Forest Ecology and Management 491 (2021) 119190

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

An evidence-based review of the effectiveness of riparian buffers to maintain stream temperature and stream-associated amphibian populations in the Pacific Northwest of Canada and the United States

Douglas J. Martin [a,*], Andrew J. Kroll [b], Jenny L. Knoth [c]

[a] Martin Environmental, Seattle, WA 98103, USA
[b] Weyerhaeuser, Springfield, OR 97478, USA
[c] PNK Consulting, Olympia, WA 98516, USA

ARTICLE INFO

Keywords:
Amphibians
Forest management
Non-fish-bearing streams
Pacific Northwest
Riparian buffers
Water temperature

ABSTRACT

We employed a systematic evidence review to evaluate empirical scientific evidence for the effectiveness of buffering headwater (typically non-fish-bearing) streams to maintain stream temperature and stream-associated amphibian populations in the Pacific Northwest of Canada and the United States. To address our synthesis objective, we identified thirteen temperature, seven amphibian, and two temperature/amphibian primary research studies that met objective inclusion criteria. We evaluated external validity for how study treatments inform or were linked to causal factors influencing temperature response (a) and how the sampled population represented or provided inference to an intended target population or landscape (b). The evidence indicated substantial variability in the temperature response to streamside buffers. The effect size for the mean 7-day maximum temperature metric showed a positive association when comparing no-buffers (clearcut) to treatments with wide buffers (≥30 m). However, this effect varied substantially and overlap existed in effect sizes among no-cut buffers, no-cut plus variable retention buffers, and no-cut patch buffers all ≤20 m wide. Large variability in effect size among treatments obscured any potential trend between effect size for the seasonal (summer) mean daily maximum temperature metric and buffer width. Shade was correlated with temperature response within several studies, but direct comparisons of treatment effectiveness among studies as a function of shade was confounded by different measurement methods. The evidence also indicated that variation in temperature response among studies may be associated with multiple factors (geology, hydrology, topography, latitude, and stream azimuth) that influence thermal sensitivity of streams to shade loss. For amphibians, we found mixed evidence for relationships between population responses and buffers maintained along streams after forest harvest. Specifically, we did not find evidence to support the contention that positive population responses are associated consistently with larger buffers. Also, considerable uncertainty exists about which environmental covariates reliably explain variation in amphibian population responses. Collectively, our results indicate that evidence is weak to address questions most relevant to policy discussions concerning effectiveness of alternative riparian management schemes. Future studies should test effectiveness of alternative treatments with either experimental or purposefully structured observational studies to develop tools and derive guidelines for how to achieve management goals based on site and landscape characteristics.

1. Introduction

Riparian ecosystems integrate aquatic and terrestrial communities and often support unique assemblages of flora and fauna (Naiman and Bilby, 1998). These areas can be more productive and compositionally and structurally diverse than adjacent uplands (Bull, 1978; Thomas et al., 1979; National Research Council, 2002). Globally, riparian ecosystems support and enhance human welfare and well-being but face numerous threats due to increasing levels of utilization and dramatic variability in climactic regimes (Pettit and Naiman, 2007; Kovach et al., 2019). As a result, many management programs emphasize effective practices to conserve riparian functions while allowing for sustainable

* Corresponding author.
E-mail address: doug@martinenv.com (D.J. Martin).

https://doi.org/10.1016/j.foreco.2021.119190
Received 12 February 2021; Received in revised form 15 March 2021; Accepted 18 March 2021
Available online 8 April 2021

human use (Ryan and Calhoun, 2010; Tiwari et al., 2016).

Due to the ecological importance of riparian ecosystems, buffers of standing trees or intact native vegetation are often left between uplands and aquatic environments to reduce potential negative effects of timber harvest, agriculture, or other land uses (Stauffer and Best, 1980; Knopf et al., 1988). Buffers may support natural processes and functions of aquatic systems such as shading, sediment capture, and inputs of large wood and leaf litter (Chamberlin et al., 1991); retain aquatic species and communities (Osmundson et al., 2002; Kiffney et al., 2004; Sabo et al., 2005); protect riparian flora and fauna (Naiman et al., 2005; Richardson et al., 2005; Pearson et al., 2015), and facilitate exchange of nutrients between aquatic and terrestrial systems (Nakano and Murakami, 2001; Chaloner et al., 2002; Helfield and Naiman, 2002; Bilby et al., 2003). Also, buffers may serve as dispersal corridors and counteract problems associated with landscape fragmentation (Wilcox and Murphy, 1985; Saunders et al., 1991; Grant et al., 2010). Although general effectiveness of buffers in reducing and ameliorating timber harvest effects is accepted, recent applications and research have focused on how variation in management prescriptions, as opposed to uniformity, can yield a broader range of positive environmental outcomes (Kreutzweiser et al., 2012; Richardson et al., 2012).

Headwater streams are sensitive to environmental disturbances given strong coupling with the surrounding environment (Richardson et al., 2005; Richardson and Danehy, 2007) and are ecologically important given their large spatial extent in the Pacific Northwest (e.g., comprise 90% of stream length within the Oregon coastal ecoregion; Benda and Dunne, 1997). Therefore, federal and state forest management agencies in this region have developed riparian strategies to provide more substantial protections of aquatic resources, with an emphasis on maintaining temperature regimes suitable for aquatic biota in non-fish bearing headwater (e.g., amphibians) and downstream fish waters (Washington Department of Natural Resources, 2005; Ryan and Calhoun, 2010; Reeves et al., 2018). The best-management practices (BMPs) for headwaters have been in place for about two decades (since 2000) and their relative effectiveness is of particular concern to forest managers and policy makers who are charged with assessing whether the BMPs achieve stream management objectives and assuring the public that the environment is being protected (Jackson et al. 2021).

Here, we examine and present empirical evidence concerning the effectiveness of buffering headwater (typically non-fish-bearing) streams to maintain stream temperature and stream-associated amphibian populations in the Pacific Northwest of Canada and the United States (PNW). The study region is characterized by rugged mountains with steep slopes and highly dissected terrain. The maritime climate is typically cool (summer days rarely exceed 26 °C) with wet winters and mild, dry summers. Conifer forests are predominantly Douglas fir, western hemlock, western red cedar, and Sitka spruce (Franklin and Dyrness, 1973).

We used systematic review guidelines from the Collaboration for Environmental Evidence (2018) to search for relevant data, critically evaluate study quality, and quantitatively/qualitatively synthesize outcomes of different studies to address three questions. First, what is the relative effectiveness of forest practices (FP) buffer rules for maintaining water temperature (effects/response) and stream-associated amphibian populations compared to alternative buffering schemes for headwater streams (synthesis question 1)? Our intent was to examine relative effectiveness of interventions to provide specific evidence for informing policy decisions. Second, because observed outcomes of buffer effectiveness on the subject response or population can be influenced or modified by a range of abiotic and biotic factors other than treatments, we asked: what factors (e.g., shade, hydrology, geology, topography, and ecological interactions such as competition) account for differences in buffer effectiveness (synthesis question 2)? Finally, riparian buffer rules are implemented at a regional scale, and we asked: how applicable are study findings to headwater streams across the Pacific Northwest of Canada and the United States (synthesis question 3)?

2. Methods

To evaluate the strength of evidence for associations between riparian buffers and post-harvest stream temperature and amphibian population responses, we employed a systematic evidence review. Briefly, evidence reviews require formulation of questions that are answerable, relevant to decision-makers, and neutral (unbiased) to stakeholder groups. Also, questions specify the subject population, treatment, comparator, and outcomes of interest (Collaboration for Environmental Evidence, 2018; www.environmentalevidence.org/information-for-authors). Given the subject and regional specificity of our synthesis, we included grey literature to capture all relevant documents, increase sample population, and reduce potential bias in findings (Collaboration for Environmental Evidence, 2018).

2.1. Search strategy

We implemented an unbiased and transparent systematic search of the literature using a search strategy to locate relevant publications that we filtered and retained for review and data extraction based on specific inclusion criteria. We searched a wide range of databases and document types (described in *Inclusion Criteria* below) to assemble a comprehensive and unbiased sample of the relevant literature (Haddaway et al., 2020). Queried databases included Web of Science, EBSCO Environment Complete, TreeSearch, Google Scholar, and ProQuest for thesis and dissertations. We developed five unique query strings using search terms relating to water temperature, amphibian, riparian forest, and harvest (Appendix A1). Based on evidence syntheses that examined riparian buffers (Bowler et al., 2012; Czarnomski et al., 2013), we tested and incrementally modified the query strings. In addition, we screened bibliographies of included studies and cited reviews for relevant references (e.g., public agency or industry reports). The search was considered complete when queries did not return unique references. Full citation information and abstracts are available in an Endnote file upon request to the authors.

2.2. Inclusion criteria

The criteria for selection of documents from the database query returns were: a study directly informs the primary review question (response includes measure of amphibians or stream temperature); only primary research (studies with original data, not reviews, modeling or *meta*-analysis); published studies or gray literature (government reports, industry studies, graduate student theses, and manuscripts in review); studies reported from 1990 to the present that examined contemporary streamside buffers; and studies performed in the target geographic area, including Pacific Northwest coastal areas west of the Cascades with mixed conifer forests in British Columbia, coastal Southeast Alaska, northern California, Oregon, and Washington.

We applied initial inclusion criteria by viewing the titles of the articles. When the titles provided insufficient information, the abstract was reviewed to determine inclusion. Finally, studies from the queried database were read to determine relevance based on whether studies had controls or un-treated references (i.e., before-after-control-impact = BACI, after-control-impact = ACI, impact gradient analysis of varying disturbance levels including non-impacted/control = IG); study results clearly reported or can be reliably determined from data in figures or tables; studies in 1st to 3rd order headwater streams (Strahler, 1957) with basin sizes similar to where FP buffer rules are implemented (typically basin areas <100 ha, channel bank full widths <6 m; Washington Department of Ecology, 2015); studies that provide accurate information for timing of pre- and post-treatment monitoring relative to timing of riparian treatments; and studies of intact riparian stands (e.g., no recent disturbance from fire or debris flow). We retained for the synthesis only those studies and specific treatment data that met all inclusion criteria including temperature studies in small fish-bearing streams. For example, only the data for small streams were extracted from a study that included findings from larger (non-qualifying) streams.

2.3. Data extraction and analysis

We extracted all information and data relevant to addressing the synthesis questions and entered these into an Excel database. The database included information about study source, regional location, design, outcome, modifiers, site descriptors, quality, and external validity (Appendix B). In some cases, additional information or key points were noted as an attribute. The full dataset is available as an additional supplementary file (Supplemental Data A and B).

The types of information and data available for assessing study outcome differed for temperature, shade, and amphibian studies. For temperature, we coded the reported outcome mean effect sizes and variability statistics by treatment group (i.e., sites with common treatment) where data were available. In a few cases where summary statistics were not reported, we coded the temperature outcomes for each study site and the average was computed for all sites within a treatment group. In one case where effect size was not reported we calculated the treatment change in temperature as the difference of differences between reference and treatment sites for the pre- and post-harvest periods. Only temperature data for the summer season (specific time period varied by study) and for metrics expressing temperature maxima were extracted from each study. Therefore, depending on how the data were presented, we coded one of the following metrics: the mean of 7-day maximum (i.e., 7-day moving average of daily maximum temperatures), the mean daily maximum, or the daily maximum.

For shade, we coded the shade measurement tool (densiometer or hemispherical photography) and a shade metric (canopy closure or effective shade) that were appropriate for the reported measurement/data analysis. The latter was necessary because the studies expressed shade using a variety of shade terminology (e.g., canopy closure, canopy topo density, canopy cover, effective shade). For consistency, we used the shade metric definitions and nomenclature described by Jennings et al. (1999). Therefore, canopy closure was assigned to all cases where the shade metric was based on estimating the proportion of the sky hemisphere obscured by vegetation and topography when viewed from a single point (e.g., canopy density, canopy topo density). The term canopy cover was coded as canopy closure because the measurement method (i.e., view from single point) did not fit the Jennings et al. (1999) canopy cover definition which refers to the proportion of the forest floor (or stream surface) covered by the vertical projection of tree crowns. Effective shade, defined as the fraction of total possible potential solar radiation that is blocked by riparian vegetation and topographic features (Allen and Dent, 2001; Teti and Pike, 2005), was assigned to cases where solar radiation was estimated from hemispherical photography. We presented results as the proportion of both direct and diffuse energy under a plant canopy relative to the available direct and diffuse energy above canopy. Effective shade is a function of riparian vegetation characteristics (height, density, buffer width) and channel orientation to the sun (Boyd and Kasper, 2003; DeWalle, 2010).

We used the term "shade" generically in reporting the study findings in the Results section and specify shade measurement metrics in associated tables and figures. In the Discussion section, we evaluated and described how the different shade metrics confounded comparisons of treatment effectiveness among studies. Also, to facilitate equivalent comparisons of shade among studies, we converted canopy closure results to estimated effective shade using a regression equation from Allen and Dent (2001) where:

$$\text{Estimated effective shade} = 0.6514\ (\text{canopy closure}) + 17.121\ (R^2 = 0.72).$$

The amphibian studies focused on two types of responses: presence/absence (probability of occupancy) and abundance (or density, if sampled and estimated as a function of a fixed area). Studies estimated outcome effect sizes as the differences between observed and predicted responses at reference and treatment sites for pre-harvest and post-harvest periods (experimental studies) or differences across treated and untreated sites (observational studies). We summarized treatment contrasts (if presented in publications) and/or calculated differences based on summary results presented in publications (Details in Supplemental Data A and B). We note that, in many cases, this information was not provided in the publications in which case the publications were not included in the review. For a detailed overview of the pervasiveness of this problem in the literature on stream-associated amphibian responses to forest management, please see Kroll (2009).

Factors (covariates) other than the intervention treatment may be associated with study outcomes. We identified potential effect modifiers and confounding variables examined or reported as plausible explanations of outcomes for each study (Appendix B). Knowledge of relationships and correlations among outcomes and covariates may help to explain differences in buffer effectiveness and confidence for addressing the synthesis questions. Similarly, we extracted study site information (e.g., stream width, gradient, riparian stand age), if available, to provide context and understanding of differences or similarities among studies.

2.4. Assessment of study quality and external validity

We based study quality on the ranking of eight attributes concerning study design, site selection, replication, pre-post treatment monitoring, statistical rigor, and peer review (Appendix B). We ranked each attribute (weighted) from low to high quality with values 1–3, respectively (e.g., Nichols et al., 2017). The sum of all quality values for each study is a measure of overall quality and strength of evidence for answering the primary synthesis question. External validity is a qualitative assessment of how well each study addressed the secondary questions. External validity for synthesis question 2, was based on how well each investigation was designed to address or examine causal factors influencing study population outcome or address plausible reasons for heterogeneity in study results. Ranking for synthesis question 3, was based on how well the study plan and sampled population represented or provided inference to an intended target population or landscape.

3. Results

3.1. Literature search

The initial query of the Web of Science database cast a wide net returning thousands of citations. Therefore, additional filter terms were applied prior to a methodical screening of titles and abstracts resulting in 121 unique citations (Appendix A2). Ten additional citations were revealed through searches of the EBSCO Environment Complete and the United States Forest Service TreeSearch databases. Three relevant dissertations were identified through ProQuest and a bibliography search returned one master's thesis that was not identified through the ProQuest query. A broad search of Google Scholar returned no new documents. Further reading of the queried documents ($n = 134$) resulted in the retention of 12 studies that met the full inclusion criteria (nine temperature and three amphibian).

Bibliography searches provided eight additional relevant studies that were not identified by the database searches and fit the inclusion criteria. Many of the additional documents were unpublished agency reports including three studies performed by the Washington Department of Natural Resources (MacCracken et al., 2018; McIntyre et al., 2018; Ehinger et al., in review). Combining all relevant documents resulted in thirteen temperature, seven amphibian, and two temperature/amphibian studies that informed the primary synthesis question. McIntyre et al. (2021) includes results from short (2-year) and extended (years 7–8) post-treatment monitoring periods.

3.2. Demographics of studies

3.2.1. Temperature

The temperature and buffer treatment effectiveness studies occurred in four state/province regions and examined 28 treatments at 117 study sites (Fig. 1 and Appendix C). Most studies occurred in Washington State and were concentrated geographically in the southwestern coastal ecoregion, with a few studies in the north and south Cascade ecoregions. Oregon had the second most studies, and these were similarly concentrated in the coastal ecoregion, with one study in the south Cascade ecoregion. All three studies from British Columbia were in the University of British Columbia Malcolm Knapp Research Forest, southwestern Cascades, BC. Both California studies occurred in the northern Siskiyou range of the Klamath ecoregion.

Buffer prescriptions were highly variable and consisted of no-cut buffers, thinned riparian stands, no-buffer (clearcut), and various combinations of all options. The no-cut buffer treatments varied from 10 to 33 m wide and included continuous and discontinuous (patch) buffers of variable lengths within a clearcut basin. Studies of thinning included thinned riparian stands directly adjacent to the stream and thinned stands outside (upslope) of a stream-adjacent no-cut buffer. A clearcut prescription was the only treatment in common among multiple (7 of 15) studies.

Study treatment replication varied from less than three replicates (50% of treatments), four to seven replicates (43% of treatments), and 15 and 18 replicates (7% of treatments). The duration of studies pre- and post-treatment ranged from zero to six years and one to nine years, respectively. Only one study had no pre-treatment monitoring.

3.2.2. Amphibians

The amphibian and buffer treatment effectiveness studies occurred in three geographic regions and examined nine treatments at 150 study sites (Fig. 1, Appendix C). The majority of the studies occurred in western Washington, with two studies each in Oregon and British Columbia.

Buffer treatments were variable and included clearcut, partial, and full buffers as well as thinning to different shade levels. Most of the studies examined no-cut buffers of varying width (10–64 m) and the clearcut prescription was the only treatment examined by multiple (5 of 7) studies.

Replication of treatments ranged from 3 to18 sites and study duration, pre- and post-treatment, ranged from zero to three years and 3–10 years, respectively

3.3. Effect sizes

3.3.1. Temperature

Among the 15 studies, five reported temperature effect size maxima with the mean daily maximum, eight studies used the mean 7-day maximum, and two studies used the daily maximum (Table 1). Effect size statistics, including average, range, confidence intervals (95% CI), were not consistently available among all studies making comparisons difficult. For example, the average effect size results from Jackson et al. (2001) and Guenther (2007) were absent and outcome variability data were not reported for several studies. Consequently, our evaluation of treatment effectiveness is focused on two temperature metrics (mean daily maximum, mean 7-day maximum) for which we have the most effect size data. Results for studies using daily maximum temperature provide supplemental support for some findings.

The average of mean daily maximum effect sizes for 10 different riparian treatments including clearcuts ranged from 0 °C to 1.7 °C (Table 1). Clearcut treatments generally resulted in the largest effect size (i.e., >1.0 °C) except for one study where the effect size decreased after treatment and averaged 0.2 °C (Kibler et al., 2013). Kibler et al. (2013) attributed the small temperature response, in part, to shading from logging slash; Jackson et al. (2001; Table 1) reported a similar finding with regards to shading provided by slash. No-cut buffers ranging from 6 to 15 m wide and some with variable retention outside the no-cut core had average effect sizes ranging from 0.7 °C to 1.1 °C. A similar effect size (0.7 °C) was observed in basins with no-cut patch buffers (Janisch et al., 2012). The smallest effect sizes (≤0.01 °C) were associated with the widest buffer treatments that included either wide no-cuts (i.e., 30 m, Gomi et al., 2006) or no-cuts with wide (52 m) variable retention zones (Groom et al., 2011). These findings suggested a trend of decreasing effect sizes in association with increasing buffer width. However, substantial variability in effect size existed among the treatments (Fig. 2).

The average of mean 7-day maximum effect sizes for 14 different riparian treatments including clearcuts ranged from −1.0 °C to 3.4 °C (Table 1). Treatment average effect sizes showed an apparent trend (Fig. 2) with clearcuts having the largest response (≥3.4 °C) and treatments with buffers ≥30 m (i.e., no-cut buffers with or without variable retention) had the smallest response (<0°C; Table 1). A suite of five treatments that included no-cut buffers, no-cut plus variable retention buffers, or no-cut patch buffers all ≤20 m wide had average effect sizes ranging from 0.6 °C to 1.4 °C. Despite differences among these five treatments, the variability in effect sizes (i.e., overlap of 95% CI for 4 of 5 treatments) indicated temperature responses were similar. Further, the effectiveness of different patch buffer lengths (i.e., ranged from 50+% to 100% of study unit) was unclear given the overlap of 95% CIs for effect sizes among treatments (McIntyre et al., 2021; Ehinger et al., in review). In contrast, effect sizes varied in association to the intensity of riparian thinning of the stream-adjacent stand. For example, low intensity thinning to the steam bank that retained 70–77% riparian closure resulted in average effect sizes of 0.5 °C and 0.2 °C, respectively (Farber and Whitaker, 2010a; MacCracken et al., 2018). Also, MacCracken et al. (2018) showed that progressively reducing canopy closure to 61% and 40%, corresponded to incremental increases in average effect size of 2.0 °C and 2.5 °C, respectively.

3.3.2. Amphibians

The seven studies evaluated several responses including abundance, catch per unit effort, density (abundance within a fixed area), and probability of occurrence. As a result, drawing general conclusions across the studies is challenging. We found no general pattern in the direction of responses across taxa. Importantly, none of the studies reported extirpation of amphibian species due to harvest with buffer treatments. However, McIntyre et al. (2021) reported declines of greater than 60% in density of tailed frogs eight years post-treatment in three buffer treatments compared to a control treatment (Table 2).

Across the seven studies, 95% CI were reported for 47 responses across taxa. In 30 cases, the 95% CI included 0 or 1 (depending on the response) indicating that no evidence of an effect was found. The 95% CI provided evidence for a negative or positive effect for eight and nine responses, respectively (i.e., the interval did not include 0 or 1, depending on the response). Generally, authors did not distinguish between statistical and biological effect sizes (but see MacCracken et al., 2018). That is, authors did not discuss whether the magnitude of change, whether positive or negative, was likely to be meaningful with regards to population performance over space and time (Nakagawa and Cuthill, 2007; Wasserstein et al., 2019).

3.4. Effect size modifiers

3.4.1. Temperature

All of the studies examined either correlations or identified plausible associations between temperature response and potential modifier variables (Appendix D1). Among 12 variables, significant correlations with effect size were observed for eight variables and plausible associations were identified for seven variables. Shade was the most frequent variable correlated with temperature response and 5 of 6 studies found shade inversely correlated with effect size. Slash cover over the stream

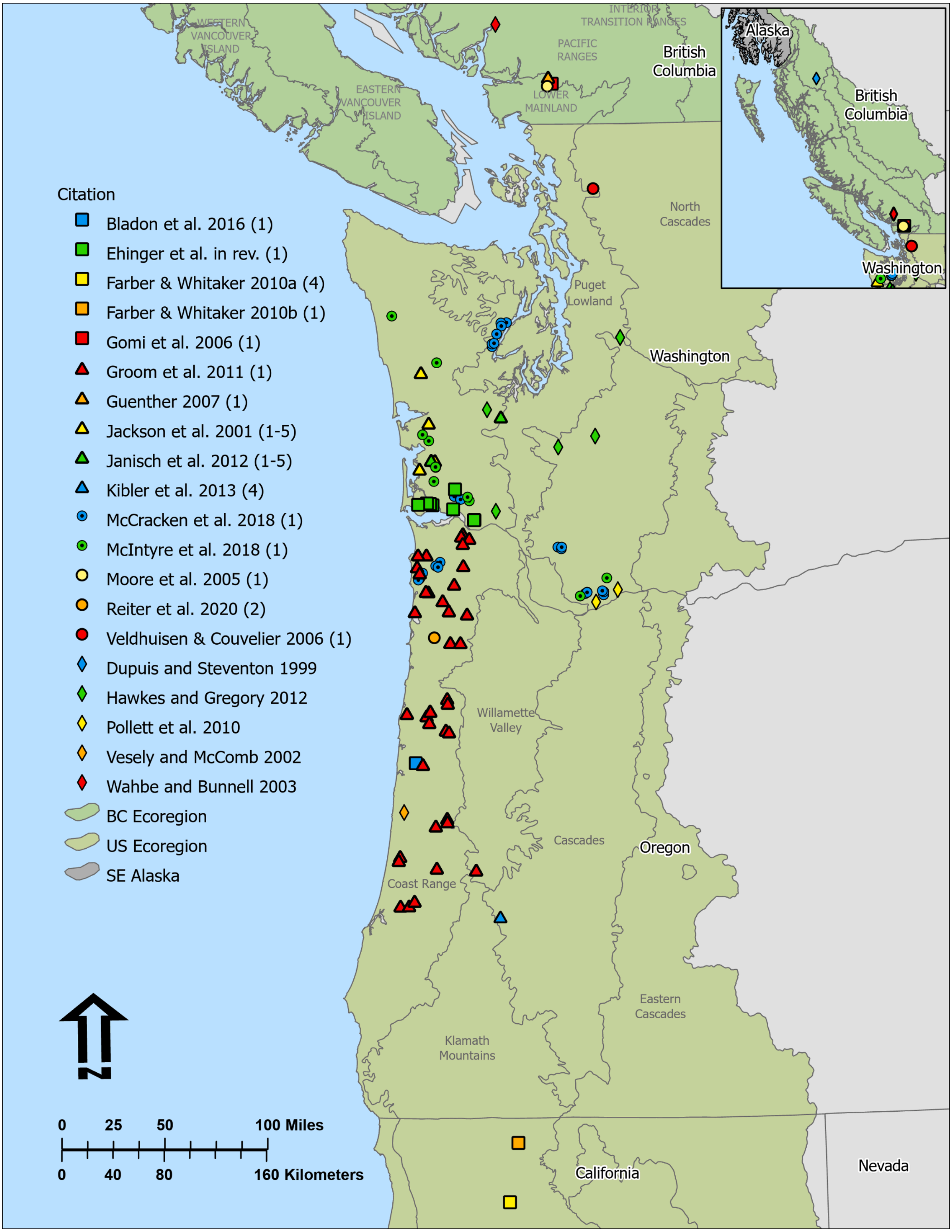

Fig. 1. Map showing locations of riparian treatment effectiveness studies of temperature and amphibian response, Pacific Northwest of Canada and United States.

Table 1
Summary of temperature effect size statistics during summer and post-harvest shade levels for each riparian treatment, Pacific Northwest of Canada and United States. Effect size is reported as the maximum value observed during study period.

Temp. metric	Riparian treatment[a]	Citation	No. sites	Avg. effect size (°C)	Outcome variability			Shade[b] (%)
					metric	min. (°C)	max (°C)	
mean daily max.	CC (Gomi)	Gomi et al. (2006)	4	1.7	range	−1.8	7.3	0[c]
	CC (Janisch)	Janisch et al. (2012)	5	1.5	range	0.8	1.5	53 cc
	CC (Moore)	Moore et al. (2005a)	1	1.3	range	−1.0	5.0	0[c]
	No-cut 10–15	Janisch et al. (2012)	6	1.1	range	0.4	1.1	86 cc
	No-cut 10	Gomi et al. (2006)	1	1.0	range	−0.6	4.1	–
	No-cut 15+, patch 50–110 m lng.	Janisch et al. (2012)	5	0.7	range	0.2	0.7	76 cc
	No-cut 6, var. ret. 15–21	Groom et al. (2011)	18	0.7	range	−0.9	2.5	78 es
	Clearcut slash cover	Kibler et al. (2013)	4	0.2[c]	range	−1.6	1.1	66 cc
	No-cut 30	Gomi et al. (2006)	2	0.1	range	−1.6	1.8	–
	No-cut 8, thin 30, var. ret. 52	Groom et al. (2011)	15	0.0	range	−0.9	2.3	89 es
mean 7-day max.	Clearcut	McIntyre et al. (2018)	4	3.4	95% CI	2.5	4.4	9 cc
	Clearcut	Reiter et al. (2020)	2	3.6[c]	–	–	–	9 es
	Thin ret. 40% canopy 10–20	MacCracken et al. (2018)	6	2.5	95% CI	–	–	40 cc
	Thin ret. 61% canopy 10–20	MacCracken et al. (2018)	7	2.0	95% CI	–	–	61 cc
	No-cut 15+, patch 55%-73% lng.	McIntyre et al. (2018)	3	1.4	95% CI	0.4	2.3	67 cc
	No-cut 12–20	Veldhuisen and Couvelier (2006)	1	1.4[c]	–	–	–	77 cc
	No-cut 15+	McIntyre et al. (2018)	4	1.2	95% CI	0.4	2.0	85 cc
	No-cut 6, var. ret. 15	Bladon et al. (2016)	1	0.7	95% CI	0.3	1.1	89 cc
	No-cut 15+, patch 53%-100% lng.	Ehinger et al., in review	7	0.6	95% CI	0.3	1.0	72 cc
	Thin ret. 77% canopy 10–20	MacCracken et al. (2018)	6	0.5	95% CI	–	–	77 cc
	Ret. 70% canopy 8, ret. 50% canopy 30	Farber and Whitaker (2010a)	4	0.2[c]	range	−1.4	1.1	79[c] cc
	No-cut 23–33	Reiter et al. (2020)	3	−0.4	–	–	–	85 es
	No-cut 15, ret. 50% canopy 46	Farber and Whitaker (2010b)	1	−0.4[c]	range	−0.6	−0.1	94 cc
	No-cut 15, thin ret. 33–55% BA∞	Reiter et al. (2020)	1	−1.0	–	–	–	84 es
daily max.	Clearcut	Jackson et al. (2001)	2	–	range	3.9	16.8	–
	No-cut 15–21	Jackson et al. (2001)	3	1.4	–	–	–	–
	Clearcut slash cover	Jackson et al. (2001)	5	–	range	−1.8	1.2	–
	Thin, ret. 50% BA∞	Guenther (2007)	1	–	range	5.0	7.0	82 cc

[a] Abbreviated description of riparian treatment: CC = clearcut, ret = retention, var. ret. = variable retention, patch = discontinuous buffer, lng = length, BA = basal area, ∞ = no outer edge. Each treatment has text followed by number indicating outer extent distance (m) of treatment (e.g., no-cut 15 = no-cut continuous buffer out to 15-m). More complex treatments with multiple treatments for stream-adjacent and outer buffer zones has multiple commas (e.g., No-cut 8, thin 30, var. ret. 52 = no-cut out to 8-m, thin out to 30-m, variable retention out to 52-m). In some cases the study unit includes a discontinuous buffer ("patch") for a portion of unit length and an upstream clearcut (e.g., No-cut 15+, patch 55%-73% lng. = no-cut buffer out to 15 + m wide that ranged 55%-73% of study unit length).
[b] Shade expressed as canopy closure (cc) or effective shade (es).
[c] Value estimated from report text or data, see synthesis data file.

(not quantified) was the most frequent variable identified as a plausible external factor influencing temperature response by providing additional shade.

The association between temperature response effect size and shade was highly variable among the six studies that reported correlation values (Appendix D2). Two studies each found weak ($r < 0.4$), moderate ($r = 0.4–0.8$), and strong ($r > 0.8$) correlations, respectively. The inverse relationship between temperature and shade was consistently observed by all studies except (McIntyre et al., 2021). McIntyre et al. (2021) found the correlation between shade and effect size for all treatments was negative during the initial study period (2 years), but not during extended monitoring (7–8 years).

Plots of temperature effect size versus shade among all studies showed an inverse relationship (Fig. 3). However, the evidence for a temperature-shade trend depended on the temperature metric and was best illustrated by the mean 7-day maximum. The latter metric shows average effect sizes may range from 2 °C to 3.6 °C when shade is lower than about 60%, from 0.5 °C to 2 °C for shade levels of 60% to 80%, and from −1.0 °C to 0.5 °C with higher shade levels. Given the variability in the data, effect sizes for the mean 7-day maximum ranged up to 4.4°, 2.3 °C, and 2.0 °C at low, intermediate, and high shade levels, respectively.

The longitudinal variability of temperature within study stream reaches was identified as another factor that may confound measures of treatment effect size. We noted that three of the 15 studies measured longitudinal variability. Further, we found other studies that examined longitudinal temperature processes, but these studies were not included in our synthesis because they did not address the primary synthesis question.

3.4.2. *Amphibians*

Three of the seven studies were designed experiments with pre- and post-treatment data collection (Hawkes and Gregory, 2012; MacCracken et al., 2018). The other four studies used a "space for time" design in which "treatments" of various ages were chosen for sampling. In the former case, the modifiers are not used under the assumption that the treatments are the major sources of variation in the responses. Similarly, in a "space for time" design, the "treatments" are assumed to be the main sources of variation although other covariates could be the source of variation in responses. For example, Wahbe and Bunnell (2003) examined variation in tailed frog density as a function of clearcut, second growth, and old growth stand characteristics around streams. In addition, the associations between density and elevation, percent of pools, sand substrates, and riffles in the stream reach, wetted width, and stream gradient were examined (although p-values, and not effect sizes, were presented for these associations). Vesely and McComb (2002) tested explicitly for an association between buffer width and torrent salamander occurrence and did not find evidence for a relationship (the 95% CI included 0 for the slope term in the logistic regression model fit). Pollett et al. (2010) presented associated evidence (proportions only) that Cascade torrent salamanders were less likely to occur in stream reaches in which temperatures exceeded 14 °C for more than 35

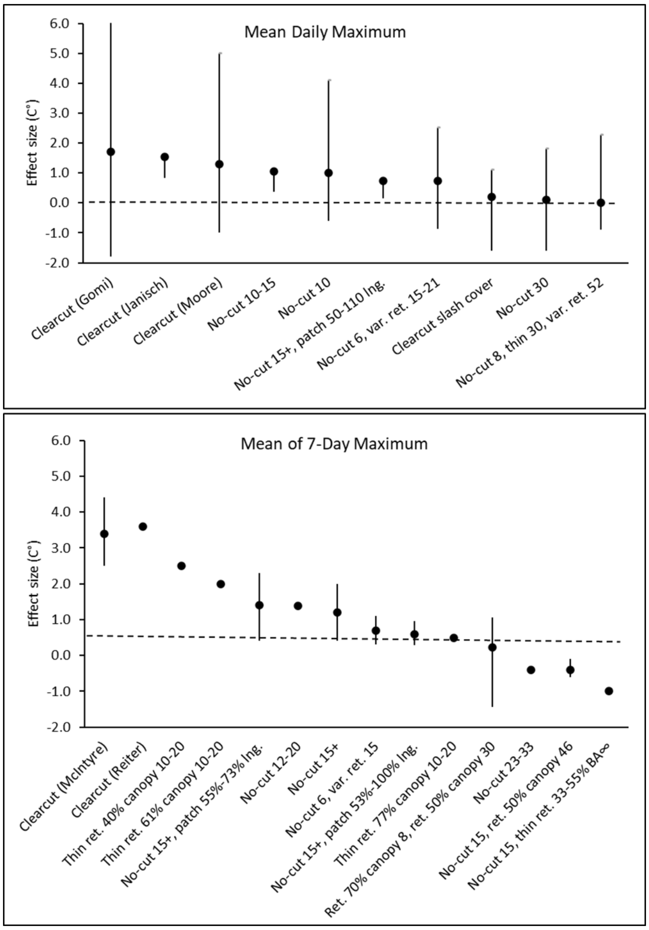

Fig. 2. Maximum post-treatment temperature effect size for the mean daily maximum (A) and mean 7-day maximum (B) during summer in relation to buffer treatments, Pacific Northwest of Canada and United States. Vertical lines express outcome variability (range or 95% CI) where data are available (Table 1). Treatments ordered by effect size. See Table 1 footnote for description of riparian treatment abbreviations.

consecutive hours. Finally, Dupuis and Steventon (1999) presented evidence that tailed frog larval density was associated negatively with the amount of fine sediments and fine organic matter in stream reaches.

3.5. Study quality

3.5.1. Temperature

Study quality scores ranged from 11 to 21 relative to a maximum possible score of 24 (Table 3). Groom et al. (2011) ranked highest by having the maximum score for 6 of 8 quality attributes and is the only study with a maximum attribute score (3) for having the largest number of treatment replicates. Nine studies ranked second overall (total scores 17–19) with a mix of medium and high scores for most attributes. Five studies ranked low (total score 11–16) due to multiple medium and low attribute scores.

The relative influence of study quality attributes on overall score varied among the eight attributes. For example, the site selection and experimental design attributes had no relative weight because scores were equal among all but one study. In contrast, several attributes (e.g., numbers of treatment replicates, pre-treatment years, post-treatment years) had stronger influence on study rankings. Scores for statistical robustness and peer review separated the first and second highest ranked studies from the third and lower ranked studies. No study was based on a random sample; thus, all studies had a low score for site selection bias.

3.5.2. Amphibians

Study quality scores ranged from 14 to 21 relative to a maximum possible score of 24 (Table 3). MacCracken et al. (2018) ranked highest by having the maximum score for 5 of 8 quality attributes. Generally, studies either did or did not include manipulative experimental designs, resulting in scores of one or three for the Experimental Design attribute. Several of the studies lacked adequate temporal replication but we note that none of the studies presented any information to indicate whether the sample size was adequate to estimate quantities of interest with sufficient precision (e.g., a pre-sampling power analysis). Finally, four of the studies received scores of one for Statistical Robustness, indicating that the statistical analysis did not present sufficient information about estimating spatial and temporal variability in quantities of interest.

4. Discussion

4.1. Question One

4.1.1. Temperature

The usability of study findings and strength of evidence for informing the primary question varied in relation to study quality, shade, and temperature metrics. In some cases, a lack of spatial and temporal replication reduced our confidence in study findings, especially for short-duration studies (e.g., 1 year post-treatment) with no spatial replication (e.g., Veldhuisen and Couvelier, 2006). Also, direct comparisons of treatment effectiveness as a function of shade were difficult because of different shade metrics. For example, canopy closure and effective shade are correlated, but are not equivalent (Kelley and Krueger, 2005; Fialaa et al., 2006; Cole and Newton, 2015). Also, measures of canopy closure tend to over-predict shade especially at higher levels (e.g., >70%, Allen and Dent, 2001). Studies reporting temperature findings with the daily maximum (e.g., Jackson et al., 2001; Guenther, 2007) have limited usefulness for this analysis because the results were based on a single value and were not comparable to others in this assessment. Similarly, studies reporting the mean daily maximum metric used different time-averaging periods for computing effect sizes. Five studies computed the mean daily maximum from data collected over the summer period and two studies computed the metric from the warmest portion (July–August) of the summer. Consequently, the temperature effect size for studies using the mean daily maximum was influenced by the time-averaging period, potentially confounding comparisons among studies using this metric. In contrast, the mean 7-day maximum, which is based on the warmest week (7 days) of summer, is a more consistent measure of treatment effect size. Given the use of different temperature metrics and study quality, we focused the synthesis on higher quality studies (Table 3) using the mean 7-day maximum, and we used studies from the mean daily maximum group to evaluate consistency, or lack of, among temperature response trends with similar treatments.

The distribution of effect sizes for studies using the mean 7-day maximum indicated that relative effectiveness of riparian treatments was associated with shade retention and poorly associated with various prescriptive components (e.g., width, no-cut, variable retention). The mean 7-day maximum effect size declined in relation to increasing shade retention for the subgroup of treatments, including clearcuts, targeting a specific level of canopy closure (shade-targeted; Fig. 4a). Consequently, the expected cause-effect relationship between effect size and shade was clear among the aggregate of studies with shade-targeted treatments (black symbols; Fig. 4b). On the other hand, the association between treatment effectiveness and shade was not readily apparent for other treatments. For example, the effect sizes were similar among the no-cut 15+ patch, no-cut 15+, and narrow combination (no-cut 6, var. ret. 15) treatments (Fig. 4a), but shade for these treatments differed substantially and ranged from 67% to 89% (plot points within polygon; Fig. 4b). Therefore, in this comparison, the increased shade (canopy closure) provided by continuous no-cut and combination buffers were apparently

Table 2
Summary of amphibian effect size statistics for each riparian treatment, Pacific Northwest of Canada and United States. Effect size is reported as the maximum value observed during study period.

Riparian treatment[a]	Citation with response	Number of sites	Taxa	Average effect size for response	Outcome variability			Shade (%)
					Range	Min.	Max.	
Small buffer 7–26 *m*	Hawkes and Gregory (2012) (abundance)	6	Tailed frog	−0.19 (pre-harvest); −0.04 (post 2 years); and −0.11 (post 10 years)	NA	NA	NA	NA
Large buffer 22–41 *m*		6		−0.18 (pre-harvest); −0.13 (post 2 years); and −0.11 (post 10 years)	NA	NA	NA	NA
Thin ret. 40% canopy 10–20	MacCracken et al. (2018) (abundance)	8	Tailed frog	−2	90% CI	−7	3	40
Thin ret. 61% canopy 10–20		9		0	90% CI	−8	2	61
Thin ret. 77% canopy 10–20		8		3	90% CI	−2	8	77
Thin ret. 40% canopy 10–20		8	Giant salamander	8	90% CI	3	13	40
Thin ret. 61% canopy 10–20		9		3	90% CI	−4	6	61
Thin ret. 77% canopy 10–20		8		<1	90% CI	−4	2	77
Thin ret. 40% canopy 10–20		8	Cascade torrent salamander	−3	90% CI	−10	0	40
Thin ret. 61% canopy 10–20		9		6	90% CI	0	10	61
Thin ret. 77% canopy 10–20		8		9	90% CI	4	14	77
Thin ret. 40% canopy 10–20		8	Columbia torrent salamander	1	90% CI	−1	4	40
Thin ret. 61% canopy 10–20		9		<1	90% CI	−4	2	61
Thin ret. 77% canopy 10–20		8		−3	90% CI	−7	3	77
Thin ret. 40% canopy 10–20		8	Olympic torrent salamander	−3	90% CI	−5	1	40
Thin ret. 61% canopy 10–20		9		−3	90% CI	−6	0	61
Thin ret. 77% canopy 10–20		8		5	90% CI	2	8	77
Clearcut	Dupuis and Steventon (1999) (density)	18	Tailed frog	0.571	95% CI	−1.16	2.297	NA
Clearcut with buffer 5–60		18		−1.47	95% CI	−3.2	0.256	NA
Clearcut	McIntyre et al. (2018) (density)	4	Larval tailed frog	1.44 (2 years post)[b]	95% CI	0.99	2.15	9
No-cut 15+, patch 100% of length		3		2.06 (2 years post)	95% CI	1.3	3.3	67
No-cut 15+, patch 55–73% of length		4		1.36 (2 years post)	95% CI	0.97	1.89	85
Clearcut		4		0.16 (7–8 years post)	95% CI	0.08	0.27	9
No-cut 15+, patch 100% of length		3		0.07 (7–8 years post)	95% CI	0.02	0.21	67
No-cut 15+, patch 55–73% of length		4		0.35 (7–8 years post)	95% CI	0.21	0.57	85
Clearcut		4	Metamorph tailed frog	10.61 (2 years post)	95% CI	4.81	25.48	9
No-cut 15+, patch 100% of length		3		0.51 (2 years post)	95% CI	0.21	1.17	67
No-cut 15+, patch 55–73% of length		4		0.43 (2 years post)	95% CI	0.27	0.69	85
Clearcut		4		0.4 (7–8 years post)	95% CI	0.12	1.38	9
No-cut 15+, patch 100% of length		3		0.03 (7–8 years post)	95% CI	0.01	0.14	67
No-cut 15+, patch 55–73% of length		4		0.29 (7–8 years post)	95% CI	0.18	0.48	85
Clearcut		4	Giant salamander	1.42 (2 years post)	95% CI	0.61	3.34	9
No-cut 15+, patch 100% of length		3		0.36 (2 years post)	95% CI	0.14	0.9	67
No-cut 15+, patch 55–73% of length		4		0.84 (2 years post)	95% CI	0.35	2	85
Clearcut		4		0.7 (7–8 years post)	95% CI	0.32	1.55	9
No-cut 15+, patch 100% of length		3		0.47 (7–8 years post)	95% CI	0.21	1.06	67

(continued on next page)

Table 2 (*continued*)

Riparian treatment[a]	Citation with response	Number of sites	Taxa	Average effect size for response	Outcome variability			Shade (%)
					Range	Min.	Max.	
No-cut 15+, patch 55–73% of length		4		0.64 (7–8 years post)	95% CI	0.28	1.48	85
Clearcut		4	Torrent salamander	2.98 (2 years post)	95% CI	1.18	7.51	9
No-cut 15+, patch 100% of length		3		0.81 (2 years post)	95% CI	0.28	2.33	67
No-cut 15+, patch 55–73% of length		4		1.12 (2 years post)	95% CI	0.44	2.88	85
Clearcut		4		0.84 (7–8 years post)	95% CI	0.37	1.92	9
No-cut 15+, patch 100% of length		3		0.36 (7–8 years post)	95% CI	0.14	0.9	67
No-cut 15+, patch 55–73% of length		4		1.2 (7–8 years post)	95% CI	0.59	2.43	85
Clearcut	Pollett et al. (2010) (density)	12	Tailed frog	−0.31	range	−1.5	0.8	NA
Buffer < 10 years		10		NA	range	NA	NA	NA
Buffer > 35 years		10		−0.5	range	−0.6	1.4	NA
Clearcut		12	Giant salamander	0	range	−0.5	0.6	NA
Buffer < 10 years		10		0.1	range	−0.5	0.7	NA
Buffer > 35 years		10		0.1	range	−0.6	0.8	NA
Clearcut		12	Cascade torrent salamander	−1	range	−2.7	0.8	NA
Buffer < 10 years		10		−1.3	range	−2.6	0.1	NA
Buffer > 35 years		10		−0.9	range	−2.7	−0.2	NA
Clearcut	Wahbe and Bunnell (2003) (density)	3	Tailed frog	−0.153	95% CI	−2.55	2.24	NA
Second growth		3		−0.463	95% CI	−2.86	1.93	NA
Clearcut with buffer 0–64 *m*	Vesely and McComb (2002) (occupancy)	17	Southern torrent salamander	5% more likely to occur for each 1 m of buffer width	95% CI	0.99	1.13	NA

[a] Abbreviated description of riparian treatment: patch = discontinuous buffer. Each treatment has text followed by number indicating outer extent distance (m) of treatment (e.g., no-cut 15 = no-cut continuous buffer out to 15-m). More complex treatments with multiple treatments for stream-adjacent and outer buffer zones has multiple commas (e.g., No-cut 8, thin 30, var. ret. 52 = no-cut out to 8-m, thin out to 30-m, variable retention out to 52-m). In some cases, the study unit includes a discontinuous buffer ("patch") for a portion of unit length and an upstream clearcut (e.g., No-cut 15+, patch 55%-73% lng. = no-cut buffer out to 15 + m wide that ranged 55%-73% of study unit length). [b]McIntyre et al. present proportional responses. For example, 0.53 indicates a 47% decline across a specific time period and 2.35 indicates a 235% increase across a specific time period.

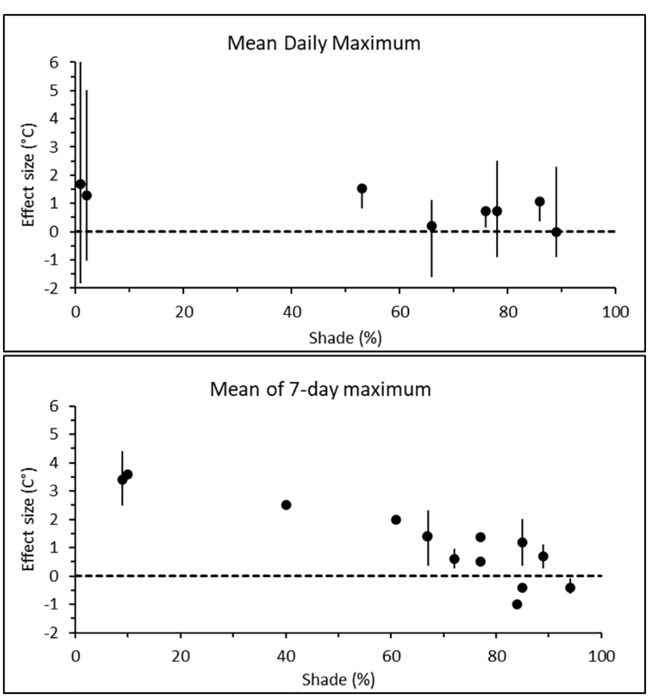

Fig. 3. Plots of summer temperature effect size in relation to shade for the mean daily maximum (A) and mean 7-day maximum (B) in riparian buffer studies, Pacific Northwest of Canada and United States. Vertical lines express outcome variability (range or 95% CI) where data are available (Table 1).

no more effective for minimizing temperature change than discontinuous patch buffers. In another comparison, we found that wide no-cut and combination treatments (i.e., no-cut 23–33; No-cut 15, thin ret. 33–55% BA ∞) have smaller effect sizes (Fig. 4a) than narrow no-cut and combination treatments (i.e., no-cut 15+, no-cut 6, var. ret 15), but all four treatments apparently have similar levels of shade (plot points within box; Fig. 4b). These inconsistencies between treatment effectiveness and shade are confounded by comparisons between treatments using different shade metrics which are not equivalent or comparisons with the same metric (canopy closure) which are highly variable and of low precision (Allen and Dent, 2001; Kelley and Krueger, 2005). However, when the canopy closure values were converted to estimated effective shade (see Data Extraction and Analysis) and plotted along with findings from studies that measured effective shade, the results clearly indicated an inverse relationship between the mean 7-day maximum effect size and shade for all treatments (Fig. 4c). Further, the plot indicated that continuous no-cut and combination buffers (blue and green points, Fig. 4c) were more effective for minimizing temperature change than discontinuous patch buffers (orange points) because the former treatments provided higher levels of effective shade. Similarly, the wider no-cut and combination treatments provided more effective shade and were more effective than the narrow no-cut and combination treatments. Also, multiple treatments existed, including shade-targeted, continuous no-cut, and combination buffers, that provided more effective shade and had smaller effect sizes than the no-cut discontinuous patch buffers.

The demonstrated relationship between temperature response and effective shade is consistent with the large body of science on stream shading. Retaining riparian vegetation to attenuate incoming solar

Table 3
Study quality assessment for temperature and amphibian responses in riparian buffer studies, Pacific Northwest of Canada and United States. See Appendix B for description of attributes and ranking criteria. Maximum possible score = 24; attribute scores ranked as low = 1, medium = 2, and high = 3.

Citation	Response	Quality score total	Experimental design	Control replicates design	Site selection bias	Number of treatment replicates	Number of pre-treatment replicates	Number of post-treatment years	Statistical robustness	Peer review
Groom et al. (2011)	Temperature	21	3	3	1	3	2	3	3	3
Kibler et al. (2013)		19	3	3	1	2	3	1	3	3
Moore et al. (2005a)		19	3	3	1	1	2	3	3	3
Janisch et al. (2012)		19	3	2	1	2	2	3	3	3
Reiter et al. (2020)		19	3	2	1	1	3	3	3	3
MacCracken et al. (2018)		18	3	3	1	2	2	2	3	2
Bladon et al. (2016)		18	3	3	1	1	2	2	3	3
McIntyre et al. (2018), McIntyre et al. (2021)		18	3	2	1	2	2	3	3	2
Gomi et al. (2006)		18	3	1	1	2	2	3	3	3
Ehinger et al., in review		17	3	2	1	2	2	2	3	2
Farber and Whitaker (2010a)		16	3	3	1	2	2	2	2	1
Farber and Whitaker (2010b)		15	3	3	1	1	2	2	2	1
Jackson et al. (2001)		15	3	3	1	2	1	1	1	3
Guenther (2007)		15	3	3	1	1	1	1	3	2
Veldhuisen and Couvelier (2006)		11	2	3	1	1	1	1	1	1
MacCracken et al. (2018)	Amphibian	21	3	3	3	3	2	2	3	2
McIntyre et al. (2018)		20	3	1	2	3	3	3	3	2
Hawkes and Gregory (2012)		19	3	3	2	2	2	3	1	3
Dupuis and Steventon (1999)		17	1	3	2	3	1	1	3	3
Vesely and McComb (2002)		16	1	3	2	3	1	2	1	3
Pollett et al. (2010)		15	1	1	2	3	1	3	1	3
Wahbe and Bunnell (2003)		14	1	1	2	2	1	3	1	3

radiation (direct and diffuse) is the best-management practice for conserving water temperature in streams (Brazier and Brown, 1973; Beschta et al., 1987; Johnson, 2004). Direct-beam solar radiation on the water's surface is the dominant source of heat energy that may be absorbed by the water column and streambed (Brown, 1969; Johnson, 2004). Therefore, riparian vegetation that blocks direct solar radiation along the sun's pathway across the sky is the most effective way to reduce radiant energy available for stream heating (Moore et al., 2005a). Previous research demonstrated that attenuation of direct beam radiation by riparian vegetation is a function of canopy height, vegetation density, and buffer width (Beschta et al., 1987; Sridhar et al., 2004; DeWalle, 2010). Light attenuation increases with increasing canopy height and increasing buffer density because of the increased solar path length and extinction of energy. Buffer width has a variable influence on light attenuation depending on stream azimuth (e.g., buffer widths for E-W streams may be narrower than for N-S streams due to shifts in solar beam pathway from the sides to the tops of the buffers; DeWalle, 2010).

4.1.2. Amphibians

In the studies included in the review, we did not find consistent amphibian population responses to riparian buffers. Responses varied within and across taxa as indicated by differences in the direction and magnitude of reported responses, as well as the broad confidence interval coverage for the responses.

Inconsistency in the patterns could result from several factors. First, six of the studies presented site level comparisons for stream reaches of relatively short length (<60 m). One study (McIntyre et al., 2021)

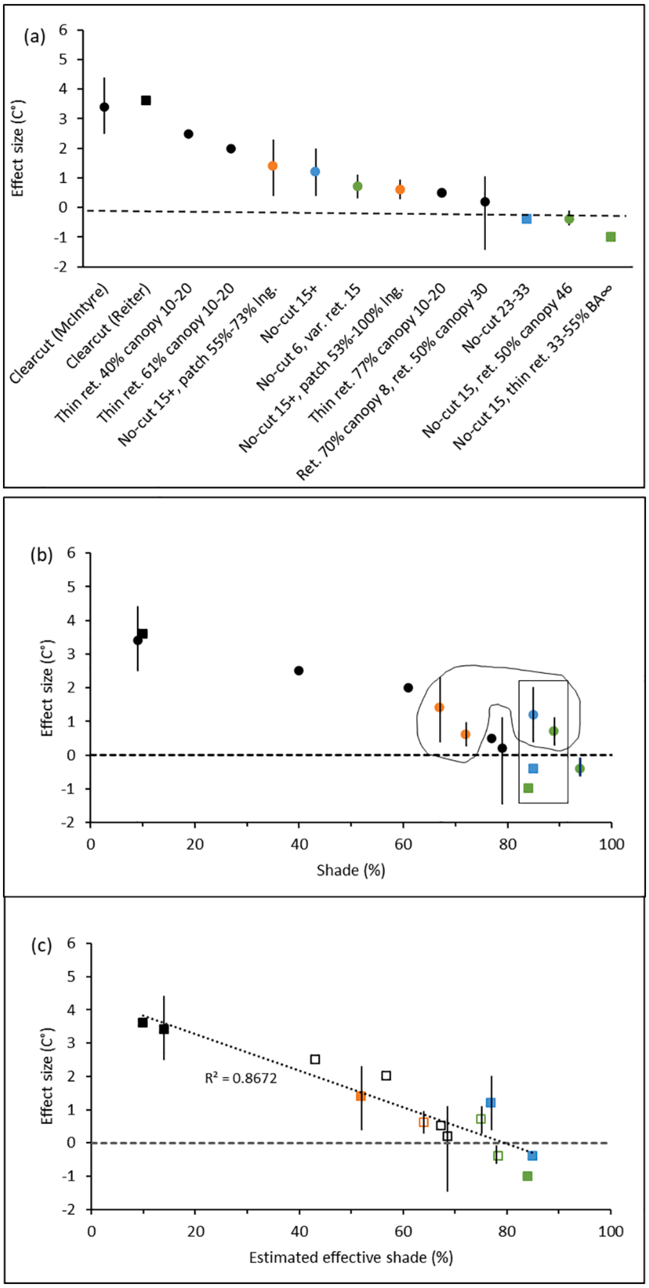

Fig. 4. Scatter plots of maximum post-treatment temperature effect size by riparian treatment, ordered by effect size (A) and temperature effect size in relation to shade (B), and estimated effective shade (C) for the mean 7-day maximum during summer in riparian buffer studies, Pacific Northwest of Canada and United States. Symbol colors denote treatment types: black = shade-targeted, orange = discontinuous patch, blue = continuous no-cut, green = combination. Symbol shapes denote shade metrics: circles = canopy closure, squares = effective shade, empty square = estimated effective shade. Points within box and polygon areas = data point comparisons (see text, Question One—Temperature). Vertical lines express outcome variability (range or 95% CI) where data are available (see Table 1). See Table 1 footnote for description of riparian treatment abbreviations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

presented estimates of population responses for 2nd order basins. Sampling programs to estimate responses at the basin scale are more costly but may provide more accurate representations of how amphibian populations respond to different prescriptions than reach-scale studies that sample only population segments at the reach scale (e.g., Kroll et al., 2010). Second, four of the studies were observational in nature and did not include multi-year time series, including pre- and post-treatment sampling to establish base-line conditions prior to treatment implementation.

By manipulating shade directly, MacCracken et al. (2018) make specific inference about how changes in shade affected ecological responses, including amphibian populations. MacCracken et al. (2018) was the only study included in the review that manipulated shade levels directly and provided stronger inference about putative mechanisms behind sampled responses; in the other six studies, shade was manipulated indirectly through general buffer prescriptions. The latter studies can be characterized as "applied" observational or experimental studies, in which a single, or multiple, responses are sampled across two or more buffer treatments. However, because the treatments may manipulate multiple factors within treatments (e.g., other covariates might change in addition to the amount of shade), understanding the exact nature of the relationship between the response and a specific treatment is challenging.

The amphibian taxa we included have broad geographic distributions from coastal northern California, USA, to the interior of British Columbia, Canada. Given the variation in riparian conditions across this region, expecting general trends to appear in responses to buffer treatments may be unrealistic. Also, although two studies evaluated old-growth (unharvested) forest as a treatment type, other studies applied treatments to second growth stands of different ages. Given heterogeneity in stand structure and composition, substantial variation could have occurred across responses that was unrelated to the treatments per se (Ohmann et al., 2007). In contrast, we documented substantial variation in temperature changes associated with different buffer treatments. How spatial and temporal variation in temperature response, in turn, causes significant biological effects in stream-associated amphibians across their geographic distributions has received little consideration. Finally, legacy effects of previous harvests may be substantial and continue to occur for decades after harvest (Harding et al., 1998). In the absence of specific information about how harvests were implemented or statistical adjustments for this possibility (e.g., treating sites as random effects in an analysis), it is difficult to preclude the possibility that variation in responses to buffer treatments is also affected by site-level factors not addressed in analyses.

Although not all stream-associated amphibian taxa were evaluated in each study included in the review, we note that information is not available about interactions (e.g., direct or indirect competition and/or predation) across the species or how these interactions may be altered by anthropogenic disturbance such as timber harvest. For example, McIntyre et al. (2021) chose study sites on the basis that all three taxa were present, a decision that may have important consequences for inference about effectiveness of buffers of different sizes for retaining species. Although overlap exists in the species distributions within a watershed, all three taxa often occur in isolation (Corn and Bury, 1989; Russell et al., 2004; Kroll et al., 2008). In these cases, responses to buffer prescriptions could differ as opposed to when distributions are sympatric and competition and other interactions occur.

Although a substantial number of studies were excluded from the review due to inadequate presentation of data and statistical summaries (e.g., no effect sizes were presented), we emphasize that concerns remain for the studies included in the review. For example, none of the studies included in the review determined *a priori* whether sample sizes were sufficient to estimate quantities of interest (an exercise often referred to as a power analysis). As a result, spatial and temporal replication may not have been as large as was needed to summarize variability in responses adequately. For studies with limited temporal replication, variation from treatments may be confounded with "annual" variation related to factors such as temperature and precipitation (although we note that, under specific study designs one year of sampling may be preferred to avoid this problem). Finally, the lack of discernment between statistical and biological effect sizes in all the studies save one restricts understanding about if and how amphibian populations respond to different buffer treatments (Kroll, 2009; Kroll et al., 2009; Wasserstein et al., 2019).

4.2. Question Two

External validity refers to the relevance of studies for informing the context of the primary question. Synthesis question 2 was intended to supplement the primary question by assessing how well each investigation and the collective body of evidence from all studies provides context concerning how study treatments inform or were linked to causal factors influencing temperature response. For example, in a well-conducted research study of causation, the intervention (riparian treatment) would be the explanatory variable and the effect (measured effect size) would be the dependent variable, and these would be linked by one or more clearly-specified cause-effect pathways (Collaboration for Environmental Evidence, 2018).

4.2.1. Temperature

Among the temperature studies, only one (MacCracken et al., 2018) implemented a gradient of shade-targeted treatments to examine directly the effects of different shade/light levels (hypothesized causal mechanism) on stream temperature and aquatic populations (see amphibian findings section). Consequently, MacCracken et al. (2018) has the highest ranking of external validity by experimentally demonstrating a clearly specified cause-effect pathway for the measured outcomes (Question 2; Appendix E). Further, knowledge gained from MacCracken et al. (2018) is transferable because it quantifies how shade alone influences relative effectiveness of riparian treatments. In contrast, transferability of findings from other studies must be done carefully because the detection of causal factors is confounded by multiple covariates that likely influenced study outcomes. For example, riparian treatments of other studies included adjacent upland timber harvests that can change hydrologic dynamics (e.g., summer stream flow, surface flow extent). As a result, separation of riparian treatment effects (shade) from other factors that influence water temperature is challenging.

Evidence from quantitative analyses of environmental covariates indicated that basin physiographic characteristics and post-harvest changes in hydrology likely influenced relative effectiveness of riparian treatments to maintain water temperature. For example, Janisch et al. (2012) reported that surface flow extent was highly correlated (0.81, $p = 0.05$) with temperature effect size and that correlations with shade were not significant ($r = -0.22$, $p = 0.57$). Both Washington studies of patch buffers (McIntyre et al., 2021; Ehinger et al., in review) also reported moderate positive correlations with surface flow extent. Other studies found that that temperature response increased with decreasing channel gradients (Veldhuisen and Couvelier, 2006; Groom et al., 2011), and decreasing stream flow (Moore et al., 2005a; Guenther, 2007). Longitudinal variability of temperature within study reaches was identified by several studies and indicated that substrate texture and channel morphology influenced hyporheic exchange and stream–groundwater interactions (Moore et al., 2005a; Janisch et al., 2012). Further, longitudinal variability could confound evaluation of treatment effectiveness because temperature data from a single monitoring station at the lower end of a study reach may not have provided an accurate measure of riparian treatment response (Moore et al., 2005b). Windthrow of buffer strips was another factor confounding assessment of treatment effectiveness as several studies indicated that post-harvest windthrow influenced shade loss (Supplemental Appendix B).

4.2.2. Amphibians

Similarly, for amphibians, only one study (MacCracken et al., 2018) implemented a gradient of shade levels as the treatment rather than implementing buffers at different levels of retention. In so doing, MacCracken et al. (2018) provide direct inference about effects of different shade, or light, levels (hypothesized causal mechanism) on stream temperature and aquatic populations. Consequently, MacCracken et al. (2018) received the highest ranking of external validating by providing experimental evidence for a cause-effect pathway for the measured outcomes (Question 2; Appendix E). In addition, MacCracken et al. (2018) is unique as riparian habitat was modified without also modifying upland habitat; in other studies, riparian buffer treatments could be confounded with upland habitat condition (e.g., controls are unmodified or buffers of varying width are left after harvest of adjacent upland areas). Harvest of adjacent upland areas can modify hydrologic dynamics, including summer stream flow and surface flow extent, that contribute to sampled responses. As a result, variation not associated with the riparian treatments themselves can increase uncertainty in treatment effect estimates.

In the review, we did note specific instances where authors evaluated if and how amphibian responses were associated with covariates. This type of effort occurs frequently in the literature for stream-associated amphibian responses to forest management (Kroll, 2009). Several potential problems exist with this approach including insufficient number of replicates compared to the number of measured covariates; potential correlations among covariates; and no pre-sampling stratification to identify how populations may be distributed across the range of categorical and continuous covariates. In contrast, experimental manipulation of factors of interest is more likely to yield strong inference to determine if populations respond to covariates. As a result, we suggest circumspect interpretations of reported associations between amphibian responses and covariates in published studies (Kroll et al., 2009).

4.3. Question Three

Synthesis question 3 is intended to supplement the primary question by assessing how well the sampled population represented or provided inference to an intended target population or landscape.

4.3.1. Temperature

No study had high external validity for question 3 because none collected a random and representative sample from a target population of headwater streams (Appendix E). Most studies that intended representative sampling from a geographic region were hampered by the reality of timber harvest schedules and site availability, resulting in non-random and opportunistic acquisition of study sites.

The transferability of study findings across the landscape is limited by their geographic distribution and by the compatibility of study designs and metrics. Several studies with multiple sample sites were concentrated within the coastal ecoregion (Fig. 1). Groom et al. (2018) provided the largest and spatially extensive sample of Oregon forest practices within the coastal ecoregion. McIntyre et al. (2018) and Ehinger et al. (in review) provided a smaller, but extensive sample of the Washington non-fish-stream forest practices from the coastal ecoregion. Collectively, all studies (Richardson and Danehy, 2007) within the coastal ecoregion provided a spatially extensive sample, but the comparability of findings, hence external validity is hindered by different treatments and study designs.

Given differences in study design, treatment comparisons between ecoregions can only be made within a single study. Among the synthesis studies, only three (Janisch et al., 2012; MacCracken et al., 2018) had multiple sample sites among ecoregions. Unfortunately, none of the studies directly address potential differences at the ecoregion scale. However, lithology may influence treatment response as Janisch et al. (2012) found that streams with fine-textured streambed sediment (i.e., marine sediment lithology) were thermally more responsive to shade loss than streams with coarse-textured substrate (i.e., basalt lithology).

4.3.2. Amphibians

We did not find any studies that had high external validity for question 3 because none collected an unbiased and representative sample from a specified target population (Appendix E). Although three studies recognized the importance of representative sampling from a geographic region, the reality of timber harvest scheduling resulted in non-random and opportunistic implementation of treatments on those sites that were available.

As a result, transferability of study findings across the landscape is mixed due to restricted geographic distributions within studies. Three studies (Hawkes and Gregory, 2012; MacCracken et al., 2018) featured blocks (in which each treatment was implemented on an independent replicate) distributed across multiple ecoregions (Fig. 1). However, these three studies had different treatment prescriptions and sampled different responses, so drawing general conclusions from the studies is challenging.

5. Conclusions

5.1. Implications for policy

5.1.1. Temperature

The evidence we reviewed indicated that relative effectiveness of riparian treatments to maintain water temperature in headwater streams depended on retaining riparian vegetation to block incoming solar radiation (i.e., provided effective shade). Also, the evidence suggested that buffer effectiveness was associated weakly with various prescriptive components (e.g., fixed-width or patch-buffer) that were not designed specifically to block direct solar radiation. Consequently, relative effectiveness of most prescriptive treatments implemented uniformly are highly variable. Finally, the evidence indicated that multiple geophysical factors (geology, hydrology, topography, latitude, and stream azimuth) can influence thermal sensitivity of streams to shade loss. To address these issues, we suggest that management schemes consider alternatives tailored to site-specific conditions.

Shade loss due to buffer windthrow also needs to be considered in management schemes for locations prone to wind damage. At the landscape scale, windthrow mortality is highly variable and can demonstrate a skewed mortality distribution (i.e., most sites have low mortality and a few have high mortality; Grizzel and Wolff, 1998; Martin and Grotefendt, 2007; Beese et al., 2019). Wind damage is strongly associated with buffer orientation relative to the predominant storm direction (i.e., southeast, south, southwest in the Pacific Northwest) and local conditions (including wind fetch length resulting from the size of clearcut units; Kramer et al., 2001; Mitchell et al., 2001; Beese et al., 2019). Reductions in windthrow mortality are feasible when site and landscape factors are considered in harvest unit plans (Kramer et al., 2001; Mitchell et al., 2001; Beese et al., 2019).

The temperature regimes of headwater streams were spatially and temporally variable. Therefore, implementation of uniform buffer prescriptions to achieve a desired temperature target is questionable given the intrinsic spatial variability of thermal loading and temperature controlling processes within and among watersheds. Further, temperature targets are static and do not consider natural cycles including disturbances at multiple temporal scales (Poole et al., 2004). Application of a "regime standard" that would describe desirable distributions of temperature conditions over space and time within a stream network may address environmental variability successfully (Poole et al., 2004). However, implementation poses a challenge as managers would need to identify and validate the suitability of temperature distributions across a spatially variable landscape. An alternative is to shift management from temperature targets to a focus on maintaining key ecological processes (Klenk et al., 2008; Reeves et al., 2018). Here, the objective is to implement flexible riparian management schemes that incorporate modern technology to derive site-specific, and potentially watershed level, prescriptions for managing thermal loading processes best suited for a particular locale.

5.1.2. Amphibians

This systematic review found mixed evidence about relationships between amphibian population responses and buffers that were maintained along streams during forest harvest in the Pacific Northwest of North America. Evidence was scant to support recommendations for specific buffer widths to achieve specific population outcomes such as maintaining abundance or density. Also, considerable uncertainty exists about which environmental covariates are associated strongly with amphibian population responses. Given the substantial variation in ecological conditions across the region, including stand structure and composition, lithology, management history, and natural disturbance regimes, it may be unrealistic to expect that a "one size fits all prescription" can be identified. We cannot reject the possibility that, on some streams, buffers may not be required to maintain adequate conditions to retain amphibian populations at pre-harvest levels or that buffers larger than those evaluated in the reviewed studies may be required. As a result, retaining a range of buffers on streams classified based on stream size, site productivity, or other covariates may be a conservative and prudent approach.

5.2. Implications for research

5.2.1. Temperature

This synthesis and related reviews of the literature (e.g., Richardson et al., 2012) demonstrated that few studies have tested the effectiveness of riparian treatments tailored to site-specific conditions for headwater streams. Further, the implementation of uniform buffers (e.g., fixed-width forest stands on both sides of a channel) has been the dominant paradigm in watershed management including federal (Forest Ecosystem Management Assessment Team [FEMAT], 1993) as well as state and private lands (Richardson et al., 2012; Jackson et al., 2021). Buffers are designed to provide overall stream protection and to standardize implementation. However, as we demonstrated in this synthesis, buffer effectiveness is highly variable and poorly linked to uniform and simplified criteria (e.g., length and width). Technical information is lacking for addressing policy-relevant questions concerning the effectiveness of alternative riparian management schemes (Reeves et al., 2018). Future studies should prioritize testing the effectiveness of alternative treatments (e.g., use models to generate predictions to be evaluated with field studies) that will result in developing tools and deriving guidelines for how to achieve management goals with consideration for site and landscape characteristics.

5.2.2. Amphibians

Here, we reiterate recommendations in Kroll (2009) to strengthen inference about amphibian population responses to buffer treatments specifically, and forest management generally, within the region. Although experimental designs with sufficient pre- and post-treatment spatial and temporal replication are recognized as critical to support strong inference, we emphasize that studies of this type are unlikely to provide the broad spatial coverage necessary to partition variation arising from background environmental conditions in the habitats occupied by stream-associated amphibian taxa. As such, experimental studies run the risk of being expensive "case studies" or studies in which outcomes fail to match objectives because the amount of replication was inadequate to estimate spatial variation arising from factors other than the treatments. As an alternative, we propose that multiple observational studies, incorporating judiciously selected and stratified covariates and using contemporary sampling programs and analytical techniques, are likely to provide the inference required to direct management prescriptions and support desired conservation outcomes (Nichols et al., 2019).

CRediT authorship contribution statement

Douglas J. Martin: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing, Supervision. **Andrew J. Kroll:** Conceptualization, Investigation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing. **Jenny L. Knoth:** Investigation, Resources, Writing - review & editing, Data curation.

Declaration of Competing Interest

Douglas Martin and Jenny Knoth are self-employed environmental consultants and have provided consultancies for the forest industry. A.J. Kroll is employed by Weyerhaeuser, which provided funding to support this research.

Acknowledgments

Reviews by A. Coble, J. Gravelle, T. Link, J. Homyack, and J. Verschuyl greatly improved the manuscript. Bernard Romey prepared the study site map. Jeremy Groom, S. Farber, and C. Veldhuisen provided supplemental information for their respective studies. The National Council for Air and Stream Improvement, Inc. provided funding to conduct this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foreco.2021.119190.

References

Allen, M., Dent, L., 2001. Shade conditions over forested streams in the Blue Mountain and coast range georegions of Oregon. Oregon Department of Forestry, Forest Practices Monitoring Program, Salem, OR, USA.

Beese, W.J., Rollerson, T.P., Peters, C.M., 2019. Quantifying wind damage associated with variable retention harvesting in coastal British Columbia. For. Ecol. Manage. 443, 117–131.

Benda, L., Dunne, T., 1997. Stochastic forcing of sediment supply to channel networks from landsliding and debris flow. Water Resour. Res. 33, 2849–2863.

Beschta, R.L., Bilby, R., Brown, G., Hoitby, L., Hofstra, T., 1987. Stream temperature and aquatic habitat: fisheries and forestry interactions. In: Salo, E.O., Cundy, T.W. (Eds.), Streamside Management: Forestry and Fisheries Interactions. University of Washington, Institute of Forest Resources, Seattle, WA, pp. 191–232.

Bilby, R.E., Beach, E.W., Fransen, B.R., Walter, J.K., Bisson, P.A., 2003. Transfer of nutrients from spawning salmon to riparian vegetation in western Washington. Trans. Am. Fish. Soc. 132, 733–745.

Bladon, K.D., Cook, N.A., Light, J.T., Segura, C., 2016. A catchment-scale assessment of stream temperature response to contemporary forest harvesting in the Oregon Coast Range. For. Ecol. Manage. 379, 153–164.

Bowler, D.E., Mant, R., Orr, H., Hannah, D.M., Pullin, A.S., 2012. What are the effects of wooded riparian zones on stream temperature? Environ. Evidence 1, 3.

Boyd, M., Kasper, B., 2003. Analytical methods for dynamic open channel heat and mass transfer: methodology for heat source model Version 7.0. Oregon Department of Environmental Quality, Salem, OR, USA.

Brazier, J.R., Brown, G.W., 1973. Buffer strips for stream temperature control. In: School of Forestry. Oregon State University, Corvallis, OR, p. 9.

Brown, G.W., 1969. Predicting temperatures of small streams. Water Resour. Res. 5, 68–75.

Bull, E.L., 1978. Specialized habitat requirements of birds: snag management, old growth, and riparian habitat. In: DeGraaf, R.M. (Ed.), Proceedings of the workshop on nongame bird habitat management in the coniferous forests of the western United States. USDA Forest Service General Technical Report PNW-64, Pacific Northwest Forest and Range Experiment Station, Portland, OR, USA, pp. 74–82.

Chaloner, D.T., Martin, K.M., Wipfli, M.S., Ostrom, P.H., Lamberti, G.A., 2002. Marine carbon and N in southeastern Alaska food webs: evidence from artificial and natural streams. Can. J. Fish. Aquatic Sci. 59, 1257–1265.

Chamberlin, T.W., Harr, R.D., Everest, F.H., 1991. Timber harvesting, silviculture, and watershed processes. Am. Fish. Soc. Special Publ. 19, 181–206.

Cole, L., Newton, M., 2015. Relationships between radiation and canopy closure estimates in streamside buffers in western Oregon. Fors. Sci. 61, 559–569.

Collaboration for Environmental Evidence. 2018. Guidelines and Standards for Evidence synthesis in Environmental Management, Version 5.0. Pullin, A.S., Frampton, G.K., Livoreil, B., Petrokofsky, G. (Eds.). www.environmentalevidence.org/information-for-authors.

Corn, P.S., Bury, R.B., 1989. Logging in western Oregon: responses of headwater habitats and stream amphibians. For. Ecol. Manage. 29, 39–57.

Czarnomski, N., Hale, C., Frueh, W., Allen, M., Groom, J., 2013. Effectiveness of riparian buffers at protecting stream temperature and shade in Pacific Northwest Forests: a systematic review. Oregon Department of Forestry, Salem, OR, USA.

DeWalle, D.R., 2010. Modeling stream shade: riparian buffer height and density as important as buffer width. JAWRA J. Am. Water Resources Assoc. 46, 323–333.

Dupuis, L., Steventon, D., 1999. Riparian management and the tailed frog in northern coastal forests. For. Ecol. Manage. 124, 35–43.

Ehinger, W.J., Schuett-Hames, D., Stewart, G., in review. Type N Experimental Buffer Treatment Study in Incompetent Lithologies: Riparian Inputs, Water Quality,and Exports to Fish-Bearing Waters, Cooperative Monitoring Evaluation and Research, Washington Department of Natural Resources, Olympia, WA, USA.

Farber, S., Whitaker, J., 2010a. Monitoring results of alternative watercourse and lake protection zones in the Etna Creek watershed in interior. Northern California, Timber Products Company, Yreka, CA, USA.

Farber, S., Whitaker, J., 2010b. Monitoring results of alternative watercourse and lake protection zones in the McKinney Creek watershed in interior. Northern California, Timber Products Company, Yreka, CA, USA.

Fialaa, A.C., Garmana, S.L., Gray, A.N., 2006. Comparison of five canopy cover estimation techniques in the western Oregon Cascades. For. Ecol. Manage. 232, 188–197.

Forest Ecosystem Management Assessment Team [FEMAT], 1993. Forest ecosystem management: an ecological, economic, and social assessment., Report of the Forest Ecosystem Management Assessment Team, U.S. Government Printing Office, Washington, D.C., USA.

Franklin, J.F., Dyrness, C.T., 1973. Natural vegetation of Oregon and Washington. In: U. S. Forest Service GTR-PNW-8, U.S. Department of Agriculture, Portland, OR, USA.

Gomi, T., Moore, R.D., Dhakal, A.S., 2006. Headwater stream temperature response to clear-cut harvesting with different riparian treatments, coastal British Columbia, Canada. Water Resour. Res. 42, W08437.

Grant, E.H.C., Nichols, J.D., Lowe, W.H., Fagan, W.F., 2010. Use of multiple dispersal pathways facilitates amphibian persistence in stream networks. Proc. Natl. Acad. Sci. U. S. A. 107, 6936–6940.

Grizzel, J., Wolff, N., 1998. Occurrence of windthrow in forest buffer strips and its effect on small streams in northwest Washington. Northwest Sci. 72, 214–223.

Groom, J.D., Dent, L., Madsen, L.J., 2011. Stream temperature change detection for state and private forests in the Oregon Coast Range. Water Resour. Res. 47, 12.

Groom, J.D., Madsen, L.J., Jones, J.E., Giovanini, J.N., 2018. Informing changes to riparian forestry rules with a Bayesian hierarchical model. Forest Ecology and Management 419/420, 17–30.

Guenther, S.M., 2007. Impacts of partial-retention harvesting with no buffer on the thermal regime of a headwater stream and its riparian zone. University of British Columbia, Vancouver, BC, Canada.

Haddaway, N.R., Bethel, A., Dicks, L.V., Koricheva, J., Macura, B., Petrokofsky, G., Pullin, A.S., Savilaakso, S., Stewart, G.B., 2020. Eight problems with literature reviews and how to fix them. Nat. Ecol. Evol. 4, 1582–1589.

Harding, J.S., Benfield, E.F., Bolstad, P.V., Helfman, G.S., Jones III, E.B.D., 1998. Stream biodiversity: the ghost of landuse past. Proc. Natl. Acad. Sci. U. S. A. 95, 14843–14847.

Hawkes, V.C., Gregory, P.T., 2012. Temporal changes in the relative abundance of amphibians relative to riparian buffer width in western Washington, USA. For. Ecol. Manage. 274, 67–80.

Helfield, J.M., Naiman, R.J., 2002. Salmon and alder nitrogen sources to riparian forests in a boreal Alaskan watershed. Oecologia 133, 573–582.

Jackson, C.R., Ice, G.G., Cundy, T.W., Tschaplinski, P.J., Martin, D.J., Groom, J.D., Ehinger, W.J., 2021. Determining the effectiveness of contemporary forestry best management practices: illustrative case studies. In: Danehy, R.J., Dolloff, C.A. (Eds.). Reflections on Forest Management: can Fish and Fiber Coexist?, American Fisheries Society, Symposium 92, Bethesda, MD, USA, pp. XX–XX.

Jackson, C.R., Sturm, C.A., Ward, J.M., 2001. Timber harvest impacts on small headwater channels in the Coast Ranges of Washington. J. Am. Water Resour. Assoc. 37, 1533–1550.

Janisch, J.E., Wondzell, S.M., Ehinger, W.J., 2012. Headwater stream temperature: Interpreting response after logging, with and without riparian buffers, Washington, USA. For. Ecol. Manage. 270, 302–313.

Jennings, S.B., Brown, N.D., Sheil, D., 1999. Assessing forest canopies and understory illumination, canopy closure, canopy cover and other measures. Forestry 72, 59–73.

Johnson, S.L., 2004. Factors influencing stream temperatures in small streams: substrate effects and a shading experiment. Can. J. Fish. Aquat. Sci. 61, 913–923.

Kelley, C.E., Krueger, W.C., 2005. Canopy cover and shade determinations in riparian zones. J. Am. Water Resour. Assoc. 41, 37–046.

Kibler, K.M., Skaugset, A., Ganio, L.M., Huso, M.M., 2013. Effect of contemporary forest harvesting practices on headwater stream temperatures: initial response of the Hinkle Creek catchment, Pacific Northwest, USA. For. Ecol. Manage. 310, 680–691.

Kiffney, P.M., Richardson, J.S., Bull, J.P., 2004. Establishing light as a causal mechanism structuring stream communities in response to experimental manipulation of riparian buffer width. J. North Am. Benthol. Soc. 23, 542–555.

Klenk, N., Bull, G., Cohen, D., 2008. What is the "END'' (emulation of natural disturbance) in forest ecosystem management? an open question. Can. J. For. Res. 38, 2159–2168.

Knopf, F.L., Johnson, R.R., Rich, T., Samson, F.B., Szaro, R.C., 1988. Conservation of riparian ecosystems in the United States. Wilson Bull. 100, 272–284.

Kovach, R.P., Muhlfeld, C.C., Al-Chokhachy, R., Ojala, J.V., Archer, E.K., 2019. Effects of land use on summer thermal regimes in critical salmonid habitats of the Pacific Northwest. Can. J. Fish. Aquat. Sci. 76, 753–761.

Kramer, M.A., Hansen, M.T., Kissinger, E., 2001. Abiotic controls on long-term windthrow disturbance and temperate rain forest dynamics in Southeast Alaska. Ecology 82, 2749–2768.

Kreutzweiser, D.P., Sibley, P.K., Richardson, J.S., Gordon, A.M., 2012. Introduction and a theoretical basis for using disturbance by forest management activities to sustain aquatic ecosystems. Freshw. Sci. 31, 224–231.
Kroll, A.J., 2009. Sources of uncertainty in stream-associated amphibian ecology and responses to forest management in the Pacific Northwest, USA: A review. For. Ecol. Manage. 257, 1188–1199.
Kroll, A.J., Hayes, M.H., MacCracken, J.G., 2009. Concerns regarding the use of amphibians as metrics of critical biological thresholds: a comment on Welsh and Hodgson 2008. Freshw. Biol. 54, 2364–2373.
Kroll, A.J., MacCracken, J.G., McBride, T.C., Bakke, J., Light, J., Peterson, P., Bach, J., 2010. Basin-scale surveys of stream-associated amphibians in intensively managed forests. J. Wildl. Manage. 74, 1580–1587.
Kroll, A.J., Risenhoover, K.L., McBride, T., Beach, E., Kernohan, B.J., Light, J., Bach, J., 2008. Factors influencing stream occupancy and detection probability parameters of stream-associated amphibians in commercial forests of Oregon and Washington, USA. For. Ecol. Manage. 255, 3726–3735.
MacCracken, J.G., Hayes, M.P., Tyson, J.A., Stebbings, J.L., 2018. Stream-associated amphibian response to manipulation of forest canopy shading. Washington Department of Natural Resources, Olympia, WA.
Martin, D.J., Grotefendt, R.A., 2007. Stand mortality in buffer strips and the supply of woody debris to streams in Southeast Alaska. Can. J. For. Res. 37, 36–49.
McIntyre, A. P., Hayes, M. P., Ehinger, W. J., Estrella, S. M., Schuett-Hames, D., Quinn, T. (2018). Effectiveness of experimental riparian buffers on perennial non-fish-bearing streams on competent lithologies in western Washington. In, Cooperative Monitoring, Evaluation and Research Report CMER 18-100, Washington State Forest Practices Adaptive Management Program, Washington Department of Natural Resources, Olympia, WA, USA.
McIntyre, A.P., Hayes, M.P., Ehinger, W.J., Estrella, S.M., Schuett-Hames, D., Quinn, T., 2021. Type N Experimental Extended Report. In. Washington State Forest Practices Adaptive Management Program. Washington Department of Natural Resources, Olympia, WA.
Mitchell, S., Hailemariam, T., Kulis, Y., 2001. Empirical modeling of cutblock edge windthrow risk on Vancouver Island, Canada, using stand level information. For. Ecol. Manage. 154, 117–130.
Moore, D.R., Spittlehouse, D.L., Story, A., 2005a. Riparian microclimate and stream temperature response to forest harvesting: A review. J. Am. Water Resour. Assoc. 41, 813–834.
Moore, R.D., Sutherland, P., Gomi, T., Dhakal, A., 2005b. Thermal regime of a headwater stream within a clear-cut, coastal British Columbia, Canada. Hydrol. Process. 19, 2591–2608.
Naiman, R.J., Bilby, R.E. (Eds.), 1998. River Ecology and Management. Springer, New York, NY, USA.
Naiman, R.J., Décamps, H., McClain, M.E., 2005. Riparia: ecology, conservation, and management of streamside communities. Elsevier Academic Press, San Diego, CA, USA.
Nakagawa, S., Cuthill, I.C., 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol. Rev. 82, 591–605.
Nakano, S., Murakami, M., 2001. Reciprocal subsidies: dynamic interdependence between terrestrial and aquatic food webs (foreststream ecotone/allochthonous prey flux). Proc. Nat Acad. Sci. 98, 166–170.
National Research Council, 2002. Riparian Areas: Functions and Strategies for Management. The National Academies Press, Washington, DC, USA.
Nichols, J.D., Kendall, W.L., Boomer, G.S., 2019. Accumulating evidence in ecology: Once is not enough. Ecol. Evol. 9, 13991–14004.
Nichols, S.J., Peat, M., Webb, J.A., 2017. Challenges for evidence-based environmental management: what is acceptable and sufficient evidence of causation? Freshw. Sci. 36, 240–249.
Ohmann, J.L., Gregory, M.J., Spies, T.A., 2007. Influence of environment, disturbance, and ownership on forest vegetation of coastal Oregon. Ecol. Appl. 17, 18–33.
Osmundson, D.B., Ryel, R.J., Lamarra, V.L., Pitlick, P., 2002. Flow-sediment-biota relations: implications for river regulation on native fish abundance. Ecol. Appl. 12, 1719–1739.
Pearson, S.F., Giovanini, J., Jones, J.E., Kroll, A.J., 2015. Breeding bird community continues to colonize riparian buffers ten years after harvest. PLoS ONE 10, e0143241.
Pettit, N.E., Naiman, R.J., 2007. Fire in the riparian zone: characteristics and ecological consequences. Ecosystems 10, 673–687.
Pollett, K.L., MacCracken, J.G., MacMahon, J.A., 2010. Stream buffers ameliorate the effects of timber harvest on amphibians in the Cascade Range of Southern Washington, USA. For. Ecol. Manage. 260, 1083–1087.
Poole, G.C., Dunham, J.B., Keenan, D.M., Sauter, S.T., McCullough, D.A., Mebane, C., Lockwood, J.C., Essig, D.A., Hicks, M.P., Sturdevant, D.J., Materna, E.J., Spalding, S. A., Risley, J., Deppman, M., 2004. The case for regime-based water quality standards. Bioscience 54, 155–161.
Reeves, G.H., Olson, D.H., Wondzell, S.M., Bisson, P.A., Gordon, S., Miller, S.A., Long, J. W., Furniss, M.J., 2018. The aquatic conservation strategy of the northwest forest plan—A review of the relevant science after 23 years. In: Spies, T.A., Stine, P.A., Gravenmier, R., Long, J.W., Reilly, M.J. (Eds.), Synthesis of science to inform land management within the Northwest Forest Plan area, Gen. Tech. Rep. PNW-GTR-966. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR, pp. 461-624.
Reiter, M., Johnson, S.L., Homyack, J., Jones, J.E., James, P.L., 2020. Summer stream temperature changes following forest harvest in the headwaters of the Trask River watershed, Oregon Coast Range. Ecohydrology 13, e2178.
Richardson, J.S., Danehy, R.J., 2007. A synthesis of the ecology of headwater streams and their riparian zones in temperate forests. Forests 53, 131–147.
Richardson, J.S., Naiman, R.J., Bisson, P.A., 2012. How did fixed-width buffers become standard practice for protecting freshwaters and their riparian areas from forest harvest practices? Freshw. Sci. 31, 232–238.
Richardson, J.S., Naiman, R.J., Swanson, F.J., Hibbs, D.E., 2005. Riparian communities associated with Pacific Northwest headwater streams: assemblages, processes, and uniqueness. JAWRA J. Am. Water Resources Assoc. 41, 935–947.
Russell, K.R., Mabee, T.J., Cole, M.B., 2004. Distribution and habitat of Columbia torrent salamanders at multiple spatial scales in managed forests of northwestern Oregon. J. Wildl. Manage. 68, 405–417.
Ryan, D.F., Calhoun, J.M., 2010. Riparian adaptive management symposium: a conversation between scientists and management. In: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR, USA, p. 135.
Sabo, J.L., Sponseller, R., Dixon, M., Gade, K., Harms, T., Heffernan, J., Jani, A., Katz, G., Soykan, C., Watts, J., Welter, J., 2005. Riparian zones increase regional species richness by harboring different, not more, species. Ecology 85, 56–62.
Saunders, D.A., Hobbs, R.J., Margules, C.R., 1991. Biological consequences of ecosystem fragmentation: a review. Conserv. Biol. 5, 18–32.
Sridhar, V., Sansone, A.L., Lamarche, J., Dubin, T., Lettnemaier, D.P., 2004. Prediction of stream temperature in forested watersheds. J. Am. Water Resources Assoc. 1213–1997.
Stauffer, D.I., Best, L.B., 1980. Habitat selection by birds of riparian communities: evaluating effects of habitat alterations. J. Wildl. Manage. 44, 1–15.
Strahler, A.N., 1957. Quantitative analysis of watershed geomorphology. Trans. Am. Geophys. Union 38, 913–920.
Teti, P.A., Pike, R.G., 2005. Selecting and testing an instrument for surveying stream shade. BC J. Ecosyst. Manage. 6, 1–16.
Thomas, J.W., Maser, C., Rodiek, J.E., 1979. Riparian zones. In: Thomas, J.W. (Ed.), Wildlife habitats in managed forests - the Blue Mountains of Oregon and Washington, U.S. Forest Service Handbook No. 553, Washington, D.C., USA, pp. 40-47.
Tiwari, T., Lundström, J., Kuglerová, L., Laudon, H., Öhman, K., Ågren, A.M., 2016. Cost of riparian buffer zones: A comparison of hydrologically adapted site-specific riparian buffers with traditional fixed widths. Water Resour. Res. 52, 1056–1069.
Veldhuisen, C., Couvelier, D., 2006. Summer temperatures of Skagit basin headwater streams: results of 2001–2003 monitoring. Skagit River System Cooperative, LaConner, WA, USA.
Vesely, D.G., McComb, W.C., 2002. Salamander abundance and amphibian species richness in riparian buffer strips in the Oregon Coast Range. Fors. Sci. 48, 291–297.
Wahbe, T.R., Bunnell, F.L., 2003. Relations among larval tailed frogs, forest harvesting, stream microhabitat, and site parameters in southwestern British Columbia. Can. J. For. Res. 33, 1256–1266.
Washington Department of Ecology, 2015. Extensive riparian status and trends monitoring program-stream temperature, Phase I: Westside Type F/S and Np Monitoring Project, Washington Department of Ecology, CMER Report, Olympia, WA, USA.
Washington Department of Natural Resources. 2005. Forest Practices Rules - Title 222 WAC. www.dnr.wa.gov/forestpractices/rules/.
Wasserstein, R.L., Schirm, A.L., Lazar, N.A., 2019. Moving to a world beyond "$p < 0.05$". Am. Statist. 73, 1–19.
Wilcox, B.A., Murphy, D.D., 1985. Conservation strategy: the effects of fragmentation on extinction. Am. Nat. 125, 879–887.