Linking the continental migratory cycle of the monarch butterfly to understand its population decline

Hidetoshi Inamine, Stephen P. Ellner, James P. Springer and Anurag A. Agrawal

H. Inamine, S. P. Ellner and A. A. Agrawal (aa337@cornell.edu), Dept of Ecology and Evolutionary Biology, Cornell Univ., Ithaca, NY 14853, USA. AAA also at: Dept of Entomology, Cornell University, Ithaca, NY, USA. – J. P. Springer, North American Butterfly Association, Morristown, NJ, USA.

Cross-continent animal migrations are some of the most spectacular ecological phenomena and are severely threatened (Wilcove and Wikelski 2008). A major hurdle in conceptualizing and conserving animal migrations is understanding the demographic connectivity and population dynamics over the migratory cycle, especially in the face of large spatial movements over long time scales. For hundreds of years the annual migration of the monarch butterfly Danaus plexippus from Mexico to the northern USA and Canada has captured the imagination of scientists and non-scientists alike (Gustafsson et al. 2015). Like many other migratory animals, monarch butterflies have a complex multigenerational cycle and changes in any one of the stages can affect their population dynamics (Fig. 1).

Monarchs have a history of 10-fold or larger annual population fluctuations (Swengel 1995, García-Serrano et al. 2004; Rendón-Salinas et al. 2014). Yet, a 2011 study based on 17 years of data revealed a precipitous long-term population decline at the overwintering sites in Mexico (Brower et al. 2012b, Rendón-Salinas et al. 2014) and the rate of decline may be increasing over time (Fig. 2) (Ries et al. 2015b). Nonetheless, two independent fall monitoring programs that enumerate returning monarchs from the northern USA and Canada did not show a decline over the same time period (Davis 2012, Badgett and Davis 2015). Understanding the complex population dynamics of monarchs over space and time therefore remains an important ecological as well as conservation challenge.

From a conservation perspective, it is critical to identify key stage(s) influencing population dynamics. The classic case of loggerhead sea turtle conservation exemplifies this issue, as initial efforts emphasized life stages (eggs and hatchlings) that were unlikely to substantially benefit the population (Crouse et al. 1987). The monarchs’ annual cycle has several potentially critical stages (Malcolm and Zalucki 1993, Flockhart et al. 2013). In late winter, overwintering butterflies mate and fly from Mexico north to the southern USA, where most individuals lay eggs on emerging milkweeds, and die (Fig. 1). The next generation migrates...
north, expanding into southern Canada on both sides of the Appalachians, but east of the Rocky Mountains. Additional generations breed in these northern regions. Beginning in late August, unmated butterflies siphon through several funneling points and migrate up to 4000 km back to the overwintering grounds in Mexico. Millions of monarchs ultimately concentrate on about a dozen mountaintops, in an area less than 800 km² (Slayback and Brower 2007).

Since the discovery of the overwintering grounds in Mexico, numerous threats to monarchs and their migration have been identified, most of which involve human activities such as logging and agriculture (Malcolm 1993, Oberhauser and Peterson 2003, Bradley and Altizer 2005, Brower et al. 2006). Recent analyses have specifically implicated the decline of milkweed host plants due to increased use of genetically modified herbicide-resistant crops, especially in the agricultural Midwest USA (the “milkweed limitation hypothesis”) (Oberhauser et al. 2001, Brower et al. 2012a, Pleasants and Oberhauser 2013). Given that monarch caterpillars exclusively feed on milkweed, it is not surprising that milkweed decline appears to be locally impacting butterfly and caterpillar abundance (Zalucki and Lammers 2010, Pleasants and Oberhauser 2013, Stenoien et al. 2015). These local demographic effects were recently incorporated in prospective models for the monarch’s migratory cycle (Flockhart et al. 2015), where a stage-structured matrix projection model for the monarch life-cycle was combined with spatial structure and migration.

Here we take an alternative approach, using multiple datasets covering 22 years of monarch monitoring programs across North America to retrospectively investigate associations between population dynamics in different regions, and to identify stages contributing to the recent population decline. Using count data reported to the North American Butterfly Association (NABA) and other citizen scientist data, we sought to follow the cycle from overwintering abundance, to spring and summer breeding populations, and finally to fall migrating butterfly counts. Our analyses
and findings overlap, in part, with other recent analyses of monarch population trends (Davis 2012, Badgett and Davis 2015, Ries et al. 2015a, b). However, our study is the first to link the entire annual migratory cycle, and the first to analyze temporal trends in population indices and stage-to-stage relationships through the migratory cycle. Thus, our analyses are uniquely positioned to address the milkweed limitation hypothesis by understanding demographic connectivity at larger temporal and spatial scales.

Material and methods

NABA citizen science data

The North American Butterfly Association (NABA) has compiled butterfly counts from participating citizens across North America since 1975. The dataset consists of thousands of observations on the number of adult monarchs, the location, number of participants, and total hours spent in the field for each census. We focus on > 6000 records from 1993–2014, as each of these years had a substantial number of counts (mean of 290 counts per year) and matches the census records available from the Mexico overwintering grounds. Although census locations are not fixed, butterflies are counted across the regions of interest (Fig. 3), and should represent regional sums of adults across small scale habitat variation. We scaled each count by total group hours (Koenig 2006, Ries et al. 2015a, b). In the Supplementary material Appendix 1, we show that our results are robust to alternative ways of normalizing counts for observer effort.

To focus on the long-range migration of the eastern monarch population, we eliminated some regions from the NABA dataset. We removed counts west of 105°W corresponding to the largely separate migratory “Californian” populations (Koenig 2006) and counts from Georgia, South Carolina, and Florida, corresponding to the mostly non-migratory “Gulf Coast” populations (Brower 1995). The NABA dataset was then subdivided into three spatially distinct regions (Fig 3A): South (south of 34.5°N and west of 79°W), Northeast (north of 34.5°N and east of 79°W, corresponding to the Appalachian mountain divide), and Midwest (north of 34.5°N and west of 79°W). Northeast and Midwest include dates from 27 March to 3 October, corresponding to the entire summer breeding season. South was further temporally subdivided into two groups: Spring South (1 March through 30 June, corresponding to reproducing migrants moving north), and Fall South (1 September to 30 November, corresponding to returning migrants moving south).

The NABA citizen science dataset has spatially and temporally variable sampling effort, and its quality has been challenged (Pleasants et al. 2016). We therefore based our population abundance indices on a temporal moving average for each region (Brown 2004). For each date, we calculated an equally weighted average of all the counts in the region falling in a seven-day window, centered around that date. This approach alleviates biases in the population index due to temporal variation in sampling intensity. Ripley’s K function shows that there is no evidence for increased clustering of census points over time (Supplementary material Appendix 1). We assessed and addressed other potential biases in the NABA dataset (varying sampling intensity, missing butterfly arrivals, etc.) in multiple ways (Supplementary material Appendix 1).

The total monarch index for a given year was calculated by summing the population index within a region during the time periods defined above. Northern populations have

![Figure 3. (A) Locations of North American Butterfly Association citizen science counts of monarch adults east of the Rocky Mountains, and separated into three regions: South (south of 34.5°N and west of Appalachians, excluding Georgia, South Carolina and Florida where monarchs are largely non-migratory), Midwest (west of Appalachians), and Northeast (east of Appalachians) (1993–2014). Red dots indicate count locations and increasing color intensity indicates overlapping count points across years. A few count points located outside of this region were included in the analyses, but are not shown on the map. (B) Mean ± SEM (across years) moving average of the relative monarch population index over days of the year in the three regions. Shown above the curves are the windows of dates for which we used data to estimate the annual population index, with colors and line types corresponding to regions. Note that South is divided into spring and fall populations. The proportional abundance reflects the regional population density, not regional total population size. The relative indices here are therefore not directly comparable.](image-url)
temporally overlapping and variable numbers of generations, so we calculated season-wide sums of observations in the midwest and northeast (99% of the counts taken June–August). The same number of days was used for a season each year, so sums and averages are equivalent.

**Assessing butterfly phenology in NABA data across years**

Seasonal population trends over the 22 years in the NABA dataset were estimated by calculating the proportional number of monarchs for each day (population index at each date / total index for that year across all regions). This value indicates the proportional abundance of butterflies seen each day. To assess whether the NABA dataset captured the known pattern of annual migratory pheno-logy, we plotted the mean daily proportional abundance trends with the standard error calculated across the 22 years.

**Additional data from repeatedly-measured sites**

Cape May Point, New Jersey, is a major funneling point for southern migrating monarchs from the northeastern USA (Walton and Brower 1996, Davis 2012). Transect counts are conducted three times daily on a defined route, and the counts are normalized by hours of observation. Weekly averages are reported over nine weeks from 1 September to 31 October (1992–2014). We summed the weekly averages to get a population size index reflecting the total number of butterflies migrating south.

Peninsula Point, Michigan, is a funneling point for southern migrating monarchs from eastern and potentially midwestern Canadian populations (Meitner et al. 2004, Davis 2012). Transect counts are conducted one or two times daily on a defined route, and the counts are normalized by hours of observation. We obtained original data sheets from the Peninsula Point Monitoring Project. We averaged the daily counts for each week and summed the weekly averages to get our population index. The count period runs from early August to late September (1996–2014), so the population index for each year is based on the number of butterflies passing through over eight weeks. Two years (1996 and 1998) were missing from the original data sheets, but were presented in another study using a slightly different estimation method (Davis 2012). We used regression analysis of the data from the 17 years included in both studies to predict values for the two missing years.

Mexican overwintering sites are monitored by the World Wildlife Fund. A December estimate of total hectares occupied by roosting butterflies is reported for 1993–2014 as a proxy for population size (Rendón-Salinas et al. 2014).

Finally, United States Department of Agriculture collects data on the adoption of herbicide resistant crops and we used the mean of adoption rates for corn and soybean from 1996–2015. We expected that increase adoption of herbicide resistant crops leads to increased herbicide use and consequently, decrease in milkweed abundance. We therefore employ this dataset as a proxy for milkweed abundance in the agricultural fields (Pleasants and Oberhauser 2013). To move beyond the simply association between herbicide resistant crop adoption and monarch populations, we test whether year-to-year changes in adoption correlate with year-to-year changes in monarch population estimates.

**Analyzing relationships between regions**

We used regression analysis to test for relationships between annual population indices at successive stages in the migratory cycle (from ‘donor’ to ‘recipient’, respectively): Mexico to Spring South (overwintering butterflies migrating north); Spring South to Northeast and Spring South to Midwest (butterflies migrating further north, while population size builds up over 2–3 generations); Northeast to Cape May and Midwest to Peninsula Point (fall migration south); Cape May to Fall South, Peninsula Point to Fall South, and Fall South to Mexico (migration further south); Midwest to Mexico, and Northeast to Mexico (bypassing Fall South). All analyses were performed in R ver. 2.15.1. (<www.r-project.org>).

The one NABA census that limited sampling was Fall South (we only have counts beginning in 2002 and 2014 was the single year with a count in November. The mean number of counts from 2002 to 2014 is 8.85). When we summarized the NABA regions, between-year variability was substantially higher for Fall South than for the Midwest and Northeast summer indices (Fig. 3B). These patterns call into question the quality of the fall South index perhaps due to much-reduced sampling. Nonetheless, we include the fall South index in analyses, but we do not base conclusions strictly on those results. In addition, we include regressions between the northern censuses and Mexico.

Given our knowledge of the migratory cycle, finding the expected positive relationship between successive pairwise stages indicates reasonable data quality and provides a basis for further analyses of the demographic links between the indices (WWF Mexico, NABA, Cape May, Peninsula Point) that were collected independently using different protocols. We tested for the link between overwintering and the spring migration by regressing the NABA Spring South population index against the Mexico population index. Spring South counts did not include March and April for 1993–2004, thereby missing the crucial first-generation migrants from Mexico (Supplementary material Appendix 1). Accordingly, for this link we focus on 2005–2014. We determined the link between summer breeding and the fall migration by regressing the Cape May index against the NABA Northeast index, and the Peninsula Point index against the NABA Midwest index. Our strong a priori expectation was to find a positive relationship between population indices of the ‘donor’ and ‘recipient’ regions based on the monarch’s known annual cycle. Therefore, although we present two-tailed p-values, we consider p-values below 0.1 to be significant for these tests. If a data point had a studentized residual over 3 in magnitude, it was considered a possible outlier (Lund 1975). In such cases, results are presented with and without the possible outlier.

**Testing for temporal trends in population relationships**

We conducted forward and backward model selection and used F-tests to address whether the cause of any observed
population decline at a ‘recipient’ stage in the cycle was due
to declining inputs (population index at the ‘donor’ stage), or
due to a decline in the relationship between population indices
at these two stages. We used ‘donor’, ‘year’ and ‘donor-
by-year’ interaction, as predictors. We performed model
selection based on AIC, and used F-tests to assess whether
a potential predictor significantly improves the model’s fit.
Additional details of the analysis and complete output of the
model selection procedure are presented in the Supplementary
material Appendix 1. In four such analyses we detected
a single outlier, and present results with and without the
outlier.

Results

Citizen science data captures the annual migratory cycle

Figure 3B shows a moving-average index of relative popula-
tion size across each year, based on the 6376 records in the
NABA dataset, separated into three geographical regions.
The daily population indices clearly captured the monarch’s
continental migration: a spring wave of population increase
and decrease in the southern USA, followed by a similar but
extended summer pattern in the Midwest and Northeast as
the butterflies move north, and then a fall wave of returning
butterflies in the south flying to Mexico. The trend varies
somewhat (standard errors indicated by shaded regions) but
the major features are consistent across years.

Next we assessed whether the population indices reflect
the known links in the migratory cycle. Despite using dis-
tinct monitoring methods, the overwintering population
index linearly predicted the Spring South population index
based on NABA counts (p < 0.001, adjusted R² = 0.753;
Table 1, Fig. 4A), reflecting the first breeding generation in
the southern USA. The annual Spring South index predicted
the subsequent Midwest and Northeast NABA-based indices
(p = 0.04, adjusted R² = 0.346 and p = 0.06, adjusted
R² = 0.291, respectively; Table 1, Fig. 4B, 4D). Although
there may be spatial biases in these NABA counts (e.g. we
lack data on whether surveys were conducted in crop fields),
the high mobility of monarchs and the large geographic area
of sampling in the Midwest and Northeast (Fig. 3) are likely
to alleviate these issues. The Midwest and Northeast indices
are also correlated with each other (n = 22, r = 0.682,
< 0.001), suggesting that variation in spatial sampling
among NABA volunteers has not greatly biased these
indices.

Links between summer indices and the single-location
counts of fall southward flying migrants again involve dist-
tinct data sets. The NABA-based summer Northeast index
predicted the Cape May, NJ fall migration count (p < 0.01,
adjusted R² = 0.283; Table 1, Fig. 4E). The Midwest index
predicted, although less strongly, the fall migration through
Peninsula Point, MI (p = 0.06, adjusted R² = 0.15 on log-
transformed data; p = 0.05, adjusted R² = 0.17 with one
potential outlier (2014) removed; Fig. 4C). These results are
what we would expect if the fall migration counts at Cape
May and Peninsula Point are representative of the total adult
butterfly source populations in the Northeast and Midwest
regions, respectively. Overall, these statistical linkages
(from Mexico through the fall migration counts) are strong
evidence for data quality and expected demographic links
between Mexico, spring and summer breeding generations,
and fall migration indices.

In contrast, our Fall South index (2002–2014), repre-
senting butterflies returning to Mexico through the southern
USA, showed no relationship with the end of summer indices
(Cape May and Peninsula Point) or with the overwintering
population in Mexico (Table 1, Fig. 4F–H). Given the lack
of relationship with fall South, we also regressed the Mexico
overwintering population index against summer indices
(Midwest, Northeast, Peninsula Point and Cape May) (e.g.,
Fig. 4I; all analyses in Table 1). We used all four indices to
independently test the relationship, yet the only significant

Table 1. Regression analyses between stages of the annual monarch migratory cycle. Statistical significance is indicated by ***p < 0.001,
**p < 0.01, *p < 0.05, and ’p < 0.1 (two-tailed tests).

<table>
<thead>
<tr>
<th>Independent</th>
<th>Dependent</th>
<th>N (years)</th>
<th>Slope</th>
<th>Adj R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring migration</td>
<td>Mexico²</td>
<td>Spring South</td>
<td>10</td>
<td>10.465***</td>
</tr>
<tr>
<td>Summer breeding (first generation)</td>
<td>Spring South</td>
<td>Northeast</td>
<td>10</td>
<td>1.855*</td>
</tr>
<tr>
<td>Summer breeding (up to 3 generations)</td>
<td>Spring South</td>
<td>Midwest</td>
<td>10</td>
<td>2.401*</td>
</tr>
<tr>
<td>Fall migration</td>
<td>Northeast</td>
<td>Cape May</td>
<td>22</td>
<td>5.713**</td>
</tr>
<tr>
<td></td>
<td>Midwest</td>
<td>Peninsula Point</td>
<td>19</td>
<td>0.598¹</td>
</tr>
<tr>
<td></td>
<td>Midwest</td>
<td>Fall South</td>
<td>13</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>Northeast</td>
<td>Fall South</td>
<td>13</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>Peninsula Point</td>
<td>Fall South</td>
<td>13</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>Cape May</td>
<td>Fall South</td>
<td>13</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>Fall South</td>
<td>Mexico</td>
<td>13</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>Midwest</td>
<td>Mexico</td>
<td>22</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Northeast</td>
<td>Mexico</td>
<td>22</td>
<td>-0.010</td>
</tr>
<tr>
<td></td>
<td>Cape May</td>
<td>Mexico</td>
<td>22</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Peninsula Point</td>
<td>Mexico</td>
<td>19</td>
<td>0.001</td>
</tr>
</tbody>
</table>

¹With a single outlying year (2014) removed, this relationship was significant at p = 0.050 (Adj R² = 0.171). The 2014 census for PP was a
statistically significant outlier (studentized residual > 3.1), see Lund 1975.
²Although this is the most direct test of a relationship between overwintering monarch numbers in Mexico and the spring populations in the
USA, it is limited by 10 years of data. If we skip this first breeding generation (Spring South) and examine the 22 year relationship between
Mexico and either the Midwest or Northeast, there is no significant relationship in either case (p > 0.4). The visual suggestion (not shown) of
a hump-shaped relationship in scatterplots of the data is not statistically supported (p > 0.35 in all cases, for quadratic regression or nonpara-
metric spline regression of Midwest or Northeast on the prior Mexico population).
To address this, we examined the temporal trends beginning then, but only found marginally significant declines since 2005 (Adjusted $R^2 = 0.15$, $p = 0.15$ and adjusted $R^2 = 0.31$, $p = 0.055$, for Northeast and Midwest respectively, Fig. 5). Additionally, over the same time period the declines were more severe in Mexico and in the first generation in the south (the fitted linear regressions of population index versus year correspond to decreases of 92.4% over the time period for Mexico, 78.5% for Spring South, 62.1% for Midwest and 64.9% for Northeast), and there was no statistically significant trend in the Cape May, Peninsula Point or Fall South indices. The steep decline in Mexico is reflected in subsequent northern indices, as we expect from the annual cycle, but the impact is progressively attenuated at each step of the cycle (Fig. 5).

The milkweed limitation hypothesis is not supported

None of the four northern indices (or the fall South index) showed a statistically significant decline across the full 22-year period covered by the NABA counts (Fig. 5). This suggests that northern populations were able to build up during the breeding generations, despite the overwinter decline. Admittedly, the Northeast and Midwest indices do show some visual indication of a decline beginning in 2005. To address this, we examined the temporal trends beginning then, but only found marginally significant declines since 2005 (Adjusted $R^2 = 0.15$, $p = 0.15$ and adjusted $R^2 = 0.31$, $p = 0.055$, for Northeast and Midwest respectively, Fig. 5). Additionally, over the same time period the declines were more severe in Mexico and in the first generation in the south (the fitted linear regressions of population index versus year correspond to decreases of 92.4% over the time period for Mexico, 78.5% for Spring South, 62.1% for Midwest and 64.9% for Northeast), and there was no statistically significant trend in the Cape May, Peninsula Point or Fall South indices. The steep decline in Mexico is reflected in subsequent northern indices, as we expect from the annual cycle, but the impact is progressively attenuated at each step of the cycle (Fig. 5).
We next asked if the year-to-year changes in overwintering monarch population estimates are correlated with year-to-year changes in adoption of herbicide-resistant crops. We use differenced data, rather than directly correlating monarch populations with herbicide-resistant crop adoption, because any two variables showing a trend over the same time will be correlated. If the relationship is causal, however, annual differences in adoption should correlate with annual differences in monarch populations. No such correlation is observed (n = 21, r = 0.03, p = 0.897, Supplementary material Appendix 1 Table A1). As a case in point, the peaks of the summer Midwest and Northeast population indices both occurred in 2006, following a 10-year period in which use of herbicide-tolerant crops rose steadily from 0 to 62.5% of corn and soybean acreage (Supplementary material Appendix 1 Table A1).

Where is the break in cycle?

We used a model selection approach to test for temporal trends in the stage-to-stage relationships between population indices (Table 2, Supplementary material Appendix S2). This procedure evaluates whether the population at ‘recipient’ stage (e.g. Midwest) was driven by inputs from the previous ‘donor’ stage (e.g. monarchs migrating from the South), or by ‘year’, which represents an unspecified directional effect on the stage being predicted (e.g. habitat degradation, decline of milkweed, etc.). A significant interaction between ‘year’ and ‘donor’ indicates a relationship changing over time. If milkweed reduction is the main factor in the monarch decline, we expect to see changing relationships in stages where breeding occurs. Instead, we found temporal trends at stages where the population is not building and not dependent on milkweed (Table 2).

Along the northern migration, the ‘donor’ stage was the most important predictor of the ‘recipient’ stage, and the interaction term was at most marginally significant (Table 2, Supplementary material Appendix 2). For example, the Midwest summer population index is best explained by the input population size from Spring South.

Similar results hold for the beginning of the southern migration, when the butterflies fly south through funnel points such as Cape May and Peninsula Point (Supplementary material Appendix 2). However, along the southern migration, ‘year’ is a significant predictor of the Mexican overwintering population, indicating a decline at this stage,
with three out of four northern USA indices failing to explain additional variation in the overwintering population index. The one exception was the relationship between migrants from Peninsula Point and Mexico with an outlier removed. Here, both the ‘donor’ and interaction terms were statistically significant, indicating a declining relationship over time. Importantly, Peninsula Point captures migrating butterflies and therefore a decline in the slope predicting the population size in Mexico is independent of milkweed. Lastly, our analyses show some evidence for an interaction between Fall South and ‘year’ in predicting monarch populations in Mexico. Again, these adult butterflies do not depend on milkweed, suggesting a temporal change in migratory success.

Ries et al. (2015b) found a significant positive correlation between summer peak populations (estimated from NABA counts) and the subsequent overwintering population in Mexico after the latter data were detrended for their annual decline. We therefore tested for associations between our summer indices and detrended Mexico data (i.e. the residuals from the nonlinear trend plotted in Fig. 2). The correlations of Mexico with our Midwest index (which is most similar to the region considered by Ries et al. 2015b) and with Peninsula Point were positive and marginally significant \( r = 0.40, p = 0.07 \) and \( r = 0.37, p = 0.12 \) respectively. In other words, the Midwestern index weakly predicts the numbers arriving to Mexico, but only after the downward trend in Mexico is removed. We found no significant correlations between detrended Mexico data and the Northeast, Cape May or Fall South indices \( (r > 0 \) but \( p > 0.2 \) in all cases). Taken together, our results are consistent with failed migration or re-establishment at the overwintering grounds impacting the population decline in Mexico.

**Discussion**

The current literature on monarch population dynamics and decline is rife with inconsistent patterns and interpretations (Brower et al. 2012b, Davis 2012, Pleasants and Oberhauser 2013, Davis and Dyer 2015, Dyer and Forister 2016). While the monarch population is clearly declining in Mexico, a similar pattern is not observed in many northern regions. We have attempted to make sense of these inconsistencies by connecting the demographic dots of the annual monarch migratory cycle. Although none of the datasets employed is perfect, they represent the bulk of the available data, and we linked them in new ways. Citizen-science data allow for investigations at "large spatial scales, where important processes not detectable at local scales may dominate dynamics" (Dickinson et al. 2010).

Although limited sample size (10–22 years for population indices) and variability in the data limit the statistical power of any one test, we used several different approaches to examine population trends and the milkweed limitation hypothesis. Several trends during the years covered by our data conflict with expectations of the milkweed limitation hypothesis. The lack of an overall decline in the two NABA summer regional indices, and the two fall migratory indices covering the same years, suggest that the milkweed decline is not limiting the production of adult butterflies. Similarly, Crewe and McCracken (2015) found that the fall migrant counts at the Long Point, Ontario funneling point decreased at roughly half the rate of the decline in Mexico. Taken together these results indicate a substantial recovery, during the breeding season, from population bottlenecks. We found no correlation between annual increases in the adoption of herbicide-resistant crops (the hypothesized causal agent of milkweed declines) and annual decreases in the Mexico overwintering population. For the summer population indices, where milkweed limitation should be most evident, monarch populations actually increased substantially over the decade that included 2/3 of the total increase in herbicide-resistant crop acreage (through the mid 2000s).

Our stage-to-stage regression analyses show that monarch population changes are predictable along the annual cycle from Mexico through to the summer breeding grounds; the annual population index at each step reflects the index of the previous step. Furthermore, model selection confirmed that the regional links are sufficient to explain the population dynamics up to the fall migration. For example, the decline in the overwintering population fully accounts for the decline in the first generation in the southern USA. However, there is a break in predictability beginning with the fall migration. That three out of four northern indices fail to predict the numbers arriving in the south conflicts with our most basic expectations about monarch population dynamics and suggests an external factor, but it occurs at a stage when milkweed is not utilized. Badgett and Davis (2015) also hypothesized that diminished fall migration success is an important factor in the overwintering population decline, but this was based solely on the lack of a decline in the Peninsula Point counts, and was predicated on the assumption that Peninsula Point is representative of the total northern breeding population. Our analysis of the successive links across the entire migratory cycle more completely addresses this issue and is concordant with the hypothesis that the population decline is most strongly driven by events after monarchs rely on milkweeds.

Our model selection analysis indicates that over the years, populations of adult monarchs consistently build up during the summer, beginning in the first (southern USA) breeding generation. There is some evidence that the density of monarch eggs has declined since 2007 (Stenoien et al. 2015), and that larval survival has also declined (Nail et al. 2015); these results, however, are inconsistent with the trends in summer adult counts and initial number of fall migrants (Fig. 4). We believe that adult counts are complementary to egg and larval counts, but are more useful in analyses such as ours because the adult stage represents migrants after egg and larval mortality (which is typically very high in the field). Our analysis indicates that an unknown, annually increasing effect, is impacting the monarch population by the time they reach Mexico, producing a consistent decline over the past two decades.

One way in which lack of milkweed could drive monarch declines is if the monarchs that reach Mexico are a small geographical subset of those breeding in the USA during the summer months. It has been suggested that the midwestern USA is the critical area for monarch breeding that populates the overwintering grounds (Pleasants and Oberhauser 2013, Flockhart et al. 2015). Although NABA counts are not...
typically conducted in agricultural fields where milkweed declines are strongest (Oberhauser et al. 2001, Pleasants and Oberhauser 2013), our regional population indices sum over large areas (Fig. 3A), and predict the numbers flying south in the fall that are drawn from all habitats (Fig. 4). Results from stable isotope work and tagging are variable, but indicate that well over half of the monarchs making it to Mexico are derived from outside the agricultural Midwest (Wassenaar and Hobson 1998, Hobson 2008, Flockhart et al. 2013). Few eastern coastal migrants reach the overwintering sites, but non-coastal migrants east and west of the Appalachians are more successful (Brindza et al. 2008, Steffy 2015).

Many factors have been suggested to explain the overwintering population decline of the monarch butterfly (Table 3). Our analyses point to the fall migration and re-establishment on the wintering grounds as key issues. Uncovering the cause of the trends at these stages may be critical to understanding the decline in Mexico. The severe “100-year” drought in Texas (2010–2015) likely had a strong impact on spring and fall migrants, corresponding to the lowest monarch numbers on record (Brower et al. 2015, Zipkin et al. 2012). Factors such as sub-lethal insecticide effects in the breeding grounds (Krischik et al. 2014) or lack of nectar sources during the fall (Brower et al. 2006) may be important in driving a wedge between summer and overwintering populations. Milkweed is typically no longer flowering during the fall migration, and other plant species (many in the Asteraceae) serve as nectar sources. However, the condition of fall migrants might be affected by the environments they experience early in life, including milkweed shortage, insecticides, or other changes in habitat quality.

Other aspects of the breeding and migratory behavior of monarchs are changing, and their roles in population dynamics are unknown. The sex ratio of monarchs at the overwintering sites has changed over the past three decades from 53% female to 43% female (Davis and Rendón-Salinas 2009). As well, the pace of the fall southern migration has sped up over the past decade (Howard and Davis 2015). We currently have few estimates of the sub-lethal impacts of poor quality summer breeding habitat, insecticide residues, and intensified agriculture on the monarchs’ migratory success. Predator, parasitoid, and disease impacts can also be severe (Bradley and Altizer 2005, Oberhauser et al. 2015). Determining the extent to which these and other factors contribute to the dynamics of eastern monarchs is a high priority.

### Conclusion: conservation and controversy

The past two years have seen tremendous media attention and scientific discourse on the population decline of the monarch butterfly (Wagner et al. 2014, Berenbaum 2015, Rubinoff 2015, Shapiro 2015, Pleasants et al. 2016). While there has been consistency in some of the analyses, other research and interpretations has called into question the extent to which we truly understand fluctuations in monarch population sizes, especially given the complex annual migratory cycle (Dyer and Forister 2016, Espeset et al. 2016, Ries et al. 2015b). We hope that our analysis linking the annual population steps has shed light on this important conservation issue, and moreover that our approach will be useful in understanding the similar challenges faced by many long-distance migrants (Wilcove and Wikelski 2008).

The monarch butterfly is far from being threatened, but the eastern USA migration, one of the most spectacular animal migrations in the world, may be an endangered phenomenon (Brower et al. 2012b). To identify and manage the risk factors associated with its decline, deeper critical analyses of the existing data are essential. We do not dispute that milkweed is essential for larval monarchs, and might serve as a buffer against further aggravation. Yet our analyses indicate that other stages are critical, so milkweed conservation alone is unlikely to be sufficient to preserve the migration. Additional resources are necessary to study and improve the transition between summer breeding in the USA and overwintering in their highland forested habitats in Mexico.

### Acknowledgements


### Funding

Funding was entirely provided by the Atkinson Center for a Sustainable Future at Cornell University.

### References


### Table 3. Proposed threats to the sustainability of the eastern monarch butterfly annual migration.

<table>
<thead>
<tr>
<th>Proposed threat</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Habitat destruction/reduced nectar availability on southern migration</td>
<td>Alonso-Mejia et al. 1997, Brower et al. 2006</td>
</tr>
<tr>
<td>Disease, predation, and parasitoids</td>
<td>Bradley and Altizer 2005, Oberhauser et al. 2015</td>
</tr>
<tr>
<td>Climate change/extreme weather</td>
<td>Oberhauser and Peterson 2003, Brower et al. 2012b</td>
</tr>
<tr>
<td>Herbicides/genetically modified herbicide tolerant crops (loss of milkweed)</td>
<td>Zalucki and Lammers 2010, Pleasants and Oberhauser 2013, Flockhart et al. 2015, Krischik et al. 2014</td>
</tr>
<tr>
<td>Insecticides/genetically modified insecticidal crops</td>
<td>McKenna et al. 2001, Guerra et al. 2014</td>
</tr>
<tr>
<td>Automobile accidents, especially during the migration</td>
<td>Casagrande et al. 2014, Batalden and Oberhauser 2015</td>
</tr>
<tr>
<td>Electromagnetic fields/microwave emissions</td>
<td></td>
</tr>
<tr>
<td>Trap plants</td>
<td></td>
</tr>
</tbody>
</table>

**Proposed threat**

Habitat destruction/logging at the overwintering sites
Habitat destruction/reduced nectar availability on southern migration
Disease, predation, and parasitoids
Climate change/extreme weather
Herbicides/genetically modified herbicide tolerant crops (loss of milkweed)
Insecticides/genetically modified insecticidal crops
Automobile accidents, especially during the migration
Electromagnetic fields/microwave emissions
Trap plants

**References**


Gustafsson, K. et al. 2015. The monarch butterfly through time and space: the social construction of an icon. – Bioscience 65: 612–622.


Krischik, V. et al. 2014. Soil-applied imidacloprid translocates to ornamental flowers and reduces survival of adult *Coleomegilla maculata*, *Harmonia axyridis*, and *Hippodamia convergens* lady beetles, and larval *Danaus plexippus* and *Vanessa cardui* butterflies. – PloS ONE 10: e0119133-e0119133.


Slayback, D. and Brower, L. 2007. Further aerial surveys confirm the extreme localization of overwintering monarch butterfly colonies in Mexico. – Am. Entomol. 53: 146–149.


Supplementary material (available online as Appendix oik-03196 at <www.oikosjournal.org/appendix/oik-03196>). Summary of annual data used in analyses (Table A1), analyses examining quality and potential biases in the NABA dataset (Appendix 1), and analyses to examine temporal change in the relationship between stages of the monarch’s annual migratory cycle (Appendix 2) are available online.
Supplementary materials

Table S1: Summary of annual data used in analyses.

Appendix S1: Summary of analyses examining quality and potential biases in the NABA dataset.

Appendix S2: Summary of analyses to examine temporal change in the relationship between stages of the monarch’s annual migratory cycle.
Table S1. A summary of the annual census data used in analyses. All data were compiled, normalized and smoothed from the raw data (see Methods; code provided in Dryad), except that of the last four columns beginning with Mexico.

<table>
<thead>
<tr>
<th>YEAR</th>
<th>Spring</th>
<th>South</th>
<th>Midwest</th>
<th>Northeast</th>
<th>Truncated Midwest</th>
<th>Truncated Northeast</th>
<th>Cape Peninsula</th>
<th>Fall</th>
<th>Mexico Change in monarch population (Mexico)</th>
<th>Average adoption of HT corn &amp; soybean</th>
<th>Change in HT adoption</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>NA</td>
<td>188.3</td>
<td>82.6</td>
<td>70.5</td>
<td>23.4</td>
<td>10.2</td>
<td>39.9</td>
<td>12.8</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1994</td>
<td>NA</td>
<td>226.5</td>
<td>59.7</td>
<td>210.5</td>
<td>39.1</td>
<td>83.9</td>
<td>544.6</td>
<td>NA</td>
<td>0.7</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1995</td>
<td>NA</td>
<td>35.7</td>
<td>43.0</td>
<td>34.4</td>
<td>35.1</td>
<td>248.5</td>
<td>12.6</td>
<td>NA</td>
<td>12.6</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1996</td>
<td>NA</td>
<td>102.2</td>
<td>37.7</td>
<td>61.3</td>
<td>32.9</td>
<td>503.6</td>
<td>104.4</td>
<td>NA</td>
<td>18.2</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1997</td>
<td>NA</td>
<td>230.1</td>
<td>108.2</td>
<td>149.5</td>
<td>70.2</td>
<td>919.6</td>
<td>254.4</td>
<td>NA</td>
<td>5.8</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1998</td>
<td>NA</td>
<td>104.9</td>
<td>40.9</td>
<td>47.7</td>
<td>25.3</td>
<td>403.1</td>
<td>63.5</td>
<td>NA</td>
<td>5.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1999</td>
<td>NA</td>
<td>255.7</td>
<td>104.1</td>
<td>126.9</td>
<td>45.1</td>
<td>2849.2</td>
<td>287.7</td>
<td>NA</td>
<td>8.9</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2000</td>
<td>NA</td>
<td>149.8</td>
<td>80.3</td>
<td>73.5</td>
<td>32.8</td>
<td>250.7</td>
<td>259.5</td>
<td>NA</td>
<td>3.8</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2001</td>
<td>NA</td>
<td>307.8</td>
<td>90.5</td>
<td>141.4</td>
<td>34.4</td>
<td>658.4</td>
<td>421.8</td>
<td>NA</td>
<td>9.4</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2002</td>
<td>NA</td>
<td>166.0</td>
<td>21.4</td>
<td>62.1</td>
<td>8.5</td>
<td>276.8</td>
<td>317.8</td>
<td>35</td>
<td>7.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2003</td>
<td>NA</td>
<td>193.0</td>
<td>41.9</td>
<td>103.5</td>
<td>17.3</td>
<td>392.3</td>
<td>466.9</td>
<td>110.8</td>
<td>1.12</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2004</td>
<td>NA</td>
<td>58.7</td>
<td>16.0</td>
<td>33.4</td>
<td>9.2</td>
<td>74</td>
<td>92.1</td>
<td>NA</td>
<td>2.19</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2005</td>
<td>44.6</td>
<td>163.3</td>
<td>59.0</td>
<td>89.6</td>
<td>20.2</td>
<td>538.2</td>
<td>401.2</td>
<td>NA</td>
<td>5.91</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2006</td>
<td>77.3</td>
<td>338.1</td>
<td>265.5</td>
<td>162.7</td>
<td>120.7</td>
<td>1743.4</td>
<td>56.6</td>
<td>NA</td>
<td>6.87</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2007</td>
<td>72.9</td>
<td>266.0</td>
<td>179.7</td>
<td>159.4</td>
<td>90.5</td>
<td>746</td>
<td>129.4</td>
<td>NA</td>
<td>4.61</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2008</td>
<td>51.3</td>
<td>170.1</td>
<td>132.0</td>
<td>76.1</td>
<td>57.1</td>
<td>265.8</td>
<td>320.0</td>
<td>NA</td>
<td>5.06</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2009</td>
<td>75.3</td>
<td>185.2</td>
<td>88.1</td>
<td>84.4</td>
<td>43.1</td>
<td>392.3</td>
<td>466.9</td>
<td>110.8</td>
<td>1.12</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2010</td>
<td>29.6</td>
<td>306.8</td>
<td>9.5</td>
<td>5.8</td>
<td>51.3</td>
<td>1026.5</td>
<td>624.6</td>
<td>NA</td>
<td>5.06</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2011</td>
<td>34.3</td>
<td>140.4</td>
<td>80.1</td>
<td>76.4</td>
<td>43.1</td>
<td>681.73</td>
<td>108.4</td>
<td>NA</td>
<td>4.61</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2012</td>
<td>20.9</td>
<td>169.6</td>
<td>178.3</td>
<td>89.0</td>
<td>43.1</td>
<td>1222.3</td>
<td>121.7</td>
<td>NA</td>
<td>4.61</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2013</td>
<td>10.3</td>
<td>41.9</td>
<td>16.8</td>
<td>17.2</td>
<td>6.5</td>
<td>112.73</td>
<td>42.4</td>
<td>NA</td>
<td>0.67</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2014</td>
<td>21.1</td>
<td>99.0</td>
<td>46.4</td>
<td>64.9</td>
<td>114.6</td>
<td>1222.3</td>
<td>121.7</td>
<td>NA</td>
<td>4.61</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Appendix S1: Summary of analyses examining quality and potential biases in the NABA dataset

Here we examine potential biases and quality issues common in citizen science datasets [1]. While there are some shortcomings, several lines of evidence and past studies [e.g., 2] suggest that this is a reliable dataset and it is appropriate for our analyses. First, we compared our complete population indices with truncated indices that only included sampling dates that had consistent data across all years. The truncated dataset constitutes a very small portion (20-25%) of the original dataset, yet we see very high correlations between the two (Pearson's $r$ in Midwest: 0.88; Northeast: 0.94). Second, to address the potential for missing data early in the season, we plotted the yearly counts for the Midwest and Northeast to ensure that censuses captured a temporal increase in butterfly abundance in late spring. Third, we addressed the relationship between sampling effort and butterfly counts by transforming party hours to test for sampling effort biases common in citizen science datasets [1]. Fourth, we used Ripley’s $K$ function [3] to assess whether the count data show a temporal bias of increased clustering over years. Finally, the potential for additional spatial biases in sampling are addressed in Results and Discussion in the main article.

A1.1

Description of NABA dataset. The North American Butterfly Association (NABA) has compiled butterfly counts from participating citizens across North America since 1975. The counts are taken from various locations throughout the year and the data includes the number of observed monarchs, the location (latitude and longitude), date, number of observers, number of parties (groups of observers), and the total hours spent. The dataset goes back to 1975 initially as July 4th counts (led by the Xerces Society for Invertebrate Conservation, later acquired by NABA), but the number of sampling dates has been increasing every year, with samples taken more widely throughout the year. The number of counts gradually increased over the years and substantial number of counts were reported 1993-2014 (mean of 290 counts per year across the USA, see Fig A1.1). Furthermore, these years correspond to the data available on the overwintering population in Mexico from the surveys by the WWF.

While the counts originally took place on July 4th, participants started to collect data more widely throughout the year. Figure A1.2 shows the fraction of data points (each colored line represents a year) taken in each month. Northeast and Midwest are concentrated while South has wider sampling range. The two to three key breeding generations during the summer occur in the Midwest and Northeast regions. Although our earliest and latest NABA samples from these regions (across the 22 years in the dataset) were taken from March 27th and October 3rd, respectively, on average there are ~74% of counts in July, with fewer samples in June (~20%) and August (~5%). These months correspond to the peak abundance and breeding period of monarchs [4] (also see Fig 3B). We used March 27th to October 3rd to capture all the information
available on the breeding populations. While these intervals are large, they again capture the regional dynamics (Fig 3B); a smaller subset of the dataset corresponding to the maximum of each peak (and with equal sampling effort across years) is highly correlated with the full dataset (see Section 1 below).
It is important to note that intense sampling does not necessarily correspond to high butterfly counts. As a case in point, the mean relative population size index of the monarchs in the south is lower in the summer compared to spring and fall (Fig. 3B), even though the number of samples are much higher in the summer than either season. Below we address potential issues with varying sampling intensity.

**1. Moving average over large spatial and temporal scale: Will varying intensity cause bias in moving average?**

NABA data points are collected in various locations throughout the USA, with different years of coverage. Furthermore, we see varying sampling intensity within a year. Not surprisingly, we see no obvious population dynamics pattern at fine spatial and temporal scales in the dataset. In order to focus on the appropriate scale that reflects continental population dynamics, we use a moving average (i.e., kernel estimation using uniform function) over 7-day windows. For each observed count within a region, let \( i \) be the day of year, and \( y_i \) the observed number of monarchs per party hour. Then, the averaged abundance assigned to day \( j \) for the specified region is

\[
\bar{y}_j = \frac{1}{n_j} \left( \sum_{i=j-3}^{j+3} y_i \right)
\]

where \( n_j \) is the number of counts that occurred during the 7-day window. If there are several counts on one day, they are both included in the sum. Conversely, a day without any counts within the 7-day window is assigned value 0.

Varying sampling intensity may bias our index, because clustered missing data results in 0, and therefore lowers the index compared to widely sampled years. For example, Figure A1.3
shows the fraction of days in NE and MW where there was at least one data point within each 7-day window; the number of samples increases over time. This varying sampling intensity could bias our results, leading to non-decreasing population index over years. We do not believe this is the case for Spring South, where the population index is decreasing over time; any increase in sampling effort over time would counteract the observed decline. The concern lies in Midwest and Northeast, however, where we see a largely stable population index across years despite decreasing abundance in Mexico. We therefore focus on these two regions for the rest of this Appendix.

To assess this potential bias, we constructed a truncated dataset for each region where the averaged days consistently included a count, across all 22 years; that is, we focused on days where \( n_j > 0 \) across all years (See Fig A1.4 for corresponding dates; the figure shows, for each date, the number of years with a data point in the 7-day window). We summed the indices from these days and compared them to the total Midwest and Northeast population indices derived by our methods.

![Graph A1.5](image)

This reduced the dataset to samples taken from June 13-Aug 1. Importantly, this truncated index is not impacted by varying sampling intensity across years because sampling intensity has been fixed (no days without counts). Our complete yearly index was highly correlated with this truncated index \((n=22, \text{Midwest Pearson's } r=0.88, p<0.001; \text{Northeast Pearson's } r=0.94, p<0.001; \text{see Fig A1.5})\). Furthermore, analyses of linkages between regions and declines were qualitatively the same if we used the yearly index or the truncated index (data provided in Table S1). We therefore conclude that varying sampling intensity across years is not affecting the population indices. Accordingly, to utilize the most available information, we include the complete index from March through October for the main analyses.

2. Census of early season butterflies

To address the potential for missing data early in the season, we plotted the yearly counts for the Midwest and Northeast to ensure that censuses captured a temporal increase in butterfly abundance in late spring. Namely, we were concerned that scarce sampling in some years could have missed some of the early migrating butterflies. In order to check that the incoming butterflies are all taken into account, we plotted the raw counts (i.e., before smoothing via moving average) for the Midwest and Northeast (Fig A1.6). Throughout the panels, the seasonal data sets consistently begin with a low count (~ 0 monarchs per hour) early in the breeding
season, and the values typically increase over time. This suggests that counts began each year
early enough to capture the timing of monarch arrival (which is somewhat variable across years).
Given the consistent sampling coverage within the time of high monarch abundance each year,
we are confident that our indices capture both the migrants and the breeding populations in
Midwest and Northeast.
Midwest

Monarchs/hr vs Date

1993-06-11 to 1993-06-28

DATE 1993

Midwest

Monarchs/hr vs Date

1994-06-13 to 1994-07-10

DATE 1994

Midwest

Monarchs/hr vs Date

1995-06-12 to 1995-07-08

DATE 1995

Midwest

Monarchs/hr vs Date

1996-06-08 to 1996-07-11

DATE 1996

Midwest

Monarchs/hr vs Date

1997-06-14 to 1997-08-12

DATE 1997

Midwest

Monarchs/hr vs Date

1998-06-13 to 1998-08-12

DATE 1998

Midwest

Monarchs/hr vs Date

1999-05-23 to 1999-07-10

DATE 1999

Midwest

Monarchs/hr vs Date

2000-06-05 to 2000-08-27

DATE 2000

Midwest

Monarchs/hr vs Date

2001-06-05 to 2001-08-18

DATE 2001
3. Are their biases in monarch censuses due to varying party hours?

A potential problem with citizen science datasets is variation in survey effort and its non-linear effect on counts (Link and Sauer 1999). As indicated in the Materials and Methods, each NABA count was normalized by dividing the number of observed monarchs by the party hours [5-7]. In some areas of citizen science analysis, as with Christmas bird counts, additional statistical methods have been used to account for potential spatial and temporal effort biases [1, 8]. For example, the number of organisms found may saturate with observation hours. These methods are used to correct for the saturating nature of count data with respect to hours spent. This bias would only appear when effort values are particularly high. Figure A1.7 shows representative graphs (from year 1997 and 2012) of how the number of observed monarchs changes with party hours for the count in both Northeast and Midwest. Specifically, we focused on July (the most intensely sampled month) under the assumption that the population size is more or less the same within a region over a month. We do not see a saturating relationship between sampling effort and butterfly observations. Similar results hold for other years.

In order to further test our dataset, we transformed our party hours to see if it affected the analyses [8, 9]. We re-ran our analyses using counts standardized by the square root of party hours (a simple method of transformation suggested by Link et al. 2006), and the patterns remain the same. Using sqrt(effort) and re-calculating the annual indices, comparisons of the transformed to the original indices yielded $R^2$ values of 0.95 to 0.99 (with the intercepts not being significantly different from zero). Thus, given the linear relationship between effort and monarch counts, the lack of an effect of further transforming the data, and to align with previous analyses [5-7], we maintain using the count data standardized by party hours.

A1.7
4. Do census points cluster more over the years?

If patches of suitable monarch habitat are disappearing (in particular, due to loss of milkweed), then it is conceivable that NABA citizen science counts in later years were done in the few remaining patches, leading to an upward bias in population indices and masking a decline in the total regional population. To test for this possibility, we asked if NABA count locations show increasing spatial clustering in later years, which would occur if the counts are being done in a smaller number of locations. We used Ripley’s $K$ function [3], a standard measure of clustering in spatial statistics, to quantify the clustering of count locations in each year. Ripley’s $K$ function calculates the number of neighboring data points present within concentric circles around a focal sampling location, as the radius/distance increases. These values are averaged over all the sampling locations present in the data set for that year. We used Mercator projection (mapproj library in R) of sampling locations (given as latitude and longitude in the NABA data set) and Ripley’s isotropic correction estimate of $K$ (spatstat library in R).

The patterns are consistent across years in both Northeast and Midwest regions (Fig A1.8, different colors and lines correspond to different years), and do not differ substantially across years. More importantly, we do not see any trends in the $K$ function with respect to year (Fig A1.9) at any spatial scale. This implies that the count locations do not cluster more over time. We conclude that geographic clustering of monarch sampling is not increasing over time, and is therefore not a source of temporal bias in the NABA dataset.

Figure A1.8

Figure A1.8. Ripley’s $K$ function for the spatial locations of NABA population counts in each year.
**Figure A1.9.** Ripley’s K function as a function of year for the Northeast and Midwest regions. The different colors and lines correspond to distances 0.01, 0.02, …, 0.11 from bottom to top.
Appendix S2. Statistical analyses to examine temporal change in the relationship between stages of the annual migratory cycle.

Inamine et al., submitted to Oikos

In the following series of analyses, we investigated the relationship between population size at one stage of the annual migratory cycle (DONOR region, independent variable) and the next time step (RECIPIENT region, dependent variable). To address temporal change in these relationships, we considered YEAR and the DONOR×YEAR interaction as additional covariates. YEAR was entered as a numerical covariate because we are interested in directional trends over time. Because the change in YEAR is small relative to its mean, DONOR and DONOR×YEAR are strongly collinear. To remove this, we centered YEAR about its mean. We considered the following models:

- Model 1: \( \text{RECIPIENT} \sim \text{DONOR} + \text{YEAR} + \text{DONOR}\times\text{YEAR} \)
- Model 2: \( \text{RECIPIENT} \sim \text{DONOR} + \text{DONOR}\times\text{YEAR} \)
- Model 3: \( \text{RECIPIENT} \sim \text{DONOR} \)
- Model 4: \( \text{RECIPIENT} \sim \text{DONOR}\times\text{YEAR} \)
- Model 5: \( \text{RECIPIENT} \sim \text{DONOR} + \text{YEAR} \)
- Model 6: \( \text{RECIPIENT} \sim \text{YEAR} + \text{DONOR}\times\text{YEAR} \)
- Model 7: \( \text{RECIPIENT} \sim \text{YEAR} \)

For each DONOR-RECIPIENT pair, we plot the relationship between regions or between region and year, with the letters on the plot indicating chronological order (a = first year of census, etc.). The table next to the graph shows the \( \Delta \text{AIC} \) value for each model, relative to the lowest AIC value.

We performed stepwise model selection based on AIC values [10], and also F-tests to evaluate the statistical significance of terms by a comparison of nested models with and without the term. We performed both backward and forward selection to check for consistency between these approaches. In backward selection, we started with the full model (Model 1) and sequentially eliminated the non-significant term (if any such exist) that resulted in the largest improvement in AIC, stopping when all terms are significant. In forward selection, we started with either DONOR (Model 3) or YEAR (Model 7), whichever had the stronger univariate correlation with the dependent variable, and sequentially added the term that gave the largest improvement in AIC, stopping when the added term was not statistically significant.

The table below each plot summarizes backward and forward model selection. The entries under Model Comparison in each row show the significance of that covariate, based on an \( F \)-test against a model with that term dropped (for Backward selection) or added (for Forward selection). The AIC of the modified model (with a term added or dropped) is also given. If an outlier was detected, the table reflects the analyses after it was removed.
1 Mexico to Spring South

Backward and Forward model selection both lead to Model 3,

Spring South ~ Mexico

AIC favors the addition of Mexico*YEAR (Model 2), but the $F$-test shows that this term is only marginal ($p = 0.07$) and the residuals from Model 3 (plotted above) do not show any visible pattern over time.

**Conclusion:** The overwintering populations in Mexico predict Spring South populations. There is marginal evidence for a small decrease in the slope of this relationship over time.
2 Spring South to Midwest

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>Model comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backward</td>
<td>Spring South</td>
<td>YEAR</td>
</tr>
<tr>
<td>1 Spring South + YEAR + Spring South*YEAR</td>
<td>91.2</td>
<td>AIC=90.28, p=0.44</td>
</tr>
<tr>
<td>5 Spring South + YEAR</td>
<td>89.22</td>
<td>AIC=88.42, p=0.38</td>
</tr>
<tr>
<td>3 Spring South</td>
<td>87.87</td>
<td><strong>AIC=91.30, p=0.04</strong></td>
</tr>
<tr>
<td>Forward</td>
<td>Spring South</td>
<td></td>
</tr>
<tr>
<td>3 Spring South</td>
<td>87.87</td>
<td>AIC=89.22, p=0.51</td>
</tr>
</tbody>
</table>

Forward selection, Backward selection, and AIC all lead to Model 3,

Midwest $\sim$ Spring South

with the donor region as the only significant predictor ($p < 0.05$).

*Conclusion:* Monarch populations in Spring South significantly predict those in the Midwest. There is no evidence for a temporal trend in this relationship.
3 Spring South to Northeast

Forward selection, Backward selection, and AIC all lead to Model 3, 

Northeast ~ Spring South

with the donor region as the marginally significant predictor ($p = 0.06$).

**Conclusion:** Monarch populations in Spring South marginally predict that in the Northeast. There is no evidence for a temporal trend in this relationship.
# 4 Midwest to Peninsula Point

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>Model comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backward</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Midwest + YEAR + Midwest*YEAR</td>
<td>186.29</td>
</tr>
<tr>
<td>5</td>
<td>Midwest + YEAR</td>
<td>184.35</td>
</tr>
<tr>
<td>3</td>
<td>Midwest</td>
<td>182.41</td>
</tr>
<tr>
<td>Forward</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Midwest</td>
<td>182.41</td>
</tr>
</tbody>
</table>

Forward selection, Backward selection, and AIC all lead to Model 3.

Peninsula Point $\sim$ Midwest

With an outlier (2014: Midwest = 98.8, Peninsula Point = 652.8; Studentized residual >3.1) included, Midwest is not a significant predictor ($p = 0.26$). However with an outlier removed, Midwest becomes a significant predictor ($p < 0.05$). The model selection table reflects the analysis after the outlier was removed.

*Conclusion:* Without an outlier, Midwest monarch populations significantly predict fall migrants through Peninsula Point, and we do not see any signatures of change in the slope over time.
5 Northeast to Cape May

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>Model comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backward</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1   Northeast + YEAR + Northeast*YEAR</td>
<td>236.29</td>
<td>Northeast, YEAR, Northeast*YEAR</td>
</tr>
<tr>
<td>5   Northeast + YEAR</td>
<td>234.36</td>
<td>AIC=254.96, p&lt;0.0001, AIC=233.74, p=0.28</td>
</tr>
<tr>
<td>3   Northeast</td>
<td>233.74</td>
<td>AIC=253.10, p&lt;0.0001, AIC=233.74, p=0.28</td>
</tr>
<tr>
<td>Forward</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3   Northeast</td>
<td>233.74</td>
<td>AIC=234.36, p=0.28, AIC=234.98, p=0.43</td>
</tr>
</tbody>
</table>

Without an outlier (1999: Northeast = 104.1, Cape May = 2849.2; Studentized residual 8.420), Forward selection, Backward selection, and AIC all lead to Model 3,

\[ \text{Cape May} \sim \text{Northeast} \]

When the outlier is included, however, we see marginally significant effect \((p = 0.09)\) of the interaction term (Model 2) with negative slope. The model selection table reflects the analysis after the outlier was removed.

**Conclusion:** Northeast monarch populations predict Cape May, and the weak evidence for a temporal trend was due to a single outlier.
Forward and Backward model selection both lead to Model 7,

\[ \text{Mexico} \sim \text{YEAR} \]

AIC favors the addition of Midwest (Model 5), but this term is not significant \((p = 0.19)\). We had the same result with and without an outlier (1996: Midwest = 102.15, Mexico = 18.19; Studentized residual = 3.93). The model selection table reflects the analysis after the outlier was removed.

**Conclusion:** YEAR is an important predictor of the Mexican overwintering population, and neither Midwest nor the interaction shows statistical significance.
Forward selection, Backward selection, and AIC all lead to Model 7,

\[ \text{Mexico} \sim \text{YEAR} \]

where YEAR is the only significant predictor \( p < 0.001 \).

**Conclusion:** YEAR is an important predictor of the Mexican overwintering population, and neither Northeast nor the interaction shows statistical significance.
8 Peninsula Point to Mexico

With an outlier included, Forward selection, Backward selection, and AIC all lead to Model 7, 
Mexico $\sim$ YEAR

However when an outlier (1996: Peninsula Point = 104.4, Mexico = 18.19; Studentized residual = 4.41) is removed, Forward selection, Backward selection, and AIC all lead to Model 2,

Mexico $\sim$ Pen Point + Pen Point*YEAR

with a negative coefficient for the interaction term ($p < 0.001$) and significant donor region ($p < 0.01$). The model selection table reflects the analysis after the outlier was removed.

Conclusion: With an outlier remove, Peninsula Point predicts Mexico and the relationship changes over time (i.e. the slope decreases over time). This effect cannot be explained by declining milkweed.

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>Peninsula Point</th>
<th>YEAR</th>
<th>Peninsula Point*YEAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backward</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Pen Point + YEAR + Pen Point*YEAR</td>
<td>26.63</td>
<td>AIC=33.43, $p=0.01$</td>
<td>AIC=25.14, $p=0.54$</td>
<td>AIC=30.20, $p=0.04$</td>
</tr>
<tr>
<td>2 Pen Point + Pen Point*YEAR</td>
<td>25.14</td>
<td>AIC=34.69, $p&lt;0.01$</td>
<td></td>
<td>AIC=41.62, $p&lt;0.001$</td>
</tr>
<tr>
<td>Forward</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Pen Point</td>
<td>41.62</td>
<td>AIC=30.2, $p&lt;0.001$</td>
<td>AIC=25.14, $p&lt;0.001$</td>
<td></td>
</tr>
<tr>
<td>2 Pen Point + Pen Point*YEAR</td>
<td>25.14</td>
<td>AIC=26.63, $p=0.54$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
9 Cape May to Mexico

Forward selection, Backward selection, and AIC all lead to Model 7,

\[ \text{Mexico} \sim \text{YEAR} \]

where YEAR is the only significant predictor \((p < 0.001)\).

**Conclusion:** YEAR is an important predictor of the Mexican overwintering population, and neither Cape May nor the interaction shows statistical significance.
10 Fall South to Mexico

AIC leads to Model2, but backward selection shows that Fall South is not significant under the F-test. Forward selection shows that the interaction term is marginally significant even when YEAR is included in the model. Taken together, we infer that

\[
\text{Mexico} \sim \text{Fall South} \times \text{YEAR}
\]

is the best model.

**Conclusion:** Interaction term is an important predictor of the Mexican overwintering population, and neither Fall South nor YEAR shows statistical significance.
Additional references