

Quantifying the strength of migratory connectivity

Emily B. Cohen^{1†}  | Jeffrey A. Hostetler^{1,2†} | Michael T. Hallworth¹ |
Clark S. Rushing^{1,3} | T. Scott Sillett¹ | Peter P. Marra¹

¹Migratory Bird Center, Smithsonian Conservation Biology Institute, National Zoological Park, Washington, DC, USA

²Fish and Wildlife Research Institute, Florida Fish and Wildlife Conservation Commission, St. Petersburg, FL, USA

³Patuxent Wildlife Research Center, U.S. Geological Survey, Laurel, MD, USA

Correspondence

Emily B. Cohen

Email: cohene@si.edu

Jeffrey A. Hostetler

Email: hostetlerj@si.edu

Funding information

United States Fish and Wildlife Service; ConocoPhillips Corporation; U.S. National Science Foundation; Smithsonian Institution for CSR's Smithsonian Fellowship

Handling Editor: Luca Börger

Abstract

1. Technological advancements have spurred rapid growth in the study of migratory connectivity, the linkage of individuals and populations between seasons of the annual cycle. The strength of migratory connectivity is a measure of the co-occurrence of populations throughout the annual cycle and can be represented by a correlation of the distances between individuals during one season and another (Mantel correlation, r_M). However, measurement of seasonal distributions most often involves incomplete sampling and use of technologies that vary in accuracy and precision. For these reasons, we expanded r_M to measure the strength of migratory connectivity (MC) with population-specific transition probabilities that can be derived from many data types and uneven sampling.
2. We explore the sensitivity of MC to possible real-world variation in input parameters: transition probabilities, abundance among regions, spatial arrangement of regions, and sample sizes. We compare MC to r_M , present a series of resampling approaches for propagating uncertainty in input values into estimation of MC and r_M , and validate the method with bird tracking data.
3. Migratory connectivity was negative when populations are further apart between seasons, positive when populations remain together between seasons, and zero when populations have no patterns in distribution between seasons. MC is most sensitive to transition probabilities and spatial arrangement of regions and performs better than r_M when sampling effort is not proportional to true abundance, and when the strength of migratory connectivity varies across the range of the species. Our estimators for MC and r_M performed well across several data types.
4. We hope that these methods and the MigConnectivity `R` package will facilitate quantitative comparisons of migratory connectivity across studies, data types, and taxa to better understand the causes and consequences of the seasonal distributions of populations. Several study design recommendations emerge from our simulations: (1) incorporate abundance among regions when sampling is not proportional; (2) measure transition probabilities across as much of the range as logistically possible; (3) define study regions with either biological information about population delineation, or use discrete study locations as centroids of regions; and (4) estimate and report uncertainty from appropriate sources of sampling and process errors.

[†]These authors contributed equally to this manuscript.

KEYWORDS

animal movement, full annual cycle, migration, movement ecology, population distribution, population ecology

1 | INTRODUCTION

Migration, the regular and repeated seasonal movements of animals between distinct geographic regions, is a taxonomically widespread global phenomenon. Where and when animals occur throughout the annual cycle is driven by complex behavioral, ecological, and evolutionary processes and can have profound consequences for populations and species. Migratory connectivity describes the spatial and temporal linkages of individuals and populations between seasons that result from migratory movements (Marra, Norris, Haig, Webster, & Royle, 2006; Webster, Marra, Haig, Bensch, & Holmes, 2002). Research on migratory connectivity has increased dramatically with the growing recognition of the importance of full annual cycle biology (Marra, Cohen, Loss, Rutter, & Tonra, 2015) and the improved analytic (e.g. Cohen, Hostetler, Royle, & Marra, 2014; Hostetler, Sillett, & Marra, 2015; Thorup, Korner-Nievergelt, Cohen, & Baillie, 2014), molecular (e.g. Hobson et al., 2014; Ruegg, Anderson, Boone, Pouls, & Smith, 2014; Rushing, Ryder, Saracco, & Marra, 2013), and technological (e.g. Bridge et al., 2011; Hallworth & Marra, 2015) tools to study the movements of migratory animals. The strength, or degree, of migratory connectivity is the extent to which individuals and populations remain together between seasons of the annual cycle. When migratory connectivity is strong, or high, animals remain close during multiple seasons of the annual cycle; when migratory connectivity is weak, or low, animals that are close during one season of the annual cycle are not close during other seasons. For example, breeding populations that remain together during the winter have strong migratory connectivity during stationary non-breeding (hereafter, non-breeding) and breeding populations that use the same migration routes at the same times have strong migratory connectivity during migration. Conversely, breeding populations that separate in winter have weak migratory connectivity during non-breeding (Finch, Butler, Franco, & Cresswell, 2017) and breeding populations that use different migration routes, or pass through them at different times, have weak connectivity during migration (Bauer, Lisovski, & Hahn, 2015).

The strength of migratory connectivity is fundamental to a full-annual cycle perspective on population limitation and conservation because it describes the extent to which populations co-occur between seasons and, therefore, are exposed to the same environmental conditions and selective pressures (Webster et al., 2002). The events and conditions that populations are exposed to throughout the year have consequences for many biological processes, including migration schedules, reproductive success, survival and abundance, range limits, and natal dispersal (Esler, 2000; Finch et al., 2017; Hostetler et al., 2015; Marra et al., 2015). Consequently, quantitative measurement of migratory connectivity strength is key to advancing full annual cycle biology. At the same time, animal movement data is increasingly

available at the spatial and temporal scales necessary to measure migratory connectivity for many species.

Methods to quantify the strength of migratory connectivity should be equally applicable among taxonomic groups, range sizes and shapes, and seasons of the annual cycle, e.g. breeding and non-breeding or breeding and migration. Measures of migratory connectivity strength also need to be comparable among all of the data types available to measure migratory movements, e.g. ring recoveries, tracking, and isotopic or genetic assignments. The true strength of migratory connectivity for a species could be calculated from the correlation of the distances between every pair of individuals during two seasons, although such unbiased and comprehensive measurement is not attainable. Migratory connectivity can be estimated using a Mantel correlation (r_M) on a sample of animals located in both seasons (Besag & Diggle, 1977). Ambrosini, Møller, and Saino (2009) used r_M and a clustering algorithm to measure the strength of migratory connectivity for birds banded on the breeding range and recovered on the non-breeding range. Band recoveries are among the most accurate available for passerine birds, but these data are notoriously sparse for many species and require accounting for considerable geographic bias in marking and detection (Korner-Nievergelt et al., 2010; Thorup et al., 2014). Tracking and molecular data are increasingly available to assign animals to regions within seasonal ranges, but are generally restricted to a limited set of individuals from a priori selected study locations. Even when migration data are available from a species' range, sample sizes per site are often small, and sampling effort may not be proportional to abundance. We know of only one study of migratory connectivity strength, that has accounted for abundance among sampling regions (Hallworth, Sillett, Van Wilgenburg, Hobson, & Marra, 2015). Further, methods used to measure the seasonal distributions of individuals vary in accuracy and precision and the regions between which migratory movements are measured are rarely defined with information about population delineation.

Because information about seasonal movement is often derived from incomplete sampling with data types that vary in accuracy and precision, we undertook an extension of r_M that accounts for population-specific movement (transition probabilities) between seasons and for sampling that is not proportional to abundance. The use of transition probabilities is broadly applicable because they can be derived from any data type between any two seasons of the annual cycle. Furthermore, biased and incomplete sampling can introduce error into the measurement of migratory connectivity, as can natural variability in animal movement such as breeding dispersal. Our objective was to develop methodology to incorporate uncertainty due to multiple sources of sampling error into estimates of migratory connectivity strength. Here, we measure the strength of migratory connectivity (MC) with population-specific transition probabilities

derived from many data types and uneven sampling. We explore the properties of MC and use simulation to test its sensitivity to possible real-world variation in input values and sampling regimes. We compare MC and r_M under many sampling scenarios, present a series of resampling approaches for propagating uncertainty in input values into estimation of MC, and validate the method with simulated and real data of several types. The accompanying R package, MigConnectivity, includes functions to calculate r_M and MC and incorporate error from multiple sources.

2 | MATERIALS AND METHODS

2.1 | Calculating migratory connectivity

Whereas r_M is the correlation between distance matrices of individuals encountered during one season of the annual cycle and re-encountered during another season (Ambrosini et al., 2009; Besag & Diggle, 1977), MC is an approximation of this correlation using discrete populations (Data S1). The population-specific transition probabilities used by MC can be defined from any available data type and are directional, with an origin and a destination season. For simplicity, we only include MC from breeding to non-breeding but the metric is equally applicable between other seasons. Because MC makes use of transition probabilities, individuals are divided into a priori defined populations and, unlike r_M , distances among individuals within those regions are not incorporated. MC is ideally defined at the species level, with data from all populations, but can also be measured when data are available for subsets of seasonal ranges that represent regional populations.

$$MC = \frac{\sum_{i=1}^B \sum_{j=1}^B \sum_{k=1}^{NB} \sum_{l=1}^{NB} [M_{ijkl} (D_{ij} - \mu_D) (V_{kl} - \mu_V)]}{\sigma_D \sigma_V} \quad (1)$$

where B is the number of breeding regions, NB is the number of non-breeding regions, D_{ij} is the distance between breeding region i and j , V_{kl} is the distance between non-breeding region k and l , μ_D and μ_V are the mean distances between individuals in the breeding and non-breeding regions, respectively, σ_D and σ_V are the standard deviations of these distances, and M_{ijkl} is the product of the proportions of individuals transitioning between i and k and between j and l , corrected to not include an individual and itself

$$M_{ijkl} = R_i \psi_{ik} \frac{R_j n \psi_{jl} - I(i=j \& k=l)}{n-1} \quad (2)$$

where R_i is the relative abundance at breeding region i , ψ_{ik} is the probability that a bird breeding at region i transitions to non-breeding region k , n is the total sample size (number of animals providing transition probabilities), and I is the indicator function. Therefore, the data needed to calculate MC (without incorporating uncertainty) are the number of and distance between the regions within the destination and origin ranges, the relative abundance among origin regions, the transition probabilities from origin to destination regions, and the sample size from which transition probabilities were estimated.

The means and standard deviations of the distances between breeding (D) and non-breeding (V) individuals can be calculated as

$$\begin{aligned} \mu_D &= \sum_{i=1}^B \sum_{j=1}^B R_i \frac{R_j n - I(i=j)}{n-1} D_{ij} \\ \mu_V &= \sum_{k=1}^{NB} \sum_{l=1}^{NB} \left[V_{kl} \sum_{i=1}^B \sum_{j=1}^B M_{ijkl} \right] \\ \sigma_D &= \sqrt{\sum_{i=1}^B \sum_{j=1}^B R_i \frac{R_j n - I(i=j)}{n-1} (D_{ij} - \mu_D)^2} \\ \sigma_V &= \sqrt{\sum_{k=1}^{NB} \sum_{l=1}^{NB} \left[(V_{kl} - \mu_V)^2 \sum_{i=1}^B \sum_{j=1}^B M_{ijkl} \right]} \end{aligned} \quad (3)$$

Negative values of MC indicate that individuals close to each other in one season are further apart in the other season. If $MC = 0$, no relationship exists between distances in one season and another; if $MC = 1$, the relative distances between individuals in one season are the same in the other, although the scale can differ. MC can be calculated in the calcMC function available in the MigConnectivity R package.

2.2 | Estimating migratory connectivity

Armed with records of animals observed in both breeding and non-breeding regions, one is tempted to simply compute the transition probabilities for Equations 1–3 as

$$\psi_{ik} = \frac{C_{ik}}{\sum_{l=1}^{NB} C_{il}} \quad (4)$$

where C_{ik} is the number of animals observed in breeding location i and non-breeding location k . However, this ignores the sampling uncertainty in those transition probabilities. Incorporating uncertainty in input values (particularly in transition probabilities and relative abundances) into MC is a key component of any assessment of migratory connectivity strength. Sources of sampling error include limited sample sizes, detection heterogeneity (detection probabilities varying between regions), and location error. Because studies and data types vary in sampling error, valid quantitative comparisons require that they be incorporated into measures of migratory connectivity strength. Therefore, we developed a series of resampling approaches for propagating uncertainty into MC and estimating confidence intervals and/or standard errors measuring this uncertainty, dependent on the types of data used to estimate the input parameters.

Procedures to estimate MC with uncertainty associated with input values are implemented in the estMC function available in the MigConnectivity R package. Process errors due to natural variability in animal movement also increase uncertainty around transition probability estimates and, therefore, influence measurement and definition of MC (see Data S1 for further details). We describe the resampling approaches available with estMC function below. We recognize that these do not encompass all of the uncertainty associated with data used to measure MC, but expect that they are broadly applicable to many studies and data types.

An estimate of sampling uncertainty in transition probabilities (e.g., from a multistate mark-reencounter model; Cohen et al., 2014), can be

propagated into an estimate of uncertainty in MC by resampling a large number of times (S) from the distribution for transition probabilities and computing MC for each iteration

$$MC^{s*} = \frac{\sum_{i=1}^B \sum_{j=1}^B \sum_{k=1}^{NB} \sum_{l=1}^{NB} \left[M_{ijkl}^{s*} (D_{ij} - \mu_D) (V_{kl} - \mu_V^{s*}) \right]}{\sigma_D \sigma_V^{s*}} \quad (5)$$

$$M_{ijkl}^{s*} = R_i \Psi_{ik}^{s*} \frac{R_j n \Psi_{jl}^{s*} - I (i=j \& k=l)}{n-1}$$

where Ψ_{ik}^{s*} is the s th sample for the transition probability between i and k , MC^{s*} is the s th sample for MC, and μ_V^{s*} and σ_V^{s*} are also calculated by substituting M_{ijkl}^{s*} for M_{ijkl} . MC can then be estimated as the mean of the samples and its standard error as their SD

$$\widehat{MC} = \sum_{s=1}^S MC^{s*} / S$$

$$\widehat{se}(MC) = sd(MC^{s*}) \quad (6)$$

Simple percentile-based confidence intervals for MC can be estimated as

$$\left[\widehat{MC}_{low(\%)}, \widehat{MC}_{high(\%)} \right] = \left[MC^{*(\alpha/2)}, MC^{*(1-\alpha/2)} \right] \quad (7)$$

where a $1 - \alpha$ confidence interval is desired and $MC^{*(\alpha/2)}$ is the $\alpha/2$ th quantile of the MC samples. This confidence interval is biased when the median of the samples is not the same as the estimated value (in this case, the mean; Efron & Tibshirani, 1994; Ellner & Fieberg, 2003).

A bias-corrected (BC) confidence interval can be estimated as

$$\left[\widehat{MC}_{low(BC)}, \widehat{MC}_{high(BC)} \right] = \left[MC^{*(\alpha_1)}, MC^{*(\alpha_2)} \right]$$

$$\alpha_1 = \Phi \left(2\hat{z}_0 + \Phi^{-1} (r\alpha/2) \right) \quad \alpha_2 = \Phi \left(2\hat{z}_0 + \Phi^{-1} (1 - \alpha/2) \right) \quad (8)$$

$$\hat{z}_0 = \Phi^{-1} \left(\frac{\#\{MC^{s*} < \widehat{MC}\}}{S} \right)$$

where Φ is the standard normal cumulative distribution function, Φ^{-1} is its inverse (the standard normal quantile function), \hat{z}_0 is a bias-correction term based on the proportion of samples less than the mean, and $\#$ indicates the number of samples.

When sampling uncertainty in the relative abundances per breeding location has also been estimated (generally separately from transition probabilities), Equation 5 can be extended to include samples from that distribution

$$MC^{s*} = \frac{\sum_{i=1}^B \sum_{j=1}^B \sum_{k=1}^{NB} \sum_{l=1}^{NB} \left[M_{ijkl}^{s*} (D_{ij} - \mu_D^{s*}) (V_{kl} - \mu_V^{s*}) \right]}{\sigma_D^{s*} \sigma_V^{s*}} \quad (9)$$

$$M_{ijkl}^{s*} = R_i^{s*} \Psi_{ik}^{s*} \frac{R_j^{s*} n \Psi_{jl}^{s*} - I (i=j \& k=l)}{n-1}$$

where R_i^{s*} is the s th sample from the distribution for the relative abundance of animals in breeding site i . Equations 6–8 can be applied in the same way to these samples.

If detection heterogeneity is not an issue, a separate model may not be needed to estimate the uncertainty in transition probabilities. However, when location error and limited sample sizes exist, a bootstrapping approach can be used

$$loc^{s*} = (loc_1^*, loc_2^*, \dots, loc_n^*) \quad (10)$$

where loc_r^* is the breeding and non-breeding coordinates (lat-long or others) for the r th animal drawn with replacement from the n animals monitored and loc^{s*} is the s th bootstrap sample. When there are estimates of bias and variance in the non-breeding coordinates, they should be incorporated in the bootstrap sampling:

$$loc_r^* \sim N_4 \left(loc_q + \begin{bmatrix} 0 \\ 0 \\ \text{bias}(x_{NB}) \\ \text{bias}(y_{NB}) \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \text{var}(x_{NB}) & \text{cov}(x_{NB}, y_{NB}) \\ 0 & 0 & \text{cov}(x_{NB}, y_{NB}) & \text{var}(y_{NB}) \end{bmatrix} \right) \quad (11)$$

where the r th bootstrap animal is the q th animal in the original dataset, loc_q are its four coordinates, N_4 is the quadrivariate normal distribution, and the bias, variance, and covariance of coordinates on the non-breeding range are estimated separately (breeding location error can be applied in a similar way). With or without location error, the sampled coordinates are assigned to the discrete breeding and non-breeding locations, censoring those that fall outside these regions. Bootstrap samples that do not include all breeding locations should be replaced. The transition probabilities for the s th bootstrap sample can be computed by modifying Equation 4

$$\Psi_{ik}^{s*} = \frac{C_{ik}^{s*}}{\sum_{l=1}^{NB} C_{il}^{s*}} \quad (12)$$

where C_{ik}^{s*} is the number of animals observed in breeding location i and non-breeding location k in the s th bootstrap sample. Equations 5–8 can then be applied to estimate MC and its uncertainty. In addition, estimates of r_M and its uncertainty can be obtained by using the sampled coordinates (Equation 10) directly in distance matrices (see Manly, 1991 for another approach to r_M confidence intervals).

Approximate p -values for the one-sided test $MC > 0$ can be calculated using the logic of percentile based confidence intervals

$$p_{(\%)} = \sum_{s=1}^S I(MC^{s*} < 0) / S \quad (13)$$

and approximate bias-corrected p -values can be calculated as

$$p_{(BC)} = \Phi \left(\Phi^{-1} (p_{(\%)}) - 2\hat{z}_0 \right) \quad (14)$$

However, presenting means and standard errors or confidence intervals are probably more informative than p -values (Anderson, Link, Johnson, & Burnham, 2001; Efron & Tibshirani, 1994).

2.3 | Sensitivity to real-world values

We used calcMC with available information about possible real-world input values to explore the range of attainable MC values in response to (1) transition probabilities, (2) distances between regions, and (3) relative abundance among regions. We used North American Bird Banding Laboratory ringing and long-distance re-encounter data, which houses over 4.5 million re-encounter records from 1914 to the present (www.pwrc.usgs.gov/bbl), to construct eight probable real-world patterns of migratory connectivity (Figure 1), including

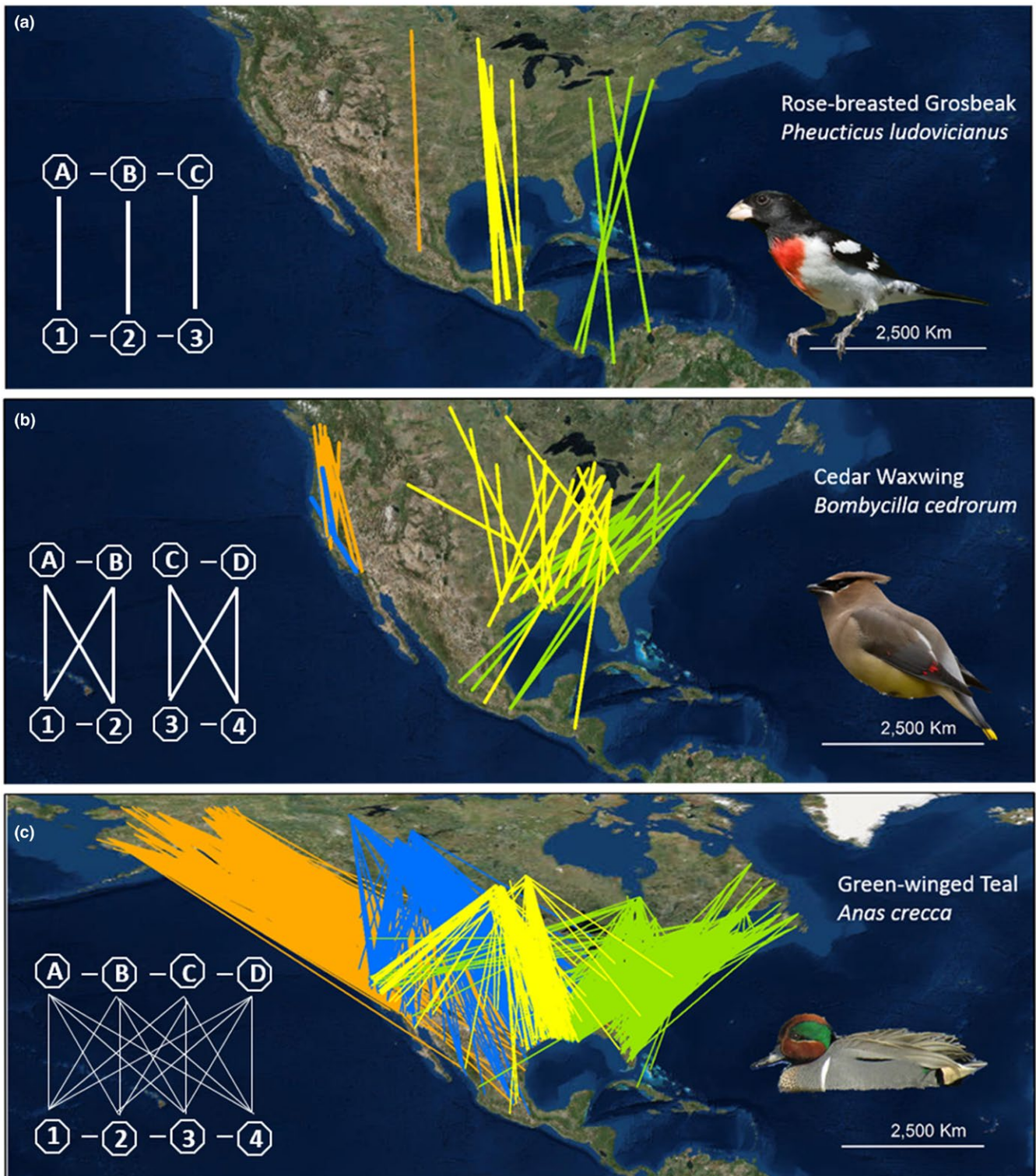


FIGURE 1 Migratory movement (transition probability) scenarios derived from North American bird banding laboratory ringing and re-encounter data (www.pwrc.usgs.gov/bbl) to reflect real-world patterns, including (a) complete connectivity between one breeding and one non-breeding region, (b) intermediate connectivity with full mixing among regions on either side of a divide, and (c) full mixing among all regions. See Figure S1 for the eight transition probability scenarios included in sensitivity analyses

complete connectivity between one breeding and one non-breeding region, full mixing among all regions, and a range of intermediate scenarios (Figure S1). We explored the range of MC values in response to

the eight transition probability scenarios, seven scenarios for the spatial arrangements regions, and five scenarios for relative abundance among regions.

2.4 | Influence of sampling error

Biased or incomplete sampling can introduce error into the measurement of MC that should be accounted for in comparisons among studies. For example, species abundance is unlikely to be even across a seasonal range. Consequently, when transition probabilities are measured among populations across a range, sampling may not be proportional to the relative abundance of those regions. Further, estimates of MC may also be influenced by the accuracy of devices used to track animals between seasons. We therefore, used calcMC and estMC functions to measure the influence of many potential sources of sampling error on MC and r_M . Simulations were used to measure the influence of incorrect delineation of populations into regions, sampling not proportional to abundance, heterogeneity in detection among regions (e.g., from mark re-encounter data), uncertainty in estimates of breeding abundance, small sample sizes, and location uncertainty on estimation of MC and r_M . We used simulations with the estMC function to explore the influence of sampling error (due to limited sample size, location error, detection heterogeneity, or uncertainty in relative abundance) on estimates of the expected value of MC , its variance, and its confidence intervals (see Data S1 for further details). We also used real tracking data from a temperate – tropical migratory passerine, the Ovenbird (*Seiurus aurocapilla*), collected with light-level and GPS geolocation (Hallworth & Marra, 2015; Hallworth et al., 2015) and a resampling technique was used to measure the influence of small sample size and location uncertainty (e.g., from tracking data) on MC and to compare MC to r_M (see Data S1 for methods).

2.5 | Comparison of r_M and MC

We define and explore the relationships between the true strength of migratory connectivity and various approximations to it, including r_M , MC , and the large sample size asymptotic version of MC . Based on the data types available, we make recommendations on the best approach to estimate the strength of migratory connectivity.

3 | RESULTS

3.1 | Sensitivity to real-world values

Migratory connectivity ranged from -0.08 to 1.0 for the eight transition probability scenarios (Figure S1). MC was most influenced by the transition probabilities between regions, reflective of the correlations of distances between individuals. The influence of the spatial arrangement and relative abundance of regions depended on the transition probability matrix. When transitions from each breeding region were broadly distributed across non-breeding regions (i.e., weak connectivity), the distances between those regions did not strongly influence MC but when transitions from each breeding region were primarily to one or two non-breeding regions (i.e., strong connectivity), the influence of distance on MC was greater (Figure 2). The spatial arrangement of regions had the greatest influence on MC when the difference between the breeding and non-breeding region arrangements were the greatest (Figure 2 scenario

F). Relative abundance was less influential than transition probabilities and spatial arrangements (Figure 3).

3.2 | Influence of sampling error

We used simulations with the calcMC function to explore the influence of sampling error on MC , including the number and arrangement of regions designated, sampling out of proportion to abundance, and sampling when migratory connectivity varies across the range (Tables 1 and 2). MC was biased low when breeding populations were grouped into fewer regions than sampling locations and when breeding or non-breeding regions were not compact (10 strips rather than 4 quadrats), even with equal abundance among breeding regions (Table 1). MC contained the greatest bias when sampling occurred in all populations and samples were grouped into fewer regions (Tables 1 and 2), although this bias was increased further by variable strength of migratory connectivity. When region definitions matched sampling locations, MC was only slightly biased by sampling effort not proportional to abundance and migratory connectivity that varied across the range (Table 2). r_M was not affected by grouping but was strongly biased when sampling effort was not proportional to abundance and when migratory connectivity was variable across the range (Table 2).

We found that MC error due to detection heterogeneity in long-distance mark and re-encounter data or uncertainty in relative abundance was negligible (Data S1). In our simulation of GPS tracking data, error in MC and r_M were also negligible (Table S2; Figure S2). Error was somewhat higher for our estimates of MC and r_M from simulated light-level geolocator data, but still minor and considerably less than the uncorrected point estimates using Equation 4 (Data S1, Table S2). Results from Ovenbird tracking were similar with GPS and light-level geolocation levels of location uncertainty; MC values were consistently lower than r_M values (Data S1, Figure S3).

3.3 | Comparison of r_M and MC

Under ideal conditions without grouping error and with sampling proportional to relative abundances, r_M and MC are identical (Data S1). The true strength of migratory connectivity, when the true locations of all individuals are known between seasons and animal locations are not grouped, is identical to MC calculated with absolute abundance (instead of sample size). A version of MC that does not account for relative abundances is also identical to r_M when animal locations are not grouped (Data S1). From simulated data, we found MC is more robust to bias issues than its other variants, but that there are conditions where r_M is preferred (Data S1, Table S2). We provide a decision tree (Figure 4) to guide researchers about how to calculate or estimate the strength of migratory connectivity, given available data.

4 | DISCUSSION

Populations with strong migratory connectivity may be adapted to local environments and at greater risk from localized threats, while

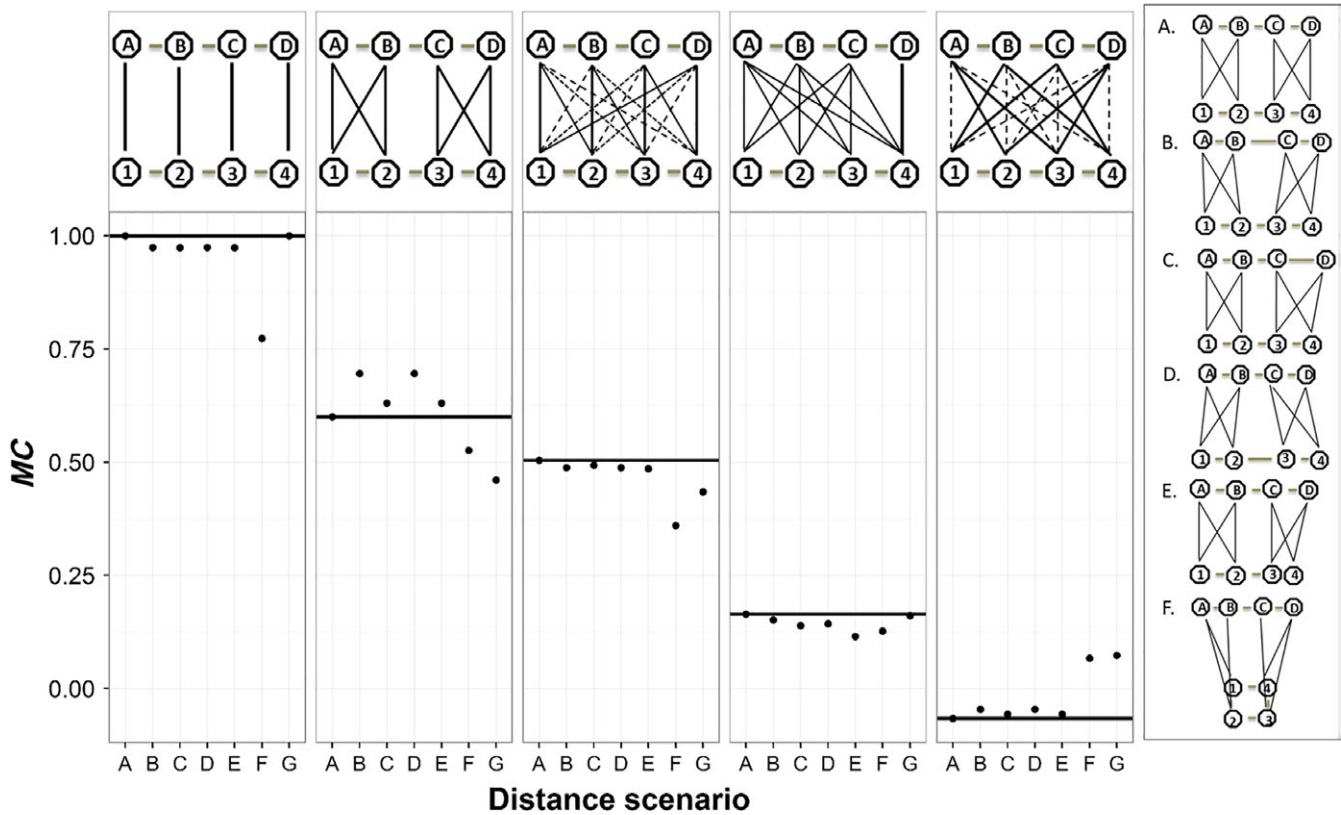


FIGURE 2 Influence of the spatial arrangement of populations into regions on migratory connectivity (MC) (calculated assuming sample sizes are large). Sensitivity scenarios varied in transition probabilities (columns; Figure S1) and the spatial arrangement of breeding and non-breeding populations into regions (distance scenarios). Bold lines indicate the value when breeding regions and non-breeding regions were arranged linearly (scenario A). Additional spatial arrangements are linear and with the distance between the regions doubled or halved (B)–(E), linear and gridded (F), or gridded and gridded (G, not pictured)

populations with weak migratory connectivity may be adapted to environments that occur across a larger geographic area and be more buffered from localized threats (Ambrosini et al., 2016; Cresswell, 2014). Therefore, the strength of migratory connectivity is a demographic metric, like age ratios, fertility rates, and genetic relatedness (e.g., F_{ST} ; Meirans & Hedrick, 2011), with profound consequences for populations and species. Yet, the role of migratory connectivity strength in evolution, ecology, and conservation has not been well studied, despite advancing analytic, molecular, and technological tools for measuring animal movement throughout the annual cycle.

Our objective here is to facilitate quantitative exploration of the strength of migratory connectivity with methodology that is comparable across species, studies, and data types. For example, most studies have not sampled migratory animals across the full extent of a species' range, nor have they sampled in proportion to regional abundance. Inferences about migratory connectivity strength derived from methods that do not account for incomplete sampling (e.g., Finch et al., 2017) may not be robust. The approach we present here explicitly addresses uneven sampling and incorporates uncertainty in estimates of migratory connectivity strength from multiple sources of sampling and process error.

We define the ideal properties of a metric to quantify the true strength of migratory connectivity, when locations of all individuals of a species are known between seasons, and define and explore relationships among

approximations to the true strength (Data S1). The most commonly used method to quantify migratory connectivity strength, r_M , makes use of a sample of individual animals located during two seasons (Besag & Diggle, 1977). MC builds on this method by incorporating population-specific transition probabilities that can be derived from many data types and accounts for uneven sampling. MC approximates r_M , which is a correlation coefficient with a range from -1 to 1 . MC may also qualify as a correlation coefficient, at least in some circumstances. In sensitivity analyses based on a range of probable real-world scenarios, MC ranged from -0.08 to 1.0 , but it can almost certainly go lower in less realistic scenarios. The mathematical and statistical properties of MC , including expected value, variance, and range, are yet to be analytically explored. Our simulation studies, however, show that our resampling approaches do a good job of estimating expected value, variance, and confidence intervals for tracking and capture mark recapture data.

Our simulations demonstrate that failure to account for small sample sizes, uneven sampling among populations, sampling across the extent of a species' range, grouping of individuals into populations, or accuracy and precision of locations derived from different data types, likely bias calculation of the strength of migratory connectivity. Sampling regime and natural variability in movement (see Data S1, Process error) also influenced estimates. Therefore, we recommend estimating and reporting uncertainty from appropriate sources of

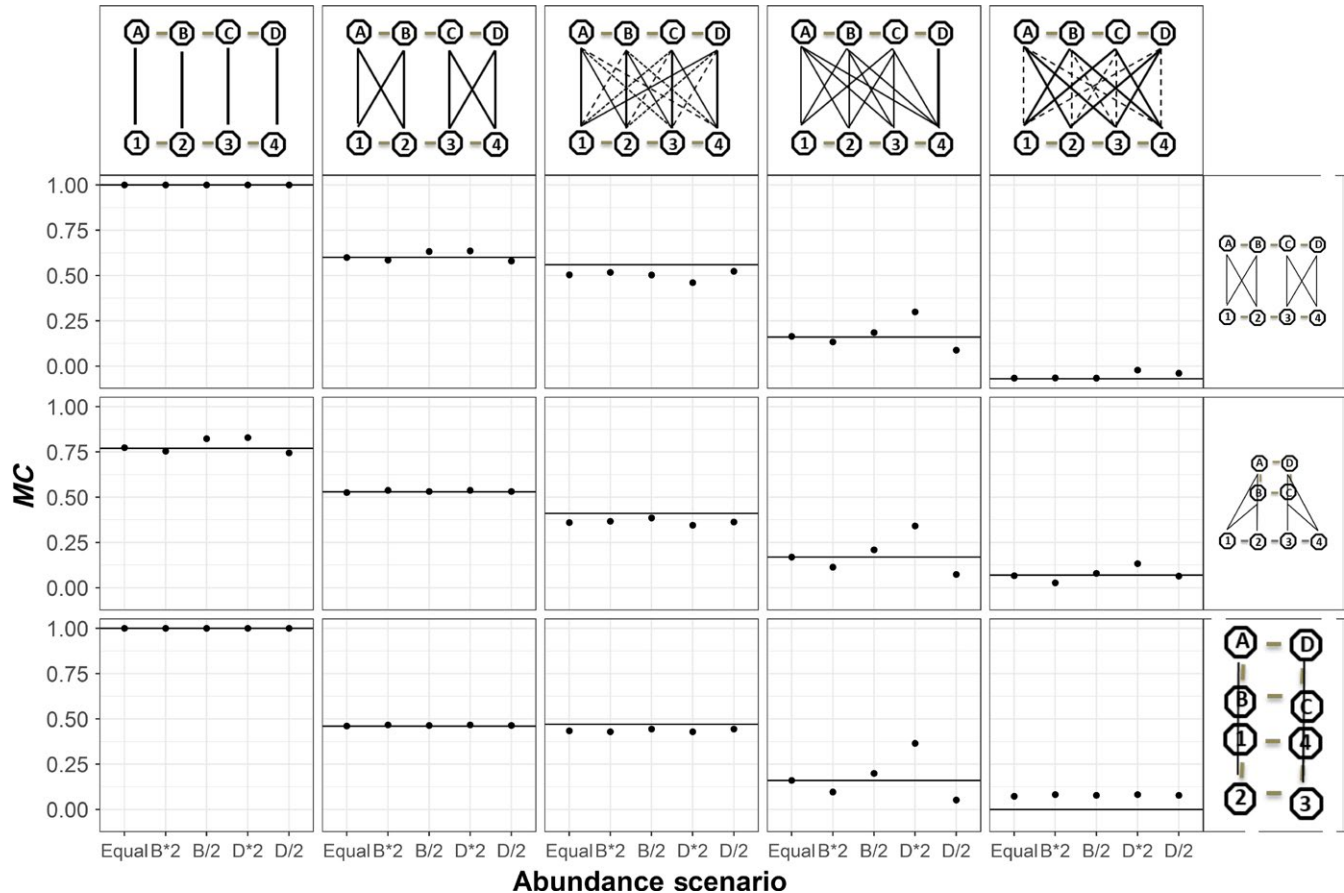


FIGURE 3 Influence of the relative abundance among regions on migratory connectivity (MC) (calculated assuming sample sizes are large). Sensitivity scenarios varied in transition probabilities between breeding and non-breeding regions (columns; Figure S1), the spatial arrangement of breeding and non-breeding regions (rows, linear - linear, grid - linear, and grid - grid), and the relative abundance between breeding regions (Abundance Scenario). In five abundance scenarios, the relative abundance among breeding regions was equal (bold line), doubled (D^*2 , B^*2), or halved ($B/2$, $D/2$), in each of two regions

sampling and process errors when possible. For example, estimates of MC from tracking studies should report uncertainty due to location error while long-distance marking and re-encounter studies should report uncertainty due to geographic variability in detection.

Migratory connectivity measures the connections of population with probabilities of movement between regions, but does not account for the distribution of populations within those regions. MC will be most accurate when there are no patterns within regions and r_M will be stronger than MC when individuals that share breeding and non-breeding regions maintain the same positions, relative to each other, within regions (Ambrosini et al., 2009). For example, r_M was higher than MC for our Ovenbird tracking dataset, suggesting that individuals in connected populations maintain the same relative positions within regions between the breeding and non-breeding seasons. MC performed better than r_M in simulations when sampling effort was not proportional to abundance and the strength of migratory connectivity varied across the range. Therefore, estimates of r_M may be appropriate when estimates of abundance are not available, movement data are for individual animals that are not naturally clustered into regions, and detection and location uncertainty are low (Figure 4).

The designation of geographical regions for the measurement of transition probabilities is necessary for MC and a key difference between MC and r_M . Artificially grouping animals into regions had the greatest influence on MC when the strength of migratory connections was region-specific, a likely scenario for many species (e.g. Cohen et al., 2014; Trierweiler et al., 2014). In practice, unless geography or habitat fragmentation create “natural” population delineations, identifying discrete populations for geographic sampling is only possible for species with range-wide demographic data (Rushing, Ryder, Scarpignato, Saracco, & Marra, 2015). We used a priori portioning of population structure, as opposed to inferring structure in a post hoc way. Most migratory connectivity studies select sampling locations in this way, usually for logistical reasons, rather than based on known information about population structure (Data S1). We found that bias in MC decreased with more compact, defined regions and with sampling at centroids defined by ringing stations or by locations where tracking devices are deployed. Measuring MC from a priori selected regions also enables hypothesis tests about the role of migratory connectivity strength in population dynamics. For example, we hypothesize that stationary non-breeding conditions shape breeding population trends when migratory connectivity is strong and the non-breeding

TABLE 1 Mean MC difference from a true value of 0.25 when breeding ($n = 100$) and non-breeding ($n = 100$) populations were grouped into fewer sampling regions ($n = 4$ arranged in a grid or $n = 10$ arranged in a row). Sampling effort ($n = 1000$) was either even across the breeding range or at the centroids of the sampling regions. Abundance was even among regions and 100 simulations were run for each scenario

Breeding regions	Non-breeding regions	Sampling locations	MC	ΔMC
100	100	100	0.25	0.00
100	10	100	0.10	-0.15
100	4	100	0.18	-0.07
10	100	100	0.13	-0.12
10	10	100	0.18	-0.07
10	4	100	0.09	-0.16
4	100	100	0.18	-0.07
4	10	100	0.07	-0.18
4	4	100	0.18	-0.07
10	100	10	0.15	-0.10
10	10	10	0.18	-0.07
10	4	10	0.09	-0.16
4	100	4	0.22	-0.03
4	10	4	0.09	-0.16
4	4	4	0.22	-0.03

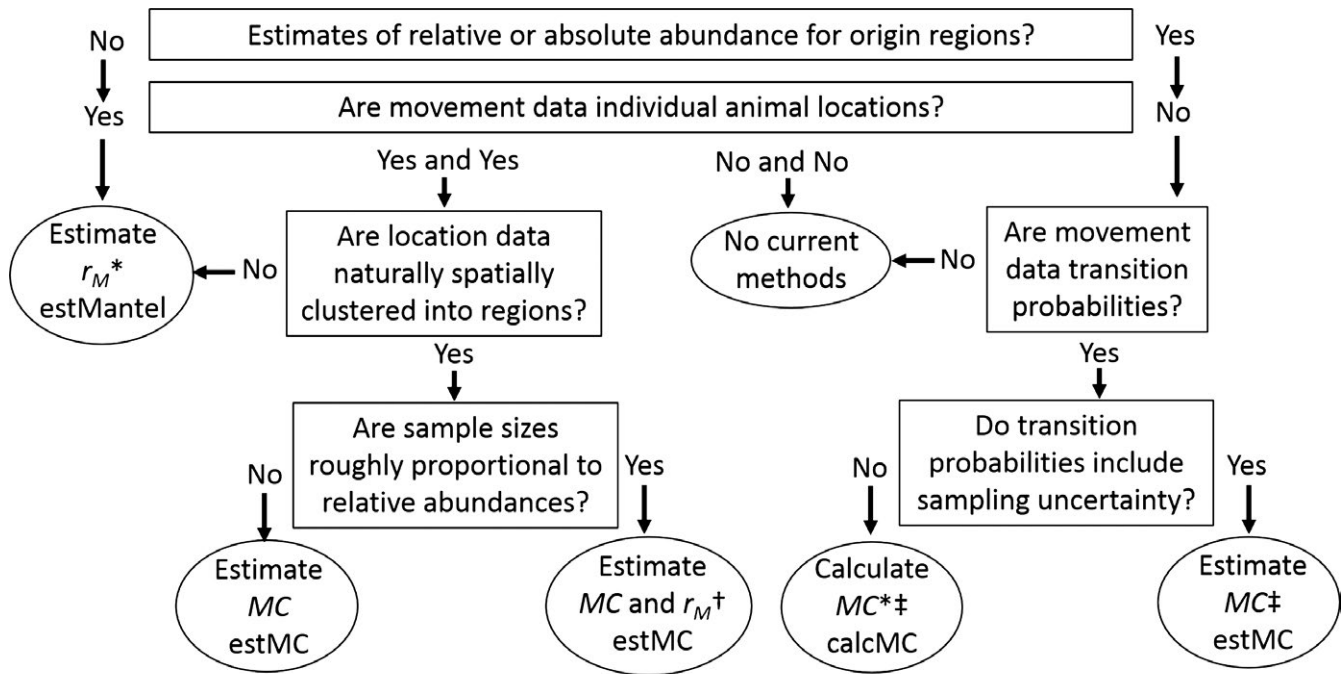
environment is locally variable (Cresswell, 2014). Methods to identify populations (Ambrosini et al., 2009; Rushing, Ryder, et al., 2015b), could be used in parallel with MC to generate testable hypotheses about the demographic consequences of migratory connectivity.

TABLE 2 Mean differences in MC and r_M when breeding populations ($n = 100$) were grouped into regions ($n = 4$) with uneven relative abundance. Sampling effort (total $n = 1,000$) was either positively (higher in high abundance regions and lower in low abundance regions) or negatively (lower in high abundance regions and higher in low abundance regions) proportional to abundance. Migratory connectivity strength was either even throughout the breeding range (range-wide MC = 0.25) or varied (range-wide MC = 0.47; =0.60, 0.30, 0.45, and 0.15 in each quadrat). Sampling occurred in each population ($n = 100$, 25 locations per region) or at centroids of regions ($n = 4$, 1 location per region). Non-breeding populations ($n = 100$) were not grouped and 100 simulations were run for each scenario

Sampling effort proportional	Variable migratory connectivity	Breeding regions	Number of sampling locations	MC	ΔMC	r_M	Δr_M
Yes	No	100	100	0.25	0.00	0.25	0.00
Yes	No	4	100	0.19	-0.06	0.25	0.00
Yes	No	4	4	0.25	0.00	0.25	0.00
Yes	Yes	100	100	0.47	0.00	0.47	0.00
Yes	Yes	4	100	0.37	-0.10	0.47	0.00
Yes	Yes	4	4	0.50	0.03	0.50	0.03
No	No	100	100	0.28	0.03	0.24	0.01
No	No	4	100	0.20	-0.05	0.24	0.01
No	No	4	4	0.26	0.01	0.24	-0.01
No	Yes	100	100	0.48	0.01	0.25	-0.23
No	Yes	4	100	0.35	-0.12	0.25	-0.23
No	Yes	4	4	0.50	0.03	0.23	-0.24

Further research is needed to understand how MC is influenced by sources of process error such as long-distance dispersal rates, annual and within-season environmental variation, and individual heterogeneity. MC assumes fidelity to regions within seasonal ranges, which could be violated by long-distance dispersal and other sources of variation in movement within and among seasons. Systematic, long-distance dispersal in particular would lower MC over time. We found that relatively small differences in long-distance dispersal rates had a large impact on the rate at which MC declined to zero (Data S1, Process error and Figure S4). The effect of long-distance dispersal on MC will depend on the size and arrangement of regions as well as the magnitude and frequency of dispersal movements, information that is poorly known for most species (Clobert, Baguette, Benton, Bullock, & Ducatez, 2012; Paradis, Baillie, Sutherland, & Gregory, 1998). In addition, movement of populations could vary annually due to evolution (Delmore, Kenyon, Germain, & Irwin, 2015), climate (Rushing, Dudash, Studds, & Marra, 2015), or habitat (Taylor & Stutchbury, 2015), further increasing uncertainty in measurement of migratory connectivity strength. Individual and population heterogeneity in annual movement may also be important sources of uncertainty in MC. However, the data necessary to measure individual heterogeneity—movement of many individuals for many years—has not been practical. As technological advances make these data available, individual heterogeneity could be incorporated into MC as covariates (e.g., estimated separately by age or sex).

Studies of the strength of migratory connectivity are applicable to migratory periods as well as multiple non-breeding areas. Although most migratory connectivity studies have measured connections between breeding and non-breeding stationary seasons of the annual cycle, including migration will become practical as finer-scale spatial and temporal data become available. Moreover, the extent to which populations



* May be biased if detection or location uncertainty are high

† Estimates likely to be similar if detection heterogeneity is low

‡ May be biased if locations were artificially clustered into regions

FIGURE 4 Decision support tree for use of R package MigConnectivity, given available data

co-occur outside of breeding depends in part on the diversity of migratory strategies within breeding populations (Gilroy, Gill, Butchart, Jones, & Franco, 2016). Breeding populations may use the same spring and autumn migration routes each year, but at different times of the season (Bauer et al., 2015; Briedis et al., 2016; Paxton & Moore, 2015). Demographic consequences could result if this variation causes populations to encounter different conditions during passage (Drent et al., 2007). Therefore, *en route* migratory connectivity has a spatial and a temporal component (Bauer et al., 2015). Weak MC would occur when the same breeding populations have more diverse timing, routes, and destinations. Further, some migratory species may move considerable distances during winter months and remain for extended periods of time within more than one region (Heckscher, Taylor, Fox, & Afanasyev, 2011; Renfrew et al., 2013; Rohwer, Hobson, & Rohwer, 2009). Increased information about the migratory movements will therefore require the consideration of multiple regions during the “stationary” non-breeding season. The locations of individuals within populations would ideally be measured continuously in space and time, making it possible to estimate how migratory connectivity changes throughout an annual cycle. The advancement of technological and analytical tools will facilitate finer-scale measurement of MC.

In conclusion, MC can reduce bias from available data types, but accurate estimates of migratory connectivity strength are ultimately limited by uncertainty about the seasonal distribution of populations. Advances in tracking methodology will reduce this uncertainty and enable us to integrate multiple, high resolution data sources, such as miniature GPS tags and genomics. We hope the MigConnectivity package, MC metric, and exploration of available

methods presented here provide a robust framework to advance migratory connectivity research.

ACKNOWLEDGEMENTS

Hormuzd Katki, Jim Nichols, Colin Studds, and Brandt Ryder provided helpful comments during the development of the MC metric. Roberto Ambrosini, Nathan Cooper, Autumn-Lynn Harrison, and two anonymous reviewers provided helpful comments on earlier drafts of the manuscript. Amy Scarpignato made Figure 1. The U.S. Fish and Wildlife Service, ConocoPhillips Corporation, the U.S. National Science Foundation, and a James Smithson Fellowship to CSR provided support for this research.

AUTHOR'S CONTRIBUTIONS

E.B.C and J.A.H. wrote the first draft of the paper. J.A.H. and M.T.H. wrote the R package with contributions from E.B.C and C.S.R. All authors conceived the ideas, designed methodology, and contributed to ideas and critically reviewed and edited the manuscript and the package and gave final approval for publication.

DATA ACCESSIBILITY

R package MigConnectivity with vignette, code, and data associated with this paper are available <https://doi.org/10.5281/zenodo.1002229> (Hostetler, & Hallworth, 2017).

Ovenbird tracking data available at Movebank Data Repository <https://doi.org/10.5441/001/1.h43t9j6b> (Hallworth & Marra, 2015).

ORCID

Emily B. Cohen  <http://orcid.org/0000-0002-4978-4278>

REFERENCES

- Ambrosini, R., Cuervo, J. J., Feu, C., Fiedler, W., Musitelli, F., Rubolini, D., ... Møller, A. P. (2016). Migratory connectivity and effects of winter temperatures on migratory behaviour of the European robin *Erithacus rubecula*: A continent-wide analysis. *Journal of Animal Ecology*, *85*, 749–760. <https://doi.org/10.1111/1365-2656.12497>
- Ambrosini, R., Møller, A. P., & Saino, N. (2009). A quantitative measure of migratory connectivity. *Journal of Theoretical Biology*, *257*, 203–211. <https://doi.org/10.1016/j.jtbi.2008.11.019>
- Anderson, D. R., Link, W. A., Johnson, D. H., & Burnham, K. P. (2001). Suggestions for presenting the results of data analyses. *The Journal of Wildlife Management*, *65*, 373–378. <https://doi.org/10.2307/3803088>
- Bauer, S., Lisovski, S., & Hahn, S. (2015). Timing is crucial for consequences of migratory connectivity. *Oikos*, *125*, 605–612. <https://doi.org/10.1111/oik.02706>
- Besag, J., & Diggle, P. J. (1977). Simple Monte Carlo tests for spatial pattern. *Applied Statistics*, *26*, 327–333. <https://doi.org/10.2307/2346974>
- Bridge, E. S., Thorup, K., Bowlin, M. S., Chilson, P. B., Diehl, R. H., Fléron, R. W., ... Wikelski, M. (2011). Technology on the move: Recent and forthcoming innovations for tracking migratory birds. *BioScience*, *61*, 689–698. <https://doi.org/10.1525/bio.2011.61.9.7>
- Briedis, M., Hahn, S., Gustafsson, L., Henshaw, I., Träff, J., Král, M., & Adamik, P. (2016). Breeding latitude leads to different temporal but not spatial organization of the annual cycle in a long-distance migrant. *Journal of Avian Biology*, *47*, 743–748. <https://doi.org/10.1111/jav.2016.v47.i6>
- Clobert, J., Baguette, M., Benton, T. G., Bullock, J. M., & Ducatez, S. (2012). *Dispersal ecology and evolution*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199608898.001.0001>
- Cohen, E. B., Hostetler, J. A., Royle, J. A., & Marra, P. P. (2014). Estimating migratory connectivity of birds when re-encounter probabilities are heterogeneous. *Ecology and Evolution*, *4*, 1659–1670. <https://doi.org/10.1002/ece3.2014.4.issue-9>
- Cresswell, W. (2014). Migratory connectivity of Palaearctic-African migratory birds and their responses to environmental change: The serial residency hypothesis. *Ibis*, *156*, 493–510. <https://doi.org/10.1111/ibi.12168>
- Delmore, K. E., Kenyon, H. L., Germain, R. R., & Irwin, D. E. (2015). Phenotypic divergence during speciation is inversely associated with differences in seasonal migration. *Proceedings of the Royal Society B: Biological Sciences*, *282*, 20151921. <https://doi.org/10.1098/rspb.2015.1921>
- Drent, R. H., Eichhorn, G., Flagstad, A., Van der Graaf, A. J., Litvin, K. E., & Stahl, J. (2007). Migratory connectivity in Arctic geese: Spring stopovers are the weak links in meeting targets for breeding. *Journal of Ornithology*, *148*, 501–514. <https://doi.org/10.1007/s10336-007-0223-4>
- Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. Boca Raton, FL: CRC Press.
- Ellner, S. P., & Fieberg, J. (2003). Using PVA for management despite uncertainty: Effects of habitat, hatcheries, and harvest on salmon. *Ecology*, *84*, 1359–1369. [https://doi.org/10.1890/0012-9658\(2003\)084\[1359:UPFMDU\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2003)084[1359:UPFMDU]2.0.CO;2)
- Esler, D. (2000). Applying metapopulation theory to conservation of migratory birds. *Conservation Biology*, *14*, 366–372. <https://doi.org/10.1046/j.1523-1739.2000.98147.x>
- Finch, T., Butler, S. J., Franco, A., & Cresswell, W. (2017). Low migratory connectivity is common in long-distance migrant birds. *Journal of Animal Ecology*, *86*, 662–673. <https://doi.org/10.1111/1365-2656.12635>
- Gilroy, J. J., Gill, J. A., Butchart, S. H., Jones, V. R., & Franco, A. (2016). Migratory diversity predicts population declines in birds. *Ecology Letters*, *19*, 308–317. <https://doi.org/10.1111/ele.2016.19.issue-3>
- Hallworth, M. T., & Marra, P. P. (2015). Miniaturized GPS tags identify non-breeding territories of a small breeding migratory songbird. *Scientific Reports*, *5*, 11069. <https://doi.org/10.1038/srep11069>
- Hallworth, M. T., Sillett, T. S., Van Wilgenburg, S. L., Hobson, K. A., & Marra, P. P. (2015). Migratory connectivity of a Neotropical migratory songbird revealed by archival light-level geolocators. *Ecological Applications*, *25*, 336–347. <https://doi.org/10.1890/14-0195.1>
- Heckscher, C. M., Taylor, S. M., Fox, J. W., & Afanasyev, V. (2011). Veery (*Catharus fuscescens*) wintering locations, migratory connectivity, and a revision of its winter range using geocator technology. *The Auk*, *128*, 531–542. <https://doi.org/10.1525/auk.2011.10280>
- Hobson, K. A., Wilgenburg, S. L. V., Wesolowski, T., Maziarsz, M., Bijlsma, R. G., Grendelmeier, A., & Mallord, J. W. (2014). A multi-isotope (δ 2H, δ 13C, δ 15N) approach to establishing migratory connectivity in palearctic-afrotropical migrants: An example using wood warblers *Phylloscopus sibilatrix*. *Acta Ornithologica*, *49*, 57–69. <https://doi.org/10.3161/000164514X682896>
- Hostetler, J. A., & Hallworth, M. T. (2017). MigConnectivity: Estimate strength of migratory connectivity for migratory animals. R package version 0.2.4. <https://doi.org/10.5281/zenodo.1002229>
- Hostetler, J. A., Sillett, T. S., & Marra, P. P. (2015). Full-annual-cycle population models for migratory birds. *The Auk*, *132*, 433–449. <https://doi.org/10.1642/AUK-14-211.1>
- Korner-Nievergelt, F., Sauter, A., Atkinson, P. W., Guélat, J., Kania, W., Kéry, M., ... Van Noordwijk, A. J. (2010). Improving the analysis of movement data from marked individuals through explicit estimation of observer heterogeneity. *Journal of Avian Biology*, *41*, 8–17. <https://doi.org/10.1111/jav.2010.41.issue-1>
- Manly, B. F. J. (1991). *Randomization and Monte Carlo methods in biology*. London: Chapman and Hall. <https://doi.org/10.1007/978-1-4899-2995-2>
- Marra, P. P., Cohen, E. B., Loss, S. R., Rutter, J. E., & Tonra, C. M. (2015). A call for full annual cycle research in animal ecology. *Biology Letters*, *11*, 20150552. <https://doi.org/10.1098/rsbl.2015.0552>
- Marra, P. P., Norris, D. R., Haig, S. M., Webster, M., & Royle, J. A. (2006). *Migratory connectivity*. In K. R. Crooks, & M. Sanjayan (Eds.), *Connectivity conservation* (pp. 157–183). London: Cambridge University Press. <https://doi.org/10.1017/CBO9780511754821>
- Meirns, P. G., & Hedrick, P. W. (2011). Assessing population structure: FST and related measures. *Molecular Ecology Resources*, *11*, 5–18. <https://doi.org/10.1111/men.2010.11.issue-1>
- Paradis, E., Baillie, S. R., Sutherland, W. J., & Gregory, R. D. (1998). Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology*, *67*, 518–536. <https://doi.org/10.1046/j.1365-2656.1998.00215.x>
- Paxton, K. L., & Moore, F. R. (2015). Carry-over effects of winter habitat quality on en route timing and condition of a migratory passerine during spring migration. *Journal of Avian Biology*, *46*, 495–506. <https://doi.org/10.1111/jav.2015.v46.i5>
- Renfrew, R. B., Kim, D., Perlut, N., Smith, J., Fox, J., & Marra, P. P. (2013). Phenological matching across hemispheres in a long-distance migratory bird. *Diversity and Distributions*, *19*, 1008–1019. <https://doi.org/10.1111/ddi.2013.19.issue-8>
- Rohwer, S., Hobson, K. A., & Rohwer, V. G. (2009). Migratory double breeding in Neotropical migrant birds. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 19050–19055. <https://doi.org/10.1073/pnas.0908121106>
- Ruegg, K., Anderson, E. C., Boone, J., Pouls, J., & Smith, T. B. (2014). A role for migration-linked genes and genomic islands in divergence of a songbird. *Molecular Ecology*, *23*, 4757–4769. <https://doi.org/10.1111/mec.12842>

- Rushing, C. S., Dudash, M. R., Studds, C. E., & Marra, P. P. (2015). Annual variation in long-distance dispersal driven by breeding and non-breeding season climatic conditions in a migratory bird. *Ecography*, 38, 1006–1014. <https://doi.org/10.1111/ecog.01196>
- Rushing, C. S., Ryder, T. B., Saracco, J. F., & Marra, P. P. (2013). Assessing migratory connectivity for a long-distance migratory bird using multiple intrinsic markers. *Ecological Applications*, 24, 445–456.
- Rushing, C. S., Ryder, T. B., Scarpignato, A. L., Saracco, J. F., & Marra, P. P. (2015). Using demographic attributes from long-term monitoring data to delineate natural population structure. *Journal of Applied Ecology*, 53, 491–500.
- Taylor, C. M., & Stutchbury, B. J. M. (2015). Effects of breeding versus winter habitat loss and fragmentation on the population dynamics of a migratory songbird. *Ecological Applications*, 26, 424–37.
- Thorup, K., Korner-Nievergelt, F., Cohen, E. B., & Baillie, S. R. (2014). Large-scale spatial analysis of ringing and re-encounter data to infer movement patterns: A review including methodological perspectives. *Methods in Ecology and Evolution*, 5, 1337–1350. <https://doi.org/10.1111/2041-210X.12258>
- Trierweiler, C., Klaassen, R. H. G., Drent, R. H., Exo, K.-M., Komdeur, J., Bairlein, F., & Koks, B. J. (2014). Migratory connectivity and

- population-specific migration routes in a long-distance migratory bird. *Proceedings of the Royal Society of London B: Biological Sciences*, 281, 20132897. <https://doi.org/10.1098/rspb.2013.2897>
- Webster, M. S., Marra, P. P., Haig, S. M., Bensch, S., & Holmes, R. T. (2002). Links between worlds: Unraveling migratory connectivity. *Trends in Ecology & Evolution*, 17, 76–83. [https://doi.org/10.1016/S0169-5347\(01\)02380-1](https://doi.org/10.1016/S0169-5347(01)02380-1)

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Cohen EB, Hostetler JA, Hallworth MT, Rushing CS, Sillett TS, Marra PP. Quantifying the strength of migratory connectivity. *Methods Ecol Evol*. 2018;9:513–524. <https://doi.org/10.1111/2041-210X.12916>