Environmental Science and Policy 87 (2018) 64–73

Contents lists available at ScienceDirect

Environmental Science and Policy

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

Review

Defining a best practice methodology for modeling the environmental performance of agriculture

Christina Tonitto[a,*], Peter B. Woodbury[b], Eileen L. McLellan[c]

[a] Cornell University, Cornell International Institute for Food, Agriculture and Development, Ithaca, NY, USA
[b] Cornell University, Soil and Crop Sciences, Ithaca, NY, USA
[c] Environmental Defense Fund, Washington DC, USA

ARTICLE INFO

Keywords:
Model methodology
Agricultural management
Payment for ecosystem services

ABSTRACT

Robust quantification of the environmental performance of agricultural management practices is critical both for ensuring regulatory compliance and for creating accountability in voluntary environmental markets and corporate sustainability commitments. Because environmental impacts cannot be measured under all conditions and on all farms, models are required. However, models must be used appropriately if predictions of environmental performance are to be reliable. To assist policymakers and stakeholders, we define a 7-step process for model selection and use, and present a case study applying this 7-step process to greenhouse gas emissions from corn (*Zea mays* L.) fields in the USA. Based on this case study and other examples from the literature, we suggest that all models are limited by the data available to validate them for different combinations of cropping systems, management practices, site conditions, and types of environmental performance. Additionally, both statistical and process models are much more reliable for making predictions of environmental performance for multiple fields and years than for predictions of a single location and year. We suggest that using this 7-step process will help improve predictions of environmental performance for regulatory and voluntary purposes at local, state, and national scales.

1. Introduction

Models representing how agriculture affects the environment have become increasing accessible via the internet, often with graphical user interfaces that make them accessible to a wide variety of users. However, non-expert users are often unaware of the limitations of model suitability resulting from both model structural uncertainty (simulation of the underlying biophysical processes) and limited availability of data for model validation. As a result, models may be selected based on user accessibility rather than due to proven performance for the specific task at hand. This is especially important because results are unlikely to receive the scrutiny accorded peer-reviewed research.

In this article, we define a best practice framework for using models in developing and implementing methods to quantify the environmental performance of agricultural management practices in a Payment for Ecosystem Services (PES), regulatory, or voluntary setting (including corporate sustainability commitments). The primary audience is policymakers and stakeholders in the agricultural and environmental management communities who seek to use models for applied (not research) purposes of quantifying diverse environmental impacts of management practices in agricultural and other managed landscapes.

We demonstrate the use of the 7-step process using a case study of agricultural greenhouse gas (GHG) emissions while noting that similar concerns and procedures are broadly applicable for quantifying the environmental impact of land management practices on soil, water, and air quality. Given the spatial and temporal context of regulatory policy (structured as payment-for-performance, e.g. Angelsen, 2017), PES markets (e.g. CAR, 2012), and corporate sustainability commitments, we focus on regional (aggregated) field-scale assessment of environmental performance and annual to multi-year time scales. Models can also be used to predict longer-term (e.g., multi-decadal) environmental impacts that cannot be fully validated, such as future effects of climate change on water quantity and quality. However, such uses of models are beyond the scope of the 7-step process defined herein.

2. Models for assessing environmental performance

2.1. Why do we need models?

Models are needed because measurement of all relevant ecosystem processes and types of environmental performance is precluded by technical and financial reasons. Models are increasingly a component of

* Corresponding author.
E-mail addresses: ctonitto@cornell.edu (C. Tonitto), peterwoodbury@cornell.edu (P.B. Woodbury), emclellan@edf.org (E.L. McLellan).

https://doi.org/10.1016/j.envsci.2018.04.009
Received 19 January 2018; Received in revised form 14 April 2018; Accepted 16 April 2018
Available online 05 June 2018

environmental planning and management for agriculture (e.g. Garrett et al., 2013; Melkonian et al., 2008). Many commonly used agricultural models were developed to assess regional environmental performance of agricultural best management practices (BMPs). For example, United States Department of Agriculture (USDA) research teams developed the SWAT (Neitsch et al., 2009) and EPIC (Williams, 1995) models to quantify the impact of management practices on nutrient cycling, soil erosion, and water quality. Traditionally these complex process models have been used by experts to test model features or apply the model to specific regions for which site-specific data are available for model calibration and validation. However, with emerging interest in various PES markets, such as carbon markets and water quality markets, model predictions are increasingly applied as a way to (1) design BMPs for PES market protocol descriptions and (2) verify that environmental performance has improved following BMP implementation. Companies pledging to reduce supply chain greenhouse gas emissions and/or water quality impacts likewise place increasing reliance on models to focus their efforts and track impacts. Despite widespread interest in using model predictions as a key component of defining and meeting land management goals, there is not a broadly accepted process for selecting appropriate models for PES market, voluntary accounting, regulatory, or policy applications. This is particularly problematic for complex process models that are increasingly applied outside of the research settings in which they were developed and tested. Specifically, such models are increasingly used by non-experts who lack a thorough understanding of the biophysical mechanisms underlying the model as well as the limitations and assumptions represented by model algorithms and parameters. For example, food supply chain sustainability initiatives may apply a process model to identify and rank potential GHG emissions reductions associated with various land management scenarios, with model results determining priorities for funding decisions. In such situations, inaccurate model predictions may divert available funding to projects with limited potential to improve environmental performance. Such use of models by non-experts and outside of the scientific peer-reviewed publication process is expected to increase as large areas of farmland enroll in PES or similar projects that require quantification of environmental performance for project design and verification. Our best practice methodology provides step-by-step guidelines to assist non-expert teams and policy makers as they select quantitative tools relevant to their land management goals.

2.2. Statistical versus process models

The simplest models are statistical (also called empirical) models. Such models predict environmental performance as a function of observed relationships between a process and one or more independent variables; complete ecosystem processes are not modeled. Statistical models define regionally-specific functional relationships based on data from field experiments or environmental surveys relevant to specific cropping systems, soil textures, or climates. For example, for agricultural GHG emissions, statistical models of N_2O emission have been used to simplify the complexity of nitrogen (N) gas emission measurement and modeling (e.g. Dalgaard et al., 2011; Leip et al., 2011; Millar et al., 2010; Tonitto et al., 2009). Statistical models are almost always simpler, more transparent, and easier to use than process models. Therefore, there is less risk of a spurious prediction from a statistical model than from a process model. However, because statistical models depend entirely on the observations used to derive the relationship, regionally-specific (Tier 2) statistical models generally have a smaller geographic range of application than process models (Smith et al., 2012). We use the term “observations” to mean reliable, publically available (preferably peer-reviewed) data, such as from field experiments or on-farm measurements, that are directly applicable to the cropping system and management practice under investigation.

In contrast to statistical models, complex ecosystem process models attempt to represent all processes affecting environmental flows and stocks. Process models are appealing to non-experts because they may, in theory, be capable of making predictions for many combinations of geography, climate, cropping system, and management practice. In practice, however, process models are limited by incomplete scientific understanding of key processes. For example, mechanistic descriptions of controls on soil organic carbon (SOC) stabilization (e.g.Giardina and Ryan, 2000; Kramer et al., 2012; Mikutta et al., 2006; Sollins et al., 2006; Torn et al., 1997, 2005) and impacts of management practices on long-term SOC sequestration (e.g. Baker et al., 2007; Eagle and Olander, 2012) remain incomplete. As a result, the full mechanistic complexity of SOC accumulation is not currently represented in ecosystem models (Tonitto et al., 2016). In addition, the ability to use a model for specific management scenarios or site conditions is always limited by the availability of field data to validate model predictions. Together, these limitations pose challenges for the use of process models. A recent synthesis of agroecosystem model applications (Brilli et al., 2017) found that: 1) 52% of N-cycle applications demonstrated poor validation due to limitations in model representation of soil physical and chemical properties, and climate (pedo-climate), 2) model representation of management practices led to poor validation outcomes in 43% of C-cycle applications, and 3) model representation of pedo-climate led to poor simulation of spatial and temporal C- and N-cycle dynamics in 20% of applications (Brilli et al., 2017). These results are consistent with other reviews of process model applications that demonstrated high variation in predictions by individual models within model ensemble studies (e.g. Ehrhardt et al., 2018), poor agreement between seasonal GHG predictions and observations (e.g. Tonitto et al., 2016), the need for site-specific calibration of water quality models (Forsberg et al., 2017), and overall poor performance of process-based fate and transport water quality models compared to simpler statistical models (Kleinman et al., 2017). For example, a recent test at 11 US sites found that a daily process model did not outperform an annual statistical model in predicting phosphorus pollution from agricultural fields (Bolster et al., 2017). Therefore, while process models may appear to be more broadly applicable and flexible they may be equally limited by data availability as are statistical models.

However, in some situations a suitable process model may offer more capacity to model different management practices or changing climatic conditions. For example, for well-studied Fluxnet-Canada cropping systems Goglio et al. (2018) demonstrated that the DNDC model produced more accurate predictions of observations than did statistical IPCC Tier 1 and Tier 2 models. While process models can simulate nuanced changes in management or site conditions, this does not guarantee improved prediction accuracy. Validation studies have demonstrated that seasonal predictions often don’t capture observed trends (e.g. Beheydt et al., 2007; Bolster et al., 2017). For example, a model may not accurately predict system response to weather patterns, resulting in prediction of peak events that do not occur or missing peak events that are observed. In addition, input data quality can affect model predictions. Bagstad et al. (2018) demonstrated that process models predicted very different magnitudes for some ecosystem properties (such as soil C and sediment export) depending on the resolution of the input data, indicating that systems with limited input observations may not be more accurately predicted using process models. Specifically, Bagstad et al. (2018) concluded that for more complex models applied to address heterogeneous sites, data and model choices could strongly influence predictions. Similarly, in a comparison of agroecosystem models for grassland, maize, rice, and wheat cropping systems, Ehrhardt et al. (2018) found that while increased site input data increased the accuracy of yield predictions, increased site observations did not improve (or for some systems reduced) the number of models with N_2O emission predictions within 1 SD of observations. Therefore, this ensemble comparison indicates that, for the example of N_2O emissions, model predictive capacity remains uncertain; uncertainty in predcitive capacity is likely due to both model structural uncertainty as well as N_2O emission observations with inadequate