Journal of Hydrology 539 (2016) 27–37

Contents lists available at ScienceDirect

Journal of Hydrology

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

Model parameter uncertainty analysis for an annual field-scale P loss model

Carl H. Bolster[a,*], Peter A. Vadas[b], Debbie Boykin[c]

[a] Food Animal Environmental Systems Research Unit, USDA-ARS, Bowling Green, KY 42101, United States
[b] Dairy Forage Research Center, USDA-ARS, Madison, WI 53706, United States
[c] USDA-ARS, Stoneville, MS 38776, United States

ARTICLE INFO

Article history:
Received 7 July 2015
Received in revised form 29 February 2016
Accepted 3 May 2016
Available online 10 May 2016
This manuscript was handled by Laurent Charlet, Editor-in-Chief, with the assistance of Antonino Cancelliere, Associate Editor

Keywords:
Annual P Loss Estimator model (APLE)
Model uncertainty
Confidence intervals
Prediction intervals
Phosphorus modeling

SUMMARY

Phosphorous (P) fate and transport models are important tools for developing and evaluating conservation practices aimed at reducing P losses from agricultural fields. Because all models are simplifications of complex systems, there will exist an inherent amount of uncertainty associated with their predictions. It is therefore important that efforts be directed at identifying, quantifying, and communicating the different sources of model uncertainties. In this study, we conducted an uncertainty analysis with the Annual P Loss Estimator (APLE) model. Our analysis included calculating parameter uncertainties and confidence and prediction intervals for five internal regression equations in APLE. We also estimated uncertainties of the model input variables based on values reported in the literature. We then predicted P loss for a suite of fields under different management and climatic conditions while accounting for uncertainties in the model parameters and inputs and compared the relative contributions of these two sources of uncertainty to the overall uncertainty associated with predictions of P loss. Both the overall magnitude of the prediction uncertainties and the relative contributions of the two sources of uncertainty varied depending on management practices and field characteristics. This was due to differences in the number of model input variables and the uncertainties in the regression equations associated with each P loss pathway. Inspection of the uncertainties in the five regression equations brought attention to a previously unrecognized limitation with the equation used to partition surface-applied fertilizer P between leaching and runoff losses. As a result, an alternate equation was identified that provided similar predictions with much less uncertainty. Our results demonstrate how a thorough uncertainty and model residual analysis can be used to identify limitations with a model. Such insight can then be used to guide future data collection and model development and evaluation efforts.

Published by Elsevier B.V.

1. Introduction

The translocation of phosphorus (P) from the landscape to surface waters via runoff, erosion, and/or subsurface leaching can lead to water quality deterioration of P-sensitive water bodies. The degradation of water quality resulting from P loading from diffuse sources is a global concern (Kleinman et al., 2015; Sharpley et al., 2015) with such notable examples as the Baltic Sea, Chesapeake Bay, Florida Everglades, Mississippi River, and Yangtze River (Boesch et al., 2006; Dai et al., 2011; Dale et al., 2010; Executive Order 13508, 2009; HELCOM, 2011; Richardson et al., 2007). In many areas, agriculture is a significant contributor of P loading to P-sensitive waters. To mitigate the effects of agricultural activities on water quality, decades of research has been devoted to better understanding the processes controlling P movement through the landscape and in developing conservation practices to minimize P losses (Radcliffe and Cabrera, 2007). Given the costs and long-term commitments associated with field-scale experiments, emphasis has also been placed on using field-scale models to test different management strategies on reducing P loss from agricultural fields. Model simulations have also been used to assess the effectiveness of various conservation practices at the watershed scale (USDA-NRCS, 2013). To corroborate the findings of any model, however, will require comparing model predictions with long-term monitoring data.

Abbreviations: AP, active P; APLE, Annual P Loss Estimator; CI, confidence interval; LP, labile P; MCS, Monte Carlo simulations; PI, prediction interval; STP, soil test P.

* Corresponding author.
E-mail address: carl.bolster@ars.usda.gov (C.H. Bolster).

http://dx.doi.org/10.1016/j.jhydrol.2016.05.009
0022-1694/Published by Elsevier B.V.

While model predictions of P fate and transport can provide useful information to researchers, land owners, regulatory agencies, and other stakeholders, uncertainties exist with all model predictions, regardless of how complex or "physically-based" a model may be (Radcliffe et al., 2009; Sharpley et al., 2002). Several sources of uncertainty exist that are inherent to all P loss models (Beck, 1987; Beven et al., 2007; Gupta et al., 2012). These include model structure errors/inadequacies that result when approximating complex physical phenomena with simplified mathematical models. These approximations result from our incomplete knowledge of the system, as well as practical limitations of including all processes and associated parameters into the model. Moreover, errors may be introduced by the numerical methods employed for solving the model equations and how the model is discretized in time and space. In addition, there is an inherent amount of randomness within natural systems (both temporally and spatially), much of which is not adequately, or cannot reasonably be, captured by models.

Measurement errors in the input variables that are required to run the model such as precipitation, evapotranspiration, soil test P, P application rates, and initial and boundary conditions will also affect the accuracy of the model predictions. Moreover, there may be errors resulting from using unrepresentative values for these input variables; for instance, the use of measured soil test P (STP) at a point to describe spatially variable STP over a large area. Errors associated with the model parameters will also affect the reliability of the model predictions. The magnitude of the errors introduced from these different sources will depend on the validity of the model assumptions, the complexity of the model, the quality of the input data, and on how well the various model parameters have been estimated. Because these errors are often interrelated, it is difficult to isolate and obtain good estimates of their magnitude, particularly errors associated with model structure (Gupta et al., 2012; Yen et al., 2014). Further complicating matters is that multiple models and parameter sets may describe a data set equally well (Beven, 2006a). Nevertheless, efforts should be taken to obtain reasonable estimates of model uncertainties (National Research Council, 2007; USEPA, 2009). These uncertainties may be estimated by model uncertainty analysis, values reported in the literature, and/or expert assessment (Uusitalo et al., 2015).

The impact of model parameter uncertainty on predictions of P loss is arguably the most common type of uncertainty analysis conducted with P loss models (Barlund and Tattari, 2001; Beven et al., 2007; Dean et al., 2009; Krueger et al., 2009; McFarland and Hauck, 2001; Smith and Wheater, 2004; Veith et al., 2010; Zhang and Haan, 1996). Model parameters are those constants incorporated into the model equations for making calculations from input data. Values for model parameters are often estimated by adjusting (i.e. calibrating) their values (either manually or by an automated computer algorithm) until the differences between modeled and observed data are minimized. A commonly used method in the hydrological sciences for model calibration is least-squares regression, where values of the model parameters are chosen that minimize the sum of the squared differences between model predictions and observations. Uncertainties in the model parameters can be estimated from the resulting parameter variance-covariance matrix provided that the assumptions of least-squares regression are valid or not too badly violated (Seber and Wild, 2003; Draper and Smith, 1998; Helsel and Hirsch, 2002). Potential sources of errors in model parameters estimated by calibration include: using incorrect calibration performance measures (i.e. optimization targets); using inaccurate, incomplete, or unrepresentative data sets during model calibration; and ignoring uncertainties in the calibration data (Bolster and Tellinghuisen, 2010; Haan, 2002; Sorooshian and Gupta, 1995).

We recently conducted an analysis investigating the effects of model input error on prediction uncertainties of P loss at the field scale using the Annual P Loss Estimator (APLE) model (Bolster and Vadas, 2013), an empirically-based spreadsheet model developed to describe annual, field-scale P loss when surface runoff is the dominant P loss pathway (Vadas et al., 2009). In this study, we extend our analysis by evaluating the effects of model parameter uncertainties on predictions of P loss for APLE. The specific objectives of this study were to estimate the model parameter uncertainty associated with five internal regression equations used in APLE and to evaluate how the parameter uncertainties affect model prediction uncertainties. We estimate the parameter uncertainties associated with the regression equations used to estimate total soil P from measurements of soil clay content, organic matter, and labile P; the P enrichment ratio calculated from erosion rates; concentration of P in runoff calculated from labile soil P; and the partitioning of P between runoff and infiltration for applied manures and fertilizers based on runoff ratio. Our analysis included calculating parameter uncertainties and 95% confidence and prediction intervals for the regression equations. We then calculated predictions of P loss using the APLE model while including uncertainties in model parameters and inputs and compared the relative magnitude of these sources of uncertainty to the overall uncertainty associated with predictions of P loss. Results from this study highlight the importance of including reasonable estimates of model parameter uncertainties when using models to predict P loss. Our results also demonstrate how the estimation of model parameter uncertainty can be used to identify model limitations.

2. Methodology

2.1. Annual P Loss Estimator (APLE) model

The APLE model calculates annual total P loss in surface runoff from agricultural fields as (Vadas et al., 2009):

$$P_{tot} = P_{sed} + DP_{soil} + DP_{man} + DP_{fert} \quad (1)$$

where P_{tot} is the total annual P loss from surface runoff (kg ha^{-1}), P_{sed} is annual sediment P loss from eroded soil (kg ha^{-1}), DP_{soil} is annual dissolved P loss in runoff from soil (kg ha^{-1}), DP_{man} is annual dissolved P loss in runoff from applied manure (kg ha^{-1}), and DP_{fert} is annual dissolved P loss in runoff from applied fertilizer (kg ha^{-1}). Each component in Eq. (1) includes one or more terms obtained by regression as detailed below (Vadas et al., 2009).

The component estimating particulate P loss is calculated as:

$$P_{sed} = SED \cdot PER \cdot TP \cdot 10^{-6} \quad (2)$$

where TP is total soil P (mg kg^{-1}), SED is the annual erosion rate (kg ha^{-1}), PER is the P enrichment rate that accounts for the preferential movement of fine soil particles enriched in P, and 10^{-6} is a unit conversion factor. The P enrichment ratio is calculated as:

$$PER = C_1 \cdot SED^{C_2} \quad (3)$$

where C_1 and C_2 are regression coefficients. The model is coded so that PER has a minimum value of 1.

Total soil P is determined by summing four simulated P pools: organic (OP), labile (LP), active (OP), and stable (SP). The organic P pool (OP) is calculated as:

$$OP = 10^4 \cdot SOC/(NP \cdot CN) \quad (4)$$

where 10^4 is a unit conversion factor, SOC is percent soil organic carbon (assumed to be equal to 58% of soil organic matter), NP is the nitrogen to phosphorus ratio in the soil organic matter (assumed to be 8), and CN is the carbon to nitrogen ratio of the soil organic matter (assumed to be 14). The active (AP) and stable P (SP) pools are determined by (Jones et al., 1984):