Tradeoffs in Model Performance and Effort for Long-Term Phosphorus Leaching Based on In Situ Field Data


Phosphorus and N are critical nutrients for agriculture but are also responsible for surface water enrichment that leads to toxic algal growth. Although P loading to surface waters has traditionally been thought to occur primarily in surface runoff, contributions from subsurface transport can also be significant. The primary objectives of this research were to evaluate several methods of representing macropore flow and transport in a finite element model using plot-scale infiltration and leaching data and to compare several models of various levels of complexity to simulate long-term P leaching. To determine flow and transport parameters, single- and dual-porosity models in HYDRUS-2D were calibrated with infiltration, Cl−, and P data from a 22-h plot-scale leaching experiment on a silt loam mantle with gravel subsoil. Both homogeneous and heterogeneous gravel profiles were simulated. The dual-porosity model with heterogeneous hydraulic conductivity best matched experimental data, with physical nonequilibrium (dual porosity) being more important than two-dimensional (2D) heterogeneity. Long-term (9 yr) P leaching to the water table (3 m below the soil surface) at the field site was simulated with both one-dimensional (1D) and 2D models using the calibrated parameters. There was little difference between analogous 1D and 2D models, suggesting that HYDRUS-1D may be sufficient to model long-term P leaching. Overall, the most important elements for accurately simulating P leaching in this silt loam and gravel soil profile were found to be (i) field-measured hydraulic conductivity of the limiting soil layer, (ii) calibrated dispersivity, and (iii) dual-porosity, in some circumstances.

Abbreviations: BTC, breakthrough curve; EPC, equilibrium phosphorus concentration; NSE, Nash–Sutcliffe efficiency; 1D, one-dimensional; 2D, two-dimensional.

Phosphorous is often the limiting nutrient in surface waters (Correll, 1999). Phosphorus transport has been assumed to take place primarily in surface runoff, although a growing collection of research indicates that subsurface P transport can be significant (Osborne and Kovacic, 1993; Cooper et al., 1995; Gburek et al., 2005; Fuchs et al., 2009). At field sites in the St. Joseph River watershed in northeastern Indiana, Smith et al. (2015) found that approximately half of the P losses occurred through tile drainage, due in part to macropore flow. There is a need to be able to perform long-term simulations of P leaching to estimate long-term loading of P to aquifers and streams through subsurface transport processes.

It has been shown that, in porous media with heterogeneous flow properties, the majority of the flow can occur in small preferential flow paths (Gotovac et al., 2009; Najm et al., 2010), with potential for rapid leaching of solutes through soil profiles to groundwater. Djodjic et al. (2004) performed experiments on P leaching through undisturbed soil columns.
and stressed the need to consider larger scale leaching processes due to soil heterogeneity. Subsurface P transport rates in Ozark floodplains have been shown to be comparable with surface runoff P transport rates (Mittelstet et al., 2011). In many gravelly floodplains, gravel outcrops and macropores are present, resulting in high infiltration rates, some of which are reported to be 10 to 74 cm h\(^{-1}\) (Heeren et al., 2015). Using plot-scale solute injection experiments, Heeren et al. (2017) found P leaching from the soil surface to the groundwater to be significant, with rapid detection of P in gravel outcrops (e.g., 4 min) and in silt loam soils with macropores. In one silt loam plot, the maximum transport velocity for soluble reactive P was 810 cm h\(^{-1}\). However, these field experiments were relatively short (3–52 h) and did not document long-term P leaching.

One way to conceptualize macroporosity in a soil is through the use of a multidomain system (Beven and Germann, 1982; Šimůnek and van Genuchten, 2008). Multidomain models split the soil profile into a fracture (macropore) domain and a matrix domain to simulate rapid flow and transport rates due to macroporosity. Multidomain models express physical transport in several ways. Mobile–immobile models define water and solute flow through the macropore space, with solute transport also occurring between the immobile and mobile domains through molecular diffusion. Dual-porosity models build on this further by allowing both water flow and solute transport (through advection as well as diffusion) to occur between the mobile and immobile domains.

HYDRUS-1D and HYDRUS-2D simulate the transport of water, solutes, and heat through simple and complex soil profiles. HYDRUS uses numerical methods (finite element in space and finite difference in time) to solve the Richards equation for variably saturated water flow and the advection–dispersion equation for heat and solute transport (Šimůnek and van Genuchten, 2008; Šimůnek et al., 2012). HYDRUS has been used to simulate preferential flow and transport both by using multidomain models that express domains as overlapping continua (e.g., dual-porosity) and by simulating a macropore as a single band of highly conductive material built directly into the finite element mesh (Akay et al., 2008; Lamy et al., 2009).

Elmi et al. (2012) used HYDRUS-1D and a single-porosity model to simulate P transport through undisturbed soil cores. Naseri et al. (2011) also performed column experiments on soils cores to measure P transport. However, neither of these studies simulated preferential flow in HYDRUS. Limited research has been performed using profile data from advanced tools, such as electrical resistivity mapping, to determine a two-dimensional (2D) saturated hydraulic conductivity (\(K_s\)) field for simulating P transport in HYDRUS. There is a need to develop long-term nutrient leaching models based on field experiments that capture the complexities of macropore flow in situ.

The objectives of this research were (i) to evaluate several methods of representing macropore flow and transport in a 2D finite element model using plot-scale infiltration and leaching data, and (ii) to compare multiple numerical models of various complexities to simulate long-term P leaching. It was hypothesized that using a dual-porosity model and accounting for spatial heterogeneity in \(K_s\), as well as using field measured data, would improve the ability of the model to perform accurate long-term simulations. Although increasing model complexity may improve the predictive capabilities of the model, the level of improvement needs to be compared with the effort required to calibrate more complex models.

Materials and Methods

Field Site

The Barren Fork Creek floodplain site (35.90° N, 94.85° W) was located in the Ozark ecoregion of northeastern Oklahoma. The Barren Fork Creek is a state-designated Scenic River and is on the Oklahoma 303(d) list for nutrient impairment (USEPA, 2015). Poultry feed, and thus poultry litter, is the largest source of P in the watershed (Mittelstet and Storm, 2016). Floodplains generally consist of coarse chert gravel overlaid by a mantle of gravelly loam or silt loam (Fig. 1). The soils were Razort gravelly loam (fine-loamy, mixed, active, mesic Mollic Hapludalfs) with the silt loam layer ranging from 30 to 200 cm thick and the chert gravel layer ranging from 3 to 5 m thick, extending down to limestone bedrock. The gravel vadose zone had \(K_s\) values ranging from 550 to 1700 cm h\(^{-1}\) according to site borehole permeameter tests (Miller et al., 2014). The gravel itself was a complex alluvial deposit that included both clean gravel lenses associated with rapid flow and transport (Fox et al., 2011) and layers of fine gravel that could cause lateral flow in the silt loam and subsequent seepage erosion (Fig. 1). The anisotropic horizontal layering resulted in a propensity for lateral flow.

Numerical simulations used data from a previous plot-scale infiltration experiment implemented at the Barren Fork Creek site. The berm infiltration method (Heeren et al., 2014) was used to confine water and solutes in an infiltration plot (1 by 1 m) within the floodplain (Fig. 2). A constant head of water and constant Cl\(^{-}\) and P concentrations were maintained within the plot. The Cl\(^{-}\) (conservative) was injected as KCl, resulting in a concentration of 50.1 mg L\(^{-1}\) Cl\(^{-}\). The P (highly sorbing) concentration of 1.68 mg L\(^{-1}\) (corresponding to 5.6 mg L\(^{-1}\) as phosphate) was used to represent poultry litter application rates (typically used as a fertilizer source in the Ozark ecoregion) in the range of 2 to 8 Mg ha\(^{-1}\). The P concentrations were achieved by adding H\(_3\)PO\(_4\), which deprotonated to H\(_2\)PO\(_4\)\(^{-}\) and HPO\(_4\)\(^{2-}\) in the slightly acidic solution. Five observation wells were installed near the plot to collect water samples to document solute breakthrough curves (BTCs). The infiltration data were presented in Heeren et al. (2015) and the transport data in Heeren et al. (2017). The current research used HYDRUS to simulate the 1- by 1-m infiltration plot that was tested on 30 June 2011 (Fig. 2).

Soil Chemistry

To determine soil chemical properties, soil core samples were collected with a Geoprobe Systems 6200 TMP (trailer-mounted probe) direct-push drilling machine using a dual-tube core sampler with a 4.45-cm opening. Before the P injection experiment, background soil cores were collected during the installation of the observation wells and were tested for water-soluble P. After the P injection experiment, soil cores were collected from within the plot.
to document the change in the soil profile water-soluble P levels due to the infiltration of P-laden water (Fig. 3). All soils were air dried and sieved with an 8-mm sieve prior to analysis. Details of the laboratory testing are presented in Heeren et al. (2017).

Isotherms were performed on samples (<8-mm fraction) from the soil cores near the infiltration plot (Table 1). The P isotherms exhibited linearity at low concentrations (<8 mg L\(^{-1}\)) and were fit with a linear isotherm (Heeren et al., 2017). The equilibrium P concentration (EPC), where neither sorption nor desorption occurred, was calculated as the \(x\) intercept of a logarithmic trend line fit to the entire data set (including high concentrations) (Table 1). Although the EPC was high (0.94–1.08 mg L\(^{-1}\)) compared with the background P concentrations in the aquifer (0.055 mg L\(^{-1}\)), the EPC was lower than the P concentration of the infiltrating water during the field experiment (1.68 mg L\(^{-1}\)).

The isotherms were performed on the fine fraction (<8 mm); however, parameters were needed that characterized the whole soil sample, since HYDRUS calculates P sorption in terms of the entire soil mass. Sorption on the coarse size fraction (>8 mm) was assumed to be negligible (Heeren et al., 2017). Therefore, “weighted” linear isotherm parameters were determined by accounting for the fraction of total sample on which testing was performed:

\[ K_{d,\text{whole}} = f_{<8\text{mm}} K_{d,<8\text{mm}} \]
where $K_{d,\text{whole}}$ is the linear sorption coefficient for the whole soil sample (L water kg$^{-1}$ soil), $f_{<8\text{mm}}$ is the fraction of the soil sample that passes through an 8-mm sieve (kg kg$^{-1}$), and $K_{d,<8\text{mm}}$ is linear sorption coefficient for the fine fraction (L water kg$^{-1}$ soil) (Table 1).

The $y_{\text{int,whole}}$ (where the line of the weighted isotherm intercepted the y axis, mg P kg$^{-1}$ soil) was also weighted according to $f_{<8\text{mm}}$. The EPC was the same for the fine fraction and the entire sample.

**Numerical Simulations**

Numerical methods were used to solve the Richards equation and the advection–dispersion equation for variably saturated flow and transport within HYRUS-1D and -2D, using both single- and dual-porosity formulations (Table 2). Model A was designed to use the level of data that would be available from a quick field site visit, including visual observation of silt loam and gravel layering on the streambank, a bucket sample of gravel (from the streambank) to determine the particle size distribution, and soil coring in the floodplain to determine depth of the silt loam, soil texture, water-soluble P, and P sorption isotherms. If successful, this single-porosity, one-dimensional (1D) model would require relatively little effort to implement at other field sites. Model B used measured $K_s$ (from an infiltration experiment for the silt loam and from a borehole permeameter for the gravel) and calibrated longitudinal dispersivity ($D_L$). Model B was 2D, although both the silt loam and the gravel were homogenous. Model C was designed to use data collected from an in-depth study of the research site. Model C accounted for heterogeneity in the gravel with three gravel layers, according to electrical resistivity imaging of the vadose zone. Due to the highly complex alluvial deposits (Fig. 1), it was expected that accounting for 2D heterogeneity in soil properties would significantly improve model performance.

Models D and E used a dual-porosity model to simulate the impact of preferential flow. Both models used measured $K_s$, calibrated $D_L$, and calibrated dual-porosity parameters. Model E required the most field data and modeling effort, using both dual-porosity and 2D heterogeneity in $K_s$.

**Soil Physical Properties**

The vertical soil profile was divided into two distinct soil layers—a 1.33-m silt loam layer and an underlying layer of

---

### Table 1. Soil chemical properties for two soil samples from soil cores near the field infiltration experiment.

<table>
<thead>
<tr>
<th>Borehole</th>
<th>Depth (cm)</th>
<th>Soil texture</th>
<th>8-mm sieve % passing</th>
<th>EPC†</th>
<th>$K_d$ (mg P kg$^{-1}$)</th>
<th>$y$ intercept</th>
<th>Weighted $K_d$ (mg kg$^{-1}$)</th>
<th>$y$ intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well B</td>
<td>64–83</td>
<td>Silt loam, some gravel</td>
<td>94</td>
<td>0.94</td>
<td>11</td>
<td>−14</td>
<td>10</td>
<td>−13</td>
</tr>
<tr>
<td>Well K</td>
<td>142–163</td>
<td>Sandy gravel</td>
<td>57</td>
<td>1.08</td>
<td>2.6</td>
<td>−6.3</td>
<td>1.5</td>
<td>−3.6</td>
</tr>
</tbody>
</table>

† EPC, equilibrium P concentration.
The silt loam $K_s$ value was estimated to be 9.6 cm h$^{-1}$ from the $K_s$ values for the three gravel layers ranged from 130 to 578 cm h$^{-1}$. Soil Chemical Properties

Gravel soil parameters were estimated using the “sand” classification below, which was developed using borehole permeameter to improve model performance by accounting for spatial heterogeneity in Rosetta Lite, since parameters were not available for gravel.

Initial conditions included soil solution P concentrations equal to the EPC for the silt loam layer (0.94 mg L$^{-1}$) and the top of the gravel layer (1.08 mg L$^{-1}$). Initial solution P concentration in the gravel below the water table was equal to the average of background P concentrations from well samples (0.055 mg L$^{-1}$). The disparity in these concentrations indicates the presence of a solute front in the soil matrix (from historical P leaching) that has not yet reached the water table, although P leaching through macropores may have reached the water table during rainfall events. Given the relative location of this solute front, which is apparent in the water-soluble P data (Fig. 3), a linear interpolation was used for the initial solution P concentration between 1.08 mg L$^{-1}$ at 160 cm and 0.055 mg L$^{-1}$ at 175 cm. Soil P was assumed to be in chemical equilibrium with the solution.

Flow and transport are highly complex, 2D or three-dimensional processes (Fig. 1) (Heeren et al., 2017). This research sought to improve model performance by accounting for spatial heterogeneity in soil properties by using a 2D simulation informed by geophysics. The gravel was divided into three layers identified using electrical resistivity imaging data (Miller et al., 2014, 2016) from the location of the plot. The $K_s$ (m d$^{-1}$) values for the gravel layers were determined using electrical resistivity data and the relationship below, which was developed using borehole permeameter data from the Barren Fork Creek site and one other floodplain site in the Ozark ecoregion (Miller et al., 2014, 2016):

$$K_s = 0.11 \rho$$

where $\rho$ is electrical resistivity (Ω m). The $K_s$ values for points within each gravel layer, as determined with electrical resistivity data, were then averaged to generate an average $K_s$ for that layer. Average $K_s$ values for the three gravel layers ranged from 130 to 578 cm h$^{-1}$.

**Soil Chemical Properties**

The units in the HYDRUS simulations were centimeters for length, grams for soil mass (i.e., bulk density in g cm$^{-3}$), micrograms for P or Cl$^-$ mass, and hours for time. Therefore, the linear sorption coefficient ($K_d$) for P was entered in units of cubic centimeters per gram (e.g., $K_d = 10.3$ L kg$^{-1} = 10.3$ cm$^3$ g$^{-1}$ for the silt loam). The measured $K_d$ for the gravel sample was applied to the whole gravel layer. Initial conditions included soil solution P concentrations equal to the EPC for the silt loam layer (0.94 mg L$^{-1}$) and the top of the gravel layer (1.08 mg L$^{-1}$). Initial solution P concentration in the gravel—based on the layering at the field site. Values for van Genuchten parameters and soil material properties for the soil layers were estimated using the Rosetta Lite (version 1.1) module. Gravel soil parameters were estimated using the “sand” classification in Rosetta Lite, since parameters were not available for gravel. The silt loam $K_s$ value was estimated to be 9.6 cm h$^{-1}$ from the field infiltration tests (Heeren et al., 2015).

**Calibration**

During calibration, simulation results from the 2D models were matched to data collected from Observation Wells C and E adjacent to the plot (Fig. 2), which were the only two wells in which P was detected. Since the field experiments involved a 2D flow field (including lateral flow at the top of the water table before intersecting a well), it was necessary to use 2D numerical models for the calibration. Observation nodes in HYDRUS were placed at the water table on either side of the plot to represent the selected observation wells. A constant head of 6 cm was applied across the plot area. Constant concentration boundary conditions of 50.1 and 1.68 mg L$^{-1}$ were used for Cl$^-$ and P, respectively. Calibration was performed for the 2D models for both Cl$^-$ and P transport. Goodness-of-fit was determined using the R² and Nash–Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970) as recommended by Moriasi et al. (2007). Although R² describes the collinearity between the observed and simulated values, it is oversensitive to outliers and insensitive to additive and proportional differences between model predictions and measured data (Legates and McCabe, 1999). Nash–Sutcliffe efficiency, with a range from $-\infty$ to 1.0, indicates how well the plot of observed vs. simulated data fits the 1:1 line.

The numerical simulations included several levels of input data (Table 2). Models B and C were performed using the default single-porosity van Genuchten–Mualem 2D model, both with a homogeneous gravel (Model B) and heterogeneous gravel (three

---

Table 2. Description of models used to simulate flow and transport.

<table>
<thead>
<tr>
<th>Model</th>
<th>Porosity†</th>
<th>Dimension‡</th>
<th>$K_s$±</th>
<th>Gravel#</th>
<th>$D_L$††</th>
<th>Application‡‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>SP</td>
<td>1D</td>
<td>PTF</td>
<td>PSD-HM</td>
<td>Lit</td>
<td>LT</td>
</tr>
<tr>
<td>B</td>
<td>SP</td>
<td>2D</td>
<td>Meas</td>
<td>Meas-HM</td>
<td>Cal</td>
<td>Cal</td>
</tr>
<tr>
<td>C</td>
<td>SP</td>
<td>2D</td>
<td>Meas</td>
<td>Meas-HT</td>
<td>Cal</td>
<td>LT, Cal</td>
</tr>
<tr>
<td>D</td>
<td>DP</td>
<td>1D</td>
<td>Meas</td>
<td>Meas-HT</td>
<td>Cal</td>
<td>LT</td>
</tr>
<tr>
<td>E</td>
<td>DP</td>
<td>2D</td>
<td>Meas</td>
<td>Meas-HT</td>
<td>Cal</td>
<td>LT, Cal</td>
</tr>
</tbody>
</table>

† SP, single porosity; DP, dual porosity.
‡ 1D, one-dimensional; 2D, two-dimensional.
§ $K_s$, saturated hydraulic conductivity.
¶ $K_s$ estimated with pedotransfer function (Rosetta Lite); Meas, measured with plot infiltration experiment.
# PSD, $K_s$ estimated by particle size distribution; HM, homogeneous with a single gravel layer; Meas, measured with borehole permeameter and electrical resistivity imaging; HT, heterogeneous with three gravel layers.
†† $D_L$, longitudinal dispersivity; Lit, $D_L$ according to literature; Cal, calibrated $D_L$.
‡‡ LT, long-term simulations; Cal, calibration simulations.
gravel layers, Model C). Additional models were included to evaluate the effects of incorporating macropore flow (dual-porosity) on arrival time and the overall shape of Cl− and P BTCs.

**Single-Porosity Parameters**

Dispersivity [L] is used to correlate pore velocity to the mechanical dispersion of solutes in porous media. Traditionally, $D_L$ has been approximated to be 10% of the sample length in the direction of flow, with transverse dispersivity ($D_T$) being ~10% of the $D_L$ (Lallemand-Barres and Peaudecerf, 1978; as presented in Fetter, 1999). The flow path length during the field experiments was $\sim 400$ cm, resulting in a first estimate of $D_L$ of 40 cm for the calibration. The $D_T$ was always calculated to be 10% of the $D_L$.

**Dual-Porosity Parameters**

For the dual-porosity models, $\theta_{s,mo}$ and $\theta_{s,im}$ [L3 L−3] are the saturated volumetric water contents of the mobile and immobile domains. Tension infiltrometer tests conducted by Heeren et al. (2015) showed that 99% of flow was directed through macropores at the Barren Fork Creek site. Simulations conducted by Šimůnek et al. (2003) suggested the possibility of such flows occurring through a mere 2.5% of total pore space, which suggested that macropores can have a dominant effect on subsurface flows. Furthermore, Haws et al. (2005) modeled 2D mobile zones as a small percentage of the total porosity. Reducing the flow domain to such a small space has dramatic effects on mean pore water velocity, causing water and solutes to arrive much sooner than arrival after flowing through simple matrix flow. Values of $\theta_{s,mo}$ and $\theta_{s,im}$ were initially set to reflect the simulation conducted by Šimůnek et al. (2003), and $\theta_{s,mo}$ was allowed to be adjusted between 0.01 and the porosity estimated by the Rosetta Lite function (Table 3).

The $\omega$ [T−1] and $\alpha$ [T−1] terms are the water and solute mass transfer coefficients, respectively, for the mass transfer function in the modified advection–dispersion equation. Values of $\alpha$ are traditionally believed to range between 0.1 and 5.0 h−1 (Radcliffe and Šimůnek, 2010); however, Alletto et al. (2006) found $\alpha$ to range between 0.0006 and 0.0424 h−1, and Cheviron and Coquet (2008) reported $\alpha$ values of 0.0192 to 0.6528 h−1. González-Delgado and Shukla (2014) reported $\omega$ values of 0.001 to 0.30 h−1 in loam and 0.20 to 1.02 h−1 in sand. Therefore, BTCs were analyzed with $\omega$ ranging by several orders of magnitude with a minimum of 0.001 for both silt loam and gravel (Table 3).

The $f[-]$ is the fraction of sites available for sorption that are governed by an equilibrium process. Given the mobile–immobile nature of this particular model, $f$ was used to denote the fraction of sites in contact with mobile water during physical nonequilibrium. We analyzed $f$ for the entire range of possible values to get a good understanding of its effect on P sorption (Table 3). Due to the conservative nature of Cl−, $f$ was not calibrated when simulating Cl− transport. The best-fit parameter values (Table 3) were used to set the baseline parameter values for the sensitivity analysis. Each solute simulation was analyzed with respect to the time taken for water at the Well C observation node to reach a concentration of 15 mg L−1 for Cl− ($t_{15}$) or 0.12 mg L−1 for P ($t_{0.12}$). Input parameters were then increased or decreased, and the percentage change in $t_{15}$ or $t_{0.12}$ was recorded. Results were plotted as the percentage change in the parameter from the baseline value vs. the percentage change in time to the target concentration.

**Long-Term Phosphorus Simulations**

Long-term P transport was simulated with both the 1D and 2D calibrated models. Long-term trials simulated water and P application to a soil profile for a 9-yr period between March 2004 and March 2013. Rainfall data were obtained through the Oklahoma Mesonet (McPherson et al., 2007). Since the focus of this research was a comparison of model performance, rather than the magnitude of P load to the aquifer, evapotranspiration was neglected in the simulations. Future research should account for root water and nutrient uptake (Šimůnek et al., 2016) when simulating P leaching.

Phosphorus from poultry litter application was simulated as P applied with infiltrating rainwater starting 1 March of each year to match traditional fertilizer application times. Each year, 0.619 mg P cm−2 of soil surface was added to the simulation, consistent with a 5 Mg ha−1 (2 t acre−1) application rate of poultry litter on grass and a P content of 12.7 kg P t−1 of litter, as recommended by the Oklahoma Cooperative Extension Service (OCES, March 2013.

---

**Table 3. Soil properties and calibration parameters.** The most optimal parameter set was achieved using Model E.

<table>
<thead>
<tr>
<th>Soil</th>
<th>$K_s$ [cm h−1]</th>
<th>van Genuchten parameter†</th>
<th>Mobile</th>
<th>Immobile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\alpha$ [cm−1]</td>
<td>$n$</td>
</tr>
<tr>
<td>Silt loam</td>
<td>9.6</td>
<td></td>
<td>0.1</td>
<td>2.00</td>
</tr>
<tr>
<td>Gravel</td>
<td>130–578</td>
<td></td>
<td>0.145</td>
<td>2.68</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calibration parameter range†</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_{s,mo}$ [cm3 cm−3]</td>
</tr>
<tr>
<td>$D_L$ [cm]</td>
</tr>
<tr>
<td>Silt loam</td>
</tr>
<tr>
<td>Gravel</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Most optimal parameter set</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_{s,mo}$ [cm3 cm−3]</td>
</tr>
<tr>
<td>$D_L$ [cm]</td>
</tr>
<tr>
<td>Silt loam</td>
</tr>
<tr>
<td>Gravel</td>
</tr>
</tbody>
</table>

† $\alpha$, parameter for the retention curve; $n$, parameter for the retention curve; $l$, pore connectivity parameter for unsaturated hydraulic conductivity; $\theta_{s,mo}$, saturated volumetric water content of the mobile domain; $D_L$, longitudinal dispersivity; $D_T$, transverse dispersivity; $\omega$, water mass transfer coefficient; $\alpha$, solute transfer coefficient; $f$, fraction of sites available for sorption.
Initial concentrations of P in the simulated infiltration started at 15 mg L$^{-1}$, which is consistent with P concentrations in the first post-litter-application runoff event found by DeLaune et al. (2004). The decay of runoff P concentrations was simplified as a linear relationship between concentration and cumulative rainfall:

\[ C = I - 0.182R \]  

where \( C \) is the concentration (mg L$^{-1}$) at the given time step, \( I \) is 15 mg L$^{-1}$, the initial leachate concentration at 1 March of each year (mg L$^{-1}$), and \( R \) is the cumulative rainfall (cm) since 1 March of each year. Once the cumulative applied P reached 0.619 mg, no additional P was added to rainwater for that year. If rainfall was insufficient to remove all P from the surface for a given year, the excess P was added to the next year and a new linear relationship was developed to reflect the additional P.

Two long-term simulations were performed in HYDRUS-1D. Model A was designed to use the level of data that would be available from a quick site visit, and Model D included dual porosity to account for preferential flow. For Model A, the soil profile featured a 1.33-m silt loam mantle and a single 1.66-m gravel layer. Most of the soil characteristics for the gravel layer were defined as sand by Rosetta Lite, although the \( K_r \) value was determined using data collected by Fuchs et al. (2009) for the Barren Fork Creek site. The \( D_L \) was set to 10% of the length of the flow path (3 m); therefore, \( D_L \) was 30 cm for Model A. For Model D, the gravel layer was broken into three distinct layers. Model D evaluated transport with calibrated values for the rate constants and \( D_L \).

For 2D modeling, a 100-cm-wide, 300-cm-deep 2D domain was developed, corresponding to the vadose zone of the soil profile directly under the 100-cm-wide plot used in calibration. Long-term P transport to the water table, situated at the bottom of the profile, was of interest. Boundary conditions were set so that the sides of the domain were no-flow boundaries, the bottom of the domain was a constant head boundary (pressure head equal to zero) at the water table elevation, and the top of the domain was set as a variable flux boundary to simulate rainfall events. Initial conditions were at hydrostatic equilibrium with the water table, and initial concentrations remained the same as the initial concentrations used for P calibration. Long-term 2D simulations were performed on Models C and E (Table 2).

**Results**

**Calibration**

For Cl$^-$, Model E achieved the best calibration (Table 4, Fig. 4 and 5). Both \( R^2 \) and NSE were higher for Well C than Well E, 0.70 and −0.96, respectively. The low NSE values were due to the under-prediction at times 0.83 and 2.1 h. The NSE increased to 0.92 and −0.36 in Wells C and E, respectively, when only considering times 8.1 and 18.8 h. Overall, HYDRUS simulated P better than Cl$^-$. In Well C, the fit was excellent, but the concentrations were overpredicted in Well E. The single-porosity model with a homogeneous gravel layer (Model B) produced BTCs with longer arrival times, reduced peak concentrations, and poor differentiation between the two observation wells (Fig. 4 and 5). The single-porosity model with heterogeneous gravel layers (Model C) performed slightly better. While still having poor arrival times and peak concentrations, this model showed better differentiation between the two observation wells (Fig. 4 and 5). Calibration parameters for these two models were limited to \( D_L \) and \( D_L \) for the silt loam; all other variables either belong to the dual-porosity model or were already set to their maximum value prior to calibration. Silt loam \( D_L \) and \( D_L \) were set to the maximum value established in Table 3 to produce these results. The best-fit parameter values for Model E are shown in Table 3.

Decreasing \( \omega \) increased arrival time for both Cl$^-$ and P, and increasing \( \omega \) had the opposite effect. Effects of \( \alpha \) were more complex. Decreasing \( \alpha \) made the Cl$^-$ BTC sharper but had little effect on arrival time; however, increasing \( \alpha \) affected both time and shape of Cl$^-$ BTC. No significant effect was seen in the P BTC for changes in \( \alpha \).

One limitation of the model was the inability to match observed data with reasonable \( f \) values. Predicted values of \( f \) were $\sim$ 0.03, which is consistent with the percentage macropore composition of the soil profile. However, parameter optimization resulted in \( f \) being close to one to achieve reasonable arrival times for P and remain consistent with Cl$^-$ calibration results. Arrival times were difficult to match for Cl$^-$ and P simultaneously. Although arrival times for Cl$^-$ were relatively short, arrival times for P were relatively long. Balancing parameters that managed water flow, such as \( \omega \), was a difficult task, as changing these parameters to better match one solute caused a poor match with the other. Solute transport parameters, such as soil isotherm properties, were not enough to balance the Cl$^-$ and P perfectly. Observation data showed that both wells received some level of Cl$^-$, but only Well C recorded any significant P increase. Although the P increase simulated in Well E was reduced in comparison with Well C, the increase simulated was still far above the trend defined by observed data (Fig. 5). These challenges may indicate the limitations of using the Richards equation to simulate the complex flow and transport processes observed in field conditions (Beven and Germann, 2013).

During the calibration, solute mass balance errors for P were <0.88% for all time steps and models. The Cl$^-$ mass balance errors

<table>
<thead>
<tr>
<th>Model</th>
<th>Statistic</th>
<th>Cl$^-$ $R^2$</th>
<th>NSE</th>
<th>P $R^2$</th>
<th>NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Well C</td>
<td>0.47</td>
<td>−3.30</td>
<td>0.77</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Well E</td>
<td>0.47</td>
<td>−4.93</td>
<td>0.83</td>
<td>−3.35</td>
</tr>
<tr>
<td>C</td>
<td>Well C</td>
<td>0.46</td>
<td>−3.44</td>
<td>0.77</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Well E</td>
<td>0.47</td>
<td>−5.13</td>
<td>0.56</td>
<td>−2.61</td>
</tr>
<tr>
<td>E</td>
<td>Well C</td>
<td>0.70</td>
<td>−0.96</td>
<td>0.82</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>Well E</td>
<td>0.34</td>
<td>−3.32</td>
<td>0.85</td>
<td>−9.56</td>
</tr>
</tbody>
</table>
were <3.7% for all time steps and models. The water mass balance errors were <2.1% for all time steps for Models B and C, although water mass balance errors reached 24% for Model E during calibration.

**Sensitivity Analysis**

For Cl⁻ modeling, $D_L$ and immobile pore fraction had an inverse relationship to $t_{15}$, whereas $\alpha$ and $\omega$ had a positive relationship to $t_{15}$, although both $\alpha$ and $\omega$ seemed to display asymptotic behaviors at large percentage increases in the variable. The most sensitive parameters for the Cl⁻ analysis were $\theta_{s,im}$ for both the silt loam and the gravel, with maximum increases in $t_{15}$ of 77 and 167%, respectively. The least sensitive parameter was $\alpha$ for the silt loam, which only produced a 4% increase in $t_{15}$ despite seeing a 400% increase in $\alpha$ (Fig. 6).

For P modeling, the gravel mobile sorption site fraction and the gravel adsorption isotherm coefficient had a positive relationship to $t_{0.12}$. Neither the mobile site sorption fraction nor the adsorption isotherm coefficient for the silt loam layer had any significant effect on $t_{0.12}$. Although soil chemical analysis showed that the soils were not close to P saturation (degree of P saturation < 16%), initial solution P concentration in the silt loam (0.94 mg L⁻¹) was high relative to the plot inflow P concentration (1.68 mg L⁻¹). This initial condition would significantly reduce the impact of silt loam-dependent parameters, as sorption sites are already mostly filled with P for the inflow concentration. The gravel mobile sorption site fraction was the most sensitive parameter, with a maximum of 70% decrease in $t_{0.12}$. The least sensitive parameters were gravel adsorption isotherm coefficients, with changes between −20 and 20% in $t_{0.12}$ across a wide percentage change in the variable (Fig. 6).
Long-Term Phosphorus Simulations

Long-term modeling from March 2004 to March 2013 was conducted using Models A, C, D, and E. During the calibration, Model E resulted in the best fit with the field data, using a dual-porosity model and field-calibrated parameters; therefore, Model E was also considered to be the optimum model for the long-term simulations, and results from the other models were compared with Model E results.

During 9 yr, ~540 kg ha$^{-1}$ P was applied to the plot area through simulated fertilizer application. Model E simulated 91.7 kg P ha$^{-1}$ being delivered to the water table, resulting in a P delivery ratio of 16.8% (Table 5). The P concentration of the flow into the water table steadily increased with time, with a final concentration of 1.74 mg L$^{-1}$. Wet years (2004, 2008, and 2009) resulted in larger increases in P concentration than average and dry years (Fig. 7). Model C, similar to Model E but with a single porosity, predicted a final P concentration at the water table of 1.64 mg L$^{-1}$.

Among the 1D models, Model D and Model E results were the most similar, with a P delivery ratio of 16.5% and a maximum P concentration of 1.67 mg L$^{-1}$. Model A (single porosity, without field-measured $K_s$) predicted that a negligible amount of P (0.2 kg P ha$^{-1}$) crossed the water table after 9 yr of simulation (Table 5). It is important to note that these trials do not take evapotranspiration into account; these results are intended for comparison of various models, rather than quantifying the magnitude of P leaching loads. It is expected that, if evapotranspiration was included in the model, deep percolation past the root zone and P leaching would be proportionally less than the simulated values, with the relative differences among the various models being similar.

During the long-term simulations, the P mass balance errors were <4.9% for all time steps and models. The water mass balance errors were <0.78% for all time steps for Models A, C, and D; the maximum water mass balance error for Model E was 8.2%.

Discussion

Calibration

During the calibration step, Cl$^-$ and P transport were modeled satisfactorily (Table 4) while still keeping the values of soil properties within accepted ranges (Table 3), except for the fraction of sites available for sorption. There was some difficulty matching simulation BTCs to observed data. It is possible that the electrical resistivity imaging data could not provide a fine enough resolution of the soil profile to catch heterogeneity that would have explained why only one well displayed P transport. Another explanation might be that the dual-porosity model does not capture all of the flow and transport processes in this system and that alternative modeling techniques may be preferable (Nimmo, 2010; Beven and Germann, 2013).

The comparison of the models suggests the necessity of using a dual-porosity model to accurately represent macropore flow. During calibration, models not featuring a dual-porosity system considerably undersimulated both P and Cl$^-$. A single-porosity model does not adequately simulate solute transport processes, especially early arrival times, for soils dominated by macropore flow.

Fig. 6. Sensitivity analysis of Model E for Cl$^-$ (left) and P (right). The model output (y axis) was selected to be comparable with the calibration dataset, with $t_{15}$ being the breakthrough time for Cl$^-$ to reach 15 mg L$^{-1}$ and $t_{0.12}$ being the breakthrough time for P to reach 0.12 mg L$^{-1}$. DL is the longitudinal dispersivity, $\theta_{s,Im}$ is the saturated water content of the immobile domain, $\alpha$ and $\omega$ are the solute and water mass transfer coefficients, $f$ is the fraction of sites available for sorption, and $K_d$ is the linear sorption coefficient.
Additional data (e.g., from a longer infiltration experiment) would have been helpful for calibrating complex models with many parameters, although long field infiltration experiments become logistically difficult. Future infiltration and leaching experiments could begin with Cl\(^-\) and P in the injection water but also add an additional conservative solute (e.g., Br\(^-\)) once the water flux reaches steady state. This would allow an additional step in the model development process: calibrate flow (e.g., \(K_s\)) using infiltration data, calibrate mobile–immobile parameters (e.g., \(D, \alpha,\) and \(q_{s,m}\)) using the data from the second conservative solute (steady flow results in no water flux between the mobile and immobile regions), calibrate the additional dual-porosity parameter (i.e., \(\omega\)) using the Cl\(^-\) data (transient flow conditions resulting in water flux between the mobile and immobile regions), and calibrate sorption parameters using the P data.

**Long-Term Phosphorus Simulations**

The 1D long-term models demonstrated the importance of several factors in long-term simulations of nutrient transport. Model A demonstrated the importance of collecting detailed soil data. Using Rosetta Lite to define soil properties, especially the silt loam \(K_s\) value, together with a single-porosity model resulted in a P load estimate over two orders of magnitude lower than the P load predicted by Model E. Conducting plot infiltration experiments or using a double-ring infiltrometer to obtain soil \(K_s\) values would be preferred over using pedotransfer functions. Model D underscored

---

**Table 5. Summary of long-term results for P leaching.** Cumulative P delivery and final P concentrations shown are at the water table at the end of each simulation. The P delivery ratio is the ratio of the cumulative P delivered (to the water table) to the cumulative P applied (to the soil surface).

<table>
<thead>
<tr>
<th>Model</th>
<th>Modeling effort</th>
<th>Cumulative P delivered (\text{kg ha}^{-1})</th>
<th>P delivery ratio</th>
<th>Final water table P concentration (\text{mg L}^{-1})</th>
<th>Difference(\dagger)</th>
<th>Difference(\ddagger)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Very low</td>
<td>0.2</td>
<td>0.04</td>
<td>0.05</td>
<td>−99.8</td>
<td>−96.8</td>
</tr>
<tr>
<td>C</td>
<td>High</td>
<td>87.1</td>
<td>16.0</td>
<td>1.64</td>
<td>−5.9</td>
<td>−5.7</td>
</tr>
<tr>
<td>D</td>
<td>Medium</td>
<td>88.5</td>
<td>16.5</td>
<td>1.67</td>
<td>−2.9</td>
<td>−4.0</td>
</tr>
<tr>
<td>E</td>
<td>Very high</td>
<td>91.7</td>
<td>16.8</td>
<td>1.74</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

\(\dagger\) Difference from the 9-yr simulations.

\(\ddagger\) Difference from Model E results.

---

**Fig. 7.** Soil water P concentration profiles for (a) Model C and (b) Model E during the 9-yr simulation. Profiles were from 1 March each year. The solution P concentration was in equilibrium with the soil P concentration. Wet years included 2004, 2008, and 2009.
the value of calibrating $D_L$, reducing the difference in the final P concentration to $-4.0\%$.

There was not a substantial difference between the 1D and 2D models when the 1D model was well calibrated. Specifically, Model D had a P delivery ratio only 2.9% lower than Model E, suggesting that HYDRUS-1D and -2D performed equally well in long-term simulations. This was surprising due to the high level of 2D heterogeneity at the field site (Fig. 1). However, flow was limited by the silt loam layer (treated as homogeneous in all models), and the 2D $K_s$ data in the gravel had mostly vertical variation and little horizontal variation. The suitability of the 1D model is a significant finding because of the large amount of effort and expense required to collect geophysical data and to develop a 2D field of $K_s$ data to inform a 2D model.

Final water table P concentrations were $\sim 2.7 \text{ mg L}^{-1}$, almost two orders of magnitude higher than the 0.037-mg L$^{-1}$ P surface water standard set for Oklahoma Scenic Rivers. These data have implications for surface water P enrichment, especially in gravel floodplains with rapid stream–groundwater interactions.

Comparing Models C and E in the long-term 2D simulations highlighted the physical process of solute flux between the mobile and immobile zones. It was expected that the dual-porosity soil profile in Model E would deliver more P to the water table than Model C (single-porosity); however, the difference between the two models was not substantial. A possible explanation is that the solute mass transfer rate ($\Gamma_m$) is high enough to move much of the solute out of the macropore and into the matrix before solute-laden water reaches the water table, resulting in a quasi-physical equilibrium. The large value of $\Gamma_m$ is influenced by two important factors in these simulations. First, the $\alpha$ term for the silt loam mantle is moderately high compared with ranges found in the literature. Second, the difference between the mobile and immobile concentrations ($c_{mo} - c_{im}$) is large. The matrix P concentration at the top of the soil profile is relatively low throughout the 9-yr simulation period ($< 1$ to $\sim 3.5 \text{ mg L}^{-1}$). In comparison, the infiltrating water of the long-term simulations contained P concentrations starting at 15 mg L$^{-1}$ at the beginning of each year and had a higher P concentration than the matrix for most of the year. In contrast, the conditions during calibration resulted in a far smaller concentration gradient, where the inflow concentration was only 1.68 mg L$^{-1}$. With the $\Gamma_m$ term being much smaller, flux from the macropores to the matrix was limited. Therefore, the dual-porosity model simulated rapid transport of solute through the macropore to the water table, explaining the large difference between Model C and Model E during calibration (Fig. 4 and 5).

Future long-term modeling attempts should simulate additional plots and sites to create a more comprehensive analysis of each of the models studied in this research. Research could determine which model components (e.g., single vs. dual porosity) would be best suited for various combinations of soil profile, initial, and boundary conditions. Future studies simulating long-term infiltration could help watershed managers better understand the lag time for multiple other conservative and highly sorbing pollutants such as nitrate and atrazine. Understanding this lag time is imperative to better manage water quality and legacy pollutants.

### Summary and Conclusions

A numerical model was calibrated to match observed data for $\text{Cl}^-$ (conservative) and P (highly sorbing). Of the three calibrated models, the Model E dual-porosity heterogeneous profile model matched the observed data for both solutes the best. The sensitivity analysis indicated that physical nonequilibrium input parameters ($\theta_{s,im}$ and $f$) were the most important, followed by dispersivity ($D_L$). Using a heterogeneous profile for gravel $K_s$ (Model C) provided only a minor improvement over a homogeneous profile (Model B), despite complex soil layering.

For long-term simulations of P leaching, the most convenient model (Model A) was inadequate (two orders of magnitude low), primarily due to the pedotransfer-function-estimated $K_s$ for the silt loam and the estimated $D_L$. Model C provided much better results, confirming the well-known fact that $K_s$ needs to be measured in the field (especially for the limiting layer), and that $D_L$ needs to be calibrated with field data. For the long-term simulations, accounting for physical nonequilibrium (the dual-porosity model) only provided a small benefit (Models D and E vs. Model C). Also, using a 2D model only provided a small improvement (Model E vs. Model D), suggesting that a well-calibrated 1D model would be sufficient for long-term simulations at this field site, especially when considering the amount of effort required for more complex models.

Modelers should evaluate their particular situation to determine whether the increased effort of 2D heterogeneity and/or dual-porosity models is needed. However, due to the poor results of the most convenient model (Model A), it is highly discouraged to conduct any long-term simulations without first calibrating the model.

### Acknowledgments

This research was funded by the Department of Biological Systems Engineering at the University of Nebraska–Lincoln, the Water for Food Global Institute at the University of Nebraska, and the USDA National Institute of Food and Agriculture, Hatch Project no. 1009760. The authors would also like to thank Dr. Gary Fox, North Carolina State University, for his input on this research. Dr. Chad Penn, USDA-ARS, provided a technical review of an earlier version of this manuscript. Ms. Ronica Stromberg provided a detailed editorial review of the manuscript.

### References


