$Roosts_t$ (r = 0.58), we independently tested their effects and retained the most strongly supported index (i.e. models with lowest Deviance Information Criterion [DIC]; see below) in subsequent models (13).

Autumn climate—We defined two annual temperature variables (average temperature and minimum temperature) at distinct spatiotemporal scales along the autumn migratory route. The time periods and regions of interest were determined via timing and location of sightings by citizen scientists over the 12-year period (10), as well as by monarch phenology (14). We restricted our calculation of autumn environmental variables (climate, NDVI) to within the central flyway (Fig. 2 in main text) because the majority of monarchs in the overwintering grounds likely use this region during migration. A smaller fraction of monarchs originating in the northeast use an eastern corridor to migrate, but they constitute less than a quarter of the overwintering population (5, 9). First, we acquired daily minimum and average temperatures throughout the first half of the migration route (Region 1) between 100°W, 40°N and 90°W, 30°N (Fig. 2 in main text) during 15 Sept – 15 Oct from Daymet (daymet.ornl.gov), which interpolates data from weather stations to produce spatially gridded estimates of daily weather (15). Second, we acquired daily minimum and average temperatures throughout the second half of the migration route (Region 2) between 100°W, 30°N and 95°W, 20°N (Fig. 2 in main text) during 15 Oct – 15 Nov. For both measures, we used daily minimum and average temperatures in a grid of points separated by 1 degree across the specified region, and averaged the values across each region to yield a single mean temperature ($avgTemp.R1_t$, $avgTemp.R2_t$) and minimum temperature ($minTemp.R1_t$, $minTemp.R2_t$) for each year *t* in each region (R1 and R2). Because regional values were highly correlated (r = 0.86 for min temp; 0.88 for mean temp), we independently tested the effects of the two temperature covariates for each region and retained the regional pair of covariates that was most strongly supported in subsequent models.

Normalized Difference Vegetation Index. – To assess the influence of nectar availability along the migration route on arrival abundances at the wintering grounds (and hence the ability to build up lipid reserves for successful migration; ref. 16), we used the Normalized Difference Vegetation Index (NDVI; lta.cr.usgs.gov/noaa_cdr_ndvi), which quantifies the density of green vegetation by calculating the visible and near-infrared light reflected by vegetation. Most of the energy (i.e. lipids) that adult butterflies use to fuel their activity is obtained from flower nectar (17, 18). Monarchs arrive at Mexican wintering sites with high lipid levels, suggesting that they accumulate lipid reserves in the southern U.S. and northern Mexico (14). NDVI is commonly used as an indicator of drought, primary biomass production, and herbivore resource availability on large spatial scales (19-21). Bottom-up effects resulting from drought can greatly reduce Lepidoptera resource availability (22, 23), and the relationship between plants and available moisture has been identified as a primary driver of population dynamics for a number of herbivorous insects (24-26).

In this study, we use the NDVI greenness index as a proxy for vegetation productivity/phenology (i.e. nectar resource availability), as has been done in several prior studies which assess NDVI associations with the distribution, abundance, and migratory condition of nectar-reliant organisms. For example, NDVI has been used as a proxy for: resource availability for butterflies (23); nectar flow for honeybees (27); vegetation heterogeneity in resources for butterfly species (19); and resource availability for migrating hummingbirds (28). NDVI has also been correlated with mosquito abundance (for those species that feed on nectar; 29) and giant honeybee abundance (16). Further, a recent study demonstrated a link between remotely-sensed productivity (EVI, enhanced vegetation index) and nectar abundance using

ground-truthed data from flowers (30). Taken together, these studies indicate that NDVI can accurately represent broad-scale nectar availability in the context of our analyses.

We defined NDVI at the same two spatiotemporal scales as the temperature covariates ($NDVI.R1_t$, $NDVI.R2_t$). We averaged these data across the two regions (Fig. 2 in main text) at monthly timescales during the autumns of 2004 – 2015. Because of a high correlation between regions (r = 0.80), we tested their effects individually and retained the most strongly supported covariate of the two measures. NDVI data were collected from the Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua (Instrument: MOD13Q1)/ Terra (Instrument: MYD13Q1 product) satellite at a 250 m spatial and 8-day temporal resolution.

OE parasitism rate—Monarchs infected by the protozoan parasite *Ophryocystis elektroscirrha* have reduced survival, flight speed, and flight endurance (31). The prevalence of infection in the population accumulates over the course of the breeding season such that monarchs of the final generation that are infected by *OE* are less likely to complete the autumn migration (32, 33). We included the annual proportion of the eastern migratory population (i.e. during the autumn migratory/early wintering period) infected with *OE* ($OEInfect_t$) from 2004 – 2014 (32, 34) with an imputed missing value for 2015 (using the mean value from 2004 – 2014; 35) because data were not available for that year (34).

Forest habitat availability.— To identify changes in forest area and condition between 2004 and 2015, we used a map series (36-38) generated for a long-term land cover monitoring project of the Monarch Butterfly Biosphere Reserve and its surrounding area (286,993 ha). We used maps which show land cover for the years of 2003, 2006, 2009, and 2012 generated at a 1:40,000 scale. To update the map series for 2015, we used Landsat ETM+ images 27/46-47 (glovis.usgs.gov), SPOT images (ERMEX-SPOT), and high resolution GeoEye images (Google Earth Pro). On-screen visual interpretation and map analysis were carried out using ArcGIS 9.0. We extracted polygons of dense forest (canopy cover > 70%, indicative of well-preserved forests) and open forest (canopy cover 40 – 70%) from the available land cover maps. We defined two covariates related to forest habitat availability at each of *j* individual colonies: the amounts of (1) dense forest cover ($DForest_{j,t}$) and (2) dense + open forest cover ($DOForest_{j,t}$) at individual colony sites. Dense forest cover provides the optimal microclimate for wintering monarchs, whereas open forest cover represents habitat that is usable but sub-optimal. Each covariate was measured in hectares and calculated annually around individual colony sites within a 100 ha critical area (i.e. historical perching area, 1994 – 2015) with an additional 500 m out from the perimeter of the critical area (i.e. adjacent areas where monarchs fly in search of water). The 500 m buffer distance was selected based on the average proximity of the critical areas to water, as well as the maximum distance monarchs typically travel from roost edges. To best capture microclimate conditions, critical area ellipses were oriented downhill and accounted for both the microbasin and the known ravine each colony used for movement.

Colony locations are not random over time; rather, monarchs tend to congregate in the same general areas year after year, which are represented by the critical area designations. All ellipses were the same area but differed in shape (Fig. 2 in main text) depending on the spatial distribution of colony location points (39, 40). Because the shape of critical area ellipses varied, the size of the buffer areas also varied depending on the area/perimeter ratio (Fig. 2 in main text), thus accounting for the total potential habitat used by monarchs at each site. Two colony sites (El Calabozo Fracción and Sierra Chincua State) shared an ellipse due to close geographical proximity. Land cover data were not available in every year of the study period, so we used linear extrapolation between available years to obtain missing estimates. Because of a high

correlation (r = 0.85) between $DForest_{j,t}$ and $DOForest_{j,t}$, we independently tested their effects and retained the most strongly supported forest cover covariate in subsequent models.

Previous year dynamics—In addition to wintering habitat availability, we considered two autoregressive covariates in our models: (i) presence/absence of individuals at a colony site ($ColPres_{j,t-1}$) in the previous year (evaluated on the first part of the two-part model; see modeling details below), and (ii) colony size ($ColSize_{j,t-1}$) at each site in the previous year (tested on the second part of the two-part model). We incorporated these variables to test for a potential spatiotemporal effect between colony occurrence and monarch abundances from year to year.

Colony presence within/outside of Monarch Butterfly Biosphere Reserve—We included a site-specific categorical variable indicating whether the colony was located within ($Reserve_j$ = 1, n = 14 sites) or outside ($Reserve_j$ = 0, n = 5 sites) of the Monarch Butterfly Biosphere Reserve. The reserve is known to hold the largest aggregations of monarchs, and therefore the majority (~70%) of the wintering population (39, 40).

Rationale for use of gamma zero-altered hurdle model

Two-part (or 'zero-altered hurdle') models applied to ecological and biological questions are useful when observational data show overdispersion and excessive zero values (41-45; Fig. S4). These models are referred to as 'hurdle' models because, regardless of the mechanisms causing an increase in the response variable, a hurdle must first be crossed before the data are observed (46, 47). Ecologically, it is relevant to consider these two processes separately because predictors that determine the presence-absence of colonies can be different from those describing abundance (conditional on presence). Statistically, ignoring the large number of non-detections (i.e. areas occupied < 0.01 ha) is problematic as it could result in exaggerated estimates of variance and biased estimates of parameters and standard errors (48, 49).

We chose a zero-altered model, as opposed to a zero-inflated mixture model, to deal with the high number of zeros in the data because we were interested in the probability of not measuring any detectable colony/area occupied versus measuring any size colony/area occupied (i.e. probability of occurrence). In contrast, the aim of zero-inflated mixture models is to discriminate between false and true zeros (i.e. count process allows for zeros). When modeling continuous data that has too many zeros, a distribution with inflated error is needed. However, the gamma distribution, which accurately captures the long tail in the colony size data, does not allow for zero values, so modeling the zeros separately from the non-zeros in a binomial-gamma hurdle model is recommended (41, 42). To account for pseudoreplication of sites, as well as any unaccounted for site-specific variation, we included colony site as a random effect on both parts of the hurdle model (46).

We estimated parameter values for all models using a Bayesian approach with JAGS (50) called from program R (R package jagsUI; 51) using flat normal priors on all of the parameter values (see below for model code). We ran three chains for 150,000 iterations after a burn-in of 100,000 iterations and adaptation phase of 5,000 iterations, and thinned the chains by five. Model convergence was assessed with the Rhat statistic (52) and visual inspection of chains.

Additional details on model selection procedure and results

We started our analyses with 14 environmental variables described in Table S1. Given the high correlation in many variables (see covariate descriptions), we used a forward selection approach to model fitting. In the first step, we evaluated support for univariate models using the Deviance Information Criterion (DIC). While DIC has limitations, model selection and parameter

estimation with DIC have been effectively applied in hierarchical analyses and demonstrated to be useful for inference (53, 54). To the null model (which included random site effects), we added each of the covariates individually and discarded any uninformative covariates that led to increased DIC values after each step. If multiple covariates yielded a reduction in DIC compared to the null model, the model with the lowest DIC was used as a base model for considering additional covariates. If two or more covariates were selected, we also considered biologically reasonable interaction terms (i.e. within the same spatial scale; see below). When additional covariates no longer led to a reduction in DIC, the best-supported model from the previous step was retained as the top-supported model.

Of the 14 covariates (i.e. univariate models) tested, the following eight covariates lowered DIC as compared to the null model: peak summer abundance index ($NABA_t$), autumn migratory roost index ($Roosts_t$), greenness in Region 1 (see Fig. 2 in main text) during the first half of migration ($NDVI.R1_t$), greenness in Region 2 (see Fig. 2 in main text) during the second half of migration ($NDVI.R2_t$), whether the colony was located within the Monarch Butterfly Biosphere Reserve ($Reserve_j$), average autumn temperature in Region 1 during the first half of migration ($avgTemp.R1_t$), dense forest cover surrounding colony locations ($DForest_{j,t}$), and dense + open forest cover surrounding colony locations ($DOForest_{j,t}$).

Given that $NDVI.R1_t$ and $NDVI.R2_t$ were highly correlated with each other ($r = 0.80$), we retained $NDVI.R1_t$ in all subsequent models because it led to a greater reduction in DIC (ΔDIC = 8.1). However, support for both greenness indices (proxy for nectar availability) indicate an association between nectar availability throughout the migration route with arrival abundances on overwintering grounds. Similarly, $DForest_{j,t}$ and $DOForest_{j,t}$ were strongly correlated ($r = 0.85$); $DForest_{j,t}$ led to a greater reduction in DIC (ΔDIC = 6.3) so was retained in all subsequent models. Stronger support for $DForest_{j,t}$ over $DOForest_{j,t}$ in our models reinforces previous evidence that forest thinning should be prevented within and adjacent to overwintering sites to minimize exposure of butterflies to inclement conditions that increase winter mortality. Finally, $NABA_t$ and $Roosts_t$ were both supported yet were correlated ($r = 0.58$); $NABA_t$ was retained in subsequent models given that it led to a greater reduction in DIC (ΔDIC = 13.2). Support for both summer/autumn indices provides further evidence of a link in population dynamics across the autumn period, although the weaker support for $Roosts_t$ over $NABA_t$ suggests that our roost index does not fully capture the autumn migratory population (i.e. the index may be too coarse).

We carried the five remaining supported covariates forward in selection of multivariate models (ensuring that correlated predictors were not included in a model simultaneously) and tested two-way interaction terms between covariates within the same spatial scale that were included in the same model. Thus, we tested the following interactions: $DForest_{j,t} \times Reserve_j$, $NABA_t \times NDVI.R1_t$, $NABA_t \times avgTemp.R1_t$, and $NDVI.R1_t \times avgTemp.R1_t$. Inclusion of these interaction terms did not yield models with lower DIC values, so interactions were not included in subsequent models. The ten most strongly supported models are provided in Table S2; note that the top model, which included only main effects of $NABA_t$, $NDVI.R1_t$, $Reserve_j$, and $DForest_{j,,}$ clearly provides the best fit to the data (with ΔDIC = 4.5 for the second ranked model).

Results on colony occurrence probabilities

The probability of annual monarch colony occurrence at each of the 19 overwintering sites ranged between 0.15 and 0.91 on average (Table S3). Mean probabilities of colony occurrence were highest at El Rosario, San Antonio Albarranes, and San Mateo Almomoloa (Table S3).

Mean colony sizes, conditional on occurrence, ranged between 0.01 and 1.25 ha for average covariate values, with the largest expected colony sizes at El Rosario (see Fig. 1 in main text).

Models fit with additional years of data at the aggregate colony level

To ensure that our estimated relationships with autumn temporal covariates in our top-supported model (i.e. NDVI, dense forest cover, NABA counts) were not a product of beginning our analysis in 2004, we fit a gamma mixed model using the aggregated colony sizes (i.e. total overwintering abundance) during 2000 (the first year NDVI data are available) to 2015 as a function of annual measures of NDVI, total dense forest cover (summed across all colony sites), NABA counts, and a random effect of year (to account for possible dependence due to temporal autocorrelation). Fig. S2 illustrates the estimated relationships between each variable and total overwintering population size – all of which follow the same trend as that of our colony-level analysis beginning in 2004. Parameter estimates for each covariate were in the same direction, albeit the effect sizes were smaller and 95% credible intervals overlapped zero (see Table S4 below). We fit the same aggregate model starting with data from 2004 to demonstrate that these species-environment relationships exist at both fine and broad spatial scales (Table S4).

Comparison of linear trends and total variation in residuals among models

We calculated residuals to evaluate model fit and to examine any trends in the residuals. Comparison of trends in the residuals (via post-hoc Bayesian linear regression with year as the explanatory variable) allowed us to assess how well our models fit over time. Comparison of variation (standard deviation [SD]) in residuals among models allowed us to assess the amount of variation explained by the predictors in each model. Greater variation in the residuals during a given time period (i.e. higher SD) represents greater variability in the response (i.e. more remaining unexplained temporal variation), indicating that the predictors in the model poorly explain fluctuations in colony sizes. If the residuals are minimally variable (i.e. lower SD), the predictors account for more of the temporal variation in colony sizes. For example, there is a much greater spread (i.e. higher SD) of residual values in the bottom row of Fig. S3 versus the top row.

We fit models using aggregate colony data starting in 2000 (the first year NDVI data are available) and 2004 (timeframe of colony-level analysis). We included the top-supported temporal covariates from our colony-level analysis in these models (peak summer index, greenness index in region 1, dense forest cover summed across colonies, and a random effect of year). Posterior means of residuals from these two models exhibited slight declines over time (top row of Fig. S3), whereas mean residuals from a model fit with only peak summer index (since 2000 and 2004) exhibited statistically significant negative trends (bottom row of Fig. S3). In comparison, mean residuals from our colony-level top model do not show a significant trend over time (Fig. S1), demonstrating adequate fit with the fine-scale model throughout the timeframe of our analysis.

Although our colony-level model underestimated certain colony sizes (i.e. the consistently large colonies) early in the time series, our fine-scale model has the lowest variation (SD range: 0.03 – 0.57) of mean residual values compared to the aggregate models. Note that we compare the mean and range of the 19 colony-specific SDs from our colony-level model (Fig. S1) with the SDs from the aggregate models for a more equivalent comparison than simply the SD of all 228 mean residual values. These results suggest that (i) summing colony-level data masks important local-scale variation and (ii) accounting for spatial heterogeneity of individual

colony sizes is key to explaining temporal variation in monarch abundances. Importantly, failure to consider effects of both autumn greenness (i.e. proxy for broad-scale nectar availability) and dense forest cover on winter population sizes results in significantly more unexplained temporal variation in residuals compared to summer-only models.

SI References

1. Brower L (1985) New perspectives on the migration biology of the monarch butterfly, *Danaus plexippus. Migration: Mechanisms and Adaptive Significance.* Contributions in Marine Science 27: 749-785.

2. Cockrell BJ, Malcolm SB & Brower LP (1993) in *Biology and conservation of the monarch butterfly,* Natural History Museum of Los Angeles County, Science Series, 38.

3. Malcolm, SB, *et al* (1993) Spring recolonization of eastern North America by the monarch butterfly: successive brood or single sweep migration? *Biology and conservation of the monarch butterfly.* Natural History Museum of Los Angeles County, pp. 253–267.

4. Wassenaar LI & Hobson A (1998) Natal origins of migratory monarch butterflies at wintering colonies in Mexico: New isotopic evidence. *Proc Natl Acad Sci USA* 95(26): 15436-15439.

5. Flockhart D, *et al* (2017) Regional climate on the breeding grounds predicts variation in the natal origin of monarch butterflies overwintering in Mexico over 38 years. *Global Change Biol* 23(7): 2565-2576.

6. Vidal O & Rendón-Salinas E (2014) Dynamics and trends of overwintering colonies of the monarch butterfly in Mexico. *Biol Conserv* 180: 165-175.

7. Inamine H, Ellner SP, Springer JP & Agrawal AA (2016) Linking the continental migratory cycle of the monarch butterfly to understand its population decline. *Oikos* 125(8): 1081-1091.

8. Swengel AB (1990) Monitoring butterfly populations using the Fourth of July butterfly count. *Am Midl Nat* 124(2): 395-406.

9. Howard E & Davis AK (2009) The fall migration flyways of monarch butterflies in eastern North America revealed by citizen scientists. *J Insect Conserv* 13(3): 279-286.

10. Howard E & Davis AK (2011) A simple numerical index for assessing the spring migration of monarch butterflies using data from Journey North, a citizen-science program. *J Lepidop Soc* 65(4): 267-270.

11. Howard E & Davis AK (2015) Investigating long-term changes in the spring migration of monarch butterflies (Lepidoptera: Nymphalidae) using 18 years of data from Journey North, a citizen science program. *Ann Entomol Soc Am* 108(5): 664-669.

12. Oberhauser K, *et al* (2017) A trans-national monarch butterfly population model and implications for regional conservation priorities. *Ecol Entomol* 42(1): 51-60.

13. Spiegelhalter DJ, Best NG, Carlin BP & Linde A (2014) The deviance information criterion: 12 years on. *J R Stat Soc* 76(3): 485-493.

14. Brower LP, *et al* (2015) Effect of the 2010–2011 drought on the lipid content of monarchs migrating through Texas to overwintering sites in Mexico. *Monarchs in a Changing World: Biology and Conservation of an Iconic Butterfly.* Cornell University Press: 117-129.

15. Thornton PE, *et al* (2014) Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2. *Oak Ridge National Laboratory Distributed Active Archive Center*, http://daac.ornl.gov.

16. Alonso-Mejía A, Rendon-Salinas E, Montesinos-Patiño E & Brower LP (1997) Use of lipid reserves by monarch butterflies overwintering in Mexico: Implications for conservation. *Ecol Appl* 7(3): 934-947.

17. Tooker JF, Reagel PF & Hanks LM (2002) Nectar sources of day-flying Lepidoptera of central Illinois. *Ann Entomol Soc Am* 95(1): 84-96.

18. Brower LP, Fink LS & Walford P (2006) Fueling the fall migration of the monarch butterfly. *Integrative and Comparative Biology* 46(6): 1123-1142.

19. Kumar S, Simonson SE & Stohlgren TJ (2009) Effects of spatial heterogeneity on butterfly species richness in rocky mountain national park, CO, USA. *Biodivers Conserv* 18(3): 739-763.

20. Teitelbaum CS, *et al* (2015) How far to go? Determinants of migration distance in land mammals. *Ecol Lett* 18(6): 545-552.

21. van Gils JA, *et al* (2016) Body shrinkage due to arctic warming reduces red knot fitness in tropical wintering range. *Science* 352(6287): 819-821.

22. Fleishman E, Mac Nally R & Murphy DD (2005) Relationships among non-native plants, diversity of plants and butterflies, and adequacy of spatial sampling. *Biol J Linn Soc* 85(2): 157-166.

23. Seto K, Fleishman E, Fay J & Betrus C (2004) Linking spatial patterns of bird and butterfly species richness with landsat TM derived NDVI. *Int J Remote Sens* 25(20): 4309-4324.

24. Price PW & Hunter MD (2005) Long-term population dynamics of a sawfly show strong bottom-up effects. *J Anim Ecol* 74(5): 917-925.

25. Pollard E (1977) A method for assessing changes in the abundance of butterflies. *Biol Conserv* 12(2): 115-134.

26. Stevens SR & Frey DF (2010) Host plant pattern and variation in climate predict the location of natal grounds for migratory monarch butterflies in western North America. *J Insect Conserv* 14(6): 731-744.

27. Nightingale JM, Esaias WE, Wolfe RE, Nickeson JE & Ma PL (2008) Assessing honey bee equilibrium range and forage supply using satellite-derived phenology. *IGARSS IEEE International Symposium* 3 doi: 10.1109/IGARSS.2008.4779460.

28. Graham CH, *et al* (2016) Winter conditions influence biological responses of migrating hummingbirds. *Ecosphere* 7(10): e01470.

29. Kumi-Boateng B, Stemn E & Mireku-Gyimah D (2015) Modelling of malaria risk areas in Ghana by using environmental and anthropogenic Variables–A spatial multi-criteria approach. *Ghana Mining Journal* 15(2): 1-10.

30. Feldman RE & McGill BJ (2014) How important is nectar in shaping spatial variation in the abundance of temperate breeding hummingbirds?. *J Biogeogr* 41(3): 489-500.

31. Altizer SM & Oberhauser KS (1999) Effects of the protozoan parasite *Ophryocystis elektroscirrha* on the fitness of monarch butterflies (*Danaus plexippus*). *J Invertebr Pathol* 74(1): 76-88

32. Bartel RA, Oberhauser KS, De Roode JC & Altizer SM (2011) Monarch butterfly migration and parasite transmission in eastern North America. *Ecology* 92(2): 342-351.

33. Bradley CA & Altizer S (2005) Parasites hinder monarch butterfly flight: Implications for disease spread in migratory hosts. *Ecol Lett* 8(3): 290-300.

34. Thogmartin WE, *et al* (2017) Monarch butterfly population decline in North America: Identifying the threatening processes. *Royal Soc Open Sci* 4(9): 170760.

35. Kéry M & Royle JA (2016) *Applied hierarchical modeling in ecology.* Academic Press, London, UK.

36. Ramírez MI, Miranda R, Zubieta R & Jiménez M (2007) Land cover and road network map for the Monarch Butterfly Biosphere Reserve in Mexico, 2003. *Journal of Maps* 3(1): 181-190.

37. Ramírez MI, Sáenz-Romero C, Rehfeldt GE & Salas-Canela L (2015) Threats to the availability of overwintering habitat in the Monarch Butterfly Biosphere Reserve: Land use and climate change. *Monarchs in a Changing World: Biology and Conservation of an Iconic Butterfly*. Cornell University Press: 157-168.

38. Ramírez M & Zubieta R (2005) Análisis regional y comparación metodológica del cambio en la cubierta forestal en la reserva de la biosfera mariposa monarca. *Reporte Técnico Preparado Para El Fondo Para La Conservación De La Mariposa Monarca, México.*

39. García-Serrano E, Reyes JL & Alvarez BXM (2004) Locations and area occupied by monarch butterflies overwintering in Mexico from 1993 to 2002. *The Monarch Butterfly: Biology and Conservation*. Cornell University Press: 129-133.

40. Vidal O, López-García J & Rendón-Salinas E (2014) Trends in deforestation and forest degradation after a decade of monitoring in the Monarch Butterfly Biosphere Reserve in Mexico. *Conserv Biol* 28(1): 177-186.

41. Thorson JT, Shelton AO, Ward EJ & Skaug HJ (2015) Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for west coast groundfishes. *ICES J Mar Sci* 72(5): 1297-1310.

42. Brilleman SL, Crowther MJ, May MT, Gompels M & Abrams KR (2016) Joint longitudinal hurdle and time-to-event models: An application related to viral load and duration of the first treatment regimen in patients with HIV initiating therapy. *Stat Med* 35(20): 3583-3594.

43. Chen Y, *et al* (2016) Predicting hotspots of human-elephant conflict to inform mitigation strategies in Xishuangbanna, southwest China. *PloS One* 11(9): e0162035.

44. Muir JE, *et al* (2016) Gray whale densities during a seismic survey off Sakhalin Island, Russia. *Endangered Species Research* 29(3): 211-227.

45. Taranu ZE, Gregory-Eaves I, Steele RJ, Beaulieu M & Legendre P (2017) Predicting microcystin concentrations in lakes and reservoirs at a continental scale: A new framework for modelling an important health risk factor. *Global Ecol Biogeogr* 26(6): 625-637.

46. Zuur A, Ieno E, Walker N, Saveliev A & Smith G (2009) *Mixed effects models and extensions in ecology with R.* Spring Science and Business Media.

47. Zuur AF & Ieno EN (2016) *Beginner's Guide to Zero-inflated Models with R.* Highland Statistics Limited.

48. Lachenbruch PA (2001) Comparisons of two-part models with competitors. *Stat Med* 20(8): 1215-1234.

49. Moulton LH, Curriero FC & Barroso PF (2002) Mixture models for quantitative HIV RNA data. *Stat Methods Med Res* 11(4): 317-325.

50. Plummer M (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International workshop on Distributed Statistical Computing.* DSC: 20-22.

51. Kellner K (2015) JagsUI: A wrapper around rjags to streamline JAGS analyses. *R Package Version* 4.2.0, https://github.com/kenkellner/jagsui

52. Gelman A & Hill J (2006) *Data analysis using regression and multilevel/hierarchical models.* Cambridge University Press.

53. Moore JE & Barlow J (2011) Bayesian state-space model of fin whale abundance trends from a 1991–2008 time series of line-transect surveys in the California current. *J Appl Ecol* 48(5): 1195-1205.

54. Sauer JR, Zimmerman GS, Klimstra JD & Link WA (2014) Hierarchical model analysis of the Atlantic flyway breeding waterfowl survey. *J Wildl Manage* 78(6): 1050-1059.

R/JAGS code for the most strongly supported hierarchical gamma hurdle model

```
#-------------------------------------------------------------------
# Hierarchical zero-altered gamma hurdle model for estimating
# overwintering monarch butterfly colony sizes at 19 sites in central
# Mexico during December 2004 – 2015.
# Author of code: Sarah Saunders (2017 - 2018)
#-------------------------------------------------------------------
sink("fall_model_best")
cat("
model{

# Priors
a1  ~ dnorm(0,0.01)
a2  ~ dnorm(0,0.01)
a3  ~ dnorm(0,0.01)
a4  ~ dnorm(0,0.01)
a5  ~ dnorm(0,0.01)

tau.a8 <- pow(sigma.a8, -2)
sigma.a8 ~ dunif(0, 10000)
tau.a9 <- pow(sigma.a9, -2)
sigma.a9 ~ dunif(0, 10000)

g0 ~ dnorm(0, 0.01)
sd ~ dgamma(2,2)

# For the ones trick
C <- 100000

for (j in 1:usite){
    # Random effect of site on both model parts
    a8[j] ~ dnorm(0, tau.a8)
    a9[j] ~ dnorm(0, tau.a9)

for (t in 1:uyear){
    # Define logistic regression model, w is probability of occurrence
    # Use the logistic transformation exp(z)/(1 + exp(z))
    logit(w[j,t]) <- zeta[j,t]
    zeta[j,t] <- g0 + a9[j]

    # Define gamma regression model for the mean using inverse link
    mu[j,t] <- pow(eta[j,t], -1)
    eta[j,t] <- a1 + a2*naba.st[t] + a3*nectar.st1[t] +
a4*forestd.st[j,t] + a5*reserve[j] + a8[j]

    # Redefine mu & sd of continuous part into shape & rate parameters
    shape[j,t] <- pow(mu[j,t], 2) / pow(sd, 2)
    rate[j,t] <- mu[j,t] / pow(sd, 2)

    # For readability, define log-likelihood of gamma
    logGamma[j,t] <- log(dgamma(y[j,t], shape[j,t], rate[j,t]))
```

```
    # Define total likelihood, where likelihood is (1 - w) if y <
    # 0.0001 (z = 0) or likelihood is w * gammalik if y >= 0.0001 (z =
    # 1). So if z = 1, then first part must be 0 and second part must
    # be 1. Use 1 - z, which is 0 if y > 0.0001 and 1 if y < 0.0001.
    # Z matrix is data and consists of dummy variable of 0s where y >
    # 0.0001 and 1s where y > 0.0001.

    logLik[j,t] <- (1 - z[j,t])*log(1 - w[j,t]) + z[j,t]*(log(w[j,t])
+ logGamma[j,t])

    Lik[j,t] <- exp(logLik[j,t])

    # Use the ones trick (matrix of ones fed in as data)
    p[j,t] <- Lik[j,t] / C
    ones[j,t] ~ dbern(p[j,t])
  }
 }
}
sink()

## PREP JAGSUI DATA ##
bugsdata <- list(uyear=length(uyear), usite=length(usite),
y=data.matrix(fallmonarchs), z=data.matrix(non_zero),ones=ones,
naba.st=naba.st, nectar.st1=nectar.st1, reserve=reserve,
forestd.st=forestd.st)

inits <- function(){
list(a1=runif(1,1,10),a2=rnorm(1),a3=rnorm(1),a4=rnorm(1),a5=rnorm(1),
g0=rnorm(1))
}

parameters<-c('a1', 'a2', 'a3', 'a4', 'a5', 'g0', 'sigma.a8',
'sigma.a9', 'w')

## RUN BUGS MODEL IN JAGSUI ##
fall.model.best <- jags(data = bugsdata, inits = inits,
parameters.to.save = parameters, model.file = 'fall_model_best.txt',
n.chains = 3, n.adapt = 5000, n.iter = 150000, n.burnin = 100000,
n.thin = 5, parallel = TRUE, store.data = TRUE)
```

Table S1. Continental- and local-scale environmental variables evaluated in our models.

Covariate category	Covariate abbreviation	Description
Continental-scale		
Monarch population size	$NABA_t$	North American Butterfly Association index of monarch counts from 19 July – 15 Aug in Midwestern U.S. [peak summer population]
	$Roosts_t$	Journey North index of roosts migrating through central flyway from 15 Aug – 31 Oct [autumn migratory population]
Autumn climate	$avgTemp.R1_t$	Daily mean temperatures averaged in Region 1 (Fig. 2) during 15 Sept – 15 Oct [first half of migration]
	$avgTemp.R2_t$	Daily mean temperatures averaged in Region 2 (Fig. 2) during 15 Oct – 15 Nov [second half of migration]
	$minTemp.R1_t$	Daily minimum temperatures averaged in Region 1 (Fig. 2) during 15 Sept – 15 Oct
	$minTemp.R2_t$	Daily minimum temperatures averaged in Region 2 (Fig. 2) during 15 Oct – 15 Nov
Greenness index	$NDVI.R1_t$	NDVI measured in Region 1 (Fig. 2) during 15 Sept – 15 Oct [proxy of nectar availability]
	$NDVI.R2_t$	NDVI measured in Region 2 (Fig. 2) during 15 Oct – 15 Nov
Parasitism	$OEInfect_t$	Proportion of larvae infected with protozoan parasite *Ophryocystis elektroscirrha*
Local-scale		
Previous year dynamics	$ColPres_{j,t-1}$	Presence/absence of a colony at each site in previous year
	$ColSize_{j,t-1}$	Size of a colony at each site in previous year
Forest habitat availability	$DForest_{j,t}$	Amount of dense forest (> 70% canopy cover) surrounding each colony, measured within 100 ha ellipse + 500 m buffer
	$DOForest_{j,t}$	Amount of dense + open forest (40 – 70% cover) surrounding each colony, measured within 100 ha ellipse + 500 m buffer
Reserve location	$Reserve_j$	Indicator of whether a colony was located within (1) or outside of (0) the Monarch Butterfly Biosphere Reserve

Subscripts *t* and *j* refer to year and colony site, respectively; covariate abbreviations in main text.

Table S2. The ten most strongly supported models of winter monarch butterfly colony sizes at all known sites in central Mexico during December 2004 – 2015. The null model is also shown.

Model	DIC	ΔDIC	Number of parameters
$NABA_t + NDVI.R1_t + DForest_{j,t} + Reserve_j$	5457.8	0.0	6
$NABA_t + NDVI.R1_t + DForest_{j,t}$	5462.3	4.5	5
$NABA_t + NDVI.R1_t + Reserve_j$	5462.8	5.0	5
$NABA_t + NDVI.R1_t$	5464.1	6.3	4
$NABA_t + DForest_{j,t}$	5467.4	9.6	4
$NABA_t + NDVI.R1_t + avgTemp.R1_t$	5467.5	9.7	5
$NABA_t + Reserve_j$	5467.9	10.1	4
$NABA_t + NDVI.R1_t + DForest_{j,t} + avgTemp.R1_t$	5468.1	10.3	6
$NABA_t$	5468.4	10.6	3
$NABA_t + avgTemp.R1_t$	5472.1	14.3	4
Null model	5500.8	43.0	2

Models were ranked according to differences in the Deviance Information Criterion (ΔDIC). All covariates shown were included on the gamma submodel only. Models also included an intercept term and colony site as a random effect on each part of the hurdle model (i.e. logistic and gamma submodels). See Table S1 for covariate descriptions.

Table S3. Parameter estimates from the most strongly supported model estimating winter monarch butterfly colony size at all 19 sites in central Mexico during 2004 – 2015.

Parameter	Mean	95% CI	85% CI	50% CI
NABA [gamma submodel]	-0.423	-0.690, -0.217	-0.607, -0.263	-0.498, -0.339
NDVI [gamma]	-0.351	-0.647, -0.056	-0.563, -0.138	-0.448, -0.252
Dense forest [gamma]	-0.530	-1.332, 0.145	-1.087, -0.028	-0.757, -0.277
Reserve [gamma]	-1.389	-2.772, -0.084	-2.388, -0.402	-1.827, -0.938
Intercept [gamma]	4.389	3.221, 5.780	-	-
Random site effect SD [gamma]	0.938	0.527, 1.578	-	-
Intercept [logistic submodel]	0.108	-0.757, 1.044	-	-
Random site effect SD [logistic]	1.701	0.984, 2.810	-	-
w E. Contepec	0.217	0.051, 0.456	-	-
w IC Carpinteros	0.647	0.386, 0.866	-	-
w Sierra Chincua Fed	0.287	0.091, 0.540	-	-
w Sierra Chincua State	0.358	0.137, 0.618	-	-
w E. Cerro Prieto	0.502	0.251, 0.751	-	-
w El Calabozo Fracción	0.218	0.052, 0.461	-	-
w IC Crescencio Morales	0.288	0.091, 0.542	-	-
w E. Nicolás Romero	0.358	0.138, 0.619	-	-
w E. El Rosario	0.914	0.729, 0.996	-	-
w E. La Mesa	0.429	0.191, 0.687	-	-
w E. San Juan Xoconusco	0.359	0.137, 0.619	-	-
w El Capulín	0.575	0.319, 0.812	-	-
w Mesas Altas de Xoconusco	0.151	0.022, 0.372	-	-
w IC San Pablo Malacatepec	0.358	0.136, 0.615	-	-
w PP San Andrés	0.647	0.385, 0.867	-	-
w Río de Parras	0.647	0.387, 0.866	-	-
w San Francisco Oxtotilpan	0.718	0.465, 0.913	-	-
w San Antonio Albarranes	0.913	0.729, 0.996	-	-
w San Mateo Almomoloa	0.913	0.728, 0.996	-	-

The 95% credible intervals (CI) for all parameters are shown; 85% and 50% CIs for covariate effects on gamma submodel also shown for reference. NABA = $NABA_t$, NDVI = $NDVI.R1_t$, Dense forest = $DForest_{j,t}$, Reserve = $Reserve_j$, w_j = probability of colony occupancy at each site *j* (estimated from logistic submodel), SD = standard deviation; see Table S1 for covariate descriptions. Negative mean values indicate a positive effect due to the use of an inverse-log link function for the gamma distribution abundance submodel (i.e. all four covariates in the top-supported model had a positive association with winter colony sizes).

Table S4. Parameter estimates from the most strongly supported model (with random year effect) fit using total winter monarch butterfly colony sizes annually (i.e. aggregated colony-level data) in central Mexico during 2000 – 2015 and 2004 – 2015.

Parameter	Mean	95% CI	85% CI
2000 – 2015			
NABA	-0.075	-0.181, 0.004	-0.145, -0.017
NDVI	-0.086	-0.225, 0.015	-0.173, -0.010
Dense forest	-0.119	-0.227, -0.036	-0.191, -0.057
Intercept	0.321	0.240, 0.431	0.258, 0.394
Random year effect SD	0.107	0.030, 0.224	0.046, 0.181
2004 – 2015			
NABA	-0.174	-0.322, -0.070	-0.270, -0.095
NDVI	-0.180	-0.439, 0.054	-0.359, -0.006
Dense forest	-0.215	-0.556, 0.066	-0.437, -0.021
Intercept	0.381	0.197, 0.609	0.250, 0.522
Random year effect SD	0.091	0.007, 0.282	0.019, 0.203

The 95% and 85% credible intervals (CI) for all parameters are shown. Negative mean values indicate a positive effect due to the use of an inverse-log link function (i.e. all three covariates had a positive association with total winter population size). SD = standard deviation.

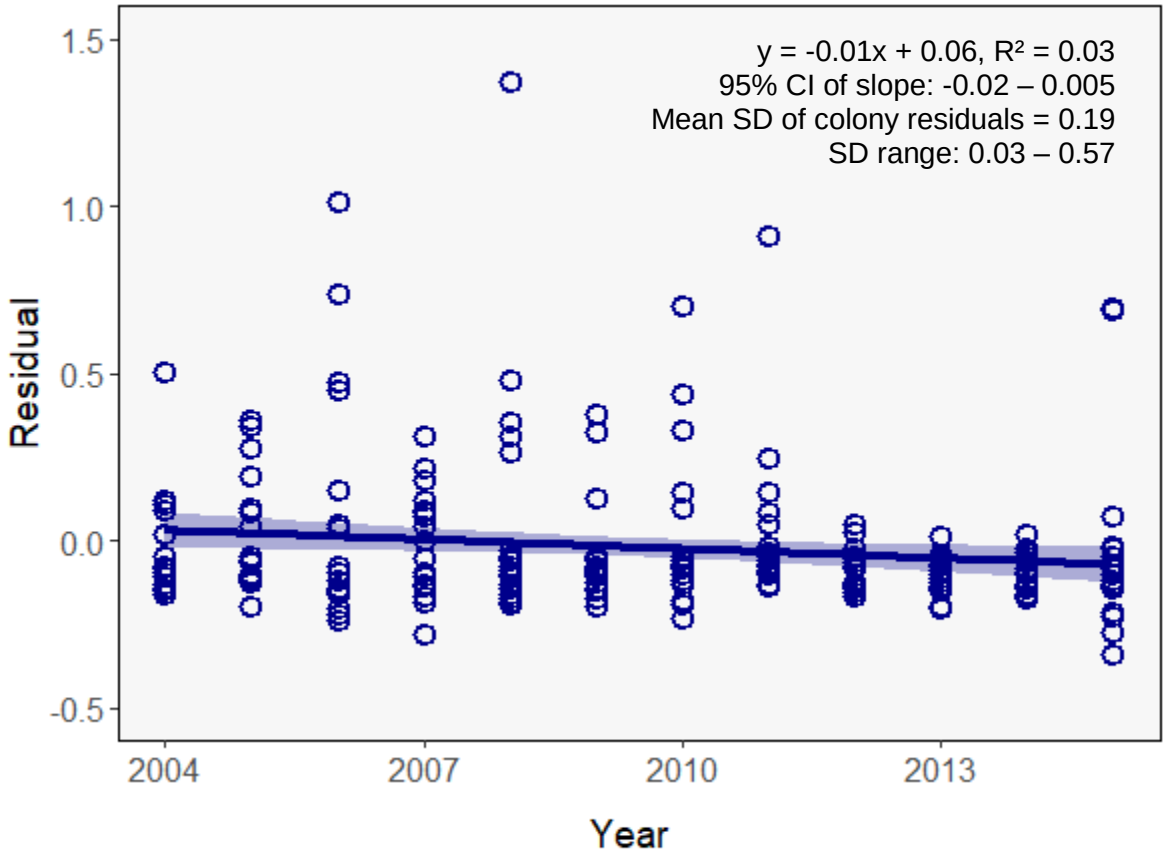

Figure S1. Posterior means (from 30,000 MCMC simulations) of estimated residuals from our top-supported hierarchical gamma hurdle model using colony-level size data (from all 19 known winter colonies) during 2004 – 2015 as a function of peak summer index (NABA counts), autumn greenness in region 1 (NDVI), dense forest cover surrounding colony locations, whether a colony was located inside/outside of the reserve, and random site effects on both parts of the hurdle model. The linear trend (solid line) and associated 95% credible interval (shading) are shown. Text indicates linear equation (for a post-hoc Bayesian regression), R^2 (amount of variance explained by year), 95% credible interval of slope, and standard deviation (SD, mean and range) of residual values.

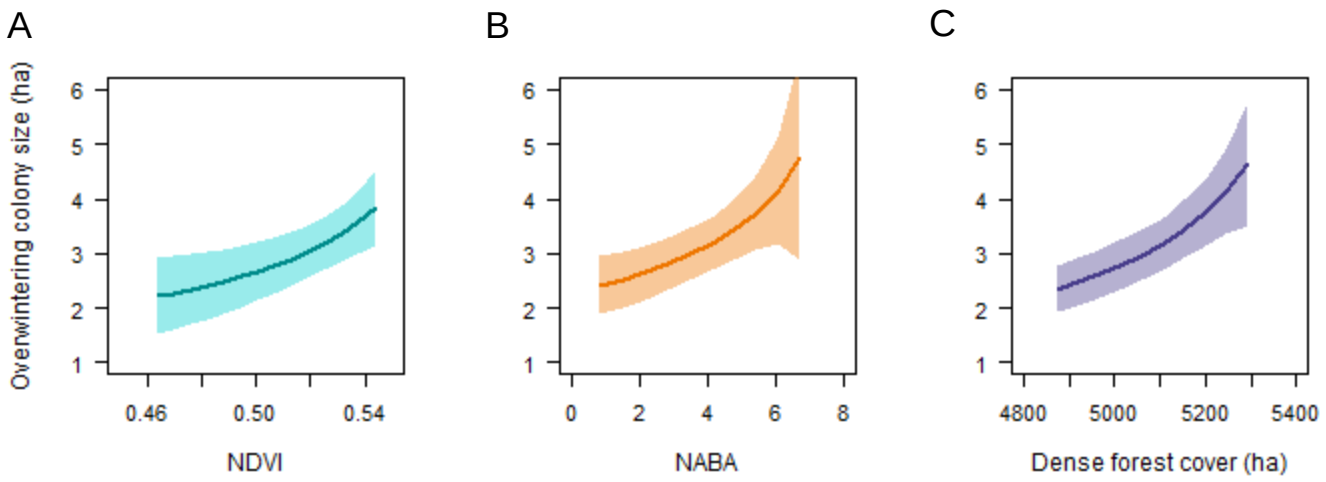

Figure S2. Covariate effects of (*A*) NDVI during the first half of autumn migration, (*B*) NABA counts (peak summer index), and (*C*) dense forest cover (ha; summed across colonies) on total overwintering colony size during 2000 – 2015, as estimated using a model with aggregated winter colony data. Solid lines show the marginal effect (with 95% credible intervals shaded) when all other covariates are held at their mean values.

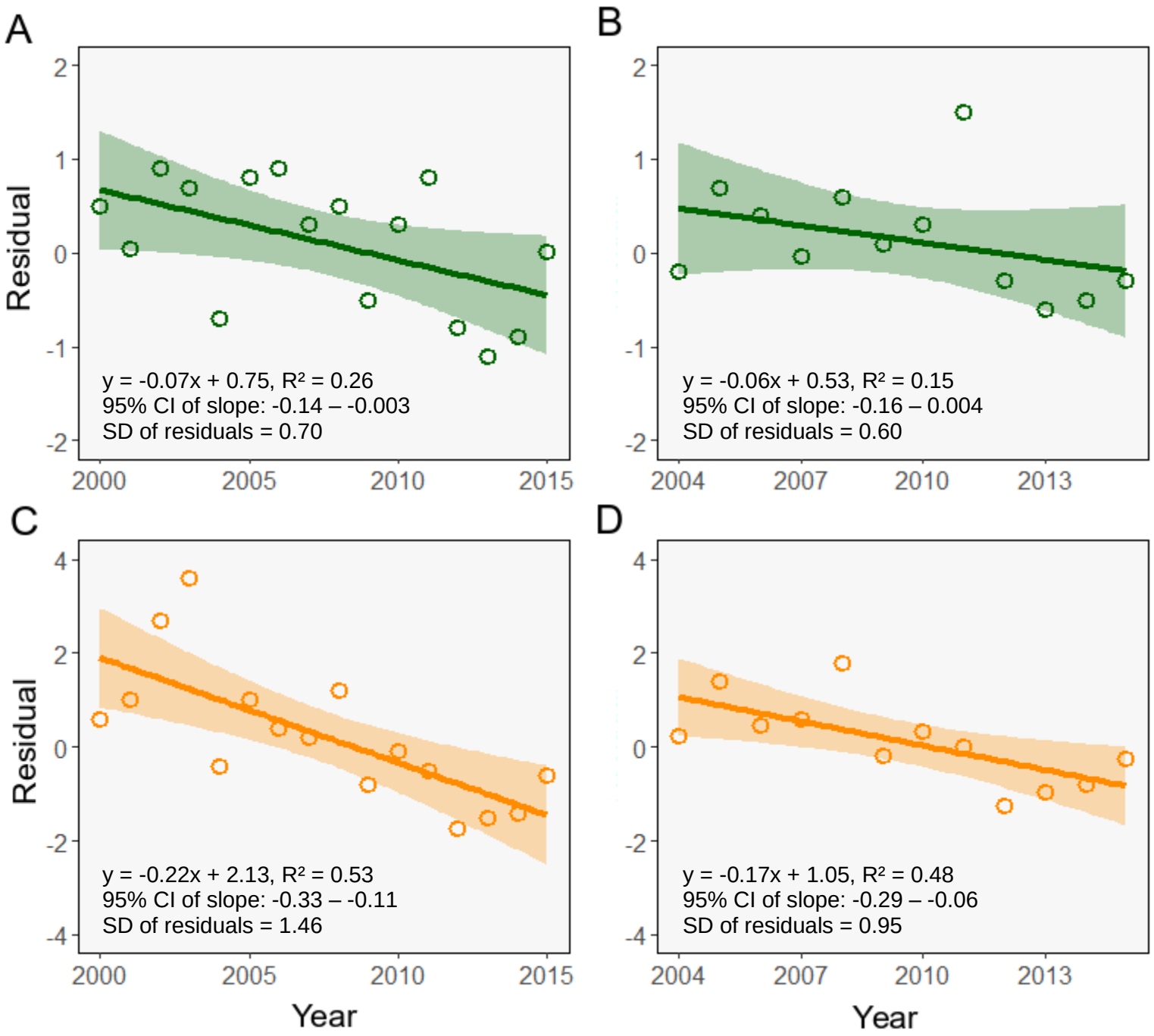

Figure S3. Estimated residuals (posterior means from 30,000 MCMC simulations) from gamma mixed models fit using aggregate colony data during 2000 – 2015 (*A* & *C*) and 2004 – 2015 (*B* & *D*). In *A* & *B*, models included peak summer index (NABA counts), autumn greenness in region 1 (NDVI), total dense forest cover surrounding colony locations (summed across colonies), and a random effect of year. In *C* & *D*, models included only peak summer index (NABA counts) and a random effect of year. The linear trend (solid line) and associated 95% credible interval (shading) are shown for post-hoc Bayesian regressions of the residuals. Text indicates linear equation, R^2 (amount of variance explained by year), 95% credible interval of slope, and standard deviation (SD) of residual values.

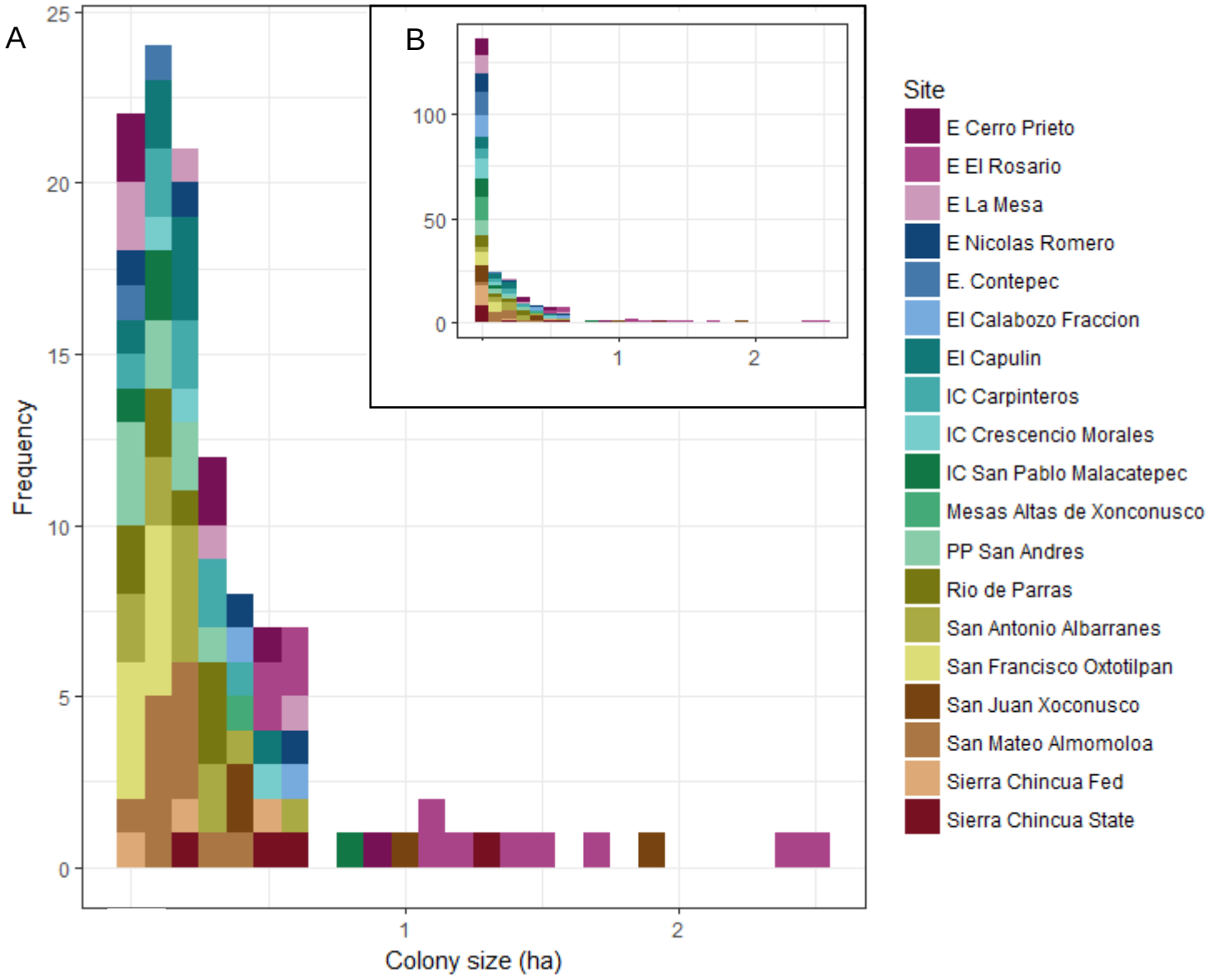

Figure S4. Histograms of monarch butterfly wintering colony sizes (hectares). *(A)* Annual observations across all 19 colony sites during December 2004 – 2015, excluding zero values (i.e. ≥ 0.01 ha; n = 114 or 50% of site-years). *(B)* Total annual observations, including zeros (n = 228). Color coding represents the colony site and matches that of Fig. 1 in main text. E. El Rosario, San Antonio Albarranes, and San Mateo Almomoloa had a colony detected every year during the study period.