Evaluation of the APEX Model to Simulate Runoff Quality from Agricultural Fields in the Southern Region of the United States

John J. Ramirez-Avila,* David E. Radcliffe, Deanna Osmond, Carl Bolster, Andrew Sharpley, Sandra L. Ortega-Achury, Adam Forsberg, and J. Larry Oldham

Abstract

The Agricultural Policy Environmental eXtender (APEX) model has been widely applied to assess phosphorus (P) loss in runoff water and has been proposed as a model to support practical decisions regarding agricultural P management, as well as a model to evaluate tools such as the P Index. The aim of this study is to evaluate the performance of APEX to simulate P losses from agricultural systems to determine its potential use for refinement or replacement of the P Index in the southern region of the United States. Uncalibrated and calibrated APEX model predictions were compared against measured water quality data from row crop fields in North Carolina and Mississippi and pasture fields in Arkansas and Georgia. Calibrated models satisfactorily predicted event-based surface runoff volumes at all sites (Nash-Sutcliffe efficiency [NSE] > 0.47, [percent bias [PBIAS]] < 34) except Arkansas (NSE < 0.11, [PBIAS] < 50) but did not satisfactorily simulate sediment, dissolved P, or total P losses in runoff water. The APEX model tended to underestimate dissolved and total P losses from fields where manure was surface applied. The model also overestimated sediments and total P loads during irrigation events. We conclude that the capability of APEX to predict sediment and P losses is limited, and consequently so is the potential for using APEX to make P management recommendations to improve P indices in the southern United States.

Core Ideas

- Calibrated APEX reasonably predicted runoff in no-tillage and conventional tillage fields.
- APEX was unsatisfactory in predicting sediment losses, especially from pasture fields.
- P loss was inadequately predicted, especially in fields with surface applied manure.
- Adding a surface manure P pool to APEX could improve P model predictions.
- Improving P model predictions allow it to be used to refine southern region P Indices.

PHOSPHORUS (P) nutrient management planning in the United States relies heavily on decision support tools that balance agronomic and water quality outcomes. Concern over the P Index, the principal P-based tool used in nutrient management planning, resulted in revisions to the USDA’s Nutrient Management Standard (USDA-NRCS, 2011). These revisions stipulated that US states verify their versions of the P Index, including by using other models to evaluate P Index ratings, and recommended consideration of the Nutrient Tracking Tool (NTT; Saleh et al., 2011), a user-friendly linkage to the Agricultural Policy Environmental eXtender model (APEX; Williams et al., 2008).

The APEX model has been widely applied to different regions in the United States under different agricultural management conditions to assess hydrology, soil erosion, crop yield, and the effect of agricultural nutrient management practices on soluble and particulate P loads in runoff (Chung et al., 1999; Ramirez-Avila et al., 2012; Ford et al., 2015). However, APEX has been reported to have limitations in correctly predicting flow, sediments, and/or nutrient losses due to model process representation (Francesconi et al., 2014; Bhandari et al., 2017). For instance, Bhandari et al. (2017) reported model limitations to predict sediment loss due to its inability to simulate erosion processes beyond sheet and rill erosion, whereas Francesconi et al. (2014) pointed out APEX’s limitations in simulating tile flow, and consequently phosphorus and nitrogen loss, from tile-drained systems.

Motivated by the reported concerns relative to the performance of APEX, the objective of this study was (i) to compare modeled predictions from calibrated, uncalibrated, and validated versions of APEX with water quality data collected at four experimental sites (Arkansas, Georgia, Mississippi, and North Carolina), and (ii) to evaluate APEX model performance. These APEX predictions were integrated along with other output from the Annual P Loss Estimator (APLE; Vadas et al., 2007) and the Texas Best Management Practice Evaluation Tool (TBET; White...
et al., 2012) models and compared with southern P Indices to determine their usefulness in assessing P loss from agricultural fields (Osmond et al., 2017).

Materials and Methods

APEX Overview

The APEX model is a semi-process-based, distributed hydrologic and water quality model that operates on a daily time step (Radcliffe et al., 2015). It was developed as an extension of the Environmental Policy Integrated Climate (EPIC) model and can be used to evaluate various land management strategies by considering hydrology, soil and water quality, wind, sheet and channel erosion, plant growth, weather, pests, and economics (Steglich and Williams, 2013). Different databases (soils, crops, tillage, fertilizer, and pesticides) have been built in APEX to facilitate model initialization (Steglich and Williams, 2013). The detailed theoretical and technical documentation of the APEX model (Version 0806) can be found in Steglich and Williams (2012, 2013).

Study Sites

This study was part of a larger southern regional project that used the same water quality and land treatment datasets from Arkansas, Georgia, Mississippi, and North Carolina to test models and P Indices. Our study used the same observed datasets as the TBET study by Forsberg et al. (2017), except we added a site in Mississippi. Study site information including crops, treatments, soil mapping units, P application amounts and sources, soil-test P, and other general and specific characteristics is shown in Table 1.

Model Setup

The information used to set up and evaluate the APEX model included 10-m digital elevation maps and climate, soils, water quality, and agricultural operations datasets. The runoff quantity and water quality datasets and the soil-test P for each individual field were directly supplied by researchers at the corresponding study sites. The information about agricultural operations (e.g., operation type and date, nutrient application rates), crop yield, and soil properties (e.g., pH, organic matter, nitrate) were obtained from Pierson et al. (2001), Yuan et al. (2013), Larsen et al. (2014), and Edgell et al. (2015).

The Mississippi watersheds represented relatively large areas (11.3 and 17 ha) compared with the sites in Arkansas, Georgia, and North Carolina, with various soil series and slopes and establishment of two crops (cotton [Gossypium hirsutum L.] and soybean [Glycine max (L.) Merr.]) on the watershed’s fields. For these reasons, different subareas were delineated when the model was set up. The sites in Arkansas, Georgia, and North Carolina (0.017 to 0.4 ha) were treated as individual homogeneous areas with unique slope, soil series, crop, and land use management. Additional soils and agricultural operation schedule databases were built to set up the model for each individual field or watershed. Soil properties that were not measured for the sites (e.g., bulk density, cation exchange capacity, saturated conductivity) were obtained from the Soil Survey Geographic Database included in the ArcAPEX package (ArcGIS interface for APEX) Version 0806.10_0.4 Beta3 (released on 5 June 2013). The datasets we built also included detailed records of agricultural operations for tillage, planting, harvesting, fertilization, manure application, chemical burning, grazing, and/or haying. For Mississippi, when the specific date for harvesting soybean was not reported, it was estimated by using the Maturity Date Calculator (SoyPheno) model (Zhang et al., 2004) available from the Mississippi State Agricultural and Forestry Experiment Station of Mississippi State University.

Precipitation datasets for the entire monitoring period for each set of fields in Georgia, North Carolina, and Mississippi were obtained from weather stations onsite. The onsite climate information for the Arkansas fields only referenced precipitation depths when runoff events were observed. Missing data for onsite precipitation and information about maximum and minimum temperatures and wind velocity were obtained from weather stations located no more than 15 km away from the different study sites and added to the onsite data. Detailed information about soils, agricultural operations, and monitoring data are contained in Table 1.

The Penman–Monteith equation (Monteith, 1965) was selected to estimate the potential evapotranspiration, the modified rational equation (Kuichling, 1889) to calculate the peak runoff rate, and the curve number (CN) method (Wilson et al., 2017) to estimate the runoff depths. The Stochastic CN Estimator and the option Variable Daily CN Soil Moisture Index (Wang et al., 2012) were selected to estimate daily CN adjustments. Runoff depths induced by irrigation were estimated by APEX from user-defined runoff fractions of irrigation water. In APEX, irrigation is simulated with a uniform distribution of the water on the irrigated area.

The Modified Universal Soil Loss Equation for Small Watersheds (MUSS), a variant of the Modified Universal Soil Loss Equation (MUSLE; Williams, 1975) particularly adapted for fields and/or small watersheds with no erosion in their channel or stream (Williams and Izaaurralde, 2005), was the method selected to estimate single-storm soil erosion loads and sediment yield. Crop growth was simulated using preexisting crop characteristics contained in the model crop parameter file. The EPIC enrichment ratio method was selected to allow APEX to estimate sediment-bound P losses in runoff, and the Groundwater Loading Effects of Agricultural Management Systems (GLEAMS; Leonard et al., 1987) equation was used to estimate the soluble P runoff. The APEX model estimates soluble P by partitioning the total P (TP) into solution and sediment phases. Calibration and validation timeframes were determined according to the magnitude of total annual rainfall for each site, considering at least 1 yr of low and high rainfall during the calibration process. For Mississippi, the timeframes also considered changes in soybean varieties, P application only during the third year (out of four) determined by the nutrient management plan for the fields, and furrow irrigation application, determined by weather conditions, during the summer of the last 2 yr only. The time periods used in the calibration datasets were the first 2 yr (out of three) for Arkansas and North Carolina, the first and third year (out of four) for Georgia, and the first 3 yr (out of four) for Mississippi. The remaining years were used for validation (Supplemental Table S1).

Uncalibrated runs were performed using the default values for the different model parameters included in the ArcAPEX package. Runs for uncalibrated and calibrated model realizations used
Model Performance Analysis

Performance of the APEX model predictions for event-based runoff, sediment, and P loads were evaluated by using the Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) and percent bias (PBIAS) as the performance statistical measures. Positive values of PBIAS indicate underprediction. On the basis of a literature review (Ramanarayan et al., 1997; Moriasi et al., 2007, 2015; Wang et al., 2012) and recognizing that simulations are more variable at reduced time steps, we chose a NSE $\geq 35$, 60, 70, and 70% for runoff, sediment, TP, and DP as representing a successful simulation. These values corresponded to the general criteria used in other modeling efforts associated with the overall southern project (Bolster et al., 2017; Forsberg et al., 2017).

The mean absolute error (MAE) and the Spearman's rank correlation coefficient ($\rho$) (Spearman, 1904) were also determined. Nonparametric estimates of linear regression coefficients between observed and APEX predicted values were calculated.
with the Kendall-Theil Robust (KTR) method using the KTRLine (Version 1.0) model (Granato, 2006). The KTR method was selected because it minimized the effects of outliers and normality in residuals, a common condition of hydrological datasets. The model estimated the slope of the line as the median of all possible slopes between points (Granato, 2006).

## Results and Discussion

### Uncalibrated Runs

Given the criteria for event-based model performance (NSE and PBIAS), uncalibrated runs of the APEX model for the default settings were, in general, unsatisfactory at all sites except Georgia for runoff and Mississippi for sediment (Table 2). Large MAE and PBIAS absolute values were measured, as well as low model efficiencies (NSE), low or no significant linear correlations (Spearman), and low KTR slopes, some of them equal or very close to zero. Saleh et al. (2004) and Winchell et al. (2011) stated that estimates using uncalibrated runs of APEX may result in large errors compared with the observed data.

#### Runoff

Uncalibrated runs of APEX predicted runoff better than any other variable for all sites (Table 2). For row crop fields (Mississippi and North Carolina), margins of error were smaller (MAE < 8 mm), whereas correlation parameters ($r > 0.64$) were higher (Fig. 1, Table 2, Supplemental Table S2) than in pasture fields (Georgia and Arkansas). Uncalibrated model predictions of runoff yielded a satisfactory model efficiency (NSE > 0.5) in Mississippi and Georgia, although runoff predictions were unsatisfactory (underestimated) in Mississippi (PBIAS = 47.22). Runoff depths were overestimated and several runoff events not reported by the observed dataset were predicted by APEX in Arkansas. Forsberg et al. (2017) reported similar performance for the same Arkansas site running the TBET model under uncalibrated conditions.

#### Sediments and Phosphorus

The APEX model underestimated sediment loss in Mississippi but severely overestimated losses in Arkansas and North Carolina (Table 2). Uncalibrated runs resulted in an unsuccessful model performance (negative NSE) of TP loads with overestimation in pasture fields and underestimations in row crop fields (Table 2). The DP fraction was underestimated by APEX for all the sites. Underestimations in row crop fields were greater than those in pasture fields. The unsuccessful model performance for sediment and TP loss in the Arkansas fields can be attributed, in part, to overestimations of runoff and the small magnitude of sediment (0.01 Mg ha$^{-1}$) and TP losses (0.03 kg ha$^{-1}$) exported from the pasture fields. Saleh et al. (2004) reported that for fields or watersheds contributing small levels of sediment or nutrient losses, any small error or difference in the magnitude

| Table 2. Agricultural Policy Environmental eXtender (APEX) model performance estimates (MAE, mean absolute error; NSE, Nash-Sutcliffe Efficiency; PBIAS, percent bias; $r$, Spearman’s rank correlation coefficient of runoff, sediment loss, Total P, and dissolved P for uncalibrated and validated predictions for fields or subwatersheds in Arkansas, Georgia, North Carolina, Mississippi. Model predictions that met the performance criteria for NSE and PBIAS are shown in bold. |
|---|---|---|---|
| **Uncalibrated** | **Calibrated** | **Validated** |
| **Arkansas** | **Georgia** | **North Carolina** | **Mississippi** | **Arkansas** | **Georgia** | **North Carolina** | **Mississippi** | **Arkansas** | **Georgia** | **North Carolina** | **Mississippi** |
| **Runoff** | | | | | | | | | | | | |
| MAE (mm) | 11.88 | 8.22 | 6.59 | 7.82 | 7.18 | 7.42 | 5.45 | 6.52 | 23.11 | 7.05 | 11.00 | 4.63 |
| NSE | −17.08 | 0.58 | 0.21 | 0.67 | 0.70 | 0.47 | 0.11 | 0.11 | −0.28 | 0.72 | 0.68 | 0.52 |
| PBIAS (%) | −3.42 | 47.22 | −49.54 | 23.03 | 23.03 | 23.03 | 23.03 | 23.03 | 23.03 | 23.03 | 23.03 | 23.03 |
| $r$ | −0.10 | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 |
| **Sediment** | | | | | | | | | | | | |
| MAE (Mg ha$^{-1}$) | −12.15 | −11.0 | −8.6 | −6.9 | −5.3 | −4.8 | −4.3 | −3.9 | −3.5 | −3.1 | −2.7 | −2.3 |
| NSE | −321,534 | −160 | 0.34 | −0.28 | −0.28 | −0.28 | −0.28 | −0.28 | −0.28 | −0.28 | −0.28 | −0.28 |
| PBIAS (%) | −13,057.9 | −752 | 49.77 | 30.65 | 30.65 | 30.65 | 30.65 | 30.65 | 30.65 | 30.65 | 30.65 | 30.65 |
| $r$ | 0.27 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| **Total P** | | | | | | | | | | | | |
| MAE (kg ha$^{-1}$) | −13,057.9 | −752 | 49.77 | 30.65 | 30.65 | 30.65 | 30.65 | 30.65 | 30.65 | 30.65 | 30.65 | 30.65 |
| NSE | −118.27 | −0.10 | −0.34 | −0.51 | −0.51 | −0.51 | −0.51 | −0.51 | −0.51 | −0.51 | −0.51 | −0.51 |
| PBIAS (%) | −296.65 | −11.10 | 78.05 | 74.17 | 74.17 | 74.17 | 74.17 | 74.17 | 74.17 | 74.17 | 74.17 | 74.17 |
| $r$ | 0.47 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 |
| **Dissolved P** | | | | | | | | | | | | |
| MAE (kg ha$^{-1}$) | −296.65 | −11.10 | 78.05 | 74.17 | 74.17 | 74.17 | 74.17 | 74.17 | 74.17 | 74.17 | 74.17 | 74.17 |
| NSE | −1.39 | 0.04 | −0.05 | −0.61 | −0.61 | −0.61 | −0.61 | −0.61 | −0.61 | −0.61 | −0.61 | −0.61 |
| PBIAS (%) | 46.62 | 5.39 | 84.27 | 89.04 | 89.04 | 89.04 | 89.04 | 89.04 | 89.04 | 89.04 | 89.04 | 89.04 |
| $r$ | 0.29 | 0.03 | 0.04 | 0.40 | 0.40 | 0.40 | 0.40 | 0.40 | 0.40 | 0.40 | 0.40 | 0.40 |
Fig. 1. Kendall-Theil Robust linear regression and 95% confidence intervals for (A) uncalibrated, (B) calibrated, and (C) validated Agricultural Policy Environmental eXtender (APEX) model predictions of runoff datasets from Arkansas, Georgia, Mississippi, and North Carolina.
1362 Journal of Environmental Quality

of observed or estimated data results in large errors and consequently low model efficiencies.

Due to its unsatisfactory performance, using uncalibrated versions of APEX (or NTT) in the southern states could lead to inappropriate nutrient management planning or management strategies such incorrect P amendment rates, the lack of implementation of best management practices, and/or the miscalculation of P credits for water quality trading.

Calibration and Validation

Runoff

Event model performance assessments for calibration and validation periods resulted in satisfactory performance for runoff datasets at all sites except for Arkansas, where calibration improved runoff predictions but model performance was still unsatisfactory (Fig. 1, Table 2). Event-based, monthly, and annual predicted runoff depths in Mississippi are shown in Fig. 2a, 2b, and 2c, respectively. There was a tendency to slightly underestimate runoff (15–30 mm) in months with intermediate precipitation (April–June) (Fig. 2b), and this was reflected in the positive value for PBIAS (20%) in Table 2. The MAE for event runoff ranged from ~5 to 7 mm for the calibration period and from 5 to 23 mm for the validation period, with the highest value being observed in the Arkansas validation period. Significant positive correlations were observed with Spearman coefficients ranging from 0.63 to 0.76 and from 0.68 to 0.92 during the calibration and validation periods, respectively. Correlations were higher for the row crop sites than for pasture sites. The KTR slopes for runoff comparisons were also higher than those observed during uncalibrated runs (0.34–0.78) for the calibration (0.63–0.78) and validation periods (0.51–1.12), respectively (Supplemental Table S3).

The inaccurate prediction of runoff depths during the calibration and validation periods in Arkansas could have been due to input precipitation data from the remote weather site different than the onsite precipitation. As observed during the uncalibrated runs, APEX predicted runoff events for dates not in the observed data, which represented up to 530 mm of additional annual runoff depth. Forsberg et al. (2017) found that the TBET model performed poorly at the Arkansas sites and also attributed this to the lack of onsite precipitation data.

Sediments

The event modeling results of sediment loss from the Mississippi fields were the only ones that yielded satisfactory performance for calibration and validation periods (NSE = 0.48 and 0.49; PBIAS = 22 and −12%). The APEX model tended to overestimate sediment loss for the plots in Arkansas and North Carolina. During the validation period (1999) in Mississippi, a slight overestimation of sediment loss was caused by summer irrigation events. Since APEX uses empirical relationships to predict erosion from user-defined runoff fractions of irrigation water and runoff depths, it can lead to overestimating the flow velocity (a main parameter in the estimation of erosion caused by applying irrigation) when runoff is high. The validation period corresponded to a dry year in which almost 60% (240 mm) of the water supplied to crops during the summer was irrigated. In contrast, during the summer of 2008, the irrigation depths represented only 25% of the water supplied to crops during that season.
The low rates of sediment loss observed for the Arkansas pastures (6–160 kg ha$^{-1}$) and the no-tillage treatments in North Carolina (50–700 kg ha$^{-1}$) affected the goodness-of-fit values (Table 2). As noted earlier, it is difficult to predict erosion accurately when sediment losses are small (Saleh et al., 2004). Kumar et al. (2011) reported poor calibration performance of APEX (NSE < 0.19) when it was evaluated on grazed pasture watersheds. Neary (1998) remarked that field- and plot-scale erosion models have difficulty predicting small-scale events.

Phosphorus

Low performance ratings for TP and DP loads were observed for the calibrated and validated periods at all sites except Arkansas during the validation period, where the TP load met the performance criteria (NSE = 0.33, PBIAS = 44%). However, since the runoff and sediment predictions for Arkansas did not meet the performance criteria, the satisfactory predictions for TP and DP at this site are suspect. According to Francesconi et al. (2014), APEX cannot accurately predict small concentrations and loads of nutrients. Under these circumstances, APEX tended to underpredict TP and DP loads by a wide margin in Georgia (PBIAS > 86%) and North Carolina (PBIAS > 73%) during calibration and validation periods. This implies that the partitioning of TP and DP in APEX underestimates DP in situations where erosion and TP losses are small, such as pastures and no-tillage. unsuccessful performance and underestimations of DP in Arkansas, Georgia, and North Carolina can also be attributed to the lack of a model subroutine in APEX for surface-applied manure. Pierson et al. (2001), Sen et al. (2012), and Collick et al. (2016) reported that P loads from sites with organic amendments were underestimated because EPIC, APEX, and Soil and Water Assessment Tool (SWAT) did not include a surface manure pool for P that is separated from the soil surface layer pools. Collick et al. (2016) added to SWAT2012 (Revision 586) the set of soil P routines proposed by Vadas et al. (2007) to simulate surface-applied manure at field and subwatershed scales. A similar revision and update of the APEX routines could improve the model’s sensitivity to the effects of timing, rate, method, and source of P application. The APEX model overestimated the DP load carried by irrigation runoff events in Mississippi during summer 1999 (validation period). Sediment loss in Mississippi from irrigation runoff events was overestimated, as noted before, but this should have led to an overestimation of TP, not DP.

Our results for APEX are similar to what Forsberg et al. (2017) found for TBET. Since both models used similar routines for predicting runoff and sediment and P losses, this is not surprising.

Conclusion

The performance of APEX to simulate P losses from agricultural systems in the southern region of the United States was evaluated. The uncalibrated model produced unsatisfactory results for all constituents except runoff in Georgia and sediment loss in Mississippi. The calibrated model predicted runoff satisfactorily in almost all situations, but predictions of sediment loss, TP, and DP were unsatisfactory in most cases. Uncertainty in the precipitation record at the Arkansas site may have contributed to model error. Sediment losses from all sites except Mississippi were low because they were pastures or no-tillage; the Mississippi site was conventional row crop agriculture. Underprediction of TP and DP losses in manured pastures could be due to the lack of a surface manure pool in APEX. Partitioning of TP into DP in APEX may also underestimate runoff losses of DP in pastures. A calibrated APEX model could be used to improve predictions of runoff in southern region P Indices, but not for prediction sediment and P losses.

Acknowledgments

The authors are grateful for funding provided by USDA-NRCS Conservation Innovation Grants (Agreement no. 69-35-175-12-182).

References


