Appendix A: Overview and updates to Agro-IBIS

The Agro-IBIS model simulates the movement of water, energy, momentum, carbon, nitrogen, and now phosphorus, in both natural and managed ecosystems. The structure of Agro-IBIS has been described in detail (Foley and others 1996; Kucharik and others 2000; Kucharik 2003), and many components of the model have been validated across a range of ecosystems at various spatial and temporal scales (El Maayar and others 2001; Kucharik and Brye 2003; Kucharik and others 2006; Kucharik and Twine 2007; Soylu and others 2014; Zipper and others 2015). Recently, Agro-IBIS was integrated with the variably saturated soil water flow model HYDRUS-1D in order to more accurately simulate the subsurface using the physically based Richards Equation (Soylu et al. 2014, Richards 1931). Soylu and others (2014) validated the new Agro-IBIS model with crop net primary production (NPP), leaf area index (LAI), and soil moisture and temperature performance with plot-level observations on silt loam soil in the YW. Subsequent work by Zipper and others (2015) validated the model's soil moisture and temperature, LAI, and corn yield results across a field with variable soil texture and groundwater depths (0-7 m), also in the YW. Since integration with HYDRUS-1D, the infiltration reduction function was modified to generate saturation-excess overland flow when the head of the surface soil layer exceeds a user-defined puddling depth, allowing for more accurate representation of surface hydrology (Dunne and Black 1970a, 1970b). Additionally, a minimum head threshold parameter from HYDRUS-1D was introduced in order to prevent extreme head gradients, which improved model stability, particularly in sandy soils.

For this study, biogeochemical cycling of P and loss of P to runoff have been added to Agro-IBIS to enable simulation of P dynamics, including interactions between surface water quality, climate, and land management. The new terrestrial P module in Agro-IBIS features P application, transformation, and loss of dissolved P to runoff; in-soil cycling of organic and inorganic forms of P; and loss of particulate-bound P with erosion (Fig. 2). The SurPhos model was integrated into Agro-IBIS to handle the simulation of inorganic P cycling in manure and soils, and loss of dissolved P to runoff (Vadas and others 2004, 2005, 2007a). SurPhos represents state-of-the-art dissolved P loss modeling for agricultural systems receiving manure. The ability of SurPhos to simulate P release has been validated in laboratory, soil-box, and field-scale settings for a variety of manure types and rainfall regimes (Vadas and others 2004, 2007a, 2007b; Sen and others 2012, Collick and others 2016). SurPhos simulates interconnected pools of dry matter and P on the soil surface, flows of manure P between pools, leaching and assimilation of manure P into soil, and loss of manure P to runoff (Fig. 2). Inorganic P fertilizer application as well as plowing are also simulated by SurPhos. Since SurPhos simulates only the soil P dynamics that are most important for dissolved P loss to runoff, it does not include organic P transformations, dynamic plant growth and uptake, or erosion. However, the model is designed to be incorporated into more complex models that simulate these processes, as is done here. Phosphorus estimation tools including APLE (Vadas and others 2012) and SnapPlus (Good and others 2012) are also based on SurPhos, making the basic methodology employed in Agro-IBIS consistent with tools currently used by managers and researchers in Wisconsin. SurPhos is described in further detail by Vadas and others (2007a) and is summarized in the following section.

SurPhos

Within SurPhos, surface manure P is divided into four pools: inorganic water-extractable P (WEP), organic WEP, inorganic stable WEP, and organic stable WEP. For manure applications containing less than 15% solids, 60% of manure P slurry is allowed to infiltrate into the soil on the day of application. Stable forms of manure P are transformed to WEP by a decomposition process that is a function of mean daily air temperature and manure moisture and age factors. SurPhos also simulates slow, physical assimilation of manure dry matter and P into soil, representing incorporation by macroinvertebrates, or rain. Required soil, climate, and hydrology input data to SurPhos include initial labile P content, soil textural properties, daily average temperature, rainfall, and runoff. For manure applications, inputs include dates of application, application area, manure dry matter and P content.

For manure slurry applications, slurry WEP that infiltrates into the soil is added to the labile P pool, where P can then be gradually transformed to soil active inorganic P. Slurry inorganic stable P that infiltrates is added to soil active inorganic P. During rain events, inorganic WEP is leached into soil, with 80% added to the surface labile pool where it may interact with runoff via desorption, and 20% infiltrating to deeper layers. For slow, physical assimilation, manure inorganic WEP is added to the soil labile pool, and manure inorganic stable P is added to the soil active inorganic pool. Loss of dissolved inorganic P in runoff from the surface soil labile pool is calculated by multiplying the concentration of labile P by 0.004. Labile P is assumed to be half of Mehlich-3 and Bray-1 soil P (Vadas and White 2010).

Soil P Pools and Organic P Cycling

Soil P pools in Agro-IBIS are represented based on the soil-plant model of Jones and others (1984) which was designed to simulate long-term dynamics of inorganic and organic P for a variety of soils under variable management practices. Other models that simulate agricultural P runoff are based on the model of Jones and others (1984), including SWAT (Neitsch and others 2011), EPIC (Williams 1995), and AGNPS (Young and others 1989). Soil P pools in Agro-IBIS contain three layers with depths of 2.5, 15, and 75 cm, respectively. The three inorganic pools include stable, active, and labile forms which are largely handled by SurPhos. Labile P represents easily desorbed P that is available to both runoff and plant uptake. The stable and active inorganic pools are not easily desorbed to runoff and maintain equilibrium based on a function of soil properties (Vadas and others 2006, 2007a).

Organic soil P pools are also based on the model of Jones and others (1984) with further modifications found in SWAT (Neitsch and others 2011) (Fig. 2). The fresh organic pool is associated with plant residue and microbial biomass. The soil humic pool is partitioned into active and stable components to account for the variation in availability of humic substances to mineralization (Neitsch and others 2011). The soil decomposition equations controlling flows of P between the organic pools and the labile pool are net decomposition equations since they account for both immobilization as well as mineralization. These transformations depend on temperature, soil moisture, and N availability (Dalal 1977; Anderson 1980). For crops, 90% of plant P is assumed to be harvested at the end of each growing season. The remaining 10%, accounting for plant residue, is transferred to the organic fresh pool. For grass PFTs, 100% of plant P is returned to the fresh organic pool on December 31st each year. For trees, 5% of plant P is returned each year (assumed litterfall), with 95% allocated to plant storage. Phosphorus taken up by plants is removed from the surface soil layer labile pool. During model development

uptake from additional soil depths was attempted but failed in non-crop systems due to the slow flow of litter-based P through the soil profile. This caused P to build up disproportionately in the surface layer. Restricting uptake to the top layer alleviates this problem in non-crop systems, but because crops may switch to non-crops and vice-versa during simulations, uptake is limited to the surface layer in all cases to maintain a consistent approach for crops and non-crops. Losses of P in runoff and soil P concentration in croplands are relatively insensitive to single versus multi-layer uptake. Future model enhancements however will focus on representing plant uptake as a function of root distribution as well as plant limitation. In the YW, soil P is generally high enough where plants are not P limited.

Daily plant uptake is simulated using a rate constant that scales daily increments of biomass for crops, and NPP for non-crops. Rate constants were calibrated to achieve typical plant P removal rates in crops (Laboski and Peters 2012). For natural ecosystems, uptake rates were calibrated to achieve realistic plant P fractions in grasses (Ryser and others 1997) and trees (Ovington and Madgwick 1958) as well as to maintain approximate soil P equilibrium, since P fluxes to and from unfertilized soils are relatively small (Walker and Syers 1976), and uptake rates in natural ecosystems are lower than in croplands (Hobbie 1992; Horrocks and others 2014). Dustfall loading of P is represented following Lathrop (1979) and Amy and others (1974), with 0.62 kg ha^{-1} yr^{-1} added to the inorganic soil P pools (Fig. 2).

Erosion, loss of particulate P, and surface runoff

Sediment yield is calculated in each grid cell using the Modified Universal Soil Loss Equation (MUSLE) (Williams 1975). The MUSLE calculates daily sediment yield as:

$$Y = 11.8 \times (Q_{surf} \times Q_{peak})^{0.56} \times K \times C \times LS \times P \quad [1]$$

where Y is daily sediment yield (tons), Q_{surf} is surface runoff (m^3), Q_{peak} is the peak runoff rate ($m^3\ s^{-1}$), K is the soil "erodibility" factor, C is the cover and management factor, LS is the topographic factor, and P is the support practice factor. K, C, LS, and P are all unitless factors ranging from 0-1. The coefficient and power constants for the MUSLE, given as 11.8 and 0.56 respectively for metric units, are location-specific parameters that are often calibrated (Sadeghi and others 2014). The soil erodibility factor, K, is computed at the beginning of each year using the sand, silt, and clay fractions of the soil as well as the organic carbon content, as done in EPIC (Williams 1995). The cover and management factor, C, is assigned based on land cover type. For corn, the C factor is calibrated to achieve reasonable rates of sediment yield at the grid-cell level when compared with local observations (Stuntebeck and others 2011). For the remaining biomes, C factors are assigned based on the relative differences between C factors observed by Panagos and others (2015) for a range of crop and non-crop types. Barren, corn, soy, wheat, alfalfa, pasture, grassland, forest, and wetland are assigned C values of 0.9, 0.8, 0.5, 0.45, 0.40, 0.20, 0.10, 0.05, and 0.01, respectively. The topographic factor LS is calculated using slope and slope length of each grid cell, as done in the USLE (Wischmeier and Smith 1978) and EPIC (Williams 1995). The support practice factor P represents conservation practices for sloping soils vulnerable to erosive rains, like contour tillage, strip-cropping, and terracing. Because slopes are generally small in the YW, and these practices are typically rare, the P factor is set to 1.0 throughout the watershed. Daily loss of sediment-bound P is calculated using the total amount of P in the surface layer, daily sediment yield, and a P enrichment ratio (PER) (McElroy and others

1976; Williams and Hann 1978, Sharpley 1985). Total P in the surface layer is the sum of P in all five soil pools including the labile pool, as recommended by Vadas and White (2010). The PER is defined as the ratio of the concentration of P transported with sediment to the concentration in the surface soil layer. It is calculated using a relationship developed by Sharpley (1985) in which there is a linear relationship between the logarithms of soil loss and the PER:

$$\ln(PER) = 1.21 - 0.16 \ln(sed) \quad [2]$$

where *sed* is daily soil loss in kg ha^{-1}.

Surface runoff is calculated based on the following governing equation for water balance in Agro-IBIS:

$$P = Q_{surf} + ET + D + \Delta S \quad [3]$$

where *P* is precipitation (mm), *ET* is evapotranspiration (mm), *D* is drainage exiting the soil column (mm), and ΔS is the change in storage (mm). The model solves the mixed-based Richards equation to simulate water flow in the unsaturated zone (Soylu and others 2014). Surface runoff accumulates as a puddle on the soil surface when the upper soil layer reaches saturation, or the infiltration rate exceeds the maximum infiltration rate allowed. Peak runoff rate, Q_{peak}, is calculated for each grid-cell following the modified rational method which assumes that for a given rainfall event having a constant and uniform intensity over a basin, the rate of runoff will increase until the time of concentration when the entire subbasin is contributing to flow at the outlet. Because hydrologic transport is handled separately by THMB (Terrestrial Hydrology Model with Biogeochemistry, Appendix B), no channel is assumed in land grid cells simulated by Agro-IBIS and time of concentration is computed for overland flow only.

Representation of land cover

Agro-IBIS has been modified to allow simulation of annually changing land cover grids that can contain both crops and non-crops. Additionally, biomes have been modified to accommodate the range of land cover classes found in the YW. Pre-existing biomes in the model that can be readily used in the YW include corn, soy, wheat, temperate deciduous forest, and grassland. No new plant functional types (PFTs) have been developed but new biomes have been added based on pre-existing PFTs. New biomes include alfalfa, hay, wetlands, and urban areas. To simulate alfalfa, C3 grasses are grown. To simulate hay, grassland is grown, consisting of both C3 and C4 grasses. Ninety percent of the aboveground biomass in both alfalfa and hay is harvested at the end of each year, with corresponding adjustments made to both carbon and phosphorus pools. To represent areas of the watershed growing small grains, spring wheat is grown as a proxy. For fruits and vegetables, soy is grown in order to best capture plant phenological trends. To represent wetlands, C4 grasses are grown.

Urban areas are represented as combinations of impervious area and well-maintained turf cover following Schneider and others (2012). Turf cover is represented by growing C4 grasses subject to a maximum canopy height of 11 cm and normal vegetation albedo. Soil albedo is set to 0.28 in all urban cells. Four classes of urban areas are simulated: high-, medium-, and low-intensity, and open space. Except for open space, each of the classes is assigned a fraction of pervious area,

assumed to be growing turf cover. These fractions are 10, 35, and 65% respectively for high-, medium-, and low-intensity urban cells. The remainder of each cell is assumed to be impervious. Each urban cell is first simulated as 100% turf cover, and an adjustment is made to water balance terms to account for fraction of the impervious area. For runoff, the Curve Number approach (NRCS 2004) is used to calculate the amount of runoff generated from the impervious area. The pervious and impervious fractions are then used to scale runoff from turf cover and impervious area, respectively. Additional model quantities are scaled according to the pervious fraction in each urban cell, including ET, drainage, biomass, NPP, LAI, net ecosystem exchange (NEE), sediment yield, and nitrate leached from the soil column. Open space grid cells are left unadjusted, with no assumed impervious cover.

Code Parallelization

Agro-IBIS simulations are costly in terms of CPU time, however the fact that there is no data exchange between land points allows for efficient process level parallelization. To accomplish this, the geographic simulation domain is divided into several smaller subdomains and Agro-IBIS is executed for each of the subdomains as a separate process on its own core. After all subdomain simulations are completed, the subdomain results are assembled to cover the original domain. Domain division, process launch and monitoring, and results assembly are performed automatically by a custom written parallelization software. A 60-core/120-thread machine with four Intel Xeon E7-4890 processors is used to run all simulations. A 250 year simulation (including spin-up) of the Yahara Watershed using a 220 m grid and 25,881 grid cells takes approximately 28 hours of real time.

Appendix B: THMB Overview and Updates

Agro-IBIS is coupled to the THMB model to enable simulation of water, sediment, and P transport at the watershed scale, including delivery of P loads to the YW lakes and the Rock River. The THMB model, formerly called HYDRA, is a physically-based hydrologic routing model that has been coupled to Agro-IBIS and validated in several previous studies (Donner and Kucharik 2003, Donner and others 2004, Donner and others 2008). By linking the topographic data of a prescribed stream network and river morphological characteristics to a set of linear reservoir functions, the model simulates the temporal variability of water flow and storage in hydrologic systems (Coe 1998, 2000; Coe and others 2008). Donner and others (2002) added a nitrogen transport module to the THMB modeling framework. In this study, sediment and P transport have been added to THMB.

Soils in the YW are mostly silt loam, indicating a fine particle size. Therefore, the sediment transport functions in the THMB model have been developed based on a fine sediment transport model from Patil and others (2012). The mass balance of sediment loads in the stream network S_r is described by the following equation:

$$\frac{d(S_r)}{dt} = \Sigma\, S_{in} + S_g\,(1 - A_w)\; +\; E\; +\; I\; -\; L\; -\; (\frac{S_r}{T_r}) \qquad [4]$$

where S_{in} is the sediment load from upstream cells; S_g is the sediment load from land; A_w is the fractional standing water area in the cell; E is the in-channel erosion rate; I is the sediment load from point sources; L is the sediment removal due to in-channel processes; and T_r is the river water residence time (Donner and others 2002; Coe and others 2008). Equations for in-channel erosion rate (E) and in-channel deposition (L) can be found in Patil et al. (2012).

The P loads are divided into two major forms, dissolved P and sediment P. These two forms have different transport mechanisms. Dissolved P is considered a conservative solute with the mass balance of dissolved P in the stream network, P_d^r, described by:

$$\frac{d(P_r^d)}{dt} = \Sigma\, P_{in}^d + \left(\frac{P_s^d}{T_s} + \frac{P_d^d}{T_d}\right)\,(1 - A_w)\; + I^d\; -\; (\frac{P_r^d}{T_r}) \qquad [5]$$

where P_{in}^d is the dissolved P flux from upstream cells; P_s^d and P_d^d are the dissolved P loads in surface runoff and subsurface drainage from the land, respectively; T_s and T_d are the residence times of cells for surface runoff and subsurface drainage, respectively; and I^d is the dissolved P load from point sources. The transport function of dissolved P is similar to the N transport function developed by Donner and others (2002), which has the capability to be adapted to simulate the transport of different kinds of semi-conservative chemicals. The transport process for sediment P is related to sediment processes. The erosion and enrichment of sediment P on the land are handled by Agro-IBIS, while the in-channel erosion and deposition of sediment P are handled by THMB. The mass balance of sediment P in the stream network, P_r^s, is described by the following equation:

$$\frac{d(P_r^s)}{dt} = \sum P_{in}^s + \left(\frac{P_s^s}{T_s} + \frac{P_d^s}{T_d}\right)(1 - A_w) + E^s + I^s - L^s - (\frac{P_r^s}{T_r}) \qquad [6]$$

where P_{in}^{s},P_{s}^{s} ,P_{d}^{s}, and $I^{s}$are different sources of sediment P loads with similar descriptions as the dissolved P loads; E^{s} is the sediment P load from in-channel erosion; and L^{s} is the sediment P removal via in-channel deposition. In-channel erosion is calculated as:

$$X_p = \frac{k_{XP}}{Y_E^b} \quad [7]$$

$$E = k_E X_p Y_E \quad [8]$$

where $X_p$is the in-channel enrichment ratio; $k_{xp}$is the enrichment coefficient; Y_E is the sediment erosion rate; *b* is the enrichment optimization parameter; and k_E is the sediment P erosion coefficient. Enrichment functions of this form are commonly used in process-based models for P transport simulation (Viney and others 2000). In-channel sediment P deposition is calculated with the following equation:

$$L = D_S\left(\frac{C_{pS}}{C_s}\right) \quad [9]$$

where D_s is the amount of sediment deposition; C_{Ps} is sediment P concentration; and C_s is sediment concentration. In general, the amount of sediment P deposition is considered to have a proportional relationship with the sediment deposition in the stream network. In addition to erosion and deposition of sediment P, the in-channel storage of sediment P is also represented in the model. Erosion of sediment P can be limited depending on the in-channel storage of sediment P.

Internally-drained basins – areas with no contribution to surface runoff but are connected on the subsurface level – are common in the Yahara Watershed (Figure B1) due to the influence of glaciation as recently as 10,000 years ago. As a result, THMB only simulates the subsurface transport processes in these areas. Surface runoff and sediment/nutrient transport are considered to be isolated from the main watershed.

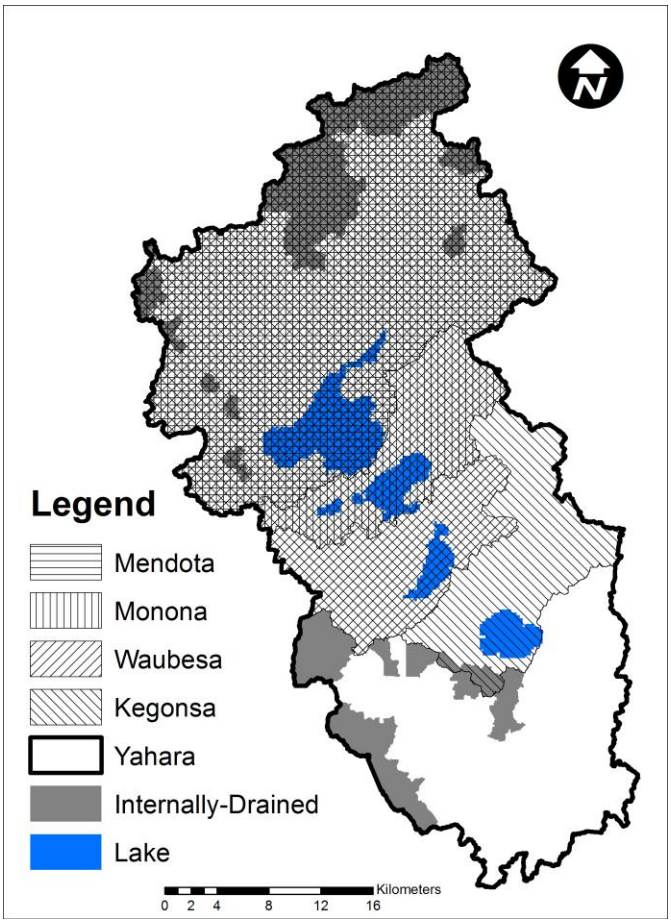

Figure B1. Map showing total drainage areas for each lake as well as internally-drained basins of the Yahara Watershed.

Appendix C: Yahara Water Quality Model

The Yahara Water Quality Model predicts summer water quality variables in the four mainstem Yahara lakes given annual direct drainage loads to each lake (Carpenter and Lathrop 2014). The model first computes P mass balances for the four lakes using mass balance principles and empirical relationships among terms of the mass balance. The mass balance is computed between 1 November and 31 October to correspond with fall-overturn estimates of P mass in the water column of each lake. Then, summer water quality variables are computed using empirical regression models. These models predict summer water quality from terms of the P mass balance. Summer is defined as 30 June to 7 September, which is reliably a period of summer stratification. Water quality variables are TP (total P concentration in surface water), DRP (dissolved reactive phosphorus concentration in surface water), chlorophyll (chlorophyll a concentration in surface water), and Secchi disc transparency. Empirical relationships are computed using 33 full years of data for lakes Mendota and Monona (1975-2008) and 28 full years of data for lakes Waubesa and Kegonsa (1980-2008). Datasets and regression models are described by Lathrop and Carpenter (2013) and Carpenter and Lathrop (2014). Here we explain how the Yahara Water Quality Model is used to compute water quality in the four lakes given direct drainage loads computed by THMB.

Annual loads and exports

The four mainstem lakes, from upstream to downstream, are Mendota (Me), Monona (Mo), Waubesa (Wa), and Kegonsa (Ke). For Lake Mendota, which has phosphorus inputs from land but no inflow of phosphorus from other lakes, the annual load for year *k*, $L_{k,Me}$ in kg yr^{-1} is the same as the direct drainage load, $M_{k,Me}$:

$$L_{k,Me} = M_{k,Me} \quad [10]$$

For the other lakes, annual loads also include transfers from upstream lakes.

$$\begin{aligned} L_{k,Mo} &= M_{k,Mo} + E_{k,Me} \\ L_{k,Wa} &= M_{k,Wa} + E_{k,Mo} \\ L_{k,Ke} &= M_{k,Ke} + E_{k,Wa} \end{aligned} \quad [11]$$

In equations [11], *E* is the annual export (kg yr^{-1}). *E* is estimated from *M* using models fitted for each lake as explained below.

Computing P dynamics over each year

For each lake *l*, P dynamics during each year *k* from 1 November to 31 October are modeled as

$$\frac{dP_{k,l}}{dt} = (1 - w_l)\, L_{k,l} - (s_l + h_l) P_{k,l} \quad [12]$$

where *P* is the mass of P phosphorus (kg) in year *k* in lake *l*, *s* and *h* are sedimentation and outflow coefficients for *P*, and *w* is the outflow coefficient for *L*. Comparisons using the Akaike information criterion (AIC) show that model fits are improved significantly by using two outflow coefficients. From equation 12 it follows that cumulative exports from a lake during each year are given by

$$\frac{dE_{k,l}}{dt} = wL_{k,l} + h_{k,l}\, P_{k,l} \quad [13]$$

For integration within each year, the initial values of *P* and *E* are the 1 November observation of P mass in the water and 0, respectively. Values of *P* and *E* after 1 year (i.e. 31 October of the next year) are obtained by integrating equations 12 and 13 by the Euler method using 30 time steps.

Appendix D: Model Calibration and Evaluation

Agro-IBIS

The goals for model performance were to generate reasonable values of sediment and P loss at the field scale in Agro-IBIS and also to generate values of in-stream sediment and P load within THMB that agreed with the long-term USGS gage record. To achieve the model performance goals with minimal calibration in Agro-IBIS, calibration was limited to three parameters. These included the MUSLE coefficient and power constants (Eq. 1) as well as the PER (Eq. 2). To achieve satisfactory simulation of sediment yield at both the field scale and in the river network, we calibrated the location-specific MUSLE coefficients, originally given as 11.8 and 0.56 for metric units (Williams 1975). The final calibrated values of these parameters were 5.9 and 0.35, respectively.

To guide P cycling and the magnitude of P loss in runoff, the soil P pools were initialized by first setting the labile pools to roughly correspond with observed values of soil test P for local croplands and prairies (Bennett 2004). As in SurPhos, the labile P pool was assumed to be half of Bray-1 soil test P (Vadas and White 2010). For the calibrated historical run, ME, the P pools were allowed to spin up for 25 years prior to the start of the simulation time period. During the spin-up, manure was applied to all cropland cells at a rate of 50 kg ha^{-1} y^{-1}. By the start of the simulation time period (1986), average surface soil P concentration in croplands was approximately 176 ppm.

To achieve reasonable values of P loss in both the dissolved and sediment form, the PER was calibrated. This was required since without calibration of the PER, simulated sediment P in the river network was significantly higher than the observed gage record and also exceeded typical edge-of-field loss rates observed locally (Stuntebeck and others 2011). The PER was the optimal parameter to calibrate since simulated sediment yield and dissolved P loads/losses performed well without calibration, but sediment P losses did not. The coefficient used in the equation for the PER (Eq. 2) was calibrated from its original value of 0.16 to 0.02, and an additional multiplier was included in the equation having a value of 0.09.

SurPhos and the soil P routines used in Agro-IBIS have been validated previously (Vadas and others 2004, 2007a, 2007b; Vadas and White 2010; Sen and others 2012, Collick and others 2016). However, we wished to verify that the calibrated version of Agro-IBIS simulated a reasonable range of soil P values within the YW. To do this, we compared simulated soil test P from the calibrated historical run (ME) with three separate sources of local observations as well as observations reported in the literature. Simulated soil P in the second soil layer (2.5-15 cm) (Fig. D1) was compared with a set of soil P observations made in three subwatersheds of the YW in the years 2011-2014 (Kyle Minks, Dane County Land and Water Resources Department, unpublished data). The measurements were made on 1,869 anonymous crop fields spanning a variety of cropping systems and management practices. The data were collected using either the Bray-1 or Mehlich-3 methods. The average soil test P measured across all three subwatersheds was 68 ppm. Simulated soil test P in the second soil layer of all YW croplands, averaged over the 1986-2013 historical period, was 73 ppm. This suggested that the model simulated a mean value of soil P in croplands consistent with local observations.

We also compared mean simulated soil test P from ME with observations made from 2003-2008 at the Discovery Farms and Pioneer Farm (DPF) sites located throughout Wisconsin (Stuntebeck and others 2011). At these sites, (at least) 39 separate soil P measurements were made on 18 fields that spanned a range of landscapes, soil textures, and farming systems typical of livestock farms in southern Wisconsin. Based on the reported values, we calculated a mean soil test P of 79 ppm, with a standard deviation of 37 ppm. Compared to the average simulated soil P value of 73 ppm, this result further verified reasonable model performance.

Andraski and Bundy (2003) measured soil test P in the 0-2 cm layer at three sites in Wisconsin, with Mehlich-3 measurements ranging from 7-325 ppm, and Bray-1 measurements ranging from 5-274 ppm. Simulated soil test P in the 0-2.5 cm layer for croplands averaged 176 ppm. falling within the ranges observed by Andraski and Bundy. Their observations across soil layers indicated stratification, which they attributed to unincorporated manure treatments in conjunction with no-till. This observation is consistent with other studies showing that soils receiving long-term broadcast application of fertilizer and manure without physical incorporation can accumulate P in the soil surface, increasing the potential for P loss in overland flow (Oloya and Logan 1980; Sharpley and others 1994; Andraski and Bundy 2003; Sharpley 2003; Smith and Warnemuende-Pappas 2015). Sharpley (2003) reported Mehlich-3 values for corn/soy rotations receiving long-term manure applications ranging between 128-961 ppm in the 0-5 cm layer which declined rapidly with depth. In a study by Smith and Warnemuende-Pappas (2015), stratification was observed in corn/soy rotations managed with both no-till and vertical tillage (227 ppm in the 0-5 cm layer compared with 102 ppm in the 5-15 cm layer for no-till; 154 ppm in the 0-5 cm layer compared with 114 ppm in the 5-15 cm layer for vertical tillage.) For our simulations, tillage to a 10 cm depth occurred each year at the time of planting, immediately following the springtime manure application. Both of the remaining manure applications (winter and fall) were unincorporated. Stratification was evident in our simulations (Fig. D1), with cropland soil P in the second soil layer averaging 59% less than the first layer (176 ppm versus 73 ppm). This helps confirm the ability of the model to capture soil P dynamics across soil layers.

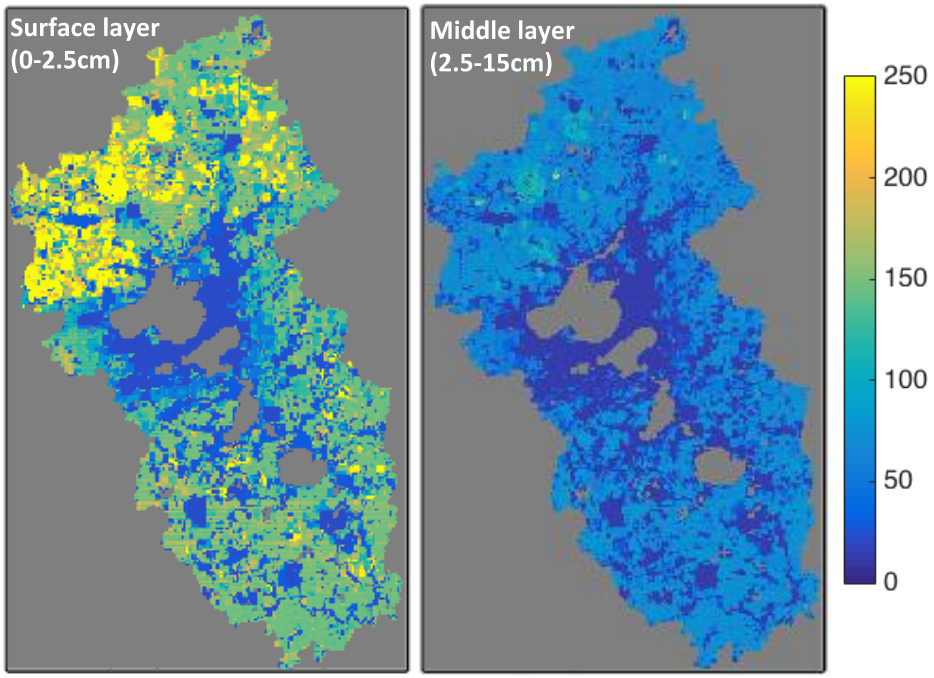

Figure D1. Mean simulated soil test P (ppm) for the calibrated historical simulation (ME) in the top two soil layers, averaged from 1986 to 2013.

Edge-of-field observations on local farms and elsewhere indicate that typical values of sediment and P yield span a wide range of values. At the Discovery and Pioneer farm (DPF) sites, annual sediment yield ranged from 3-5,604 kg ha^{-1}, with a mean of 748 kg ha^{-1} (Stuntebeck and others 2011). Observed average annual total P in runoff ranged from 0.03-7.8 kg ha^{-1}, with a mean of 2.2 kg ha^{-1}. Sediment and dissolved forms of P yield were roughly in proportion (50:50). Simulated values of sediment yield (ME) ranged from 670-2,310 kg ha^{-1}, with a mean of 1,325 kg ha^{-1}. Mean annual simulated TP was 1.0 kg ha^{-1}, and spanned a range of 0.5-2.1 kg ha^{-1}, with the ratio of sediment and dissolved forms roughly 53:47. Modeled values of P yield were generally lower than measurements made at DPF. Results from a modeling study conducted for grazing-based dairy farms in Wisconsin also suggest that simulated values of P yield from Agro-IBIS tend to be low yet reasonable. In that study, average annual whole-farm P losses were estimated to be 0.5-1.8 kg ha^{-1} (Vadas and others 2015). These values were consistent with values of P yield simulated by Agro-IBIS yet were cited as being low due to low erosion rates on vegetated pastures. Local observations made for the Wisconsin Buffer Initiative observed a wide range in P yield on a single field between just two years, with 22.5 kg ha^{-1} in the first year and <0.1 kg ha^{-1} in the following year (UW-CALS 2005).

The Agro-IBIS simulation (ME) was not designed to mimic the cropping systems and management practices employed at DPF or any other validation site, thus these comparisons serve only as a general check for simulated magnitudes. Further studies from agricultural fields in the U.S. and beyond indicate a wide range in observed sediment and P yield observations, sometimes explained by variation in local environmental or management-based conditions

(Wischmeier and Smith 1978; Laflen and others 2004)**,** but often due to unexplained sources of variation (Nearing and others 1999).

THMB

Observations of streamflow, sediment load and P load collected from six USGS gages were used to calibrate and validate THMB (Fig. D2). Continuous monthly streamflow data were split into the periods 2004-2013 for calibration and 1994-2003 for validation. Calibration and validation of sediment load was performed using data from three of the six gages (05427718, 05427965 and 05427948), and for P load from two of the six gages (05427718 and 05427948). THMB was calibrated using a stepwise procedure. First, the simulation time step was chosen to optimize model performance while considering computational expense. The time step was chosen to be twelve minutes, shorter than the time step of one hour used in previous THMB studies (Coe, 1998; Coe et al., 2008). Second, the streamflow was calibrated based on historical stream gauge data. Major parameters adjusted in this step included: u_{o1}, the minimum effective river flow velocity (u_{o1}=0.35 m s^{-1}); i_o, the reference gradient (i_o=0.0001 m m^{-1}); and p_o, the reference wetted perimeter (p_o=25 m). Physical explanations of these parameters have been previously given by Coe and others (2008). Third, we calibrated sediment transport using the calibrated streamflow. Major parameters adjusted in this step included: e_0, the reference erosion rate (e_0=2.0x10^{-4} kg m^{-2} s^{-1}); and n_p, the sediment porosity (n_p=0.5). Finally, P transport was calibrated based on the calibrated streamflow and sediment loads. Major parameters adjusted in this step included b, a phosphorus enrichment optimization parameter (b=0.27), and k_E, a sediment P erosion coefficient (k_E=0.0002).

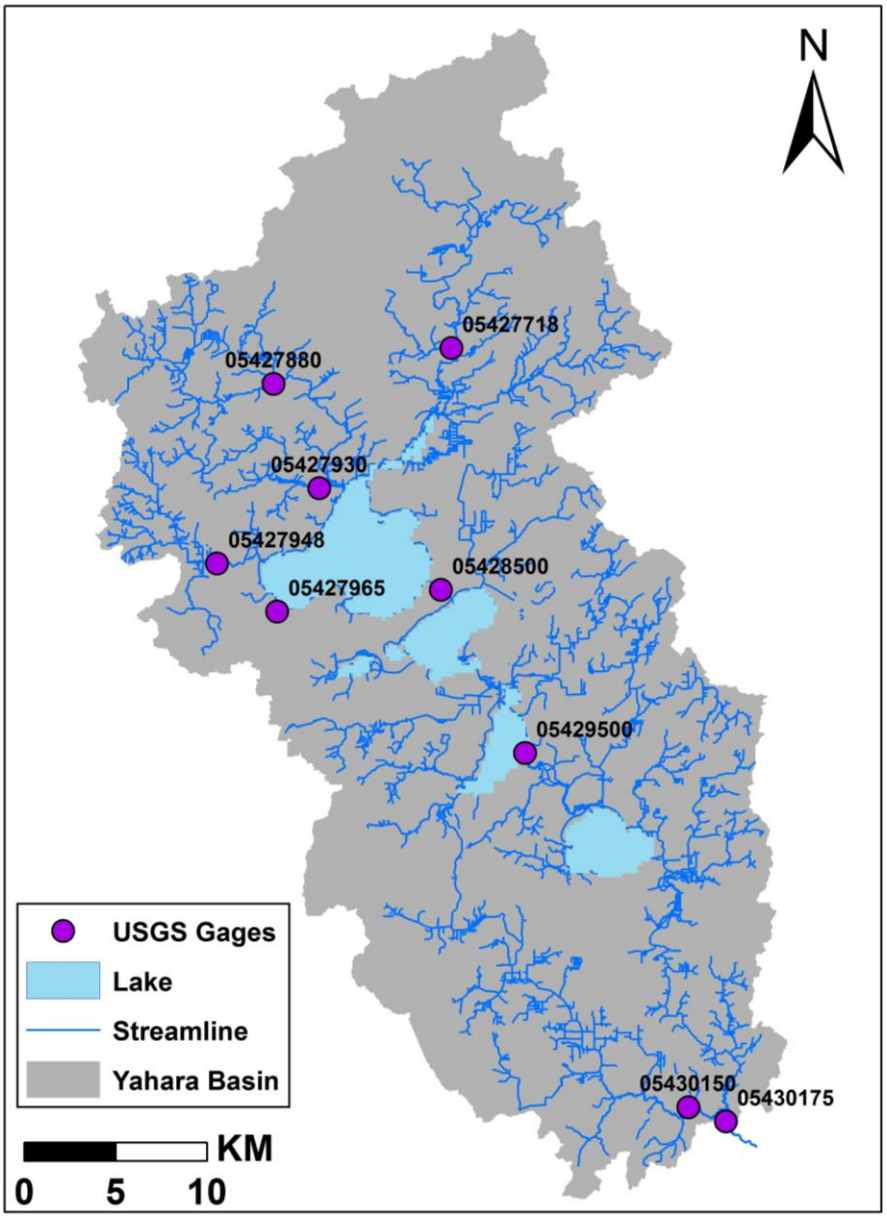

Figure D2. Map of USGS gages in the Yahara Watershed

USGS Gage ID	Calibration			Validation		
	NSE	R^2	RMSE ($m^3 s^{-1}$)	NSE	R^2	RMSE ($m^3 s^{-1}$)
05427718	0.55	0.55	0.46	0.09	0.46	0.31
05427965	0.63	0.78	0.03	0.59	0.77	0.03
05428500	0.49	0.53	1.93	-	-	-
05429500	0.51	0.53	2.75	0.26	0.45	2.40
05430150	0.73	0.82	0.61	0.29	0.60	0.65
05430175	0.66	0.67	3.55	0.44	0.56	3.27

Table D1. Monthly streamflow simulation performance in the Yahara Watershed using Nash-Sutcliffe efficiency (NSE), coefficient of determination (R^2), and root-mean-squared-error.

Table D1 lists monthly streamflow simulation performance at different locations in the Yahara Watershed. The performance varied across locations and was generally acceptable. Simulations of sediment and P loads were closely related to streamflow. As shown in Table D2, the performance of sediment and P transport was worse than streamflow due to the accumulation of errors from streamflow to sediment and P loading. Future model improvements should focus on increasing the accuracy of surface runoff simulation since both sediment and P loads are highly sensitive to changes in surface runoff.

We evaluated simulated P transport using records of direct drainage P loads to Lake Mendota (Lathrop and others 1998; Carpenter and Lathrop 2008; Lathrop and Carpenter 2013). The comparison showed that the model is able to capture high and low P loads with no significant bias (Figure D3). Overall performance was deemed acceptable because annual TP loads to the lakes were required to evaluate water quality outcomes using the Yahara WQ Model. Given that the magnitude of annual loads was well-captured by the model, this satisfied the minimum performance criteria required for our study.

USGS Gage ID	Sediment Mass Flux				Phosphorus Mass Flux			
	Calibration		Validation		Calibration		Validation	
	R^2	RMSE (kg s^{-1})	R^2	RMSE (kg s^{-1})	R^2	RMSE (kg/s)	R^2	RMSE (kg s^{-1})
05427718	0.24	0.27	0.35	0.30	0.29	4.51e-4	0.16	3.99e-4
05427965	0.29	0.16	0.31	0.25	0.24	1.93e-4	0.43	2.07e-4
05427948	0.58	0.012	0.61	0.025	-	-	-	-

Table D2. Monthly sediment and phosphorus transport simulation performance.

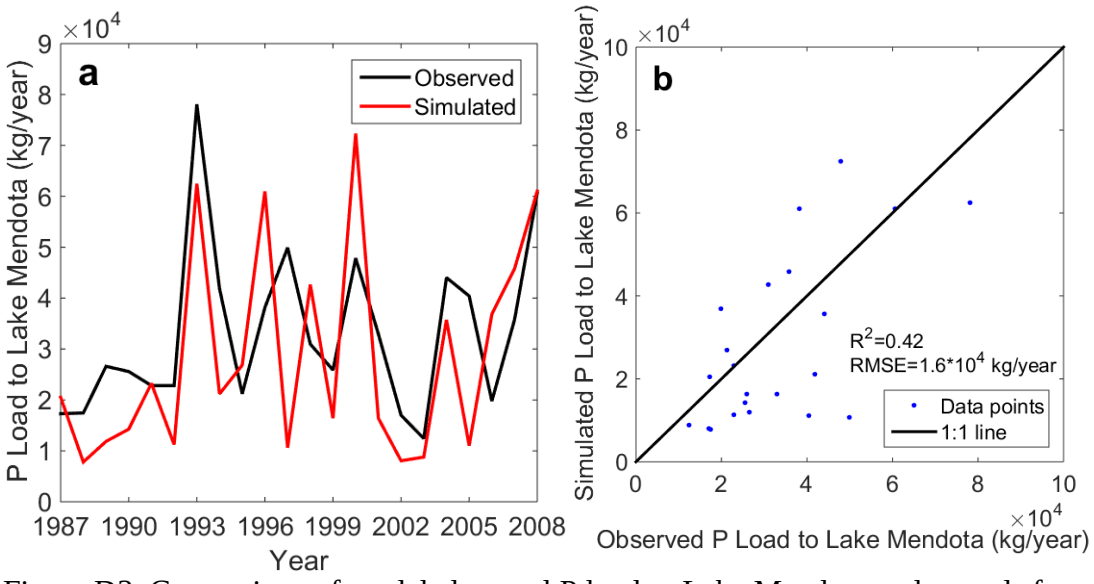

Figure D3. Comparison of modeled annual P load to Lake Mendota and records from previous studies.

Yahara WQ Model

Observations of the P budget and water quality in the four lakes were used to estimate the empirical coefficients used in the Yahara WQ Model. *P*, the mass of phosphorus in each lake at fall overturn, was inventoried for each lake on 1 November of each year (Lathrop and Carpenter 2013; Carpenter and Lathrop 2014). Annual values of *M* and *E* for each lake were measured between 1 November and 31 October for each year (Carpenter and Lathrop 2014). Summer water quality (TP, DRP and Secchi transparency) were measured between 30 June and 7 September each year (Lathrop and Carpenter 2013). Years of observation were 1975-2008 for lakes Mendota and Monona and 1980-2008 for lakes Waubesa and Kegonsa. Datasets are archived by the North Temperate Lakes Long-Term Ecological Research site (http://lter.limnology.wisc.edu).

Summer chlorophyll was computed from TP using the model of Filstrup and others (2014). Multiple regressions for each lake were used to predict summer water quality (TP, Secchi transparency and chlorophyll) from terms of the annual phosphorus budget (Lathrop and Carpenter 2013). Summer DRP data were used to estimate the probability of hypereutrophy, defined as DRP > 0.005 mg/L. Probability of hypereutrophy was predicted from summer TP using logistic regression.

Regressions to predict summer water quality from terms of the annual phosphorus budget were computed with the lm() package in R. Results of these models are presented in Lathrop and Carpenter (2013) and Carpenter and Lathrop (2014).

Two new model fits were carried out for this paper. Logistic regression to predict the probability of hypereutrophy from summer TP was computed with the glm() package in R using the binomial error distribution. Parameters of equations 12 and 13 (Appendix C) were estimated by

fitting the one-year ahead projections of the two equations to the observed annual time series of *E*, *L*, and *P* using maximum likelihood (Hilborn and Mangel 1997).

Comparisons based on AIC showed that the probability of hypereutrophy was best predicted using a single model fit to data from all four lakes, rather than individual models for each lake. Parameter estimates for the logistic regression were intercept = 8.89 (s.e. = 1.91) and log[TP] effect 3.42 (s.e. = 0.69).

Maximum likelihood estimates of *s*, *h*, and *w* obtained by fitting equations 12 and 13 to data are presented in Table D3.

Lake	**s (y^{-1})**	**h (y^{-1})**	**w (y^{-1})**
Mendota	0.342	0.688	0.220
Monona	0.893	0.021	0.538
Waubesa	0.216	0.466	0.879
Kegonsa	0.557	0	0.892

Table D3. Maximum likelihood estimates for parameters *s*, *h*, and *w* of the Yahara

All computations for the YWQM were performed with the base package of R 3.2.3 (R core team 2015).

Appendix E: Climate Inputs

Inputs of daily precipitation, maximum and minimum air temperature, solar radiation, relative humidity, and wind speed are required by Agro-IBIS. Daily precipitation observations from the official NOAA automated weather observation station at the Madison, Wisconsin airport in the north-central part of the watershed were used for 1986-2000 and applied across the model domain (NCDC 2015). Spatially-variable radar-derived daily precipitation estimates (4 km x 4 km resolution) compiled by the National Centers for Environmental Prediction were used for 2001-2013 (NCEP 2015). Daily maximum and minimum air temperature and wind speed from the Madison airport were used for 1986-2013 and applied across the model domain (NCDC 2015). Daily solar radiation and relative humidity data were used from a weather station in Arlington, Wisconsin, representing the closest geographic source of daily solar radiation measurements. These data were used for the 1986-2013 period and applied across the model domain (UW-Extension 2014). Finally, random years between 1986 and 2013 were selected to create the meteorological variables used during the model spin-up period (1786-1985).

Appendix F: Soil Inputs and Groundwater Specifications

Soil layers simulated by Agro-IBIS included a 0.5 cm surface layer followed by 99 deeper layers of linearly increasing thickness. Soil textural class in each grid cell was assigned based on 11 possible categories in the USDA Soil Survey Geographic Database (SSURGO) (USDA 2013). Associated physical properties including percent sand, silt, and clay were then obtained for each class based on data in Campbell and Norman (1998). For each soil class, Van Genuchten soil water characteristic curve relationships (Van Genuchten 1980) were used with retention parameters from Carsel and Parrish (1988). Soils in the YW are primarily silt loam, but to minimize convergence issues within the soil physics routines of HYDRUS-1D, the soil texture was assumed constant throughout the soil profile using properties of the top 2.5 cm surface layer, and clay and sandy soils were substituted with silt loam and loam, respectively, which account for about 15% of grid cells in the watershed. To set the pressure head bottom boundary condition in Agro-IBIS, a constant soil profile depth of 10 m was set across the model domain. The bottom boundary condition for the variably-saturated groundwater flow equation was either (1) a constant head equal to the difference between the soil column depth and the water table depth (WTD) if the WTD was less than 10 m; or (2) free drainage if the WTD was greater than or equal to 10 m. Values for the WTD were determined based on a steady-state, present day (2014) groundwater flow simulation for Dane County (Parsen and others 2016).

Appendix G: LULC Determination

To generate land cover grids for the watershed in each year of simulation, land-use/land-cover (LULC) was divided into 17 biophysically distinct categories and interpolated to a 220 m by 220 m spatial resolution. For recent years (2001-2013), maps were based on land use data from Dane, Columbia, and Rock counties, and national-level datasets including the National Land Cover Database (Homer and others 2015), and the Cropland Data Layer (USDA 2014) (see Gillon and others 2016 for more details). We estimated LULC in 1786 using a pre-settlement vegetation map of Wisconsin that was created using U.S. General Land Office Notes from the mid-1800s, compiled by Finley in 1976 and later digitized (WDNR 2011). The landscape at the time of the land surveys was dominated by oak forest, oak openings, prairie, and wetlands. While Euro-American settlement of the watershed began in the 1830s, we assumed, based on historical U.S. agricultural census data for Dane County (UVGSDC 2004), that LULC changed from the pre-settlement condition to peak agricultural development in 1870. This LULC condition remains until 1890 when urban areas begin to emerge. We digitized urban areas from Dane County plat maps for 1890, 1911, and 1953 as well as from Wisconsin Land Economic Inventory for 1929. We also used digitized tax parcel maps from 1962, 1973, and 1981 to estimate urban areas based on a maximum parcel area threshold of 6 hectares. Between 1992 and 2000, urban areas were assigned using the Wisconsin Land Cover database derived from satellite imagery captured in 1992 (WDNR 2013). Prior to 1920, all non-urban areas except open water and wetlands were classified as corn. From 1920 to 2000, LULC for these non-urban areas was rotated using LULC from 2008 to 2014. While non-urban LULC was likely different between these two periods, we made this assumption in order to institute crop rotations within the model and capture its important effect on the spin-up of carbon and nitrogen.

Appendix H: Determination of manure and fertilizer rates

For nutrient inputs, historical manure P and N production was determined based on estimates of watershed animal units and milk production. First, livestock numbers (cattle, swine, poultry, horses, and sheep) and milk production were determined and interpolated for each year between 1870 and 2014 based on historical agricultural census data for Dane County (Hibbard 1904; USDA-NASS 2013, 2014). Next, county animal units were calculated for each livestock type based on the interpolated animal numbers and historical estimates of animal production weights (USDA-ERS 2015).

In order to scale the county estimates to the Yahara watershed and spatially distribute manure production, we developed a livestock operation inventory for 2001-2014 based on unpublished spatial data from the Dane County Land & Water Resources Department, Concentrated Animal Feeding Operation data from the Wisconsin Department of Natural Resources (WDNR 2015), and milk producer locations from the Wisconsin Department of Agriculture, Trade, and Consumer Protection (WDATCP 2015). This inventory includes livestock type, estimated number of animal units, and estimated manure hauling distances. In addition, we used historical aerial imagery from Google Earth to document whether operations were active or inactive in a given year. Due to data limitations, we assumed that all operations in 2001 existed from 1870 to 2000 even though the total number of farms has decreased through time (UVGSDC 2004). Based on the 2001-2014 analysis, it was determined that the number of animal units in the Yahara

watershed is approximately 38% of that for Dane County. We assumed that this percentage is constant from 1870 to 2000.

Manure P and N production (as excreted) per animal unit for each livestock type (excluding dairy cows) was then calculated for each livestock operation based on guidelines from the (USDA 2008). Dairy cow manure P and N production were separately calculated using regression equations based on milk production (Nennich and others 2005), which has in addition to manure steadily increased per cow over the past half-century (Gillon and others 2016). The estimated manure N was reduced by 70% at all operations to account for ammonia volatilization and first year crop availability (Laboski and Peters 2012) which can vary substantially due to storage/collection characteristics and weather conditions (Rotz and others 2014; Powell and Rotz 2015). Land available for spreading manure P and N from a single operation included all crop and pasture land within a radius equal to the estimated manure hauling distance. More details on the method for land application including treatment of overlapping hauling areas can be found in Appendix I.

Following the determination of manure P and N application rates, fertilizer P and N application rates were calculated for the following LULC types: alfalfa, corn, fruits/vegetables, small grains, and soybeans. Recent (2007-2013) fertilizer application rates were determined based on University of Wisconsin – Extension nutrient application guidelines (Laboski and Peters 2012) assuming high yield potential soils, optimal soil nutrient status, and high yield goals. If manure had already been applied in a grid cell, additional fertilizer was applied at a lower rate only if the LULC type was corn. Application rates for 1945 to 2006 were determined based on county-level fertilizer sales data compiled by Alexander and Smith (1990) and Gronberg and Spahr (2012). These values were used to create a ratio of total fertilizer sold in a given year to that in 2006; this ratio was then multiplied by the 2007-2013 rates.

For all simulations, annual manure rates were divided into three applications per year. In the first application, 10% was applied on February 15th. In the second application, 45% of the annual rate was applied just prior to spring planting, a date that is determined in Agro-IBIS based on climate. In the third application, the remaining 45% was applied on October 1st. Fertilizer was applied in a single application also just prior to planting. Plowing to a 10 cm depth also occurred at the time of planting, immediately following the springtime manure application. The winter and fall manure applications were unincorporated.

Appendix I: Algorithm for determining spatial application of manure and fertilizer rates

This appendix documents the logic used to determine manure and fertilizer applications of nitrogen and phosphorus to grid cells in the YW. Manure is estimated for each livestock operation based on animal units and livestock type. This manure is spread on available land within the estimated hauling distance radius. Overlapping of manure by different operations is accounted for and explained. Fertilizer application rates are based on University of Wisconsin-Extension recommendations.

Step-by-step description of MATLAB script

- The modeling grid is loaded with a grid cell size of 219.456 m by 219.456 m.
- The historical land-use, land cover (LULC) data are loaded.

The following categories of LULC are available for manure and/or fertilizer:

Code	LULC Category
11	Alfalfa
12	Corn
13	Fruits/Veg
14	Small Grains
15	Soybeans
16	Hay
17	Pasture

- The livestock operation characteristics dataset is loaded.
- Data on total number of animal units in YW and total milk production per year for the historical time period (1870-2014) is loaded. The baseline estimate for 2013 uses data from USDA-Agricultural Census and USDA-NASS surveys on Dane County livestock numbers and milk production (USDA-NASS 2013). The conversion from Dane County animal units to Yahara watershed animal units is based on the 2010 ratio, which was calculated to be 38%.
- Historical animal production weights – based on data from the USDA-Economic Research Service (USDA-ERS 2015) – are loaded.
- Minimum manure (wet) application rate is defined to be 23,000 kg ha^{-1}. This variable is based on estimates from surveys of south-central Wisconsin dairy farmers by (Powell and others 2005).
- The number of livestock operations for each year is loaded. The active livestock operations are determined for each year.
- The first nested *for* loop calculates the sum of all animal units (the 2010 estimate) at all active livestock operations in the watershed. This value is then used to calculate the animal unit ratio which is the ratio of the current year Yahara animal units to the 2010 Yahara animal units. This ratio later scales the animal units at each livestock operation to account for changes through time at the county scale.
- The next nested *for* loop is started and loops through each active livestock operation. First, manure (wet, dry, phosphorus, and nitrogen) production is estimated for each operation. Second, the manure is spread according to the estimated hauling distance and

the LULC available with some modifications based on the minimum manure application rate threshold and overlapping manure hauling areas (explained further below).

- The production of manure (wet, dry, phosphorus, nitrogen) from livestock (except lactating cows) is calculated based on the USDA-Agricultural Waste Management Field Handbook (USDA 2008). The production of lactating cow manure is estimated based on a regression model using milk production as a predictor from Nennich and others (2005). Note that manure is calculated on an as-excreted basis.
- The spatial distribution of manure is done over the following steps:
    - The grid cells that 1) fall within the circular area with radius equal to the estimated manure hauling distance; and 2) are cropland or pasture are identified ('haulCells')
    - If, for some reason, no cells are identified, then the closest cell that is cropland or pasture is identified (a 'flag' variable is used to identify whether this ever happens)
    - The available haul area is calculated
    - A series of if statements handle the different cases…
    - if no manure has been spread in the haul area yet, then the manure is spread everywhere evenly; but if the minimum manure application rate threshold is not reached, then the available haul area is iteratively "shrunk" from the outside-in until the threshold is reached
    - else, some manure has been spread somewhere in the haul area (overlapping case)
        - if at least half of the available haul area is still "dry", then spread manure evenly to just the dry cells (but the minimum manure application rate must be satisfied…if not then the dry cells are iteratively "shrunk" from outside-in until threshold is reached)
        - elseif there are no dry cells at all (all available haul cells have at least some manure applied already), first sort the cells by increasing manure N, then increasing distance. Manure is added proportionally more to the beginning of the list compared to the end so that the first gets twice as much as evenly spread and the last gets nothing.
        - else (if some dry cells exist but they sum to less than half of the available haul area), manure is applied to the dry cells at twice the rate as if all the haul cells were dry and the remaining amount is added to the wet cells (but first, the rate for the dry cells needs to be checked to see if it is above the minimum rate threshold…if not, then the available wet and dry areas are iteratively "shrunk" from outside-in).
    - the livestock operation loop is closed
- Next, the fertilizer N and P application rate is determined
    - The next nested *for* loop is started that loops through every grid cell
    - if the cell has manure applied, then fertilizer is only applied if it is corn
    - else (no manure has been applied), fertilizer is applied to all crops
    - Application baseline rates are taken from UW-Extension recommendations (A2809) and determined based on the crop type and an assumption of "optimum" soil conditions. These rates are then multiplied by the relative fertilizer application rate for each year.

Appendix J: Parameter Sensitivity Analysis

Agro-IBIS

To identify parameters that exert a strong influence on P yield, a simple, one-at-a-time parameter sensitivity analysis was conducted for Agro-IBIS. Thirty-three parameters used to calculate sediment, dissolved, and total P yield were examined (Table J1). Two single-cell simulations were conducted for each parameter, both using a grid-cell located in the northern region of the YW. This cell was chosen to represent soil texture and cropland management practices typical to the northern YW. Continuous corn was simulated from 1986-2013 using nominal annual fertilizer and manure applications of 12.9 and 28.9 kg ha^{-1}, respectively. Historical climate data were used and silt loam was the simulated soil texture. The two simulations were conducted by varying the nominal value of each parameter by +1 and -1%, for a total of 66 simulations. Changes in the response variables (sediment, dissolved, and total P yield) greater than 1% could indicate over-sensitivity to a particular parameter with the influence of such a parameter becoming amplified within the model. Changes in the response variables were analyzed for each simulation, and those parameters inducing a response greater than 1% were identified.

The results showed that no parameters induced a change in total P yield greater than 1% (Table J1), implying that none of the parameters exerted a disproportionate influence on modeled P losses. This was also the case for dissolved P yield, for which the amount of manure applied (MANPAPP) exerted the largest influence (0.70%). For sediment P yield, two parameters induced a change slightly greater than 1%. These were the puddle normalization constant (1.03%) and the MUSLE power (1.18%). The puddle normalization constant represents a maximum puddle depth for which standing water is allowed to collect. Water collecting in excess of the puddle depth is treated as runoff; therefore the puddle depth is an important control on the amount of runoff generated. The influence of runoff on the transport of P from soil, in both dissolved and sediment forms, explains why this parameter had a relatively large impact on total P yield (0.90%), dissolved P yield (0.64%), and sediment P yield (1.03%). Within the MUSLE, daily runoff as well the peak runoff rate are used to calculated sediment yield (Eq. 1). The original value of the MUSLE power given by Williams (1975) was 0.56. For this study, calibration of this location-specific parameter yielded a value of 0.35, a value that helped generate both reasonable amounts of sediment yield at the field scale and sediment load within streams (Appendix D). The sensitivity analysis showed that this parameter exerted an important influence on sediment P yield (1.18%) as well as total P yield (0.79%). As expected, it had no influence on dissolved P yield (0.01%). The other MUSLE parameters, including the coefficient (originally given as 11.8 by Williams (1975) for metric units, and calibrated here to 5.9), and the *C*, *K*, and *LS* parameters, also exerted a relatively strong influence on sediment P yield (0.97%) and total P yield (0.65%). Since these parameters all act as simple multipliers in the MUSLE (Eq. 1), their influence was identical in this experiment.

Table J1. Parameter sensitivity results for Agro-IBIS showing the maximum absolute percent change induced in P yield for a 1% change in each parameter.

Agro-IBIS Parameter	Description	Nominal Value	Total P Yield (%)	Sediment P Yield (%)	Dissolved P Yield (%)
INITLAB	Initial labile P concentration, top layer	40 ppm	0.19	0.24	0.09
CLAY	Clay percentage in soil	20%	0.01	0.01	0.00
SOM	Organic matter percentage in soil	7%	0.00	0.00	0.00
BD	Soil bulk density	1.5 g cm^{-3}	0.60	0.75	0.30
FERTPKG	Annual fertilizer application.	12.9 kg ha^{-1}	0.17	0.18	0.14
MANPAPP	Annual manure application.	28.9 kg ha^{-1}	0.47	0.35	0.71
MWIPFRAC	Manure inorganic water extractable P fraction	0.50	0.10	0.05	0.19
MWOPFRAC	Manure organic water extractable P fraction	0.05	0.01	0.00	0.02
DRYMAT	Dry matter content of manure applied.	9%	0.00	0.00	0.00
PCTCOV	Percent of field covered by manure application	99%	0.01	0.01	0.04
EFFINC	Surface incorporation efficiency	30%	0.09	0.02	0.25
FMIX	Soil mixing efficiency	30%	0.38	0.48	0.18
DEPTIL	Depth of tillage	10 cm	0.00	0.00	0.00
betares	Fresh P mineralization coefficient	0.06	0.00	0.00	0.00
betamin	Humic P mineralization coefficient	0.002	0.00	0.00	0.00
cnrres	C:N ratio of residue	25	0.00	0.00	0.00
fracnsta	Fraction of soil stable N.	1%	0.00	0.00	0.00
fracnact	Fraction of soil active N.	99%	0.01	0.01	0.00
pctpharv	Fraction of plant P returned to organic fresh pool at harvest.	10%	0.03	0.04	0.02
uptake	Plant uptake rate coefficient.	0.004	0.34	0.42	0.17
INFIL	Slurry P infiltration fraction	0.40	0.07	0.14	0.50
PER	P enrichment ratio	1.0	0.66	0.99	0.00
K	MUSLE K factor	0.30	0.65	0.97	0.00
LS	MUSLE LS factor	0.35	0.65	0.97	0.00
C	MUSLE C factor	0.80	0.65	0.97	0.00
manning	Mannings roughness coefficient	0.05	0.00	0.00	0.00
hydcond	Saturated hydraulic conductivity	1.89e-6 m s^{-1}	0.68	0.73	0.71
puddle	Puddle normalization constant	30 mm	0.90	1.03	0.64
zmin	Runoff threshold	0.03	0.74	0.75	0.70
zmax	Runoff threshold	0.30	0.44	0.65	0.08
PSP	P sorption parameter	0.17	0.13	0.35	0.32
POW	MUSLE power	0.35	0.79	1.18	0.01
COEFF	MUSLE coefficient	5.9	0.65	0.97	0.00

THMB

The sensitivity analysis conducted for THMB used streamflow, sediment load, and P load as the output variables. For the THMB sensitivity simulations, the entire watershed was simulated over the historical time period using climate and land use inputs identical to those used in ME. Seven parameters that have major influence on streamflow, sediment, and P processes were examined. For streamflow, parameters included u_{o1}, the minimum effective river flow velocity (u_{o1}=0.35 m s^{-1}); i_o, the reference gradient (i_o=0.0001 m m^{-1}); and p_o, the reference wetted perimeter (p_o=25 m). For sediment load, parameters included e_0, the reference erosion rate (e_0=1.0x10^{-5} kg $m^{-2}s^{-1}$) and n_p, the sediment porosity (n_p=0.1). For P load, parameters included b, a phosphorus enrichment optimization parameter (b=0.27), and k_E, a sediment P erosion coefficient (k_E=0.0005). Each parameter was varied by ± 1% and the percentage change in each output variable was calculated. For streamflow and sediment load, the changes were calculated based on mean annual watershed values. For P load, the changes were based on mean annual P loads to Lake Mendota. The results are shown in Table J2.

Table J2. Parameter sensitivity results for THMB showing the percent change induced in streamflow, sediment load, and direct drainage P load for a 1% change in each parameter.

THMB Parameter	Description	Nominal Value	Streamflow (%)	Sediment Load (%)	Phosphorus Load (%)
u_{o1}	velocity	0.35 m s^{-1}	0.0012	0.22	-0.11
i_o	Reference gradient	1e-4 m m^{-1}	-0.0002	0.25	-0.11
p_o	Reference wetted perimeter	25 m	-0.0002	0.08	-0.13
e_0	Erosion rate	2e-4 kg $m^{-2}s^{-1}$	0	0.25	0.01
n_p	Sediment porosity	0.5	0	2.72	0.27
b	Phosphorus enrichment optimization parameter	0.27	0	0	-0.08
k_E	Sediment P erosion coefficient	2e-4	0	0	0.07

All parameters showed a reasonable amount of influence on their related variables, with most of them lower than 1%. The influence of streamflow parameters on sediment load and P load, and the influence of sediment load parameters on streamflow and P load, were included only as a matter of interest. Among the three output variables examined, changes in sediment and P loads were larger than changes in streamflow, especially when the parameters related to streamflow were changed. Sediment loads in the stream network have power law relationships with streamflow. Additionally, sediment P transport processes follow sediment transport, in terms of erosion and deposition. Therefore, changes in streamflow will be magnified in sediment and P loads. Parameter adjustments did not always cause same-sign changes among the output variables. This highlighted the nonlinearity of the relationships between the parameters and the variables.

References

Alexander RB, Smith RA. 1990. County-level estimates of nitrogen and phosphorus fertilizer use in the United States, 1945 to 1985. Open-file report 90-130. Reston, VA: U.S. Geological Survey http://pubs.usgs.gov/of/1990/ofr90130/

Amy G, Pitt R, Rameshawar-Singh, Bradford WL, LaGraff MB. 1974. Water quality management planning for urban runoff. Washington, D.C.: U.S. Environmental Protection Agency

Anderson G. 1980. Assessing organic phosphorus in soils. In: The Role of Phosphorus in Agriculture. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America. pp 411–31.

Andraski TW, Bundy LG. 2003. Relationships between phosphorus levels in soil and in runoff from corn production systems. J Environ Qual 32:310–6.

Bennett EM, Carpenter SR, Clayton MK. 2004. Soil phosphorus variability: scale-dependence in an urbanizing agricultural landscape. Landsc Ecol 20:389–400.

Campbell GS, Norman JM. 1998. An Introduction to Environmental Biophysics. Second Edition. Springer Science & Business Media

Carpenter SR, Lathrop RC. 2008. Probabilistic estimate of a threshold for eutrophication. Ecosystems 11:601–13.

Carpenter SR, Lathrop RC. 2014. Phosphorus loading, transport and concentrations in a lake chain: a probabilistic model to compare management options. Aquat Sci 76:145–54.

Carsel RF, Parrish RS. 1988. Developing joint probability distributions of soil water retention characteristics. Water Resour Res 24:755–69.

Coe MT. 1998. A linked global model of terrestrial hydrologic processes: Simulation of modern rivers, lakes, and wetlands. J Geophys Res 103:8885–99.

Coe MT. 2000. Modeling Terrestrial Hydrological Systems at the Continental Scale: Testing the Accuracy of an Atmospheric GCM. J Clim 13:686–704.

Coe MT, Costa MH, Howard EA. 2008. Simulating the surface waters of the Amazon River basin: impacts of new river geomorphic and flow parameterizations. Hydrol Process 22:2542–53.

Collick AS, Veith TL, Fuka DR, Kleinman PJA, Buda AR, Weld JL, Bryant RB, Vadas PA, White MJ, Harmel RD, Easton ZM. 2016. Improved Simulation of Edaphic and Manure Phosphorus Loss in SWAT. J Environ Qual. https://dl.sciencesocieties.org/publications/jeq/abstracts/0/0/jeq2015.03.0135

Dalal RC. 1977. Soil organic phosphorus. Adv Agron 29:83–117.

Donner SD, Coe MT, Lenters JD, Twine TE, Foley JA. 2002. Modeling the impact of hydrological changes on nitrate transport in the Mississippi River Basin from 1955 to 1994. Global Biogeochem Cycles 16:1–19.

Donner SD, Kucharik CJ. 2003. Evaluating the impacts of land management and climate variability on crop production and nitrate export across the Upper Mississippi Basin. Global Biogeochem Cycles 17:1–16.

Donner SD, Kucharik CJ. 2008. Corn-based ethanol production compromises goal of reducing nitrogen export by the Mississippi River. Proc Natl Acad Sci U S A 105:4513–8.

Donner SD, Kucharik CJ, Foley JA. 2004. Impact of changing land use practices on nitrate export by the Mississippi River. Global Biogeochem Cycles 18:1–21.

Dunne T, Black RD. 1970a. An experimental investigation of runoff production in permeable soils. Water Resour Res 6:478–90.

Dunne T, Black RD. 1970b. Partial area contributions to storm runoff in a small New England watershed. Water Resour Res 6:1296–311.

El Maayar M, Price DT, Delire C, Foley JA, Black TA, Bessemoulin P. 2001. Validation of the Integrated Biosphere Simulator over Canadian deciduous and coniferous boreal forest stands. J Geophys Res 106:14339.

Filstrup CT, Wagner T, Soranno PA, Stanley EH, Stow CA, Webster KE, Downing JA. 2014. Regional variability among nonlinear chlorophyll—phosphorus relationships in lakes. Limnol Oceanogr 59:1691–703.

Foley JA, Prentice IC, Ramankutty N, Levis S, Pollard D, Sitch S, Haxeltine A. 1996. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. Global Biogeochem Cycles 10:603–28.

Gillon S, Booth EG, Rissman AR. 2016. Shifting drivers and static baselines in environmental governance: challenges for improving and proving water quality outcomes. Regional Environ Change 16:759–75.

Good LW, Vadas P, Panuska JC, Bonilla CA, Jokela WE. 2012. Testing the Wisconsin Phosphorus Index with year-round, field-scale runoff monitoring. J Environ Qual 41:1730–40.

Gronberg JAM, Spahr NE. 2012. County-level estimates of nitrogen and phosphorus from commercial fertilizer for the conterminous United States, 1987–2006. Scientific Investigations Report 2012-5207. Reston, VA: U.S. Geological Survey http://pubs.usgs.gov/sir/2012/5207/

Hibbard BH. 1904. The history of agriculture in Dane County, Wisconsin. Madison, WI: University of Wisconsin http://digital.library.wisc.edu/1711.dl/WI.HistAgDane

Hilborn R, Mangel M. 1997. The ecological detective: confronting models with data. Princeton, N.J.: Princeton University Press

Hobbie SE. 1992. Effects of plant species on nutrient cycling. Trends Ecol Evol 7:336–9.

Homer CG, Dewitz JA, Yang L, Jin S, Danielson P, Xian G, Coulston J, Herold ND, Wickham JD, Megown K. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. Photogramm Eng Remote Sens 81:345–54.

Horrocks CA, Dungait JAJ, Cardenas LM, Heal KV. 2014. Does extensification lead to enhanced provision of ecosystems services from soils in UK agriculture? Land Use Policy 38:123–8.

Jones CA, Cole CV, Sharpley AN, R WJ. 1984. A simplified soil and plant phosphorus model: I. Documentation. Soil Sci Soc Am J 48:1–3.

Kucharik CJ. 2003. Evaluation of a Process-Based Agro-Ecosystem Model (Agro-IBIS) across the U.S. Corn Belt: Simulations of the Interannual Variability in Maize Yield. Earth Interact 7:1–33.

Kucharik CJ, Barford CC, Maayar ME, Wofsy SC, Monson RK, Baldocchi DD. 2006. A multiyear evaluation of a Dynamic Global Vegetation Model at three AmeriFlux forest sites: Vegetation structure, phenology, soil temperature, and CO2 and H2O vapor exchange. Ecol Modell 196:1–31.

Kucharik CJ, Brye KR. 2003. Integrated BIosphere Simulator (IBIS) yield and nitrate loss predictions for Wisconsin maize receiving varied amounts of nitrogen fertilizer. J Environ Qual 32:247–68.

Kucharik CJ, Foley JA, Delire C, Fisher VA, Coe MT, Lenters JD, Young-Molling C, Ramankutty N, Norman JM, Gower ST. 2000. Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure. Global Biogeochem Cycles 14:795–825.

Kucharik CJ, Twine TE. 2007. Residue, respiration, and residuals: Evaluation of a dynamic agroecosystem model using eddy flux measurements and biometric data. Agric For Meteorol 146:134–58.

Laboski CAM, Peters JB. 2012. Nutrient application guidelines for field, vegetable, and fruit crops in Wisconsin. University of Wisconsin Extension

Laflen JM, Flanagan DC, Engel BA. 2004. Soil erosion and sediment yield prediction accuracy using WEPP. JAWRA 40:289–97.

Lathrop RC. 1979. Dane County Water Quality Plan. Dane County Regional Planning Commission

Lathrop RC, Carpenter SR. 2013. Water quality implications from three decades of phosphorus loads and trophic dynamics in the Yahara chain of lakes. Inland Waters 4:1–14.

Lathrop RC, Carpenter SR, Stow CA, Soranno PA, Panuska JC. 1998. Phosphorus loading reductions needed to control blue-green algal blooms in Lake Mendota. Can J Fish Aquat Sci 55:1169–78.

McElroy AD, Chiu SY, Nebgen JW, Aleti A, Bennett FW. 1976. Loading functions for assessment of water pollution from nonpoint sources. Environ Prot Tech Serv, EPA 600/2-76-151.

NCDC. 2015. Global Historical Climate Network. National Climatic Data Center, National Oceanic and Atmospheric Administration http://www.ncdc.noaa.gov/data-access

NCEP. 2015. National Stage IV Quantitative Precipitation Estimate Mosaic, National Centers for Environmental Prediction. Center for Data Analytics, Office of Water Information, U.S. Geological Survey http://cida.usgs.gov/thredds/catalog.html?dataset=cida.usgs.gov/thredds/ncep_stageiv

Nearing MA, Govers G, Norton LD. 1999. Variability in soil erosion data from replicated plots. Soil Sci Soc Am J 63:1829–35.

Neitsch SL, Arnold JG, Kiniry JR, Williams Grassland JR, Labora WR. 2011. Soil and Water Assessment Tool Theoretical Documentation Version 2009.

Nennich TD, Harrison JH, VanWieringen LM, Meyer D, Heinrichs AJ, Weiss WP, St-Pierre NR, Kincaid RL, Davidson DL, Block E. 2005. Prediction of manure and nutrient excretion from dairy cattle. J Dairy Sci 88:3721–33.

NRCS. 2004. National Engineering Handbook - Hydrology Chapters: Chapter 10 Estimation of Direct Runoff from Storm Rainfall. USDA-NRCS

Oloya TO, Logan TJ. 1980. Phosphate desorption from soils and sediments with varying levels of extractable phosphate. J Environ Qual 9:526–31.

Ovington JD, Madgwick HAI. 1958. The sodium, potassium and phosphorus contents of tree species grown in close stands. New Phytol 57:273–84.

Panagos P, Borrelli P, Meusburger K, Alewell C, Lugato E, Montanarella L. 2015. Estimating the soil erosion cover-management factor at the European scale. Land Use Policy 48:38–50.

Parsen, MJ, Bradbury, KR, Hunt, RJ, and Feinstein, DT, 2016. The 2016 groundwater flow model for Dane County, Wisconsin: Wisconsin Geological and Natural History Survey Bulletin 110, 56 p.

Patil S, Sivapalan M, Hassan MA, Ye S, Harman CJ, Xu X. 2012. A network model for prediction and diagnosis of sediment dynamics at the watershed scale. J Geophys Res 117:F00A04.

Powell JM, McCrory DF, Jackson-Smith DB, Saam H. 2005. Manure Collection and Distribution on Wisconsin Dairy Farms. J Environ Qual 34:2036.

Powell JM, Rotz CA. 2015. Measures of nitrogen use efficiency and nitrogen loss from dairy production systems. J Environ Qual 44:336–44.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Richards LA. 1931. Capillary conduction of liquids through porous mediums. J Appl Phys 1:318–33.

Rotz CA, Montes F, Hafner SD, Heber AJ, Grant RH. 2014. Ammonia emission model for whole farm evaluation of dairy production systems. J Environ Qual 43:1143–58.

Ryser P, Verduyn B, Lambers H. 1997. Phosphorus allocation and utilization in three grass species with contrasting response to N and P supply. New Phytol 137:293–302.

Sadeghi SHR, Gholami L, Darvishan AK, Saeidi P. 2014. A review of the application of the MUSLE model worldwide. Hydrol Sci J 59:365–75.

Schneider A, Logan KE, Kucharik CJ. 2012. Impacts of Urbanization on Ecosystem Goods and Services in the U.S. Corn Belt. Ecosystems 15:519–41.

Sen S, Srivastava P, Vadas PA, Kalin L. 2012. Watershed-level comparison of predictability and sensitivity of two phosphorus models. J Environ Qual 41:1642–52.

Sharpley AN. 1985. The selection erosion of plant nutrients in runoff. Soil Sci Soc Am J 49:1527–34.

Sharpley AN. 2003. Soil mixing to decrease surface stratification of phosphorus in manured soils. J Environ Qual 32:1375–84.

Sharpley AN, Chapra SC, Wedepohl R, Sims JT, Daniel TC, Reddy KR. 1994. Managing agricultural phosphorus for protection of surface waters - issues and options. J Environ Qual 23:437–51.

Smith DR, Warnemuende-Pappas EA. 2015. Vertical tillage impacts on water quality derived from rainfall simulations. Soil Tillage Res 153:155–60.

Soylu ME, Kucharik CJ, Loheide, II, Steven P. 2014. Influence of groundwater on plant water use and productivity: Development of an integrated ecosystem – Variably saturated soil water flow model. Agric For Meteorol 189-190:198–210.

Stuntebeck T, Komiskey M, Peppler M, Owens D, Frame D. 2011. Precipitation-runoff relations and water-quality characteristics at edge-of-field stations, Discovery Farms and Pioneer Farm, Wisconsin, 2003–08.

USDA. 2008. Agricultural Waste Characteristics. In: Agricultural Waste Management Field Handbook, 210-VI-AWMFH, March 2008. Washington D.C.: Natural Resources Conservation Service, U.S. Department of Agriculture

USDA. 2013. Soil Survey Geographic (SSURGO) Database for Columbia, Dane, and Rock counties, Wisconsin. http://websoilsurvey.nrcs.usda.gov/. Last accessed 22/09/2011

USDA. 2014. United States Department of Agriculture, National Agricultural Statistics Service Cropland Data Layers. http://nassgeodata.gmu.edu/CropScape/

USDA-ERS. 2015. Livestock and poultry live and dressed weights. Economic Research Service, U.S. Department of Agriculture http://www.ers.usda.gov/data-products/livestock-meat-domestic-data.aspx#26070

USDA-NASS. 2013. USDA Census of Agriculture Historical Archive. Albert R. Mann Library at Cornell University and the National Agricultural Statistics Service, U.S. Department of Agriculture http://agcensus.mannlib.cornell.edu/

USDA-NASS. 2014. Census of Agriculture. National Agricultural Statistics Service, United States Department of Agriculture http://www.agcensus.usda.gov/

UVGSDC. 2004. Historical Census Browser. University of Virginia, Geospatial and Statistical Data Center http://mapserver.lib.virginia.edu/

UW-CALS. 2005. The Wisconsin Buffer Initiative. A report to the Natural Resources Board of the Wisconsin Department of Natural Resources by the University of Wisconsin-Madison, College of Agricultural and Life Sciences.

UW-Extension. 2014. Automated Weather Observation Network Data. University of Wisconsin-Extension Agricultural Weather http://agwx.soils.wisc.edu/uwex_agwx/awon

Vadas PA, Gburek WJ, Sharpley AN, Kleinman PJA, Moore PA Jr, Cabrera ML, Harmel RD. 2007a. A model for phosphorus transformation and runoff loss for surface-applied manures. J Environ Qual 36:324–32.

Vadas PA, Haggard BE, Gburek WJ. 2005. Predicting dissolved phosphorus in runoff from manured field plots. J Environ Qual 34:1347–53.

Vadas PA, Harmel RD, Kleinman PJA. 2007b. Transformations of soil and manure phosphorus after surface application of manure to field plots. Nutr Cycling Agroecosyst 77:83–99.

Vadas PA, Joern BC, Moore PA. 2012. Simulating soil phosphorus dynamics for a phosphorus loss quantification tool. J Environ Qual 41:1750–7.

Vadas PA, Kleinman PJA, Sharpley AN. 2004. A simple method to predict dissolved phosphorus in runoff from surface-applied manures. J Environ Qual 33:749–56.

Vadas PA, Krogstad T, Sharpley AN. 2006. Modeling Phosphorus Transfer between Labile and Nonlabile Soil Pools. Soil Sci Soc Am J 70:736.

Vadas PA, Mark Powell J, Brink GE, Busch DL, Good LW. 2015. Whole-farm phosphorus loss from grazing-based dairy farms. Agric Syst 140:40–7.

Vadas PA, White MJ. 2010. Validating Soil Phosphorus Routines in the SWAT Model. Transactions of the ASABE 53:1469–76.

Van Genuchten MT. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci Soc Am J 44:892–8.

Viney NR, Sivapalan M, Deeley D. 2000. A conceptual model of nutrient mobilisation and transport applicable at large catchment scales. J Hydrol 240:23–44.

Walker TW, Syers JK. 1976. Fate of phosphorus during pedogenesis. Geoderma 15:1–19.

WDATCP. 2015. DATCP Public FTP Site. Wisconsin Department of Agriculture, Trade, and Consumer Protection ftp://ftp.datcp.state.wi.us/GIS/DATCP_FOOD_SAFETY.gdb.zip

WDNR. 2011. Original Vegetation Cover of Wisconsin. Madison, WI: Wisconsin Department of Natural Resources ftp://dnrftp01.wi.gov/geodata/orig_veg_cover/orig_veg_cover.zip

WDNR. 2013. Wisconsin Land Cover database - Version 1. Wisconsin Department of Natural Resources ftp://gomapout.dnr.state.wi.us/landcover/

WDNR. 2015. Searchable database of CAFO WPDES permittees. Wisconsin Department of Natural Resources http://dnr.wi.gov/topic/AgBusiness/data/CAFO/

Williams J. 1995. The EPIC Model. In: Singh VP, editor. Computer Models of Watershed Hydrology. Littleton, CO: Water Resources Publications. pp 909–1000.

Williams JR. 1975. Sediment-yield prediction with Universal Equation using runoff energy factor. In: Present and Prospective Technology for Predicting Sediment Yield and Sources. Vol. ARS-S-40. U.S. Dept. Agric. pp 244–52.

Williams JR, Hann RW. 1978. Optimal operation of large agricultural watersheds with water quality restraints. Texas Water Resources Institute, Texas A&M Univ, Tech Rept No 96.

Wischmeier WH, Smith DD. 1978. Predicting rainfall erosion losses. A guide to conservation planning. U.S. Department of Agriculture, Washington, D.C.

Young RA, Onstad CA, Bosch DD, Anderson WP. 1989. AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds. J Soil Water Conserv 44:168–73.

Zipper SC, Soylu ME, Booth EG, Loheide, II, Steven P. 2015. Untangling the effects of shallow groundwater and soil texture as drivers of subfield-scale yield variability. Water Resour Res 51:6338–58.