WebTable 1. Agricultural management scenarios were run in five separately configured SWAT models; potential differences among models are described here.

Aspect of modeling in order of development	Potential differences among models in this study	Further details
Spatial discretization & resolution *in initial model setup through ArcGIS interface*	• Size of sub-watersheds as dictated by stream threshold • Definition of HRU slope classes • Lumping of HRUs	Initial model setup is determined based on the goals of the project, and once completed is difficult to change. These model differences were retained in the Baseline models. This source of uncertainty is referred to as structural uncertainty.
Model/Sub-model algorithms *chosen within SWAT*	• Model release version and source code updates • Tile drainage routine • In-stream processing • Evaporation method • Water table method • Runoff method • Carbon model • Soil phosphorus model	SWAT is a compilation of multiple sub-models, and the user can choose which sub-models to use. The algorithms used in the model introduce structural uncertainty.
Model inputs *including data sources, spatial resolution, and preprocessing*	• Land use data: NLCD versus NASS CDL • Point source data: None included versus included based on emissions caps versus based on measured data • Weather data	Model inputs are also chosen early in the modeling process. In this study we chose to control for some of these input differences by homogenizing point sources and climate forcing across Baseline models. These choices introduce input and measurement uncertainty.
Land management operations *include a host of assumptions based on disparate sources*	• Spatial distribution/ heterogeneity of operations • Timing of operations • Crop rotations • Fertilizer applications • Manure applications • Inclusion of existing conservation practices in the watershed	Assumptions made about cropland management operations are critically important for realistically simulating current agricultural practices in the watershed, many of which are difficult to determine using publicly available datasets. Cropland management differences were retained in the Baseline models. This is a form of input uncertainty, and addressing this was a primary goal of the study.
Model parameterization *in choosing realistic parameter values to calibrate a model*	• Parameters changed in calibration • Bounds on parameter values • Methods for assessing model performance during calibration	Modelers all changed different sets of parameters to calibrate their models, and the final parameter values span a wide range. Multiple parameter sets can achieve a calibrated model, which leads to parameter uncertainty. These differences were retained in the Baseline models.
Measured data *for calibration*	• Extent of water quality calibration • Extent of hydrology considered at upstream monitoring stations • Method to fill in or ignore missing data	Measured data provide a reality check against which we assess how well our models perform. It is easy to forget that measured data are only a snapshot of true events and there can be considerable uncertainty in them.

Notes: While all models were developed using the same base SWAT framework, they are distinct from one another in many ways, from initial model setup and model subroutines activated to assumptions about farmland management and model calibration. The main themes in model differences are shown below with examples of differences among models, explanations for the type of uncertainty derived from these differences, and a description of how models were homogenized in Baseline models for this effort. A model-specific list of differences among SWAT models can be found in WebTables 2 and 3. CDL = Cropland Data Layer; HRU = hydrologic response unit; NASS = National Agricultural Statistics Service; NLCD = National Land Cover Dataset; SWAT = Soil and Water Assessment Tool.