Applied Geography 93 (2018) 25–36

Contents lists available at ScienceDirect

Applied Geography

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

The sensitivity of ecosystem service models to choices of input data and spatial resolution

Kenneth J. Bagstad[a,b,*], Erika Cohen[c], Zachary H. Ancona[a], Steven G. McNulty[c], Ge Sun[c]

[a] Geosciences & Environmental Change Science Center, U.S. Geological Survey, Denver, CO, USA
[b] Wealth Accounting and Valuation of Ecosystem Services (WAVES) Program, The World Bank, Washington, DC, USA
[c] Eastern Forest Environmental Threat Assessment Center, USDA Forest Service Southern Research Station, Raleigh, NC, USA

ARTICLE INFO

Keywords:
Ecosystem accounting
InVEST
Model comparison
Spatial scale
Uncertainty
WaSSI

ABSTRACT

Although ecosystem service (ES) modeling has progressed rapidly in the last 10–15 years, comparative studies on data and model selection effects have become more common only recently. Such studies have drawn mixed conclusions about whether different data and model choices yield divergent results. In this study, we compared the results of different models to address these questions at national, provincial, and subwatershed scales in Rwanda. We compared results for carbon, water, and sediment as modeled using InVEST and WaSSI using (1) land cover data at 30 and 300 m resolution and (2) three different input land cover datasets. WaSSI and simpler InVEST models (carbon storage and annual water yield) were relatively insensitive to the choice of spatial resolution, but more complex InVEST models (seasonal water yield and sediment regulation) produced large differences when applied at differing resolution. Six out of nine ES metrics (InVEST annual and seasonal water yield and WaSSI) gave similar predictions for at least two different input land cover datasets. Despite differences in mean values when using different data sources and resolution, we found significant and highly correlated results when using Spearman's rank correlation, indicating consistent spatial patterns of high and low values. Our results confirm and extend conclusions of past studies, showing that in certain cases (e.g., simpler models and national-scale analyses), results can be robust to data and modeling choices. For more complex models, those with different output metrics, and subnational to site-based analyses in heterogeneous environments, data and model choices may strongly influence study findings.

1. Introduction

Spatial modeling of ecosystem services (ES)—the value nature provides to people—is a key step in ES assessments (Burkhard, Kroll, Nedkov, & Müller, 2012; Schröter, Remme, Sumarga, Barton, & Hein, 2015) and an increasingly common area of research in sustainability science (Burkhard and Maes 2017). ES modeling is useful to inform national ES assessments (e.g., Rabe, Koellner, Marzelli, Schumacher, & Grêt-Regamey, 2016), ecosystem accounting within the System of Environmental-Economic Accounting (U.N. et al., 2014), and other regional, subnational, and global assessments. A large body of literature, including modeling tools, has developed over the last decade to quantify ES (Bagstad, Semmens, Waage, & Winthrop, 2013a; Martinez-Harms & Balvanera, 2012; Schröter et al., 2015). Meanwhile new data sources derived through remote sensing (Araujo Barbosa, Atkinson, & Dearing, 2015), in combination with sensor networks and crowdsourcing (Johnson & Iizuka, 2016), offer additional data sources to populate models. Modelers now have a diverse body of feasible assessment approaches and an increasing number of global- and national-scale datasets to populate the models. Yet in both data-rich and data-limited environments, determining the most appropriate combination of data and tools for an ES assessment can be challenging.

This challenge also raises the question of replicability in ES assessment: how much difference would the use of different modeling tools and data sources make in an ES assessment for the decision-making process? In response to this challenge, scientists have called for inter- and intra-model comparative studies testing the sensitivity of ES models to choices of input data (Bagstad et al., 2013a; Sharps et al., 2017). Others have recommended the standardization of approaches, while remaining aware of the difficulty of doing so in a still-evolving field (Polasky, Tallis, & Reyers, 2015). Before such standards can be reached, better guidance is needed on navigating the choice and proper use of data and models for ES mapping in support of assessments. More broadly, such model comparison, calibration (where needed data are available), and sensitivity analysis can improve trust in environmental models (Bennett et al., 2013). Similar studies have evaluated the

* Corresponding author. Geosciences & Environmental Change Science Center, U.S. Geological Survey, P.O. Box 25046, DFC, MS 980, Denver, CO, USA.
E-mail address: kjbagstad@usgs.gov (K.J. Bagstad).

https://doi.org/10.1016/j.apgeog.2018.02.005
Received 7 November 2017; Received in revised form 6 February 2018; Accepted 6 February 2018
0143-6228/ Published by Elsevier Ltd.

impacts of data and model choices for the simulation of ecological phenomena (Martin, Brabyn, & Potter, 2011), hydrologic systems (Bell & Moore, 2000; Geza & McCray, 2008; Koren et al., 1999), and landscape pattern (Rendenieks, Tērauds, Nikodemus, & Brūmelis, 2017).

While ES research has grown substantially in the last 10–15 years, assessments of how data and model choices influence estimates of ES are relatively new. This issue is particularly important when ES assessments are conducted in developing countries, which may have limited data availability and modeling expertise. In this paper, we evaluate the effects of using different input data and spatial resolution when using two different ES modeling tools to conduct a terrestrial/freshwater ES assessment in Rwanda.

Past studies, which we review below, have addressed a number of important questions about model, data input, and data resolution choices in ES assessments, but have most commonly addressed only one, and occasionally two, of these three issues. Additionally, we are unaware of previous studies that make multiple comparisons across multiple modeling tools and ES. Nearly all authors have suggested the need for further research across more diverse study contexts, to better assess the range of application of their findings.

In this study, we modeled carbon sequestration and storage, sediment regulation, and annual and seasonal water yield as part of a national-scale ecosystem accounting project in Rwanda, using the Integrated Valuation of Ecosystem Services and Tradeoffs (InVEST, Sharp et al., 2016) and Water Supply Stress Index (WaSSI, Caldwell et al., 2013; Sun et al., 2011) modeling tools. Below, we reviewed past studies on the effects of data and model choices on ES assessment results. Next, we tested the similarity of conclusions drawn about ES trends in Rwanda from 1990 to 2010 using the InVEST and WaSSI modeling tools. We then compared the results of InVEST and WaSSI models using input data of varying spatial resolution (30 and 300 m) and three different input land cover datasets to test whether coarser resolution and/or global data give similar results. We compared all results at the national scale, the provincial scale (for Rwanda's five provinces), and used statistical analyses to compare mean values and rank-order correlation at the subwatershed scale. By evaluating the effects of ES data and model choices, we tested whether previous authors' conclusions about data and model selection hold for Rwanda, a small, heterogeneous, and relatively data-limited developing nation in central Africa. We also provided further instruction to guide data and model choice in ES mapping and modeling elsewhere.

1.1. Past studies on the effects of model and data choices on ES assessments

As the ES modeling literature has grown, an increasing number of studies have tested the effects of using different models, data inputs, spatiotemporal resolution, and uncertainty analysis in ES assessments, though such findings have not been broadly synthesized. First, *models* differ widely in their purpose, approach, and output metrics. Given this range, fit for purpose is an important consideration (Schröter et al., 2015). Simpler models may be adequate for addressing screening-level policy questions, while detailed models may be required for high-resolution spatial planning and prioritization. To understand when and where complex models produce more reliable results or whether simpler approaches are satisfactory, it can be useful to compare the results of models that use different methods but share the same purpose (Schulp, Burkhard, Maes, van Vliet, & Verburg, 2014; Sharps et al., 2017; Tallis and Polasky 2011; Willcock et al., in press). Model calibration remains a critical, and often overlooked, aspect of model performance evaluation, especially in data-limited environments (Baveye, 2017).

Second, modelers must choose which *data sources* to use as inputs to ES models. National datasets for key attributes like land cover, soils, or climate may not be available in all countries (particularly in developing nations), raising the question of the adequacy of global data for ES modeling and how much agreement ES model results have when using different global and local datasets as inputs. For instance, Dong, Bryan, Connor, Nolan, and Gao (2016) reported 60–65% per-pixel agreement between different global land cover datasets. Benítez, McCallum, Obersteiner, and Yamagata (2007) found differences of up to 45% in global carbon sequestration estimates for model results that used different global datasets. In a study of crop and fodder production in northern Germany, Kandziora, Burkhard, and Müller (2013) found that European input data overestimated ES provision relative to local data, while Redhead et al. (2016) found that U.K. data produced a better calibration of water models than global data. Finally, Schulp and Alkemade (2011) compared pollination model outputs using national, two European, and two global land cover input datasets, and found results generated using GlobCover to yield the best agreement with those from national data.

Third, choices must be made about the *spatiotemporal resolution* on which to run models. Generally, high-resolution analyses are assumed to be more accurate (though true accuracy assessments require model calibration), but potential gains from progressively higher resolution analysis must be weighed against greater storage and processing requirements, and could reach a point of diminishing returns (Grêt-Regamey et al., 2014; Hamel et al., 2017; Schulp & Alkemade, 2011). Decision-maker needs for both spatial resolution and model accuracy, both of which may be context dependent, should also be considered (Willcock et al., 2016). Continual improvements in data storage and computer processing power mean that moderate-to high-resolution ES analysis is increasingly feasible in developed nations and for many smaller developing countries. Yet for larger middle-income and developing countries, questions of the optimal spatial resolution on which to run ES models remain.

Fourth, data are of different quality, and *data uncertainty* is major source of variability and error in ES modeling (Hamel & Bryant, 2017). At least two recent studies have evaluated the effects of uncertainty related to error in land cover datasets (Dong et al., 2016; Foody, 2015), and further work on this and other types of uncertainty in ES analysis is needed.

Nineteen recent studies focus on the first three types of data and model choices that we address in our study. Each study's characteristics and findings are summarized below (Table 1). We exclude studies from this table that rely on land cover-based benefit transfers (Konarska, Sutton, & Castellon, 2002; Whitham, Shi, & Riordan, 2015) due to this method's well-known limitations (Bockstael, Freeman, Kopp, Portney, & Smith, 2000). We also excluded papers that conducted mapping at different scales but either aggregated fine-scale results (Larondelle & Lauf, 2016) or used different indicators for analysis at different scales (Rabe et al., 2016).

Taken together, these studies reach several broad conclusions. When using different approaches, local ES differences may be evident for small geographic regions but disappear when averaged across larger regions (Bagstad, Semmens, & Winthrop, 2013b; Dong et al., 2016). At national and continental scales, proxy-based results often perform poorly when compared to those of primary ES data or models (Eigenbrod et al., 2010; Schulp et al., 2014). Willcock et al. (in press) generally support this, but found that some complex ES models do not always have the best predictive power.

Geographic aggregation means that infrequent and/or dispersed values (e.g., scattered wetland or forest patches) will be “lost” as they are averaged into coarser scale data, particularly for categorical data like land cover. We thus generally expect fine-resolution data to produce more accurate ES assessment results than coarse-resolution data. This is typically the case (Grêt-Regamey et al., 2014 for all services but carbon sequestration; Grafius et al., 2016). Less divergence is expected in homogeneous environments than in heterogeneous ones, meaning that coarser-resolution analyses may be adequate in relatively homogeneous settings (Grêt-Regamey et al., 2014; Schulp & Alkemade, 2011; Willcock et al., in press). Additionally, a comparison of ES results at very coarse resolutions (1 vs. 10 km for sub-Saharan Africa) found