

Appendix S2. Sampling effort to minimize false absences

Results of logistic regression occupancy models can be biased if used with datasets that include large numbers of “false absences” (i.e., instances when it is incorrectly concluded that a species is absent; Tyre et al. 2003; Comte and Grenouillet 2013). We minimized that bias by ensuring that a sufficient amount of sampling effort was conducted within each CWH to determine occupancy. That effort was estimated using the spreadsheet calculator developed by Peterson and Dunham (2003). The user inputs two values—a species-specific site-level detection efficiency and the prior probability that a sampling unit is occupied—to estimate the probability of a false absence given that a species was not detected (Bayley and Peterson 2001).

Estimates of detection efficiency were derived from a subset of CWHs with ≥ 10 fish samples that were occupied by juvenile native trout. The regional fish database included 118 such CWHs for BT and 181 for CT. Site-level detection efficiency estimates based on those habitats were 0.53 for BT and 0.81 for CT. Estimates of the prior probability of habitat occupancy were derived from simple summaries relating CWH size to the proportion of habitats occupied (Fig. 1). Entering those estimates into the calculator indicated that 2–6 sites needed to be sampled to reliably determine the absence of BT from a CWH while maintaining the false absence rate at ≤ 0.1 . Two samples were needed in $\text{CWH} \leq 10$ km because prior probabilities were ≤ 0.2 ; six samples were needed for $\text{CWH} \geq 50$ km because prior probabilities were ≥ 0.9 . For CT, 2–3 samples met the same false absence rate threshold because detection efficiency was higher for this species. Two samples were needed for $\text{CWH} \leq 3$ km because prior probabilities were ≤ 0.2 ; three samples were needed for $\text{CWH} \geq 8$ km because prior probabilities were ≥ 0.9 . In most cases, the number of fish site samples within the 512 BT CWH and 566 CT CWH used to develop the occupancy models significantly exceeded minimum sample sizes (Fig. 2). For BT, the average number of samples within a CWH was nine (range, 2-81; total sites sampled within BT CWH was 4,608) and for CT, six (range, 2-47; total sites sampled within CT CWH was 3,396).

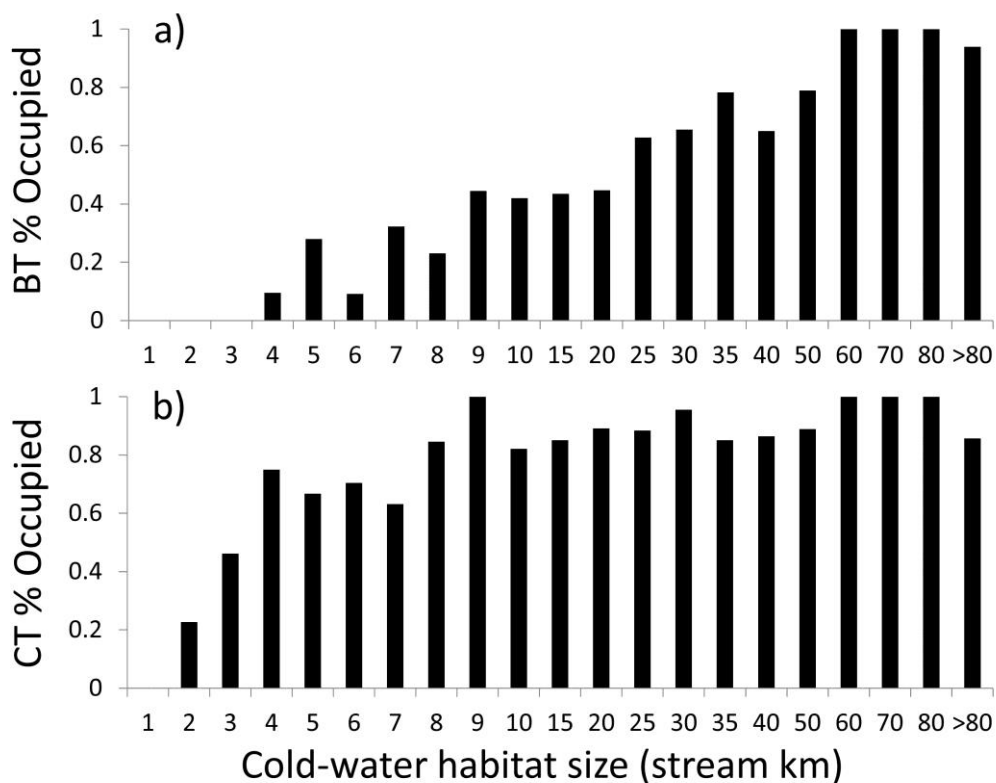


Fig. 1 Juvenile native trout occurrence relative to cold-water habitat size for 512 Bull Trout habitats (a) and 566 Cutthroat Trout habitats (b) that were screened for false absences and used to develop logistic regression occupancy models.

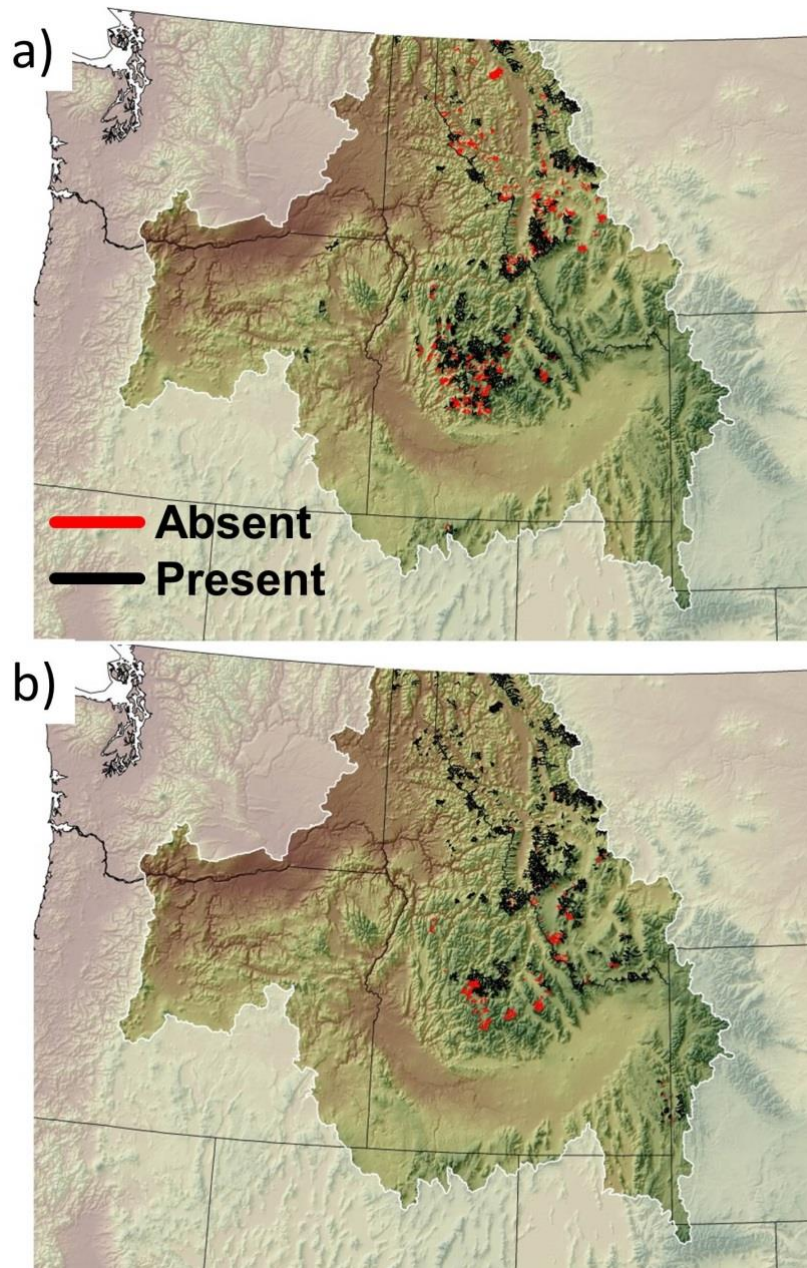


Fig. 2 Locations of cold-water habitats that were screened for false absences and used to develop logistic regression models for predicting occupancy by juvenile Bull Trout (a; $n = 512$) and juvenile Cutthroat Trout (b; $n = 566$).

References

- Bayley PB, Peterson JT (2001) Species presence for zero observations: an approach and an application to estimate probability of presence of fish species and species richness. *Transactions of the American Fisheries Society*, **130**, 620–633.
- Comte L, Grenouillet G (2013) Species distribution modelling and imperfect detection: comparing occupancy versus consensus methods. *Diversity and Distributions*, **19**, 996–1007.
- Peterson JT, Dunham J (2003). Combining inferences from models of capture efficiency, detectability, and suitable habitat to classify landscapes for conservation of threatened bull trout. *Conservation Biology* **17**, 1070–1077.
- Tyre AJ, Tenhumberg B, Field SA et al. (2003) Improving precision and reducing bias in biological surveys: estimating false-negative error rates. *Ecological Applications*, **13**, 1790–1801.