

The Promise, Practice, and State of Planning Tools to Assess Site Vulnerability to Runoff Phosphorus Loss

P. J. A. Kleinman,* A. N. Sharpley, A. R. Buda, Z. M. Easton, J. A. Lory, D. L. Osmond, D. E. Radcliffe, N. O. Nelson, T. L. Veith, and D. G. Doody

Abstract

Over the past 20 yr, there has been a proliferation of phosphorus (P) site assessment tools for nutrient management planning, particularly in the United States. The 19 papers that make up this special section on P site assessment include decision support tools ranging from the P Index to fate-and-transport models to weather-forecast-based risk calculators. All require objective evaluation to ensure that they are effective in achieving intended benefits to protecting water quality. In the United States, efforts have been underway to compare, evaluate, and advance an array of P site assessment tools. Efforts to corroborate their performance using water quality monitoring data confirms previously documented discrepancies between different P site assessment tools but also highlights a surprisingly strong performance of many versions of the P Index as a predictor of water quality. At the same time, fate-and-transport models, often considered to be superior in their prediction of hydrology and water quality due to their complexity, reveal limitations when applied to site assessment. Indeed, one consistent theme from recent experience is the need to calibrate highly parameterized models. As P site assessment evolves, so too do routines representing important aspects of P cycling and transport. New classes of P site assessment tools are an opportunity to move P site assessment from general, strategic goals to web-based tools supporting daily, operational decisions.

Core Ideas

- Corroboration of decision support tools for controlling phosphorus loss is limited.
- Phosphorus Index is often better related to water quality than fate-and-transport models.
- New site assessment tools for phosphorus must undergo rigorous verification.
- Informed calibration of fate and transport models often improves model performance.
- Care must be taken not to overparameterize fate-and-transport models.

CHANGING management to limit phosphorus (P) losses in runoff from agricultural fields is a complex process in which decisions about crop response, cost, and impact of offsite pollution are but a few of the variables considered. Over the past 20 yr, there has been a proliferation of decision support tools in the nutrient management planning arena intended to provide objective assessment of the potential for P loss in runoff from agricultural fields. The most notable of these tools is the P Index, which has been widely modified and implemented in the United States (Sharpley et al., 2003) and has also been adapted, largely on an experimental or proposed basis, to the physiographic and management conditions of other countries (Heathwaite et al., 2003; Bechmann et al., 2005). Other tools to infer runoff P loss potential and inform field management options have been advanced with varying degrees of attention and adoption (Vadas et al., 2009; White et al., 2012). Evaluation of these tools is extremely complicated, as empirical information and even underlying science are often missing. In this special collection of papers (Fig. 1a), we traverse the state of the science on P site assessment, highlighting a broad range of site assessment approaches, efforts to corroborate and revise P site assessment tools, and findings that shed light on questions surrounding P site assessment.

A Plethora of Options

Although site assessment tools range widely, few are as well documented and widely implemented as the P Index (Lemunyon and Gilbert, 1993). In the United States, implementation of the P Index was prompted by federal law, required widespread research coordination, involved statutory action by state agriculture and environmental departments, and produced legions of consultants and bureaucrats charged with implementing and certifying nutrient management plans developed with the P Index (USDA-USEPA, 1999; Sharpley et al., 2003). Despite this level of activity, the US experience with the P Index has included persistent concerns over differences in P Index formulations and nutrient

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. 46:1243–1249 (2017).

doi:10.2134/jeq2017.10.0395

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 9 Oct. 2017.

Accepted 23 Oct. 2017.

*Corresponding author (peter.kleinman@ars.usda.gov).

P.J.A. Kleinman, A.R. Buda, and T.L. Veith, USDA-ARS, Pasture Systems and Watershed Management Research Unit, University Park, PA 16802; A.N. Sharpley, Dep. Crop, Soil and Environmental Sciences, Univ. of Arkansas, Fayetteville, AR 72701; Z.M. Easton, Dep. Biological Systems Engineering, Virginia Tech, Blacksburg, VA 24061; J.A. Lory, Univ. of Missouri, Division of Plant Sciences, Columbia, MO 65211; D.L. Osmond, Dep. Crop and Soil Sciences, North Carolina State Univ., Raleigh, NC 27695; D.E. Radcliffe, Dep. Crop and Soil Sciences, Univ. of Georgia, Athens, GA 30602; N.O. Nelson, Dep. Agronomy, Kansas State Univ., Manhattan, KS 66506; D.G. Doody, Agrifood and Biosciences Institute, Belfast, BT9-5PQ, UK.

Abbreviations: APEX, Agricultural Policy Environmental eXtender; APLE, Annual Phosphorus Loss Estimator; SWAT, Soil and Water Assessment Tool; TBET, Texas Best Management Practice Evaluation Tool.

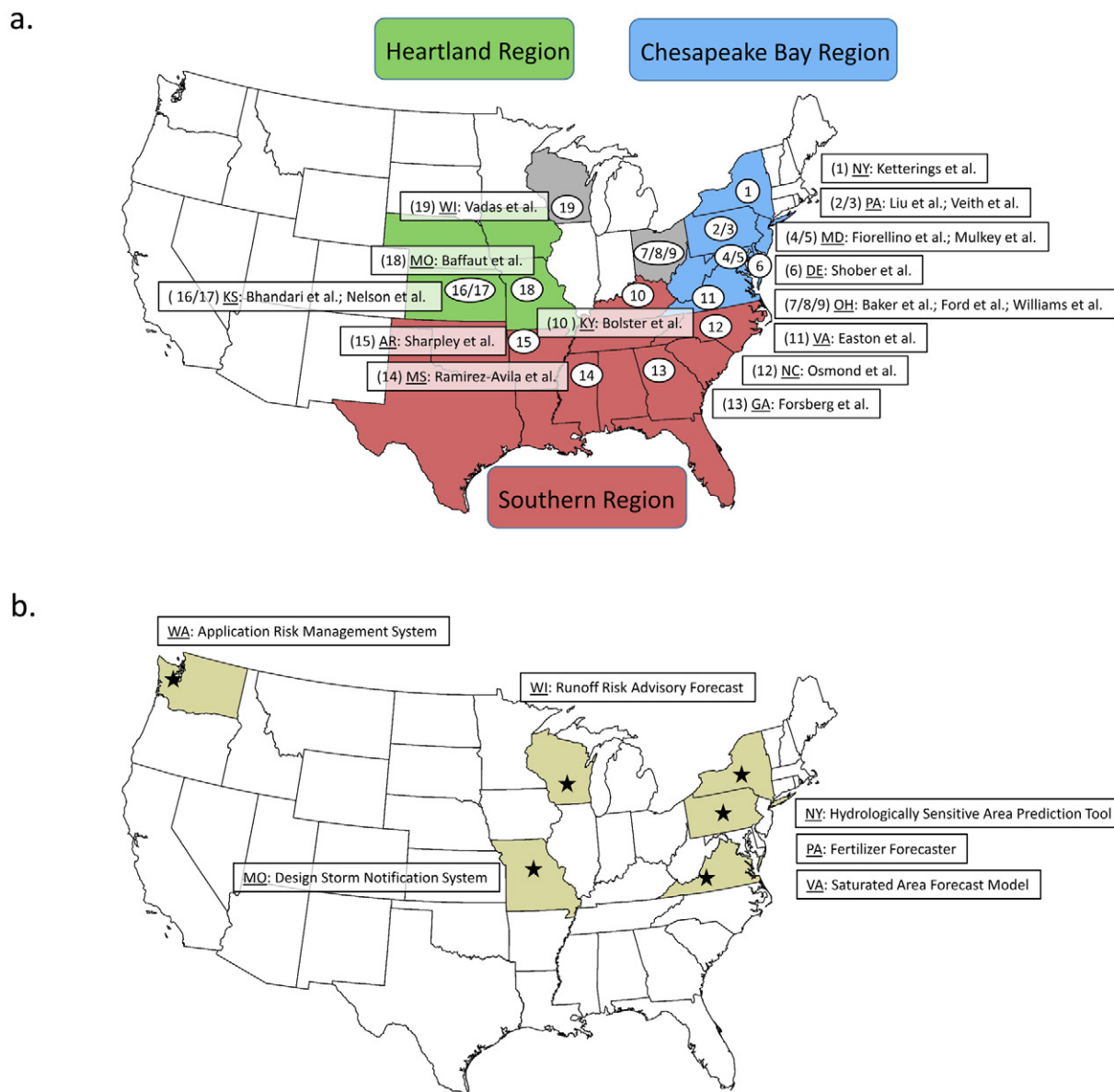


Fig. 1. Geographic distribution of papers on P site assessment included as part of the special issue, including (a) three regional initiatives described under Sharpley et al. (2017), and (b) the location of next-generation decision support tools based on weather forecasts, as reviewed by Easton et al. (2017).

management recommendations between states (Osmond et al., 2006; Sharpley et al., 2012).

There is little question that the P Index has served to educate farmers, policymakers, and others in the implications of nutrient management decisions on water quality. However, there has been mixed evidence of water quality improvement after P Index implementation (Sharpley et al., 2012). In addition, end-user reaction to the usefulness of the strategic recommendations made by the P Index has been mixed—in part due to difficulty in connecting their behavior with downstream water quality outcomes, but also because the P Index typically does not address daily, operational decisions that must react to site conditions and weather patterns (Osmond et al., 2012). Indeed, there has long been interest in developing dynamic P site assessment tools that directly estimate changes in runoff P loadings associated with different field management practices (White et al., 2010). Thus, opportunity has been ripe for new approaches to P site assessment.

Site assessment approaches other than the P Index have been applied, either as an alternative to or as a complement to the P Index assessment process by offering additional nutrient management decision support. Alternatives to the P Index for nutrient management planning often are derived from fate-and-transport models, particularly the Agricultural Policy Environmental eXtender (APEX; Williams and Izaurralde, 2006) and the Soil and Water Assessment Tool (SWAT; Arnold et al., 2012), as well as from empirical models such as the Annual P Loss Estimator (APLE; Vadas et al., 2009). These tools have even been subsumed into the P Index framework. For example, the P Index of White et al. (2010) provides a recognizable interface for SWAT calibrated to watersheds in Oklahoma, and the Wisconsin P Index uses P loss routines from APLE (Good et al., 2012). Other proposals to use fate-and-transport models in nutrient management planning include the Nutrient Trading Tool, an APEX derivative (Saleh et al., 2011). Notably, there has been an absence of P site assessment tools, from indices to process-based models, for

subsurface transport of P from agricultural sources (Kleinman et al., 2015; Shober et al., 2017).

Soil testing can also be adapted to P site assessment for water quality inference. There is long-standing debate over the use of soil P thresholds alone to guide P application rates, in lieu of a full assessment of P source and transport factors, as found in the abovementioned tools. Indeed, in the United States, federal nutrient management policy allows states to manage agricultural P using several soil P testing approaches (agronomic and environmental thresholds), but the majority of states have elected to use versions of the P Index (USDA-USEPA, 1999; Sharpley et al., 2003). In the current issue, Baker et al. (2017) use soil testing to highlight the extent to which soil P is vertically stratified in soils, a function of the common practices of surface application of fertilizers and no-till, albeit as part of rotations that periodically included tillage. They provide compelling evidence that sampling for vertical P stratification should be included in nutrient management planning as a check on excessive P buildup near the soil surface, within the “Effective Depth of Interaction” layer between soil and runoff water (Sharpley, 1985). The findings of Baker et al. (2017) point to the need for vertical P stratification to be included in soil health assessment (Idowu et al., 2009).

Elsewhere, a new class of decision support tools is emerging, one that harnesses high-resolution, short-term weather forecasting models to support decisions on when and where to apply nutrients to avoid “incidental transfers” or “wash-off” of recently applied nutrients from the soil surface by runoff water (Preedy et al., 2001; Buda et al., 2013). Easton et al. (2017) review six short-term decision support tools (Fig. 1b), all of which rely on the US NOAA’s National Weather Service forecasts to develop management recommendations: (i) Wisconsin’s Runoff Risk Advisory Forecast System, (ii) Washington’s Application Risk Management System, (iii) Pennsylvania’s Fertilizer Forecaster (Fig. 2), (iv) Virginia’s Saturated Area Forecast Model, (v) New York’s Hydrologically Sensitive Area

Prediction Tool, and (vi) Missouri’s Design Storm Notification System. Although many of these tools are experimental, several have been implemented as part of state nutrient management advisory programs (Missouri and Wisconsin), whereas others are in various stages of β -testing. Certainly, short-term site assessment tools are not intended to replace the P Index but instead are seen as complementary to the longer-term planning horizons of current P site assessment tools.

Phosphorus Site Assessment and Water Quality

The utility of P site assessment tools derives from their ability to inform practical field management decisions over a wide range of conditions and, with time, improve or protect water quality. However, performance of P site assessment tools varies widely, as documented in various comparisons between tools and their recommended actions for the end user (Osmond et al., 2017). Consequently, there have been calls to corroborate P site assessment tools with water quality data, ensuring, at a minimum, that ratings provided by different tools do indeed correspond to risk potential of runoff P loss (USDA-NRCS, 2011). Because edge-of-field water quality monitoring data that would be used to substantiate P site assessment tools are limited in availability and scope, there have been proposals to employ output from fate-and-transport models as verification datasets for simpler tools, such as the P Index (Osmond et al., 2017; Sharpley et al., 2017). Despite their limited availability, edge-of-field water quality datasets have helped reveal striking disparities in the ability of different P Indices to estimate the direction and magnitude of field-scale P loss (Osmond et al., 2012).

The Phosphorus Index and Water Quality

Nowhere has use and testing of the P Index been more concentrated than in the United States, where the P Index has spawned

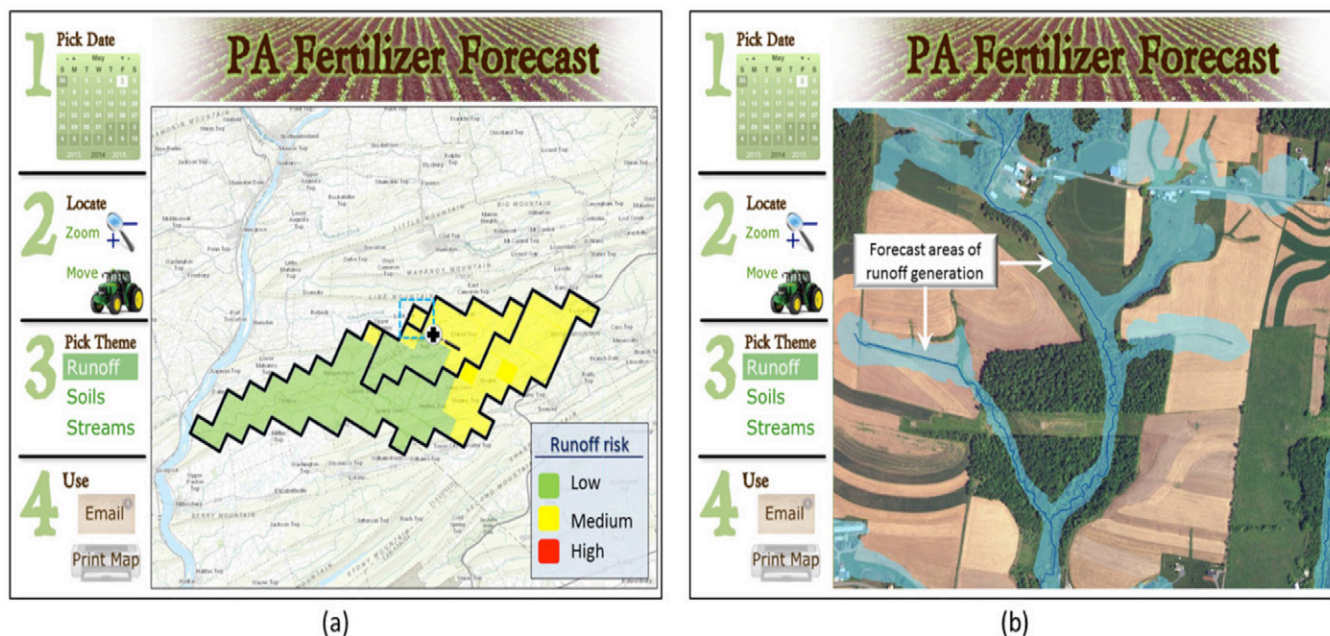


Fig. 2. An example of a short-term site assessment tool, Pennsylvania’s Fertilizer Forecaster, showing (a) forecasted runoff risk levels (low, medium, and high) for 88.2- by 2-km grid cells, and (b) a field-scale view of predicted runoff contributing areas for a moderate risk event. Adapted from Easton et al. (2017).

up to 47 variants as part of state nutrient management programs. A variety of studies have, over the years, demonstrated relationships between different versions of the P Index and runoff from plots under artificial rainfall (Eghball and Gilley, 2001; Sharpley et al., 2001), or from field-scale (Harmel et al., 2005; Butler et al., 2010) or small (Veith et al., 2005) to large (Birr and Mulla, 2001) multifield catchments. These studies have not been widely replicated and often evaluate source factors of the P Index (e.g., rate, method, timing of fertilizer and manure applications) more comprehensively than P Index transport factors (e.g., hydrologic and proximity of field to stream). Thus, Sharpley et al. (2017) review a variety of research activities to respond to concerns raised by the USDA over inconsistencies in state P Indices and the need for systematic verification of those indices to ensure public confidence in their role at the center of nutrient management recommendations for agriculture (USDA-NRCS, 2011).

Three regional programs were developed with the participation of 22 states: the Southern United States (12 states), the Heartland (four states), and the Chesapeake Bay (six states). A unique activity here was to move beyond limited monitoring datasets to investigate the use of fate-and-transport models in generating the edge-of-field water quality data needed to evaluate the P index. In the Southern United States, different state P Indices were evaluated against water quality monitoring data and estimates of water quality generated by several field-scale models (APEX, APLE, and the Texas Best Management Practice Evaluation Tool [TBET]). Osmond et al. (2017) confirm widespread differences in the performance of P Indices of southern states under various benchmark conditions, when P Index ratings of “low,” “moderate,” and “high” are categorized according to the recommendations of USDA-NRCS (2011): low (<2.2 kg total P ha⁻¹ yr⁻¹), moderate (2.2–5.5 kg total P ha⁻¹ yr⁻¹), high (>5.5 kg total P ha⁻¹ yr⁻¹). On the whole, southern P Indices perform remarkably well, correlating as well or better than fate-and-transport models to measured P losses in runoff. Their findings suggest that the P Index, a decision support tool designed for simplicity and compliance with nutrient management regulations, may often provide water quality inferences at the field edge that are comparable with more complex and difficult to use fate-and-transport models.

To date, most P Index assessments have fixated on P losses in surface runoff. Delineating the association between P Index ratings and measures of subsurface P loss is an important yet oft ignored aspect of recent P Index assessments. Williams et al. (2017) seek to address this gap by evaluating the Ohio P Index in the Corn Belt region of western Ohio, where intensive tile drainage greatly amplifies the risk of subsurface P loss to surface waters. Although the Ohio P Index lacks a formal subsurface P component, its predictions of P risk agree reasonably well with average annual P loads in edge-of-field runoff and tile drainage. Still, Williams et al. (2017) acknowledge that P leaching to tile drains is an endemic pathway of subsurface P loss that should be added to future iterations of the Ohio P Index. Although the Ohio P Index is one of 27 indices that does not consider the prospect of subsurface P loss, there are 21 P indices that do possess subsurface P risk components. In a study examining the subsurface P risk routines of five P Indices on the Lower Delmarva Peninsula, Shober et al. (2017) find that simpler, semiquantitative indices provide better estimates of P export in leachate and

ditch drainage than more quantitative P indices based on mechanistic modeling. More importantly, Shober et al. (2017) demonstrate that when the formulations of source and transport factors in these indices are revised according to current knowledge of P fate-and-transport processes on the Delmarva Peninsula, estimation of subsurface P losses by these indices improves.

Fate-and-Transport Models and Water Quality

Fate-and-transport models are usually not intended for general application as P site assessment tools, despite a few successful examples of their transformation (e.g., White et al., 2010). Although models such as APEX, SWAT, APLE, the Agricultural Non-Point Source Pollution model (AGNPS; Young et al., 1989), and the Hydrological Simulation Program Fortran (HSPF; Bicknell et al., 1997) have all been shown to accurately predict P loss in runoff at various scales, their accuracy in predicting the outcomes of various field managements at finer scales is limited. Bolster et al. (2017) compare APLE, a simple, empirically based model, with TBET, a more parameter-intensive model based on SWAT. Neither APLE nor TBET perform consistently better in predicting P loss when compared with edge-of-field monitoring data. Additionally, at these finer scales, key natural system processes and interactions might be better understood, reducing the number of uncertainties that the model needs to capture, but the impact of precision error around each uncertainty in the model or the input data is enhanced (Refsgaard et al., 2007). Accordingly, Baffaut et al. (2017), Bhandari et al. (2017), Bolster et al. (2017), Forsberg et al. (2017), Nelson et al. (2017), and Ramirez-Avila et al. (2017) report that fate-and-transport models can perform well when applied in locations where they have been carefully calibrated and corroborated, but they perform poorly for P and sediment when applied in locations where they have not been adequately calibrated or corroborated. Elsewhere, Bhandari et al. (2017) and Ramirez-Avila et al. (2017) find that calibrated versions of APEX perform better in predicting total P losses than uncalibrated versions of APEX. Nelson et al. (2017) seeks to develop regional parameters to enable the widespread transfer of APEX to different sites. Although regional parameters could be used to accurately predict runoff from field sites, this approach is unsuccessful in predicting sediment and P loss in runoff.

The role of calibration in modeling can be divisive, with one camp arguing that calibration is an imperative step to model implementation and another camp raising concerns over calibration due to the “overparameterization” of models (i.e., that adjustment of too many parameters in a model can obfuscate poor simulation of important processes; Beven, 1993, 2006; Refsgaard and Henriksen, 2004; Baffaut et al., 2017). Experience with fate-and-transport models indicates that these are not mutually exclusive perspectives, confirming the insistence by Refsgaard (1997) that prudent model application entails careful verification of model performance, balancing optimization of performance metrics with parsimonious, defensible adjustment of model parameters. Regardless, general concerns were raised with the prediction of edge-of-field sediment loss in both calibrated and uncalibrated models, with important implications to estimating particulate P loss from fields (Forsberg et al., 2017).

Efforts to Improve Phosphorus Site Assessment

As reviewed by Sharpley et al. (2017), strong impetus exists to revise or update P site assessment tools, either to improve their performance as indicators of water quality improvement, or to bring new models and new understanding of P transport processes to bear. In the Chesapeake Bay region, an array of efforts has been underway to facilitate the revision of P Indices. Cela et al. (2016) surveyed nutrient management planners in New York and Pennsylvania to identify areas in which revisions should be targeted, as well as to prioritize nutrient management practices that should be promoted after P site assessment. To ensure that P Index revisions in Pennsylvania represents all major field conditions found in that state, Veith et al. (2017) perform cluster analysis of publicly available topographic and soils data, identifying key physiographic conditions where new versions of the P Index should be evaluated.

One of the few proposals to update the P Index is described by Ketterings et al. (2017). Their proposal builds on the successful advance of so-called “component” P Indices (i.e., those P Indices that assess dissolved and particulate P loss potentials separately), as well as the aforementioned feedback from nutrient management planners in New York State (Cela et al., 2016). A unique

aspect of the proposed P Index of Ketterings et al. (2017) is the identification of site transport factors that are relatively fixed and therefore difficult to reverse through management (e.g., hydrologically active and connected areas near streams). This approach applies well to the variable source area hydrology of New York’s sloping landscapes, where little can be done to alter the landscape processes dominating surface runoff generation. Real potential exists to apply this framework to other regions as well. Other novel features of the proposed P Index for New York are a strong, overt tie to mitigating practices (a complaint of many P Indices is the indirect nature in which nutrient management recommendations are derived) and a soil P threshold that serves as a ceiling for additional P applications (Table 1).

Considerable success has also been documented in using or applying the APLE model, which was developed from empirical studies on P transport (Vadas et al., 2009), to P site assessment around the Chesapeake Bay region. Fiorellino et al. (2017) describe successful application of the APLE model to assess the Maryland P Index, and Mulkey et al. (2017) apply APLE to improve the identification of agricultural P sources within the Chesapeake Bay Model. Previously, Bolster et al. (2012) demonstrated how to apply APLE to revise the Pennsylvania P Index. Although APLE does not simulate runoff and is only applicable to P loss in surface runoff, its simple format, principal focus on representing P

Table 1. Proposed Phosphorus Index for New York. Adapted from Ketterings et al. (2017).

Overall interpretation (landscape factor score × best management practice score × 10)						
Management implication						
P-loss risk	P Index score	Cornell soil-test P				
		<20 mg kg ⁻¹	20–50 mg kg ⁻¹	51–80 mg kg ⁻¹	>80 mg kg ⁻¹	
Low	<50	N-based	N-based	P-based	Zero	
Medium	50–74	N-based	P-based	Zero	Zero	
High	75–99	P-based	P-based	Zero	Zero	
Very high	≥100	Zero	Zero	Zero	Zero	
Transport factors						
Factor	Option	Coefficient	Factor	Option	Coefficient	
Flow distance to stream or ditch (m)	>153 m (500 ft)	0	Soil drainage class	Well	0	
	>91–153 m (300–500 ft)	6		Moderately well	2	
	>33–91 m (100–300 ft)	9		Somewhat poor	4	
	≤33 m (100 ft)	12		Poor	8	
	Never	0		None	0	
Erosion (Mg ha ⁻¹)	Occasionally	2	Subsurface drainage	Random	1	
	Frequent	5		Pattern	2	
	≤2.24	0		Vegetated buffer	Absent	0
	>2.24–6.73	1			Present	–4
	>6.73–11.2	3			Untreated concentrated flow	No
≥11.2	5	Yes	4			
Best management practice score (method × ground coverage and timing scores)					Coefficient	
Method of application						
Surface spread without setback					1.0	
Surface spread with 33-m (100-ft) setback					0.8	
Surface spread with managed (sod, harvested) vegetative setback					0.7	
Incorporation (within 24 h and 5-m [15-ft] setback)					0.6	
Injection (with 5-m [15-ft] setback)					0.5	
Ground cover and timing						
Bare ground outside 1–2 wk of planting					1.0	
Bare ground within 1–2 wk of planting					0.8	
Dormant sod or cover crop after corn harvest					0.8	
Growing sod or row crop					0.6	

transport alone, and ease of use are factors in its growing integration into P loss decision support tools (Benskin et al., 2014).

Advances in fate-and-transport models continue to enhance our abilities to simulate P transformations and predict P movement under varieties of agricultural conditions. Vadas et al. (2017) and Liu et al. (2017) evaluate several approaches to winter manure spreading guidelines using SurPhos and SWAT, respectively, to project water quality outcomes. In comparing winter spreading bans with more flexible approaches that allow manure application depending on site conditions, they report that both strategies reduce field-edge P losses relative to no restrictions. Over the long-term, Liu et al. (2017) show that winter spreading bans can affect greater reductions in watershed P loads, albeit while periodically exacerbating peaks in P concentrations of watershed discharge. Ford et al. (2017) describe efforts to adapt macropore flow routines in APEX to better simulate preferential P losses to tile drains in the Western Lake Erie Basin. This represents one of the few successful applications of an agriculturally based simulation model to estimate subsurface P loss, an area identified as a priority need in areas with extensive artificial drainage (Radcliffe et al., 2015).

Conclusions

The testing of fate-and-transport models with runoff monitoring data highlights various traps and pitfalls that must be avoided to ensure that these models achieve their potential in predicting phosphorus loss. Despite greater complexity of fate-and-transport models, many recent studies suggest that simple tools such as the P Index perform equally well in representing the potential for P loss from agricultural fields. Although there is debate over the potential for over parameterization of some models to compensate for errors in their simulation of certain processes, experience with using APEX in various settings points to the need for calibration of process-based models if they are to be useful in P site assessment. New versions of the P Index and daily decision support tools all show promise in improving the utility of site assessment tools to agriculture, affecting desired changes in nutrient management. Indeed, the greater availability of high-resolution topographic maps promises to improve delineation of hydrologic flow pathways and designation of hydrologically active areas relative to field P loss. As P site assessment evolves, it is clear that better edge-of-field water quality monitoring databases are needed to ensure consistent evaluation of new tools. Over time, wider documentation of watershed-scale response to the use of decision support tools will help to fuel their implementation and adoption worldwide.

Acknowledgments

The authors are grateful to USDA's Natural Resource Conservation Service for support of many of the studies reported in this special collection.

References

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.A. Senaviratne, A.B. Bhandari, R.P. Udawatta et al. 2017. Multisite evaluation of APEX for water quality: I. Best professional judgment parameterization. *J. Environ. Qual.* 46:1323–1331. doi:10.2134/jeq2016.06.0226

Baker, D.B., L.T. Johnson, R.B. Confesor, and J.P. Crumrine. 2017. Vertical stratification of soil phosphorus as a concern for dissolved phosphorus runoff in the Lake Erie Basin. *J. Environ. Qual.* 46:1287–1295. doi:10.2134/jeq2016.09.0337

Bechmann, M., T. Krogstad, and A. Sharpley. 2005. A phosphorus index for Norway. *Acta Agric. Scand., Sect. B* 55:205–213. doi:10.1080/09064710510029088

Benskin, C.M.W., W.M. Roberts, Y. Wang, and P.M. Haygarth. 2014. Review of the Annual Phosphorus Loss Estimator tool: A new model for estimating phosphorus losses at the field scale. *Soil Use Manage.* 30:337–341. doi:10.1111/sum.12128

Beven, K. 1993. Prophecy, reality, and uncertainty in distributed hydrological modeling. *Adv. Water Resour.* 16:41–51. doi:10.1016/0309-1708(93)90028-E

Beven, K. 2006. A manifesto for the equifinality thesis. *J. Hydrol.* 320:18–36. doi:10.1016/j.jhydrol.2005.07.007

Bhandari, A.B., N.O. Nelson, D.W. Sweeney, C. Baffaut, J.A. Lory, A. Senaviratne et al. 2017. Calibration of the APEX model to simulate management practice effects on runoff, sediment, and phosphorus loss. *J. Environ. Qual.* 46:1332–1340. doi:10.2134/jeq2016.07.0272

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, Jr., A.S. Donigan, Jr., and R.C. Johanson. 1997. Hydrological Simulation Program—Fortran, User's manual for version 11. USEPA, Natl. Exposure Res. Lab., Athens, GA.

Birr, A.S., and D.J. Mulla. 2001. Evaluation of the Phosphorus Index in watersheds at the regional scale. *J. Environ. Qual.* 30:2018–2025. doi:10.2134/jeq2001.2018

Bolster, C.H., A. Forsberg, A. Mittelstet, D. Radcliffe, D. Storm, J. Ramirez-Avila, A.N. Sharpley, and D. Osmond. 2017. Comparing an annual and daily time-step model for predicting field-scale phosphorus loss. *J. Environ. Qual.* 46:1314–1322. doi:10.2134/jeq2016.04.0159

Bolster, C.H., P.A. Vadas, A.N. Sharpley, and J.A. Lory. 2012. Using a P loss model to evaluate and improve P Indices. *J. Environ. Qual.* 41:1758–1766. doi:10.2134/jeq2011.0457

Buda, A.R., P.J.A. Kleinman, R.B. Bryant, G.W. Feyereisen, D.A. Miller, P.G. Knight, and P.J. Drohan. 2013. Forecasting runoff from Pennsylvania landscapes. *J. Soil Water Conserv.* 68:185–198. doi:10.2489/jswc.68.3.185

Butler, D.M., D.H. Franklin, M.L. Cabrera, L.M. Risse, D.E. Radcliffe, L.T. West, and J.W. Gaskin. 2010. Assessment of the GA Phosphorus Index on farm at the field scale for grassland management. *J. Soil Water Conserv.* 65:200–210. doi:10.2489/jswc.65.3.200

Cela, S., Q.M. Ketterings, K.J. Czymmek, J.L. 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

Easton, Z.M., P.J.A. Kleinman, A.R. Buda, D. Goering, N. Emberston, S. Reed et al. 2017. Short-term forecasting tools for agricultural nutrient management. *J. Environ. Qual.* 46:1257–1269. doi:10.2134/jeq2016.09.0377

Eghball, B., and J.E. Gilley. 2001. Phosphorus risk assessment index evaluation using runoff measurements. *J. Soil Water Conserv.* 56:202–206.

Ford, W., III, K. King, M. Williams, and R. Confesor. 2017. Modified APEX model for simulating macropore phosphorus contributions to tile drains. *J. Environ. Qual.* 46:1413–1423. doi:10.2134/jeq2016.06.0218

Fiorellino, N., J. McGrath, P. Vadas, C. Bolster, and F. Coale. 2017. Use of Annual Phosphorus Loss Estimator (APLE) model to evaluate a phosphorus index. *J. Environ. Qual.* 46:1380–1387. doi:10.2134/jeq2016.05.0203

Forsberg, A., D. Radcliffe, C. Bolster, A. Mittelstet, D. Storm, and D. Osmond. 2017. Evaluation of the TBET model for potential improvement of southern P Indices. 46:1341–1348. *J. Environ. Qual.* doi:10.2134/jeq2016.06.0210

Good, L.W., P. 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., H.A. Torbert, P.B. DeLaune, B.E. Haggard, and R.L. Haney. 2005. Field evaluation of three phosphorus indices on new application sites in Texas. *J. Soil Water Conserv.* 60:29–42.

Heathwaite, L., A. Sharpley, and M. Bechmann. 2003. The conceptual basis for a decision support framework to assess the risk of phosphorus loss at the field scale across Europe. *J. Plant Nutr. Soil Sci.* 166:447–458. doi:10.1002/jpln.200321154

Idowu, O.J., H.M. Van Es, G.S. Abawi, D.W. Wolfe, R.R. Schindelbeck, B.N. Moebius-Clune, and B.K. Gugino. 2009. Use of an integrative soil health test for evaluation of soil management impacts. *Renew. Agric. Food Syst.* 24:214–224. doi:10.1017/S1742170509990068

Ketterings, Q.M., S. Cela, A.S. Collick, S.J. Crittenden, and K. Czymmek. 2017. Restructuring the P index to better address P management in New York. *J. Environ. Qual.* 46:1372–1379. 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
- Lemunyon, J.L., and R.G. Gilbert. 1993. The concept and need for a phosphorus assessment tool. *J. Prod. Agric.* 6:483–486. doi:10.2134/jpa1993.0483
- Liu, J., T.L. Veith, A.S. Collick, P.J.A. Kleinman, D.B. Beegle, and R.B. Bryant. 2017. Seasonal manure application timing and storage effects on field and watershed level phosphorus loss. *J. Environ. Qual.* 46:1403–1412. doi:10.2134/jeq2017.04.0150
- Mulkey, A.S., F. Coale, P. Vadas, G. Shenk, and G. Bhatt. 2017. Revised method and outcomes for estimating soil phosphorus losses from agricultural land in the Chesapeake Bay watershed model. *J. Environ. Qual.* 46:1388–1394. doi:10.2134/jeq2016.05.0201
- Nelson, N.O., C. Baffaut, 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: II Regional parameterization. *J. Environ. Qual.* 46:1349–1356. doi:10.2134/jeq2016.07.0254
- Osmond, D., C. Bolster, A. Sharpley, M. Cabrera, S. Feagley, A. Forsberg et al. 2017. Southern phosphorus indices, water quality data, and modeling (APEX, APLE, and TBET) results: A comparison. *J. Environ. Qual.* 46:1296–1305. doi:10.2134/jeq2016.05.0200
- Osmond, D., M. Cabrera, S. Feagley, G. Hardee, C. Mitchell, P. Moore et al. 2006. Comparing southern P indices. *J. Soil Water Conserv.* 61:325–337.
- Osmond, D., D. Meals, D. Hoag, M. Arabi, A. Luloff, G. Jennings et al. 2012. Improving conservation practices programming to protect water quality in agricultural watersheds: Lessons learned from the National Institute of Food and Agriculture–Conservation Effects Assessment Project. *J. Soil Water Conserv.* 67:122A–127A. doi:10.2489/jswc.67.5.122A
- Preedy, N., K. McTiernan, R. Matthews, L. Heathwaite, and P. Haygarth. 2001. Rapid incidental phosphorus transfers from grassland. *J. Environ. Qual.* 30:2105–2112. doi:10.2134/jeq2001.2105
- 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
- Ramirez-Avila, J.J., D.E. Radcliffe, D. Osmond, C. Bolster, A. Sharpley, S.L. Ortega-Achury et al. 2017. Evaluation of the APEX model to simulate runoff quality from agricultural fields in the southern region of the United States. *J. Environ. Qual.* 46:1357–1364. doi:10.2134/jeq2017.07.0258
- Refsgaard, J.C. 1997. Parameterisation, calibration and validation of distributed hydrological models. *J. Hydrol.* 198:69–97. doi:10.1016/S0022-1694(96)03329-X
- Refsgaard, J.C., and H.J. Henriksen. 2004. Modelling guidelines: Terminology and guiding principles. *Adv. Water Resour.* 27:71–82. doi:10.1016/j.advwatres.2003.08.006
- Refsgaard, J.C., J.P. van der Sluijs, A.L. Højberg, and P.A. Vanrolleghem. 2007. Uncertainty in the environmental modelling process: A framework and guidance. *Environ. Model. Softw.* 22:1543–1556. doi:10.1016/j.envsoft.2007.02.004
- Saleh, A., O. Gallego, E. Osei, H. Lal, C. Gross, S. McKinney, and H. Cover. 2011. Nutrient Tracking Tool: A user-friendly tool for calculating nutrient reductions for water quality trading. *J. Soil Water Conserv.* 66:400–410. doi:10.2489/jswc.66.6.400
- Sharpley, A., D. Beegle, C. Bolster, L. Good, B. Joern, Q. Ketterings et al. 2012. Phosphorus Indices: Why we need to take stock of how we are doing. *J. Environ. Qual.* 41:1711–1719. doi:10.2134/jeq2012.0040
- Sharpley, A., P. Kleinman, C. Baffaut, D. Beegle, C. Bolster, A. Collick et al. 2017. Evaluation of phosphorus site assessment tools: Lessons from the USA. *J. Environ. Qual.* 46:1250–1256. doi:10.2134/jeq2016.11.0427
- Sharpley, A.N. 1985. Depth of surface soil-runoff interaction as affected by rainfall, soil slope, and management. *Soil Sci. Soc. Am. J.* 49:1010–1015. doi:10.2136/sssaj1985.03615995004900040044x
- Sharpley, A.N., R.W. McDowell, J.L. Weld, and P.J.A. Kleinman. 2001. Assessing site vulnerability to phosphorus loss in an agricultural watershed. *J. Environ. Qual.* 30:2026–2036. doi:10.2134/jeq2001.2026
- Sharpley, A.N., J.L. Weld, D.B. Beegle, P.J.A. Kleinman, W.J. Gburek, P.A. Moore, and G. Mullins. 2003. Development of phosphorus indices for nutrient management planning strategies in the U.S. *J. Soil Water Conserv.* 58:137–152.
- Shober, A.L., A.R. Buda, K.C. Turner, N.M. Fiorellino, A.S. Andres, J.M. McGrath, and J.T. Sims. 2017. Assessing Coastal Plain risk indices for sub-surface phosphorus loss. *J. Environ. Qual.* 46:1270–1286. doi:10.2134/jeq2017.03.0102
- USDA-NRCS. 2011. Conservation practice standard, Nutrient management 590. USDA-NRCS. https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1192371.pdf (accessed 28 Sept. 2017).
- USDA-EUSEPA. 1999. Unified national strategy for animal feeding operations. March 9, 1999. U. S. Gov. Print. Off., Washington, DC. <https://www3.epa.gov/npdes/pubs/finafost.pdf> (accessed 28 Sept. 2017).
- Vadas, P., L. Good, W. Jokela, K. Karthikeyan, F. Arriaga, and M. Stock. 2017. Quantifying the impact of seasonal and short-term manure application decisions on phosphorus loss in surface runoff. *J. Environ. Qual.* 46:1395–1402. doi:10.2134/jeq2016.06.0220
- 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
- Veith, T.L., S.C. Goslee, D.B. Beegle, J.L. Weld, and P.J.A. Kleinman. 2017. Analyzing within-county hydrogeomorphological characteristics as a precursor to phosphorus index modifications. *J. Environ. Qual.* 46:1365–1371. doi:10.2134/jeq2016.10.0416
- Veith, T.L., A.N. Sharpley, J.L. Weld, and W.J. Gburek. 2005. Comparison of measured and simulated phosphorus losses with indexed site vulnerability. *Trans. ASAE* 48:557–565. doi:10.13031/2013.18330
- 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
- White, M.W., R.D. Harmel, and R.L. Haney. 2012. Development and validation of the TX Best Management Practice Evaluation Tool (TBET). *J. Soil Water Conserv.* 67:525–535. doi:10.2489/jswc.67.6.525
- Williams, J.R., and R.C. Izaurralde. 2006. The APEX model. In: V.P. Singh and D.K. Frevert, editors, *Watershed models*. CRC Press, Boca Raton, FL. p. 437–482.
- Williams, M.R., K.W. King, G.A. LaBarge, R.B. Confesor, Jr., and N.R. Fausey. 2017. Edge-of-field evaluation of the Ohio phosphorus risk index. *J. Environ. Qual.* 46:1306–1313. doi:10.2134/jeq2016.05.0198
- Young, R.A., C.A. Onstad, D.D. Bosch, and W.P. Anderson. 1989. AGNPS: A nonpoint source pollution model for evaluating agricultural watersheds. *J. Soil Water Conserv.* 44:168–173.