

THESIS

USING COMMUNITY DETECTION ON NETWORKS TO IDENTIFY MIGRATORY BIRD
FLYWAYS IN NORTH AMERICA

Submitted by

Michael G. Buhnerkempe

Department of Statistics

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2012

Master's Committee:

Advisor: Jennifer A. Hoeting

Geof H. Givens
Colleen T. Webb

ABSTRACT

USING COMMUNITY DETECTION ON NETWORKS TO IDENTIFY MIGRATORY BIRD FLYWAYS IN NORTH AMERICA

Migratory behavior of waterfowl populations in North America has traditionally been broadly characterized by four north-south flyways, and these flyways have been central to the management of waterfowl populations for more than 80 years. However, recent desires to incorporate uncertainty regarding biological processes into an adaptive harvest management program have underscored the need to re-evaluate the traditional flyway concept and bring uncertainty in flyways themselves into management planning. Here, we use bird band and recovery data to develop a network model of migratory movement for four waterfowl species, mallard (*Anas platyrhynchos*), northern pintail (*A. acuta*), American green-winged teal (*A. carolinensis*), and Canada Goose (*Branta Canadensis*) in North America. A community detection algorithm is then used to identify migratory flyways. Additionally, we compare flyway structure both across species and through time to determine broad applicability of the previous flyway concept. We also propose a novel metric, the consolidation factor, to describe a node's (i.e., small geographic area) importance in determining flyway structure. The community detection algorithm identified four main flyways for mallards, northern pintails, and American green-winged teal with the flyway structure of Canada geese exhibiting higher complexity. For mallards, flyway structure was relatively consistent through time. However, consolidation factors and cross-community mixing patterns revealed that for mallards and green-winged teal the presumptive Mississippi flyway was potentially a zone of high mixing between flyways. Additionally, interspersed throughout these major flyways were smaller mixing zones that point to added complexity and uncertainty in the four-flyway concept. Not only does the incorporation

of this uncertainty due to mixing provide a potential alternative management strategy, but the network approach provides a robust, quantitative approach to flyway identification that fits well with the adaptive harvest management framework currently used in North American waterfowl management.

ACKNOWLEDGEMENTS

Funding for this work was provided by USDA Cooperative Agreements 07-7100-0228-CA9, 09-7100-0305-CA, 11-9208-0269-CA and 09-9208-0235-CA. I would like to thank Andrew Merton who cleaned the data and performed the preliminary analyses from which this work drew inspiration. I would also like to thank Colleen Webb, Shweta Bansal, Dan Grear, and Clint Leach for useful discussions on conceptual and methodological issues related to network analyses and John Buhnerkempe who helped provide the management implications of this work. Additionally, I would like to thank my committee members, Geof Givens and Colleen Webb, and especially my advisor, Jennifer Hoeting, for the useful feedback on this manuscript and work in general. Finally, I am deeply grateful for the continued support of my family and especially to my wife, Billie Buhnerkempe, for her love and encouragement without which this would not have been possible.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iv
Introduction.....	1
Methods	4
Data.....	4
Network Construction.....	5
Community Detection.....	5
Node Importance.....	7
Metric Comparison.....	9
Analysis.....	10
Results.....	10
Interspecific variation	10
Temporal Stability	13
Metric Comparison	13
Discussion.....	14
Table 1	18
Table 2	18
Figure 1	19
Figure 2:.....	20
Figure 3	21
Literature Cited.....	22
Appendix A.....	24
Table A1.....	24
Table A2:.....	24
Table A3:.....	25
Table A4:.....	25
Figure A1	26
Figure A2	27
Figure A3	28

Introduction

Migratory bird populations represent a significant natural resource in North America. In the U.S. alone, an estimated 2.3 million people hunted migratory birds for a total of 20 million days in 2006 spending approximately \$1.3 billion (U.S. Depts. Interior and Commerce 2006). When the spending on hunting is coupled with the potential spending from another 15.4 million away-from-home migratory bird watchers (U.S. Depts. Interior and Commerce 2006), migratory birds have a sizable economic impact. However, of all migratory birds in North America, migratory waterfowl represent one of the most important groups encompassing more than half of the migratory bird hunting efforts (U.S. Depts. Interior and Commerce 2006). Mallards alone numbered 12.78 million in 2012 with an expected harvest of 1.55 million birds (USFWS 2012). Given the economic importance of migratory waterfowl species, focused management actions and plans are needed to ensure continued population viability.

Management of migratory waterfowl has centered on the use of flyways. Initially introduced by Frederick Lincoln (1935), migratory flyways divide North America into four historically consistent zones: the Atlantic, Mississippi, Central, and Pacific flyways. The identification of these flyways relied on qualitative analyses and mapping of bird banding data from the 1930's (Lincoln 1935). An update to the flyway concept combined bird banding and recovery data with additional observations of bird movement and abundance (i.e., visual sightings of waterfowl, radar surveillance, waterfowl censuses, and winter inventories; Bellrose 1968). Although the increased depth and breadth of data used in this update provided more realistic and consequently, more complex, pictures of waterfowl migration in North America, flyway patterns were still based primarily on qualitative interpretations of the data (Bellrose 1968). Despite the lack of a rigorous quantitative analysis, the migratory flyways identified by

Lincoln (1935) and updated by Bellrose (1968) have continued to be used largely unchanged in setting waterfowl management strategies and regulation in North America (USFWS 2012).

However, recent emphasis on incorporating uncertainty in management planning may necessitate a more quantitative approach to flyway identification. In particular, an adaptive harvest management framework has been implemented to regulate waterfowl populations in the U.S. (USFWS 2012). By iteratively evaluating management goals and performance, management strategies can be maximized even in the face of multiple sources of uncertainty (Nichols et al. 1995, Williams and Johnson 1995, Williams et al. 1996). Among these sources, environmental variation is an inescapable driver of waterfowl population dynamics given its impact on breeding and forage conditions (Williams et al. 1996, USFWS 2012). With the growing impact of climate change, efforts to incorporate environmental uncertainty are crucial to waterfowl decision-making (Nichols et al. 2011). Additionally, managers are interested in structural uncertainties in waterfowl management (Nichols et al. 1995, Williams et al. 1996, USFWS 2012). These uncertainties are associated with underlying biological processes (e.g., the importance of additive vs. compensatory mortality for harvested waterfowl populations; Nichols et al. 1995, Williams et al. 1996). To this point, however, all uncertainty has been interpreted within the previously identified flyway framework without any apparent recognition that environmental and inherent biological uncertainty may affect the flyway structure on which waterfowl management is based (USFWS 2012) and that this uncertainty may be particularly important within certain regions of the flyways. Thus, a more quantitative approach is needed to provide updated flyway assessments that coincide with adaptive management timeframes and to target areas of flyways that are important for management.

Network models bridge the gap to quantitative analyses of flyway structure and have been used to study animal movements in other species (e.g., sheep, Kiss et al. 2006; and cattle, Ortiz-Pelaez et al. 2006). Network models generally consist of two components: nodes and edges. Nodes are individual units that we seek to describe relationships between, and edges describe interactions between nodes. Thus, as a whole, networks describe the flow of information from one node to another and can be used to comprehensively study the movement of birds (i.e., edges) between spatially distinct regions of North America (i.e., nodes). In the network context, flyways are analogous to communities, where communities are subsets of nodes that are internally well connected with minimal connections to other subsets. Thus, community detection can identify flyways determined by movement patterns across North America in its entirety.

However, community detection alone does not identify those nodes that are important for determining community structure and, consequently, any indication of the regions of flyways where characterizing uncertainty is most important. Early community detection algorithms use network centrality measures, like the degree or betweenness of a node, to identify communities (Girvan and Newman 2002), but these metrics provide better information on the importance in the network as a whole as opposed to communities specifically. More recent work on community roles has sought to bridge this gap (Guimerà and Amaral 2005). However, these community roles are currently defined using relatively arbitrary groupings of two metrics, the within-community degree z-score and the participation coefficient making the identification of nodes important for community structure a system-specific problem (Guimerà and Amaral 2005). To more generally address the question of community importance, we propose the consolidation factor, a novel descriptive statistic describing a node's ability to consolidate other nodes into a single community. The consolidation factor provides a single metric to compare community importance

between nodes. In addition, we provide a novel algorithm for a blocked-permutation test to identify areas of the country that have a greater impact on consolidating the network into communities.

Here, we use network analyses to identify migratory waterfowl flyways quantitatively. This provides an ability to develop alternative definitions of flyways that can be used to characterize uncertainty in the flyway concept within current adaptive management models. Using our novel consolidation factor and associated permutation test, we also identify regions of North America that are important when modeling flyway structure, and consequently, in developing management strategies. Additionally, we compare flyway structure both across species and through time to determine broad applicability of the previous flyway concept. Implications for waterfowl management are also discussed.

Methods

Data

To quantitatively assess migratory flyways we used band-recovery records from the United States Geological Survey Patuxent Wildlife Research Center Bird Banding Laboratory (BBL; USGS Bird Banding Laboratory 2008). These publicly available data contain information on the date and location (to the 10-min. block) of banding and, if the bird was recovered, date and location of recovery for multiple waterfowl species. Because we are interested in waterfowl movement, we limit our analysis to those records that had recovery information. We also restricted our analysis to four species: mallard (*Anas platyrhynchos*; MALL), northern pintail (*A. acuta*; NOPI), American green-winged teal (*A. carolinensis*; AGWT), and Canada goose (*Branta Canadensis*; CAGO). These species were chosen based on their importance in waterfowl

management as well as the availability of consistent high quality band-recovery data. Due to a relative dearth of data for Canada geese, northern pintails and green-winged teal, we compiled recovery data across years 2000-2007, irrespective of banding date, to form one static network for each species allowing us to look at interspecific variation in migratory behavior (Table 1). However, mallards had enough data to allow for construction of separate yearly networks as well which enabled us to study temporal stability of flyways in this species (Table 1). In this case, years were determined by the date of recovery.

Network Construction

We created an approximately 200km x 200km square grid across North America and used the grid cells as nodes for our network. To create the grid, we used the point (40° N, 100° W), an approximate to the center of the United States, as a starting point. We then used Great Circle Distance (i.e., the Haversine function; Sinnott 1984) to find points that were approximately 200km away to the North, South, East, and West. This process was repeated until our grid covered North America. Banding and recovery locations were then assigned to the nearest grid cell, and weighted edges were defined by the number of birds banded in one node and recovered in another. Thus, the network models describe bird movement between geographically distinct areas of North America.

Community Detection

In the network sense, communities are defined as distinct subsets of nodes that are internally well connected but share relatively little connection between subsets. Because we defined nodes to be geographic regions with edges describing bird movement, we sought a

community detection algorithm that would find regions of North America with similar movement patterns in the migratory networks that were distinct from other areas, a concept that equates well with the identification of flyways.

Most community detection algorithms assign nodes to communities to maximize within-community connectivity while minimizing between community connections, or modularity, in the network (e.g., Newman and Garvin 2004, Guimerà and Amaral 2005). Modularity is more formally defined by the quantity Q ,

$$Q = \sum_{i=1}^K (e_{ii} - a_i^2) \quad (1)$$

where e_{ii} is the proportion of edges in the entire network which link between nodes found in community i , a_i is the proportion of edges in the network that link to community i regardless of the community membership of the node on the other end, and K is the total number of communities (Newman and Garvin 2004, Newman 2004). More intuitively, e_{ii} represents the observed within-community connectivity, and a_i^2 represents the expected within-community connectivity if the edges in the network are randomly reassigned. When $Q = 0$, community structure is approximately random. Positive values of Q indicate that the proportion of within-community edges is greater than expected, while negative values of Q indicate that the proportion of between-community edges is greater than expected (Newman and Garvin 2004, Newman 2004). Values of Q above 0.3 are assumed to indicate significant community structure (Newman and Garvin 2004, Newman 2004).

We used a two-stage algorithm proposed by Blondel et al. (2008) to detect communities. The first stage begins by placing nodes into separate communities (i.e., there are as many communities as nodes). Nodes are then sequentially reassigned to different communities

according to the maximum gain in modularity. If there is no gain possible by reassigning a node, it remains in its community. After all nodes have been updated, the second stage begins. Here, a new network is constructed using the communities found in the first stage as nodes with edges representing community-to-community links. The first stage is then repeated on this new network completing one step of the algorithm. This process is then repeated through multiple steps until the modularity, Q , reaches a maximum. For a more complete description of the algorithm see Blondel et al. (2008).

Node Importance

Although the method of Blondel et al. (2008) provides an efficient way to identify community membership, it does not produce a measure of the relative importance of nodes in the formation of communities. To do this, we propose a statistic that measures a node's importance in consolidating other nodes into a single community. We call this the consolidation factor of node j in community i , or CF_i^j .

If we define the nodes belonging to community i as the set $N_i = \{1, \dots, j, \dots, n_i\}$, where n_i is the number of nodes in community i , then the number of node pairs in community i , ignoring node j , is $\rho_i = \binom{n_i - 1}{2}$. We then remove node j and its connections from the network and determine community structure in this reduced network. We then define $N_i^{(-j)} = \{1, \dots, j - 1, j + 1, \dots, n_i\}$ to be the set of nodes in the reduced network that originally belonged to community i . $\rho_i^{(-j)}$ is then the number of unique node pairs in $N_i^{(-j)}$ that can still be found in a community together after reassignment on the reduced network. We define CF_i^j as,

$$CF_i^j = \frac{\rho_i - \rho_i^{(-j)}}{\rho_i} \quad (2)$$

Theoretically, CF_i^j can vary between 0 (i.e., no effect on community structure) and 1 (i.e., node j links all nodes that otherwise would not be found in a community together). Thus, CF_i^j can be interpreted as the proportion of node pairs in community i that depend on node j for their community association.

Given the potential spatial dependencies, standard frequentist tests did not apply for CF_i^j , so we constructed a blocked permutation test to carry out statistical inference. This test determines whether or not the consolidation factors in the block centered at node j , where a block is determined by the grid cells directly surrounding node j for which we have data, are significantly larger than the values found elsewhere in the network. First, we calculated the residual, r_j , for each node j (i.e., $r_j = CF_i^j - \overline{CF}$, where \overline{CF} is the mean consolidation factor of all nodes). We then computed the difference in the mean value of squared residuals within the block from the mean value of the squared residuals outside of the block giving the test statistic,

$$TS_j = \frac{1}{\tau_j} \sum_{k=1}^{\tau_j} r_k^2 - \frac{1}{\tau_{(-j)}} \sum_{k=1}^{\tau_{(-j)}} r_k^2 \quad (3)$$

where τ_j is the number of grid cells directly surrounding cell j for which we have data, and $\tau_{(-j)}$ is the total number of other grid cells. We then permute the original residuals and recalculate TS_j . We use this distribution of TS_j 's to test the hypotheses:

H_0 : Residuals are distributed randomly over geographic space.

H_A : High or low residuals are geographically clustered around node j .

We determine significance by the proportion of permuted TS_j 's which are greater than or equal to the observed value.

Metric Comparison

To assess the relationship between CF_i^j and other node characteristics, we calculated Pearson's correlation coefficients between CF and four other metrics. The metrics explored were,

1. Weighted Degree – the total number of birds banded and recovered at a node
2. Betweenness - the number of shortest paths between any two nodes that go through the node in question, where a path describes any combination of edges that link two nodes.
3. Within-community degree z-score – proposed by Guimerà and Amaral (2005) as a measure of within-community connectedness. This measure is given by,

$$z_j = \frac{\kappa_j - \bar{\kappa}_i}{\sigma_i} \quad (4)$$

where κ_j is the number of within-community edges going to/from node j , $\bar{\kappa}_i$ is the average κ over all nodes in community i , and σ_i is the standard deviation of κ in community i .

4. Participation coefficient – proposed by Guimerà and Amaral (2005) as a measure of a node's connectedness across all communities. This measure is given by,

$$P_j = 1 - \sum_{i=1}^{N_c} \left(\frac{\kappa_{ji}}{k_j} \right)^2 \quad (5)$$

where N_c is the number of communities, κ_{ji} is the number of links node j shares with community i , and k_j is the total degree of node j . Values close to 1 indicate uniform participation in all communities, while 0 indicates only within-community connections (Guimerà and Amaral 2005).

Analysis

To identify species-specific flyways, we determined community structure using the aggregated 2000-2007 BBL data from our four species of interest. In addition, we assessed temporal stability in community structure for mallards by creating separate networks for years 2000-2007. In all community analyses, we also identified nodes that were important in creating community structure as well as calculating weighted degree, betweenness, within-community degree z-score, and the participation coefficient for each node.

Results

Interspecific variation

Mallards had the most band and recovery information accounting for the large size of the observed network (Table 1). The MALL band-recovery network showed significant community structure with eight identified communities (Table 1; Figure 1A), although only six of these communities contained more than ten nodes (Figure 2). Of these, we found the community along the west coast of North America to be the largest. Despite its size, two locations in Alberta, Canada were primarily responsible for shaping this community (i.e., highly significant consolidation factors). The community found in the central plains regions of North America was similar in size to the west coast community, but this community did not show a similar dependence on a limited number of regions for its structure. Instead, several regions in the center of the community (i.e., Minnesota, Iowa, and Missouri) showed relatively smaller, non-significant, consolidation factors (Figure 1A). In between the central and west coast communities, we also found a smaller community that was highly dependent on all of its members for its structure (Figure 1A). Moving to the east, we found two communities in the

areas north and south of the Great Lakes (Figure 1A). Although the community found to the south of the Great Lakes was about twice as large as its northern counterpart, both communities contained a large proportion of nodes with significant consolidation factors indicating the potential break down of these communities through small changes in membership. The last major community in North America was found in the eastern part of Canada and down the east coast of the United States. Within this community, only two nodes, one in Maryland and the other in Connecticut, were found to be important in holding this community together, although neither were significant (Figure 1A).

Northern pintails had considerably less band and recovery information than mallards leading to a smaller movement network (Table 1). The NOPI network also had less modularity than the MALL network but still showed significant community structure with seven total communities and four with more than ten nodes (Table 1; Figure 2). The largest community of NOPI was in the central plains region of the U.S. extending across to the Great Lakes (Figure 1B). Three nodes in the north-central U.S. and south-central Canada had the highest importance within this community, although these consolidation factors were non-significant. The west coast community of NOPI was the second largest community for this species and again only three nodes showed increased, but non-significant, importance in the structure of the community, one in California and two in Alaska (Figure 1B). As seen in mallards, we observed another community at the intersection of the western and central communities in NOPI (Figure 1B). However, this intervening community was much larger in the NOPI network (Figure 2) and extended along the Rockies and into Mexico creating much more separation than observed in the mallards (Figures 1A and 1B). The Rocky Mountain community was structured largely by one node in the Northwest Territories. Additionally, we again observed an eastern community that

extended from southeastern Canada down the east coast of the U.S., but as mentioned before, the extension of the central community through the Great Lakes region eliminated any separate communities in between it and the east coast community (Figure 1B).

American green-winged teal had the least band-recovery information and the smallest movement network (Table 1). Despite this, the AGWT network exhibited relatively high modularity with four observed communities, all of which contained more than 10 nodes (Figure 2). The locations of the four communities were similar to those observed in the MALL network with large west coast, central, and east coast communities and a smaller Great Lakes community. However, in general, the consolidation factors in these communities are close to zero except in the Great Lakes community. Here, we found significant importance for the nodes around the Great Lakes in structuring this community.

Canada geese had a moderate amount of band-recovery data but still had a similar sized network to NOPI and AGWT (Table 1). The CAGO network had the highest modularity (Table 1) but also had the most communities with ten of these having more than ten nodes (Figure 2). With more communities, we found more evenness in community size (Figure 2). We also found that the smaller communities are still found in relatively narrow bands across North America, except in the southeastern U.S. where small communities were relatively localized (Figure 1D). For the most part, nodes showed non-significant importance in community structure, but the community extending from Canada down the Mississippi River to Texas was made up almost exclusively of nodes that were important to structure (Figure 1D). In total, the CAGO network exhibited the most complex community structure of the four species considered.

Temporal Stability

We also analyzed how community structure changed through time for the MALL network. For the MALL network, communities were similar across years from 2000-2007 and resembled those found by aggregating across years (Figures A2-3). When we looked at the number of times pairs of nodes appeared in the same community across all eight years, we found that the communities were relatively constant through time with only eleven unique communities identified and only seven with more than ten nodes (i.e., the color blocks along the diagonal in Figure 3). Despite the tendency for nodes to appear in the same community through time, we also observed some variation in community structure with some nodes switching between communities over time (i.e., the upper right and lower left regions of Figure 3). This variation was more pronounced in some years than others. In particular, in 2000 and 2002-2004, the west coast community split into two separate communities (Figures A2A, A2C-D, and A3A). Similarly, in 2002, 2003, and 2006, the east coast community split as well (Figures A2C- D and A3C).

Metric Comparison

For the species for which we had the most data, MALL and CAGO, *CF* was only weakly correlated with the weighted degree and within-community degree z-score (Table 2). However, the correlation between these metrics and *CF* increased in the species with less data, NOPI and AGWT (Table 2). In all species, *CF* had a slight correlation with betweenness that was always less than that seen with weighted degree (Table 2). The correlation between *CF* and the participation coefficient was negligible in all species (Table 2).

Of the previously proposed metrics of communities roles (Guimerà and Amaral 2005), within-community degree z-score was almost always highly correlated with more traditional network characteristics (i.e., degree and betweenness; Tables A1-A4). The participation coefficient, however, was never well correlated with any metric (Tables A1-A4).

Discussion

Management of migratory waterfowl in North America relies heavily on the concept of flyways that was originally developed over 80 years ago (Lincoln 1935, USFWS 2012). Despite a growing desire to incorporate uncertainty in management decisions (USFWS 2012), flyways continue to be based on fixed qualitative interpretations of regional waterfowl movements. In this paper, we have constructed network models of waterfowl movement from bird banding data and used a recent community detection technique (Blondel et al. 2008) to identify communities. This approach provides robust, quantitative descriptions of migratory flyways that are based on continental, as opposed to regional, movement patterns that can be re-interpreted over management timeframes to incorporate uncertainty in flyway structure into management strategies.

North America has traditionally been divided into four flyways, the Pacific, Central, Mississippi, and Atlantic (Lincoln 1935). At a glance, our analysis appears to support this concept. Mallards, northern pintails, and green-winged teal all exhibited four, more or less, distinct north-south bands in their community structure (Figures 1A-C). However, the relative sizing of the flyways and the presence of more local communities, especially with mallards, underscores the need to understand uncertainty in flyway structure. The more localized community structure and obvious uncertainty in flyways was readily apparent in Canada geese,

where we observed the most complex community structures. This complexity is most likely due to the presence of resident geese that do not migrate leading to more numerous local communities.

In addition to identifying potential flyways, we also analyzed how individual nodes impacted communities through the consolidation factor. When we look at the consolidation factors in the various flyways, two general patterns emerge. First, a flyway may contain relatively few highly important nodes with significant consolidation factors (see the Pacific flyway in Figure 1A and Figure 1B in general). From a management perspective, these nodes are often found in the northern regions of flyways and most likely represent breeding habitats that consistently provide a starting location for southward migration through a single flyway. Other flyways, however, contain a large proportion of nodes with high consolidation factors. In this case, removal of any of these nodes will cause a break down in the structure of the community and a change in the affiliation of community members. This sensitivity in community affiliation is most significant in our interpretation of the community that represents the Mississippi flyway in mallards and green-winged teal (Figures 1A and C) causing us to question whether these communities in particular are distinct as flyways.

If these communities around the Great Lakes are not distinct flyways, then what do they represent? Node participation coefficients are high in these communities suggesting that the supposed Mississippi flyways over the Great Lakes are regions with significant cross-community mixing (Figure A1). Not only are participation coefficients high in these communities, but they are universally high (Figure A1). Other communities show regions of low participation indicating the presence of a strong flyway pattern (Figure A1). Consequently, our analysis may only detect three significant flyways interspersed with one large and several smaller zones of

increased mixing. The presence of three distinct flyways with a fourth major flyway as a mixing zone was also supported in the temporal analysis of mallard communities. Yearly communities for this species were consistent, except in two key areas (Figures 3 and A2-3). First, the Pacific flyway split on several occasions (Figures A2-3). Perhaps more importantly, however, communities east of the Great Lakes showed considerable variation in number, geographic extent, and importance patterns (Figures A2-3). This variation lends support to the notion that considerable mixing across flyways occurs around the Great Lakes. The recognition of flyways as potential mixing zones, while a departure from the traditional flyway concept (Lincoln 1935), may actually fit well with stock assessments of mallards, where three breeding stocks (i.e., the Western, Mid-continent, and Eastern) are officially recognized and appear to match well with our three significant flyways (USFWS 2012). Thus, our quantitative approach to flyway identification may provide a more direct link between the underlying population dynamics of waterfowl and management actions than the previously identified flyways.

Given the benefits of the consolidation factor in determining node importance, we also wanted to assess its novelty in relation to other network metrics. As a result, our *CF* metric conveys information on community structure that cannot be explained entirely by more traditional node centrality measures. In general, *CF* had weak or intermediate relationships with weighted degree and betweenness, more traditional measures of node centrality (Table 2). Relationships with more community-specific metrics, like within-community degree z-score and participation coefficient (Guimerà and Amaral 2005), was variable. Here, *CF* was more highly correlated with the within-community z-score but had no relationship with participation (Table 2). The relationship with z-score, however, ultimately comes from the tight relationship between weighted degree and z-score (Tables A1-4). However, the relationship between *CF* and the

notion of community centrality (Newman 2006) is still unclear, and more work is needed to elucidate the relationship between *CF*, community centrality, and community stability.

Our network approach to flyway identification provides useful insights into the flyway structure of North American waterfowl. We found support for three major distinct flyways with a fourth major flyway that showed considerable mixing of birds across flyways. This interpretation held across multiple species and coincided well with officially recognized breeding stocks of mallards. The ability to reassess flyway structure using robust, quantitative methods over management timescales should fit well with the structural (i.e., biological process) uncertainty already found within the adaptive harvest management strategy for waterfowl (Nichols et al. 1995, Williams and Johnson 1995, Williams et al. 1996, USFWS 2012). This approach also benefits from limited data requirements. For this study, we used only bird banding and recovery data, which is easily accessible and continuously updated (USGS Bird Banding Laboratory 2008). However, improvements to management recommendations could be made via the incorporation of spatial and temporal bird abundance data in the flyways (Crissey 1955). Fine-scale information on bird abundances in this framework would allow managers to predict the volumes and time-scales over which birds move in the flyways thus increasing the effectiveness of harvest policies. Despite this limitation, we view the network approach as a significant improvement over previous qualitative interpretations of flyways that can be used to enhance the adaptive management of North American waterfowl.

Table 1: Comparison of networks constructed using aggregated 2000-2007 BBL data for the four species of interest.

Species	Number of records	Number of nodes	Number of communities	Modularity (Q)
MALL	553,779	629	8	0.62
NOPI	7,836	340	7	0.36
AGWT	6,178	309	4	0.47
CAGO	83,797	345	15	0.77

Table 2: Pearson's correlation coefficients between CF and various metrics of a node's role in the community. Correlations were calculated separately for each species of interest.

Species	Weighted Degree	Betweenness	Within-comm. degree z-score	Participation coefficient
MALL	0.35	0.26	0.31	0.11
NOPI	0.42	0.27	0.53	0.11
AGWT	0.43	0.21	0.49	0.23
CAGO	0.38	0.30	0.37	-0.08

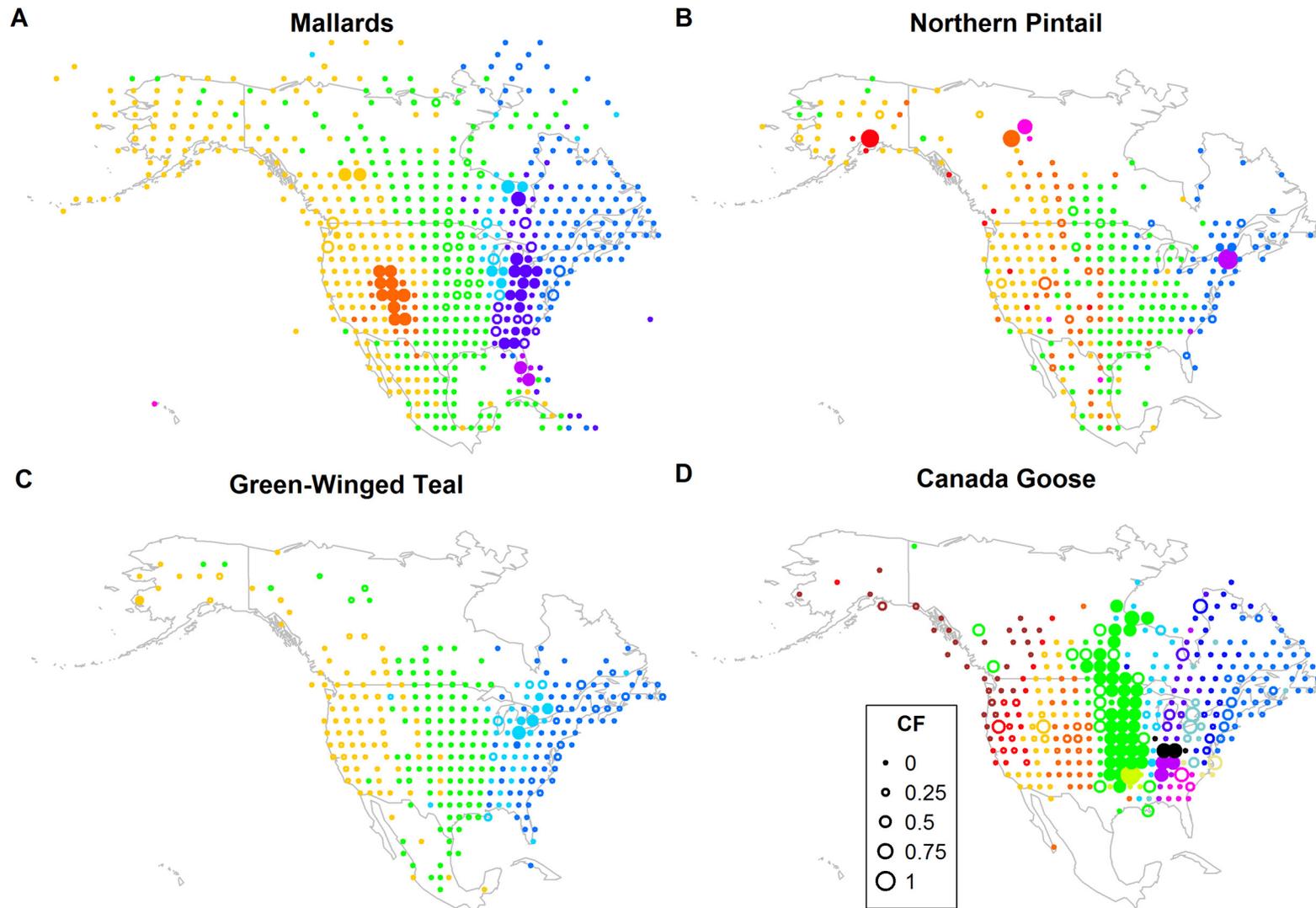


Figure 1: Spatial distribution of species networks for (A) mallards, (B) northern pintails, (C) American green-winged teal, and (D) Canada geese. The color represents the community membership of the node, and the size of the node is scaled to the node's consolidation factor. Filled circles indicate nodes with a statistically significant consolidation factor at the $\alpha = 0.05$ level.

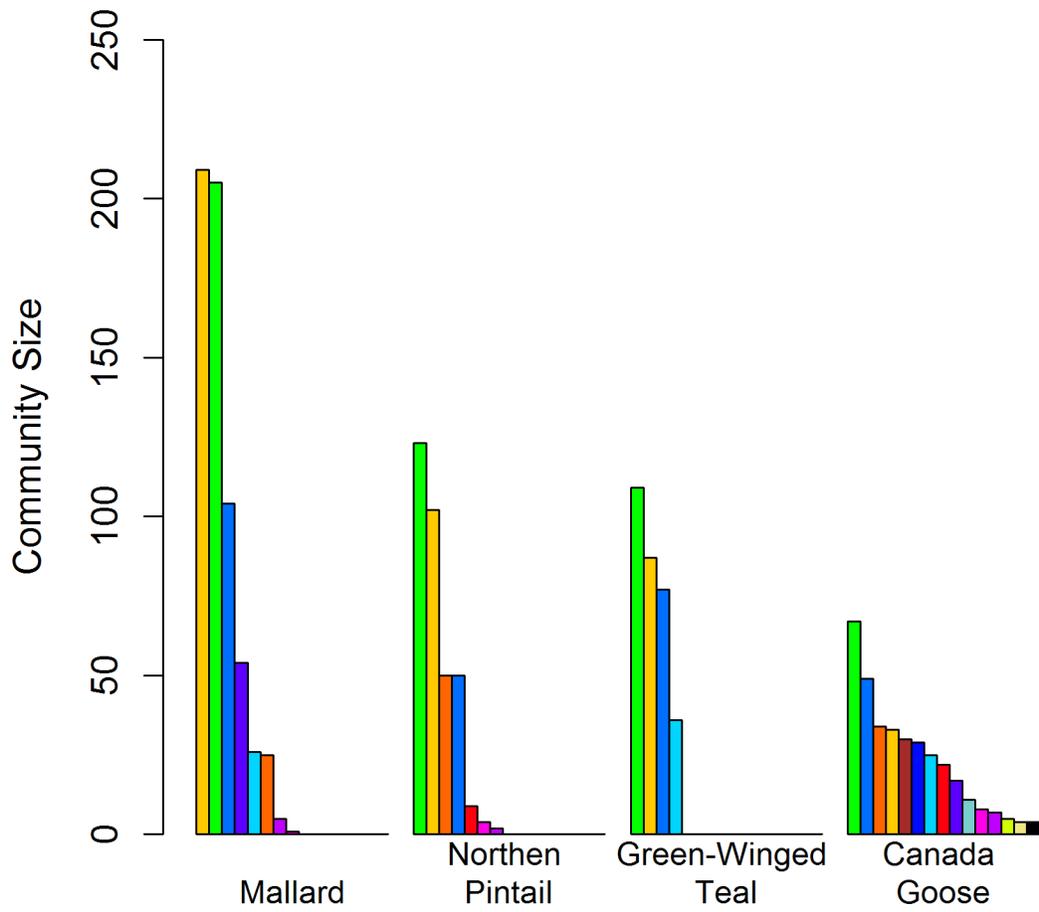


Figure 2: Community size distribution for the four species of interest. Colors correspond to community colors found in Figure 1.

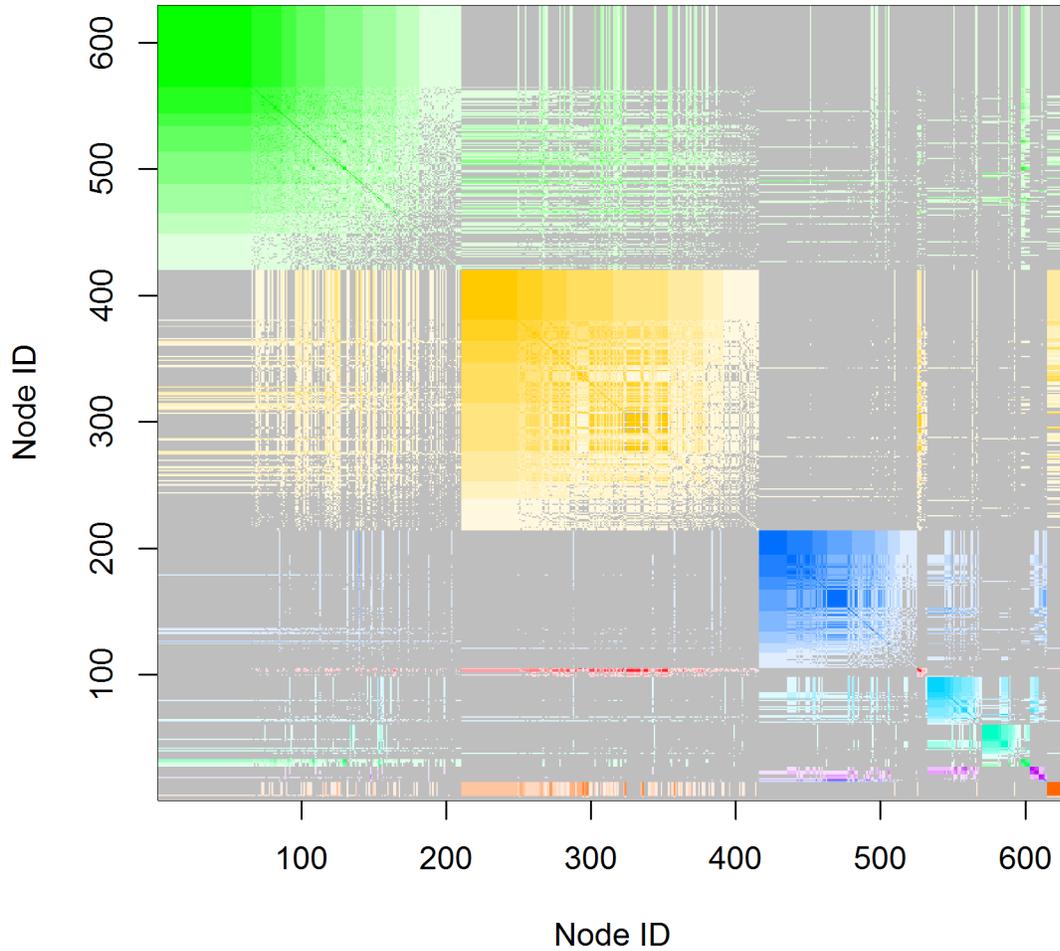


Figure 3: Affinity matrix for the MALL network showing the number of times pairs of nodes appeared in the same community from 2000-2007. The saturation of the color exhibits the number of times a pair of nodes appeared in the same community (i.e., the more frequent they were seen together, the stronger the color). The color of the node pair is determined by the community the node along the y-axis was most often found in. Grey indicates the pair of nodes was never found in the same community.

Literature Cited

- Bellrose, F.C. 1968. Waterfowl migration corridors: East of the Rocky Mountains in the United States. Illinois Natural History Survey Biological Notes No. 61. 24pp.
- Blondel, V.D., Guillaume, J.L., Lambiotte, R., and Lefebvre, E. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* P10008.
- Crissey, W.F. 1955. The use of banding data in determining waterfowl migration and distribution. *Journal of Wildlife Management* 19: 75-84.
- Girvan, M., and Newman, M.E.J. 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* 99: 7821-7826.
- Guimerà, R., and Amaral, L.A.N. Functional cartography of complex metabolic networks. *Nature* 433 (2005): 895-900.
- Kiss, I.Z., Green, D.M. and Kao, R.R. 2006. The network of sheep movements within Great Britain: network properties and their implications for infectious disease spread. *Journal of the Royal Society Interface* 3: 669-677.
- Lincoln, F. 1935. The migratory flyways of North America. USDA Circular 342. 12 pp.
- Newman, M.E.J. 2004. Fast algorithm for detecting community structure in networks. *Physical Review E* 69.6: 066133.
- Newman, M.E.J. 2006. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences* 103: 8577-8582.
- Newman, M.E.J., and Girvan, M. Finding and evaluating community structure in networks. *Physical Review E* 69.2 (2004): 026113.
- Nichols, J.D., Johnson, F.A., and Williams, B.K. 1995. Managing North American waterfowl in the face of uncertainty. *Annual Review of Ecology and Systematics* 26: 177-199.
- Nichols, J.D. Koneff, M.D., Heglund, P.J., Knutson, M.G., Seamans, M.E., Lyons, J.E., Morton, J.M., Jones, M.T., Boomer, G.S., and Williams, B.K. 2011. Climate change, uncertainty, and natural resource management. *Journal of Wildlife Management* 75: 6-18.
- Ortiz-Pelaez, A., Pfeiffer, D.U., Soares-Magalhães, R.J. and Guitian, F.J. 2006. Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK. *Preventive Veterinary Medicine* 76: 40-55.
- Sinnott, R.W. 1984. Virtues of the Haversine. *Sky and Telescope* 68: 158.

- U.S. Department of the Interior, Fish and Wildlife Service, and U.S. Department of Commerce, U.S. Census Bureau. 2006. National Survey of Fishing, Hunting, and Wildlife-Associated Recreation.
- U.S. Fish and Wildlife Service. 2012. Adaptive Harvest Management: 2012 Hunting Season. U.S. Department of the Interior, Washington, D.C. 58 pp.
- U.S. Geological Survey Bird Banding Laboratory. 2008. North American bird banding and band encounter data set. Patuxent Wildlife Research Center, Laurel, MD. Data created January 17, 2008.
- Williams, B.K., Johnson, F.A., and Wilkins, K. 1996. Uncertainty and the adaptive management of waterfowl harvests. *Journal of Wildlife Management* 60: 223-232.
- Williams, B.K., and Johnson, F.A. 1995. Adaptive management and the regulation of waterfowl harvests. *Wildlife Society Bulletin* 23: 430-436.

Appendix A

Table A1: Pearson's correlation coefficients between all non-CF metrics for mallards.

MALL	Weighted Degree	Betweenness	Within-comm. degree z-score	Participation coefficient
Weighted Degree	1.00	0.75	0.94	-0.06
Betweenness		1.00	0.78	-0.04
Within-comm. degree z-score			1.00	-0.14
Participation coefficient				1.00

Table A2: Pearson's correlation coefficients between all non-CF metrics for northern pintails.

NOPI	Weighted Degree	Betweenness	Within-comm. degree z-score	Participation coefficient
Weighted Degree	1.00	0.79	0.90	0.12
Betweenness		1.00	0.76	0.03
Within-comm. degree z-score			1.00	0.09
Participation coefficient				1.00

Table A3: Pearson’s correlation coefficients between all non-CF metrics for American green-winged teal.

AGWT	Weighted Degree	Betweenness	Within-comm. degree z-score	Participation coefficient
Weighted Degree	1.00	0.35	0.81	0.04
Betweenness		1.00	0.29	0.06
Within-comm. degree z-score			1.00	0.14
Participation coefficient				1.00

Table A4: Pearson’s correlation coefficients between all non-CF metrics for Canada geese.

CAGO	Weighted Degree	Betweenness	Within-comm. degree z-score	Participation coefficient
Weighted Degree	1.00	0.77	0.83	-0.06
Betweenness		1.00	0.77	-0.09
Within-comm. degree z-score			1.00	-0.19
Participation coefficient				1.00

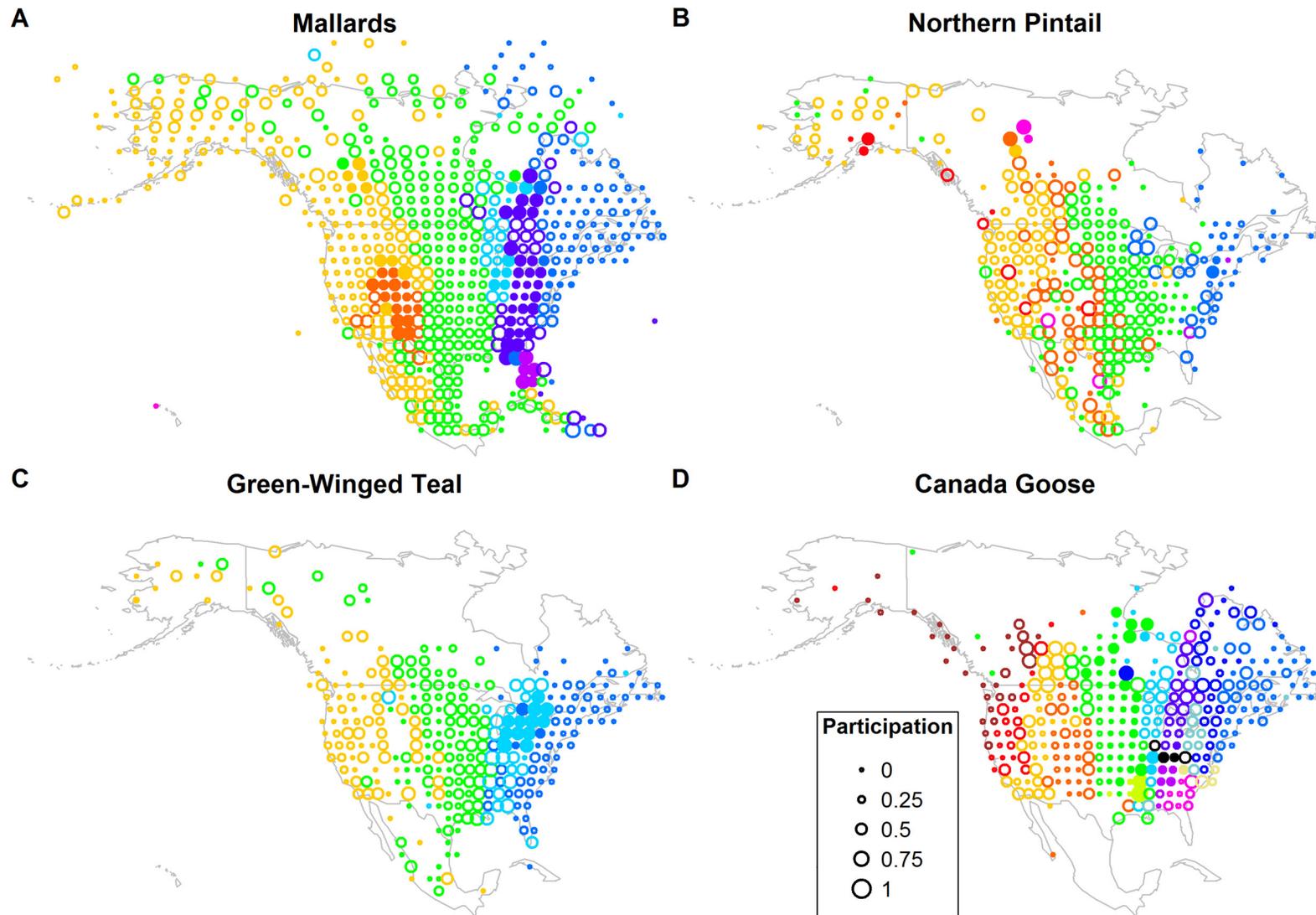


Figure A1: Cross-community mixing patterns for (A) mallards, (B) northern pintails, (C) American green-winged teal, and (D) Canada geese. The color represents the community membership of the node, and the size of the node is scaled to the node's participation coefficient. Filled circles indicate nodes with a statistically significant consolidation factor at the $\alpha = 0.05$ level.

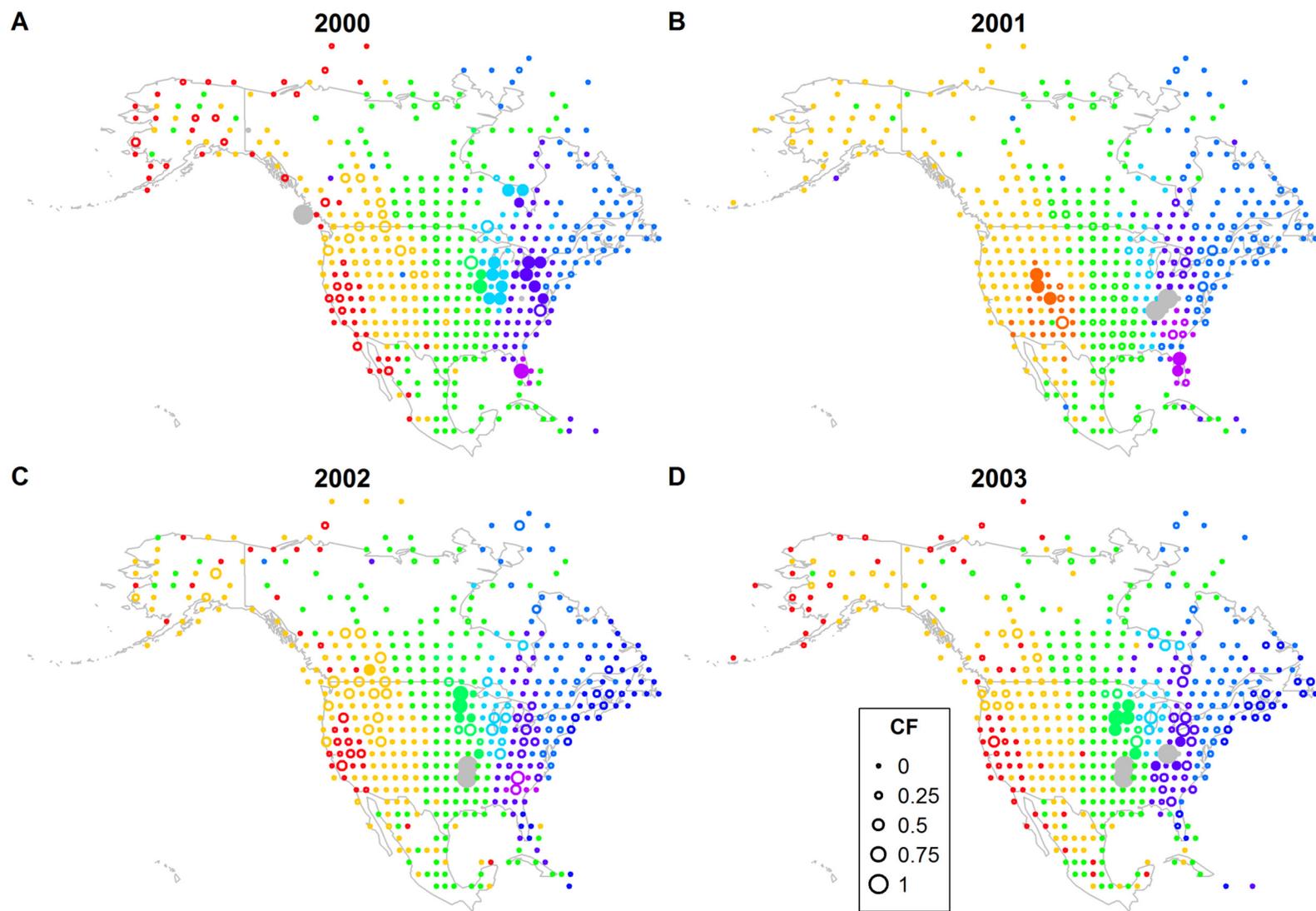


Figure A2: Yearly mallard networks for (A-D) 2000-2003. The color represents the community membership of the node, and the size of the node is scaled to the node's consolidation factor. Filled circles indicate nodes with a statistically significant consolidation factor at the $\alpha = 0.05$ level. Grey indicates community with only one or two members.

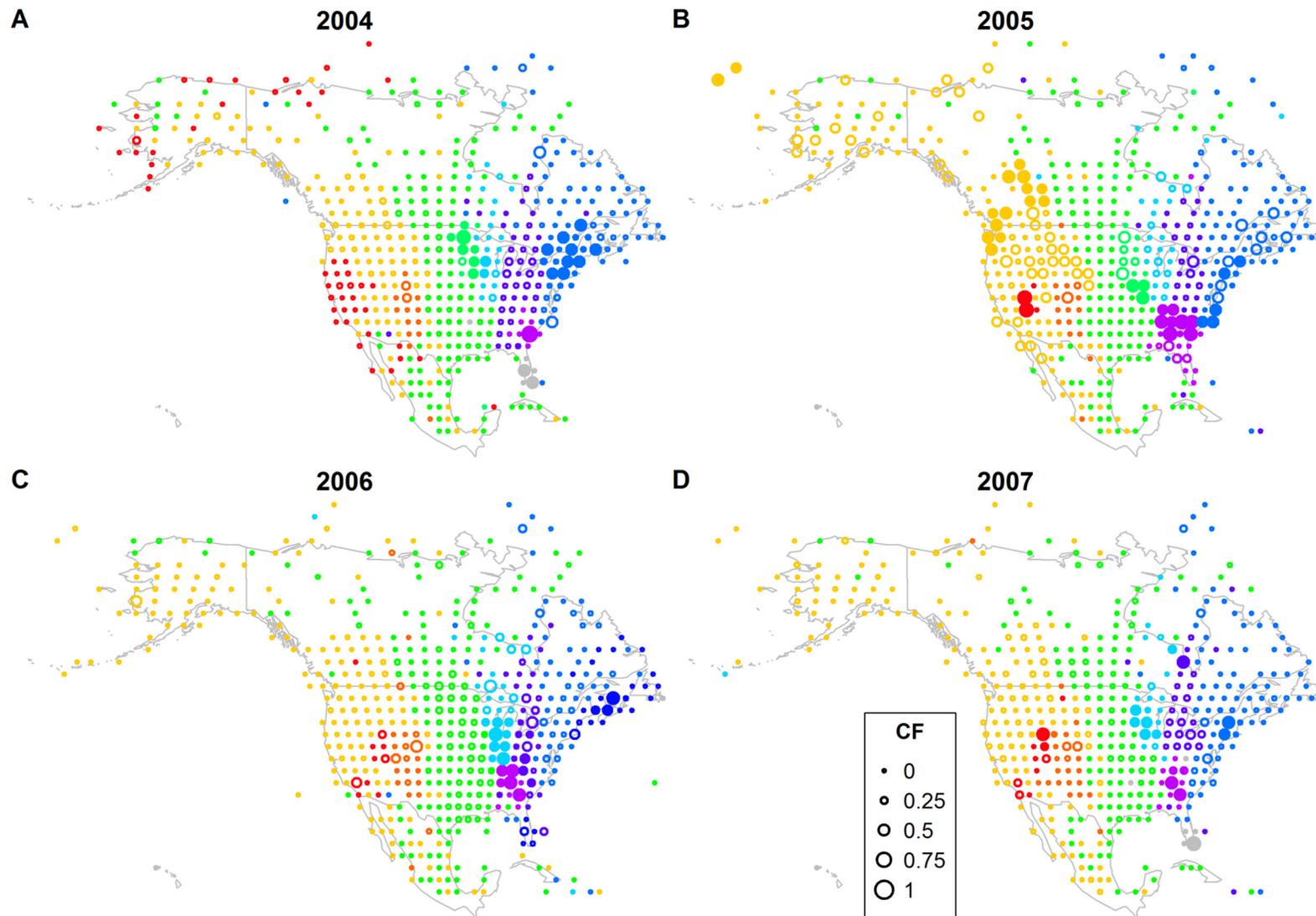


Figure A3: Yearly mallard networks for (A-D) 2004-2007. The color represents the community membership of the node, and the size of the node is scaled to the node's consolidation factor. Filled circles indicate nodes with a statistically significant consolidation factor at the $\alpha = 0.05$ level. Grey indicates community with only one or two members.