esa ECOSPHERE

SPECIAL FEATURE: UNCERTAINTY ANALYSIS

Uncertainty analysis: an evaluation metric for synthesis science

Mark E. Harmon,[1,†] Becky Fasth,[1] Charles B. Halpern,[2] and James A. Lutz[3]

[1]*Department of Forest Ecosystems and Society, Oregon State University, Corvallis, Oregon 97331 USA*
[2]*School of Environmental and Forest Sciences, College of the Environment, Box 352100, University of Washington, Seattle, Washington 98195 USA*
[3]*Wildland Resources Department, Utah State University, Logan, Utah 84322-5230 USA*

Citation: Harmon, M. E., B. Fasth, C. B. Halpern, and J. A. Lutz. 2015. Uncertainty analysis: an evaluation metric for synthesis science. Ecosphere 6(4):63. http://dx.doi.org/10.1890/ES14-00235.1

Abstract. The methods for conducting reductionist ecological science are well known and widely used. In contrast, those used in the synthesis of ecological science (i.e., synthesis science) are still being developed, vary widely, and often lack the rigor of reductionist approaches. This is unfortunate because the synthesis of ecological parts into a greater whole is critical to understanding many of the environmental challenges faced by society. To help address this imbalance in approaches, we examine how the rigor of ecological synthesis science might be increased by using uncertainty as an evaluation metric—as a parallel to methods used in reductionist science. To estimate and understand uncertainty we propose that it be divided into four general classes: (1) measurement uncertainty (i.e., experimental error) as defined by precision and accuracy, (2) sampling uncertainty that reflects natural variation in space and time as quantified by classical statistical moments (e.g., mean and variance), (3) model prediction uncertainty which relates to the transformation of measurements into other variables of interest (e.g., plant dimensions to biomass), and (4) model selection uncertainty which relates to uncertainty about the form of the relationships used in models. Of these sources of uncertainty, model selection is the least understood and potentially, the most important, because it is integral to how components of a system are combined and it reflects imperfect knowledge about these relationships. To demonstrate uncertainty in synthesis science, we examine each source of uncertainty in an analysis that estimates the live tree biomass of a forest and how knowledge of each source can improve future estimates. By quantifying sources of uncertainty in synthesis science, it should be possible to make rigorous comparisons among results, to judge whether they differ within the bounds of measurement and knowledge, and to assess the degree to which scientific progress is being made. However, to be accepted as a standard method, best practices analogous to those used in reductionist science need to be developed and implemented.

Key words: aboveground biomass; H. J. Andrews Experimental Forest; measurement uncertainty; model parameter uncertainty; model selection uncertainty; Oregon, USA; sampling uncertainty; Special Feature: Uncertainty Analysis.

Received 15 July 2014; revised 4 December 2014; accepted 18 December 2014; final version received 16 March 2015; **published** 22 April 2015. Corresponding Editor: J. R. Taylor.

Copyright: © 2015 Harmon et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. http://creativecommons.org/licenses/by/3.0/

† **E-mail:** mark.harmon@oregonstate.edu

Introduction

Today, ecological studies are proceeding on two complementary tracks: the traditional use of reductionist science and a more recent approach, which we term synthesis science. Each has its strengths and weaknesses, hence the value of pursuing both simultaneously. The reductionist

approach is well known and practiced throughout most, if not all, fields of science (Gallagher and Appenzeller 1999). It entails simplifying a system, controlling confounding factors to isolate the essential parts and mechanisms, and conducting controlled experiments. This approach works well when a system is inherently simple and when the overall structure is additive (i.e., the whole is the sum of the parts). However, ecological systems rarely comprise simple, additive structures, and this has led to attempts to synthesize the parts in ways that retain non-additive interactions and inherent complexity (Odum 1977, Holling 2001). Examples of synthetic science include ecosystem budgets (Sollins et al. 1980), simulation modeling (Shugart 1998), and analysis of biocomplexity (Michener et al. 2001), each of which embraces—rather than eliminates—complex interactions, multiple controls, and confounding factors. Despite the need for synthesis science in the field of ecology, compared to reductionist science there is not a well-established methodology for achieving it. This may stem, in part, from the more recent advent of synthesis science and from the existence of multiple, viable approaches. The latter may pose a barrier to standardization in synthesis science. Synthesis science also lags behind reductionist science in its limited use of evaluation metrics. Evaluation of reductionist science relies on well-established criteria for experimental design and a well-established set of statistical methods. Comparable sets of evaluation metrics are lacking in synthesis science. Here we describe how uncertainty analysis, if applied rigorously, could serve in this role. For example, it would facilitate rigorous comparison of multiple estimates resulting from synthesis and whether they differ—similar to more traditional statistical tests.

Although quantitative uncertainty analysis has been incorporated into past synthesis efforts (e.g., Harmon et al. 2004, Yanai et al. 2010), it has not, unfortunately, been a standard practice. This may reflect the complexity of synthesis science, limitations in classical analytical error- propagation methods, and lack of computational power. As recent advances reduce these constraints, opportunities exist to add rigor to synthesis science. Specifically, while the complexity of synthesis efforts is unlikely to diminish, greater understanding of the sources of uncertainty can aid in interpreting this complexity (Harmon et al. 2007). Furthermore, increased computing power permits use of Monte Carlo approaches to estimate uncertainty when traditional analytical methods are too challenging.

The objective of this paper is to provide a general framework for considering uncertainty as an evaluation metric in synthesis science. We illustrate this concept using an example that estimates live tree biomass over a 30-year period following clear-cutting of a forested watershed. We conclude with a set of general thoughts on future challenges to creating an uncertainty-based evaluation metric for synthetic sciences in ecology. Although we emphasize uncertainty as an evaluation metric, it can play a broader role in risk analysis (Bartell et al. 1992), decision support (Kangas 2010), methods evaluation (Lauenroth et al. 2006), and ecological modeling (Li and Wu 2006).

General Sources of Uncertainty in Ecological Sciences

Among the many sources of uncertainty that can be identified, we suggest there are four general classes: (1) measurement, (2) sampling, (3) model prediction, and (4) model selection. Of these uncertainties, those related to measurements and sampling are the best understood and model selection, the least. Model prediction uncertainty may be as well understood as measurement and sampling uncertainty, but it is often not quantified and rarely considered outside of simulation modeling. Below we describe each of these general classes in more detail.

Measurement uncertainty (usually termed measurement or experimental error) reflects the limitations of the instrument used to make a measurement, of those using the instrument or, in some cases, of the situations in which the instrument is used. There are two general components of measurement uncertainty: accuracy and precision. Accuracy is the degree that a measurement matches the actual value; precision quantifies the degree that measurements are repeatable (i.e., the variation in measuring the same object repeatedly). Accuracy and precision of the measurement instrument is generally

available (particularly for instruments that measure aspects of climate or chemistry). How accuracy and precision are influenced by users interacting with an instrument, is less well understood, because it requires additional effort (evaluating multiple users). It is more tractable when personnel do not change, but it can be difficult when they do. Additionally, accuracy and precision can vary for the same instrument and user. For example, presence of ice can affect precipitation and hydrologic measurements, thick growth of bryophytes on tree boles can influence diameter measurements, and high concentrations of some elements in water can interfere with measurements of others. Although many of these situations are well known and addressed by seasoned practitioners, they are not routinely quantified.

Sampling uncertainty is probably the best understood and quantified aspect of uncertainty in ecology. It reflects the natural variation in a variable, either in space or time. Quantifying and using sampling uncertainty is the basis of classical statistics; it is widely taught and practiced. Moreover, it is almost impossible to publish an ecological analysis (including simulation modeling) in which this aspect of uncertainty is not addressed in some way. An interesting facet of sampling uncertainty is that while there are more efficient ways to characterize it, this aspect of uncertainty cannot be reduced to zero unless the system of interest has no variation in space and time. Sampling uncertainty is also scale dependent which makes it difficult to compare ecological data collected in different ways. In general, for a given population, sampling uncertainty is reduced as the spatial and temporal extent of measurements increase. For example, estimates of tree mortality in smaller (e.g., 0.05 ha) plots will be inherently more variable than those in larger plots (e.g., >2 ha). Similarly, estimates of mortality conducted annually will be inherently more variable than those spanning a decade. Experienced practitioners generally understand the sampling designs and scales that provide useful results; however, these are not always documented. Moreover, given the wide variation in spatial structure and temporal dynamics of ecosystems (e.g., algal versus forest systems), it may not be possible to standardize all sampling designs.

There are two general aspects of model-related uncertainty that can be quantified: prediction and selection. There are merits to separating them, but we acknowledge that there is overlap that cannot be ignored. "Model" refers to any calculation that transforms the original, measured quantity that involves uncertainty. It can be a simple conversion of one quantity to another using one parameter (e.g., conversion of organic matter to carbon) or a complex calculation that involves many parameters and equations. Thus, this general class of uncertainty includes conversion and regression uncertainty (Phillips et al. 2000, Harmon et al. 2007), but could also involve complex simulation models. Model prediction uncertainty arises when an individual deviates from the mean estimate for a fixed value of the predictor, or independent, variable. For some transformations, however, there is no prediction uncertainty. For example, the calculation of basal area involves two parameters (i.e., the exponent 2 and Π), but their values are either constant or known to a very large number of decimal places. In contrast, a model that converts stem diameter to biomass has several parameters that have to be estimated empirically, hence the associated uncertainty. This form of uncertainty can be addressed in at least two ways. The first involves quantifying model residual uncertainty (e.g., the mean square error of the model). The second involves model parameter uncertainty, quantifying the effect of not knowing the exact value of a model parameter. Although model prediction uncertainty is influenced by sampling uncertainty (as mean square error and variability of parameter estimates typically reflect sampling), it differs from the latter because it involves application, not development and evaluation, of the model. Unfortunately, when models are reported, there is a tendency to document the overall goodness of fit in the form of the coefficient of determination and parameter estimates. However, one needs either the mean square error of the model or the standard errors of the individual parameters and their correlations, to incorporate model prediction uncertainty in an uncertainty analysis (but see our example analysis of uncertainty, below, for an approximation based on the coefficient of determination).

Model selection error is the least understood

class of uncertainty and involves either choosing alternative model structures or models with very different parameter values (Rowe 1994, Draper 1995, Lindenschmidt et al. 2007, Melson et al. 2011). Model selection involves knowledge uncertainty or, as Ferson and Ginzburg (1996) put it, ignorance. Simply stated, if there is limited information to guide model selection, then those models that cannot be eliminated should be considered. Unlike other forms of uncertainty estimated with standard analytical or Monte Carlo approaches, there is no random component to model selection uncertainty because the errors are exclusively systematic. This makes model selection error difficult to estimate or characterize. Although, in a general sense, the entire scientific enterprise seeks to eliminate model selection error, this form of uncertainty is usually hidden in synthesis science. Typically, models are selected based on perceived realism, adequacy, convenience, and familiarity. Reliance of model selection on expert opinion can be counterproductive because it can hide a major form of uncertainty, reducing the motivation to understand and reduce it.

There are other forms of uncertainty, as well. For example, there can be uncertainty associated with data entry and "version control" (i.e., whether the correct or most recent data or model are used). In addition, we assume that our calculations are correct, but even after verification, some errors may not be caught or recorded. Finally, we can combine these forms of uncertainty to estimate "total" uncertainty. Because it is difficult, if not impossible, to identify all forms of uncertainty (thus our use of quotes), we use the term "overall" uncertainty to represent the combined set of sources considered in an analysis.

Andrews Forest Watershed 1: An Example

We use estimates of live, aboveground tree biomass in a 100-ha watershed (WS01) within the H. J. Andrews Experimental Forest, Oregon as an example of how the four general classes of uncertainty can be estimated, combined, evaluated, and used. This assessment is part of a broader effort to quantify the carbon budget of this gauged watershed. WS01 contained an old-growth Douglas-fir/western hemlock forest that was clear-cut logged between 1962 and 1966 (Halpern and Franklin 1990). Post-harvest regeneration was very uneven despite multiple attempts at seeding and planting of Douglas-fir, resulting in large variation in the spatial distribution of biomass accumulation (Lutz and Halpern 2006, Halpern and Lutz 2013). Within the watershed, a total of 138, 0.025 ha plots was systematically arrayed along 6 widely-spaced transects. The plots were measured at 3- to 6-year intervals between 1980 and 2007 (a total of 7 measurements). At each measurement, permanently tagged trees (>1.4 m tall) were assessed for species and status (live, dead, missing, or ingrowth) and measured for diameter either at the ground surface (DAG, for smaller trees) or at breast height (DBH, for larger trees). Live, aboveground biomass was estimated from diameter (or for some species, diameter and estimated height).

To assess measurement, sampling, and model parameter uncertainty, we used biomass equations from Biopak (Means et al. 1994). Each of these forms of uncertainty was estimated using Monte Carlo methods and expressed as two standard errors of the mean (i.e., the 95% confidence bounds). We used the standard error of the mean because we are interested in the uncertainty of the overall estimate for the watershed, not of the individual measurements, trees, or plots.

To assess model selection error, we considered in addition to the biomass equations from Biopak, two other sets of equations: (1) those from Lutz and Halpern (2006) (henceforth, Lutz equations), adjusted to total aboveground biomass using the ratio of branches and leaves to boles (derived from Biopak); and (2) the national set of equations from Jenkins et al. (2003). Uncertainty associated with model selection was represented as the difference between the minimum and maximum estimates of these equations. Additional details relevant to each aspect of uncertainty are described below.

Measurement uncertainty was based on the average variation in tree diameter measured by multiple, experienced crew members. Repeated measurements of the same trees indicated that crew members measured diameters with a precision of 2% as represented by two standard

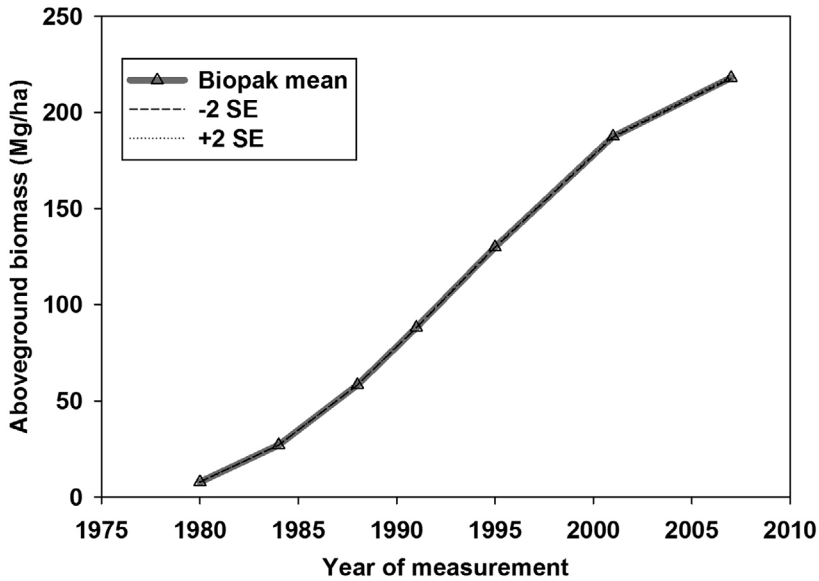

Fig. 1. Estimate of measurement uncertainty for live tree biomass in WS01 in the H. J. Andrews Experimental Forest, Oregon. Note that at this scale, the bounds indicated by two standard errors (SE) are too small to be visible.

deviations of the mean (Harmon et al. 2007), reflecting variation in the placement of the diameter tape and measurement technique of each crew member. We did not account for measurement accuracy because the diameter tapes are usually accurate to within 1 mm—the lowest unit of measurement recorded. To estimate measurement uncertainty, 10 of the 138 plots were randomly selected and the diameter of each tree in each plot was randomly varied using a distribution with a standard error of 1%. A total of 3,000 iterations was used to compute the mean and standard deviation of biomass at the time of measurement using the Biopak equations. This analysis indicated that although measurement uncertainty was as high as 2% per tree, it was only 0.09% for estimated biomass when all trees were considered (Fig. 1). This large reduction in measurement error reflects the offsetting effect of one random error by another, a general pattern with measurement error (Phillips et al. 2000).

Sampling uncertainty was calculated by "setting" all other sources of uncertainty to zero. That is, we assumed no uncertainty in the diameter measurements, model parameters, and model selection (this is usually assumed tacitly); here, too, we used the Biopak equations. Sampling uncertainty thus reflects only spatial variation among the 138 plots. In absolute terms, sampling uncertainty was relatively constant (13–14 Mg/ha over the measurement period; Fig. 2). However, in relative terms (as a percentage of total biomass), it declined from 50% in 1980 to 4% in 2007. The initially high value likely relates to substantial spatial variation in the establishment of trees (Lutz and Halpern 2006), when total biomass was low. The relative variability declined over time because the absolute variability remained constant as the average biomass increased 16-fold.

Model parameter uncertainty could only be approximated: as is too often the case, we lacked documentation of the uncertainty of the biomass-equation parameters and the mean square error of the models. Because indices of parameter uncertainty were not available, we used the coefficient of determination as a guide to the range of variation possible in the parameter estimates. Lacking knowledge of the correlation among biomass model parameters, we assumed no correlation among parameters and varied all parameters simultaneously the same relative amount until the variation in biomass estimates for hypothetical trees was consistent with the coefficient of determination reported. While the estimate of parameter variation differed some-

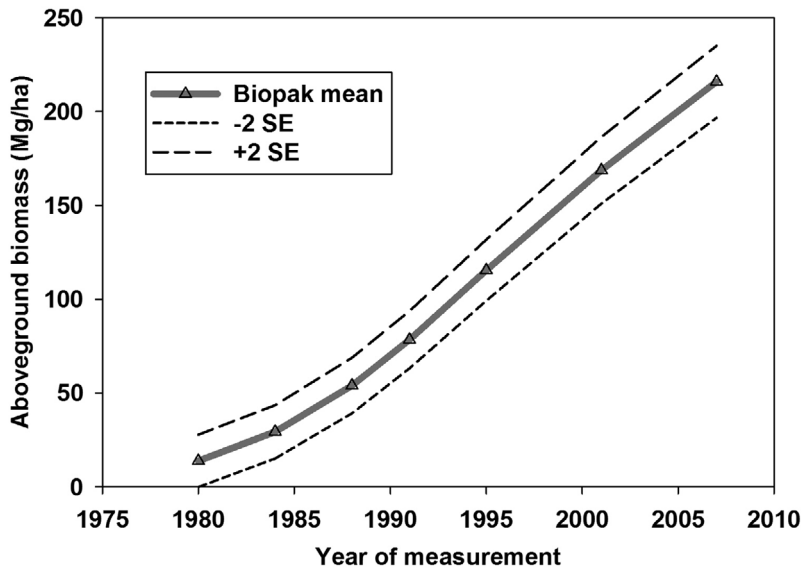

Fig. 2. Estimate of sampling uncertainty indicated by two standard errors (SE) for live tree biomass in WS01 in the H. J. Andrews Experimental Forest, Oregon. All other sources of uncertainty were set to zero.

what among tree species, ≈5% variation in the equation parameters produced a level of variation consistent with the coefficients of determination. For height-diameter equations, we used the parameter standard errors provided by Garman et al. (1995). For the same 10 plots used to assess measurement uncertainty, we estimated the biomass of all trees, 3,000 times. This analysis indicated that model parameter uncertainty increased in absolute terms from ≈0 in 1980 to 8 Mg/ha in 2007 (Fig. 3). In contrast to sampling uncertainty, the relative expression of model parameter uncertainty remained fairly constant among measurement periods, at ≈1.5%. Similar to measurement uncertainty, however, there was a substantial reduction in uncertainty of the total biomass estimates, due to the counterbalancing effects of trees with lower- and higher-than-average biomass for a given diameter.

Model selection uncertainty was estimated by comparing models with the lowest and highest live biomass estimates (Fig. 4), with the other sources of uncertainty set to zero. Estimates based on equations from Jenkins et al. (2003) and Biopak were similar (213 and 216 Mg/ha, respectively, in 2007) and both were higher than the estimate of Lutz and Halpern (2006) adjusted for non-bole components (192 Mg/ha in 2007), resulting in an average model selection uncertainty of 12% of the median value of 204 Mg/ha in 2007. As with model parameter uncertainty, model selection uncertainty increased in absolute terms over time, largely due to the cumulative nature of biomass. To determine if the Lutz and Jenkins models differed from those of Biopak, we used a series of modified t-tests. Specifically, for each measurement period we generated a ratio, with the numerator computed as the difference between the mean aboveground biomass for Biopak and the model in question (Lutz or Jenkins), and the denominator computed as the combined uncertainty in measurements and parameters for Biopak. Given the large sample size (3,000), we compared the ratios to the critical value of t when the degrees of freedom were infinite and $p = 0.05$ ($t = 1.96$). These tests indicated that the sets of Biopak and Lutz models differed "significantly" over the entire sampling period (1980 to 2007), with the difference-to-uncertainty ratio ranging from 5.71 to 7.50. Thus, there is no basis for assuming that the models are interchangeable. In contrast, the Biopak and Jenkins equations differed "significantly" only in 1980 (ratio of 2.27), indicating that for most measurement times, either could be used.

To compute the "overall" uncertainty of biomass estimates (i.e., the combination of measurement, sampling, model parameter, and

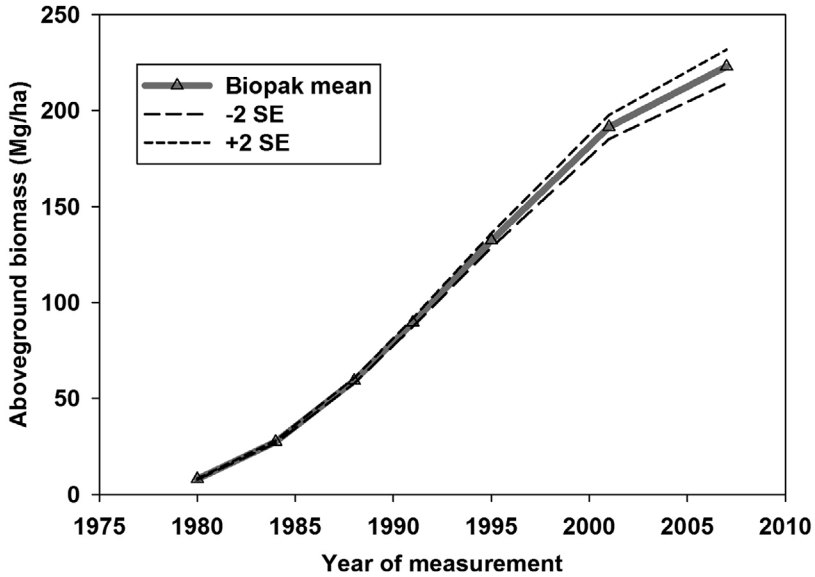

Fig. 3. Estimate of model parameter uncertainty indicated by two standard errors (SE) for live tree biomass in WS01 in the H. J. Andrews Experimental Forest, Oregon. All other sources of uncertainty were set to zero.

model selection uncertainty) we used the equations yielding the lowest (Lutz and Halpern 2006) and highest estimates (Biopak) to estimate lower and upper bounds. For this analysis we used classic error-propagation methods, i.e., summing component variation and calculating the square root. These methods assumed that (1) the relative measurement and model parameter uncertainties in the 10 plots were representative of the entire sample (138 plots), and (2) all

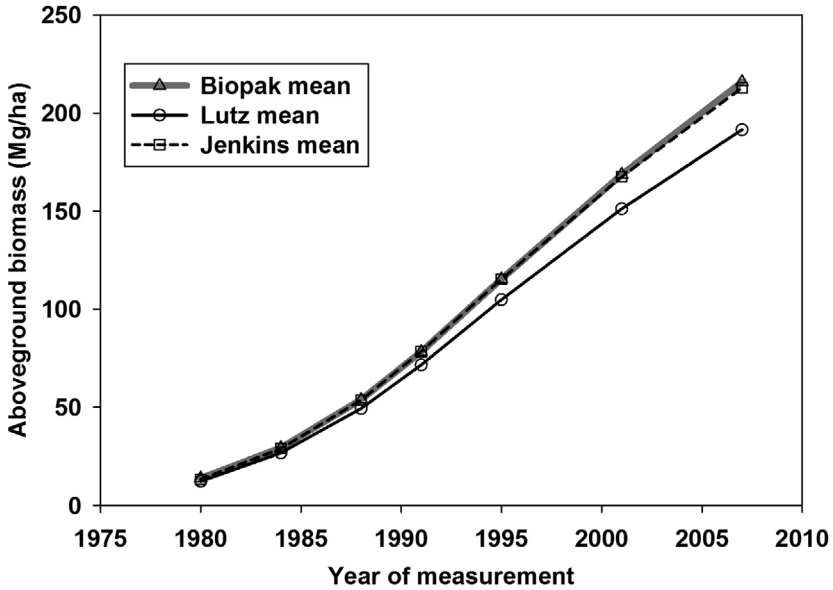

Fig. 4. Estimate of model selection uncertainty for live tree biomass in WS01 in the H. J. Andrews Experimental Forest, Oregon. All other sources of uncertainty were set to zero.

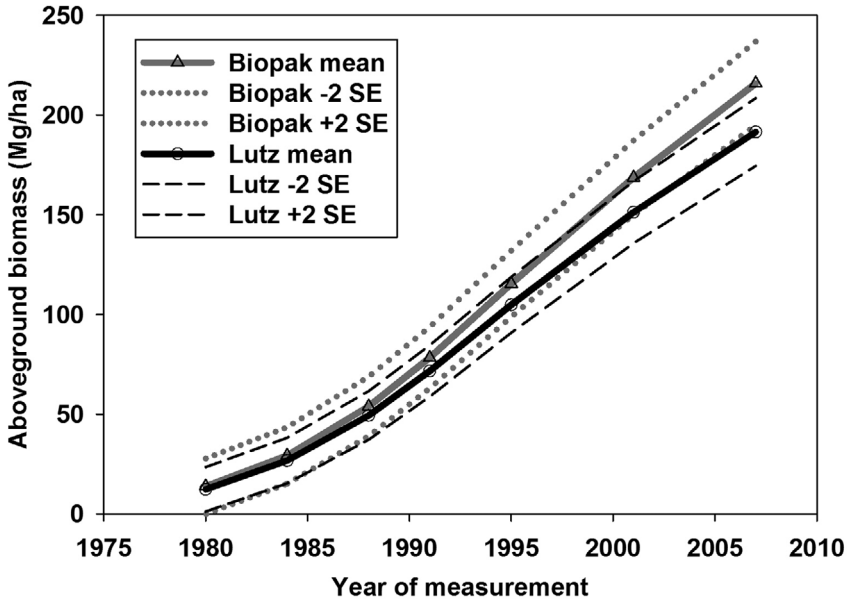

Fig. 5. Estimate of overall uncertainty for live tree biomass in WS01 in the H. J. Andrews Experimental Forest, Oregon. For each model (Biopak and Lutz) the bounds indicated by two standard errors (SE) included measurement, sampling, and model parameter uncertainty.

sources of uncertainty were uncorrelated. There was considerable overlap in the predictions of these equations (Fig. 5), with estimates in 2007 ranging from 175 to 235 Mg/ha or 15% of the average. Overall uncertainty also varied over time.

To assess the relative contributions of each of the four major forms of uncertainty to the overall uncertainty, we used the "variance", not the standard error, as the metric. (When uncertainties are not positively correlated, e.g., random, the combined standard error is lower than the arithmetic sum of the standard errors). For model selection uncertainty, we assumed that the difference between the lowest and highest model estimates approximated that of 4 standard errors. We recognize the limits to this approach; however, it is one way to establish a common basis for assessing uncertainty. Among them, sampling uncertainty was greatest in 1980 when it was >90% of the overall uncertainty "variance" (Fig. 6). By 2007, it still accounted for the largest share of the total (62–64% depending on the set of biomass models considered). Most of the remaining uncertainty was attributable to model selection and, to a lesser extent, model parameter uncertainty (e.g., 25–35% vs. 2–12%, respectively, of the overall uncertainty "variance" in 2007).

There are several ways in which these estimates of uncertainty are useful. First, they permit comparisons with other estimates of live, aboveground biomass accumulation. For example, in our system, we are interested in whether the aboveground biomass in the relatively young forests of WS01 is lower than that estimated in a nearby, paired old-growth watershed (WS02) (Acker et al. 2002). The average aboveground biomass in WS02 was estimated at ≈590 Mg/ha with a standard error of ≈32 Mg/ha (which captures only sampling uncertainty). The ratio of the difference between watersheds (385 Mg/ha) to the overall uncertainty estimated for WS01 (60 Mg/ha)—somewhat analogous to a t-test based on the range (Lord 1947)—is 6.4, a highly significant result based on the critical t-statistic of 0.126 for 20 degrees of freedom, the largest degrees of freedom published for this statistic. Although the same conclusion could have been reached without an evaluation using uncertainty analysis, the analysis would be more critical had the biomass estimates been more similar between watersheds.

Second, it is possible to use this information to develop a strategy for reducing uncertainty in

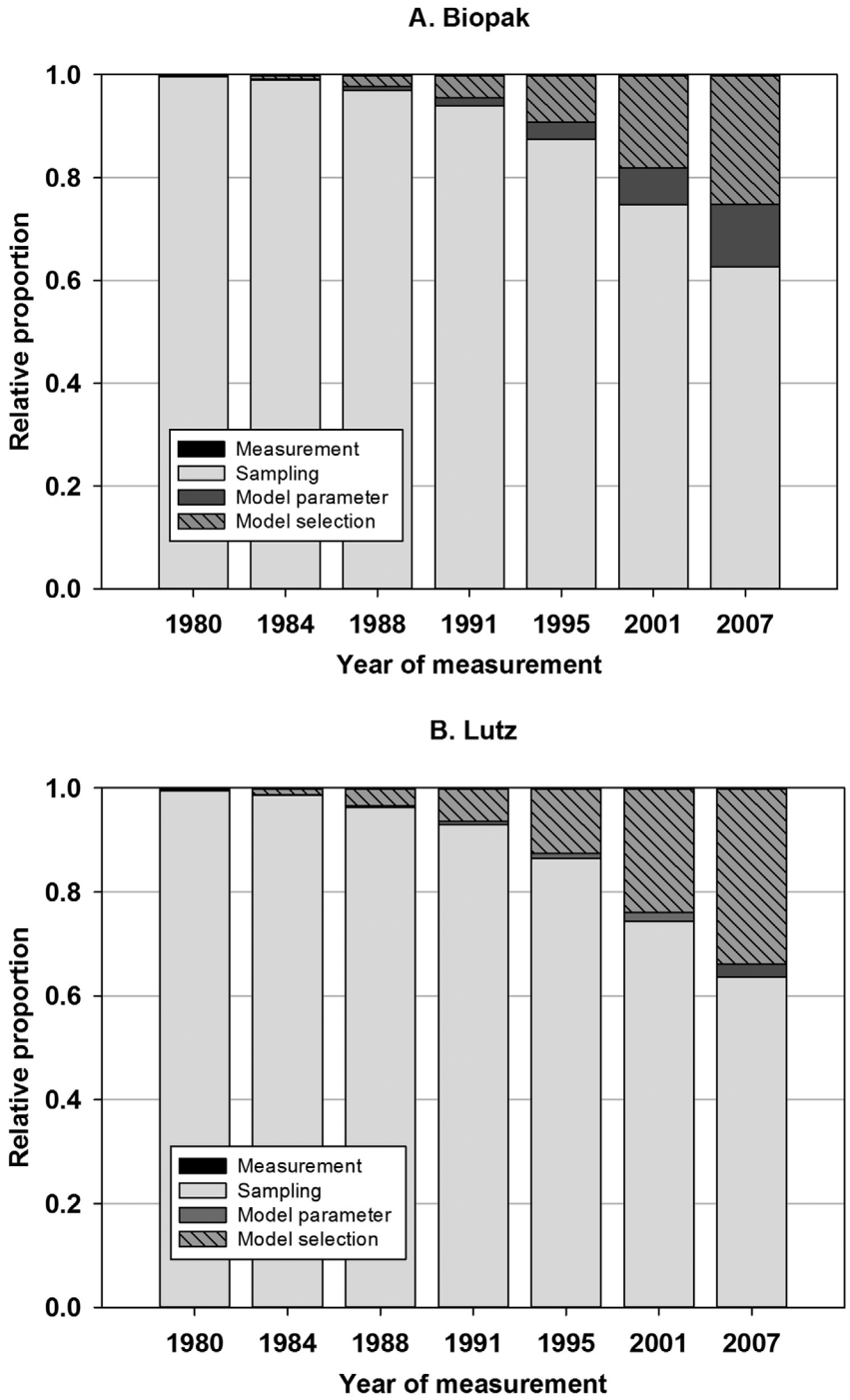

Fig. 6. Relative proportions of different forms of uncertainty "variance" for live-tree biomass in WS01 in the H. J. Andrews Experimental Forest, Oregon, for (A) Biopak set of equations and (B) Lutz set of equations. Note that at this scale, measurement uncertainty variance is too small to be visible.

future estimates or increasing sampling efficiency (Levine et al. 2014). Because forms of uncertainty are not strictly additive, we reduced each in turn to assess the effects on overall uncertainty. Measurement uncertainty was very small: even a 50% reduction would have little influence on overall uncertainty (<0.01%). Although sampling uncertainty constituted a major share of overall uncertainty, it would be difficult to reduce given the large number of plots currently sampled. For example, doubling the number of plots would likely reduce the sampling uncertainty by ≈30% and reduce overall uncertainty by 18%, but it would also double the cost of sampling. By determining which set of biomass equations is most appropriate (reducing model selection error), ≈13–19% of the overall uncertainty could be eliminated. Reducing uncertainty in model parameters could also reduce overall uncertainty, but because part of this uncertainty is related to natural variation among trees and the variation explained by the current set of equations is generally high (>90%), it would be difficult to reduce it further. For example, to reduce model parameter uncertainty by ≈50%, sampling intensity would have to increase 4-fold. This would reduce the overall uncertainty by 1–5% depending on the set of biomass models used. In this example, the best strategy to reduce overall uncertainty would target model selection. Reducing uncertainty could be achieved with relatively small effort, by sampling a small number of the principal tree species in WS01 to identify the most appropriate equations—or with substantially greater effort, by developing site-specific equations based on extensive, destructive sampling.

Although ours can be viewed as a relatively simple example of a larger set of more complex problems (Räty et al. 2011), it elucidates some of the fundamental steps to estimating and using uncertainty in a synthesis context. However, it does not address other aspects of uncertainty, e.g., the extent to which uncertainty components are correlated, multiplicative effects (which involve a covariance term in error propagation), and serial autocorrelation. For example, had the variable of interest been net primary production, the autocorrelation of individual tree biomass estimates over time would have to be considered to estimate net change in biomass over time. Similarly, had the focus been on litter production of fine roots and leaves, estimates might involve a proportion of the biomass dying per year multiplied by biomass, thus creating a multiplicative effect.

Future Challenges and Opportunities

Quantifying uncertainty more rigorously represents a powerful step for strengthening the ecological sciences (Yanai et al. 2010). Although uncertainty is not desirable, it is even less desirable to remain ignorant of uncertainty in synthesis (even if the estimate is approximate). Without a metric that allows for comparison, it is difficult to determine whether a new estimate is different, whether it reduces uncertainty, and ultimately, whether science is progressing. In a sense, synthesis science is now where reductionist science was before inferential statistics were developed, with no means of rigorously testing whether estimates differ. This could be remedied easily by adopting an agreed-upon metric, because many of the tests used in reductionist science (e.g., the t-test, which examines differences in means divided by the standard error of the mean) likely have analogues in synthesis science (e.g., the difference between alternative estimates divided by overall uncertainty).

If uncertainty becomes a metric to rigorously evaluate synthesis science, several challenges lie ahead. First, some aspects of uncertainty may be difficult to quantify. Some relationships may not be completely understood, but that is true of science in general and can be captured to some degree by model selection uncertainty. Even with complete knowledge about relationships, some quantities are difficult to measure or estimate without bias or with great precision. The question is, whether uncertainty estimates need to be perfect to be useful. There are ecological phenomena or processes that are difficult to measure and estimate, including net primary production, heterotrophic respiration, net ecosystem production, and net ecosystem carbon balance. Despite these limitations, these are routinely and productively used in ecosystem science. Although there are limits to estimating uncertainty, estimates can be useful if they represent a best attempt to incorporate current knowledge and methodology.

Second, because uncertainty is "undesirable", it is all too easy to equate the presence of uncertainty with the quality of synthesis. This leads to a tendency to deliberately underestimate, or fail to report, uncertainty, which undermines the progress of science. We suggest a radically different view of uncertainty: synthesis lacking inclusion, or involving an unexplained, deliberate underestimate of uncertainty, should be deemed unacceptable as quantitative science. Although it is difficult to address all forms of uncertainty, acknowledging and justifying those that are included vs. excluded should become routine. Accepting the positive aspects of synthesis uncertainty will likely require a major shift in attitudes among journals and reviewers.

Perhaps the most challenging aspect of uncertainty in synthesis involves model selection. It is rarely quantified and largely ignored, but it can constitute the largest portion of overall uncertainty. Synthesis scientists thus need to be more cognizant or careful about this form of uncertainty. For example, when two viable model structures are possible, it is important to test whether predictions significantly differ and, if they do, to compare the effect. This would create a strong inferential framework and identify alternative hypotheses to be evaluated (Platt 1964). Without such a framework, progress in science is impeded because critical uncertainties are hidden from view. Although model selection uncertainty is essential, we acknowledge that it does not readily fit within the realm of classical statistical methods; thus, more thought needs to be devoted to this topic.

To fully embrace uncertainty as a useful evaluation metric in synthesis science, several developments are needed: (1) improved access to (i.e., ability to retrieve) the information necessary to conduct uncertainty analysis, including standard estimates of measurement and model prediction uncertainty, (2) more effective and efficient methods to estimate and express uncertainty in all its forms, including model selection uncertainty, (3) standard guidelines for analyzing and reporting uncertainty, and broad acceptance of these guidelines, and (4) revised expectations for publication of synthesis efforts, including similar levels of rigor for synthesis as for reductionist science. This list may be daunting, but the potential benefits are vast. The time to begin this change is now, so let us start.

Acknowledgments

The authors thank QUEST (Quantifying Uncertainty in Ecosystem Studies) and Dr. Ruth Yanai for their encouragement in pursuing this important topic. Dr. Lisa Ganio provided valuable statistical advice. Ruth Yanai and two anonymous reviewers provided helpful suggestions. This research was funded by grants from the National Science Foundation (QUEST RCN: DEB1257906, Andrews LTER: DEB0823380) and the Pacific Northwest Research Station.

Literature Cited

Acker, S. A., C. B. Halpern, M. E. Harmon, and C. T. Dyrness. 2002. Trends in bole biomass accumulation, net primary production and tree mortality in *Pseudotsuga menziesii* forests of contrasting age. Tree Physiology 22:213–217.

Bartell, S. M., R. H. Gardner, and R. V. O'Neill. 1992. Ecological risk estimation. Lewis, Boca Raton, Florida, USA.

Draper, D. 1995. Assessment and propagation of model uncertainty. Journal of the Royal Statistical Society 57:45–97.

Ferson, S., and L. R. Ginzburg. 1996. Different methods are needed to propagate ignorance and variability. Reliability Engineering and System Safety 54:133–144.

Gallagher, R., and T. Appenzeller. 1999. Beyond reductionism. Science 284:79.

Garman, S. L., S. A. Acker, J. L. Ohmann, and T. A. Spies. 1995. Asymptotic height-diameter equations for twenty-four tree species in western Oregon. Research Contribution 10. Oregon State University, College of Forestry, Forest Research Laboratory, Corvallis, Oregon, USA.

Halpern, C. B., and J. F. Franklin. 1990. Physiognomic development of *Pseudotsuga* forests in relation to initial structure and disturbance intensity. Journal of Vegetation Science 1:475–482.

Halpern, C. B., and J. A. Lutz. 2013. Canopy closure exerts weak controls on understory dynamics: A 30-year study of overstory-understory interactions. Ecological Monographs 83:221–237.

Harmon, M. E., K. Bible, M. J. Ryan, D. Shaw, H. Chen, J. Klopatek, and X. Li. 2004. Production, respiration, and overall carbon balance in an old-growth *Pseudotsuga/Tsuga* forest ecosystem. Ecosystems 7:498–512.

Harmon, M. E., D. L. Phillips, J. J. Battles, A. Rassweiler, R. O. Hall, Jr., and W. K. Lauenroth.

2007. Quantifying uncertainty in net primary production measurements. Pages 238–260 *in* T. J. Fahey and A. K. Knapp, editors. Principles and standards for measuring net primary production in long-term ecological studies. Oxford University Press, Oxford, UK.

Holling, C. S. 2001. Understanding the complexity of economic, ecological, and social systems. Ecosystems 4:390–405.

Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. A. Birdsey. 2003. National-scale biomass estimators for United States tree species. Forest Science 49:12–35.

Kangas, A. 2010. Value of forest information. European Journal of Forest Research 129:863–874.

Lauenroth, W. K., A. A. Wade, M. A. Williamson, B. E. Ross, S. Kumar, and D. P. Cariveau. 2006. Uncertainty in calculations of net primary production for grasslands. Ecosystems 9:843–851.

Levine, C. R., R. D. Yanai, G. G. Lampman, D. A. Burns, C. T. Driscoll, G. B. Lawrence, J. A. Lynch, and N. Schoch. 2014. Evaluating the efficiency of environmental monitoring programs. Ecological Indicators 39:94–101.

Li, H., and J. Wu. 2006. Uncertainty analysis in ecological studies: An overview. Pages 45–66 *in* J. Wu, K. B. Jones, H. Li, and O. L. Loucks, editors. Scaling and uncertainty analysis in ecology: methods and applications. Springer, Netherlands.

Lindenschmidt, K.-E., K. Fleischbein, and M. Baborowski. 2007. Structural uncertainty in a river water quality modelling system. Ecological Modelling 204:289–300.

Lord, E. 1947. The use of range in place of standard deviation in the t-test. Biometrika 34:41–67.

Lutz, J. A., and C. B. Halpern. 2006. Tree mortality during early stand development: a long-term study of rates, causes, and consequences. Ecological Monographs 76:257–275.

Means, J. H., H. A. Hansen, G. J. Koerper, P. B. Alaback, and M. W. Klopsch. 1994. Software for computing plant biomass—BioPak users' guide. PNW-GTR-340. USDA Forest Service, Pacific Northwest Research Station, Portland, Oregon, USA.

Melson, S. L., M. E. Harmon, J. S. Fried, and J. Domingo. 2011. Estimates of live-tree carbon stores in the Pacific Northwest as sensitive to model selection. Carbon Balance and Management 6:1–16.

Michener, W. K., T. J. Baerwald, P. Firth, M. A. Palmer, J. L. Rosenberger, E. A. Sandlin, and H. Zimmerman. 2001. Defining and unraveling biocomplexity. BioScience 51:1018–1023.

Odum, E. P. 1977. The emergence of ecology as a new integrative discipline. Science 195:1289–1293.

Phillips, D. L., S. L. Brown, P. E. Schroeder, and R. A. Birdsey. 2000. Toward error analysis of large-scale forest carbon budgets. Global Ecology and Biogeography 9:305–313.

Platt, J. R. 1964. Strong inference. Science 146:347–353.

Räty, M., J. Heikkinen, and A. Kangas. 2011. Kriging with external drift in model localization. Mathematical and Computational Forestry and Natural-Resource Sciences 3:1–14.

Rowe, W. D. 1994. Understanding uncertainty. Risk Analysis 14:743–750.

Shugart, H. H. 1998. Terrestrial ecosystems in changing environments. Cambridge studies in ecology. Cambridge University Press, Cambridge, UK.

Sollins, P., C. C. Grier, F. M. McCorison, K. Cromack, Jr., R. Fogel, and R. L. Fredriksen. 1980. The internal element cycles of an old-growth Douglas-fir ecosystem in western Oregon. Ecological Monographs 50:261–285.

Yanai, R. D., J. J. Battles, A. D. Richardson, C. A. Blodgett, D. M. Wood, and E. B. Rastetter. 2010. Estimating uncertainty in ecosystem budget calculations. Ecosystems 13:239–248.