

Evaluation of Phosphorus Site Assessment Tools: Lessons from the USA

Andrew Sharpley,* Peter Kleinman, Claire Baffaut, Doug Beegle, Carl Bolster, Amy Collick, Zachary Easton, John Lory, Nathan Nelson, Deanna Osmond, David Radcliffe, Tamie Veith, and Jennifer Weld

Abstract

Critical source area identification through phosphorus (P) site assessment is a fundamental part of modern nutrient management planning in the United States, yet there has been only sparse testing of the many versions of the P Index that now exist. Each P site assessment tool was developed to be applicable across a range of field conditions found in a given geographic area, making evaluation extremely difficult. In general, evaluation with in-field monitoring data has been limited, focusing primarily on corroborating manure and fertilizer “source” factors. Thus, a multiregional effort (Chesapeake Bay, Heartland, and Southern States) was undertaken to evaluate P Indices using a combination of limited field data, as well as output from simulation models (i.e., Agricultural Policy Environmental eXtender, Annual P Loss Estimator, Soil and Water Assessment Tool [SWAT], and Texas Best Management Practice Evaluation Tool [TBET]) to compare against P Index ratings. These comparisons show promise for advancing the weighting and formulation of qualitative P Index components but require careful vetting of the simulation models. Differences among regional conclusions highlight model strengths and weaknesses. For example, the Southern States region found that, although models could simulate the effects of nutrient management on P runoff, they often more accurately predicted hydrology than total P loads. Furthermore, SWAT and TBET overpredicted particulate P and underpredicted dissolved P, resulting in correct total P predictions but for the wrong reasons. Experience in the United States supports expanded regional approaches to P site assessment, assuming closely coordinated efforts that engage science, policy, and implementation communities, but limited scientific validity exists for uniform national P site assessment tools at the present time.

Core Ideas

- Concern over the accuracy of P Indices has led to calls for their evaluation.
- Insufficient field data exist to evaluate P site assessment tools comprehensively.
- Researchers can use nonpoint source models as surrogates to field data if vetted first.
- There is no scientific justification for a single national P Index.

Copyright © American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America. 5585 Guilford Rd., Madison, WI 53711 USA. All rights reserved.

J. Environ. Qual. 46:1250–1256 (2017)
doi:10.2134/jeq2016.11.0427

This is an open access article distributed under the terms of the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Received 3 Nov. 2016.

Accepted 5 Apr. 2017.

*Corresponding author (sharpley@uark.edu).

MODERN nutrient management planning seeks to identify critical source areas of phosphorus (P) loss—single fields within watersheds that are disproportionately responsible for P export from the watershed. A multitude of decision support tools have been developed to enable critical source area identification, but no site assessment tool has been as widely adopted as the P Index. The P Index is founded on a well-documented framework of “source” and “transport” factors that, when combined, quantify the relative vulnerability of a field to P loss in runoff (Sharpley et al., 2003). Development and adoption of different versions of the P Index in the United States dates back to the early 2000s after an historic agreement between the USEPA and USDA to address nutrient runoff concerns from animal feeding operations (USDA and USEPA, 1999).

Despite widespread implementation of the P Index, P continues to be a major contributor to the impairment of a large proportion of surface waters in the United States (Dubrovsky et al., 2010; USEPA, 2015). For instance, harmful algal blooms have been linked to excess P in Western Lake Erie (Michalak et al., 2013; Scavia et al., 2014) and Florida (Reddy et al., 2011; USEPA Scientific Advisory Board, 2011), as well as to hypoxia in the Northern Gulf of Mexico (Alexander et al., 2008; Dale et al., 2010; Rebich et al., 2011). These concerns, along with an inability to meet eutrophication mitigation goals in areas where the P Index has been implemented, such as in the Chesapeake Bay Watershed (Chesapeake Bay Program, 2013; USEPA, 2010), have heightened attention on the need to improve P management strategies, particularly those that identify and target more effective conservation practices. Furthermore, differences between P Indices of individual states produce different nutrient

A. Sharpley, Crop, Soil and Environmental Sciences, Division of Agriculture, Univ. of Arkansas, Fayetteville, AR; P. Kleinman and T. Veith, USDA-ARS, Pasture Systems and Watershed Management Research Unit, University Park, PA; C. Baffaut, USDA-ARS, Cropping Systems and Water Quality Research Unit, Columbia, MO; D. Beegle, Plant Sciences, The Pennsylvania State Univ., University Park, PA; C. Bolster, USDA-ARS, Food Animal Environmental Systems Research Unit, Bowling Green, KY; A. Collick, Dep. of Agriculture, Food and Resource Sciences, Univ. of Maryland—Eastern Shores, Princess Anne, MD; Z. Easton, Biological Systems Engineering, Virginia Tech, Blacksburg, VA; J. Lory, Plant Science Division, Univ. of Missouri, Columbia, MO; N. Nelson, Agronomy, Kansas State Univ., Manhattan, KS; D. Osmond, Crop and Soil Sciences, North Carolina State Univ., Raleigh, NC; D. Radcliffe, Crop and Soil Sciences, Univ. of Georgia, Athens, GA; J. Weld, Ecosystem and Science Management, The Pennsylvania State Univ., University Park, PA. Assigned to Associate Editor Cole Green.

Abbreviations: APEX, Agricultural Policy Environmental eXtender; APLE, Annual P Loss Estimator; EPIC, Erosion-Productivity Impact Calculator; RUSLE2, Revised Universal Soil Loss Equation version 2; SWAT, Soil and Water Assessment Tool; SWAT-VSA, Soil and Water Assessment Tool-Variable Source Area; TBET, Texas Best Management Practice Evaluation Tool.

management recommendations for fields with similar P loss potential (Osmond et al., 2012).

In 2010, the USDA recognized the need for systematic testing and revision of existing P Indices to ensure that recommendations derived from their application were scientifically defensible (USDA-NRCS, 2011). This USDA decision was based largely on an in-depth review and assessment under the auspices of the Southern Extension and Research Activity-17 (SERA-17, <https://sera17.org/>) of P Indices across the United States (Sharpley et al., 2012). In response, US scientists formed three regional teams (Fig. 1) to address distinct concerns surrounding P site assessment. Some regions, such as the Southern United States and Chesapeake Bay, evaluated state P Indices with edge-of-field water quality data and modeling data as a means to ensure consistency in P site assessment outcomes. Elsewhere, scientists in the Heartland region placed a greater emphasis on evaluation of fate and transport models to determine if the models could be used to test or even to replace some of the functions of a P Index.

Evaluation of P site assessment tools, which are expected to apply to a wide range of site and management conditions, requires multiple lines of evaluation data. However, limited edge-of-field data are available to evaluate P site assessment tools in any given state or region. Furthermore, P Index evaluation may require water quality data from extended temporal spans. In general, measured water quality and weather data for a given location typically do not match the length of time required to test P site assessment tools appropriately. Most P Indices use 30-yr records of weather data, along with soil-test P and Revised Universal Soil Loss Equation version 2 (RUSLE2) estimates of long-term annual average soil loss to estimate particulate P loss potential. Edge-of-field, water quality datasets tend to skew toward nonwinter months and frequently have only a few years of data representing a specific management practice (Harmel et al., 2006, 2008). Even in cases with comparatively extensive management and water quality data monitoring (Veith et al., 2015, 2017), the inability to account for all confounding variables precludes identifying clear causal relationships.

A Role for Fate and Transport Models

Given the scarcity of available monitoring data, a major objective of US efforts to evaluate P site assessment tools has been

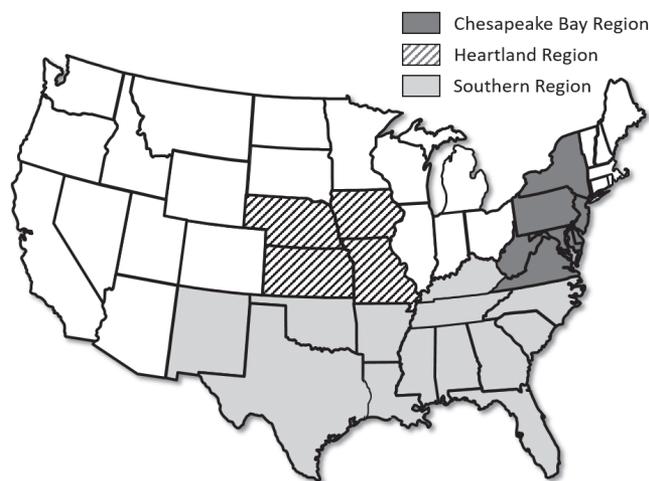


Fig. 1. The National P Index research assessment comprised assessments by the Chesapeake Bay, Heartland, and Southern regions of the United States.

to use fate and transport models to generate edge-of-field data (Sommerlot et al., 2013; Baffaut et al., 2017; Bhandari et al., 2017; Bolster et al., 2017; Forsberg et al., 2017; Nelson et al., 2017). However, this evaluation only applies for conditions where users have carefully vetted the models with regard to hydrologic and P estimation responses. Comprehensive evaluation of P site assessment tools requires that fate and transport models estimate both the spatial and temporal processes of P loss with an accuracy equal to or greater than that of the P site assessment tools, which requires a variety of models.

Heartland and Southern region teams assessed both calibrated and uncalibrated versions of the Agricultural Policy Environmental eXtender model (APEX), whereas the Chesapeake Bay and Southern region teams used the Agricultural Phosphorus Loss Estimator (APLE; Vadas et al., 2009). Also, Chesapeake Bay and Southern region teams used two different versions of the Soil and Water Assessment Tool (SWAT; Arnold et al., 2012), namely SWAT-Variable Source Area (SWAT-VSA; Easton et al., 2008) and the Texas Best Management Practice Evaluation Tool (TBET; White et al., 2012), respectively. The Soil and Water Assessment Tool-Variable Source Area extends SWAT by incorporating a topographic index into the standard input layers to enhance simulation of saturation runoff, whereas TBET uses a simplified user interface of SWAT to estimate P runoff. With regard to complexity, APEX and SWAT-VSA simulate crop growth, water movement, and nutrient fate and transport continuously throughout the simulation period, at a daily or hourly time-step. The Annual P Loss Estimator is an empirically based, annual time-step model that requires minimal data and expertise to run but does not predict runoff or erosion, and thus these variables must be generated with additional models and entered as inputs into APLE. In each region, APEX0806 version was used, and thus recent advancements in P routines for APEX described in Francesconi et al. (2016) were not used.

Accurately Modeling P Cycling

Fundamental to the use of fate and transport models as a surrogate for edge-of-field data is the accurate representation of processes governing P cycling and transport. In the Chesapeake Bay region for instance, fundamental problems were identified with SWAT P routines (Collick et al., 2016), which are derived from the Erosion-Productivity Impact Calculator (EPIC). The original EPIC P routines (Jones et al., 1984) were based on the addition of soluble mineral fertilizer P, which is immediately plant available for the most part, and thus may not accurately reflect P cycling of animal manures or biosolids, which can slowly release P (Vadas et al., 2005a). Additionally, the P routines did not consider impacts of manure rate on effective depth of interaction at the soil-air interface or of dynamic transition ratios among soil P pools (Vadas et al., 2006).

The effect of the EPIC-derived P routines in simulating short-term processes of incidental transfer, or wash off, on P loss were illustrated by Collick et al. (2016), who compared versions of SWAT containing the conventional P routines of Jones et al. (1984) with versions containing updated P routines using algorithms modified from Vadas et al. (2007). Figures 2a, 2b, and 2c illustrate estimated edge-of-field loss of dissolved P in runoff after the simulated application of manure, using original and revised SWAT P routines. Little difference in dissolved P loss

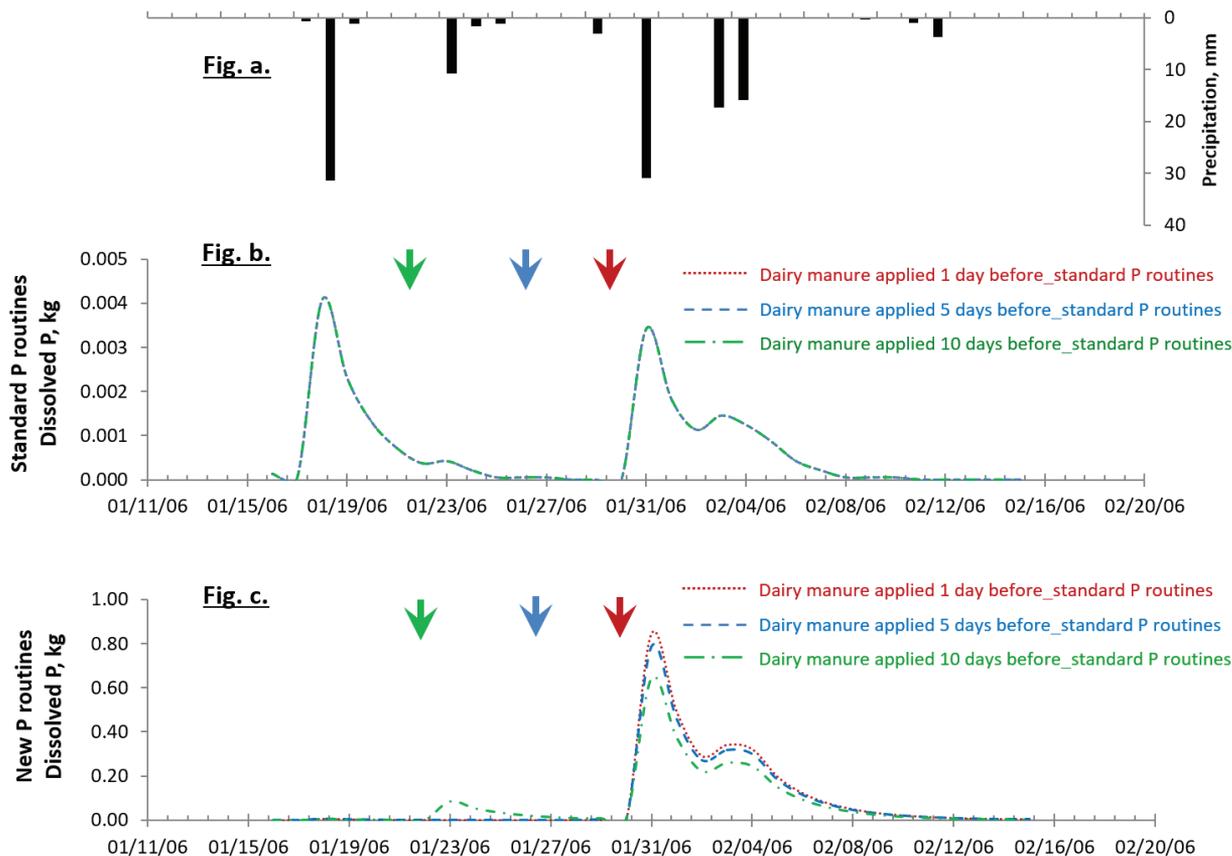


Fig. 2. (a) Simulated effect of precipitation and application timing of manure applied 1 d before a storm (red arrow), 5 d before a storm (blue arrow), or 10 d before a storm (green arrow) on dissolved P loss for (b) standard and (c) new phosphorus (P) routines, adapted with permission from Collick et al. (2016). The new routines were developed as part of US efforts to improve model representation of the effects of manure application timing, rate, method, and source on P fate and transport to enable better comparison with P Index.

as a function of manure application timing relative to rainfall-induced runoff was apparent with the original P routines (Fig. 2b). With revised P routines, estimated dissolved P loss in runoff was sensitive to the timing of recent manure application, particularly in events immediately after manure application.

The findings of Collick et al. (2016) clearly show that use of SWAT without updated P cycling routines may produce correct results, but for the wrong reason. Evaluation efforts of fate-and-transport models do not commonly report event-based output as shown in Fig. 2. However, when models are evaluated with greater temporal resolution, flaws are often revealed in short-term precision, which is important for P site assessment evaluation. Analyses over longer periods, with simulations spanning multiple seasons or years, show that the original routines of Jones et al. (1984) accurately describe manure management effects on average P loss trends. Other process-based models will likely benefit from similar improvements to the P-cycling routines. For example, Forsberg et al. (2017) found that TBET accurately simulated total P loss in runoff, but there were clear indications that it may have been for the wrong reasons. In this case, TBET underpredicted dissolved P losses but overpredicted sediment losses and thus sediment-bound P, resulting in reasonable predictions of total P loss (Forsberg et al., 2017). Similarly, Bolster et al. (2017) showed that relatively poor predictions of total P loss with APLE were due, in part, to poor predictions of runoff and erosion. When using measured runoff and erosion data, model efficiencies for APLE increased from 0.52 to 0.62 for dissolved P and from -0.13 to 0.43 for total P. However,

compared with measured P loads from 31 sites across Arkansas, Georgia, Mississippi, North Carolina, Oklahoma, and Texas, loads estimated by APLE and TBET, which were similar, were almost always higher than APEX estimates (Osmond et al., 2017).

Accurately Modeling Hydrology

Fate and transport models provide sophisticated representation of hydrology, a strength when compared with simpler tools such as the P Index. For example, in the Heartland region, APEX performance was acceptable in terms of runoff simulation, even without calibration (Baffaut et al., 2017). Agricultural Policy Environmental eXtender and conventional versions of SWAT generate surface runoff using routines that describe “infiltration excess” runoff. For low-permeability soils, such as those found in the Heartland region, these routines correctly represent the driving hydrologic processes. In sloping landscapes with high-permeability soils underlain by a restrictive layer, variable source area (VSA) hydrology predominates (i.e., “saturation excess” runoff generation) and, considering only infiltration excess, can incorrectly capture surface runoff potential between fields and thus mischaracterize P loss potential (Easton et al., 2008). By incorporating the effect of topography and how it influences the soil water content, SWAT-VSA more correctly predicts critical P source areas (Easton et al., 2008; Collick et al., 2014; Woodbury et al., 2014).

The Chesapeake Bay region used SWAT-VSA to differentiate between areas of the landscape where saturation and infiltration excess processes predominate (Easton et al., 2008). Collick

et al. (2014) found that SWAT-VSA substantially improved the representation of P loss at field scales over conventional forms of SWAT, allowing for delineation of critical source areas of P loss. Fuka et al. (2016) showed that accounting for the influence of field-scale topography on soil properties in SWAT resulted in more accurate field characterization and further improved model predictions of P loss.

Despite good runoff simulation, issues were uncovered in APEX with the routing of flow through buffers. Although APEX adequately described runoff without site-specific calibration (Baffaut et al., 2017), a regional calibration developed by Nelson et al. (2017) improved runoff simulation from multiple management practices, landscapes, and soil conditions. Others could use this regional calibration to evaluate simplified runoff models that could be included within P Indices.

Accurately Modeling Erosion

Erosion estimation also represented a consistent source of error in predicting P loss. For example, APEX frequently failed to simulate edge-of-field sediment loss. Possible reasons included difficulties in calibrating the model under low-sediment conditions and an inability to simulate erosion processes beyond sheet and rill erosion (Bhandari et al., 2017; Forsberg et al., 2017; Nelson et al., 2017). Prior versions of the RUSLE2 (Foster et al., 2001; Foster, 2013) have been found to overestimate soil erosion, especially from pastures. This overestimation of sediment was due to low biomass estimates in RUSLE2 crop management routines (Dabney and Yoder, 2012).

Accurately Modeling P in Artificial Drainage

In the United States, research has implicated artificial drainage (both open ditches and subsurface tile lines) in many high-profile cases of eutrophication (e.g., Lake Champlain, Chesapeake Bay, Lake Erie, and the Lower Mississippi River Valley). Unlike surface-runoff P loss, field P loss includes subsurface flow pathways (Kleinman et al., 2015). Subsurface drainage systems detour often-sizable portions of the overland flow, particularly in the spring, decreasing high-erosivity events. Critical source areas heavily influence P transport in these systems, with as much as 80% of the total P loss coming from 25% of the watershed area (Ghebremichael et al., 2010; Ghebremichael and Watzin, 2011). Radcliffe et al. (2015) reviewed subsurface P transport routines in major fate and transport models, finding all of them deficient, particularly in their representation of P transport via macropore flow, the dominant pathway for subsurface P transport that serves to bypass native subsoil P sinks. Although long-term simulations with watershed-scale models can obscure such problems, the deficiencies in subsurface P transport routines are manifest when we apply fate and transport models at the scales required of P site assessment tools (i.e., field to subfield).

Updating P Site Assessment Tools

Field Metrics

It can be very difficult to translate some common management factors into appropriate model parameters that ensure comparisons between model output and P Index calculations are directionally and magnitudinally correct (Vadas et al., 2005b; Vadas and Kleinman, 2006). For example, P Indices use many

different extraction methods to represent soil-test P. However, there is no direct way to relate a soil-test P concentration to soil P pools used in models, which describe P flux among soluble, labile, and stable pools. This makes it difficult to ensure that the model and P Index are using the same soil P concentration. It may also partially explain why most P Indices used in the southern United States were better correlated with measured P than the three models evaluated in the Southern region (i.e., APEX, APLE, and TBET; Osmond et al., 2017).

In Arkansas, information gathered from the farm nutrient management planning process between 2004 and 2015 was used to compare risk assigned by the original and revised P Indices (Sharpley et al., 2012). Revision of the Index involved increasing the sensitivity of runoff P concentration to soil-test P using information from plot-scale, simulated rainfall–runoff research and assigning increased risk to P applications made during the rainy spring and early summer months (e.g., March–June). In addition, the revised index now considers the mineralization of manure to provide additional source P after land application. Between 2004 and 2015, researchers assessed 18,300 fields, with assigned risk using the revised index being, on average, 1.1 times greater than the original index (Fig. 3). The average risk assigned over the 12-yr period was 66 for the original and 83 for the revised index, which changes the average site interpretation from a “medium” to “high” risk category and reduces the associated allowable manure application from an N- to P-based rate (i.e., typically 7–3 Mg broiler litter ha⁻¹).

Although not universal among the three regions, strong interest exists in certain areas of the Chesapeake Bay region to revise state P Indices, as evidenced by the recent revision of the Maryland P Management Tool (McGrath et al., 2015; Maryland Department of Agriculture, 2016). The Chesapeake Bay region team established stakeholder groups for each physiographic province in the region, surveyed stakeholders in Delaware, New York, Pennsylvania, and West Virginia, and used feedback from these meetings and surveys to identify site conditions and practices of priority concern. Stakeholders recommended that, where landscape properties indicated high runoff and transport risks, irrespective of nutrient source management, the application of manure should be discouraged and adoption of risk-reducing conservation practices encouraged (e.g., cover crops, manure

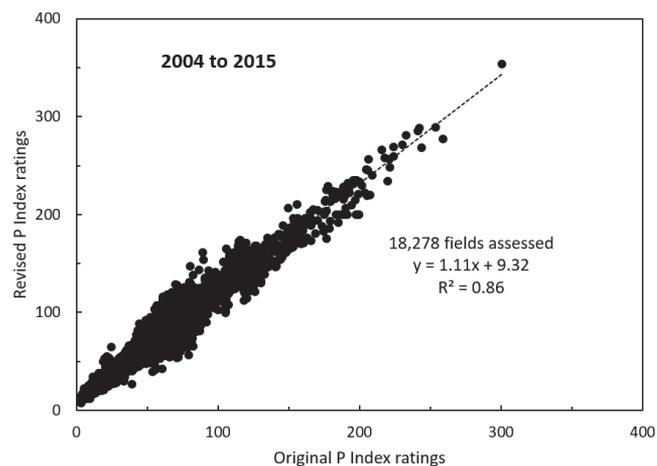


Fig. 3. Relationship between the risk of phosphorus (P) loss assigned by original and revised indices for Arkansas, based on nutrient management planning information for 2004 to 2015.

incorporation or injection, application setbacks, and vegetative buffers) (Cela et al., 2016). Ultimately, the New York stakeholder surveys, along with analysis of nutrient management planner-supplied database of >33,000 fields, led to a proposed revision of the structure of the New York P Index (Ketterings et al., 2017).

Using Fate and Transport Models as Site Assessment Tools

The potential of fate and transport models to provide accurate estimates of P loss for multiple locations, climates, and scenarios has led to widespread interest in evaluating water quality policy and program outcomes at a watershed scale (USEPA, 2010; Whittaker et al., 2015). Indeed, experience exists with “quantitative” P Indices, which leverage fate and transport to predict edge-of-field runoff under different management and site conditions (White et al., 2010; Good et al., 2012; Osmond et al., 2017). In these situations, nutrient management planners who lack expertise in fate and transport modeling use stripped-down versions of fate and transport models for site assessment.

In every region (Chesapeake Bay, Heartland, and South), fundamental problems were identified in applying fate and transport models to areas where they had neither been carefully calibrated nor corroborated. Specifically, parameterization and calibration of SWAT, TBET, and APEX was fraught with uncertainty and was time consuming, whereas APLE was less time consuming but did not provide all the options available with SWAT and APEX (Bolster et al., 2017). Circumventing or avoiding the calibration process, or relying on expert opinion to set model parameters, resulted in poor model estimates for P and sediment in several applications (White et al., 2010; Baffaut et al., 2017; Forsberg et al., 2017). Clearly, caution is required when applying calibrated models outside the systems (e.g., soils, geography, climate, and management) used during calibration and evaluation (Fig. 4; Bhandari et al., 2017; Nelson et al., 2017). Care is also needed to correctly use and interpret terminology for the various forms of soil and water P (Haygarth and Sharpley, 2000).

Implications to National P Site Assessment Approaches

Given the diversity of P site assessment tools used in the United States, there is persistent interest in developing a single, national approach to P site assessment. There is a large body of research conducted over the last 20 yr that has assessed the impacts of nutrient and land management on P runoff at a field scale, addressing all physical characteristics, management combinations, and spatial and temporal scales. Despite this, there were insufficient data across the Southern, Chesapeake, and Heartland regions to evaluate indices with sufficient technical rigor using edge-of-field data alone.

One of the most consistent findings of P site assessment efforts was that, at present, there is no scientific justification to implement a single, national P Index in the United States. The differences in regional and statewide nutrient and land management priorities, landscape properties, climatic regimes, and dominant hydrologic process were so great as to render any attempt at a single, national P Index exceedingly difficult. Fate and transport models remain research tools that are not yet capable of providing accurate estimates of P loss under the diverse set of management scenarios and locations necessary to test or replace the current state-by-state

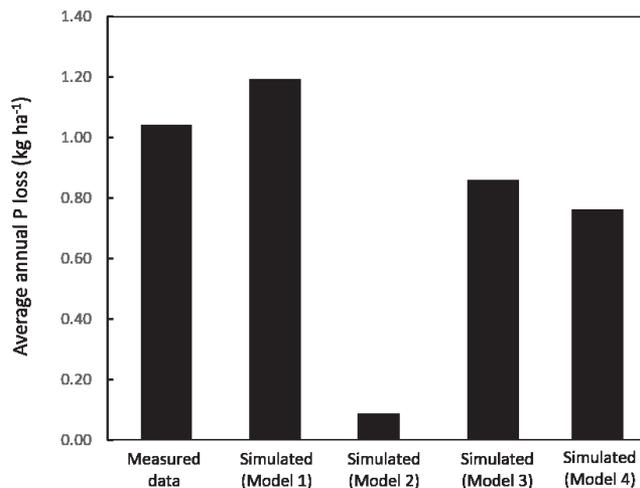


Fig. 4. Average annual total phosphorus (P) loss from a no-till sorghum [*Sorghum bicolor* (L.) Moench.]–soybean [*Glycine max* (L.) Merr.] rotation with surface-applied P fertilizer compared with Annual P Loss Estimator (APEX) model simulation results from models calibrated using datasets with management similar or different from the measured data. Model 1: APEX calibrated with data from the same management system as measured data. Model 2: APEX calibrated with data from conventional-till sorghum–soybean rotation with incorporated P fertilizer. Model 3: APEX calibrated with data from no-till sorghum–soybean rotation with subsurface-applied P fertilizer. Model 4: APEX calibrated with all data used for Models 1, 2, and 3 (Bhandari et al., 2017).

system of P Indices. For instance, results from this assessment suggest that southern P Indices are just as robust as the harder-to-use fate and transport models (Osmond et al., 2017). In addition, several states have implemented revised indices in the last 5 yr: for example, Maryland (Maryland Department of Agriculture, 2016), Arkansas (DeLaune et al., 2006; Sharpley et al., 2010), Kentucky (Bolster et al., 2014), Tennessee (Walker and Hawkins, 2016), and Texas (White et al., 2012). All the revised indices have provided more restrictive nutrient management for similar site conditions than their preceding versions.

Another consistent conclusion is that there are still limitations to the predictive capability of models regarding the effects of management practices on P runoff. The use of fate and transport models to estimate the impacts of agricultural management on P and sediment loss in runoff, including accessing data, parameterizing models, and evaluating estimates, is a labor intensive and a complex process. Although it is possible to run a model with a limited understanding of the input data and without calibration, this research confirmed that such practices result in poor water quality predictions and undermine the validity of the resulting model outcomes and subsequent recommendations.

Our assessment of the models should not, however, be taken as a lack of confidence in the potential of these models to contribute to a better understanding of P loss from agricultural systems. Instead, our results emphasize that progress in understanding water quality from small agricultural fields remains a “three-legged stool” approach. Future success requires continued support for:

1. Collection of soil, water, and land management data through field-scale watershed studies.
2. Application of models to extend the conclusions of measured data, particularly through variations in climate and management, and to provide broader understandings of individual and combined uncertainties in the natural system.

3. Updating of existing process-based models on the basis of experimental knowledge, particularly for sediment loss and P transport processes at a field scale.

Finally, there remains a critical disconnect between assigned P loss risk and biological response of any given receiving water resource (Hirt, 2016). Including site-specific variables that account for the sensitivity or biological response of a water body to P inputs into the indexing framework adds a complexity that few states have been willing to adopt. A first step to overcome these challenges would be to base the P Indexing framework on ecoregions. However, state agencies and institutions establish, approve, and legislate P Indices, making it technically and politically problematic to move ownership to ecoregion, physiographic regions, or watershed boundaries.

Acknowledgments

This research was supported by four USDA-NRCS Conservation Innovation Grants to assess the accuracy of P Indices to determine the risk of P runoff, to identify shortcomings, and to provide options to refine and improve indices. Mention of trade names or commercial products in this publication is solely to provide specific information and does not imply recommendation or endorsement by the USDA. The USDA is an equal opportunity provider and employer.

References

Alexander, R.B., R.A. Smith, G.E. Schwarz, E.W. Boyer, J.V. Nolan, and J.W. Brakebill. 2008. Differences in phosphorus and nitrogen delivery to the Gulf of Mexico from the Mississippi River. *Environ. Sci. Technol.* 42:822–830. doi:10.1021/es0716103

Arnold, J.G., D. Moriasi, P.W. Gassman, K.C. Abbaspour, M.J. White, M.S. Srinivasan et al. 2012. SWAT: Model use, calibration, and validation. *Trans. ASABE* 55:1491–1508. doi:10.13031/2013.42256

Baffaut, C., N.O. Nelson, J.A. Lory, G.M.M.M.A. Senaviratne, A.B. Bhandari, R.P. Udawatta et al. 2017. Multisite evaluation of APEX for water quality: 1. Best professional judgment parameterization. *J. Environ. Qual.* doi:10.2134/jeq2016.06.0226

Bhandari, A.B., N.O. Nelson, D.W. Sweeney, C. Baffaut, J.A. Lory, G.M.M.M.A. Senaviratne et al. 2017. Calibration of the APEX model to simulate management practice effects on runoff, sediment, and phosphorus loss. *J. Environ. Qual.* doi:10.2134/jeq2016.07.0272

Bolster, C.H., A. Forsberg, A. Mittelstet, D.E. Radcliffe, D. Storm, J. Ramirez-Avila et al. 2017. Comparing an annual and daily time-step model for predicting field-scale phosphorus loss. *J. Environ. Qual.* doi:10.2134/jeq2016.04.0159

Bolster, C.H., T. Horvath, B.D. Lee, S. Mehlhope, S. Higgins, and J.A. Delgado. 2014. Development and testing of a new phosphorus index for Kentucky. *J. Soil Water Conserv.* 69:183–196. doi:10.2489/jswc.69.3.183

Cela, S., Q.M. Ketterings, K.J. Czymmek, J. Weld, D.B. Beegle, and P.J.A. Kleinman. 2016. Nutrient management planners' feedback on New York and Pennsylvania phosphorus indices. *J. Soil Water Conserv.* 71:281–288. doi:10.2489/jswc.71.4.281

Chesapeake Bay Program. 2013. Bay barometer 2011–2012. Spotlight on health and restoration in the Chesapeake Bay and its watershed. Chesapeake Bay Program, Annapolis, MD. http://www.chesapeakebay.net/documents/10-Pg_CBP_Barometer_for_web.pdf (accessed 1 May 2017).

Collick, A.S., D.R. Fuka, P.J.A. Kleinman, A.R. Buda, J.L. Weld, M.J. White et al. 2014. Predicting phosphorus dynamics in complex terrains using a variable source area hydrology model. *Hydrol. Processes* 29:588–601. doi:10.1002/hyp.10178

Collick, A.S., T.L. Veith, D.R. Fuka, P.J.A. Kleinman, A.R. Buda, J.L. Weld et al. 2016. Improved simulation of edaphic and manure phosphorus loss in SWAT. *J. Environ. Qual.* 45:1215–1225. doi:10.2134/jeq2015.03.0135

Dabney, S.M., and D.C. Yoder. 2012. Improved description of herbaceous perennial growth and residue creation for RUSLE2. *Agron. J.* 104:771–784. doi:10.2134/agronj2011.0356

Dale, V.H., C.L. Kling, J.L. Meyer, J. Sanders, H. Stallworth, T. Armitage et al. 2010. Hypoxia in the Northern Gulf of Mexico. Springer Series on Environmental Management. Springer Sci., New York. doi:10.1007/978-0-387-89686-1

DeLaune, P.B., B.E. Haggard, T.C. Daniel, I. Chaubey, and M.J. Cochran. 2006. The Eucha/Spavinaw Phosphorus Index: A court mandated index for litter management. *J. Soil Water Conserv.* 61:96–105.

Dubrovsky, N.M., K.R. Burow, G.M. Clark, J.M. Gronberg, P.A. Hamilton, K.J. Hitt et al. 2010. The quality of our nation's waters—Nutrients in the nation's streams and groundwater, 1992–2004. USGS Circ. 1350. USGS. <http://water.usgs.gov/nawqa/nutrients/pubs/circ1350> (accessed 1 May 2017).

Easton, Z.M., D.R. Fuka, M.T. Walter, D.M. Cowan, E.M. Schneiderman, and T.S. Steenhuis. 2008. Re-conceptualizing the Soil and Water Assessment Tool (SWAT) model to predict runoff from variable source areas. *J. Hydrol.* 348:279–291. doi:10.1016/j.jhydrol.2007.10.008

Forsberg, T.A., D.E. Radcliffe, C.H. Bolster, A. Mittelstet, D.E. Storm, and D.L. Osmond. 2017. Evaluation of the TBET model for potential improvement of southern P Indices. *J. Environ. Qual.* doi:10.2134/jeq2016.06.0210

Foster, G.R. 2013. Science documentation user's guide version 2 RUSLE2. USDA–ARS, Washington, DC. http://www.ars.usda.gov/sp2UserFiles/Place/60600505/RUSLE/RUSLE2_Science_Doc.pdf (accessed 1 May 2017).

Foster, G.R., D.C. Yoder, G.A. Weeseis, and T.J. Toy. 2001. The design philosophy behind RUSLE2: Evolution of an empirical model. In: J.C. Ascough, II, and D.C. Flanagan, editors, Soil erosion research for the 21st century. Proceedings of the ASAE International Symposium, Honolulu, HI. 3–5 Jan. 2001. ASAE, St. Joseph, MI. p. 95–98.

Francesconi, W., C.O. Williams, D.R. Smith, J.R. Williams, and J. Jeong. 2016. Phosphorus modeling in tile drained agricultural systems using APEX. *J. Fertil. Pestic.* 7:166. doi:10.4172/2471-2728.1000166

Fuka, D.R., A.S. Collick, P.J.A. Kleinman, D. Auerbach, D. Harmel, and Z.M. Easton. 2016. Improving the spatial representation of soil properties and hydrology using topographically derived initialization processes in the SWAT model. *Hydrol. Processes* 30:4633–4643. doi:10.1002/hyp.10899

Ghebremichael, L.T., T.L. Veith, and M.C. Watzin. 2010. Determination of critical source areas for phosphorus loss: Lake Champlain Basin, Vermont. *Trans. ASABE* 53:1595–1604.

Ghebremichael, L.T., and M.C. Watzin. 2011. Identifying and controlling critical sources of farm phosphorus imbalances for Vermont dairy farms. *Agric. Syst.* 104:551–561. doi:10.1016/j.agsy.2011.04.004

Good, L.W., P.A. Vadas, J.C. Panuska, C.A. Bonilla, and W.E. Jokela. 2012. Testing the Wisconsin phosphorus index with year-round, field-scale runoff monitoring. *J. Environ. Qual.* 41:1730–1740. doi:10.2134/jeq2012.0001

Harmel, R.D., R.J. Cooper, R.M. Slade, R.L. Haney, and J.G. Arnold. 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Trans. ASABE* 49:689–701.

Harmel, D., S. Qian, K. Reckhow, and P. Casebolt. 2008. The MANAGE database: Nutrient load and site characteristic updates and runoff concentration data. *J. Environ. Qual.* 37:2403–2406. doi:10.2134/jeq2008.0079

Haygarth, P.M., and A.N. Sharpley. 2000. Terminology for phosphorus transfer. *J. Environ. Qual.* 29:10–15. doi:10.2134/jeq2000.00472425002900010002x

Hirt, C.R. 2016. Stream condition and nutrient runoff: Linking the Soil and Water Assessment Tool (SWAT) model with empirical ecological measures in an agricultural watershed in central Pennsylvania. Master's thesis, The Pennsylvania State Univ., State College, PA.

Jones, C.A., C.V. Cole, and A.N. Sharpley. 1984. A simplified soil and plant phosphorus model. I. Documentation. *Soil Sci. Soc. Am. J.* 48:800–805. doi:10.2136/sssaj1984.03615995004800040020x

Ketterings, Q.M., S. Cela, A. Collick, S. Crittenden, and K.J. Czymmek. 2017. Restructuring the P Index to better address P management in New York. *J. Environ. Qual.* doi:10.2134/jeq2016.05.0185

Kleinman, P.J.A., D.R. Smith, C.H. Bolster, and Z.M. Easton. 2015. Phosphorus fate, management and modeling in artificially drained systems. *J. Environ. Qual.* 44:460–466. doi:10.2134/jeq2015.02.0090

Maryland Department of Agriculture. 2016. The agricultural phosphorus initiative. Maryland Dep. of Agric. <http://mda.maryland.gov/Pages/PMT.aspx> (accessed 1 May 2017).

McGrath, J.M., F.J. Coale, and N.M. Fiorellino. 2015. University of Maryland Phosphorus Management Tool: Technical users guide. Ext. Bull. EB-405. Univ. of Maryland Ext. and The Dep. of Environ. Sci. and Technol. http://extension.umd.edu/sites/default/files/_docs/articles/EB-405%20UMD%20Phosphorus%20Management%20Tool-Technical%20Users%20Guide.pdf (accessed 1 May 2017).

Michalak, A.M., E.J. Anderson, D. Beletsky, S. Boland, N.S. Bosch, T.B. Bridgeman et al. 2013. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proc. Natl. Acad. Sci. U. S. A.* 110:6448–6452. doi:10.1073/pnas.1216006110

- Nelson, N.O., C. Baffaut, J.A. Lory, G.M.M.A. Senaviratne, A.B. Bhandari, R.P. Udawatta et al. 2017. Multisite evaluation of APEX for water quality: II. Regional parameterization. *J. Environ. Qual.* doi:10.2134/jeq2016.07.0254
- Osmond, D.L., C. Bolster, A.N. Sharpley, M. Cabrera, S. Feagley, A. Forsberg et al. 2017. Southern Phosphorus Indices, water quality data, and modeling (APPEC, APLE, and TBET) results: A comparison. *J. Environ. Qual.* doi:10.2134/jeq2016.05.0200
- Osmond, D.L., A.N. Sharpley, C. Bolster, M. Cabrera, S. Feagley, B. Lee et al. 2012. Comparing phosphorus indices from twelve southern USA states against monitored phosphorus loads from six prior southern studies. *J. Environ. Qual.* 41:1741–1750. doi:10.2134/jeq2012.0013
- Radcliffe, D.E., D.K. Reid, K. Blomback, C.H. Bolster, A.S. Collick, Z.M. Easton et al. 2015. Applicability of models to predict phosphorus losses in drainage fields: A review. *J. Environ. Qual.* 44:614–628. doi:10.2134/jeq2014.05.0220
- Rebich, R.A., N.A. Houston, S.V. Mize, D.K. Pearson, P.B. Ging, and C.E. Hornig. 2011. Sources and delivery of nutrients to the northwestern Gulf of Mexico from streams in the south-central United States. *J. Am. Water Resour. Assoc.* 47:1061–1086. doi:10.1111/j.1752-1688.2011.00583.x
- Reddy, K.R., S. Newman, T.Z. Osborne, J.R. White, and H.C. Fitz. 2011. Phosphorus cycling in the Greater Everglades ecosystem: Legacy phosphorus implications for management and restoration. *Crit. Rev. Environ. Sci. Technol.* 41:149–186. doi:10.1080/10643389.2010.530932
- Scavia, D., J.D. Allan, K.K. Arend, S. Bartell, D. Beletsky, N.S. Bosch et al. 2014. Assessing and addressing the re-eutrophication of Lake Erie: Central basin hypoxia. *J. Great Lakes Res.* 40:226–246. doi:10.1016/j.jglr.2014.02.004
- Sharpley, A.N., D.G. Beegle, C. Bolster, L.W. Good, B. Joern, Q. Ketterings et al. 2012. Phosphorus Indices: Why we need to take stock of how we are doing. *J. Environ. Qual.* 41:1711–1718. doi:10.2134/jeq2012.0040
- Sharpley, A.N., P.A. Moore, Jr., K. VanDevender, M. Daniels, W. Delp, B. Haggard et al. 2010. The Arkansas Phosphorus Index. Fact Sheet FSA 9531. Coop. Ext. Serv., Div. of Agric., Univ. of Arkansas. http://www.uaex.edu/Other_Areas/publications/PDF/FSA-9531.pdf (accessed 1 May 2017).
- Sharpley, A.N., J.L. Weld, D.B. Beegle, P.J.A. Kleinman, W.J. Gburek, P.A. Moore, Jr., and G. Mullins. 2003. Development of Phosphorus Indices for nutrient management planning strategies in the United States. *J. Soil Water Conserv.* 58:137–152.
- Sommerlot, A.R., A.P. Nejadhashemi, and S.A. Woznicki. 2013. Evaluating the impact of field-scale management strategies on sediment transport to the watershed outlet. *J. Environ. Manage.* 128:735–748. doi:10.1016/j.jenvman.2013.06.019
- USDA and USEPA. 1999. Unified national strategy for animal feeding operations. March 9, 1999. U. S. Gov. Print. Off., Washington, DC.
- USDA-NRCS. 2011. Conservation practice standard, nutrient management 590. USDA-NRCS. http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1046177.pdf (accessed 1 May 2017).
- USEPA. 2010. Draft Chesapeake Bay phase 5.3; Community watershed model. EPA 903S10002- CBP/TRS-303-10. USEPA, Chesapeake Bay Prog. Off., Annapolis MD. <http://ches.communitymodeling.org/models/CB-Phase5/documentation.php> (accessed 1 May 2017).
- USEPA. 2015. Watershed assessment, tracking & environmental results. National probable sources contribution to impairments. USEPA. http://iaspub.epa.gov/waters10/attains_nation_cy.control#prob_source (accessed 1 May 2017).
- USEPA-Scientific Advisory Board. 2011. Review of EPA's draft approaches for deriving numeric nutrient criteria for Florida's estuaries, coastal waters, and southern inland flowing waters. USEPA Scientific Advisory Board. [http://yosemite.epa.gov/sab/sabproduct.nsf/DCC3488B67473BDA852578D20058F3C9/\\$File/EPA-SAB-11-010-unsigned.pdf](http://yosemite.epa.gov/sab/sabproduct.nsf/DCC3488B67473BDA852578D20058F3C9/$File/EPA-SAB-11-010-unsigned.pdf) (accessed 1 May 2017).
- Vadas, P.A., W.J. Gburek, A.N. Sharpley, P.J.A. Kleinman, P.A. Moore, Jr., M.L. Cabrera, and R.D. Hamel. 2007. A model for phosphorus transformation and runoff loss for surface-applied manures. *J. Environ. Qual.* 36:324–332. doi:10.2134/jeq2006.0213
- Vadas, P.A., L.W. Good, P.A. Moore, Jr., and N. Widman. 2009. Estimating phosphorus loss in runoff from manure and fertilizer for a phosphorus loss quantification tool. *J. Environ. Qual.* 38:1645–1653. doi:10.2134/jeq2008.0337
- Vadas, P.A., B.E. Haggard, and W.J. Gburek. 2005a. Predicting dissolved phosphorus in runoff from manured field plots. *J. Environ. Qual.* 34:1347–1353. doi:10.2134/jeq2004.0424
- Vadas, P.A., and P.J.A. Kleinman. 2006. Effect of methodology in estimating and interpreting water-extractable phosphorus in animal manures. *J. Environ. Qual.* 35:1151–1159. doi:10.2134/jeq2005.0332
- Vadas, P.A., P.J. Kleinman, A.N. Sharpley, and B.L. Turner. 2005b. Relating soil phosphorus to dissolved phosphorus in runoff: A single extraction coefficient for water quality modeling. *J. Environ. Qual.* 34:572–580. doi:10.2134/jeq2005.0572
- Vadas, P.A., T. Krogstad, and A.N. Sharpley. 2006. Modeling phosphorus transfer between labile and nonlabile soil pools: Updating the EPIC model. *Soil Sci. Soc. Am. J.* 70:736–743. doi:10.2136/sssaj2005.0067
- Veith, T.L., S.C. Goslee, D.B. Beegle, J.L. Weld, and P.J.A. Kleinman. 2017. Analyzing within-county hydrogeomorphic characteristics across Pennsylvania as a precursor to Phosphorus Index modifications. *J. Environ. Qual.* doi:10.2134/jeq2016.10.0416
- Veith, T.L., J.E. Richards, S.C. Goslee, A.S. Collick, R.B. Bryant, D.A. Miller et al. 2015. Navigating spatial and temporal complexity in developing a long-term land use database for an agricultural watershed. *J. Soil Water Conserv.* 70:288–296. doi:10.2489/jswc.70.5.288
- Walker, F., and S. Hawkins. 2016. Guidelines for using the revised Tennessee Phosphorus Risk Index. Fact Sheet W372. Univ. of Tennessee Ext. Inst. of Agric. <https://extension.tennessee.edu/publications/Documents/W372.pdf> (accessed 1 May 2017).
- White, M.J., R.D. Harmel, and R.L. Haney. 2012. Development and validation of the Texas Best Management Practice Evaluation Tool (TBET). *J. Soil Water Conserv.* 67:525–535. doi:10.2489/jswc.67.6.525
- White, M.J., D.E. Storm, P.R. Busteed, M.D. Smolen, H. Zhang, and G.A. Fox. 2010. A quantitative phosphorus loss assessment tool for agricultural fields. *Environ. Model. Softw.* 25:1121–1129. doi:10.1016/j.envsoft.2010.03.017
- Whittaker, G., B.L. Barnhart, R. Srinivasan, and J.G. Arnold. 2015. Cost of areal reduction of gulf hypoxia through agricultural practice. *Sci. Total Environ.* 505:149–153. doi:10.1016/j.scitotenv.2014.09.101
- Woodbury, J., C.A. Shoemaker, D. Cowan, and Z.M. Easton. 2014. Application of SWAT with and without variable source area hydrology to a large watershed. *J. Am. Water Resour. Assoc.* 50:42–56. doi:10.1111/jawr.12116