Bird captured on the nest in late spring that continued migration after nest failure | Set a known stopover state for the first location in the HMM |
| RI-2018-11 | 2019 | Spring (female) | Bird recaptured at the terminal site the next fall, so it is known to have settled | Set a known post-migration state for the final location in the HMM |
| NS-2019-02 | 2020 | Spring (female) | Late termination of fall migration | Removed points prior to Jan. 28th from the HMM |
| RI-2020-42 | 2021 | Spring (female) | Late termination of fall migration | Removed points prior to Jan. 23rd from the HMM |
| RI-2021-47 | 2022 | Spring (female) | Fall migration does not terminate | Removed from spring HMM classification |
| RI-2021-52 | 2022 | Spring (female) | Fall migration does not terminate | Removed from spring HMM classification |
| SC-2019-03 | 2019 | Spring (female) | Late termination of spring migration | Used Jul. 30th, instead of Jun. 30th, as the last date of consideration for the HMM |
| VA-2019-36 | 2020 | Spring (female) | Late termination of spring migration | Used Jul. 30th, instead of Jun. 30th, as the last date of consideration for the HMM |
| VA-2020-66 | 2021 | Spring (female) | Foray loop caused early initiation of migration | Removed points on January 24th - 29th from the HMM |
| NY-2022-42 | 2022 | Spring (female) | Bird captured on the nest in late spring that continued migration after nest failure | Set a known stopover state for the first location in the HMM |

**S2 – HMM ruleset bug fixes**

Each seasonal HMM was set to follow certain rules regarding possible state designations based on step length. For example, the only state in any HMM that was allowed to have step lengths ≥16.1 km was migration. The pre-migration state could only occur before the first ≥16.1 km step was observed, and the post-migration state could only occur after the last ≥16.1 km step was observed. These rules were enforced within the HMMs using 2 mechanisms. The first was a fixed, large negative value for certain state transition coefficients, which effectively acted to prohibit movements between states that would not correspond to the diagram shown in Fig. 1 in the main text (e.g. a bird could move from pre-migration into migration, but not from migration into pre-migration). The second was a fixed probability of ~0 that birds in any state other than migration would have a step length ≥16.1 km, and a fixed probability of ~1 that birds in the migration state would have a step length ≥16.1 km. These fixed parameters effectively prohibited most state assignments that fell outside of our ruleset, but in 2 circumstances the HMM failed to enforce these rules. The first occurred when birds that stayed in the pre-migration state later into the year than expected by the HMM, which would sometimes result in the HMM classifying later pre-migratory movements as stopover locations, despite the lack of a ≥16.1 km step between the interpolated pre-migration and stopover states. The second occurred when birds entered the post-migration state earlier in the year than expected by the HMM, which occasionally resulted in the HMM classifying earlier post-migration locations as stopover locations despite the lack of an intervening ≥16.1 km movement. These two issues were both fixed after HMM classification using code to identify transitions between pre-migration and stopover, as well as stopover and post-migration, without an intervening ≥16.1 km step, and manually assigning pre-migration or post-migration states to the erroneously classified locations. This resulted in reclassifications for 35 seasonal migratory tracks, including 8 changes to pre-migratory classifications and 27 changes to post-migratory classifications.

**S3 – Modeling bird mortality**

The Pinpoint GPS transmitters used during this study usually stopped transmitting upon bird mortality due to attenuation of the signal when the antenna touched the ground. However, there were some circumstances in which transmitters continued to transmit when antennae remained upright after the bird’s death. To recognize and filter out these occurrences, we designed a two-stage process for recognizing and removing the locations of deceased birds from our dataset. The first step was an automated process, which used a hidden Markov Model (HMM) to recognize locations from birds that had ceased making normal movements. We trained this HMM using a subset of 413 training locations gathered during transmitter testing, when 10 transmitters were left on the ground to gather 1 location per minute for ~40 minutes. During this test, we placed 2 transmitters under negligible vegetation cover (short grass, height: ~10 cm), 3 transmitters under low cover (tall grass, height: ~100 cm), 2 transmitters under medium cover (early successional aspen stand, canopy height: ~8 m), and 3 transmitters under high cover (mature deciduous forest, canopy height: ~15 m). We collected these data to provide a balanced sample size from each cover type and demonstrate the likely patterns of locations produced by stationary transmitters on deceased birds under a variety of vegetative cover circumstances.

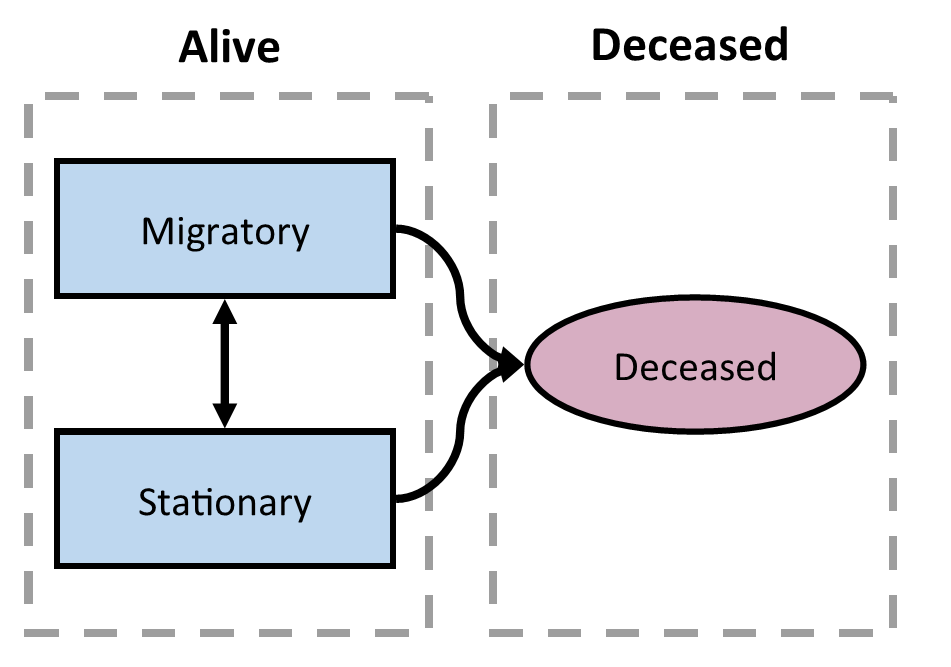


Figure S1. Movement state transition diagram for the hidden Markov Model used to identify potential mortalities among tagged American Woodcock (*Scolopax minor*). Woodcock were allowed to begin their track in any state, and transition freely between either of the two living states (migratory and stationary). Birds that entered the deceased state, however, were forced to remain in this state for the remainder of their track.

The HMM trained on these data used step length, turn angle, and mean distance to the nearest 15 points in a 3-state model to determine whether a bird was likely deceased when locations were recorded. The mean distance to the nearest 15 points metric was decided based on an exploratory analysis of the training dataset, where we determined that the mean distance to the nearest 15 points metric was more consistent between individuals than alternative metrics (mean distance to the nearest 5 and 10 points). The HMM was trained to identify 3 states in the data: migratory, stationary, and deceased. The migratory and stationary states were both living states, and birds were allowed to transition between them freely. Deceased was a terminal state, which birds could enter from either of the two living states (Fig. S1). Among 512 birds in our dataset, the HMM identified 137 individuals that had potentially experienced mortality and continued transmitting. We confirmed these mortalities during a second step, in which we manually checked and adjusted the dates for all potential mortalities identified by the HMM. We only confirmed a mortality during the second step if the following 2 criteria were met:

1. The bird had ≥15 mortality locations
2. At least half of mortality points fell within a threshold distance of the centroid, with the threshold distance varying based on the dominant land cover type as shown in Table S2.

Table S2. 50% threshold values for 4 land cover types, demonstrating the distance (m) from the centroid within which at least half of GPS points fall when a GPS transmitter is taking locations while stationary. The 50% threshold values are represented by the variable X in the criterion “At least half of mortality points fell within X m of the centroid, with X varying based on the dominant land cover type”.

|  |  |
| --- | --- |
| **Dominant land cover** | **50% threshold distance (m)** |
| Young forest | 4.93 |
| Mature forest | 10.85 |
| Short grass | 5.33 |
| Tall grass | 3.93 |

The threshold distance ensured that we had enough locations to determine that the bird was indeed stationary, and the distance to centroid threshold allowed us to ensure that the bird's movements occurred within a distance consistent with the GPS error associated with the locations’ land cover type. We required that 50% of all mortality points fell within the distance to centroid threshold, with the 50% value chosen based on consistency between individuals in the training dataset. Distance to centroid thresholds were set based on the mean values observed among all individuals in that cover type in the training dataset. Dominant land cover type was assessed via publicly available satellite imagery [1] for both the test and training datasets.

We manually classified a mortality event when both criteria were met. In certain circumstances where mortalities met one threshold but came just shy on the other, we made the final determination regarding whether a mortality had indeed occurred. Of 137 potential mortalities, we determined 20 to be true mortality events. The code used in this delineation is publicly available at https://github.com/EWMRC/mortality-detection.

**S4 – Simplified movement state designations for Movebank**

Prior to use in other studies, we simplified our hidden Markov model (HMM) classification states to 3 classes: stationary, fall migration, and spring migration. The stationary class included all pre-migration and post-migration locations, while the fall and spring migration classes included migration and stopover locations. We added 3 additional classification states to represent long-distance movements outside of spring and fall migration: summer migration, foray loops, and directional movements (described in the section of this manuscript titled *Long-distance movements and non-migratory individuals*). This class structure provided a simplified framework for delineating woodcock migratory phenology and long-distance movements throughout the full annual cycle, which could be applied to habitat use or survival analyses (Fig. S2). These classes were uploaded to Movebank (repository 351564596) to facilitate collaborator access to data.

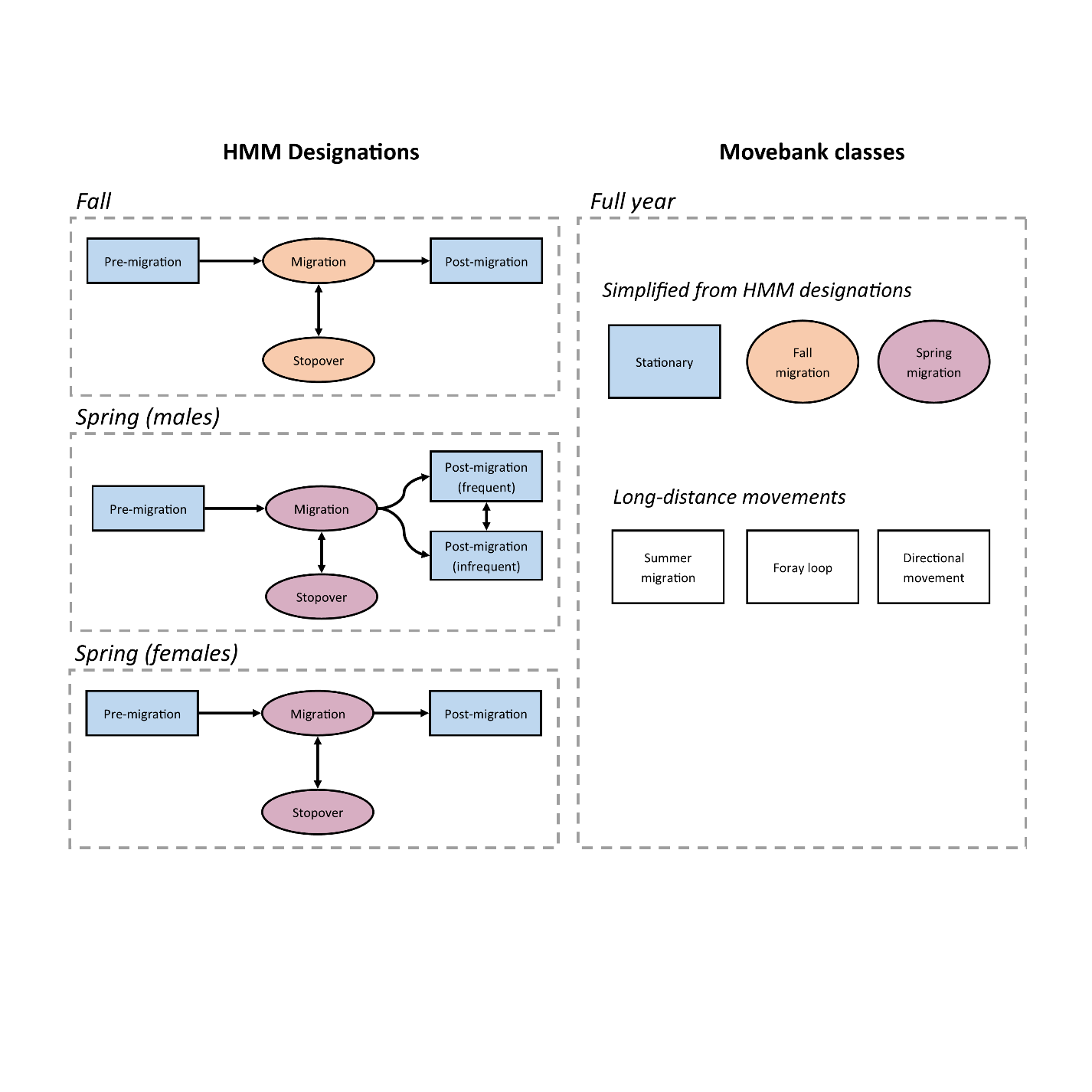


Figure S2. Movement state transition diagram for each hidden Markov Model (HMM), and corresponding classes uploaded to the Movebank repository.

**References**

1. OpenStreetMap contributors. OpenStreetMap. 2023. Available from: www.openstreetmap.org