University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

USGS Staff -- Published Research | US Geological Survey

2014

A Bayesian Network Model for Assessing Natural Estrogen Fate and Transport in a Swine Waste Lagoon

Boknam Lee
Duke University, leeboknam@gmail.com

Seth W. Kullman
North Carolina State University, swkullma@ncsu.edu

Erin Yost
North Carolina State University, eeyost@gmail.com

Michael T. Meyer
U.S. Geological Survey, mmeyer@usgs.gov

Lynn Worley-Davis
North Carolina State University, lwdavis@ncsu.edu

See next page for additional authors

Follow this and additional works at: http://digitalcommons.unl.edu/usgsstaffpub

Lee, Boknam; Kullman, Seth W.; Yost, Erin; Meyer, Michael T.; Worley-Davis, Lynn; Williams, C. Michael; and Reckhow, Kenneth H., "A Bayesian Network Model for Assessing Natural Estrogen Fate and Transport in a Swine Waste Lagoon" (2014). *USGS Staff -- Published Research*. Paper 833.
http://digitalcommons.unl.edu/usgsstaffpub/833

This Article is brought to you for free and open access by the US Geological Survey at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in USGS Staff -- Published Research by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors
Boknam Lee, Seth W. Kullman, Erin Yost, Michael T. Meyer, Lynn Worley-Davis, C. Michael Williams, and Kenneth H. Reckhow

This article is available at DigitalCommons@University of Nebraska - Lincoln: http://digitalcommons.unl.edu/usgsstaffpub/833

Original Article Integrated Environmental Assessment and Management

DOI 10.1002/ieam.1538

A Bayesian Network Model for Assessing Natural Estrogen Fate and Transport in a Swine Waste Lagoon**

Boknam Lee,[†] Seth W. Kullman,[*‡] Erin Yost,[‡] Michael T. Meyer,[§] Lynn Worley-Davis,[‡] C. Michael Williams,[‡] Kenneth H. Reckhow[†]

[†] Ph.D. Candidate, Nicholas School of the Environment, Duke University, Durham, North Carolina, 27708, E-mail: leeboknam@gmail.com
[‡] Associate Professor, Program in Environmental and Molecular Toxicology, Department of Biological Sciences, North Carolina State University, Raleigh, North Carolina, 27695, E-mail: swkullma@ncsu.edu
[‡] Ph.D. Candidate, Department of Environmental and Molecular Toxicology, North Carolina State University, Raleigh, North Carolina 27695, E-mail: eeyost@gmail.com
[§] Organic Geochemist, Organic Geochemistry Research Laboratory, U.S. Geological Survey (USGS), Lawrence, Kansas, 66049, E-mail: mmeyer@usgs.gov
[‡] Lecturer, Prestage Department of Poultry Science, North Carolina State University, Raleigh, North Carolina, 27695, E-mail: lwdavis@ncsu.edu
Professor; Department Head; Director, Prestage Department of Poultry Science, North Carolina State University, Raleigh, North Carolina, 27695, E-mail: mike_williams@ncsu.edu
[†] Professor Emeritus, Nicholas School of the Environment, Duke University, Durham, North Carolina, 27708, E-mail: reckhow@duke.edu

[*] (Corresponding author: Seth W. Kullman, PhD, Associate Professor, Program in Environmental & Molecular Toxicology, Department of Biological Sciences, Campus Box 7633. North Carolina State University, Raleigh, NC 27695-7633, Phone: (919) 515-4378
Fax: (919) 515-7169)

**This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/ieam.1538]

All Supplemental Data may be found in the online version of this article at the publisher's website.

© 2014 SETAC This article is a U.S. government work, and is not subject to copyright in the United States.

Submitted 5 September 2013; Returned for Revision 31 March 2014; Accepted 31 March 2014

Accepted Manuscript

ABSTRACT

Commercial swine waste lagoons are regarded as a major reservoir of natural estrogens, which have the potential to produce adverse physiological effects on exposed aquatic organisms and wildlife. However, there remains limited understanding of the complex mechanisms of physical, chemical, and biological processes that govern the fate and transport of natural estrogens within an anaerobic swine lagoon. To improve lagoon management and ultimately help control the offsite transport of these compounds from swine operations, a probabilistic Bayesian network model was developed to assess natural estrogen fate and budget and then compared against data collected from a commercial swine field site. In general, the model was able to describe the estrogen fate and budget in both the slurry and sludge stores within the swine lagoon. Sensitivity analysis within the model, demonstrated that the estrogen input loading from the associated barn facility was the most important factor in controlling estrogen concentrations within the lagoon slurry storage, while the settling rate was the most significant factor in the lagoon sludge storage. The degradation reactions were shown to be minor in both stores based on prediction of average total estrogen concentrations. Management scenario evaluations demonstrated that the best possible management options to reduce estrogen levels in the lagoon are either to adjust the estrogen input loading from swine barn facilities or to effectively enhancing estrogen bonding with suspended solids through the use of organic polymers or inorganic coagulants.

Keywords: Natural estrogen hormones, Bayesian network model, Swine anaerobic lagoon, Estrogen fate and transport, Model evaluation

Accepted Preprint

Accepted Preprint

INTRODUCTION

The number of concentrated animal feeding operations (CAFOs) has increased significantly in the United States in the last two decades. For instance, North Carolina has become the second largest swine production state in the US, containing approximately 2,400 major swine facilities that generate an estimated 19 million tons of waste per year (NCDWQ 2009; NCDA & CS 2011). In North Carolina, the predominant system of waste management on commercial swine operations involves the storage of swine manure in anaerobic lagoons. These lagoons provide a means of storage and primary treatment of the swine waste through anaerobic digestion before the waste slurry is applied onto croplands as a nutrient management practice (Cheng 2003).

This waste management strategy has raised environmental concerns due to the potential for manure-borne contaminants, including natural steroidal estrogen hormones, to be transported in runoff from these fields and into nearby surface waters (Bradford et al. 2008). Swine wastes contain appreciable amount of natural estrogen hormones, including 17β-estradiol (E2β), 17α-estradiol (E2α), estrone (E1), and estriol (E3), comparatively, artificial or supplemental hormones are not used in rearing swine other than gonadotropins for synchronization of estrus in (Yost et al. 2013). Natural estrogens are well known to adversely affect the physiology of aquatic organisms, even at exceedingly low concentrations, and thus the high levels of these compounds found in land-applied lagoon slurry present a potential ecotoxicological risk (Leet et al. 2011). Therefore, a systematic approach to lagoon management is essential to identify and assess the relative importance of the major mechanisms to reduce estrogen compounds in swine waste lagoons during anaerobic treatment, before the lagoon wastes are applied onto crop fields.

During the storage of swine waste in an anaerobic lagoon, estrogens undergo a series of complex physical and chemical processes. These include partitioning between solid and liquid phases, chemical and microbial transformation, settling, and burial. These processes occur between two different storage compartments within the lagoon: the slurry storage, which is predominantly liquid, and the sludge storage, which is predominantly solid. In order to describe and quantify all of the processes that govern the stability and fate of estrogen compounds in the lagoon, a mass balance integrated Bayesian network model was developed as a decision support system in this study. This hybrid approach integrates both mechanistic and probabilistic elements and can be used to better

understand and assess the fate and budget of estrogen compounds within the swine lagoon treatment stores. This approach utilizes mass balance equations to account for the mechanisms of flushing, sorption, transformation, settling, and burial of estrogen compounds in both the anaerobic lagoon slurry and sludge treatment stores. These equations are integrated within the fundamental structure of a Bayesian network (BN), and the BN was then used to account for the uncertain input variables and to propagate this uncertainty through the model. Output variables are thus represented by probability distributions rather than by point estimates. The details of model development and a case study for a prototypic swine farrowing lagoon system is described herein, including model evaluation, sensitivity analysis, and what-if scenario simulations to assess lagoon management actions that may reduce the transport of estrogen compounds into the environment.

DESCRIPTION OF STUDY AREA AND ESTROGEN VARIABILITY ANALYSIS

Study Area

The study site is a commercial swine farrowing facility in southeastern North Carolina, centrally located in the swine production region of the state. This CAFO consists of two farms, Farm 1 and Farm 2, and houses exclusively female swine (approximately 2500 sows on each farm), all of which are breeding, gestating, or farrowing. In this study, the investigation focused exclusively on Farm 2 (Supplemental Data, Figure S1), where swine wastes are flushed from the barns and stored in a rectangular anaerobic treatment lagoon, which is located adjacent to the barns. The lagoon is designed for 10 million gallons of waste, including accumulated slurry and sludge wastes from the 2500 sows as well as for temporary storage for rainfall and wastewater inputs. This wastewater storage capacity is designed to store and treat manure for 180 days, for approximately 44,000 gals/d of liquid waste accumulation and 23 ft^3/d of sludge accumulation. The lagoon depth is 10 feet with a 1-foot freeboard, and it was constructed with 3% side slopes. The rectangular length and width of the lagoon are 547 feet by 309 feet. Lagoon slurry is applied to adjoining spray fields in accordance with a producer-approved waste management plan during the growing season in order to keep the lagoon at or below the maximum wastewater storage level.

Accepted Preprint

Estrogen Variability Analysis in Lagoon

Extensive sampling of the lagoon was conducted on June 15, 2009 and April 14, 2010. Eight coordinates on the lagoon were chosen in order to create a representative cross-section of the site: 3 locations were near the outflow pipes from the barns, 4 were in the middle of the lagoon, and 1 was at the far end of the lagoon. At each coordinate, 1-liter samples were collected at 3 different depths of slurry (6 inches below the surface, 2 feet below the surface, and 6 inches above the level of the sludge) using a horizontal beta water sampler (Wildco). Sludge samples were also collected from a level of 1 foot below the sludge/slurry interface, using a specialized sludge sampler that was constructed by the North Carolina State University (NCSU) Dept. of Biological and Agricultural Engineering. In sum, a total of 48 slurry samples and 16 sludge samples were collected during sampling trips.

All lagoon slurry and sludge samples were analyzed for 17β-estradiol (E2), estrone (E1), and estriol (E3) using LC/MS-MS (Waters Micromass Quattro Micro API). The detailed procedure for the LC/MS-MS analysis is described in the Supplemental Data for sample analysis. In order to verify the significance of the differences in the estrogen concentrations across depth and location, two-way ANOVA tests were performed. However, the interaction effects between depth and location could not be performed due to limited sample size to support this estimation. Also, notice that the slurry analyte compositions at two locations on 2009 were not included due to no LC/MS-MS analyte outputs. The ANOVA results indicated no significant difference (5% level) in the estrogen concentrations between lagoon depths and locations (Table S1). Therefore, the lagoon slurry layer was considered homogenous and modeled as a continuously stirred tank reactor (CSTR) system. For the lagoon sludge layer, a statistical analysis could not be performed due to the fact that we acquired only a single sample at each location of the sampling grid. We thus considered the sludge layer to be represented by a CSTR system for modeling purposes.

MASS BALANCE INTEGRATED BAYESIAN NETWORK MODEL APPROACH

A major difficulty in developing the estrogen budget model is the considerable uncertainty in understanding how the estrogen fate and budget are determined by key mechanisms affecting estrogen compounds in the lagoon stores; understanding major mechanisms will facilitate an assessment of how particular lagoon management actions will help control the level of estrogen compounds in the lagoon. To facilitate a better understanding of the

Accepted Preprint

effects of these variables, we adopted a hybrid modeling approach by combining a Bayesian network with mass balance equations to model the estrogen fate and budgets as well as to assist for the management of estrogen levels in swine waste lagoons.

A Bayesian network (BN) provides a graphical and probabilistic modeling framework for representing and reasoning within domains involving uncertainty. Graphically, a BN represents the system of interest with a set of nodes and arrows via directed acyclic graphs (Pearl 1988; Neapolitan 1990; Lauritzen 1996; Jensen and Nielsen 2007). Nodes are associated with variables in the model, while the arrows represent a conditional probabilistic dependence between the nodes and indicate a certain causal direction. Each node is described by several discrete states of probability distributions (either categorical or interval). Causal dependencies are represented by a set of conditional probability tables (CPTs), describing the strength of the relationship between a node and its parent nodes. The probabilities that populate the CPTs describe the chances of a node being in a specific state given the states of the parent nodes. In the event a node does not have any parents, marginal probabilities are used to define its states (Pollino et al. 2007). The probabilities for the CPTs can be generated from data, expert opinion, process-based models, and/or empirical models (Reckhow 1999). This feature of a BN allows integrating the mass balance models as a meta-model. In this study, we used the mass balance model equations to generate the CPTs for the intermediate nodes due to lack of data for the relationships among intermediate nodes (Supplemental Data, Figure S2 and Table S2-S6). In particular, the object-oriented BN model approach (Jensen and Nielsen 2007) was adopted to populate the CPTs for interrelated physical mechanisms, such as settling and burial reactions, between lagoon slurry and sludge stores.

The use of BNs as tools for characterizing and predicting probabilistic relationships in complex environmental systems has increased over the past decade. Some of the applications outside artificial intelligence include severe weather forecasting (Abramson et al. 1996), fisheries management (Varis 1997), climate change prediction (Varis and Kuikka 1999), and water quality modeling (Reckhow 1999). BNs have been successfully applied to characterize and model water quality impairments in complex systems such as estuaries (Stow and Borsuk 2003; Borsuk et al. 2004; Alameddine et al. 2011) and streams (Reckhow 2010).

As a decision support system, the BN model was used to identify "good" lagoon management actions to control the estrogen budget. Two types of inference are supported with a Bayes network: bottom-up reasoning and

Accepted Preprint

top-down reasoning (Castelletti and Soncini-Sessa 2007). In bottom-up reasoning, new information on the level/category for a particular node is propagated "upward" in the Bayes network to revise the probabilities for the causal parent nodes; this serves as a diagnostic tool. The top-down reasoning is used either to forecast the outcomes when new evidence is available over time, or to perform a "what-if analysis" by propagating the network with different management scenarios and then compare their effects on the outcomes.

MODEL FRAMEWORK

There are two main tasks in developing a mass balance integrated BN model: 1) the construction of mass balance equations for the estrogen budget in both slurry and sludge lagoon stores and 2) the development of Bayesian network model. The following sections describe these tasks in turn.

Estrogen Mass Balance Budget Model Development

Based on the variability analysis of estrogen compounds in the lagoon, the lagoon system was assumed to be a well-mixed slurry liquid layer underlain by a well-mixed sludge layer, as depicted in Figure 1. In this study, the toxic loading model described by Chapra (1997) was adopted to describe the estrogen compounds. The mass balance equations for the estrogen fate and budget models describe the major mechanisms of flushing, sorption, transformation, settling, and burial of estrogen compounds in both the anaerobic lagoon slurry and the sludge treatment systems.

Since a lagoon operation is subject to constant waste loads from swine farrowing barns within a month, the lagoon system is considered in steady-state by month for a certain season. Also, a local equilibrium condition was assumed since the transformation rates of estrogen metabolites is likely to be faster than the input-output or purging rates.

Estrogen Budget Mass Balance Equations in the Lagoon

For a completely mixed lagoon system, the mass balance equations for the total natural estrogen compounds in the slurry and sludge compartments were based on the major mechanisms depicted in Figure 1. Estrogen is partitioned into dissolved and particulate fractions, and transformation is considered between the estrogen metabolites. The settling and burial processes are net losses inside the lagoon. Figure 1 represents the

Accepted Preprint

differential equations describing the change in total natural estrogen compounds and the steady-state total natural estrogen concentrations in slurry and sludge lagoon storage. These mass balance-based models attempt to integrate the significant factors that determine the fate of estrogens in the lagoon system. For the prediction of each estrogen concentration, the equilibrium ratio equations were derived using the reaction mass balances for each estrogen compound (Supplemental Data, Figure S3) and are embedded into the BN model to predict an individual estrogen budget for both slurry and sludge compartments. In this way, each estrogen fate and budget can be predicted by incorporating the interrelated effects of both environmental factors and mechanistic processes. Each component of the model that consists of a set of mass balance equations is represented as a BN node as shown in the next section (Figure 2 and 3).

Suspended Solids Budget Balance Equations in the Lagoon

The waste-flow rate allowed the determination of how natural estrogens were carried along with the wastewater; the suspended solid waste budget model is needed to estimate the levels of estrogen that would be transported along with the suspended solid wastes. Figure S4 in the Supplemental Data represents the differential equations describing the change of suspended solids budget, M1 and M2, inflow, outflow, settling, and burial processes in the slurry and sludge lagoon stores, respectively. These simultaneous equations provide a means for predicting the values of settling and burial rates. At steady-state, the settling rate is derived from the solids budget in the slurry compartment, while the burial rate is obtained by the addition of two of budget equations and then is expressed in terms of sludge porosity and density. For the prediction of settling and burial reaction rates, the derived equations were embedded into the BN model to populate the conditional probabilities (Figure 2 and 3).

Bayesian Network Model Development

The BN model structure was derived by combining the major mechanisms depicted in Figure 1, which are components of the mass balance equations for the estrogen budget and fate in the lagoon. The quantitative relationships between the BN model nodes were established using probability distributions of the derived equations in the two layered lagoon stores.

Accepted Preprint

Bayesian Network Structural Development

The causal structure for the Bayesian network (BN) was constructed using Hugin software (Hugin EXPERT 1998). Figure 2 and 3 show the BN model structure in the lagoon sludge and slurry layers, respectively. The BN consists of three interlinked segments for: management information, reaction information, and fate information. Most of the main model nodes were derived from the components of the mass balance equations and connected by the causal relationships. Within the management information segment, the waste handling nodes (e.g. irrigation, flushing, and barn input) evaluate how the waste management strategies impact estrogen fate in the lagoon. Environmental variables were also included to incorporate the effects of rainfall and evaporation on the lagoon volume. With this structure, we can assess the fate of estrogens under different combinations of environmental conditions and lagoon operations. The reaction information segment includes the nodes for chemical and physical processes, describing the influence of transformations, settling, and burial mechanisms on estrogen fate. These nodes probabilistically determine and compare how each mechanism contributes to purging estrogen compounds in the anaerobic lagoon system. In the fate information segment, the model endpoint node predicts the estrogen concentration in both lagoon layers. The object-oriented BN model approach (Jensen and Nielsen 2007) was used to encode relationships between the lagoon slurry and sludge stores, highlighted by the gray color nodes in Figures 2 and 3.

Parameterization of Model Variables

After development of the causal structure from the barn estrogen loading through the processes affecting estrogen fate in the lagoon, the next step was to quantify the strength of causal relationships between the variables by assigning marginal and conditional probabilities. To accomplish this, in this study, all variables were first discretized to have numeric intervals. The marginal variables were discretized into four to twelve states with equal or quantile intervals under the consideration of observed data distribution as well as expert opinion. For the intermediate nodes, quantile intervals were used for a better accuracy of model prediction (Alameddine et al 2011). After discretization, all nodes were assigned to have discrete probability distributions. The marginal probabilities were assigned for the parentless nodes (e.g. precipitation, irrigation, flushing, etc.) to represent the prior knowledge about frequencies of each state. The frequency data for environmental nodes were obtained from the National

Accepted Preprint

Climate Data Center (NCDC 2011) and operational variables were obtained from the field site facility manager. Where no data were available for the marginal nodes in the reaction information segment, such as estriol (E3) degradation rate, the expected range of state values were derived by interpolating the reaction rates based on given the ranges of state values of other model variables such as input, output estrogen loading, and settling rates in the mass balance equations in Figure 1. The probabilities for each state were then generated under the assumption of a lognormal distribution (Ott 1995) given the estimated range of state. The conditional probabilities were generated for the remaining intermediate nodes using the mass balance equations in Table S3 and S5. Using Monte Carlo simulation within Hugin, probabilities of each state of the marginal and conditional nodes were calculated by generating a number of samples (25 by default) within each interval and estimating the frequency for the corresponding states. With this method, the inherent variability of environmental variables and parameter uncertainty can be captured in both marginal and conditional probability distributions. Overall, the link strength enables the accounting for the structural and parameter uncertainties of the model. A more detailed explanation of the quantification of each node is provided in the Supplemental Data [Table S2-S6] and the example of parameter propagation using the Hugin software is presented in Figure S5.

MODEL APPLICATION AND EVALUATION

The developed Bayesian network (BN) model was implemented for the selected lagoon system and used to assess management options to reduce natural estrogen loading to the anaerobic swine lagoon. The model performance was evaluated to determine the ability of a BN to produce probability distributions that capture the behavior of a deterministic mass balance model using two metrics: i) an analysis of likelihood findings and ii) sensitivity analysis.

The analysis of likelihood findings was performed to evaluate the BN model prediction. Since there is not a single set of data to evaluate the entire BN model, we used a likelihood evidence inference function in Hugin (Hugin EXPERT 1998) to evaluate the prediction of the Bayesian network model on three individual estrogen metabolite concentrations using the evidence in the total estrogen concentration node. The likelihood evidence function provides a way to assess probabilities for all of the other variables via belief updating, once new evidence is entered into the BN. To insert this likelihood evidence for the total estrogen concentration sampled from the

Accepted Preprint

lagoon storages into the BN, the measured total estrogen concentrations were converted into probability distributions and added to the Hugin 'Enter likelihood evidence' function. Once this new evidence for total estrogen concentration from two different lagoon storages was entered into the network, the posterior probabilities for each of the estrogen metabolite concentration nodes were updated through a message propagation mechanism within the Hugin BN algorithm. In other words, parent nodes (e.g., the total slurry and sludge estrogen concentration nodes) pass "π messages" to their child nodes (e.g., each estrogen metabolite node) as well as to other adjoining nodes, such as waste outflow rate, settling reaction, and E3 degradation rate (Figure 2 and 3). The π messages characterize the evidence for total estrogen concentrations in the lagoon storages within the network. Likewise, child nodes (each estrogen metabolite node) and the other adjacent nodes pass "λ messages" back to their parents encapsulating all of the likelihood values. The updated probability distributions of estrogen metabolite concentration nodes are then estimated by combining the π and λ messages with the conditional probabilities matrix given all the possible combination values of total slurry and sludge estrogen nodes as well as the equilibrium ratios. The mathematical background of the approach is extensively covered by Neapolitan (2003), Jensen and Nielsen (2007), and Pearl (1988).

A sensitivity analysis was performed to measure the sensitivity of a target node in the posterior distribution to variations in the evidence entered in other nodes of the network. The target node in this study was the total estrogen concentration in the lagoon storages. The evidence sensitivity was measured as the mutual information, I(X,Y), which represents the effect of one variable (X) on another variable (Y) and is calculated as follows:

$$I(X,Y) = H(Y) - H(X|Y) \quad (1)$$

where H(Y) is entropy, which is the measure of the uncertainty or randomness of a variable (Y) represented by a probability distribution (Pearl 1998; Korb and Nicholson 2004; Pollino et al. 2007).

Accepted Preprint

RESULTS and DISCUSSION

Evaluation of Model Prediction

The predictive accuracy of the BN model was assessed indirectly based on the estrogen budget, using the analysis of likelihood findings. This was determined from the evidence of the total estrogen concentrations in the slurry and sludge storages sampled in 2009 (Figure 4). The likelihood distributions for each estrogen concentration were then compared with the frequency of each estrogen concentration in both the lagoon slurry and the sludge layers (Figure 5 and 6).

Figures 5 and 6 demonstrate that our model tends to over-predict the estrogen concentrations for the higher levels of observed estrogen in both the slurry and the sludge layers. This discrepancy could be caused by the uncertainties in mechanisms such as settling and degradation reactions. However, the overall model predictions are relatively consistent with field observations and the average of model performance error was between 14 and 24 ng/L in terms of a mean absolute error. This indicates that the model is relatively adequate for prediction of the probability distribution of each estrogen concentration.

Sensitivity Analysis

A sensitivity analysis was performed to find the most influential variable for prediction of the total estrogen budgets in the lagoon storage. We used the concept of mutual information to highlight sensitivity. In this analysis, we calculated the mutual information for the variables affecting total estrogen concentration under the different lagoon models (Table 1 and 2). Table 1 illustrates that for the lagoon slurry layer where the barn estrogen input concentration node has the largest effect on the determination of the estrogen budgets, followed by the settling rate.

In the sludge layer, the settling rate has the largest effect, followed by the barn estrogen input (Table 2). This indicates that adjusting the swine barn facility management practice as well as enhancing the physical process of the settling process could reduce the estrogen mass in the lagoon. The results also suggests that the effect of the physical environment such as rainfall, irrigation, and lagoon volume as well as management factors including irrigation and flushing were less significant as compared to the effect of estrogen input from the swine barn

Accepted Preprint

facilities. The biological degradation mechanism was also shown to have a minor effect on reducing the estrogen compounds from both lagoon layers.

Prediction of Average Total Estrogen Budget in the Lagoon

The average total estrogen budget with uncertainty predicted by the BN models is presented in the context of the major sources and sinks of estrogen compounds in both lagoon compartments along with the percent total mass accounted for by each mechanism (Figure 7). It is clearly shown that the settling reaction accounts for the highest removal mechanism in the lagoon slurry storage, while the burial reaction proved to be the most significant factor to eliminate the estrogen levels in the sludge storage. The result also indicates that the estrogen losses by degradation reactions are minor in both storages. The results thus concur with the findings of the sensitivity analysis.

What-If Scenarios for Lagoon Management

Based on the sensitivity analysis, it appears that the estrogen levels in the lagoon could be reduced by controlling the estrogen inputs from swine barn facilities or enhancing the settling reaction in the lagoon. Thus, "what-if" scenarios were developed to verify the effect of these management options on the total estrogen budgets and identify the best management option.

Effect of Swine Barn Estrogen Inputs on Estrogen Budgets in the Lagoon

The effect of barn estrogen inputs was assessed by simulating "what-if" scenarios considering three ranges of barn estrogen inputs (low:0-10000 ng/L, medium: 30000-40000 ng/L, and high: 60000-70000 ng/L) (Figure 8). The model findings demonstrate that the levels of total estrogens differed by the magnitude of the barn estrogen inputs in both lagoon slurry and sludge compartments. The more barn estrogens enter into the lagoon, the higher the estrogen in both lagoon layers. This confirms that adjusting estrogen input concentrations will reduce the estrogen levels in the lagoon. Figure 8 illustrates that the low level of estrogen input would drastically lower the estrogen levels in both lagoon storages.

Accepted Preprint

Effect of Inflow TSS Concentration on Estrogen Budgets in the Lagoon

Since the settling reaction is a function of the inflow total suspended solids (TSS) concentration, which is a controllable variable in the context of lagoon management, the effect of inflow TSS concentration on the settling reaction was assessed for both lagoon layers by simulating three ranges of inflow TSS concentrations. The model results show that a higher inflow TSS concentration tended to reduce the estrogen levels in both slurry and sludge storages (Figure 9). This indicates that estrogen compounds in the slurry layer are more likely to bind the suspended particles and settle into the sludge layer. Higher TSS reduces the estrogen compounds in the sludge layer via the burial reaction, which is a function of inflow TSS concentration. In the context of best management practices, aiming at reducing estrogen levels in the lagoon, this finding implies that the use of organic polymers or inorganic coagulants could enhance the binding and thus reduce estrogen levels effectively.

CONCLUSIONS

Given the high levels of natural estrogens that are known to occur in anaerobic commercial swine lagoons, and the potential for these compounds to adversely impact aquatic communities, the development of a probabilistic decision support system that can be used to evaluate management options for these compounds is a high priority. In this study, a mass balance integrated Bayesian network (BN) model was developed to assess the estrogen fate and budget in a swine waste lagoon system. Since marginal probability distributions reflect the uncertainty in parent nodes, and conditional probability distributions represent the uncertainty in child nodes (which includes the uncertainty in the parent-child relationship), a BN provides a systematic way to account for the uncertainties in the relationships for the major physical, chemical, and biological processes via these marginal and conditional probability distributions.

As we have shown, our BN provides a useful framework to assist in decision making to reduce estrogen transport from commercial swine operations. Through a “what-if scenario analysis” as well as with sensitivity analysis, we demonstrated that the BN model could assist decision makers to identify management actions that have the greatest influence on the distribution and fate of estrogen compounds in the anaerobic lagoon.

It is vital to recognize that data for mechanistic model development for this problem, or for development of an entirely data-based model, do not presently exist. In addition, owners and operators of CAFOs are

Accepted Preprint

understandably quite reluctant to invite researchers to monitor their facilities; we experienced that reluctance during the course of our study. In our view, these realities necessitated a novel strategy to allow decisions to be, not simply based on best professional judgment, but on a probabilistic model that allowed data gaps to be filled with quantified scientific assessment. We believe that this study represents an important first step in a comprehensive modeling analysis to address decisions concerning estrogens and other chemicals found in animal feedlot operations.

ACKNOWLEDGMENTS

This work was supported by Environmental Protection Agency Science to Achieve Results (STAR) grant (R833420 to SWK) and Environmental Protection Agency Science to Achieve Results (STAR) graduate fellowship (FP917151 to EY). We wish to thank the swine facility plant manager and crew for their assistance in sample collection and the facility owner operator for providing access to the field site.

Accepted Preprint

REFERENCES

Abramson B, Brown J, Edwards W, Murphy A, Winkler RL. 1996. Hailfinder: A Bayesian system for forecasting severe weather. *Int J Forecasting* 12:57-71.

Alameddine I, Cha Y., Reckhow KH. 2011. An evaluation of automated structure learning with Bayesian networks: An application to estuarine chlorophyll dynamics. *Environ Model Softw* 26:163-172.

Borsuk ME, Clemen R, Maguire L, Reckhow KH. 2001. Stakeholder values and scientific modeling in the Neuse River watershed. *GPDN* 10:355-373.

Borsuk ME, Stow CA, Reckhow KH. 2004. A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. *Ecol Model* 173:219–239.

Bradford SA, Seqal E, Zheng W, Wang Q, Hutchins SR. 2008. Reuse of concentrated animal feeding operation wastewater on agricultural lands. *J Environ Qual* 37:S97-S115.

Castelletti A, Soncini-Sessa R. 2007. Bayesian networks in water resource modeling and management. *Environ Modell Softw* 22:1073-1074.

Chapra SC. 1997. Surface water-quality modeling. Long Grove, IL, USA: Waveland Press. 844 p.

Cheng J. 2003. Challenges of CAFO waste management. *J Environ Eng* Editorial:391-392.

Korb KB, Nicholson AE. 2004. Bayesian artificial intelligence. London, UK: Chapman and Hall/CRC Press. 364 p.

Hugin EXPERT. 1998. Hugin Help Pages. [cited 2010 January 13]. Available from : http://www.hugin.com.

Jensen FV, Nielsen TD. 2007. Bayesian networks and decision graphs (second edition). New York, NY, USA: Springer Science + Business Media LLC. 447 p.

[NCDA & CS] North Carolina Department of Agriculture & Consumer Services. 2011. Annual statistics book. [cited 2011 April 28]. Available from: http://www.ncagr.gov/stats/2011AgStat/index.htm

[NCDC] National Climatic Data Center. 2011. National Oceanic and Atmospheric Administration. [cited 2011 January 20]. Available from: http://ncdc.noaa.gov.

[NCDWQ] North Carolina Division of Water Quality. 2009. Aquifer Protection Section. [cited 2011 March 6]. Available from: http://portal.ncdenr.org/c/document_library.

Neapolitan RE. 1990. Probabilistic reasoning in expert systems: theory and algorithm. New York, NY, USA: John Wiley. 433 p.

Neapolitan RE. 2003. Learning Bayesian networks. Upper Saddle River, NJ, USA: Prentice Hall. 674 p.

Lauritzen S. 1996. Graphical models. Oxford, UK: Clarendon Press. 308 p.

Leet JK, Gall HE, Sepulveda MS. 2011. A review of studies on androgen and estrogen exposure in fish early life stages: effects on gene and hormonal control of sexual differentiation. *J Appl Toxicol* 31: 379-398.

Ott WR. 1995. Environmental statistics and data analysis. Boca Raton, FL, USA: CRC Press L.L.C. 336 p.

Pearl J. 1988. Probabilistic reasoning in intelligent systems. San Mateo, CA, USA: Morgan Kaufmann. 552 p.

Pollino CA, Woodberry O, Nicholson A, Korb K, Hart BT. 2007. Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. *Environ Modell Softw* 22:1140-1152.

Reckhow KH. 1999. Water quality prediction and probability network models. *Can J Fish Aquat Sci* 56:1150-1158.

Reckhow KH. 2010. Bayesian networks for the assessment of the effect of urbanization on stream macroinvertebrates. In: Proceedings International Conference on System Sciences; 2010 Jan 5-8; Kauai (Hawaii), USA: System Sciences (HICSS). p 1-8.

Stow CA, Borsuk ME. 2003. Enhancing causal assessment of estuarine fishkills using graphical models. *Ecosystems* 6:11-19.

Varis O. 1997. Bayesian decision analysis for environmental and resource management. *Environ Modell Softw* 12: 177–185.

Varis O, Kuikka S. 1999. Learning Bayesian decision analysis by doing: lessons from environmental and natural resources management. *Ecol Model* 119:177-195.

Yost EE, Meyer MT, Dietze JE, Meissner BM, Worley-Davis L, Williams CM, Lee B, Kullman SW. 2013. Comprehensive assessment of hormones, phytoestrogens, and estrogenic activity in an anaerobic swine waste lagoon. *Environ Sci Technol* 47:13781-13790

Accepted Manuscript

TABLE LEGENDS:

Table 1. Sensitivity analysis for total slurry estrogen concentrations

Table 2. Sensitivity analysis for total sludge estrogen concentrations

FIGURE LEGENDS:

Figure 1. Mass balance equations for total estrogens with major mechanisms in the lagoon slurry and sludge layers (E: slurry total estrogen concentration, Es: sludge total estrogen concentration, Q: Inflow/Outflow, V: lagoon volume, k: E3 degradation rate, F_3: fraction of E3 compound, A: lagoon surface area, F_3: fraction of particulate concentration, V_s: settling rate, k_b: burial rate).

Figure 2. Bayesian network structure of estrogen budget in the lagoon slurry layer. Gray boxes represent the model nodes that link interrelated relationships between the slurry storage and the sludge storage.

Figure 3. Bayesian network structure of estrogen budget in the lagoon sludge layer. Gray boxes represent the model nodes that link interrelated relationships between the slurry storage and the sludge storage.

Figure 4. Analysis of likelihood evidence of total estrogen concentration in the slurry storage given the observed 2009 total slurry estrogen concentrations. The blue bars present the likelihood probability distribution of total estrogen concentrations, while the green bars indicate the output probability distributions of likelihood findings analysis.

Figure 5. Comparison BN model outputs with observed estrogen data in the slurry storage given the likelihood of total slurry estrogen concentrations on 2009.

Figure 6. Comparison BN model outputs with observed estrogen data in the sludge storage given the likelihood of total sludge estrogen concentrations on 2009.

Figure 7. Average total estrogen budgets predicted in lagoon slurry and sludge stores. The uncertainties are given as ± values which represent the standard deviation.

Figure 8. Effects of barn estrogen inputs on total estrogen budgets in slurry (a) and sludge (b) lagoon layers.

Figure 9. Effect of inflow total suspended solids (TSS) on total estrogen budgets in slurry (a) and sludge (b) lagoon layers.

Accepted Manuscript

Table 1. Sensitivity analysis for total slurry estrogen concentrations

Model variables	Mutual information
Barn estrogens input	0.37
Settling rate	0.16
Inflow TSS concentration	0.13
Fraction of particulate estrogens	0.004
Lagoon waste outflow rate	0.001
Slurry TSS concentration	0.0005

TSS = total suspended solids.

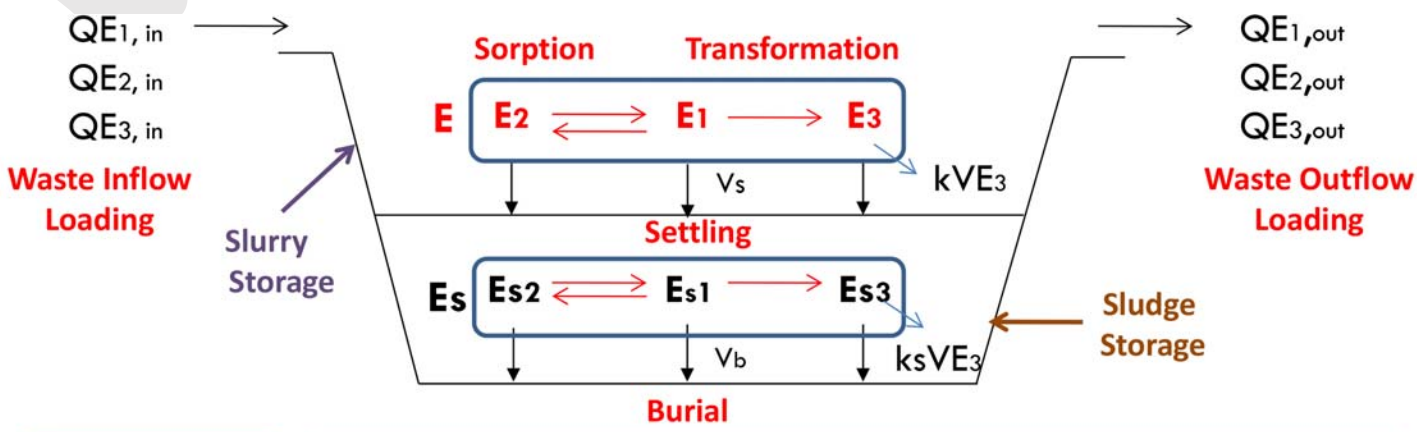
Accepted Preprint

Table 2. Sensitivity analysis for total sludge estrogen concentrations

Model variables	Mutual information
Settling rate	0.24
Barn estrogens input	0.23
Slurry total estrogens	0.22
Slurry E1	0.17
Inflow TSS concentration	0.14
Burial reaction	0.14
Slurry E3	0.05
Slurry E2	0.04
Fraction of particulate estrogens	0.02
Sludge degradation reaction	0.003
Slurry TSS concentration	0.003

TSS = total suspended solids.

Accepted Print



Total Estrogen Budget in Slurry

- $$V\frac{dE}{dt} = Q_{in}E_{in\ (urine+feces+slurry)} - Q_{out}E - kVEF_3 - V_sAEF_p$$
- $$E = \frac{Q_{in}E_{in}}{Q_{out} + kVF_3 + VsAF_p}$$

Total Estrogen Budget in Sludge

- $$V\frac{dE_s}{dt} = V_sAEF_p - kVEF_3 - k_bVEs$$
- $$E_s = \frac{VsAEF_p}{k_bV + ksVF_3}$$

Figure 1

Accepted

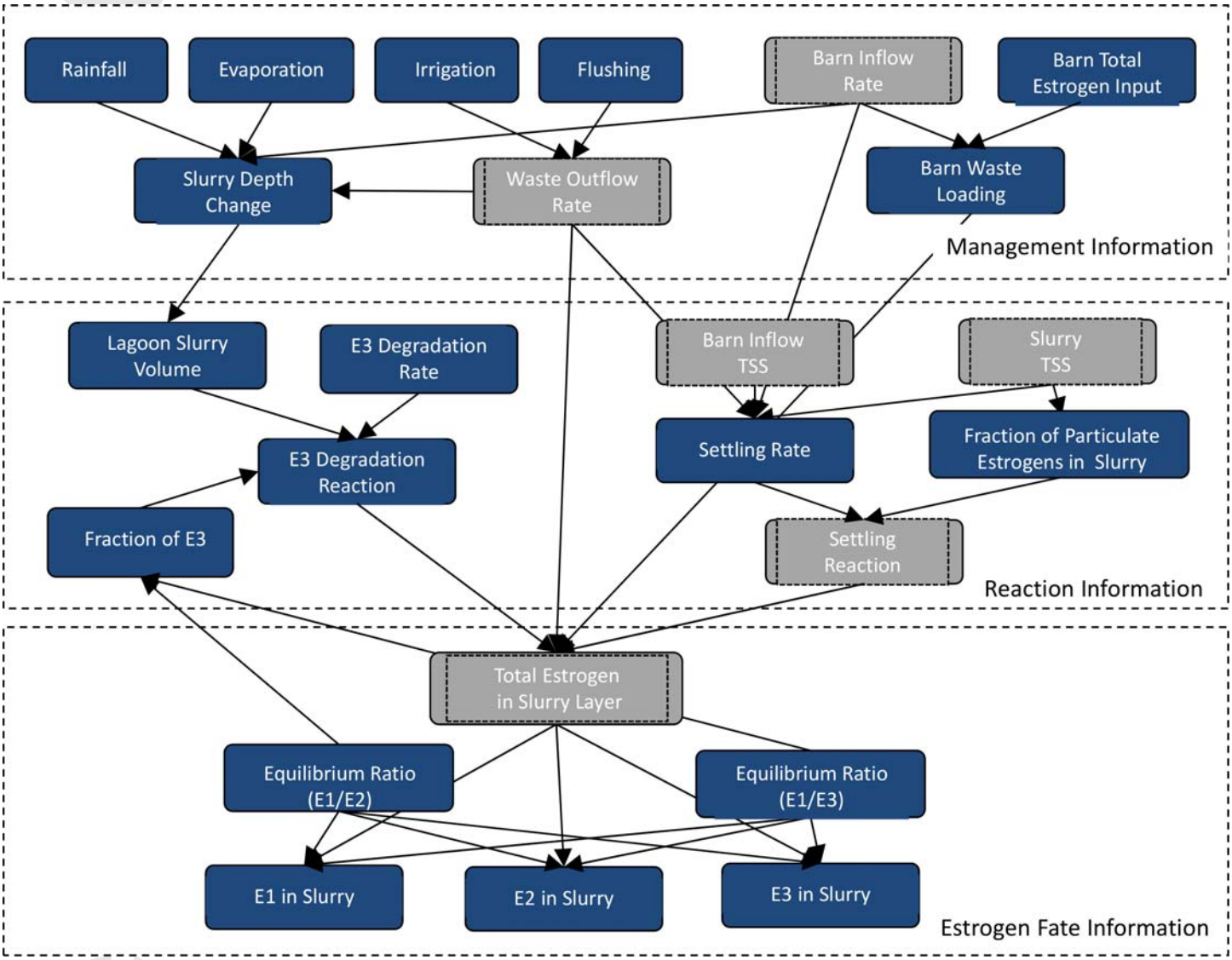


Figure 2

Accepted

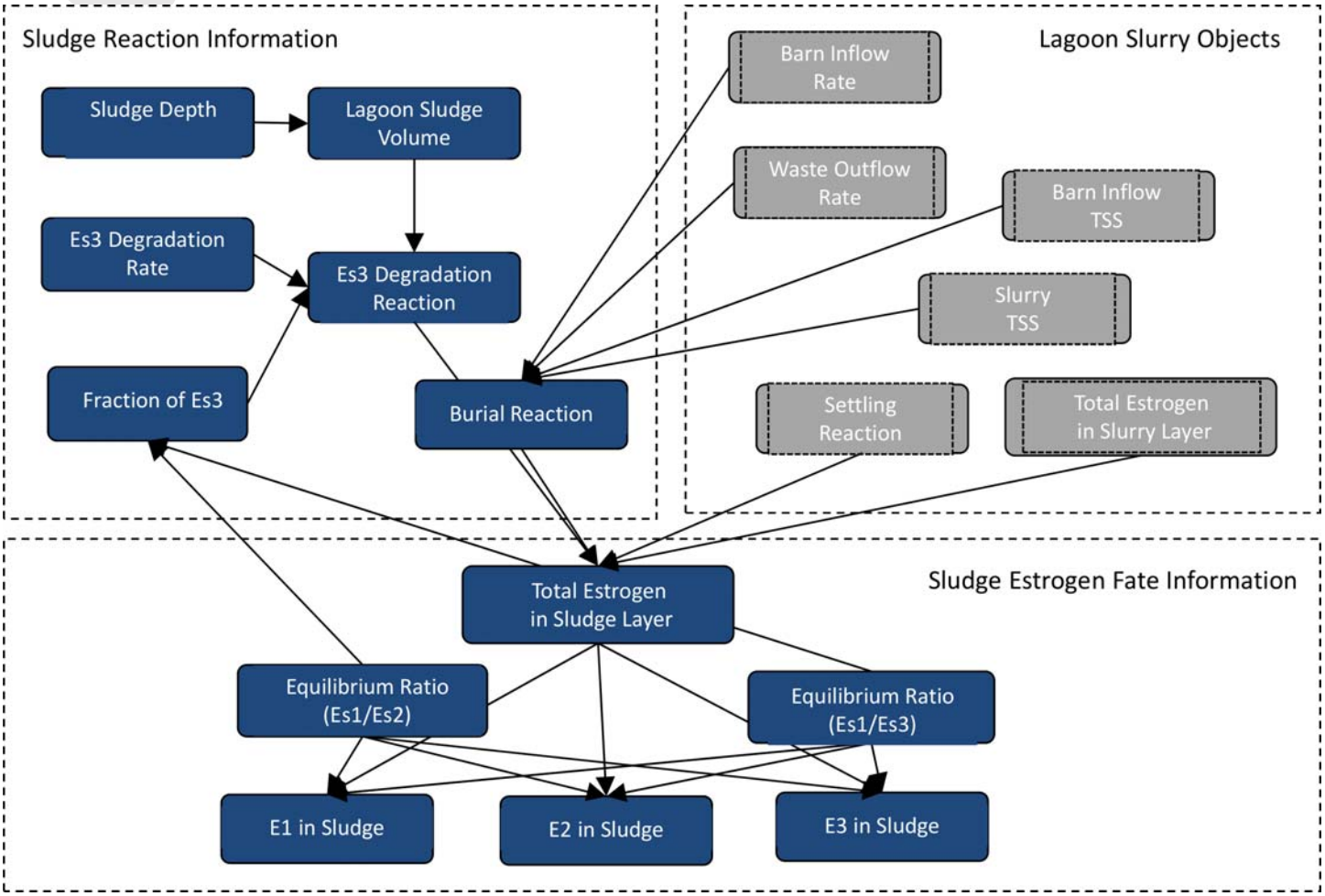


Figure 3

Accepted

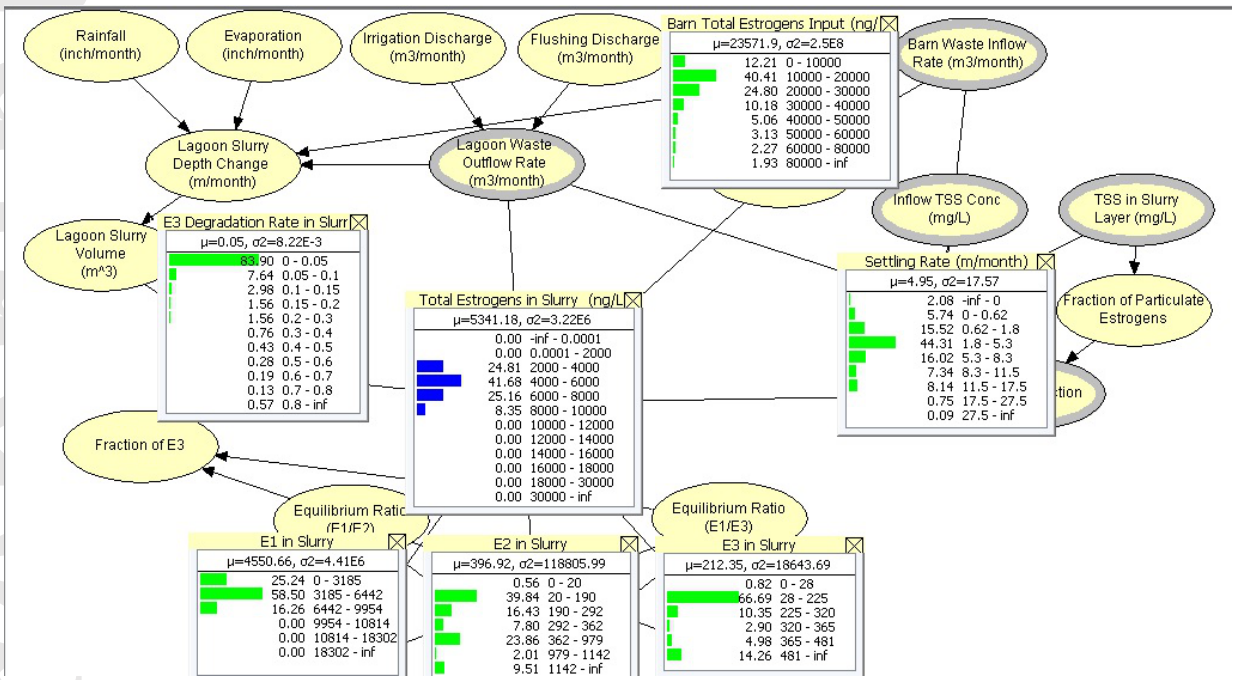


Figure 4

Accepted Preprint

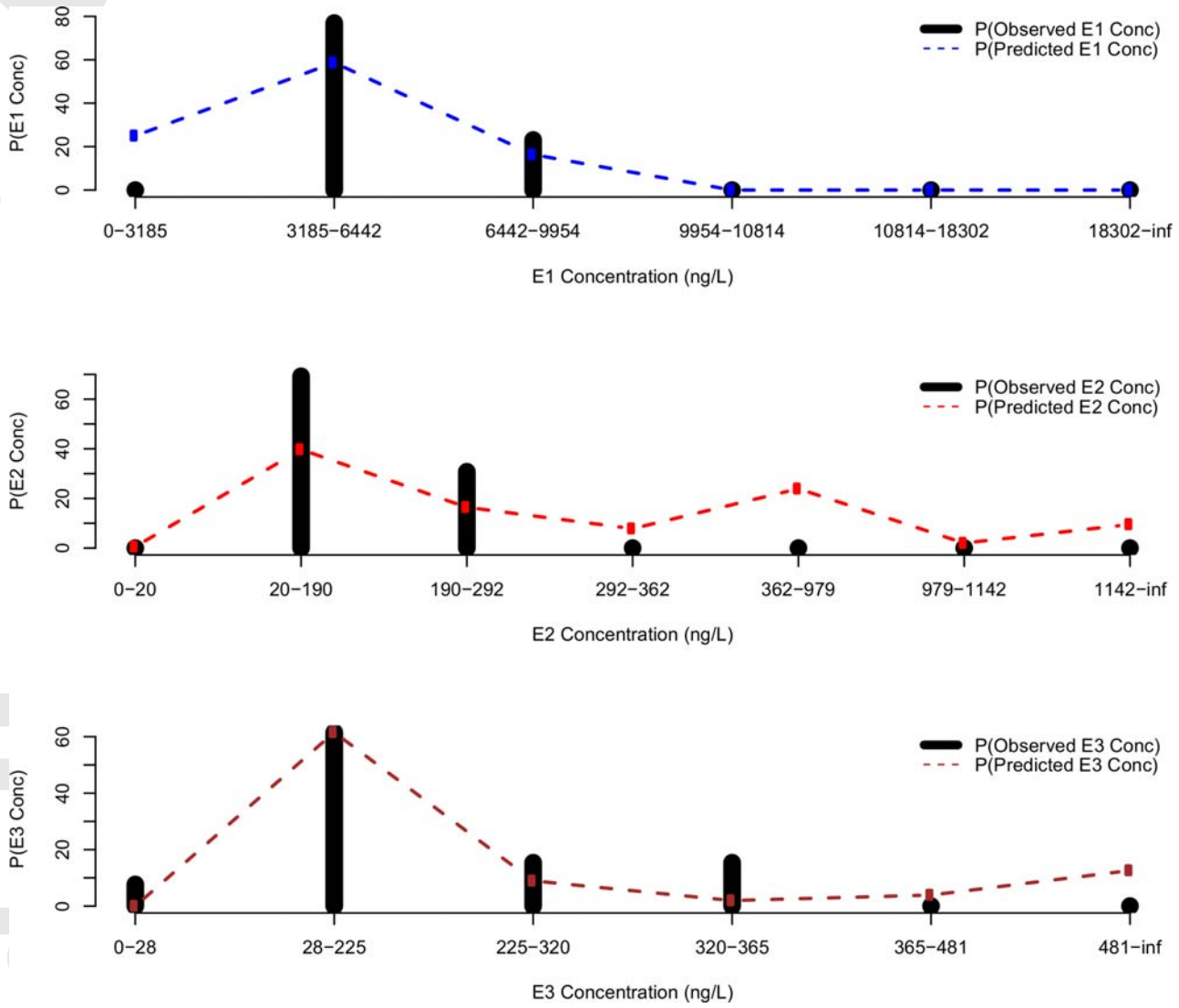


Figure 5

Accepted

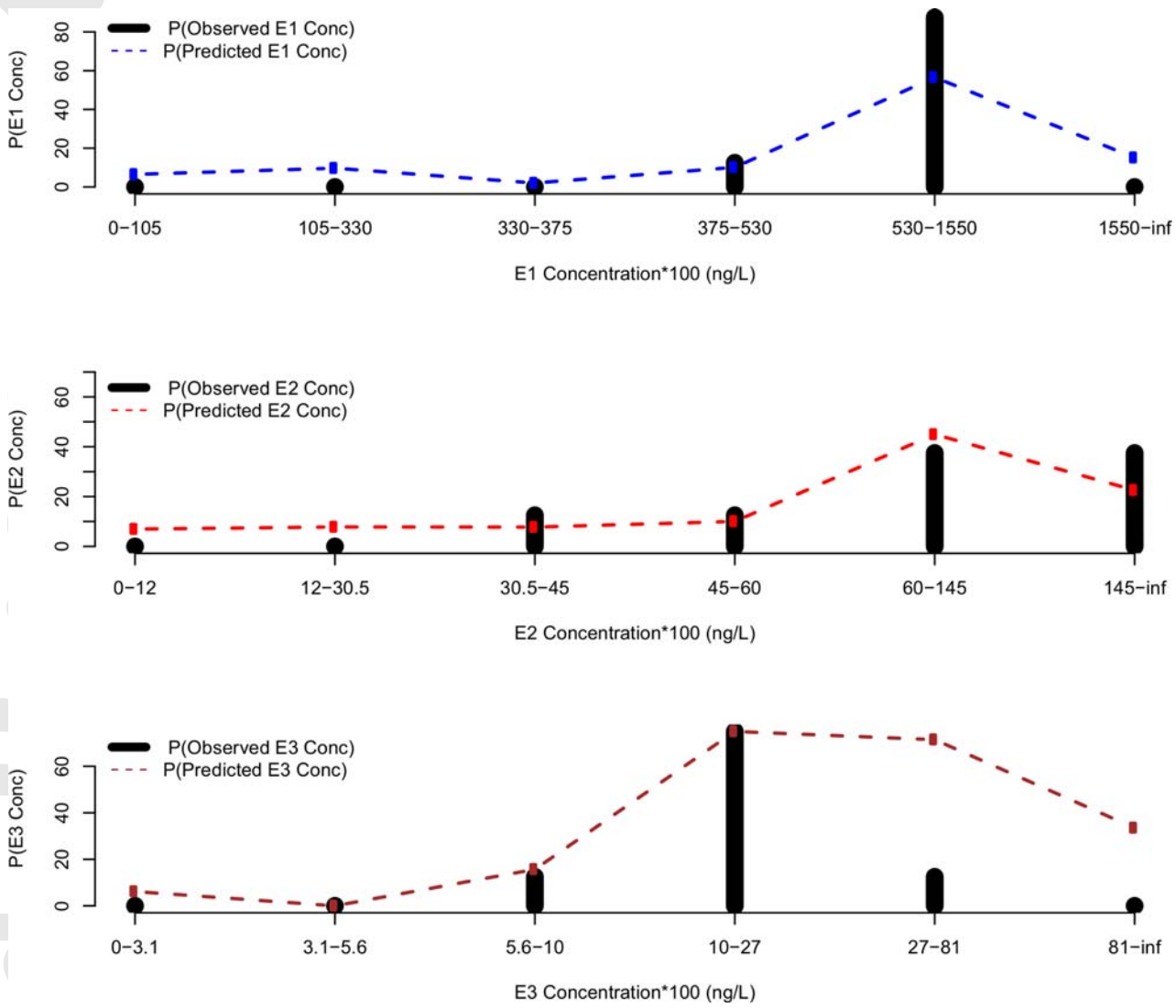


Figure 6

Accepted

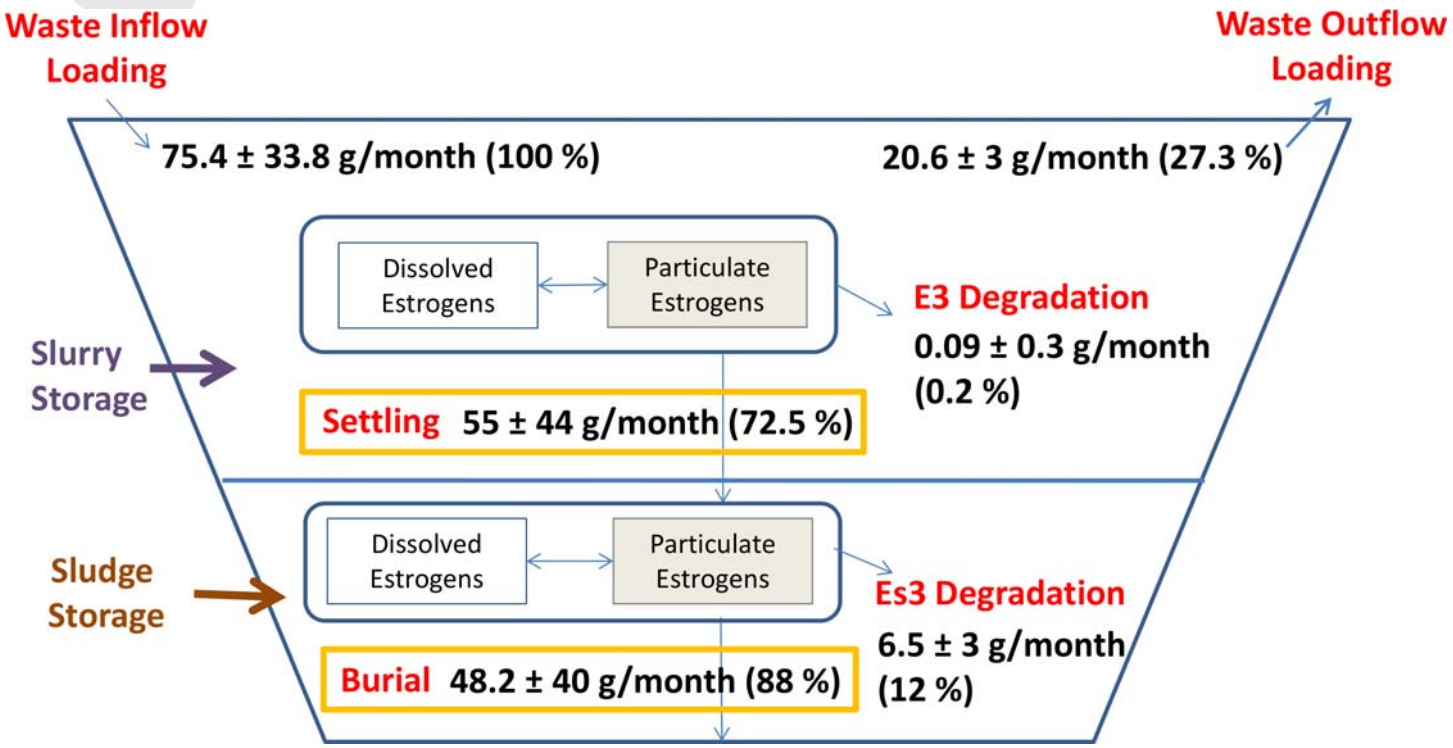


Figure 7

Accepted P

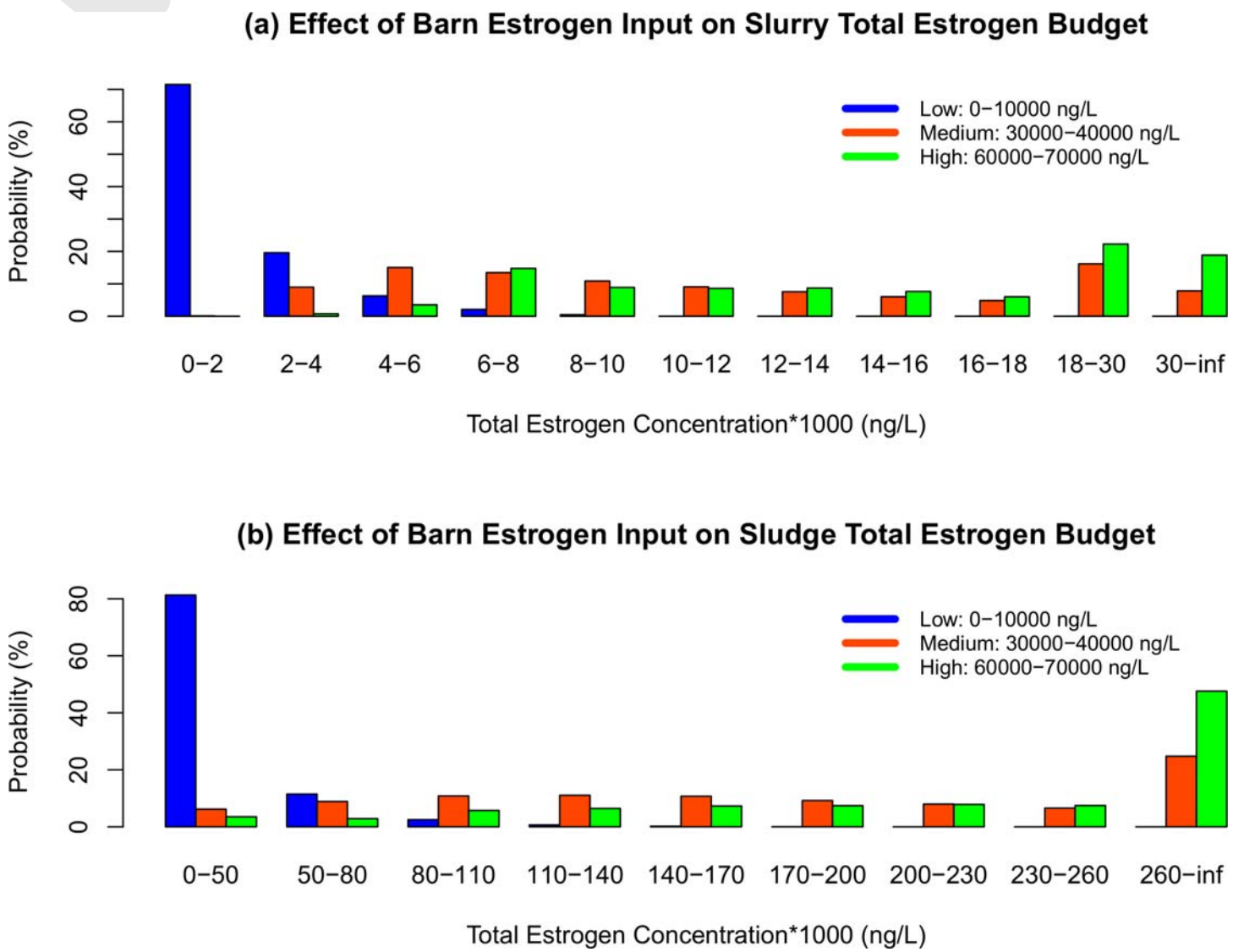


Figure 8

Accepted

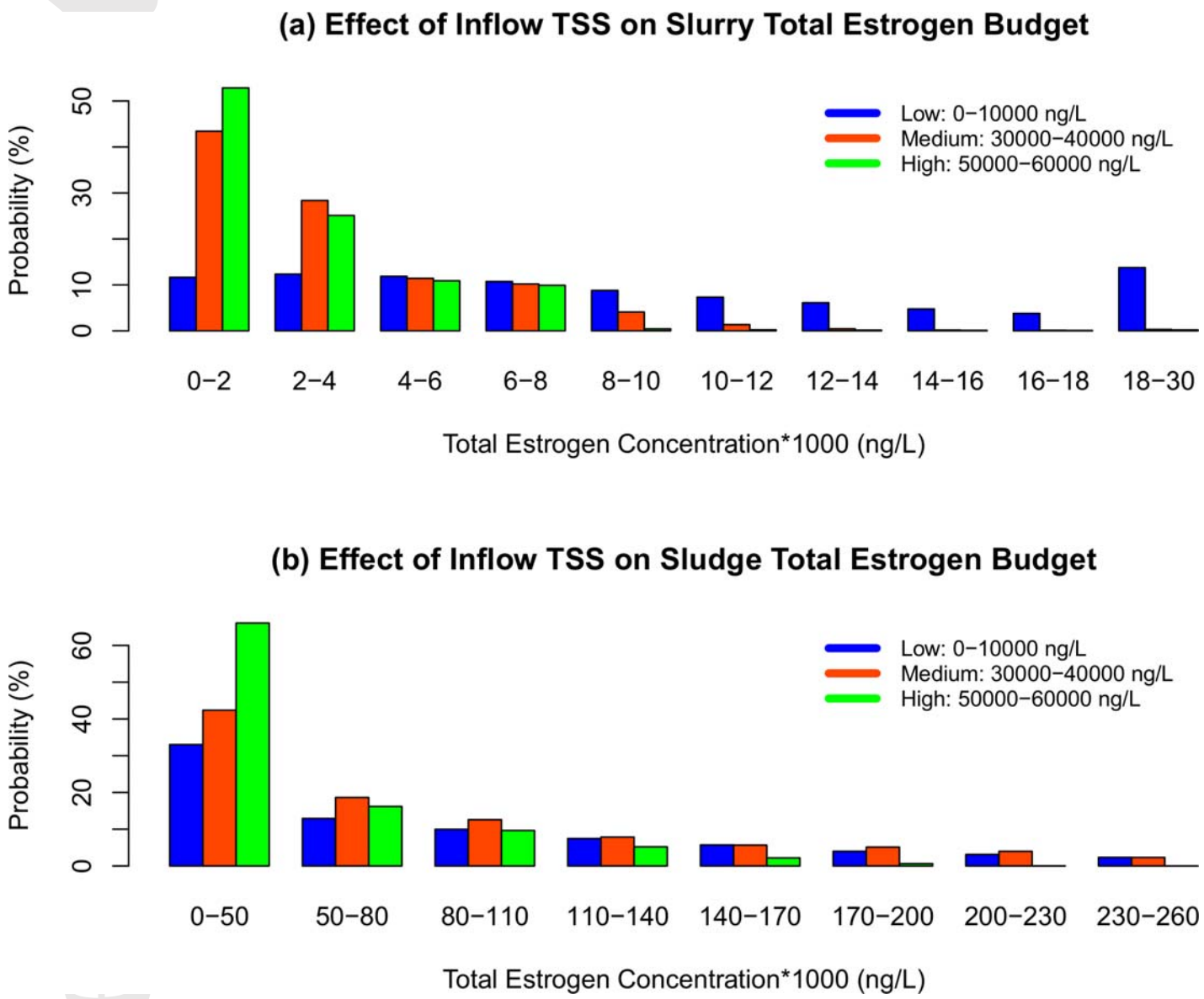


Figure 9

Accepted