Short communication

How reliable are amphibian population metrics? A response to Kroll et al.

Hartwell H. Welsh Jr. *, Karen L. Pope, Clara A. Wheeler

USDA Forest Service, Pacific Southwest Research Station, Redwood Sciences Laboratory, 1700 Bayview Dr. Arcata, CA 95521, USA

A R T I C L E  I N F O

Article history:
Received 8 May 2009
Accepted 30 May 2009
Available online 26 June 2009

Keywords:
Terrestrial salamanders
Age structure
Body condition
Forest monitoring
Plethodon elongatus
Plethodon stormi

A B S T R A C T

Kroll et al. [Kroll, A.J., Runge, J.P., MacCracken, J.G., 2009. Unreliable amphibian population metrics may obfuscate more than they reveal. Biological Conservation] criticized our recent advocacy for combining readily attainable metrics of population status to gain insight about relationships between terrestrial plethodontid salamanders and forest succession [Welsh, Jr., H.H., Pope, K.L., Wheeler, C.A., 2008. Using readily attainable metrics of population status to gain insight about relationships between terrestrial plethodontid salamanders and forest succession in the US Pacific Northwest. Biological Conservation 141, 1149–1160]. They argue that each of our metrics has the potential to be biased or flawed and, therefore, any inferences so derived are likely to be uncertain and misleading. One of our main goals was to highlight the value of combining multiple quantitative approaches with logic, knowledge of species' biology and parsimony when interpreting findings, as a means to counter uncertainty that might occur with any single metric. We applied a sampling design that minimizes the chance of bias in detections and counts, included multiple statistical methods to test relationships, and incorporated past research and species' biology when interpreting findings. We disagree with Kroll et al. [Kroll, A.J., Runge, J.P., MacCracken, J.G., 2009. Unreliable amphibian population metrics may obfuscate more than they reveal. Biological Conservation] that the potential biases associated with the individual metrics mean that the inferences from the study are too uncertain to be useful by managers. We examined metrics that are more informative than simple site occupancy, including relative abundance, life stage distributions, and body condition, to inform our knowledge of population status. Combining metrics allows us to use readily attainable data to yield deeper insights into population structure and related spatial aspects. We agree on the value of having the most detailed and accurate data possible, but do not believe this standard should limit one to conducting only intensive experimental studies to provide detailed information on vital rates of terrestrial salamanders. The controlled experiment approach is extremely difficult and would be considerably less feasible than using population metrics to assess the status of plethodontid salamanders on forest landscapes at scales relevant to land management in the Pacific Northwest.

Published by Elsevier Ltd.

1. Introduction

In this issue, Kroll et al. (2009) criticizes our recent paper (Welsh et al., 2008), which advocates for the use of a combination of readily attainable metrics of population status to gain insights about the relationships between two terrestrial plethodontid salamanders and forest succession in the Pacific Northwest of North America. They argued that each of our metrics is biased or flawed and, therefore, any inferences that might be derived are likely to be uncertain and misleading (Kroll et al., 2009).

Error, variability, unmet assumptions, and doubts are inherent in ecological field studies. We agree that some or all of these problems may be present to some degree in various aspects of our multiple studies (Welsh et al., 2008). One of our primary goals was to highlight the value of combining multiple quantitative approaches with logic, knowledge of species' biology and parsimony when interpreting findings. We disagree with Kroll et al. (2009) that studies, analysis, and interpretation should not be conducted and reported due to potential uncertainty. Applying such a standard would leave little room for investigations that discern patterns and develop hypotheses. Our goal was not to demonstrate cause-effect relationships, but to advocate for the position that simple site occupancy is insufficient to determine the suitability of environmental conditions (e.g., Kroll et al., 2008). Finding a species at a site or predicting that one should be there is not a demonstration of habitat suitability.

Welsh et al. (2008) proposed looking at metrics beyond occupancy including relative abundance, life stage distributions, and body condition to provide more relevant information on population status. We do not argue that these metrics are the best or the only approaches or that they “provide conclusive evidence” about population dynamics as suggested by Kroll et al. (2009).
Rather, we acknowledge in both our introduction and discussion that optimally “the study of survival and recruitment requires intensive demographic investigations where individuals are tracked throughout their lives” (Welsh et al., 2008). Our position is that using a combination of readily attainable metrics allows one to gain valuable insights into the population status and related spatial aspects of these secretive amphibians, information that is infeasible and impractical to obtain by tracking them throughout their lives at scales applicable to land management.

Perhaps the most fundamental difference of opinion between us and our critics regards the quality of the data. They argue that we used inappropriate sampling programs, the data were collected for different objectives, and that we relied on untested assumptions (Kroll et al., 2009), concluding that the patterns we observed may have too much bias and uncertainty to be worth assessing. We counter that correlative analyses have great value for comparing patterns of population status across broad landscapes and for developing and testing general hypotheses that more rigorous and labor-intensive studies cannot (e.g., Knapp and Matthews, 2000). Although we used prior collected data, our objectives were readily applicable and appropriate. New analyses of previously collected data are a valid and common practice when the original study design allows for new research questions to be addressed with new analyses (e.g., Arneck and Wooster, 1995). Our sampling protocols were based on rigorously tested field methods that incorporated known seasonal surface activities of the salamanders (Ollivier and Welsh, 1999; Clayton et al., in press). We surveyed under specific seasonal and climatic conditions when salamanders were most likely to be surface active, a fact that was verified when possible with reference sites with known populations. Highly trained biologists conducted all the sampling. Stands in different forest age classes were surveyed randomly so variability from diel or environmental conditions would be unlikely to reflect any patterns associated with stand age classes. Surveys were focused on specific within-stand habitat features (i.e., rocky substrates) where salamanders are most likely to occur (Welsh and Lind, 1995; Welsh et al., 2006, 2007b). In addition, an independent test of the protocols of Ollivier and Welsh (1999) found a detection probability of >90% (Welsh et al., 2006). Given the random sampling and similarity of searched substrates, it is difficult to support the Kroll et al. (2009) argument that differences in the detection of animals were due to poor visibility or conditions in younger forest stands rather than the fact that fewer animals were present. We did not apply repeated counts at each site to seek evidence of population increases or declines (Welsh et al., 2007a); rather, we compared relative count data across multiple sites \((n = 311)\) in different forest age classes.

Given that young forests had fewer sites with salamanders, fewer total salamanders, fewer adult salamanders and different body conditions, we discussed possible reasons for these patterns and how they relate to the results of other studies and the biology of these organisms. Kroll et al. (2009) expressed specific concerns about each of our metrics and we comment on these below.

### 2. Occupancy

We acknowledge that we sampled at only a subset of sites \(\geq 2\) times. Our study design was executed prior to the advent of occupancy models and, consequently, our data do not meet all the assumptions of the mathematical models. However, we only sampled when conditions were appropriate for surface activity, a method that greatly increases detectability of salamanders at a site. Using subsets of sites that were re-sampled, we calculated high detection probabilities for both *Plethodon elongatus* \((p = 0.75)\) and *Plethodon stormi* \((p = 0.80)\). Regarding this analysis, Kroll et al. (2009) argued that we violated the assumption that population status remains the same between sampling occasions to estimate occupancy rates because we sampled once a year for 2–3 years. Given that site conditions changed little over the course of our study and given the long-term stability of plethodontid populations (Hairston and Wiley, 1993; Welsh and Droege, 2001; Welsh et al., 2007a), the likelihood that our sampling violated the assumption of “a closed population” in relation to presence or absence of a species at a site is remote.

### 3. Relative abundance

We also acknowledge that we did not conduct a mark–recapture study to estimate individual capture probabilities. We agree that a failure to adjust for differential capture probabilities can result in biased data. For example, Bailey et al. (2004a) explicitly estimated capture probabilities and temporary movement of plethodontid salamanders with mark–recapture methods, and found that numbers of animals varied both spatially and temporally. Spatial differences were related to past forest disturbance with fewer animals found in forests that had been logged ~60 years prior (Hyde and Simons, 2001; Bailey et al., 2004b). The Bailey et al. (2004b) sampling design included regularly timed surveys across a 60–80 day period rather than only when conditions were appropriate for surface activity as we did. This sampling design would be expected to result in high temporal variability in counts. In fact, the original study (Hyde and Simons, 2001) included night surveys conducted when conditions were appropriate for salamander surface activity, a method that resulted in the lowest sampling variability. This variability was apparently too low to have the night sampling approach evaluated in the follow-up study examining differences in site occupancy and detectability by sampling method (Bailey et al., 2004a). Regardless, in our case, we were not attempting to attain population estimates because we were not tracking site-specific population status over time. The relevant question is what level of accuracy is required to detect a meaningful signal in a given parameter (i.e., relative abundance) such that it represents true and accurate differences in numbers in response to differences along an environmental gradient. Our design involved minimizing variability while visiting a large number of sites during appropriate sampling windows (e.g., Salvia et al., 1999; Adams et al., 2005). Our sampling was thorough, with an even effort across treatments (i.e., forest age classes), and we have little concern that variability in encounter rates due to local, unmeasured conditions would influence the overall relative abundance patterns that we found across forest seral stages. We are confident that our large sample sizes captured and accounted for natural site variations and that our conclusions are sound.

### 4. Salamander life stage distributions

Kroll et al. (2009) expressed concerns about using age ratios as estimates of survival. Their concerns, based on Conn et al. (2005), were that critical assumptions are difficult to test and are often unmet and that important parameters such as immigration and emigration were not considered. We agree that using age ratios is a simplified method for assessing survival and do not claim that our analyses provided “conclusive evidence” about population equilibrium, as suggested by Kroll et al. (2009). Nonetheless, we also agree with Ricklefs and Rohwer (2005) that age ratios revealed from count data can be very useful in comparative studies to estimate average adult survival in broadly sampled populations. As discussed (Welsh et al., 2008, pp. 1153), the potential biases associated with growing or declining populations are unlikely to be strong, although we admit we could not test this assumption. However, it has been demonstrated with time-series data from
35 different studies that variations in counts of plethodontid salamanders are highly stable, showing less variability than lepidop- tera, passerine birds, small mammals, and other amphibians (Welsh and Droege, 2001).

Kroll et al. (2009) claimed that we failed to consider the alternative hypothesis that the high proportion of young salamanders in young forest stages may be due to in situ recruitment; therefore, they concluded that this lack of “consideration of such alternate hypotheses shows that without valid estimates of reproduction, survival, and dispersal processes, it is impossible to gauge source-sink dynamics”. We discussed this alternative hypothesis (Welsh et al., 2008, pp. 1156–1157) and determined that it is unlikely given that overall salamander abundances on these sites were low. The only logical way this alternative hypothesis would make sense would be if salamander populations in all the young stands were undergoing rapid population growth at the time of sampling. Given that the young forest stands ranged from 30 and 99 years and were in different recovery states, we deemed it unlikely that eight different sites would all be experiencing a similar rapid population growth post-disturbance; dispersal simply made more sense.

Kroll et al. (2009) also contended that we should have addressed how a potential imbalance in immigration and emigration could bias our estimates of survival. They cite Conn et al. (2005) as demonstrating that an imbalance can significantly bias estimates of survival. They do not, however, refer to Ricklefs and Rohwer (2005) who showed that Conn et al. (2005) used an equation that inappropriately confounded population change with immigration and emigration. Ricklefs and Rohwer (2005) concluded that the uncertainty analysis conducted by Conn et al. (2005) overstated sources of variation in estimates of annual survival from age ratios. We agree that it is difficult to discern between despotic dispersal and differential survival and have admitted to not being able to discern between the two to explain why young forest stands had higher proportions of young salamanders (Welsh et al., 2008, pp. 1157). Furthermore, we see no reason why both processes would not be at play as young forests recover the conditions required to support similar populations of salamanders as unaltered sites (Petranka, 1999). Given the patchwork nature of managed forest landscapes, despotic interactions between adults and subordinates would likely create differences in age ratios at scales within the relatively low dispersal capabilities of plethodontid salamanders. Whereas at larger scales beyond the typical dispersal distances of these highly sedentary animals (e.g., Karraker and Welsh, 2006), differential survival better explains the differences in age ratios we found.

5. Body condition

We based our choice of body condition metrics on Schulte-Hostedde et al. (2005) who empirically tested the use of residuals as indices of condition. Schulte-Hostedde et al. (2005) determined that using residuals from ordinary least squares (OLS) regression as indices of body condition is preferred over major axis (MA) or reduced major axis (RMA) regressions. They found that “residuals from the body mass-body size OLS regressions consistently explained significant amounts of variation in fat, water, and lean dry mass” (Schulte-Hostedde et al., 2005). Given their demonstration of the superiority of this approach compared to others, we followed their recommendation.

In addition, Kroll et al. (2009) state that we incorrectly “log-transformed mass and length data without any assessment of need or the success of the transformation”. In allometric equations of the general form \( y = \beta x^b \), log transforming both sides of the equation is a general practice to derive a linear form where \( \log y = \log \beta + b \log x \), with \( b \) representing the slope of the line. Numerous relationships between biological variables and body size conform to this general equation (Schmidt-Nielsen, 1995). In this case, we were not transforming normally distributed data as suggested by Kroll et al. (2009), but were plotting an allometric equation on logarithmic coordinates to achieve a straight line.

6. Assessing the impacts of forest management

Kroll et al. (2009) provided several suggestions for monitoring programs, the first being that they should be designed so that unbiased estimates of occupancy and abundance can be calculated. We agree with this contention. However, we believe that past studies not designed for estimating detection probabilities can also provide useful ecological insights and should not be dismissed.

Kroll et al. (2009) suggested using mark-recapture or ratio estimation (given that plethodontid salamanders are cryptic) to estimate population sizes. They also suggested that it is necessary to demonstrate the accuracy of a sampling method where capture probabilities are <1. We are unaware of a logistically feasible and cost-effective technique for sampling terrestrial salamander populations at a relevant scale where capture probabilities might approach 1 given that ~2/3 of any population is likely underground at any time (e.g., Taub, 1961). Because of these fossorial habits, the resulting low recapture probabilities, and often depressed numbers on altered landscapes, Bailey et al. (2004b) considered precise populations estimates nearly impossible to achieve. Their intensive population estimates, however, did support the findings of Hyde and Simons (2001) based on encounter rates and relative abundance that more plethodontid salamanders occurred in older forest seral stages (Bailey et al., 2004b). Similar to most other statistical techniques for estimating population sizes, ratio estimation and mark-recapture are also known to have biases such as with small sample sizes or low capture probabilities (e.g., Menkens and Anderson, 1988; Bailey et al., 2004b). Regardless of sampling intensity, all population estimation techniques must be used with care given potential biases, uncertainties, and limitations inherent in these data.

Kroll et al. (2009) claimed that unmanaged old growth forest is not a true experimental control for assessing forestry impacts. This is true for experiments designed to determine the effect sizes of specific forestry actions on managed landscapes, but this was not the goal of our study. We referred to our old growth stands as “reference sites” instead of “controls” for this reason (Welsh et al., 2008). Regardless, we submit that to understand the overall effects of timber harvesting, old growth reference forests are necessary, otherwise impacts will likely be understated as a result of the shifting baseline syndrome (Pauly, 1995). To understand the relationships between plethodontid populations and forest succession, one needs to study the range of post-forest and forested sites, including un-entered old-growth reference sites; only the later comprise the highest or oldest end of the environmental gradient of forest succession. This is exemplified by the positive patterns of increased abundances and greater species richness with advanced succession evident in studies across North American forests (e.g., deMaynadier and Hunter, 1995; Wyman, 2003; Davic and Welsh, 2004). A comprehensive approach to studying succession is critical to understand and evaluate historical influences on forest ecosystems (Foster et al., 2003; Willis and Birks, 2006). This is also essential if we hope to manage these ecosystems sustainably in the future (Rhemtulla and Mladenoff, 2007). Downes et al. (2002, pp. 41) specifically warns about the value of data being limited by the lack or poor quality of control localities.

Kroll et al. (2009) claimed that without vital rates for potentially impacted species, managers “are placed in the unenviable
scenarios of having to act without adequate knowledge about critical population characteristics. There is a large body of literature on the biology and ecology of plethodontid salamanders and their relationships with forest ecosystems (e.g., Davic and Welsh, 2004; Bruce, 2008; and citations therein). Without lungs, plethodontid salamanders require moist environments to facilitate cutaneous respiration (Spight, 1968), making them highly sensitive to changes in temperature and moisture (Grover and Wilbur, 2002). Microclimatic soil conditions within disturbed forests with reduced canopy cover often exhibit warmer temperatures and lower moisture and humidity relative to undisturbed forests (Chen et al., 1999). From the available literature, land managers can easily gain insight as to how salamander populations may respond to different timber harvest treatments. For example, we believe that there is sufficient research and knowledge about these species (even accounting for uncertainties) to conclude that clear-cutting is detrimental to their populations regardless of what data on vital rates are available. We agree it would be nice to have more detailed and accurate data, but, given real-world conditions, this situation is unlikely and certainly not necessary to make sound decisions to protect salamander populations.

There is always a tradeoff between level of effort and sample size, thus, there is no single “best” approach to all questions. Van Horne (2002) related this idea to purchasing a car. Customers weigh tradeoffs between size, gas mileage, stability, and precision of handling according to their larger objectives such as minimizing cost or maximizing comfort and safety. Similar to Bailey et al. (2004b), experience has shown us (e.g., Welsh and Lind, 1992) that the level of effort to obtain enough capture–mark–recapture data on these secretive, fossorial amphibians to develop accurate population models at a scale and across habitats relevant to land managers is simply not worth the labor and expense. However, this certainly does not negate their value as useful metrics (e.g., Welsh and Droge, 2001; Highton, 2005). We are encouraged by the Kroll et al. (2009) support of labor and cost intensive projects to conduct in-depth plethodontid population-level studies across different forest management conditions. We hope that they take the next step to initiate such ambitious studies. A combination of extensive, coarse sampling with more intensive local population monitoring can elucidate scale, pattern and mechanisms associated with important changes across landscapes.

Wells and Droge (2001) noted that low variation in repeated counts of these salamanders compared to many other well studied animal taxa, including other amphibians, gave them a statistical advantage for monitoring forest health. Davic and Welsh (2004) reviewed the ecological role of salamanders, with a focus on plethodontids, noting their enormous biomass and extremely high abundances in temperate forest ecosystems compared to other small vertebrates (e.g., Burton and Likens, 1975a; Peterman et al., 2008) supported the view that they are keystone species (Paine, 1969; Davic, 2003). Their large numbers and high assimilation efficiencies make them an important protein-rich food source for predators (Burton and Likens, 1975b). Recent evidence suggests a significant positive role in the carbon cycle in these forests as a result of their impact on invertebrate shredders (Wyman, 1998). The status of planetary forest ecosystems is no longer simply an academic question; these forests may prove to be vital in helping to address the climate crisis now facing humanity. Old-growth forests accumulate carbon in both trees and soil (Zhou et al., 2006), and there is compelling evidence that old-growth forests have the capacity to sequester the most carbon (Luysaert et al., 2008; Hufriedt et al., 2009). It follows then that understanding forest ecosystem dynamics is critical, and it is essential that we develop and test meaningful biometrics that track and monitor forest health and ecosystem status (e.g., Welsh and Droge, 2001; Nichols and Williams, 2006).

References


