Peromyscus responses to alternative forest management systems in the Missouri Ozarks, USA

Robert A. Gitzen a,*, Rochelle B. Renken b, Debby K. Fantz c, Alexander J. Wolf d, 1, Randy G. Jensen d, 2, Jean Favara d, 3, Joshua J. Millspaugh e

aSchool of Forestry and Wildlife Sciences, Auburn University, Auburn, AL 36849, United States
bMissouri Department of Conservation, Jefferson City, MO 65109, United States
cMissouri Department of Conservation, Columbia, MO 65201, United States
dMissouri Department of Conservation, Ellington, MO 63638, United States
eDepartment of Ecosystem and Conservation Sciences, W. A. Franke College of Forestry and Conservation, University of Montana, Missoula, MT 59812, United States

This is the accepted manuscript of a paper published in Forest Ecology and Management. The final version may be accessed at https://doi.org/10.1016/j.foreco.2018.07.004.

This paper is available under a CC-BY-NC-ND license.

* Corresponding author.
E-mail addresses: rag0012@auburn.edu (R. A. Gitzen); Rochelle.Renken@mdc.mo.gov (R. B. Renken); Debra.Fantz@mdc.mo.gov (D. K. Fantz); alex.j.wolf@gmail.com (A. J. Wolf); jfavara@audubon.org (J. Favara); joshua.millspaugh@mso.umt.edu (J. J. Millspaugh)

1 Present address: Scenic Hudson, Poughkeepsie, NY 12601, United States
2 Deceased
3 Present address: Audubon Riverlands Center, West Alton, MO 63386, United States
Abstract
Operational-scale forest management experiments are long-term investments because harvest
treatment effects may be dynamic throughout one or more rotation lengths. We examined
deermouse (*Peromyscus* spp.) abundance over the first 20 years of the Missouri Ozark Forest
Ecosystem Project (MOFEP), which assesses ecological responses to even-aged, uneven-aged,
and no-harvest forest management systems applied at landscape scales in the Missouri Ozarks. In
the spring of each of 11 sampling years, we sampled *Peromyscus* populations on two permanent
trap grids on each of the nine study sites (*n* = 3 sites per management system). Management
entries occurred in 1996 and 2011, with small mammal sampling conducted during two pre-
16–18 (2012–2014) after the first entry. We estimated abundance for each grid in each sampling
year with Bayesian closed-population mark-recapture models, and modeled variation in
abundance with negative binomial log-linear mixed effects models. Uneven- and even-aged
management systems caused similar increases in *Peromyscus* abundance that were detectable
shortly after the first management entry [proportional effect of even-aged management on
*Peromyscus* abundance vs. no-harvest management: posterior median = 1.8, 95% credible
interval 1.0–3.2; proportional effect of uneven-aged vs. no harvest management = 1.7 (1.0, 2.8)].
These effects were not surprising given positive effects of harvest treatments on understory cover
and food resources. However, the consistency of this increase was less expected, as we observed
no conclusive dissipation of harvest effects even in Years 13–14 after the first entry or
amplification of harvest effects after the second entry. Observing extremely high system-wide
yearly variation in *Peromyscus* abundance, we did not detect evidence of increasing divergence
in effects of these three management systems or of any area-wide trends in abundance during
1994–2014. However, over subsequent decades, we expect higher potential for divergence in *Peromyscus* abundance as the three management systems differentially shape forest structure and tree species composition. Thus, the MOFEP study offers a unique framework for building and testing hypotheses about patterns and mechanisms of long-term changes in Ozark forests and effects on vertebrate communities.

**Keywords**

Deermice; Even-aged management; Missouri Ozarks; MOFEP; Small mammals; Uneven-aged management

---

### 1. Introduction

Controlled experiments at operational scales allow direct statistical inference about effects of forest management treatments, using randomization and pre-treatment measures to reduce effects of potential confounding factors that may be impossible to address adequately in non-experimental studies (Ganio, 2006; Monserud, 2002; Sheriff and He, 1997). However, when experimental treatments involve harvesting mature forests, such studies face the challenge that the full responses to forest disturbances can unfold over temporal scales ranging from short-term immediate effects and year-to-year early successional dynamics, to decadal- and century-scale changes (Franklin et al., 2002; Gill et al., 2017; Lindenmayer and Laurance, 2012; Oliver, 1980; Olson and Saunders, 2017). This challenge is especially relevant for understanding responses by forest-floor small mammals, which are a frequent research component of multidisciplinary operational experiments (Converse et al., 2006; Fantz and Renken, 2005; Kellner et al., 2013; Lehmkuhl et al., 1999). Although small mammal populations can show especially strong changes
in the first few years after canopy-removing disturbances (Gashwiler, 1970b; Kirkland, 1990; Sullivan et al., 2008), responses are dynamic over subsequent stages of stand development (Ecke et al., 2002; Fisher and Wilkinson, 2005; Sullivan et al., 1999; Zwolak, 2009). Decades after a disturbance, small mammal populations may continue to be shaped by the disturbance's long-term effects on key habitat characteristics such as abundance of large down logs (Carey and Johnson, 1995) and abundance and species composition of mast-producing trees (Ostfeld et al., 1996; Wang et al., 2009). Even in the absence of physical disturbances, populations of small mammal taxa such as deermice (*Peromyscus* spp.) can show enormous temporal variation in abundance due to extrinsic variability in food resources and weather along with intrinsic effects of density dependence at intra- to interannual time scales (e.g. Lewellen and Vessey, 1998; Luis et al., 2010; Wang et al., 2009). Such variation creates the potential for corresponding temporally heterogeneous population responses to manipulative treatments (sensu, for example, Seamon and Adler, 1996; Wheatley et al., 2002). Conversely, examining treatment effects in the context of background spatial and temporal variation may provide a better understanding of the relative importance of forest management as an agent of population change.

One of the longest-active multidisciplinary forestry experiments using large-scale operational treatments is the Missouri Ozark Forest Ecosystem Project (MOFEP) in southeastern Missouri, USA (Knapp et al., 2014). This Missouri Department of Conservation (MDC) study was initiated in the early 1990s to help managers better understand the impacts of harvest practices in state-owned forests of the region. Intended to be a multi-century study, MOFEP compares long-term application of three operational management systems (i.e. three experimental treatments: no-harvest management, uneven-aged management, and even-aged management) applied on a 100-year rotation length to several-hundred ha experimental sites.
(Brookshire and Shifley, 1997; Shifley and Kabrick, 2002). Individual management entries occur periodically, with each entry affecting a subset of the stands within each site. Compared to stand-level studies, a long-term perspective is especially critical in MOFEP given that experimental treatments are long-term applications of management systems at the scale of local landscapes. After two decades of data collection in core research areas (Knapp et al., 2014), the experiment has reached a uniquely valuable stage in which researchers and managers can begin comparing long-term effects of the three management systems (e.g. Morris et al., 2013; Olson et al., 2015; Olson et al., 2017; Rota et al., 2017).

During the first five years post-treatment in MOFEP, uneven- and even-aged management had similar positive effects on relative abundance of *Peromyscus* spp., the primary taxa captured in the study’s small mammal component (Fantz and Renken, 2005). Specifically, harvest treatments appeared to buffer populations from an apparent regional multi-year decline in abundance documented on no-harvest sites. It was unknown whether this multi-year decline was part of a longer term trend, but this possibility highlights an additional temporal complexity that can be assessed with long-term studies. Operational experiments such as MOFEP take place in forest ecosystems that may be undergoing broader directional changes due to factors such as climate change, introduced diseases, and altered disturbance patterns (Gillen and Hellgren, 2013; Hanberry et al., 2014; Olson et al., 2014; Zenner et al., 2006). Thus, there is potential for interactions between experimental treatments and larger scale agents of change (Olson et al., 2017).

Building on previous work by Fantz and Renken (2005), we examined patterns of change in *Peromyscus* abundance for 11 years of sampling during the first 20 years of the MOFEP small mammal study. Timing of sampling was tied to the first (1996) and second (2011) management
entries. Sampling occurred in four groups of years: pre-treatment (1994–1995), and Years 2–5 (1998–2001), Years 13–14 (2009–2010), and Years 16–18 (2012–2014) after the first management entry, with the last group covering the first three years after the second management entry. We focused on several questions about comparative effects of MOFEP treatments on *Peromyscus* abundance. First, we examined whether the positive initial effects of uneven- and even-aged harvest treatments noted by Fantz and Renken (2005) persisted over a decade (13–14 years) after the first management entry. Second, we assessed whether short-term population changes after the second management entry showed similar patterns as those observed after the first entry. Third, to give context for the relative importance of harvest treatments as drivers of *Peromyscus* population abundance, we compared treatment effects to the range of background variation among years and sites in this study system. Finally, we assessed whether there were detectable system-wide trends in *Peromyscus* abundance during the first 20 years of the MOFEP study, particularly treatment-specific trends that would indicate divergence among management approaches.

2. **Materials and methods**

2.1. **MOFEP design**

We provide a brief description of MOFEP's experimental design, study sites, and small mammal sampling protocol; other publications provide more extensive details (Brookshire and Shifley, 1997; Knapp et al., 2014; Shifley and Kabrick, 2002). The study occurs in the south-central Missouri Ozark Highlands, in mature, second-growth forests dominated by oaks (*Quercus* spp.), hickories (*Carya* spp.), other hardwoods, and shortleaf pine (*Pinus echinata*). Any prior timber harvesting on study sites occurred before 1950. The experiment uses a complete randomized
block design, with three blocks each consisting of three, 312–514-ha sites (Sheriff and He, 1997). Each of three experimental treatments [no harvest (NH), uneven-aged (UA), or even-aged (EA) management systems] was randomly assigned to one site per block. Uneven-aged and even-aged management systems follow MDC management guidelines current at the time of each entry (e.g. MDC, 2014), with management entries planned every 15 years such that sites are managed on a 100-year rotation for uneven- and even-aged systems.

Each study site comprises a local landscape of 44–82 individual forest stands. Manipulations applied at the stand level are components of each management system (e.g. clearcut harvests in even-aged management), but are not applied to all stands within a site at each entry. Rather, decisions about whether and how to treat each stand during each entry are based on stand and site characteristics as in standard operational management, albeit within the bounds of the randomly assigned site-level management approach (Sheriff, 2002). Thus, the primary focus of MOFEP is the comparative landscape-scale effects of long-term application of these management systems rather than stand-level effects of individual entries.

To date, treatment entries have occurred in 1996 and 2011. In even-aged treatments, during each of the two entries 10–15% of each site was regeneration harvested (clearcut) in openings 3–12 ha in size, with up to 2.3 m² ha⁻¹ residual basal area retained in clearcut areas (Knapp et al., 2014). In 2011, 11% of one even-aged treatment site was shelterwood harvested. In both entries, other stands on even-aged sites received intermediate thinning (average of 15.7% of site area per entry, Knapp et al., 2014) or were not treated. In sites assigned to uneven-aged management, during the first entry, single-tree and small group selection harvests occurred on much of each site (average area harvested = 57%, Kabrick et al., 2002). Group selection harvests were implemented on approximately 5% of the non-reserved area of each uneven-aged site.
during the 1996 entry, producing 123–267 small (21–43 m diameter) openings throughout these sites. In the second management entry, only single-tree selection was implemented on uneven-aged sites (averaging 40% of each site; Knapp et al., 2014). Variation in the specific actions implemented as part of each management system (e.g. group-selection harvests with uneven-aged management) reflect that experimental treatments are the long-term, site-level application of operational systems, which allow management flexibility within and between entries (Knapp et al., 2014). The average site-level residual basal area after each entry was similar between uneven- and even-aged treatments (approx. 17–19 m² ha⁻¹ immediately after each entry vs. average approx. 23–24 m² ha⁻¹ in no-harvest treatment sites; Olson et al., 2017). During each entry approximately 10% of each uneven- and even-aged site was reserved and not subject to harvest in that entry.

2.2. Small mammal sampling

The MOFEP small mammal design restricted inference to north- and east-facing side slopes within each site (Fantz and Renken, 2005). Within this land type, two 12 × 12 permanent trapping grids were established per site, with 25 m between trap stations, resulting in 18 grids of 7.6 ha each. Because of MOFEP's long-term focus on management systems applied at the site level, these permanent grids were established in each site without respect to where stand-level manipulations occur during individual management entries. Trap stations of grids in uneven-aged treatment sites were mostly in areas harvested during the first entry (X = 92% of stations per grid, range 75–100%, including 6–11% of stations per grid exposed to group-selection harvests), but exposure was lower and more variable in the second harvest entry (X = 30% of stations, range 0–83%). In even-aged treatment sites, 0–98% of stations per grid entry were exposed to
regeneration harvests or intermediate thinning (first entry: $\bar{X} = 28\%$, range = 0–60\% of stations per grid; second entry: $\bar{X} = 56\%$, range = 0–98\%). All grids were separated by at least 300 m; throughout all sampling we recorded between-grid movements by <5 mice.

During spring (April and May) of each sampling year, each grid was sampled for a single trap session during which traps were open continuously for six consecutive days and nights, and checked daily. Sites within a block were trapped simultaneously, while the three blocks were trapped non-simultaneously during each sampling year. A single Sherman live trap ($7.7 \times 8.9 \times 22.9 \text{ cm}$) (H.B. Sherman Trap, Inc., Tallahassee, Fla., USA) was placed at each station, baited with a mixture of peanut butter and oats, and rebaited as needed during each daily trap check. At each trap check, captured live animals were identified to species, marked, and released. *Peromyscus* were uniquely marked with toe clipping (1994–2010) or numbered ear tags (2012–2014); other taxa were marked with these methods or marked temporarily with a non-toxic permanent black marker.

2.3. Analytical methods

Analyses focused on combined abundance of *Peromyscus maniculatus* (North American deermouse) and *P. leucopus* (white-footed deermouse), as other taxa were captured infrequently. As discussed by Fantz and Renken (2005), morphological field discrimination of these two *Peromyscus* species was problematic. Moreover, both species appeared to be common throughout our sampling; in this system, patterns of combined abundance were considered to be of higher management relevance than their species-specific responses (R. Renken, pers. obs.). Naïve apparent survival between years was close to 0 for *Peromyscus*, so we treated each yearly six-night trap session as a closed-population mark-recapture sample, resulting in 198 sessions
(i.e. 9 sites × 2 grids per site × 11 sampling years). To quantify changes in abundance, we first attempted a fully integrated Bayesian hierarchical modeling approach following Schofield and Barker (2014) and Converse and Royle (2012), but experienced prohibitively slow convergence and long processing time during extensive preliminary modeling of the large MOFEP data set with these approaches (see Appendix). Therefore, we performed a two-stage analysis, first estimating abundance via mark-recapture modeling, then using these abundance estimates as the response variable for modeling variation in expected abundance.

2.3.1. Mark-recapture estimation

We implemented Bayesian closed-population mark-recapture modeling to estimate abundance for each of the 198 sessions in a unified analysis (Kéry and Schaub, 2012; Royle and Dorazio, 2008; Royle et al., 2007). Full details are available in the Appendix. The primary detectability model we used was a form of the general model \( M_{tb} \) (Otis et al., 1978) incorporating within-session variability in capture probabilities due to temporal effects and behavioral response to prior capture. Posterior distributions were estimated via Markov Chain Monte Carlo (MCMC) sampling using JAGS 4.2.0 and R library rjags (Plummer, 2016). We report posterior distributions for session-specific abundance, across-session average per-night capture probability, and the across-session average recapture effect. The recapture effect was reported on the odds ratio scale (odds of recapture if captured on the previous trap night / odds of capture if not captured on the previous trap night), with the posterior distribution estimated by exponentiating the logit-scale average recapture effect parameter within each MCMC iteration. For all analyses, posterior distributions for each parameter were summarized by the posterior median and a 95% credible interval formed by taking the (0.025, 0.975) percentiles of the
posterior samples. Our second stage of analyses used these mark-recapture abundance posterior
distributions to form response variables for analyzing patterns of abundance across sites and
years.

2.3.2. Abundance modeling

For modeling expected abundance, we used alternative forms of a general negative binomial
mixed effects model, tailoring the parameterization of each form to better address specific
research questions. Our primary second-stage analysis estimated overall and yearly effects of the
tree MOFEP experimental treatments, using estimated abundance for each grid for the nine
post-treatment years as the response variable and average pre-treatment estimated abundances
for each grid as a covariate (sensu Senn, 2006). The model included fixed effects of average pre-
treatment abundance and block, and random effects of site, year, and block × year (see
Appendix). Log-scale treatment effects were modeled with fixed-effects parameters for the
overall effects of uneven-aged and even-aged management relative to no-harvest management;
these could be interpreted as estimates of average effects of uneven-aged and even-aged
management across all nine post-treatment sampling years. We added a treatment × year random
effect allowing for yearly variation in effects of the harvest management systems. We
exponentiated posterior samples of log-scale treatment-contrast parameters to form posterior
distributions for overall and year-specific proportional effects of uneven-aged and even-aged
management on expected Peromyscus abundance. We compared harvest treatments by
calculating posterior distributions for the ratio of even-aged to uneven-aged average and year-
specific proportional effects.
To examine the implications of ignoring mark-recapture abundance estimation uncertainty, we conducted two alternative sets of analyses with this treatment-effects model. In the first analysis ($N_{\text{posterior \_median}}$ analysis), we did not incorporate mark-recapture estimation uncertainty, but fit the model to the single vector of posterior median abundance estimates from Bayesian mark-recapture estimation. The pre-treatment abundance covariate was calculated as the average of the two pre-treatment posterior median estimates for each grid (i.e. average of 1994 and 1995 values). To assess whether posterior uncertainty about model parameters increased meaningfully when mark-recapture estimation uncertainty was incorporated, we performed a second analysis ($N_{\text{posterior \_sample}}$ analysis) using individual MCMC sample vectors from the posterior distributions for session-specific abundance from mark-recapture modeling, rather than just the single posterior median vector. The Bayesian mark-capture modeling of capture histories from the 198 trap sessions produced 30,000 posterior abundance vectors, each of length 198. We used each of 3,000 of these posterior sample vectors (every 10th multivariate sample) as input data for the negative binomial treatment-effects model described above, with the pre-treatment average abundance for each grid calculated using the relevant pre-treatment abundance values from that sample vector. We analyzed each vector separately to produce 3,000 sets of posterior samples of negative binomial model parameters, then pooled these sets to form hyper-posterior distributions that incorporated mark-recapture estimation uncertainty. For all negative binomial abundance models, posterior distributions were estimated with MCMC in Stan via R packages brms (Buerkner, 2017) and RStan (Stan Development Team, 2016). More complete details about each analysis are in the Appendix.

To better estimate background spatial and temporal effects in the MOFEP system, we ran a modified form of the negative binomial model with mark-recapture posterior median
abundance as the response variable. In this model form, we incorporated pre-treatment abundance into the response variable vector rather than as a covariate, thus allowing estimation of 11 rather than 9 year effects and fully incorporating estimates of persistent spatial variation (other than among-treatment variation) into block and site effects. We calculated posterior summaries for block-, site- and year-specific effects (see Appendix), and compared their magnitudes with those of estimated treatment effects from the $N_{\text{posterior\_median}}$ analysis.

To assess evidence of long-term trends in *Peromyscus* abundance and whether treatments showed an increasingly diverging pattern, we focused on a third form of the general negative binomial mixed effects model using posterior median abundance estimates for years 1994 through 2014 as the response variable. The model incorporated year as a numeric covariate, a treatment-specific log-linear trend parameter, block fixed effects, and random effects as described above (see Appendix). Samples for log-scale trend parameters were exponentiated to produce posterior distributions for proportional trends.

3. Results

During 11 sampling years with a combined trap effort of 171,072 trap nights across 198 trap sessions, we recorded 7,135 captures of at least 4,377 unique individuals. We captured 16 species (Table 1), with *Peromyscus* comprising 88% of individuals captured (3,871 mice). Individual mice were rarely captured in more than one year, with only 48 *Peromyscus* individuals recaptured between years for the entire study. Throughout all sampling, 50% of *Peromyscus* were identified in the field as *P. leucopus* and 43% as *P. maniculatus*; 7% were not identified to species. Based on morphological identifications, both species were commonly captured throughout the study; for example, mice identified as *P. maniculatus* comprised 50% of
Table 1. Total individual animals captured by taxon during small mammal live trapping on nine MOFEP study sites in Missouri Ozarks, USA, 1994–2014. Nomenclature follows Bradley et al, 2014. (Note: morphological discrimination of *P. maniculatus* from *P. leucopus* was of uncertain accuracy; please see Section 2.3.)

<table>
<thead>
<tr>
<th>Species</th>
<th>Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elliot's short-tailed shrew (<em>Blarina hylophaga</em>)</td>
<td>51</td>
</tr>
<tr>
<td>Least shrew (<em>Cryptotis parva</em>)</td>
<td>1</td>
</tr>
<tr>
<td>Eastern cottontail (<em>Sylvilagus floridanus</em>)</td>
<td>1</td>
</tr>
<tr>
<td>Eastern gray squirrel (<em>Sciurus carolinensis</em>)</td>
<td>2</td>
</tr>
<tr>
<td>Southern flying squirrel (<em>Glaucomys volans</em>)</td>
<td>14</td>
</tr>
<tr>
<td>Eastern chipmunk (<em>Tamias striatus</em>)</td>
<td>208</td>
</tr>
<tr>
<td>Eastern woodrat (<em>Neotoma floridana</em>)</td>
<td>160</td>
</tr>
<tr>
<td>Fulvous harvest mouse (<em>Reithrodontomys fulvescens</em>)</td>
<td>4</td>
</tr>
<tr>
<td>Western harvest mouse (<em>Reithrodontomys megalotis</em>)</td>
<td>11</td>
</tr>
<tr>
<td>White-footed deermouse (<em>Peromyscus leucopus</em>)</td>
<td>1912</td>
</tr>
<tr>
<td>North American deermouse (<em>Peromyscus maniculatus</em>)</td>
<td>1675</td>
</tr>
<tr>
<td>Deermouse, not identified to species (<em>Peromyscus spp.</em>)</td>
<td>284</td>
</tr>
<tr>
<td>Golden mouse (<em>Ochrotomys nuttalli</em>)</td>
<td>42</td>
</tr>
<tr>
<td>Prairie vole (<em>Microtus ochrogaster</em>)</td>
<td>1</td>
</tr>
<tr>
<td>Woodland vole (<em>Microtus pinetorum</em>)</td>
<td>9</td>
</tr>
<tr>
<td>Southern bog lemming (<em>Synaptomys cooperi</em>)</td>
<td>1</td>
</tr>
<tr>
<td>House mouse (<em>Mus musculus</em>)</td>
<td>1</td>
</tr>
</tbody>
</table>
Captures per six-night trap session ranged from 0–184 individual *Peromyscus* per grid (\(\bar{x} = 19.6,\) median = 11, SD = 24.4; Fig. 1). Average overall probability of capture for an individual not previously captured in a trap session was low [posterior median = 0.109, 95% credible interval = 0.087, 0.135)], with 63% of all individuals captured in only one trap night during a session, and 84% captured in only one or two trap nights. However, as expected for *Peromyscus*, average recapture probability increased sharply [posterior median odds ratio, odds of being captured on a given trap night for an individual captured the previous trap night divided by odds for an individual not captured the previous trap night, = 2.72 (2.39, 3.10)]. Posterior median abundance estimates ranged from 0–290 mice per grid per session (\(\bar{x} = 38.0,\) median = 21.0). Based on the standard deviation and mean of the 30,000 posterior abundance samples for the 198 trap sessions, the median coefficient of variation (CV) within posterior abundance samples was 25% (range 5–162%; all sessions with CVs >35% were sessions with 5 or fewer individuals captured).

Both uneven-aged and even-aged management systems increased *Peromyscus* abundance relative to no-harvest management (Table 2, Fig. 2). Based on the \(N_{\text{posterior\_median}}\) negative binomial model analysis, across the nine post-treatment sampling years the estimated average effect of uneven-aged management was to increase *Peromyscus* expected abundance by 73% compared to no-harvest management [UA vs NH: posterior median = 1.73; 95% credible interval = (1.11, 2.66); a value of 1.0 would correspond to no effect on abundance]. The estimated average effect of even-aged management was an 84% increase in expected abundance compared to no-harvest management ([EA vs NH: 1.84 (1.09, 3.00)]. There was not a detectable difference between the two active management systems (EA vs UA: 1.07 (0.64, 1.72)]. Incorporating mark-
recapture estimation uncertainty increased overall uncertainty about treatment effects in terms of credible interval widths [Table 2; Fig. 2; average overall proportional effect of treatments based on $N_{posterior\_sample}$ analysis which incorporated mark-recapture estimation uncertainty: UA vs NH: 1.70 (1.01, 2.84); EA vs. NH: 1.80 (1.00, 3.18); EA vs UA: 1.06 (0.59, 1.85)].
Figure 1. Estimated abundances of *Peromyscus* on MOFEP small mammal grids in the Missouri Ozarks for 11 sampling years during 1994–2014. Each panel corresponds to a single block × year combination, with three sites per block (NH = no-harvest management UA = uneven-aged management; EA = even-aged management) and two trap grids per site. Closed symbols are median estimates from MCMC-estimated posterior distributions for each abundance parameter from Bayesian closed-population mark-recapture modeling; error bars are 0.95 credible intervals based on the (0.025, 0.975 percentiles) of the MCMC posterior samples. Open circles are the number of unique *Peromyscus* individuals captured during each trapping session. Years 1994–1995 were pre-treatment years; the first and second management entries occurred in 1996 and 2011.
Table 2. Summaries of posterior distributions (median and 95% credible interval) for parameters in Bayesian mixed effects model of *Peromyscus* expected abundance across 11 sampling years in the MOFEP study. For the $N_{\text{posterior\_median}}$ analysis, the negative binomial model was fit to posterior median values of MCMC samples of abundance from mark-recapture estimation. For the $N_{\text{posterior\_sample}}$ analysis, we fit the negative binomial model to each of 3,000 abundance vectors from the mark-recapture MCMC estimation, then pooled posterior samples for each model parameter, with resulting credible intervals incorporating mark-recapture estimation uncertainty. All parameters except $r$ are log-scale effects; see Eq. 1 in Appendix.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$N_{\text{posterior_median}}$</th>
<th>$N_{\text{posterior_sample}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_0$</td>
<td>3.01 (2.05, 3.96)</td>
<td>3.11 (2.14, 4.09)</td>
</tr>
<tr>
<td>$B_{_block2}$</td>
<td>-0.41 (-0.96, 0.13)</td>
<td>-0.44 (-1.07, 0.19)</td>
</tr>
<tr>
<td>$B_{_block3}$</td>
<td>-0.94 (-1.48, -0.42)</td>
<td>-0.94 (-1.58, -0.32)</td>
</tr>
<tr>
<td>$B_{_pre}$</td>
<td>0.01 (0.00, 0.01)</td>
<td>0.01 (0.00, 0.01)</td>
</tr>
<tr>
<td>$B_{_UAave}$</td>
<td>0.55 (0.10, 0.98)</td>
<td>0.53 (0.01, 1.04)</td>
</tr>
<tr>
<td>$B_{_EAvave}$</td>
<td>0.61 (0.09, 1.10)</td>
<td>0.58 (-0.01, 1.15)</td>
</tr>
<tr>
<td>$\sigma_{site}$</td>
<td>0.15 (0.01, 0.53)</td>
<td>0.13 (0.01, 0.66)</td>
</tr>
<tr>
<td>$\sigma_{year}$</td>
<td>1.04 (0.54, 1.93)</td>
<td>0.93 (0.49, 1.92)</td>
</tr>
<tr>
<td>$\sigma_{blockyear}$</td>
<td>0.38 (0.21, 0.60)</td>
<td>0.38 (0.18, 0.68)</td>
</tr>
<tr>
<td>$\sigma_{UA_year}$</td>
<td>0.18 (0.01, 0.55)</td>
<td>0.17 (0.01, 0.63)</td>
</tr>
<tr>
<td>$\sigma_{UA_year}$</td>
<td>0.34 (0.02, 0.83)</td>
<td>0.30 (0.02, 0.99)</td>
</tr>
<tr>
<td>$r$</td>
<td>3.66 (2.71, 4.83)</td>
<td>3.00 (2.18, 4.18)</td>
</tr>
</tbody>
</table>
Figure 2. Estimated proportional effects of MOFEP treatments on *Peromyscus* expected abundance by sampling year. Values are posterior median and 95% credible interval for each parameter from negative binomial mixed-effects model contrasting pairs of treatments by year and analysis (NH = no-harvest management, UA = uneven-aged management, EA = even-aged management). Upper row summarizes results from analysis that does not incorporate mark-recapture estimation uncertainty ($N_{\text{posterior\_median}}$ analysis which analyzed posterior median mark-recapture abundance estimates for each trap session); bottom row summarizes results for alternative analysis incorporating mark-recapture estimation uncertainty ($N_{\text{posterior\_sample\ pooled}}$ replicate analyses of individual samples from the posterior distribution of mark-recapture abundance estimates). Vertical dashed lines indicate harvest treatment entries in 1996 and 2011;
horizontal dashed line is given for reference at 1.0 (proportional effect of 1.0 = no difference between treatments).
We observed upward shifts in yearly posterior distributions for even-aged treatment effects in 1999 and 2014, vs. potentially reduced effect of even-aged management in Years 13–14 after the first entry. This pattern suggested potential dissipation of the initial treatment effects followed by a return in magnitude after the second entry (Fig. 2). However, these were not conclusive patterns. Variance components capturing yearly variation in treatment effects had small estimated values (Table 2). Associated posterior medians and credible intervals for estimated deviations of year-specific treatment effects from the overall average for each treatment \( (b_{UA_y}, b_{EA_y} \) in Appendix Eq. 2) indicated that none of the year-specific effects differed conclusively from the overall post-treatment average effect of each treatment.

Uneven-aged and even-aged management had large effects on expected abundance compared to other within-block persistent site effects (Fig. 3). Estimated persistent differences among blocks were similar to or greater than estimated effects of treatments. For example, expected *Peromyscus* abundance in Block 1 was estimated to be 2.5 (1.28, 4.73) times higher than in Block 3 on average. However, during several of the 11 sampling years, estimated area-wide yearly deviations from the multi-year average appeared to have stronger effects on expected abundance than block or treatment effects (Fig. 3). Contrasting the years of highest (975 mice) vs. lowest (102 mice) numbers of *Peromyscus* captured across all sites, expected abundance in 2012 was estimated to be 11.22 (4.14, 27.00) times higher than expected abundance in 2010.

During 1994–2014 no systematic log-linear long-term trend was evident for any treatment [proportional change in expected abundance per year: NH = 0.98 (0.89, 1.06), UA = 1.00 (0.92, 1.10), EA = 1.00 (0.91, 1.09); value of 1.0 corresponds to no trend] nor across all study sites [refitting model to incorporate a single trend parameter common to all three
treatments, the yearly proportional change was 0.99 (0.91, 1.08)]. This analysis reinforced results
from the yearly treatment-effects analysis, as both analyses indicated no detectable pattern of
increasing divergence among treatments after two management entries, at least in terms of
expected *Peromyscus* abundance. However, as the wide credible intervals summarize, high
system-wide yearly variation produced high uncertainty about the existence of trends in
*Peromyscus* abundance during the first 20 years of MOFEP.
Figure 3. Comparison of MOFEP treatment effects on *Peromyscus* expected abundance vs. background spatial (block and site within block) and year effects. Values are estimated proportional effects (posterior median and 95% credible interval) of harvest treatments (UA = uneven-aged management, EA = even-aged management, relative to no harvest management) vs. block-, site- and year-specific effects (deviations from long-term, system-wide average expected abundance) from negative binomial mixed effects modeling of expected abundance, using posterior median mark-recapture abundance as the response variable. Horizontal line at 1.0 indicates no effect; effects strengthen with increasing distance above or below 1.0.
4. Discussion

Both uneven- and even-aged management treatments increased *Peromyscus* abundance compared to no-harvest management. These increases were detectable shortly after the first entry and persisted through Year 18 post-treatment. A positive effect of treatments is unsurprising given results of past studies and MOFEP treatment effects on *Peromyscus* habitat attributes. Although *P. leucopus* sometimes declines in larger clearcut openings (Kellner et al., 2013; Kirkland, 1990), both *P. leucopus* and *P. maniculatus* are considered habitat and dietary generalists (Adler and Wilson, 1987; Cramer; 2014; Olsen et al., 2000; Seamon and Adler, 1996) that often respond positively to higher levels of ground cover, soft and hard mast, other seeds, insects, overall vegetation structural complexity (Fisher and Wilkinson, 2005; Kellner et al., 2013; Perry and Thill, 2005), and, in some cases, volume of down logs (Lee, 2004; Loeb, 1999; Manning and Edge, 2004; Urban and Swihart, 2011; but see also Bowman et al., 2000; Greenberg et al., 2006; Craig et al., 2006). At the site level, both uneven-aged and even-aged management increased ground-level cover of herbaceous vegetation and woody vines; in clearcut stands, average ground cover increased by 41% (absolute change) from pre-treatment to Years 3–4 post-treatment (Grabner and Zenner, 2002; Zenner et al., 2006). Treatments also affected food and cover resources through increased berry production (Elizabeth Olson, MDC, unpublished data), development of high levels of small woody stem density from post-harvest stump sprouting (Kabrick et al., 2002), and addition of large amounts of woody debris to the forest floor of harvested areas. While the even-aged treatment reduced red oak acorn crops during the first 14 years post-treatment (Olson et al., 2015), the degree of reduction at the site level appeared minor compared to the magnitude of year-to-year variations in total acorn crops.
From the standpoint of MDC forest management on state lands, the observed effects of MOFEP treatments on *Peromyscus* abundance to date fit into an overall desirable pattern of responses by the vertebrate community, with no alarmingly negative effects observed for small mammals in this study or for amphibians, reptiles, or birds in other MOFEP studies (Kendrick et al., 2015; Rota et al., 2017; Wolf et al., 2016). For *Peromyscus*, treatment effects are not trivial, in that credible intervals indicate that uneven-aged and even-aged management each likely increase expected abundance by 50 to >100% compared to no-harvest management. Large changes in deermouse abundance could affect numerous other community attributes and processes such as vegetation demography and composition (Bricker et al., 2010; Gashwiler, 1970a; Ostfeld et al., 1997; Tallmon et al., 2003; Zwolak et al., 2010), nest predation (Schmidt et al., 2001), predator populations (Jensen et al., 2012), invertebrate composition and abundance (Elkington et al., 1996; Jones et al., 1998), and transmission of parasites and diseases (e.g. Calisher et al., 2011; Ostfeld et al., 2006). Still, treatment effects on *Peromyscus* abundance are comparable in magnitude to the difference among MOFEP study blocks, and perhaps an order of magnitude less than the area-wide difference between years of peak vs. lowest abundance. At least under current conditions, there seems to be little risk that any MOFEP management alternative will drive *Peromyscus* abundance beyond its background range of spatial-temporal variability.

We observed relatively consistent effects of uneven-aged and even-aged management within groups of sampling years. This pattern is in line with Fantz and Renken (2005), who did not detect strong year-to-year variation in treatment effects during Years 2–5 post-treatment. However, less clear is why we did not detect conclusive variation in treatment effects among the three post-treatment groups of sampling years. Given rapid development of a dense sapling
midstory in openings created by MOFEP harvests (Good et al., 2017; Jensen and Kabrick, 2008; Kabrick et al., 2002), the pulse of ground-level resources observed initially after disturbance might have been expected to dissipate by Years 13–14 after the first entry as understory light levels declined. Yet, while berry production in Years 13–14 returned to baseline conditions comparable to pre-treatment levels, average graminoid cover of uneven-aged and even-aged treatment sites remained higher in Year 14 than in pre-treatment sampling (Elizabeth Olson, MDC, unpublished data). Increases in abundance of down logs due to thinning and harvest during the first entry likely persisted through the next two decades, with such increases potentially important given that pre-treatment MOFEP sites were near a predicted low point in the U-shaped pattern of deadwood volume vs. stand age in hardwood-dominated forests of the central U.S. (Spetich et al., 1999).

With positive effects of the first management entries on Peromyscus populations persisting, it is unclear why the second MOFEP entry did not produce a detectable amplification of treatment effects. Any initial dissipation of treatment effects in Years 13–14 may have been reversed by an additional round of positive effects from the second management entry in Year 15. Another potential factor is that MOFEP small mammal grids within uneven-aged and even-aged treatment sites received heterogeneous exposure to management actions within and across individual entries. For example, in even-aged treatment sites, sampling grids had more overlap overall with stands clearcut harvested in the first entry than with clearcut areas in the second entry, while overlap with intermediate thinning areas was higher in the second entry. Lack of conclusive differences in treatment effects across the 11 post-treatment sampling years could also reflect statistical imprecision associated with the level of replication feasible in this operational experiment and the high temporal variability of Peromyscus populations. As more
entries and rounds of sampling occur, MOFEP researchers will have increased ability to assess
temporal and spatial patterns in effects of individual entries. There will also be increased ability
to integrate data across core research areas (e.g. data on understory vegetation and small
mammal abundance) so as to better understand mechanisms driving responses to treatments.

For MOFEP’s overarching focus on long-term, landscape-scale effects of the three
management systems, assessing short-term responses to individual entries is useful mainly as a
step in understanding and predicting potential broader scale effects. Based on results to date from
MOFEP studies of tree species composition and stand dynamics, currently a major overall
question of MOFEP is how the three alternative management systems will interact with other
factors affecting forest composition across many portions of the central U.S. (Olson et al., 2017).

One important uncertainty is whether maples (particularly *Acer rubrum* and *A. saccharum*) will
become an important overstory species on any MOFEP sites. Although maples historically were
likely absent from upland forests of the Missouri Ozarks (Guyette et al., 2007; Hanberry et al.,
2014; Olson et al., 2017), over the last 20 years maples have become a major component of the
MOFEP seedling and sapling size classes (Olson et al., 2014; Olson et al., 2017). Still, while
replacement of oaks by maples and other shade-tolerant species is an ongoing transition
elsewhere in the eastern U.S. (Nowacki and Abrams, 2008; Rogers et al., 2008; Shotola et al.,
1992) low-quality soils and frequent droughty conditions of the Missouri Ozarks may inhibit
recruitment of maples to the overstory (Olson et al., 2014). To the extent such recruitment
occurs, managers will have greater ability to affect the relative abundance of maples under
uneven-aged and even-aged systems compared to no-harvest management.

A less speculative pathway of divergence among MOFEP treatments is in the species
composition of oaks, specifically the relative numbers of red oak species (with black oak, *Q.*
velutina, and scarlet oak, *Q. coccinea* most abundant on MOFEP sites) vs. white oak species
(particularly white oak, *Q. alba*, and post oak, *Q. stellata*) (Jensen and Kabrick, 2008). On no-
harvest sites, oak decline and resulting senescence of the shorter-lived red oaks has shifted
overstory dominance to white oaks (Olson et al., 2017). With current low densities of shade-
intolerant and intermediate-tolerance red oak species as sapling or midstory trees, replacement
by white oaks and other species on no-harvest sites likely will continue. The experiment's
uneven-aged management treatment may not maintain red oaks as a major species group unless
harvests reduce overstory densities and increase light levels to a greater degree than in single-tree
selection harvests of the first entry (Fan et al., 2015; Jensen and Kabrick, 2008). However, given
observed successful regeneration in clearcut openings, higher site-level abundance of red oaks
may be maintained under even-aged management (Fan et al., 2015; Jensen and Kabrick, 2008).

Shifting dominance of white oaks vs. red oaks and potential addition of maples to the
overstory assemblage are important factors for considering future patterns of abundance of
*Peromyscus* among the MOFEP treatments. Abundance of hard mast can strongly affect
*Peromyscus* populations (Clotfelter et al., 2007; McCracken et al., 1999; McShea, 2000, Ostfeld
et al., 1996; Schnurr et al., 2002; Wolff, 1996). Even if combined density of red and white oaks
does not decline, replacement of red oaks by white oaks could reduce size and year-to-year
consistency of overall acorn crops (Greenberg and Parresol, 2002; Greenberg et al., 2014; Olson
et al., 2015). Moreover, acorns from the two groups may not have interchangeable value. Higher
tannic acid levels of red oak acorns may reduce their palatability, but the resulting slower rate of
rotting, along with their pattern of overwintering before germinating, also may make red oak
acorns a more reliable over-winter food source (Kirkpatrick and Pekins, 2002; Olson, 1974;
Wang et al., 2009).
However, there is high uncertainty regarding the potential responses of *Peromyscus* populations to changes in forest composition and hard mast crops of MOFEP sites. During first 15 years post-treatment, even-aged MOFEP treatments reduced overall acorn crop size moderately, with further reduction likely during the next few entries until acorn production begins by cohorts resulting from early entries (Olson et al., 2015). Based on our results, this decrease either was not significant enough to affect *Peromyscus* spring abundances or was counterbalanced by other effects of the initial MOFEP management entries. Links between *Peromyscus* abundance and acorn crops, particularly of red oak species, are strongest at more northern latitudes, with decoupling of *Peromyscus* abundance and acorns crops in some areas further south, conjecturally due a more diverse food base (Gillen and Hellgren, 2013; Wang et al., 2009). In the southern Central Hardwoods region, *Peromyscus* abundance may not show a strong relationship with stand-level density of hard mast species or with hard mast biomass (Gillen and Hellgren, 2013). Maple seeds are also consumed by *Peromyscus* (e.g. Cramer, 2014) and the magnitude of maple seed crops may contribute to *Peromyscus* population changes (McCracken et al., 1999; but see Schnurr et al., 2002). Speculatively, this additional food resource could compensate for some reduction in hard mast or even further diversify the food base.

5. **Conclusions**

Across studies of forest floor small mammal community responses to disturbances, the most consistently observed pattern is the one we observed with MOFEP *Peromyscus*: generalist species remaining numerically dominant across successional states, frequently with increased abundance after disturbance (Bogdziewicz and Zwolak, 2014; Kirkland, 1990). In light of
*Peromyscus* responses so far to MOFEP treatments and expected positive effects of periodic management entries on important food and cover resources, uneven-aged and even-aged management likely will continue to maintain increased abundance of *Peromyscus* compared to no-harvest management. Whether this difference changes over time or whether divergence in abundance under uneven-aged vs. even-aged management occurs may depend partly on the sensitivity of these *Peromyscus* populations to differential changes in overstory composition, particularly in terms of the comparative densities of red oaks, white oaks, and maples (Gillen and Helgren, 2013). In addition, although *Peromyscus* populations may respond positively to increased habitat heterogeneity and structural complexity (Carey and Johnson, 1995), an open question with broad relevance to forest management is the relative importance of higher within-stand vertical heterogeneity expected under long-term application of uneven-aged management (but see Angers et al., 2005) vs. higher between-stand heterogeneity and availability of larger early successional patches produced by even-aged management (e.g. Holmes et al., 2012; Joelsson et al., 2018; Schall et al., 2017; Thompson et al., 1995).

Our statements of general expectations are not intended to downplay high uncertainty about potential changes in MOFEP forests and wildlife communities over the next several decades. A unique value of long-term studies is the chance to iteratively improve understanding of a system, quantify unexpected changes, and assess effects of infrequent and rare events (Hobbie et al., 2003; Lindenmayer and Likens, 2010:8; Turner et al., 1998). As additional years of data are collected, more insights will be possible about the relative magnitude of treatment effects vs. system-wide year-to-year variation, and about interactions between harvest treatments and other factors (e.g. mast crops; weather) potentially driving temporal variation in *Peromyscus* abundance. Understanding drivers of high year to year variation will increase our ability to detect
and explain differential long-term trends in abundance (Urquhart et al., 1998) and forecast future potential changes. The MOFEP experiment has an especially long time horizon, but has reached a point of providing a unique understanding of cumulative and long-term forestry impacts at a landscape scale.

Acknowledgments

We thank the Missouri Department of Conservation for supporting this research. We also recognize the work of MDC employees who initially proposed, designed, and implemented MOFEP, and the three dozen technicians who collected small mammal data for this study. The University of Missouri and Auburn University's School of Forestry and Wildlife Sciences provided additional support. We thank two anonymous reviewers for constructive comments.

Funding

This work was supported by internal MDC funding and grants from MDC to the University of Missouri for 2009–2014 sampling. Animal handling methods were approved by the MDC MOFEP Steering Committee and University of Missouri Institutional Animal Care and Use Committee (protocols 6745 and 7875).

Literature cited


Appendix

A.1. Analysis overview

To examine effects of MOFEP experimental treatments (i.e. alternative forest management systems) on *Peromyscus* abundance, we first attempted a Bayesian hierarchical modeling approach following Schofield and Barker (2014) and Converse and Royle (2012), which integrates modeling of the observation process (via mark-recapture estimation) and the underlying ecological process of interest (variability in abundance across sites and years). For such integrated modeling of abundance in mark-recapture studies, Converse and Royle (2012) outlined Markov Chain Monte Carlo estimation (MCMC) via parameter-expanded data-augmentation, with abundance as a derived parameter, while Schofield and Barker (2014) outlined a trans-dimensional MCMC approach with explicit inclusion of abundance as a model parameter. In extensive pilot analyses, regardless of which approach we used, computer processing time and extremely slow model convergence made the analysis unfeasible using JAGS and BUGS (WinBUGS, OpenBUGS) when we attempted to model more than a few years of data at a time. We believe this issue was due to the large number of mice captured (3,871) and trap sessions (198) conducted in MOFEP, possibly exacerbated by relatively low capture probabilities, as we have used these approaches successfully in other studies with smaller multi-session mark-recapture data sets for *Peromyscus* (R. A. Gitzen, unpublished data; see also Sollmann et al., 2015). Therefore, we performed a two-stage analysis, first estimating abundance via mark-recapture modeling, then modeling variation in abundance as a function of treatment and other factors.
A.2. Bayesian Mark-recapture Modeling

We used a zero-augmentation approach to Bayesian closed-population mark-recapture modeling (Kéry and Schaub, 2012; Royle and Dorazio, 2008; Royle et al., 2007). Each session’s observed capture histories were augmented with additional all-zero capture histories. The 198 sets of resulting capture histories were combined into a single response-variable matrix ($C$) for unified mark-recapture analysis. With the general zero-augmented approach, each capture history is associated with a partially latent session "occupancy" parameter denoting whether the capture history is for an individual animal actually present during the session or represents a non-existent individual. That is, the approach formulates the abundance-analysis problem as that of estimating how many of the augmented all-zero capture histories for a session correspond to individuals present but not captured. In each MCMC iteration, the abundance estimate for a session was the sum of the number of individuals actually captured plus the number estimated to be present but not captured.

The primary detectability model we used was a form of the general model $M_{0b}$ (Otis et al., 1978) incorporating within-session variability in capture probabilities due to temporal and behavioral (i.e. response to prior capture) effects (Eq. 1).
Here, the capture history value $C_{i,n}$ equaled 1 if individual $i$ was captured in trap night $n$ of the trap session of block $b$, site $s$, grid $g$, and year $y$, and equaled 0 otherwise. The session occupancy parameter, $Z_i$, equaled 1 for captured individuals and other individuals estimated to be present during the session but not captured, and equaled zero otherwise. As is standard with the zero-augmented approach, $Z_i$ was modeled as a Bernoulli outcome with session occupancy probability $\Omega_{bsgy}$, with the latter given a Uniform$(0,1)$ prior distribution. Logit-scale detectability for individual $i$ in trap night $n$ of the session was modeled as an additive function of an overall logit-scale fixed-effect mean ($Lp\_mu$), a block × year random effect ($lp\_blockyear$), a nested block × year × night random effect ($lp\_blockyearnight$) allowing night-to-night variation within a session, a nested session-specific random effect ($lp\_session$) to allow additional grid-to-grid variation within a block × year combination, an overall fixed behavioral (recapture) effect ($Lrp\_mu$), and a session-specific random effect adjustment ($lrp\_session$) to this average recapture effect. The variable $Recap_{i,n}$ equaled 1 if individual $i$ was captured in the previous trap night of the session, and equaled 0 otherwise. Logit-scale night-to-night variability in detection probability within a session was constrained to be have homogenous effects on detectability across the six grids sampled in block $b$ during year $y$. This approach allowed within-session...
variation but reduced model complexity. This approach was appropriate because the six grids within each block were trapped simultaneously each year, thus subject to parallel variation in weather, moon phase, etc.

In the Bayesian model, random effects standard deviation parameters were given separate \text{abs}[\text{Normal}(\text{mean} = 0, \text{precision} = 0.1)] prior distributions, while the fixed-effect average recapture effect \((Lrp\_mu)\) was given a \text{Student’s}_t(\text{mean} = 0, \text{precision} = 0.16, \text{df} = 5) prior distribution. These were weakly informative prior distributions that avoided putting significant prior probability mass outside a liberally plausible range for each logit-scale parameter. Average probability of first capture was given a \text{Uniform}(0.0, 0.5) distribution and logit-transformed to \(Lp\_mu\); its upper bound was constrained at 0.5 to speed convergence after preliminary modeling with a \text{Uniform}(0.0, 1.0) distribution indicated probability of \(\sim 1.0\) that the parameter was \(< 0.5\). Therefore, constraining the range of the prior did not limit the resulting posterior distribution.

Posterior distributions were estimated with MCMC sampling using program JAGS and R package \text{jags} (Plummer, 2016). Following an adaptive phase of 2500 samples and 105,000 burn-in samples, with convergence confirmed via traceplots and the Gelman-Rubin convergence diagnostic (Gelman and Rubin, 1992), the posterior distribution for each parameter was estimated with 30,000 samples (every 10th sample retained of 100,000 post-convergence samples for each of three chains). We estimated the posterior distribution for the average recapture odds ratio (odds of capture in trap night \(n\) if an individual was captured on the previous trap night / odds of capture if not captured on the previous trap night) by exponentiating \((Lrp\_mu)\) within each MCMC iteration. For this and all subsequent analyses in this manuscript, posterior distributions were summarized by the posterior median MCMC sample value for each parameter.
and a 95% credible interval formed by taking the (0.025, 0.975) percentiles of the posterior samples.

A.3. Modeling abundance of *Peromyscus*

Our second stage of analyses used these mark-recapture abundance posterior distributions to form response variables for modeling variation in abundance of *Peromyscus* across sites and years. Modeling focused on our objectives of estimating treatment effects, comparing the magnitude of treatment effects to background spatial and temporal variation in the system, and assessing evidence of long-term trends in abundance. For these three objectives, we examined three alternative forms of a general negative binomial log-linear mixed effects model. This allowed us to tailor the general model to each objective.

A.3.1. Estimating treatment effects on *Peromyscus* abundance

We estimated overall and yearly site-level effects of the three MOFEP management systems on expected abundance ($\lambda$), using estimated abundances by grid for the nine post-treatment years as the response variable and average pre-treatment (1994 and 1995) estimated abundances as a covariate (sensu Senn, 2006) (Eq. 2).
Here, $N_{bsgy}$ equaled grid-level abundance for grid $g$ in block $b$, site $s$, and year $y$, grid $g$, limited to the 162 post-treatment (1998–2014) trapping sessions; $r$ designated the negative binomial dispersion parameter; covariate $\bar{N}_{pre,bsg}$ equaled average pre-treatment (1994–1995) abundance for grid $bsg$; and $UA_s$ and $EA_s$ indicated the experimental treatment to which site $s$ was assigned. The log-linear model included block effects (incorporated as a fixed effect to avoid estimating a variance component based on only three blocks) and random site, year, and block × year effects. The year random effect was crossed with block and site effects, such that the year effect modeled the component of year-to-year variation that was consistent across the study area, while the block and site effects represented block- and site-identity effects that were persistent across years (Urquhart et al., 1998). Other potential random effects were omitted to avoid further model complexity when preliminary analyses indicated that the effects (e.g. persistent effects of grid; site × year grouping) were essentially zero or not estimable given other effects incorporated.
Log-scale effects of uneven-aged and even-aged management were modeled with fixed-effects parameters for the average effect of each harvest treatment relative to the no-harvest treatment ($B_{UA_{ave}}$, $B_{EA_{ave}}$) and additional random effects allowing for yearly variation in the effects of each treatment ($b_{UA_y}$, $b_{EA_y}$) for the nine post-treatment sampling years. For interpreting model estimates, posterior distributions for proportional effects of harvest treatments relative to the no-harvest treatment were estimated by exponentiating each posterior sample of the log-scale overall and year-specific treatment effects; the same approach was used for all subsequent analyses to transform log-scale additive effects into proportional effects. To compare uneven-aged vs. even-aged management effects, we calculated posterior distributions for the ratio of even-aged to uneven-aged proportional effects (ratios of average effects and of year-specific effects).

Posterior distributions for parameters in the negative binomial mixed effects model were estimated with MCMC in Stan via R packages brms (Buerkner, 2017) and RStan (Stan Development Team, 2016). We conducted two sets of analyses with the yearly treatment model. Firsat, we analyzed the single vector of posterior median abundance estimates from Bayesian mark-recapture modeling; the pre-treatment covariate values were the average of the two pre-treatment posterior median estimates for each grid. We estimated posterior distributions for this $N_{posterior\_median}$ analysis using the default 4 chains in package brms, with 1,000 warm-up samples and 1,000 retained samples per chain (4,000 total posterior samples retained). Convergence was assessed as in the Bayesian mark-recapture analysis but far fewer samples were needed because of high efficiency and rapid convergence in Stan.

Although this analysis accounted for among-session variation in capture probabilities because it used absolute abundance estimates, the analysis did not incorporate mark-recapture
estimation uncertainty. To assess whether posterior uncertainty about model parameters increased meaningfully when this mark-recapture estimation uncertainty was incorporated, we performed a second analysis of the treatment-effects model ($N_{\text{posterior\_sample}}$) using individual samples from the 198 posterior distributions for session-specific abundance from mark-recapture modeling. The Bayesian mark-recapture sampling produced 30,000 posterior abundance vectors, each containing a sample vector of length 198. For each of 3,000 of these vectors (every 10th multivariate sample), we used the posterior abundance vector as input data for the negative binomial abundance model described above, with the pre-treatment average abundance for each grid calculated using the relevant pre-treatment abundance values from that sample vector. Using the brms package, we obtained 1,000 samples from the posterior distributions for the abundance model parameters (a single MCMC chain with 1,000 warm-up samples and 4,000 posterior samples thinned to retain every 4th value; initial analyses of multiple chains indicated consistently adequate convergence after 1,000 warmup samples, and running one chain reduced computer run time by several days). We then pooled these 400 sets of posterior samples of negative binomial model parameters to obtain an overall posterior sample for each parameter of length $400 \times 1000 = 400,000$, which incorporated uncertainty in the 198 abundance estimates. For both the $N_{\text{posterior\_median}}$ and $N_{\text{posterior\_sample}}$ analyses, we used default prior distributions of the brms package (Buerkner, 2017).

A.3.2. Relative magnitude of treatment effects on Peromyscus abundance

To compare the relative magnitude of effects of uneven-aged and even-aged harvest management on abundance vs. effects of background spatial and temporal variation in the MOFEP system, we used posterior summaries for treatment effects from the above analysis. That analysis used the
pre-treatment average abundance covariate to increase precision of treatment-effect estimates. However, incorporating pre-treatment abundance as a covariate meant that year effects were modeled based on (and estimated for) 9 rather than 11 years, and the model essentially partitioned persistent spatial variation into a component explained by pre-treatment abundance and components capturing any additional block and site effects not explained by pre-treatment abundance. To better estimate background block, site, and year effects for comparison with treatment effects, we ran a modified form of the $N_{\text{posterior\_median}}$ analysis. The modified approach included all 11 sampling years and omitted pre-treatment abundance as a covariate. We calculated posterior summaries for block-, site- and year-specific proportional effects, where each effect was a proportional change from the long-term, area-wide average expected abundance. We visually compared estimated effects of year, block, and site effects with estimated treatment effects from the $N_{\text{posterior\_median}}$ analysis. The treatment-effects analyses ($N_{\text{posterior\_median}}$ vs. $N_{\text{posterior\_sample}}$) indicated that incorporating estimation uncertainty increased the width of posterior credible intervals somewhat (Table 2; Fig. 2), but not enough to affect our interpretations of the general relative strength of different effects. Therefore, we did not conduct time-intensive $N_{\text{posterior\_sample}}$ analyses to incorporate estimation uncertainty for addressing this objective.

A.3.3. Trend analysis

The treatment-effects analysis examined average and year-specific relative differences among management systems, but to avoid additional model complexity we did not include temporal trend parameters in that model. To assess whether *Peromyscus* abundance was exhibiting
system-wide or treatment-specific log-linear trends within the first 20 years of the MOFEP study, we analyzed a modified form of the negative binomial abundance models (Eq. 3).

**Eq. 3**

\[
\ln(\lambda_{\text{bagy}}) = B_0 + B_{\text{block}2_{b=2}} + B_{\text{block}3_{b=3}} + B_{\text{treatment}(s)} \cdot Year_y + b_{UA_y} \cdot UA_y + b_{EA_y} \cdot EA_y + b_{year_y} + b_{blockyear_{by}} + b_{site_s}
\]

\[N_{\text{bagy}} \sim \text{NegBinom}(\lambda_{\text{bagy}}, r)\]

In this modified model, \(Year_y\) was a numeric covariate (scaled to be 0 for 1994); \(B_{\text{treatment}}\) indicated a treatment-specific log-linear trend parameter allowing separate trends for no-harvest, uneven-aged, and even-aged management, respectively; and other parameters and random effects distributions were as described above. See, for example, Urquhart et al. (1998) for development of a general multi-site trend model in which year is included as both a numeric covariate for modeling systematic long-term trends and as a categorical random effect to capture additional yearly variation not accounted for by the systematic trend. In the above model, site- and year-specific treatment random effects modeled among-treatment differences apart from those captured by the log-linear trend effects. Using posterior median abundance estimates for 1994–2014 as input data, we estimated posterior distributions for the trend-model parameters using package brms as in previous \(N_{\text{posterior_median}}\) analyses (i.e. four chains, with 4,000 total samples retained for inference). Posterior samples for log-scale trend parameters were exponentiated to estimate posterior distributions for proportional trends (e.g. value of 1.02 = 2% increase in expected abundance per year). Again, we did not conduct time-intensive \(N_{\text{posterior_sample}}\) analyses to incorporate estimation uncertainty in trend modeling as it was apparent that minor increases in posterior uncertainty would not change conclusions from the \(N_{\text{posterior_median}}\) trend analysis.
With all abundance-modeling analyses, we did not explicitly account for correlation among abundance estimates induced by unified mark-recapture modeling (sensu and Converse et al., 2006; Skalski and Robson, 1992:153). However, we believe this had little or no effect on our results. The primary grouping structure assumed for variation in capture probability among sessions within the mark-recapture analysis was block × year variation. However, raw capture counts exhibited high variation among blocks, years, and block × year groups, imposing a strong correlation in capture counts in sessions within groups independent of mark-recapture modeling. Unified mark-recapture modeling appeared to contribute little to this already strong structure, based on preliminary comparisons of the block × year variance component estimates from negative binomial treatment-effects abundance analyses using raw capture counts vs. those using posterior median abundance estimates. Incorporation of block, year, and block × year effects in the negative binomial abundance modeling therefore addressed both high natural spatial-temporal within-group correlation and additional within-group correlation in abundance estimates induced by the unified mark-recapture modeling. Independent estimation of abundance for each session was not feasible given the large number of trap sessions with few individual animals captured (e.g. 51 of 198 sessions had 5 or fewer mice captured).