Effectiveness of conservation measures to support biodiversity in boreal timber-production forests

Ane Christensen Tange\textsuperscript{1,2}, Hanne K. Sjølie\textsuperscript{1} and Gunnar Austrheim\textsuperscript{3}

Abstract

Large parts of the boreal forest ecosystems have been greatly affected by human use, and the current timber-oriented forest management practice that dominates boreal forests is proven to cause biodiversity and ecosystem services declines. These negative effects are mitigated in various ways, including in-situ measures implemented upon harvest. The measures comprise trade-offs between economic and ecological aims; thus, requiring solid knowledge of their effectiveness. However, comprehensive literature review of the effectiveness of such measures is scarce. We aim to fill part of this void by reviewing the scientific literature that have gauged effects of four in-situ conservation measures: green tree retention (GTR), patch retention (PR), dead wood retention (DW) and riparian buffer zones (RB). Two outcomes were considered, species richness and species abundance across taxa.

From a total of 3012 initial papers, 48 met our inclusion criteria that generated 238 unique results. Results were grouped according to control. 178 studies used mature, unlogged forest as control. Out of those, 68\% of the findings were not significant, i.e., suggesting no significant impact of harvest with biodiversity measures on species richness and species abundance compared to no harvest. Eighteen percent of the observations showed negative effects and 14\% of the observations showed positive effects compared to no harvest. Sixty studies used harvest with no measures as control, of which 45\% showed significant positive effects, meaning that compared to harvest with no measures, harvest with conservation measures has positively effects on species richness and abundance. However, 43\% of the studies found no significant effect of the implemented conservation measures compared to harvest with no measures taken.

The relatively few significant results reported restrain distinct conclusions on the effectiveness of the assessed conservation measures, but some degree of conservation measure is likely to have positive effects on biodiversity in timber-production forest. However, the scientific basis does not allow for pointing to threshold levels. Higher transparency of study design and statistical results would allow us to include more studies. There is a clear need for more research of effectiveness of
common conservation measures in timber-production forests in order to strengthen the knowledge basis. In particular, there are few studies that employ harvest without any conservation measure as control. This is pivotal knowledge for forest managers as well as for policymakers for preserving biodiversity and the ecosystems in forest.

**Keywords** dead wood enhancement; forest certification; forestry; green tree retention; riparian buffer zone; sustainable forest management

**Addresses** 1Inland Norway University of Applied Sciences, Faculty of Applied Ecology, Agricultural Sciences and Biotechnology, Department of Forestry and Wildlife Management, Evenstad, Norway; 2Glommen Mjøsen Skog SA, Elverum, Norway; 3University Museum Norwegian University of Science and Technology, Department of Natural History, Trondheim, Norway

E-mail ane.tange@inn.no / at@glommen-mjosen.no

Received 6 October 2023 Revised 14 February 2024 Accepted 19 February 2024

1 Introduction

Intensive forest management aimed towards timber production alters forests’ composition of species, age classes and structures in addition to minimizing dead wood quantity (Franklin 1993; Gauthier et al. 2015). These changes are proven to cause decline in biodiversity and other ecosystem services (Lindenmayer et al. 2012; Pohjannies et al. 2021). During the 1980s and the 1990s, increased awareness of the impact of industrialised forestry on biodiversity and pressure from *i.a.* environmental non-governmental organizations (NGOs) led to shifts in silvicultural management practices and to the introduction of forest certification schemes and changes in legislation (Larsson and Danell 2001; Arnesen et al. 2004; FAO 2020). The most used certification schemes in the world are the Forest Stewardship Council (FSC) and the Programme for the Endorsement of Forest Certification (PEFC) (FSC® International 2023a; McDermott et al. 2023; PEFC Council 2023). Globally, 160 million and 292 million hectares of forestland are certified according to the FSC and PEFC standards, respectively (FSC® International 2023a; PEFC Council 2023). The FSC and the PEFC cover respectively 14% and 25% of the 1.15 billion hectares of forestland that is managed primarily for the production of wood and non-wood forest products (FAO 2020).

The certification systems are pivotal strategies for maintaining biodiversity and ecosystem services in production forests through sets of detailed, country-specific requirements for forest management (Clark and Kozar 2011; Rametsteiner and Simula 2003). Most of the certification schemes’ focus is on measures to support ecological function and habitats that contain specific features (Lindenmayer et al. 2006; PEFC Norge 2022). Several requirements are related to sustaining or “life-boating” old-growth forest-specialised species through the early succession stages or mitigating negative impacts on biodiversity (Franklin 1993; Lindenmayer et al. 2012; Gustafsson et al. 2020). Measures are typically categorised based on their conservation aim, e.g., old trees as habitats for saproxylic species. However, the effects of these measures to support biodiversity and ecological services within timber-production forests are debated (Palik and D’Amato 2017; Lehtonen et al. 2021; McDermott et al. 2023).

Several studies have investigated the effects of conservation measures to preserve biodiversity (Fedrowitz et al. 2014; Seibold et al. 2015; Mori et al. 2017; Chellaiah and Kuglerová 2021). Fedrowitz et al. (2014) found in their meta-analysis of retention forestry that most species groups (forest-depending, open-habitat and generalist) benefited from retention of green trees, both in species richness and abundance. Sverdrup-Thygeson et al. (2008) was an early evaluation of the measures set by the Norwegian PEFC (Programme for the Endorsement of Forest Certification) a few years after its introduction. Other studies have assessed the effects of the FSC (Forest Stewardship Council) and other certification systems on biodiversity (Schlyter et al. 2009 (Sweden); Clark
and Kozar 2011 (Canada); Elbakidze et al. 2011 (Sweden and NW Russia); Johansson et al. 2013 (Sweden); Mikulková et al. 2016 (Czech Republic); Blumröder et al. 2020 (Arkhangelsk, Russia); Lehtonen et al. 2021 (Finland, Sweden, Estonia and Latvia)). To the best of our knowledge, no systematic reviews of the effectiveness of the conservation measures commonly taken in current forestry operations subject to certification in the boreal biome exist and we aim to fill part of this gap.

The objective of the review is to assess the effectiveness of certification requirements’ conservation measures to maintain species richness and abundance in boreal forest subject to management. Species richness and abundance are frequently used as biodiversity measures in analyses of impacts of forest management and conservation measures (Lassauce et al. 2011; Fedrowitz et al. 2014; Savilaakso et al. 2021). Species richness measure community and regional diversity by the count of the number of unique species in an area (Blowes et al. 2022). Species richness is a straightforward measure of diversity on different scales. Species richness is often used for conservation purposes, but it does not consider proportional abundance of species (Tuomisto 2010). Abundance is the frequency of each species within an area (Orians and Groom 2006). Factors like resource availability, competition and disturbance influence the number of species present in the landscape and their distribution (Allaby 2010). Losses of species and skewed abundances, with a dominance of a few species, may be symptoms of stressed ecological communities (Orians and Groom 2006). Forest management for timber production is a disturbance factor that can result in stressed communities (Gauthier et al. 2015; Blowes et al. 2022) and may drive changes in forest species abundance and richness (Fedrowitz et al. 2014). Our overall research question is: what is the effectiveness of conservation measures commonly applied upon timber harvest in boreal forests in supporting biodiversity?

With this basis, we draw the following hypotheses of effects of conservation measures on forest biodiversity: 1) compared to harvest with no conservation measures, harvest with conservation measures supports species richness and abundance better (Gustafsson et al. 2012, 2010; Fedrowitz et al. 2014), 2) the level of a conservation measure is positively correlated to the effects of the measure (Gustafsson et al. 2020, 2012; Lindenmayer et al. 2012; Fedrowitz et al. 2014). 3) forest undergone harvest with conservation measures has lower species richness and abundance than mature, non-harvested forest (Lindenmayer and Hunter 2010; Lindenmayer et al. 2012).

2 Methods

2.1 Conservation measures

In forests subject to environmental certification, a set of conservation measures to support biodiversity and non-wood ecosystem services must be undertaken upon forest planning and management. Albeit the FSC and PEFC certification schemes have general requirements, the operational requirements are set in national standards. Thus, the requirements vary across countries but share core characteristics (PEFC 2018; FSC 2023). The operational requirements for conserving biodiversity can be divided into specific conservation measures (Lehtonen et al. 2021), each designed to support ecological functions and/or maintain key habitats for forest specialists (Sverdrup-Thygeson et al. 2014, 2008; Lindenmayer et al. 2012). We focus on four such conservation measures that are commonly applied in boreal timber-production forests under PEFC (PEFC Council 2018) and FSC (FSC® International 2023b): green tree retention (GTR), patch retention (PR), dead wood retention (DW) and riparian buffer zones (RB) (Table 1).
2.2 Literature search and study selection

We retrieved the literature by a wide search in Web of Science and Scopus based on a systematic review protocol according to the guidelines of the Collaboration for environmental evidence (Pullin et al. 2022). The search strings were built up around the key terms “forest*” and “biodiversity” and restrained by limiting the search with exact key words like “forestry”, or “boreal forest” in the inbuilt search panel offered by Web of Science and Scopus. The search was limited to the boreal vegetation zone and peer-reviewed papers.

The search terms were spelled in following manners:

( TITLE-ABS-KEY ( forest* ) AND TITLE-ABS-KEY ( biodiversity ) ) AND ( LIMIT-TO ( EXACTKEYWORD, “Forestry” ) OR LIMIT-TO ( EXACTKEYWORD, “Boreal Forest” ) ) AND ( LIMIT-TO ( AFFILCOUNTRY, “United States” ) OR LIMIT-TO ( AFFILCOUNTRY, “Canada” ) OR LIMIT-TO ( AFFILCOUNTRY, “Sweden” ) OR LIMIT-TO ( AFFILCOUNTRY, “Finland” ) OR LIMIT-TO ( AFFILCOUNTRY, “Norway” ) OR LIMIT-TO ( AFFILCOUNTRY, “Russian Federation” ) )

Similar searches were conducted with the following key words: “clear-cut AND biodiversity,” “continuous cover forestry AND biodiversity”, “green tree retention AND biodiversity”, “retention forestry AND biodiversity”, “patch retention AND biodiversity”, “riparian buffer AND biodiversity”, “woodland key habitat AND biodiversity” and “woody debris AND biodiversity”. The search was conducted on 23 February 2022. This resulted in 3012 papers including field studies, reviews, simulations, and modelling in the boreal vegetation zone. In addition, relevant studies were included from reference lists of other studies (Fig. 1).

A first screening of the studies was done by reading headlines to discard articles out of scope (Pullin and Stewart 2006). Second, the procedure was repeated for abstracts. We maintained a conservative approach and only excluded studies that obviously were out of scope. After this filtering, 188 studies were retained. Third, we read the articles in full, with emphasis on the study area, study design and presentation of results to deem their relevance and quality. Studies were

| Table 1. Description and rationale of the four analysed conservation measures. |
|---------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Conservation measure            | Description                                                                                                                                                                                                 |
| Green tree retention (GTR)      | This conservation measure aims to ‘lifeboat’ species over the regeneration phase after harvest, and to mimic natural disturbances such as storm felling, fire and insect outbreak (Vanha-Majamaa and Jalonen 2001; Gustafsson et al. 2010). By retaining a number of live trees (“green”) after harvest, this measure secures habitats for old tree-dependent species and connectivity (Gustafsson et al. 2012; Lindenmayer et al. 2012). Retention is a term used on a variety of measures to retain some pre-harvest forest structures after harvest (Fedrowitz et al. 2014). Often, 10 trees ha⁻¹ are left in groups or as single trees (Vanha-Majamaa and Jalonen 2001; Gustafsson et al. 2020). |
| Patch retention (PR)            | Forest patches with green trees are left after harvest, with the aim to maintain landscape-level biodiversity (Lindenmayer et al. 2012). The definition and use of PR differ greatly between countries (Timonen et al. 2010; Hakkila et al. 2019), varying in size, biological value and administration. |
| Dead wood retention (DW)        | Dead wood can be constructed artificially as high stumps, cut-logs or retention of any dead wood present before logging. The aim is to provide substrate for saproxylic species after harvest (Juutilainen et al. 2011; Lassauce et al. 2011; Seibold et al. 2015; Hekkala et al. 2016). Dead wood is especially important for many threatened forest species (Nieto and Alexander 2010). |
| Riparian buffer zone (RB)       | A forested zone along watersheds is retained after harvest and aims at sustaining the important ecotone between a water body and the upland forest (Gundersen et al. 2010; Kuglerová et al. 2014). The riparian zone holds different kinds of species than the upland forest and contributes to regional increase in species richness (Sabo et al. 2005). The width of the buffer affects its ability to sustain the riparian buffer’ ecological function (Kuglerová et al. 2014). |
included based on the following criteria: (1) conducted in the boreal vegetation zone, (2) empirical field study on the effects of at least one of the targeted conservation measures on at least one species group (taxa), (3) measured outcomes on biodiversity either as species richness, species abundance or both, and (4) has retrievable statistics and a clear control. To meet the last point, p-values or other metrics like Tukey’s post hoc test that shows statistically significant effects of conservation measures compared to a control had to be described in text, figures, or tables.

We required clearly defined controls in the studies, and two groups of studies were included based on control: “no-harvest” and “clear-cut”. The no-harvest control includes mature, unlogged forest. However, this category includes a wide range of forest types, from managed forest that is biologically young to old-growth forest with a high degree of naturalness. To analyse the impact of type of mature forest, the no-harvest control was divided in two sub-groups of managed and unmanaged forest (Table 2). The no-harvest control provides insight into the capacity of production forests with the conservation measures to support equally rich communities as unlogged stands (Spake and Doncaster 2017). To test differences in attributes and impacts of no-harvest controls, chi-squared and t-tests were run. The clear-cut control is harvest with no conservation measures and gives a baseline for comparing the effects of measure compared to harvests where measures are not being taken. The included indicators of richness and abundance were identical for the two control groups. Further, the articles are divided into short-term and long-term studies based on time span between logging and sampling, short-term referring to up to ten years and long-term ten years and beyond.

2.3 Quantification of effects

The effects of the conservation measures were categorized as positive, negative or no significant effects (NS.) (Munn et al. 2018; Pullin and Stewart 2006) allowing for drawing conclusions across the variable set of studies. The positive and negative categories include significant reported results of a conservation measure from the included articles, where one observation gives one count. The NS. category counts reported non-significant results.

<table>
<thead>
<tr>
<th>Sub-groups</th>
<th>Definition</th>
<th>Number of articles</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mature, unmanaged forest</td>
<td>Mature forest regenerated after natural disturbance e.g. wind, fire, or pests (Kuuluvainen and Gauthier 2018)</td>
<td>25</td>
<td>140</td>
</tr>
<tr>
<td>Mature, managed forest</td>
<td>Mature forest regenerated after silvicultural practice e.g. selective logging or clear-cutting, regenerated naturally or by planting (Chazdon et al. 2016)</td>
<td>21</td>
<td>38</td>
</tr>
</tbody>
</table>
To investigate any threshold in the conservation measures, they were grouped into levels, except the dead-wood category where we only considered enhancement and not the level of enhancement. The GTR studies exhibit great variety in the number of retained trees, varying from approximately 2% to 75% retention, as well as in the patterns of retention that include both scattered single trees and groups (Vanha-Majamaa and Jalonen 2001). We categorized retention up to 10 trees ha\(^{-1}\) or 2–10% as low retention, 10–30% as medium retention and above 30% as high retention level. The GTR studies were further separated in single tree retention and patch retention. We only considered patch size and not the quality of the PRs. Patches were categorised as small (< 0.4 ha), medium (0.4–0.8 ha) and large (> 0.8 ha). RB width was categorised as narrow (1–10 m) or wide (> 10 m). Several articles had multiple observations, such that the combination of conservation measure level, timeframe and species groups could be examined. To compare impacts in higher trophic levels, we grouped species based on groups derived from the reported species in the articles.

To test for independency between the effects of conservation measures and control types (McHugh 2013), chi-squared tests were run in R studio with base r package. The test was run separately for species richness and abundance, across all conservation measures and taxa. A substantial part of the articles originates from one study location, and we therefore did separate analyses of these studies’ impacts. Chi-squared tests were conducted to test dependencies between outcomes and whether an observation belongs to this study location or not. The observations from this study location were also removed from the observation pool to check whether the original results of dependency between control category and effects of conservation measures were maintained.

Figures and summaries were also conducted with R studio with the packages ggplot2, dplyr and tidyr (R Development Core Team 2022).

3 Results

3.1 Articles and general observations

Table 3 presents the distribution of the retained articles in the third and fourth rounds of the selection process, and the number of articles per conservation measure and of observations. The 48 articles that met our selection criteria gave 238 observations on the effects of conservation measures on species richness and abundance. All observations are given in the Supplementary file S1.

<table>
<thead>
<tr>
<th>Conservation measure</th>
<th>Number of articles after screening titles and abstract (n = 188)</th>
<th>Number of articles after screening methods and result section (n = 48)</th>
<th>Number of observations (n = 238)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green tree retention (GTR)</td>
<td>71</td>
<td>26</td>
<td>157</td>
</tr>
<tr>
<td>Patch retention (PR)</td>
<td>20</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>Dead wood enhancement (DW)</td>
<td>65</td>
<td>8</td>
<td>23</td>
</tr>
<tr>
<td>Riparian buffer zone (RB)</td>
<td>32</td>
<td>9</td>
<td>26</td>
</tr>
</tbody>
</table>
There were fewest articles on PR (6) and most articles on GTR (26). Further, most of the articles used no-harvest as control (40) and only a minority of the articles used harvest with no conservation measures as a defined control (8). However, in several articles where no-harvest was the defined control, harvest with no conservation measures was one of the treatments and could be used as a control in our review as long as the statistical results were clearly displayed. Thus, we have 60 observations with harvest with no conservation measures as control.

Among the original 188 papers, several articles met most criteria, but were excluded due to lack of a clearly defined control, no logging activity in the study, unclear documentation of conservation measures or insufficient reporting on statistics. Further, articles examining the effects on growth rate, movement of mammals, or territorial or other type of behaviour were excluded from our review.

Out of the ten boreal forest countries, only three were represented in the final selection of articles (Fig. 2). The majority of articles were from Canada (27), while 13 and eight articles were from Sweden and Finland, respectively. With our criteria no articles could be included from Mongolia, Russia, China, Japan, the US or Norway. The search resulted in several studies conducted in the temperate biome, which all were left out. We also excluded studies from hemi-boreal regions, due to difficulties in distinguishing between temporal and hemi-boreal forest.

Fig. 2. The extent of the boreal forest (green) (Patakov et al. 2008) and the countries included in the 48 final articles in the review (grey shading). Number of articles given by country per conservation measure. Conservation measures: GTR = Green tree retention, PR = Patch retention, DW = Dead wood retention and RB = Riparian buffer zone.
3.2 Effects across all conservation measures

Divided by control, 178 (75%) of the 238 observations had no-harvest as control and 60 observations (25%) had harvest with no conservation measures as control (Fig. 3). Summing over abundance and richness, within each control type, it appears that the type of control impacts on the effects of the conservation measures. With no-harvest as control (n = 178), 14% (25) of the observations showed positive impacts while 68% (121) of the observations showed non-significant and 18% (32) of the observations showed negative effects on biodiversity (abundance or richness). By contrast, when harvest with no conservation measures was used as control (n = 60), 45% (27) of the observations were positively significant, 43% (26) non-significant and 12% (7) negatively significant. Dividing observations into the two biodiversity measures richness and abundance, the pattern that a higher proportion of observations were not significant with no-harvest as control than harvest with no conservation measures as control, was consistent. Chi-squared tests were used to test the dependency between control type and biodiversity effect, testing whether the effect varied with control across all conservation measures (Fig. 3). The tests revealed that the effects of conservation measures for species richness were depending on control type (χ² = 18.99, df = 2, p < 0.01) with more of the observed effects being non-significant when using no-harvest as control as compared to using harvest with no conservation measures as control as control. A similar pattern was unveiled for abundance (χ² = 5.62, df = 2, p ≈ 0.06).

The large number of observations on species richness with no-harvest as control was driven by the GTR studies, which count to 95 observations. These studies held 76 (80%) of the not-significant observations, 12 (80%) of the positive and 22 (85%) of the negative observations out of all species-richness studies with no-harvest as control (Fig. 4). A similar trend was observed.

![Figure 3](image-url)
Fig. 4. Observed effects (Positive, NS, (Not significant), Negative) on biodiversity measures (Species richness and abundance) per conservation measure, by control (No harvest; No measures (Harvest with no measures)). Conservation measures include green tree retention (GTR), Patch retention (PR), Dead wood retention (DW) and Riparian buffer zone (RB).
for abundance with no-harvest as control. However, the trend was disrupted for harvest with no conservation measures as control, where GTR findings accounted for 47% of the positive, 28% of the not significant and 33% of the negative observations. Chi-squared test was used to test dependency between control type and effect of conservation measures for GTR on species richness, the only conservation measure and biodiversity effects with a sufficient number of observations to run this test (Table 3) (McHugh 2013). Dependency between control type and effect of conservation measures was established for species richness ($\chi^2=16.47$, df=2, $p<0.01$), but not for abundance ($\chi^2=0.94$, df=2, $p\approx0.63$).

Further, we applied chi-squared tests to check dependency between effect and type of no-harvest control, where we tested the two sub-categories of no-harvest: mature, managed forest versus mature, unmanaged forest. With mature, unmanaged forest as control, harvest with conservation measures had significantly more negative impacts than in cases where mature, managed forest was control, on richness ($\chi^2=11.083$, df=2, $p<0.01$) and abundance ($\chi^2=5.9567$, df=2,

**Fig. 5.** Number of observations (n) with mean stand age and geographical location in the no-harvest control, divided by sub-category (mature managed forest and mature unmanaged forest).
p ≈ 0.051). We found no significant difference in mean stand age between the two no-harvest sub-categories (Fig. 5). The mean stand age was 102 years for the mature managed control and 110 years for the mature unmanaged control (t = 1.0751, df = 29.284, p ≈ 0.29). This suggests that other forest characteristics separate the two no-harvest sub-controls. Fig. 5 reveals the difference in geographical location of the two sub-controls. Out of the 32 observations with mature, managed forest as control, 17 were from Finland, 12 from Sweden and three from Canada. While in the mature, unmanaged control, far more observations were from Canada (137) than from Finland (five) and Sweden (two). Thus, the significant dependency between outcomes and geographical location ($\chi^2 = 118.77$, df = 2, p < 0.01) followed the patterns of dependency between outcomes and no-harvest sub-control.

In total, only 12 articles examined long-term effects of the conservation measures, giving 39 observations across all taxa, on species richness and abundance independent of control type. The number of observations was not sufficient to evaluate long-term effectiveness of conservation measures.

### 3.2.1 Green-tree retention (GTR)

There was in total 27 articles on green tree retention (GTR), that resulted in 157 observations on effects of GTR on species richness and abundance across all taxa (Table 4 and Fig. 4). There were 110 observations on species richness with no-harvest as control. Most of these (69%) showed no significant effect of retention level with patterns largely maintained across GTR level. The distribution of impacts was similar for abundance with same control (n = 25) with 68% showing no significant effect of GTR. There were more negative than positive results for species richness (20% and 11% respectively), but opposite for abundance (8% and 24% respectively).

In contrast there were only 13 observations on species richness and nine observations on abundance that had harvest with no measure as control. While the findings for no-harvest as control were maintained for impacts of abundance, 54% of the observations of species richness were significantly positive when compared to harvest with no measures, indicating the GTR improves biodiversity upon harvest.

One-hundred and thirty-two of the GTR observations were Canadian studies that include all considered species groups with most studies using no-harvest as control. The study designs

<table>
<thead>
<tr>
<th>Control</th>
<th>Effect</th>
<th>Species richness</th>
<th>Abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Harvest with no measures</td>
<td>Positive</td>
<td>3 (43%)</td>
<td>3 (75%)</td>
</tr>
<tr>
<td></td>
<td>NS.</td>
<td>4 (57%)</td>
<td>1 (25%)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sum</td>
<td>7 (100%)</td>
<td>4 (100%)</td>
<td>2 (100%)</td>
</tr>
<tr>
<td>No-harvest</td>
<td>Positive</td>
<td>2 (5%)</td>
<td>7 (19.5%)</td>
</tr>
<tr>
<td></td>
<td>NS.</td>
<td>28 (70%)</td>
<td>22 (61%)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>10 (25%)</td>
<td>7 (19.5%)</td>
</tr>
<tr>
<td>Sum</td>
<td>40 (100%)</td>
<td>36 (100%)</td>
<td>34 (100%)</td>
</tr>
</tbody>
</table>
varied greatly, from leaving the minimum required number of trees to high retention levels and continuous-cover forestry (50–75% retention). Further, eight of the 27 articles stemmed from one study location in Canada, the EMEND (Ecosystem-based research into boreal forest management) (Craig and Macdonald 2009). These counted for 85% of the Canadian observations and 48% of all observations in the GTR category. All the EMEND studies used no-harvest as control and all clear-cuts had 1–2% retention level and out of the 64 observations on species richness, 67% were not significant. Since the EMEND studies dominated the GTR observations, we performed chi-squared test of dependency between outcome and whether an observation belonged to EMEND studies or not, but found no such dependency ($\chi^2=2.55$, df=2, $p \approx 0.28$). We also excluded the EMEND studies from the observation pool to check whether the original results of dependency between control category and effects of conservation measures prevailed and found that the original findings were maintained for species richness ($\chi^2=13.85$, df=2, $p \approx 0.001$) and abundance ($\chi^2=4.79$, df=2, $p \approx 0.09$).

Only four articles examined effects of low levels of retention (about 10 trees per hectare), resulting in 24 observations; all studies being short-term. Out of the 11 GTR observations of harvest with no conservation measures as control, three were positive and the others were not significant. Ovaska et al. (2016) found a short-term positive effect of low retention (5%) on terrestrial gastropods abundance, but the effect was not significant at medium retention. Sullivan and Sullivan (2014) found a positive effect both on mammal species richness and abundance at one location but no significant effect on the two other locations when the effect of 5–15 trees per hectare retention was examined. When single tree, low-level retention was compared to no-harvest (n=14), Sullivan and Sullivan (2014) found one positive observation on mammal abundance at one location, and Hyvärinen et al. (2005) found a positive effect on beetle species richness and abundance, but Ovaska et al. (2016) found in contrast a negative effect on gastropods species richness and abundance. The other nine observations were not significant.

### 3.2.2 Patch retention (PR)

Six articles had observations that matched our criteria on patch retention (Table 3). The results are shown in Table 5, in total 32 observations. Five articles were from Canada and one from Finland. In summary, there were 11 observations on species richness and nine observations on abundance

<table>
<thead>
<tr>
<th>Control</th>
<th>Effect</th>
<th>Species richness</th>
<th>Abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Harvest with no measures</td>
<td>Positive</td>
<td>1 (33%)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>NS.</td>
<td>2 (67%)</td>
<td>2 (100%)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Sum</td>
<td><strong>3 (100%)</strong></td>
<td>2 (100%)</td>
</tr>
<tr>
<td>No-harvest</td>
<td>Positive</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>NS.</td>
<td>3 (75%)</td>
<td>3 (100%)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>1 (25%)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Sum</td>
<td><strong>4 (100%)</strong></td>
<td>3 (100%)</td>
</tr>
</tbody>
</table>
with no-harvest as control. There were few significant effects of PR across controls. For species richness, 82% of the observations reported no significant effects and 18% showed negative effects compared to no-harvest as control; corresponding number for harvest with no measure as control was 71% not significant effects and 29% positive effects. Thus, the trend from GTR studies was maintained for harvest with no measure as control.

The four positive observations on species richness and abundance compared to harvest with no measure taken were all from Ovaska et al. (2016) that examined the effect of PR size on gastropod species richness and abundance. In contrast, when compared to no-harvest, the effect was negative on both species richness (two observations) and abundance (four observations). This indicates that given harvest is undertaken, PR can support gastropod species richness and abundance, but the number of observations is not sufficient to draw clear conclusions. The studies of Franklin et al. (2018) and Lee et al. (2017) were from the same study location (EMEND) and were the only long-term PR studies in our dataset. No-harvest was used as control in both cases, but different species groups were examined. Franklin et al. (2018) reported two positive observations on small and medium PR size on beetle abundance and Lee et al. (2017) found no significant effect on medium PR size on vascular plant richness.

### 3.2.3 Dead wood (DW) enhancement

Out of the 66 screened articles on DW conservation measures, only eight articles met our inclusion criteria (Table 3). Six articles examined beetles and two articles examined fungi with five of the studies being conducted in Sweden and three in Finland. All articles examined methods to enhance managed forest with DW, like low and high stumps, cut logs and downed wood retention, resulting in a total of 23 observations (Table 6 and Fig. 4), of which 16 observations were of harvest with no measure as control and seven with no-harvest as control. The results were too scarce to point to any general patterns of effects of dead wood enhancement on species richness or abundance.

Suominen et al. (2019) was the only article examining long-term effects of low-stumps retention on fungi species richness and abundance. Compared to no-harvest, the effect was positive for species richness and abundance; compared to harvest with no conservation measures, the effect was negative for species richness and abundance.

<table>
<thead>
<tr>
<th>Control</th>
<th>Effect</th>
<th>Species richness</th>
<th>Abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harvest with no measures</td>
<td>Positive</td>
<td>3 (27%)</td>
<td>3 (60%)</td>
</tr>
<tr>
<td></td>
<td>NS.</td>
<td>6 (55%)</td>
<td>1 (20%)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>2 (18%)</td>
<td>1 (20%)</td>
</tr>
<tr>
<td></td>
<td>Sum</td>
<td>11 (100%)</td>
<td>5 (100%)</td>
</tr>
<tr>
<td>No-harvest</td>
<td>Positive</td>
<td>2 (50%)</td>
<td>2 (67%)</td>
</tr>
<tr>
<td></td>
<td>NS.</td>
<td>0</td>
<td>1 (33%)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>2 (50%)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Sum</td>
<td>4 (100%)</td>
<td>3 (100%)</td>
</tr>
</tbody>
</table>
3.2.4 Riparian buffer zone (RB)

Nine articles on RB zones met our criteria, resulting in 26 observations (Tables 3 and 7). The main result was that there were no negative observations on the effect of buffer, regardless of width, time since harvest or control. Compared to no-harvest as control, ten out of eleven species richness observations and five out of five abundance observations were not significant.

The studies from Canada mostly investigated buffer width of at least 30 meters using no-harvest as control. The Swedish and Finnish studies investigated narrower buffers, 10–40 m, with exception of Marker et al. (2022) who characterized small buffers as 1–40 m and wide buffers as 41–120 m after inspiration from Lind et al. (2019). Marker et al. (2022) found no significant effect of a wide buffer (> 40 m) on spiders’ and vascular plants’ species richness compared to no-harvest (two observations). There was a positive effect of a small buffer (< 40 m) on spider species richness but no significant effect on vascular plant species richness. Most of the observations on positive effects of buffer width were compared to harvest with no buffer (seven out of eight observations). This include both species richness and abundance and short and long-term studies (Dynesius and Hylander 2007; Lavallee and Richardson 2010; MacDonald et al. 2014; Johansson et al. 2018). Originating from the same article, only two observations were long-term, reporting a positive effect of narrow buffer width in comparison to harvest with no conservation measures taken (Johansson et al. 2018). These effects were observed for both lichen species richness and abundance. The buffer was 10 m and the effect was studied in short (1 yr) and long time (16 yr) since logging. The species groups used to examine the buffer width represent specialist (lichens, bryophytes, beetles, spiders, and benthic macroinvertebrates) and generalist species (vascular plants and mammals).

Table 7. Number of observations per effect (Positive, NS. (Not significant), and Negative) of riparian buffer zone (RB) as conservation measure, per control (Harvest with no measures, No-harvest) and biodiversity measure (Species richness and abundance). RB levels (width) include Narrow (1–10 m) and Wide (> 10 m).

<table>
<thead>
<tr>
<th>Control</th>
<th>Effect</th>
<th>Species richness</th>
<th>Abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Narrow</td>
<td>Wide</td>
</tr>
<tr>
<td>Harvest with no measures</td>
<td>Positive</td>
<td>3 (100%)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>NS.</td>
<td>0      2 (100%)</td>
<td>2 (33%)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Sum</td>
<td>3 (100%)</td>
<td>2 (100%)</td>
</tr>
<tr>
<td>No-harvest</td>
<td>Positive</td>
<td>1 (25%)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>NS.</td>
<td>3 (75%)</td>
<td>7 (100%)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Sum</td>
<td>4 (100%)</td>
<td>7 (100%)</td>
</tr>
</tbody>
</table>
Fig. 6. Observed effects (Positive, NS (Not significant), Negative) on biodiversity measures (Species richness and abundance) summed over all conservation measures, by control (No harvest: No measures (harvest with no measures)) and taxa. Only taxa with sufficient number of observations are included.
3.3 Species groups

Suppl. file S2 displays the number of observations per species group and provides an overview of the conservation measures, number of observations and associated references. Vascular plants and arthropods had the highest number of observations. The arthropod group consisted mostly of saproxylic beetles and spiders, but included also bumblebees and aquatic arthropods, and was the most studied taxa with its 96 observations. The fungi group consisted of polypore fungi and soil-inhabiting fungi, and the mammal group of voles and other rodents. Articles on vascular plants and bryophytes were seldom differentiated on the basis of habitat preferences or substrate-specific niches. Several of the studied species groups were chosen to represent species that utilizes old-forest structures such as old trees, dead wood or dense canopy cover (Hämäläinen et al. 2014; Ylisirniö et al. 2016; Lee et al. 2017; Suominen et al. 2019). We tested dependency between control type and outcome per species group summed over all conservation measures for taxa with a sufficient number of observations, i.e. vascular plant richness, arthropod richness and abundance, and bryophyte richness (Fig. 6). The chi-squared test of dependency between control type and outcome was significant for bryophyte species richness ($\chi^2 = 7.09, df = 2, p \approx 0.03$) and arthropod abundance ($\chi^2 = 37.83, df = 2, and p < 0.01$), but not for vascular plant richness ($\chi^2 = 1.39, df = 2, p \approx 0.5$) or arthropod richness ($\chi^2 = 3.209, df = 2, p \approx 0.2$). The patterns from the results across all taxa were maintained for bryophyte species richness and arthropod abundance, meaning that no-harvest as control gave more non-significant observations and fewer positive observations than harvest without measures as control.

Arthropods was the only species group examined for all conservation measures. DW had only been examined with two species groups, arthropods (beetles) and fungi. Vascular plants were often examined in addition to another species group (Sullivan et al. 2008, 2010; Bescond et al. 2011; MacDonald et al. 2014; Sullivan and Sullivan 2019; Simard et al. 2021; Marker et al. 2022).

4 Discussion

We scanned 3012 articles of which 188 were scrutinized to unveil the effects of conservation measures commonly applied upon timber harvest in boreal forest. However, 137 (73%) of the reviewed articles did not meet our criteria of experimental design, transparency of findings or clearly defined control. Thus, we ended up with 48 articles that provided 238 useable observations of biodiversity effects of the four conservation measures. The low number of studies that strictly have investigated effects of conservation measures upon timber harvest is striking, given that two-thirds of the boreal forests are available for timber production (Gauthier et al. 2015). Originally, we aimed at depicting threshold effects of the conservation measures, in the short and long term. However, it was not possible due to few useable studies and the great variation in study design. Numerous studies provide relevant insight into impacts of forest structures for ecosystem functioning yet fell outside the scope of this review. For instance, albeit the importance of dead wood for forest ecosystem functioning is well established (Lassauce et al. 2011), we found few studies that examined the effectiveness of dead wood retention in combination with timber harvest and most of such studies were therefore excluded.

Inadequately reported statistics was a excluding factor of many studies even though we did not conduct a meta-analysis. This phenomenon has been reported in similar reviews and meta-analyses (Fedrowitz et al. 2014; Ekholm et al. 2023). Given the various methods and presentation of statistics in the articles, including lacking indication of variation of results in numerous articles, we applied a qualitative summary of the findings (Pullin and Stewart 2006; Spake and Doncaster...
2017; McKenzie and Brennan 2023) to avoid reducing the number of articles further. There was also a substantial difference in the number of observations of each control type, measures complicating meta-analyses (Deeks et al. 2023). Systematic reviews in applied ecology aim to gather higher level of evidence to guide management decisions (Pullin and Stewart 2006). To ensure the quality of the review, it is crucial that the included studies describe study design sufficiently and have retrievable statistics (Stanhope and Weinstein 2023). For instance, in some riparian buffer zone studies it was not possible to separate effects of buffer width and logging based on the study design and results description (Darveau et al. 2001; Holmes et al. 2017).

In the review, no-harvest controls included a broad set of forest types, from forests managed for timber production that is mature, yet biologically young, to pristine forest. Separation of the no-harvest into sub-controls revealed differences. The geographical distribution of study sites between mature, managed control and mature, unmanaged control was clear, with the large share of mature, unmanaged forests controls in Canada while Nordic studies mainly including mature forest that had been managed as control. The wide geographical distribution of study sites, over continental and vegetation zones, with differences in geology, forest history and silvicultural management (Gauthier et al. 2015; Pohjanmies et al. 2017; Spake and Doncaster 2017) necessitate caution in drawing conclusion about ecological impacts across continents. Thus, we cannot conclude whether the observed difference between the no-harvest sub-controls is due to differences in management history or natural attributes between these two groups. As pointed to by Mori et al. (2017), substantial bridges should be built between applied ecology and forest management sciences in order to strengthen the knowledge base for sustainable forest management. We can only adhere to that idea as the evident lack of studies meeting our criteria may suggest that adjacent fields of research do not cooperate sufficiently. Numerous studies of impacts of forest management on biodiversity have been carried out without meeting the criteria of control, management, and conservation measure. Several studies could have been included in our final sample with minor adaptations in study design (Hågvar et al. 2004; Lindhe et al. 2004; Fauteux et al. 2012; Olden et al. 2019), indicating that in many cases improving the knowledge basis of conservation measures in forestry does not require much additional resources. We can only speculate why presuming adequate studies did not consider such effects, but it might be due to various factors like prevailing ideas about relevant research questions, a lack of research funding towards these questions and barriers to interdisciplinary research (Lélé and Norgaard 2005).

Species richness and abundance were the chosen biodiversity indicators due to their broad usage and clearness. Orians and Groom (2006) and Spake and Doncaster (2017) argue that species richness is a poor indicator to describe effects of disturbance factors, such as logging, on biological communities. Even with an equal number of species between control and treatment, there can be substantial differences in species composition. Total species count can be supplemented with abundance, but neither measure informs on species composition in different habitats. Community change could be a better indicator to determine whether the conservation measure is an adequate habitat to secure the “old” community of species throughout the early successional stages to the “new” mature forest (Hylander et al. 2005; Hyvärinen et al. 2005; Hylander and Weibull 2012). However, the small number of such studies prevented inclusion in this review.

Differences of effects between the two controls, no-harvest and harvest with no measures, were maintained across taxa, with more negative results in the no-harvest control group. This underlines that differences within taxonomic groups such as habitat preferences and ecological niches would be more important to address (Fedrowitz et al. 2014). Employing biodiversity measure as species richness or abundance does not ensure conservation of species with special conservation concerns (e.g. dead-wood dependent species). The small number of observations also hindered the inclusion of red-list species in our review, in line with the findings of Fedrowitz et al. (2014).
Red-list species in forest are often related to old-growth forest and structures. Regarding species groups, bacteria and other microorganisms were scarcely investigated while arthropods and vascular plants accounted for the largest species group in the review. The arthropods group consisted mostly of beetles, mainly saproxylic beetles, a species group sensitive to forest management and containing several red-listed species (Nieto and Alexander 2010). Vascular plants were in many cases investigated alongside other species groups and would typically represent more generalist species or open-habitat species (Fedrowitz et al. 2014). Interestingly, we could not detect any effect of the conservation measures on vascular plants when compared to harvest with no measures, supporting the idea that community change would be a better measure to assess the effectiveness of conservation measures in forestry (Spake and Doncaster 2017).

As the primary aim was to investigate the effects of the current conservation measures, our priority was studies with harvest with no conservation measures as control (Christie et al. 2019). However, it soon became clear that the number of such studies is very small compared to studies that have no-harvest as control and we therefore decided to include the latter group. However, these studies gauge to which extent the biodiversity is maintained in a forest that undergoes harvest which include conservation measures compared to a situation with no management (Spake and Doncaster 2017). To unveil effects of current forestry management practices with conservation measures, it is valuable to examine differences in effects between current practices and management practices without any conservation measure (Spake and Doncaster 2017).

Sixty-eight percent of the observations of any conservation measure were not significant compared to no-harvest. This indicates that species richness and abundance may be maintained upon harvest if conservation measures are taken which may oppose to hypotheses of biodiversity harm of timber harvest. In contrary, 45% of the observations found positive effects on species richness and abundance when compared to harvest with no conservation measures while 43% of the observations found non-significant effects. Thus, the evidence that the commonly applied conservation measures support biodiversity and that species richness and abundance are improved when these are taken is deficient. Further, with the small number and lack of clear patterns in the significant results, we cannot recommend any level of conservation measure. However, undertaking some biodiversity conservation measure may be positive, as put forward by Gustafsson et al. (2020). Gustafsson et al. (2020) advise that more than 50% GTR, PR at least as large as 0.6 ha, a few tens of meters of RB zones and retention of high stumps, snags or logs on clear-cuts are all needed measures to sustain forest-depending species upon harvest which resonate with the findings of Fedrowitz et al. (2014). Through meta-analysis, they found a weak but positive effect of increased retention level on species richness when compared to clear-cuts but no significant effect on abundance, thus finding some support for retention tree level being important to sustain forest biodiversity.

Long-term effects and successional recovery after disturbance are still under-studied topics in forest ecology research. Only five articles examined long-term effects of the conservation measures. Long term-studies are needed to understand the overall effect of a conservation measure and differentiate between forest successional stages and real loss of biodiversity (Pretzsch et al. 2019). There is a lack of long-term studies across conservation measures, in particular in combination with harvest without conservation measures as control. Thus, based on this review, we have very little evidence to state that the conservation measures commonly undertaken in boreal timber-production forests have positive effects on the long-term biodiversity compared to no measures being taken. Also, we cannot conclude whether the long-term effects of harvesting with conservation measures impacts the biodiversity in any direction compared to no-harvest as control. With this conclusion, we can only encourage more long-term studies of effects of harvest and various conservation measures on biodiversity. Ideally, studies should aim for forest management relevance.
and be transparent by clearly displaying descriptions of controls, statistical results and ideally data to facilitate further use and analyses. Regarding our hypotheses, our review does partly support Hypothesis 1. We were not able to conclude on Hypothesis 2, while Hypothesis 3 was rejected.

5 Conclusions

We have found that the impacts on biodiversity of timber harvest with conservation measures are in most studies not significant, compared to unlogged forests. Regarding the impacts of conservation measures upon harvest on biodiversity compared to harvest with no measures, the reviewed studies are divided in two parts of about equal size, of which one part shows positive effects and the other part shows non-significant effects. Thus, it is hard to draw sharp conclusions, but some degree of measures does likely have positive effects on biodiversity. However, the scientific basis does not allow for pointing to thresholds levels. The unambiguous conclusion is, however, the need for ecology and forestry sciences to enhance collaboration in order to strengthen the basis for knowledge-based forest management.

Data availability

All data in the final selection of articles are given in the supplementary. A list with papers excluded from the final selection is available upon reasonable request to the corresponding author.

Authors’ contributions

ACT gathered data, selected articles, carried out all numerical analyses and drafted the manuscript. GA gave input on methodology, interpreted data and results and revised the manuscript. HKS gave input on methodology and selection of articles, interpreted data and results and revised the manuscript. All researchers formulated the research questions.

Acknowledgments

We are grateful to Vegard Lien at Glommen Mjøsen Skog SA for developing Fig. 2 and to Nataliya Pavlyuk in assisting sorting and organizing original article searches.

Funding

Funding for this project was provided by the Norwegian Research Council (grant number #329485).

Supplementary files

S1.xlsx; Information about the included observations in the review, S2.pdf; Overview of conservation measures and species groups, available at https://doi.org/10.14214/sf.23057.
References


by small mammals in a boreal balsam fir forest. Forest Ecol Manag 143: 95–104. https://doi.org/10.1016/S0378-1127(00)00509-0.


Lavallee S, Richardson J (2010) Relative abundance and movement of the carabid beetle Scaphi


Ovaska K, Sopuck L, Robichaud D (2016) Short-term effects of variable-retention logging practices...


PEFC Council (2023) PEFC global statistics. https://pefc.org/resources/publications?filter_category%5B0%5D=10000127.


R Development Core Team (2022) RStudio 2022.12.0+353 “Elsbeth Geranium.”


Schlyter P, Stjernquist I, Bäckstrand K (2009) Not seeing the forest for the trees? The environmental


Sverdrup-Thygeson A, Bendiksen E, Birkemoe T, Larsson KH (2014) Do conservation measures...

Total of 120 references.