LABORATORY TOXICITY AND FIELD EFFECTS OF A COMPLEX MIXTURE: OIL-FIELD PRODUCED WATER

By

JONATHAN C. FISHER

Bachelor of Science in Agriculture
Louisiana State University
Baton Rouge, LA
2003

Master of Science in Fisheries
Louisiana State University
Baton Rouge, LA
2005

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2010

LABORATORY TOXICITY AND FIELD EFFECTS

OF A COMPLEX MIXTURE: OIL-FIELD

PRODUCED WATER

Dissertation Approved:

Dr. Jason B. Belden

Dissertation Adviser

Dr. Kristen A. Baum

Dr. Andrew R. Dzialowski

Dr. Daniel E. Shoup

Outside Committee Member

Dr. Mark E. Payton

Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to thank my major advisor, Jason Belden, for nurturing me through my dissertation and helping my mind go where my interests lay. My committee members, especially Kristen Baum, provided all kinds of assistance from lab space, sage advice, conversation, and teaching experience. Joe Bidwell gave me the opportunity to come to OSU in the first place and I thank him for his initial support.

I would like thank Melissa Hutchinson, as well as Billy and Camilla, for their loving support and for pushing me along when I wanted to give up. I would not be here if not for the loving example and inspiration provided by Debra and Steve Kelly.

Finally, I am dedicating my dissertation to a childhood friend who passed recently. Gordon Jones and I grew up across the street from each other, played ball and swam in his pool more times than I can remember. We went to different schools, eventually grew apart, and I never saw him after I left for college. Gordon passed on April 20, 2010 when the Deepwater Horizon offshore oil rig exploded off the coast of Louisiana. Gordon was taken from his wife Michelle, his son Stafford, and his (at the time of his passing) unborn son Maxwell. I will always cherish the joy that Gordon and his family brought to my childhood.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

GENERAL INTRODUCTION

This dissertation investigated how organisms in the field and the laboratory respond to complex mixtures or combinations of stressors. Organisms are continually exposed to natural and anthropogenic stressors in the environment. Traditional toxicology deals with the two types of stressors in different manners. Natural stressors are typically evaluated in ordered treatments, such as high and low dissolved oxygen or temperature. Anthropogenic stressors, especially chemicals, are often the primary focus of toxicologists and are generally considered across a wide range of concentrations. This range of concentrations is then statistically modeled to produce a dose response curve that relates chemical concentration to a biological effect or endpoint. Toxicology is applied through the use of environmentally relevant chemical concentrations and dose response curves.

Dose response curves can take many forms, but the probit, a cumulative distribution function, and the logit, a probability function, are the most commonly used curves. An endpoint is the dependent or response variable in a dose response relationship, and can be any measureable biological response that can be modeled by a dose response curve. Endpoints are recorded during and at the conclusion of toxicity tests, which are experiments in which organisms are exposed to specific concentrations of a chemical of interest. The most common endpoints involve lethality, or how many organisms survive after some time period in the various treatment concentration

replicates. The toxicity tests in Chapter 2 and Chapter 3 of this dissertation primarily use lethality or survival as an endpoint. Sublethal endpoints encompass a wider variety of biological responses, including growth, reproduction, enzyme activity, or behavior. The *in situ* and laboratory experiments in Chapter 3 utilize individual growth, and the algae experiments in Chapter 4 uses population growth as endpoints.

Toxicologists derive specific effective concentrations of regulatory importance from dose response curves. These derived values are generated using a rarely used, outside of toxicology, field of statistics originally developed by R. A. Fisher called fiducial statistics or inverse probability (Aldrich 2000). Inverse probabilities are distinct from traditional statistics in that the dependent variable (effect) from a regression (dose response curve) is used to derive the independent variable (concentration) value that would cause that effect. Derived values have the following four parts: duration, endpoint type, percent effect, and concentration. For example, in Chapter 4 atrazine was found to have a 96h EC50 of 342µg/L to the algal test organism. This derived value is interpreted by saying that after 96 hours of exposure the concentration of atrazine that results in a 50% reduction in the algal population is 342µg/L. The duration can be any relevant duration, but typically range from 24h to 30d. The endpoint type is limited to LC (lethal concentration), EC (effective concentration), and IC (inhibitive concentration). The percent effect is any whole number percent effect and 50% is the most commonly used effect level because of reduced variability at this inflection point of the probit and logit dose response curves.

A large part of this dissertation is the derivation and evaluation of models that can predict effect of toxicants (chemicals that cause toxicity) occurring together in a mixture. This dissertation utilizes two types of predictive models: statistical or empirical models and theoretical models. Statistical or empirical models are regression models derived using specific organisms for a set number of contaminants. For example, Chapters 2 and 3 rely heavily on a logistic regression to relate the concentrations of major ions (Na, Ca, K, Mg, Cl, SO_4, and HCO_3) in a solution with survival of

three standardized bioassay organisms (Mount et al. 1997). Empirical models are most commonly used on metal mixtures as the total number of potential contaminants is limited and feasibly testable in a laboratory setting. Theoretical models are based on knowledge about how stressors cause toxicity and use the similarity or dissimilarity in mode of action to combine single component dose responses into multi-component responses. Organic compounds and different types of stressors occur in a logistically un-testable number of combinations resulting in very little knowledge about their combined effects and necessitating the use of theoretical mixing models rather than empirical ones (Lydy et al. 2004).The two theoretical models used in Chapter 4 are concentration addition (CA, Altenburger et al. 2000) and independent action (IA, Backhaus et al. 2000).

CA is used to create mixture effects for compounds that act similarly on an organism (Altenburger et al. 2000). This model treats all individual stressors as if they were all the same stressor. This is achieved by first standardizing the concentrations of each stressor using its relative toxicity. This can be done by dividing the toxicant concentration by that toxicant's EC50. If all toxicant concentrations are transformed in this manner, then all the stressor concentrations are now in the same units (percent of a concentration that would cause 50% effect to the organism or toxicity unit) and can be added together. The sum of the toxicity units (TUs) can then be used to derive the expected effect from the combination of those similarly acting stressors. For example, in Chapter 4 a chemical that inhibits an algal cell from making cellular energy from light energy is combined with limiting the amount of light energy available to the cell. Because both stressors limit the energy available to the cell, CA is tested to see if that model accurately predicts the effects of mixtures of those two stressors. This analysis also represents the first instance in the toxicological literature that natural and chemical stressors are combined using the CA model.

IA is used to combine effects for compounds that act dissimilarly on an organism (Backhaus et al. 2000). This model treats the effects of each stressor as completely separate events that happen to occur simultaneously. The probability of effect for all stressors are multiplied together to get a

joint probability of effect. For example, if two stressors occur at levels that would cause 60% and 50% mortality respectively, then the combined effect using IA would be 80% mortality. This means that the first stressor kills off 60% of the exposed organisms leaving 40%. The second stressor then kills off 50% of the remaining organisms leaving 20% of the organisms not killed by either stressor, or 80% mortality. Another example occurs in Chapter 5 where salts (ionic stressors) are combined with polycyclic aromatic hydrocarbons (PAHs, cause narcosis) and the IA model is tested for fit to those mixture combinations. Because those compounds cause toxicity in completely different manners, IA is expected to best model that interaction.

Finally, stressors sometime combine in manners that differ from any mixture modeling predictions. Effects that are greater (more toxic) than predicted by any model are called synergistic. Effects that are less (less toxic) than predicted by any model are called antagonistic. When these conditions (synergism or antagonism) occur, there is most likely some other process or interaction between the stressors occurring in addition to the expected interaction that affects the final mixture effects. For example, when fathead minnows are exposed to the pesticides chlorpyrifos and esfenvalerate, measured effects are greater (more toxic) than predicted by either CA or IA (Belden and Lydy 2006). This synergistic effect was the result of chlorpyrifos inhibiting the enzyme system that detoxifies esfenvalerate.

Literature Cited

Aldrich, J. 2000. Fisher's "Inverse Probability" of 1930. International Statistical Review 68:155-172.

Altenburger, R., T. Backhaus, W. Boedeker, M. Faust, M. Scholze, and L.H. Grimme. 2000. Predictability of the toxicity of multiple chemical mixtures to *Vibrio fischeri*: mixtures composed of similarly acting chemicals. Environmental Toxicology and Chemistry 19:2341-2347.

Backhaus, T., R. Altenburger, W. Boedeker, M. Faust, M. Scholze, and L.H. Grimme. 2000. Predictability of the toxicity of a multiple mixture of dissimilarly acting chemicals to *Vibrio fischeri*. Environmental Toxicology and Chemistry 19:2348-2356.

Lydy, M., J. Belden, C. Wheelock, B. Hammock, and D. Denton. 2004. Challenges in regulating pesticide mixtures. Ecology and Society 9: 1.

Mount, D., Gulley, D., Hockett, J., Garrison, T., and Evans, J., 1997. Statistical models to predict the toxicity of major ions to *Ceriodaphnia dubia*, *Daphnia magna* and *Pimephales promelas* (fathead minnows). Environmental Toxicology and Chemistry 16:2009-2019.

CHAPTER II

CAN SITE-SPECIFIC HEURISTIC TOXICITY MODELS PREDICT THE TOXICITY OF PRODUCED WATER?

Published in: Fisher, J.C., J.B. Belden, and J.R. Bidwell. 2010. Can site-specific heuristic toxicity models predict the toxicity of produced water? Chemosphere 80: 542-547.

Abstract

An empirically derived model of major ion toxicity was combined with other toxicity assessments to account for the observed toxicity in field-collected produced water and produced-water contaminated groundwater. The accuracy and precision of the ion toxicity model, calculated using model deviation ratios (MDR) and simple linear regressions, was determined for fathead minnows, *Ceriodaphnia dubia*, and *Daphnia magna*. Model accuracy for produced water fell within a factor of two for all three organisms. The precision, or variability explained by the model, was 47.9%, 56.1%, and 0.00% for fathead minnows, *C. dubia*, and *D. magna*, respectively. Incorporating other measured potential toxicants improved predictive precision for fathead minnows to 67.0% using ion toxicity and pH and to 30.9% for *D. magna* using ion toxicity, pH, and total ammonia. The observed toxicity to *D. pulex* was also evaluated using *D. magna* model predictions and other measured parameters, but no consistent relationship was found. Dissimilar results were found for produced water contaminated groundwaters with model predictions for *D. magna* falling within a factor of two of and explaining 53.8% of the observed variability in *D. pulex* responses. These results indicate that predicted major ion toxicity, combined with other measured parameters, can accurately and precisely account for observed responses in test organisms to field collected samples.

Keywords

Model deviation ratio, Ion toxicity, fathead minnow, *Ceriodaphnia dubia*, *Daphnia magna*, *Daphnia pulex*

Introduction

Produced water is water that is co-produced during crude oil recovery. It is a complex mixture containing many potential toxicants including major ions (Na, Ca, K, Mg, Cl, SO_4), ammonia, hydrogen sulfide, petroleum hydrocarbons, BTEX (benzene, toluene, ethyl-benzene, and xylenes), phenols, naphthalenes, PAHs, zinc, and other heavy metals (Fucik 1992; Schiff et al. 1992; Kharaka et al. 2005; Benko and Drewes 2008). These constituents vary between and within different geologic basins depending on geology and hydrology (Collins 1985; Daly and Mesing 1995; Benko and Drewes 2008).

During inland oil production, produced water is commonly reinjected back into the oil seam, potentially allowing the contaminants contained in the produced water to enter nearby groundwaters and related surface water. For example, this process has been reported to result in major ion, bromide, arsenic, cadmium and copper contamination in groundwater within the area of injection (Hudak and Blanchard 1997; Sadiq and Alam 1997; Okandan et al. 2001). Organic contaminants, including oil, grease, and aromatic hydrocarbons, also occur in produced waters and can be found in contaminated groundwaters (Kharaka et al. 2005; Benko and Drewes 2008). At the field site for this study, along Skiatook Lake, Oklahoma, excess major ion contamination occurs at a near-surface aquifer and the lake itself (Herkelrath and Kharaka 2002; Kharaka et al. 2005; 2007).

As produced water leaches into surface water, there is potential for toxicity to aquatic organisms, and characterization of this toxicity may help direct management and use of the water. Prediction of the toxicity of produced water to aquatic organisms can be challenging because produced water is a complex mixture that can be composed of different potential toxicants at different sites and that composition may vary both spatially and temporally at the same site. Therefore, an approach to predict the toxicity based on concentrations of the contaminants in a

specific sample of produced water requires flexible modeling approaches that can account for variability in contaminants and concentrations.

Several models exist to predict the effects of toxicant mixtures. Predictive models that only require individual toxicity testing of each contaminant are available for contaminants that generally act similarly (concentration addition, Altenburger et al. 2000) and for contaminants that act independently (independent action, Backhaus et al. 2000). Other models have been developed by empirically testing mixtures of contaminants and establishing regressions to predict toxicity. For example, derived models have been established to estimate the joint toxicity of metals (Di Toro et al. 2001), major ion salts (Mount et al. 1997), and narcotic organics such as PAHs (Swartz et al. 1995).

The model derived by Mount et al. (1997) is an empirically-derived statistical model to predict major or essential ion (Na, K, Ca, Mg, Cl, SO_4, and HCO_3) toxicity to three standardized bioassay organisms (fathead minnows (*Pimphales promelas*), *Ceriodaphnia dubia*, and *Daphnia magna*). Because these ions may account for much of the toxicity of produced water, an ion toxicity model has been suggested as a technique to predict toxicity of produced water. The results of the model can then be used to supplement toxicity identification evaluations (Tietge et al. 1997). In their study, Tetge et al. (1997) used modeled ion toxicity to screen produced waters for toxicity and suggested using modeled toxicity as a screening tool for future produced water studies. Produced waters where observed toxicity was greater than predicted (fell outside 95% confidence intervals) were then subjected to phase I toxicity identification evaluation (TIE, US EPA 1991) procedures to determine other toxic constituents.

The goal of this study was to use the ion toxicity model and other mixture and statistical modeling approaches to assess the toxicity of produced water. The specific objectives were to 1) further evaluate the accuracy and precision of the ion toxicity model for predicting the toxicity of

produced water and produced water-contaminated groundwater specific to a field site at Skiatook Lake, Oklahoma, 2) utilize TIE and regression analysis to determine the relative role of non-ion stressors within the produced water samples, and 3) evaluate the applicability of the toxicity predicted by the ion toxicity model for produced water to species other than those used to develop the model.

Methods

Produced Water

Sampling and Chemical Analyses - Samples were collected at a site along Skiatook Lake, a 4249 hectare reservoir in Osage County, northwest of Tulsa, Oklahoma. Salt scars, characterized by bare soil and salt crystals on exposed rocks, ran to the lake from two evaporation ponds located immediately down slope of a storage tank battery and a produced water reinjection well.

Produced water was collected from onsite storage tanks in one-liter amber bottles with Teflon coated lids, filled from the bottom with a hose to reduce volatile chemical loss and minimize headspace, and held at 4°C. Bimonthly monitoring of produced water occurred from January 2006 to May 2007 and included basic water quality measures (total ammonia, pH, and conductivity). Hydrocarbons were extracted and analyzed (US EPA method 8015B, Accurate Labs Inc., Stillwater, OK) for BTEX (benzene, toluene, ethyl-benzene, and xylenes), gasoline range organics (GRO) and diesel range organics (DRO). Samples were also sent to the Oklahoma Cooperative Extension Service's Soil, Water, and Forage Analytical Laboratory at Oklahoma State University to measure major ions (APHA 1995).

Toxicity Assessment - Static renewal acute (48 h) testing of the produced water samples collected for water analyses were run within 48 hours of collection and in addition to chemical analyses during bimonthly sampling events. Bioassays utilized the fathead minnow, *Pimephales promelas*, *Ceriodaphnia dubia*, *Daphnia magna*, and *D. pulex* with a two-fold serial dilution scheme (US

EPA 2002). During acute 48 h experiments for fathead minnows and *C. dubia*, organisms were fed during the experiment, unlike *D. magna* and *D. pulex* tests. However, feeding during acute toxicity tests did not affect results during development of the ion toxicity model (Mount et al. 1997). Bioassays were also conducted using all four test organisms in duplicate on laboratory major ion salt mixtures that represented the most concentrated (September 2007) and least concentrated (January 2007) produced water samples. The mean MDR's (model deviation ratios) for these laboratory salt mixtures were then used to calibrate (by dividing field MDR's by laboratory derived MDR's) the magnitude of model predictions for field-collected samples.

Toxicity Identification Evaluation (TIE) is a US EPA (1991) standardized series of water manipulations and toxicity tests meant to identify toxic constituents of a complex mixture or effluent. Evaluations of produced water were conducted using *D. pulex* in February 2006, March 2007, and September 2007. These manipulations (listed in Table 1) were used to identify contaminants that may be contributing to the observed toxicity.

Experiments utilizing *D. magna* were performed on raw produced water, produced water with nonpolar components removed (using C18 solid phase extraction, SPE), and in laboratory water spiked with the material retained on the SPE columns in June 2007 (following US EPA 1991). All derived endpoints were converted to represent percent concentration of the original produced water. *D. magna* (instead of *D. pulex* which has no ion toxicity model) was used so that predicted major ion salt toxicity could be coupled with the effects of the eluate and eluant to better assess the role of organic contaminants in the overall toxicity of the produced water.

Groundwater

Sampling and Chemical Analyses - There were 8 shallow (one to three meters deep) groundwater wells at the study site that were used for the characterization of uncontained groundwater. During August of 2006, February and May of 2007, one well volume (the volume of water

contained within and extractable from a well) of water from the eight selected groundwater wells was collected using a hand pump and filtered using Whatman GF/A glass fiber filters to remove sediment. Well water samples were analyzed for the same ion concentrations as those measured in produced water.

Toxicity Assessment - Water quality monitoring and static renewal acute toxicity testing (48 h) with *D. pulex* (US EPA 2002) on selected groundwater wells also occurred during February and May of 2007. *D. pulex* was chosen as the test organism because Oklahoma Department of Environmental Quality regulations (Oklahoma Administrative Code § 252:690-3-29(1)) specify this organism for acute whole effluent toxicity (WET) testing. *D. magna* was used for additional toxicity tests during August 2006 to coincide with major ion analyses. The volume (<500 mL) of sampled water from the groundwater wells was insufficient to measure a variety of potential contaminants or conduct a battery of bioassays on multiple test species as was done for produced water samples.

Statistics and Models

Most statistical analyses were conducted using SAS 9.1 (The SAS Institute, Cary, NC). Simple and multiple linear regressions were conducted using PROC REG with residual diagnostics analyzed using PROC UNIVARIATE, while Pearson correlations were conducted using PROC CORR. Toxicological endpoints such as LC50's and confidence limits were generated using the Comprehensive Environmental Toxicity Information System (CETIS) version 1.1.2 (Tidepool Scientific Software, McKinleyville, CA).

Model Deviation Ratios (MDR) were used to compare ion toxicity model predictions with observed bioassay endpoints. MDRs were calculated by dividing the predicted 48 h LC50 by the observed 48 h LC50 (Belden et al. 2007). Ratios that fell within a factor of two (between 0.5 and 2.0) were considered to show adequate agreement between predicted and observed

values. In cases where predicted values were not available for *D. pulex*, *D. magna* predictions were used as a surrogate. Simple linear regressions were used to supplement MDRs to quantify the variability explained by model predictions. These regressions used the toxicity predicted from the ion toxicity model as the independent variable and observed toxicity as the dependent variable. The use of MDRs and simple linear regressions allowed the simultaneous assessment of model prediction accuracy and precision.

Multiple linear regressions were used to further describe observed toxicity using predicted major ion toxicity, pH, total ammonia, BTEX, GRO, and DRO. Parameters for the final models were selected by first removing collinear variables and then backward selection based on the adjusted r-squared.

Results

Produced Water

Acute 48 h LC50's ranged from 7.44% to 11.2% for fathead minnows, 2.06% to 2.74% for *C. dubia*, 2.68% to 5.36% for *D. magna*, and 0.94% to 4.13% for *D. pulex*. Comparisons between observed 48 h LC50's and those predicted from major ion concentrations differed depending on the test organism, however those differences were similar to those found using laboratory derived salt mixtures (Figure 1). Fathead minnow lethal concentrations were most similar to predicted values with a raw MDR of 1.15, which when calibrated using the salt only tests was 1.69, and a nearly significant positive relationship for the simple linear regression model which plotted toxicity as a function of predicted toxicity (P=0.0849, R^2=0.479). Model Deviation Ratios for *C. dubia* also averaged within a factor of two of that predicted based on major ions alone with a raw value of 1.89 and a calibrated value of 0.72, indicating less toxicity than predicted. The simple linear regression resulted in a significant positive relationship for *C. dubia* (P=0.0325, R^2=0.561). *D. magna* had an average raw ratio of 3.16, but when calibrated was 0.83,

indicating less observed toxicity than predicted and no relationship for the simple linear regression (P=0.987, R^2<0.0001).

Multiple linear regressions were used to further explain variability in observed toxicities based on predicted major ion toxicity, total ammonia, pH, BTEX, GRO, and DRO. For fathead minnows, the relationship between observed toxicity and predicted toxicity was improved by incorporating pH as a second parameter (P=0.0483, adjusted R^2=0.670). However, fathead minnow observed toxicity was best explained by two other factors, pH and total ammonia (P=0.0222, adjusted R^2=0.777). For *C. dubia*, the original simple linear regression model based only on predicted toxicity from major ions had a better fit than any of the multiple regression models. The best fitting model for *D. magna* used toxicity predicted from major ions, pH, and total ammonia; however, the fit was still very poor (P>0.10, adjusted R^2=0.309).

Toxicity identification evaluation of produced water using *D. pulex* resulted in reduced toxicity due to chelation by EDTA, oxidant reduction with sodium thiosulfate, filtration at pH 11, and aeration at pH 11 when compared using multiple logistic regression to the baseline toxicity of unmodified produced water (Table 1). The table includes 95% confidence intervals for baseline toxicity to demonstrate that all treatments fell within baseline intervals and inclusion of treatment intervals was unnecessary. However, the magnitude of the reduced toxicity was minimal for all treatments.

Experiments using *D. magna* and C18 separation of produced water resulted in no observed decrease in toxicity with the removal of nonpolar compounds. The raw produced water elicited a 48 h LC50 of 3.03%, while the C18 filtered water had a 48 h LC50 of 2.87%, and the material eluted from the column had a 48 h LC50 of 15.9%. Toxicity predicted from major ions alone was 11.2% resulting in an MDR of 3.70.

Observed *D. magna* toxicity was the only bioassay endpoint that was significantly correlated to observed *D. pulex* toxicity (P=0.0167, R^2=0.643). Further, *D. pulex* toxicity predicted using the ion toxicity model for *D. magna* was nearly significant, but the relationship was negative (P=0.0517, R^2=0.564, Figure 2). Additionally, the mean MDR for this relationship was 2.03 (when corrected using the laboratory salt experiments), while the MDR using observed *D. magna* toxicity to predict observed *D. pulex* toxicity was 2.01 with a significant regression (P=0.0353, R^2=0.621, Figure 2).

Groundwater

Groundwater toxicity to *D. pulex* varied depending on well, with 48 h LC50's ranging from 10.4% to greater than 100%. Toxicity to *D. magna* during August 2006 was closely related to predicted major ion toxicity and predicted toxicity explained 61.0% of the variability in observed toxicity using simple linear regression (Figure 3). Raw Model Deviation Ratios for all wells fell between 0.82 and 1.92, but ranged from 0.21 to 0.50 when calibrated with laboratory salt experiments.

The comparisons between *D. magna* predicted ion toxicity and observed *D. pulex* toxicity yielded corrected MDRs averaging 0.54, exhibited a significant positive relationship (P=0.0012, R^2=0.538), and 87.5% of observations (N=16) had MDRs that fell between 0.3 and 2.0 (Figure 3).

Discussion

Model Accuracy and Precision

In this study, predicted major ion toxicity alone was representative of observed toxicity for both fathead minnows and *C. dubia*. Model Deviation Ratios, indicative of effect magnitude similarity between predicted and observed or accuracy, fell within a factor of 2 for three species. This degree of deviation is similar to that found for other mixture models. For example, Belden

et al. (2007) found that for pesticide mixture experiments where concentration addition was expected, greater than 90% had MDR's within a factor of 2. This degree of deviation is also in the same magnitude of the uncertainty usually associated with toxicity testing. The U.S. EPA (2002) reported the coefficient of variation for intra-laboratory and inter-laboratory bioassays ranged from 8% to 41% and 13% to 38.5% respectively, or within a factor of 2. Simple linear regressions, indicating model precision, revealed that for fathead minnows and *C. dubia*, predicted major ion toxicity explained 47. 9% and 56.1% of the variability in observed toxicity, respectively. These two techniques clearly indicate that major ion salts are the primary source of toxicity observed in these two test species.

Observed toxicity in *D. magna* fit poorly with any of the variables measured. After calibrating these experiments to the ion toxicity model using laboratory made mixtures, the response magnitude fell within a factor of two (0.83) of predicted values. Correcting the magnitude of the model predictions did not affect the precision of the predictive model and the low explained variability indicates that for produced water samples from the study site the toxicity model had poor precision. Instead, additional toxicants or other factors are likely responsible. However, groundwater testing with *D. magna* demonstrated a close relationship between predicted and observed toxicity. Both the raw magnitude (MDR within a factor of two) and the variability (61.0% of variability explained) indicate both accuracy and precision for the model when applied to site groundwaters. This may suggest that during the transport of the produced water to groundwater, secondary toxicants are dissipated leaving the major ions as the primary source of toxicity.

Toxic contributions of other components

The observed toxicity 95% confidence intervals for fathead minnows, *C. dubia*, and *D. magna* for four out of six produced water samples subjected to TIE procedures by Tietge et al.

(1997) contained the values predicted by major ions alone. Further, in the two samples with toxicity greater than predicted by major ions, observed toxicity averaged within 20% of model predictions after treatment with solid phase extraction to remove nonpolar compounds. Produced waters from oil production along Skiatook Lake (Khararka et al. 2005) and the nearby Anadarko Basin (Benko and Drewes 2008) contained predominantly sodium chloride salts with total dissolved solids levels exceeding 130,000 mg L^{-1}. Neither study reported nonessential metals at concentrations likely to cause toxicity to test organisms, nor did our TIE indicate toxicants in addition to what was reported here. Toxicity derived from other metals is possible based on other produced water studies (Sadiq and Alam 1997), however the amount of toxicity reduced using EDTA chelation for our TIEs was the maximum toxicity reduction before EDTA toxicity became evident. This low toxicity reduction compared with the high amount of chelating agent was likely an artifact of high concentrations of major metal ions.

Multiple linear regressions indicated that fathead minnow toxicity could also be correlated using pH and total ammonia concentration by explaining 77.7% of the observed toxicity variability. The approach using pH and total ammonia, however, was not applicable to the magnitude of toxicity observed because even though produced water samples averaged 49.7 mg L^{-1} total ammonia, pH during toxicity tests at the highest tested concentrations (12.5% produced water) averaged 7.03. At this pH, 0.49% or 0.24 mg L^{-1} of the ammonia was in the unionized form, before correcting for dilution. The corrected unionized ammonia concentration was almost 50x less than the reported 96h LC50 for fathead minnows at 1.50 mg L^{-1} (Mayes et al. 1986). Based on the relatively low concentrations of the unionized ammonia compared to the concentrations of excess major ions, the other multiple regression, using predicted ion toxicity and pH, which explained 67.0% of the observed variability, better represented the observed bioassay responses.

Despite a significant multiple linear regression predicting *D. magna* toxicity using predicted major ion toxicity, pH, and total ammonia, 70% of the variability was left unexplained indicating that other unmeasured variables could account for observed effects. The additional experiments using C18 separation indicated that toxicity was primarily related to polar constituents that would not be retained by C18.

Due to the high accuracy (MDRs within a factor of two) and precision (53.8% of variability explained using regression) for the ion toxicity model on site groundwaters, toxicity to *D. magna* was primarily derived from excess major ions. Furthermore, the unknown factors found in the produced water apparently did not persist in the groundwater at levels that would result in a measureable effect.

Model Applicability to D. pulex

Although produced water toxicity observed in *D. pulex* was closely related to observed *D. magna* effects (Figure 2) the *D. magna* ion toxicity model did not precisely predict observed produced water toxicity in *D. pulex.* Based on that evidence, the *D. magna* ion toxicity model should not be used in future produced water assessments in lieu of *D. pulex* bioassays. In contrast, the *D. magna* ion toxicity model described the observed toxicity of groundwater to *D. pulex* much more accurately both in trend (Figure 3; $R^2 = 0.54$) and in magnitude (MDRs between 1 and 4). The apparent discrepancy between produced water and groundwater results may indicate differences in mixture composition, i.e. additional produced water toxicants, and suggest that the *D. magna* ion toxicity model may have usefulness for *D. pulex* after further investigation into the effects of other produced water constituents on major ion toxicity.

Applicability of Research

Both this study and previous laboratory studies of produced water effects indicated greater sensitivity of *C. dubia* to produced water in general, and particularly major ions, when

compared to fathead minnows (Boelter et al. 1992; Fucik 1992). Laboratory studies were also shown to be valid measures of stream quality when compared with benthic macroinvertebrate communities in streams receiving produced water (Mount et al. 1992). Field studies of produced water receiving streams resulted in findings ranging from few significant effects (O'Neil 1992), to increases in salt tolerant diatoms (Olive et al. 1992) and loss of mayfly larvae (Ephemeroptera) at salinities of 2 ‰, and loss of Dipteran larvae and fishes at higher (10 ‰) salinities (Short et al. 1991).

The methods employed in this study can add a meaningful and low effort step in the site assessment process. Combined use of existing predictive models with site specific contaminant data not accounted for in the original model allows both the screening of multiple sites for potential multiple organism effects and the ranking of those potential effects. This combined use gives immediate and quantitative biological relevance to contaminant measurements during field studies of produced water effects.

Conclusions

The existing model for major ion toxicity to fathead minnows and *C. dubia* was representative of observed responses and accounted for both the magnitude and variability observed in bioassays for these two organisms. Predicted ion toxicity to *D. magna* was only representative of observed *D. magna* in groundwater samples and not produced water samples, possibly because of other toxic constituents in the produced water samples. Though the specific models developed here only apply to the study site, this study does support the limited use of heuristic toxicity models that incorporate the ion toxicity model at other produced water contaminated sites. Based on the relationships found at the studied site, this study also further supports the future use of major ion toxicity alone to predict potential produced water effluent toxicity to fathead minnows and *C. dubia*. However, *D. magna* predictions are less likely to be

useful and should be limited to ranking the potential toxicity of multiple samples due to the loss of predictive precision when other unknown factors were present in the produced water. The similarities between the responses of *D. magna* and *D. pulex* allowed a model derived for one organism to act as a relative surrogate for an assessment of groundwater using the other organism; however the extension was not valid for produced water. By extending the ion toxicity model to other species and utilizing other statistical techniques to evaluate the significance of other non-major ion contaminants, this study adds to our ability to estimate the toxicity of field collected oil-field produced water and produced water contaminated groundwater.

Acknowledgements

We gratefully acknowledge the assistance of N. Cooper, M. Abbott, C. Boeckman, and K. Burgess for assistance with data collection. This research was funded through a cooperative agreement with the US Department of Energy (DE-FC26-04NT15544).

Literature Cited

Altenburger, R., Backhaus, T., Boedeker, W., Faust, M., Scholze, M., and Grimme, L., 2000. Predictability of the toxicity of multiple chemical mixtures to *Vibrio fischeri*: mixtures composed of similarly acting chemicals. Environ. Toxicol. Chem. 19, 2341-2347.

APHA, 1995. Standard Methods, 19th Edition. American Public Health Association, Washington, DC.

Backhaus, T., Altenburger, R., Boedeker, W., Faust, M., Scholze, M., and Grimme, L., 2000. Predictability of the toxicity of a multiple mixture of dissimilarly acting chemicals to *Vibrio fischeri*. Environ. Toxicol. Chem. 19, 2348-2356.

Belden, J., Gilliom, R., and Lydy, M., 2007. How well can we predict the toxicity of pesticide mixtures to aquatic life? Integrated Environmental Assessment and Management 3, 364-372.

Benko, K. and Drewes, J., 2008. Produced water in the western United States: geographical distribution, occurrence, and composition. Environ. Eng. Sci. 25, 239-246.

Boelter, A., Lamming, F., Farag, A., and Bergman, H., 1992. Environmental effects of saline oil-field discharges on surface waters. Environ. Toxicol. Chem. 11, 1187-1195.

Collins, A., 1985. Geochemistry of Oilfield Waters: Developments in Petroleum Science. Elsevier Scientific Publishing Company, New York.

Daly, D., and Mesing, G., 1995. Atlas of Gas Related Produced Water for 1990. Gas Research Institute Topical Report 95/0016, Chicago.

Di Toro, D., Allen, H., Bergman, H., Meyer, J., Paquin, P., and Santore, R., 2001. Biotic ligand model of the acute toxicity of metals. 1. Technical basis. Environ. Toxicol. Chem. 20, 2383-2396.

Fucik, K., 1992. Toxicity identification evaluation and characteristics of produced water discharges from Colorado and Wyoming. In: Ray, J., Engelhart, F. (Eds.). Produced Water. Plenum Press, New York, pp. 187-198.

Herkelrath, W.N. and Y.K. Kharaka. 2002. Hydrologic controls on the subsurface transport of oil-field brine at the Osage-Skiatook Petroleum Research "B" Site, Oklahoma. Proc. Annu. Int. Petrol. Environ. Conf., Integr. Petrol. Environ. Consortium. 9.

Hudak, P., and Blanchard S., 1997. Land use and groundwater quality in the Trinity group outcrop of north-central Texas, USA. Environ. Int. 23, 507-517.

Kharaka, Y., Kakouros, E., Thordsen, J., Ambats, G., and Abbott, M., 2007. Fate and groundwater impacts of produced water releases at OSPER "B" site, Osage County, Oklahoma. Appl. Geochem. 22, 2164-2176.

Kharaka, Y., Thordsen, J., Kakouros, E., and Herkelrath, W., 2005. Impacts of petroleum production on ground and surface waters: results from the Osage-Skiatook petroleum environmental research A site, Osage county, Oklahoma. Environ. Geosci. 12, 127-138.

Mayes, M., Alexander, H., and Hopkins, D., 1986. Acute and chronic toxicity of ammonia to freshwater fish: a site-specific study. Environ. Toxicol. Chem. 5, 437-442.

Mount, D., Drottar, K., Gulley, D., Fillo, J., and O'Neil, P., 1992. Use of laboratory toxicity data for evaluating the environmental acceptability of produced water discharge to surface waters. In: Ray, J., Engelhart, F. (Eds.). Produced Water. Plenum Press, New York, pp. 175-185.

Mount, D., Gulley, D., Hockett, J., Garrison, T., and Evans, J., 1997. Statistical models to predict the toxicity of major ions to *Ceriodaphnia dubia*, *Daphnia magna* and *Pimephales promelas* (fathead minnows). Environ. Toxicol. Chem. 16, 2009-2019.

Okandan, E., Gumrah, F., and Demiral B., 2001. Pollution of an aquifer by produced oil field water. Energ. Source. 23, 327-336.

Olive, J., Jackson, J., Keller, D., and Wetzel, P., 1992. Effects of oil-field brines on biological integrity of 2 tributaries of the Little Muskingum River, southeastern Ohio. Ohio J. Sci. 92, 139-146.

O'Neil, P., Harris, S., Mettee, M., Isaacson, H., and Evans, J., 1992. Biological fate and effect of coalbed methane produced waters discharged into streams of the Warrior Basin,

Alabama. In: Ray, J., Engelhart, F. (Eds.). Produced Water. Plenum Press, New York, pp. 315-327.

Sadiq, M., and Alam, I., 1997. Metal concentrations in a shallow groundwater aquifer underneath petrochemical complex. Water Resour. 31, 3089-3097.

Schiff, K., Reish, D., Anderson, J., and Bay, S., 1992. A comparative evaluation of produced water toxicity. In: Ray, J., Engelhart, F. (Eds.). Produced Water. Plenum Press, New York, pp. 199-207.

Short, T., Black, J., and Birge, W., 1991. Ecology of a saline stream: community responses to spatial gradients of environmental conditions. Hydrobiolgia 226, 167-178.

Swartz, R., Schults, D., Ozretich, R., Lamberson, J., Cole, F., DeWitt, T., Redmond, M., and Ferraro, S., 1995. ΣPAH: A model to predict the toxicity of polynuclear aromatic hydrocarbon mixtures in field-collected sediments. Environ. Toxicol. Chem. 14, 1977-1987.

Tietge, J., Hockett, J., and Evans, J., 1997. Major ion toxicity of six produced waters to three freshwater species: application of ion toxicity models and TIE procedures. Environ. Toxicol. Chem. 16, 2002-2008.

U.S. EPA, 1991. Methods for aquatic toxicity identification evaluations: Phase I toxicity identification procedures, 2nd ed. EPA/600/6-91/003, Washington, DC.

U.S. EPA, 2002. Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms. EPA/821/R-02/012, Washington, DC.

Table 1. Effects of TIE manipulations on produced water toxicity to *D. pulex*. Values are 48 h LC50's for each manipulation. Significant differences (α=0.05) from baseline toxicity are indicated in bold.

	% LC50 (95% CI)		
TIE Manipulation	Feb 2006	Mar 2007	Sep 2007
Baseline	2.73 (1.17, 7.48)	0.66 (0.0099, 13.62)	1.99 (1.92, 2.05)
pH 6	NA	0.95	NA
pH 7	2.93	0.51	2.01
pH 8	3.39	0.78	1.95
EDTA	**3.76**	0.82	**2.99**
Oxidant Reduction	3.41	1.38	**2.99**
pH 3	2.07	NA	2.25
pH 11	2.90	0.51	1.98
Filtration	3.51	0.76	**1.50**
pH 3 filtration	1.56	NA	**1.50**
pH 11 filtration	**3.77**	1.34	1.95
Aeration	2.56	0.66	2.05
pH 3 aeration	2.81	NA	2.01
pH 11 aeration	**4.07**	0.52	**2.99**
SPE	2.35	0.66	2.05
pH 3 SPE	2.07	NA	**1.50**
pH 9 SPE	3.22	2.99	**2.99**

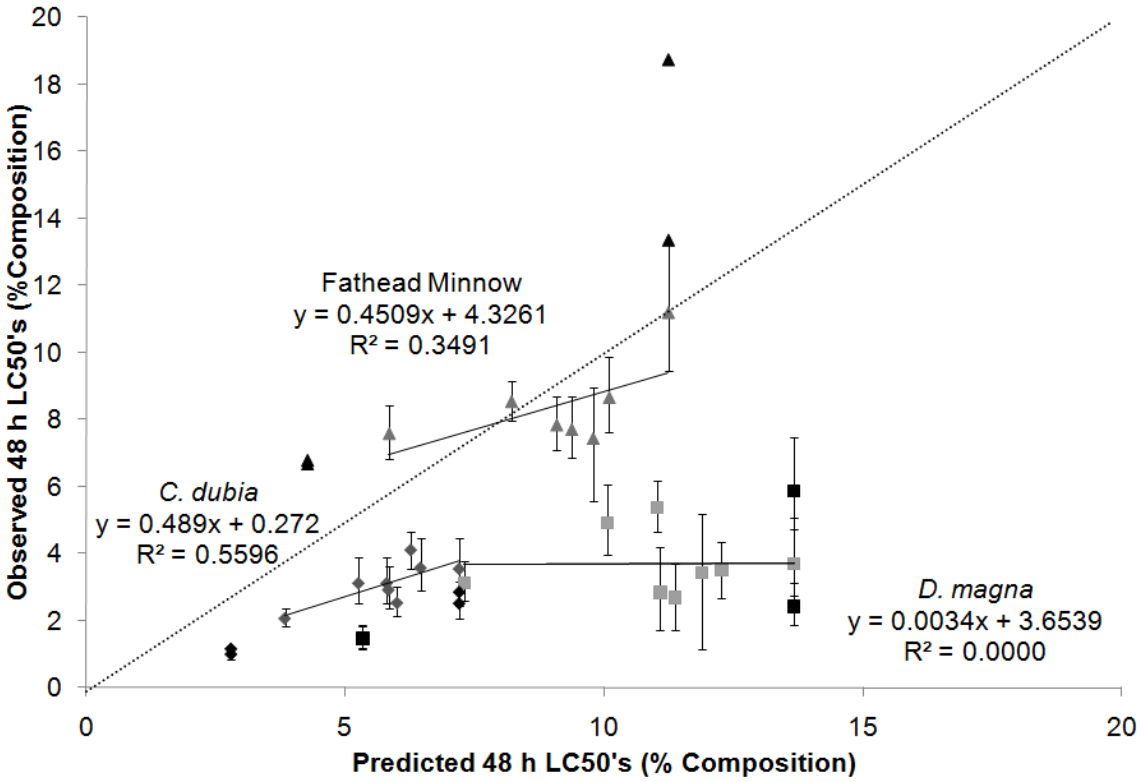

Figure 1. Comparison of LC50's measured from standardized toxicity tests and LC50's predicted based solely on the ion toxicity model. Error bars are 95% confidence intervals. Test organisms plotted are *D. magna* (■), *C. dubia* (◆), and fathead minnows (▲). Grey points are endpoints of field samples, whereas solid black points are from laboratory salt mixtures. The dotted line has a slope equal to one and represents agreement between observed and predicted values. Values that fall below the dotted line are bioassays where toxicity was observed to be greater than that predicted by the model. Values above the line exhibited toxicity less than predictions. Simple linear regressions of model toxicity predicting observed toxicity are shown for each species.

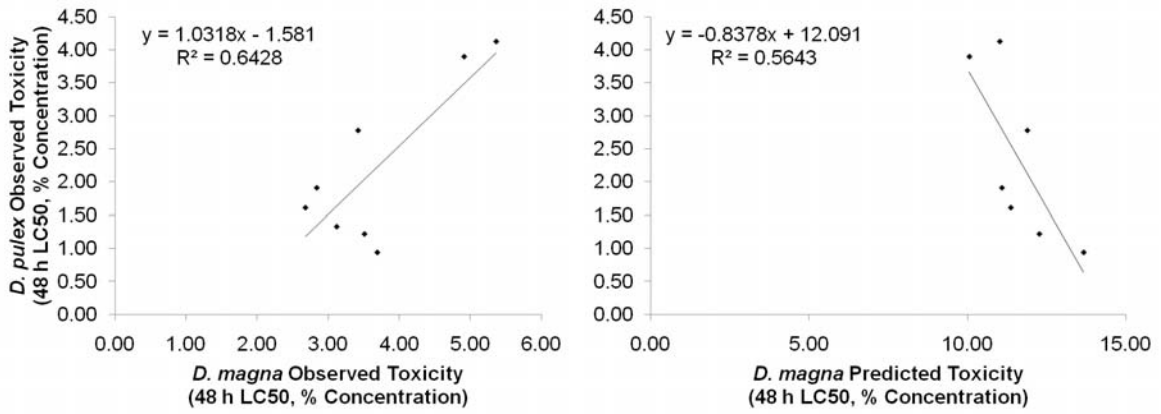

Figure 2. Simple linear regressions showing relationship between *D. magna* and *D. pulex* observed produced water toxicity (left) and between *D. pulex* observed and *D. magna* predicted ion toxicity (right).

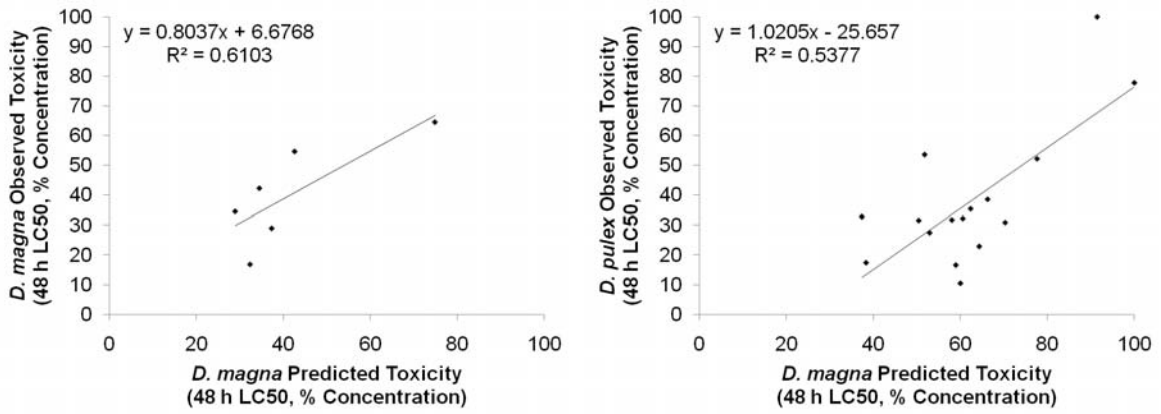

Figure 3. Major ion toxicity contribution to groundwater well toxicity to *D. magna* and *D. pulex*. Simple linear regressions showing relationship between *D. magna* observed and predicted groundwater toxicity (left) and between *D. pulex* observed and *D. magna* predicted ion toxicity (right).

CHAPTER III

LABORATORY AND FIELD MEASURED EFFECTS OF PRODUCED WATER INFILTRATION OF A LAKE BENTHOS

Abstract

Petroleum production is vital to the economy of the United States, though the environmental costs of this production are poorly documented. Produced water is a byproduct of that production and contains many potential toxicants, the most dominant of which are major ion salts. This study examined the potential for and measured aquatic effects of a near-shore petroleum operation along Skiatook Lake, Osage County, Oklahoma, USA. Produced water contamination was found at levels, based on major ion concentrations, which would potentially cause toxicity to standardized toxicity bioassay organisms in both site groundwaters and lake sediment porewaters. Laboratory bioassays utilizing the midge *Chironomus dilutus* and the amphipod *Hyallela azteca* did not demonstrate any contaminant effects of impact site sediments compared with a reference site. However, *in situ* experiments using *H. azteca* and the Asian clam *Corbicula fluminea* resulted in significantly decreased growth for organisms placed at the impact site compared with a reference site. Macroinvertebrate community composition significantly differed between sites, but that difference was confounded by the fact that the reference community exhibited lower diversity and increased overall tolerance values to pollution than the impact site. This study showed that produced water from a near-shore petroleum production operation travelling through the groundwater into the lake was sufficiently diluted by the lake to result in few observable biotic effects. Additionally, *in situ* experiments were the most sensitive to the produced water intrusion without the interferences of other site-specific effects found in macroinvertebrate community samples.

Introduction

Petroleum production is vital to the economy of the United States, though the environmental costs of this production are poorly documented. Produced waters vary widely across the U.S. , but major ions, especially chloride, represent the majority of contaminants of concern when placed in the context of relevant NPDES (National Pollution Discharge Elimination System) discharge limits and toxicity to sensitive aquatic organisms (Benko and Drewes 2008; Johnson et al. 2008). During the decade from 1993 to 2003, 12,863 fluid releases related to oil exploration and production were reported to the Oklahoma Corporation Commission. Of those releases, 76.1% involved oil or brine (produced water) and had median release volumes of 10 bbl (barrels; 1590 l) and 40 bbl (6359 l) respectively, for a total of 1.46 million barrels (232 million liters) of brine released in Oklahoma (Fisher and Sublette 2005).

These produced water releases affect the landscape for decades after production ends. For example, Otton et al. (2005) reported persistent, unvegetated salt scars at a site along Skiatook Lake, Oklahoma where oil production ceased in 1981. Sodium chloride concentrations at this site were elevated to a depth of at least eight meters into the bedrock. Blackjack oaks, *Quercus marilandica*, growing next to the salt scar exhibited elevated leaf chloride levels, indicative of salt contamination. A large groundwater plume about 3 ha in area contaminated with excess salts (electrical conductance of 44,000 μS/cm or 30,000 mg/L total dissolved solids) and soluble organic petroleum by-products that extended under the lakebed were also found beneath this site (Kharaka et al. 2005).

Several methods can be used to assess site sediment quality including laboratory toxicity bioassays of produced and groundwaters, *in-vitro* and *in-situ* sediment toxicity assessments, and macroinvertebrate community sampling. Laboratory experiments using standardized protocols enhance repeatability and reproducibility, but may lack environmental relevance due to strict

experimental controls. *In situ* studies incorporate experimental controls into field conditions to better represent site conditions while retaining some of the advantages of laboratory experiments. Unfortunately, *in situ* studies do not always mirror field or laboratory study results, confounding applicability. The most environmentally relevant method to measure disturbance effects on resident communities is to compare the community structure of potentially impacted communities with unaffected communities, though this approach is time consuming and subject to environmental variability that may overwhelm treatment effects. Sediment-dwelling or benthic macroinvertebrates are ideal study organisms because they are ubiquitous, diverse, relatively immotile (compared with freshwater vertebrates), and have lifecycles that, depending on species, can be from one month to multiple years (Rosenberg and Resh 1993).

Produced water discharges have been shown to affect standardized test organisms (Tietge et al. 1997) as well as resident macroinvertebrate communities (Short et al. 1991; Mount et al. 1992; Olive et al. 1992). For example, Olive et al. (1992) reported a lower percent Trichoptera and decreased invertebrate density in a stream impacted by brine discharges. Previous studies of the environmental effects of produced water on benthic communities have been limited to lotic systems. However, macroinvertebrates are still useful tools in biomonitoring for lentic systems. The difference between lotic and lentic systems necessitates different sampling devices and implies a change in overall community composition, but the ecological theory behind the application remains unchanged.

Laboratory toxicity tests were shown to be valid measures of stream quality when compared with benthic macroinvertebrate samples in Alabama streams receiving produced water (Mount et al. 1992). Despite these findings, laboratory toxicity testing may overestimate produced water effects because dilutions used for such tests may represent effluent concentrations only present at the immediate discharge point (Douglas and Veil 1996). Specifically, the freshwater cladoceran, *Ceriodaphnia dubia*, was more sensitive to produced water than the

fathead minnow, *Pimephales promelas* (Boelter et al. 1992; Fucik 1992). Additionally, Boelter et al. (1992) tested water from streams receiving produced water in Wyoming and found that *C. dubia* was sensitive to major inorganic ions present in the discharge while fathead minnows did not exhibit a response.

Macroinvertebrates are the most commonly used group of organisms for biomonitoring programs (Hellawell 1986). The following reasons were suggested by Rosenberg and Resh (1993) as to why macroinvertebrates are so commonly employed: macroinvertebrates are found in virtually all freshwater habitats, there is great species diversity within this group, they allow relatively easy quantification of local impacts due to their sedentary nature, and they have long enough life cycles for analyses of temporal changes to be assessed. Currently, 49 of 50 US states employ benthic macroinvertebrate surveys as a part of routine biomonitoring activities (Carter et al. 2006). Benthic macroinvertebrate monitoring has been used to assess the impact of short-term saline inputs on stream communities (Marshall and Bailey 2004) and to determine the effects of metals (cadmium, copper, lead, and zinc) along salinity gradients (Peeters et al. 2000).

The goal of this project was to examine the aquatic effects of a near-shore petroleum operation. The specific objectives were to: 1) evaluate if significant produced water contamination is present in sediments, 2) determine if groundwater, porewater, or sediment is contaminated enough to elicit a biological effect based on toxicity modeling and laboratory bioassays, and 3) compare the results of predictive models and laboratory bioassays with field experiments and benthic community assessments.

Methods

Site Description

This study was conducted at Skiatook Lake, a 4249 hectare reservoir in Osage County, northwest of Tulsa, Oklahoma. The impact and reference sites were located along the southern

side of the lake approximately 2 km apart. The impact site was characterized by active petroleum production with two produced water evaporation ponds, one on the southern end of the site and one submerged in Skiatook Lake at the eastern terminus of transect two, and a produced water reinjection well on the north side of the site (Figure 1). Salt scars, characterized by bare soil and salt crystals on exposed rocks, run from the exposed evaporation pond and from the injection well to the lake. As part of an on-going study at the time, the United States Geological Survey (USGS) drilled 29 shallow (one to three meters deep) groundwater wells at the impact site that allowed for characterization of uncontained groundwater and assessment of produced water intrusion at the site (Kharaka et al. 2005). Throughout the manuscript, the wells are referred to by the USGS naming system used on site. Six 33m transects were established at the impact and reference sites. These transects were placed in line with and downslope of the produced water sources on land. Three transects running west to east (originating from groundwater wells BE-07 and BE-62, and approximately 8 meters north of the BE-07 transect) and one running north to south (originating from groundwater well BE-62) were established at the impact site with two transects running north to south at the reference site. The transects averaged one meter in depth at the beginning or near shore end and ranged between three (transects one, four, and reference transects) and four (transects two and three) meters in depth when the lake was at conservation pool level (the pool elevation management goal set by the US Army Corps of Engineers, Figure 2).

Sampling and Chemical Analyses

During February and May of 2007, one well volume (the volume of water contained within and extractable from a well, approximately 500ml) from every groundwater well was collected using a hand pump and filtered using Whatman (Piscataway, NJ, USA) GF/A glass fiber filters to remove sediment. Produced water was collected from onsite storage tanks. Lake, well water, and produced water samples were analyzed for major ions at the Oklahoma Cooperative

Extension Service's Soil, Water, and Forage Analytical Laboratory at Oklahoma State University (APHA 1995). Major ion concentrations were then used in a principle component analysis (PCA) to show relationships between produced water, groundwater samples, and Skiatook Lake. Major Ion measurements were then used to generate surfaces (Arc GIS, Version 9.1, ESRI, Redlands, CA) using Universal Kriging showing predicted major ion toxicity to *Ceriodaphia dubia* and fathead minnows derived from empirical models of major ion toxicity (Figure 4, Mount et al., 1997).

Three aquatic sediment cores (30cm long by 1 cm wide) were taken using a hand corer driven into the sediment at the away from shore terminus of each transect at the two study sites during field sampling to a depth of 30 cm. Samples were transported on ice from the field and frozen until analysis. Porewater was extracted from 5-cm sections of the cores via centrifugation at 4100 rpm for 30 minutes at 23°C. Extracted pore water was then analyzed for major ions by the Oklahoma Cooperative Extension Service's Soil, Water, and Forage Analytical Laboratory at Oklahoma State University, as was done with groundwater samples. The analytical laboratory measured as many ions as feasible (depending on sample volume and ion content) starting with chloride, then measuring cations, and then anions. Ion concentrations were used to predict potential porewater toxicity to three standardized bioassay organisms: *Daphnia magna*, C. *dubia*, and fathead minnows using an ion toxicity model (Mount et al., 1997; Fisher et al. 2010). The porewaters collected were of insufficient volume to allow quantification of bicarbonate concentrations and these concentrations were conservatively set to ambient lake levels, unless sample pH was low (pH < 3.5) enough to limit bicarbonate concentrations below lake levels and bicarbonate was set to that pH limited maximum. If bicarbonate concentrations were greater than the assigned value, then predicted toxicity would increase. During a previous study of the current study site Zielinski et al. (2007) reported a “hot spot” of conductivity that was then

chemically characterized. Porewater ion profiles from this previous study have been included with the current porewater analyses.

Bioassays and Field Experiments

Static renewal acute toxicity testing (48hr) with *D. pulex* (USEPA, 2002) was conducted bimonthly from September 2005 to May 2007 on sampled groundwater wells. Wells were selected to capture gradients in groundwater quality from potential produced water sources (evaporation pond and injection well) to the lakeshore. The wells selected were BE-53, BE-18, BE-11, BE-08, BE-07, BE-59, BE-61, and BE-16 (Figure 1). The first five listed wells were located within visible salt scars that ran from a produced water source to the lake. Wells BE-07 and BE-16 were located in less than one meter of water when the lake was at conservation pool levels. The latter three wells were located within a vegetated area.

Sediment for laboratory bioassays was collected from the top 5 cm of the lake bottom using plastic sample containers and placed on ice for transport back to the laboratory where they were held at 4°C until testing. Samples were collected from the reference site, the middle of transect 2, and in the submerged evaporation pond or brine pit at the end of transect 2.

In situ toxicity tests with the midge *Chironomus dilutus*, the amphipod *Hyallela azteca* (US EPA 2000), and the Asian Clam *Corbicula fluminea* coincided with seasonal field sampling. Midges and amphipods were collected from existing laboratory cultures at the Ecotoxicology and Water Quality Research Laboratory, Oklahoma State University, while clams were collected from a local stream (tail waters of Canton Lake, Blain County, Oklahoma, USA).

Midge and amphipod *in situ* exposures were conducted in cylindrical, clear plastic, exposure chambers with mesh vents along two sides (12.5 cm long by 7 cm wide, following Burton et al. 2005), whereas clams were exposed in orange polyethylene mesh bags. Both tests utilized ten organisms per experimental unit (an exposure chamber or mesh bag) and were

conducted at the same three locationswhere sediment was collected for laboratory toxicity tests. For the amphipod test, there were also three position treatments of four enclosures per treatment at each location: one in which the exposure chamber contained the top 2 cm of surficial sediment, one in which the exposure chamber did not contain sediment but was laying on the sediment with one vent directly in contact with the sediment, and one in which the exposure chambers were suspended in the water column using a galvanized steel basket supported above the sediment. Burton et al. (2005) also suggested burying exposure chambers in sediment as a way to conduct *in situ* exposures with amphipods. However, this technique was attempted and abandoned due to poor test organism survival at all locations due to anoxia. Midge and Amphipod enclosures were harvested after ten-days at which time survival and mean growth per chamber were determined. Growth endpoints were mean ash free dry weight for midges and individual length for amphipods. These experiments coincided with benthic macroinvertebrate community sampling in May (high water) and September (low water) each year. Amphipod experiments were conducted on August 14, 2007, September 9, 2007, and October 6, 2007.

Laboratory sediment toxicity tests with amphipods were conducted in addition to *in situ* bioassays (USEPA, 2000). Briefly, these experiments utilized 100ml of wet sediment in a 600ml beaker for each experimental unit, with ten organisms per beaker and four replicate units per location treatment. Overlying water was exchanged daily and organisms were fed at the same time. These experiments lasted 10d and utilized the same endpoints as the *in situ* experiments. Sediment tests occurred within eight weeks of sample collection (USEPA, 2000). Sediment samples were taken from the same locations as the field experiments to facilitate comparisons between laboratory and field experiments.

Asian Clams were individually marked (8 x 4 mm type FPN shellfish tags, Hallprint Pty Ltd., Victor Harbor, South Australia) such that each clam from each experimental unit was measured before and after incubation so that there were ten replicate organisms per experimental

unit. Marked clams of approximately two centimeters in length measured using a digital caliper (to the nearest 0.005 mm) along the longest shell dimension (anterior to posterior) were placed into orange mesh bags (experimental unit) and then secured on the lake bottom with marking flags. Two mesh bag experimental units were placed at each location used for the previously described *in situ* experiments and incubated for thirty days after which growth and survival endpoints were measured (Soucek et al. 2001). Asian clam experiments were conducted on August 17, 2005, May 16, 2006, October 5, 2006, and September 8, 2007.

Hester-Dendy (Florence, KY, USA) square eight-plate samplers were employed to assess benthic community response to field conditions. Four multi-plate samplers were deployed 0.5m apart at the terminus of each transect plus four more in the submerged evaporation pond for six weeks, twice a year, during historic high and low water levels from 2005 through 2007, though sampling depths along each transect were within 1.5 meters of each other. Samplers were harvested by placing samplers in individual plastic bags with as little disturbance as possible, transporting them to the surface, and preserving all macroinvertebrates in 70% ethanol. Macroinvertebrate samples were sorted and identified to the lowest practical taxon (Merrit and Cummins 1996; Smith 2001). Macroinvertebrate community composition, richness, and Shannon-Wiener Diversity (Hauer and Resh 2006) were compared between sampling locations.

Statistical Models and Analyses

This paper utilized an empirically-derived statistical model (Mount et al., 1997) to predict major or essential ion (Na, K, Ca, Mg, Cl, SO_4, and HCO_3) toxicity to three standardized bioassay organisms (*Pimphales promelas*, *Ceriodaphnia dubia*, and *Daphnia magna*). Because these ions may account for much of the toxicity of produced water, the Ion Toxicity Model has been used to supplement toxicity identification evaluations (Tietge et al. 1997) and as a technique to predict produced water toxicity (Fisher et al. 2010).

Principle Component Analysis was performed using CANOCO 4.5 (Microcomputer Power, Ithaca, NY). The principle components analysis created a linear combination of all measured water variables for each sample. These linear combinations were then used to create new axes on which the samples were plotted. Porewater major ion profiles were plotted onto a Piper diagram using GW_Chart 1.21.0.0 (Winston 2000). All other statistical analyses were conducted using SAS 9.1 (PROC GLIMMIX, The SAS Institute, Cary, NC). Midge and amphipod data were analyzed using analysis of variance (ANOVA) with Fisher's least squared means post-hoc tests to test for treatment differences. Asian clam growth data were analyzed using analysis of covariance (ANCOVA) with initial organism size as the covariate. Macroinvertebrate community abundance data were natural log transformed and analyzed using a multivariate ANOVA to explore differences in sample location and sample season. Only taxa that were found in at least 10% of samples were included in the community analyses. Residual normality for all ANOVA type analyses was checked using ODS graphics residual diagnostics.

Results

Produced Water Contamination

Data from groundwater wells, a produced water sample, and a lake water sample were plotted using ion concentrations along the first two PCA axes and these axes explained 89.5% of the variation in ion profiles (Figure 3). Most ions were positively correlated with the produced water sample and negatively correlated with the lake sample. However, sulfate was uncorrelated with produced water and negatively correlated with the lake, and bicarbonate was uncorrelated with lake samples and negatively correlated with the produced water.

Porewater samples from the impacted site increased in total dissolved solids with sediment depth, while the reference site samples remained constant across depths and contained more than an order of magnitude lower solids concentrations than the impacted site (Figure 4).

Additionally, the reference site samples contained a lower relative concentration of chloride ions than lake samples by containing a lower relative concentration of chloride ions (Figure 5). Impact site samples, had greater relative concentrations of sodium, potassium, and chloride ion than the lake and reference samples, although these trends lessened with decreasing sediment depth.

Potential Biological Effects

Predicted toxicity surfaces for impacted site groundwater were similar for the two species modeled (Figure 6). These surfaces indicate that the theoretical toxicity of the groundwater to bioassay organisms would be greatest at two locations, downslope of the two produced water sources (injection well on the north side and evaporation pond on the south side. Furthermore, these plumes of predicted toxicity were of similar maximum intensity, but the injection well plume was of greater area.

Two locations at the impact site, transect three and the submerged evaporation pond, had major ion concentrations sufficient to result in toxicity to bioassay organisms (Figure 7). Predicted toxicity decreased with sediment depth for all organisms, though *C. dubia* was predicted to show the greatest negative effects. No toxicity was predicted for the first five centimeters of sediment depth, only *C. dubia* was predicted to experience toxicity in the second five centimeters, and all organisms were predicted to show effects after 10 cm of sediment.

Laboratory Bioassays

Bimonthly monitoring of groundwater acute (48 hr) toxicity to *D. pulex* was greatest near the brine injection well (BE-53) and lowest in a well adjacent to the evaporation pond in a vegetated area (BE-59). Toxicity (48 hr LC50 expressed as percent original groundwater) at the northernmost wells ranged from 2.1% to 31.5% in BE-53 and 13.5% to 36.8% in BE-18. Toxicity in the wells that ran from the evaporation pond to the lake along a salt scar ranged from

5.8% to 35.4% in BE-11, 12.0% to 35.4% in BE-08, and 3.0% to 30.8% in BE-07. Observed toxicity in the wells located in vegetation running from the evaporation pond to the lake ranged from 48.1% to no observed effects in BE-59, 11.7% to 56.8% in BE-61, and 9.3% to 65.6% in BE-16.

Laboratory sediment bioassays with *C. dilutus* and *H. azteca* revealed no significant differences in mortality between any treatments in lab experiments. Two of three the experiments run resulted in significantly increased growths of organisms exposed to sediment from transect two, but no other significant differences were found.

Field Experiments

The initial *in situ* experiment with *C. dilutus* resulted in no mortality and very few significant treatment differences between reference and impact site treatments (Figure 8). Though complicated by a significant three-way interaction between sediment treatment, location, and experiment run, amphipods at both locations at the impact site exhibited lower growth than those placed at the reference site, regardless of experiment run.

Corbicula fluminea growth rates (Table 1) were significantly less in 2006 than in 2005 ($P = 0.0302$) and 2007 ($P < 0.0001$) regardless of placement location. Clams placed at the reference site grew significantly larger compared to clams at all impact sites for the first two years of the study (2005 $P < 0.0001$; 2006 $P < 0.0001$) but not during the final year of the study (2007 $P = 0.5948$ for the submerged evaporation pond and $P = 0.4498$ for transect two). When all data from all years were combined, clams at the reference site grew more than those placed at the impact site ($P < 0.0001$ for the submerged evaporation pond and $P = 0.0100$ for transect two), but no differences were observed between locations at the impact site ($P = 0.4109$).

There were few significant differences in the two benthic macroinvertebrate community metrics that were compared between the sites. Shannon-Wiener diversity ranged from 1.05 to

1.82 and did not differ between impact and reference sites, though richness and diversity tended to be greater at the impact than at the reference site (Table 2). Of the nineteen organisms found in more than 10% of samples, only the oligochaete, *Lumbriculus* spp. was found in greater abundance at the reference site, whereas mayflies, *Stenonema femoratum* (Heptageniidae) and *Paracloedes* spp. (Baetidae), and at one location the caddisfly, *Cernotina* spp. (Polycentropodidae), were found in greater abundances at the impact site (Figure 9).

Discussion

Produced Water Contamination

The Principle Components Analysis revealed two separate influences on groundwater composition that coincided with the first two axes. The first (horizontal) PCA axis represents biogeochemical processing of intruded produced water (such as bicarbonate buffering, chloride exchange for sulfate via hematite, bacterial removal of nitrate, or other processes; Kharaka et al. 2007) or dilution from other groundwater sources. The second (vertical) PCA axis represents the dilution gradient from produced water to lake water. Wells located farthest from produced water sources and the lake (such as BE-01, BE-51, and BE-17) plotted on the left side of the horizontal axis away from the produced water and lake samples, although they are still likely contaminated with produced water because previous research found produced water contamination in all site wells (Kharaka et al. 2007). The vertical axis confirms the previous report of Kharaka et al. (2007) showing produced water contamination by showing correlation with major cations and chloride found in produced water and wells located near produced water sources (such as BE-11, BE-08, BE-10, and BE-53). Produced water can contain large concentrations of sodium, chloride, and calcium ions (Kharaka et al. 2005; Kharaka et al. 2007; Benko and Drewes 2008). This has also been previously documented at our study site (Kharaka et al. 2005; Kharaka et al. 2007).

Both the total dissolved solids plots (Figure 4) and the Piper diagram (Figure 5) demonstrate a gradient in both total dissolved solids and relative ion concentrations from deeper sediment samples to those near the surface and lake water. This trend represents mixing of the lake sediment porewaters with the overlying lake water. A previous study of lake sediment porewater at the study site indicated that produced water contamination is present in the lake bed and that contaminant concentrations increase with sediment depth (Zielinski et al. 2007). That study also indicated that produced water-derived chloride ions and metals such as selenium, lead, copper and nickel were present in lake sediments at levels that exceed US EPA criteria and may cause toxicity to aquatic life. Zielinski et al. (2007) and this study also found major ion concentrations in lake sediments that exceeded lake concentrations and those contaminants increased with increasing sediment depth.

Potential Biological Effects

Produced water contamination of site groundwaters was sufficient that toxicity to bioassay organisms was predicted by toxicity models. The plumes of predicted toxicity were confirmed by observed toxicity at the wells and that toxicity was accurately represented by the predicted toxicity from the groundwater major ion concentrations (Fisher et al. 2010). In particular, greatest toxicity was observed in well BE-53 which is immediately downslope of the injection well and was on the upslope end of the northern plume. Also, very little to no toxicity was observed in waters from BE-59, which was located in a vegetated area next to the evaporation pond and was on a low toxicity area of the predicted toxicity surfaces. This result suggests that a location upslope, but immediately next to a produced water contamination source has little toxicity. Further, the two plumes documented in this study corresponded to two of the three toxic groundwater plumes reported by Kharaka et al. (2007). In this study, the third plume reported by Karaka et al. (2007) did not result in predicted toxicity, based on measured major

ions, likely because the plume source was an abandoned evaporation pond and not an active source.

Two of the five lake bed locations sampled for porewater major ions had concentrations that, based on toxicity modeling, would likely result in toxicity to bioassay organisms. The predicted toxicity based on both our data and that from Zielinski et al. (2007) suggested a decrease in toxicity as the sample neared the sediment surface, reflecting dilution by the overlying lake water.

Laboratory and Field Comparison

I*n situ* studies should ideally mirror results of other assessment methods, such as laboratory experiments and field studies, if these other assessment methods are to be considered valid. However, results of experiments comparing laboratory bioassays, *in situ* experiments, and field studies can vary greatly. Compared to laboratory bioassays, *in situ* exposures have been found to underestimate toxicity responses (Hose and Van den Brink 2004), overestimate toxicity responses (Sasson-Brickson and Burton 1991; Kater et al. 2001; Anderson et al. 2004), both over- and underestimate responses depending on the stressor (Tucker and Burton 1999), or agree with laboratory results (Schroer et al. 2004). Comparative studies have also found that effects measured in field experiments more closely resembled benthic community responses than laboratory bioassays, though laboratory studies are still relevant to field conditions (Hose and Van den Brink 2004; Ingersoll et al. 2005).

The laboratory bioassays on sediment taken from the reference site and the impact site did not result in any observed negative effects of the impact site. In contrast, *in situ* bioassays using amphipods and Asian clams resulted in consistent negative growth effects at the impact site. While the benthic communities differed between sites, those differences could be interpreted to indicate improved conditions at the impact site compared to the reference site. Of the

macroinvertebrates that differed between sites, the three found more commonly at the impact site had lower pollution tolerance values (Ephemeroptera: Heptageniidae and Baetidae = 4, Trichoptera: Polycentropodidae = 6) than the one found more often at the reference site (Oligochaeta = 8) indicating possibly improved water quality at the impact site due to the decreased tolerance of the resident community to pollution (Carter et al. 2006). However, those four taxa only represent a fraction of the taxa found and analyzed, indicating that the remaining 79% of analyzable taxa did not differ between sites. The results of this study indicate a general agreement between laboratory bioassays and the benthic community assessment with no observed impact site effects.

Aquatic studies of the biological effects of produced water discharges have in some cases indicated few or no significant effects on resident biota O'Neil (1992a). However, some chronic brine discharges from oil operations significantly change benthic communities. For example, Short et al. (1991) found that while fishes and Dipteran larvae were tolerant of salinities as high as 10 ‰, mayfly larvae (Ephemeroptera) were absent when salinities exceeded 2 ‰. Similarly, biological integrity of invertebrate communities in two Ohio streams receiving brine discharge were unchanged compared to reference sites, though greater percentages of salt-tolerant diatoms were present and chloride concentrations did not exceed 74 mg/L (Olive et al. 1992).

Experiments comparing the response of chronic *H. azteca* laboratory bioassays with field collected benthic community data demonstrated similar responses to increasing concentrations of polycyclic aromatic hydrocarbons (Ingersoll et al. 2005). The invasive Asian clam (*Corbicula fluminea*) has also been effectively employed to monitor both sediment and water quality in previously invaded environments and accurately reflects resident benthic community responses (Doherty 1990; Cataldo et al. 2001; Soucek et al. 2001). Given that the responses of the test organisms used have consistently corresponded with benthic community data in other studies, then either the benthic community at the impact site is different from the reference because of

produced water contamination or there is another unmeasured factor more important than produced water contamination that is structuring the communities at both sites. In this study, field bioassays were an integral part of the assessment strategy, and by demonstrating a slight, but consistent negative effect of the impact site have proven to be a more accurate measure of site quality than the laboratory bioassays.

Conclusions

Experiments presented here suggested that produced water at this site is toxic to aquatic life. Furthermore, due to leaching of produced water, the groundwater at sites near the lake is contaminated enough that toxicity to aquatic organisms would be expected if direct exposure to the groundwater were occurring. However, the groundwater appears to be sufficiently diluted as it enters the lake as laboratory and *in situ* experiments indicate limited direct toxicity or negative shifts in community structure. Despite the lack of in-lake environmental effect at the current study site, this study demonstrates the potential for produced water to pollute aquatic systems. Of particular concern would be bodies of water where limited dilution could occur.

Acknowledgements

We gratefully acknowledge the assistance of N. Cooper, M. Abbott, C. Boeckman, and K. Burgess for assistance with data collection. This research was funded through a cooperative agreement with the US Department of Energy (DE-FC26-04NT15544).

Literature Cited

Anderson, B.S., J.W. Hunt, B.M. Phillips, P.A. Nicely, R.S. Tjeerdema, and M. Martin. 2004. A comparison of in situ laboratory toxicity tests with the estuarine amphipod *Eohaustorius estuarius*. Archives of Environmental Contamination and Toxicology 46: 52-60.

APHA, 1995. Standard Methods, 19th Edition. American Public Health Association, Washington, DC.

Benko, K. and Drewes, J., 2008. Produced water in the western United States: geographical distribution, occurrence, and composition. Environ. Eng. Sci. 25, 239-246.

Boelter, A. M., F. N. Lamming, A. M. Farag, and H. L. Bergman. 1992. Environmental effects of saline oil-field discharges on surface waters. Environmental Toxicology and Chemistry 11: 1187-1195.

Burton Jr., G.A., M.S. Greenberg, C.D. Rowland, C.A. Irvine, D.R. Lavoie, J.A. Brooker, L. Moore, D.F.N. Raymer, and R.A. McWilliam. 2005. In situ exposures using caged organisms: a multi-compartment approach to detect aquatic toxicity and bioaccumulation. Environmental Pollution 134: 133-144.

Carter, J.L., V.H. Resh, M.J. Hannaford, and M.J. Meyers. 2006. Macroinvertebrates as biotic indicators of environmental quality. In Methods in Stream Ecology, F. R. Hauer and G. A. Lamberti eds. Academic Press, New York, P. 805.

Cataldo, D. H., D. Boltovskoy, J. Stripeikis, and M. Pose. 2001. Condition index and growth rates of field caged *Corbicula fluminea* (Bivalvia) as biomarkers of pollution gradients in the Paraná river delta (Argentina). Aquatic Ecosystem Health and Management, 4: 187-201.

Doherty, F. G. 1990. The Asiatic clam, *Corbicula* spp., as a biological monitor in freshwater environments. Environmental Monitoring and Assessment, 15: 143-181.

Douglas, W. S. and J. A. Veil. 1996. Do produced water toxicity tests accurately measure produced water toxicity in marine environments? 3rd International Petroleum Environmental Conference, ANL/EA/CP--89257, CONF-9609237--4.

Fisher, J. B. and K. L. Sublette. 2005. Environmental releases from exploration and production operations in Oklahoma: type, volume, causes, and prevention. Environmental Geosciences, 12: 89-99.

Fisher, J.C., J.B. Belden, and J.R. Bidwell. 2010. Can site-specific heuristic toxicity models predict the toxicity of produced water? Chemosphere 80:542-547.

Fucik, K. W. 1992. Toxicity identification evaluation and characteristics of produced water discharges from Colorado and Wyoming. In Produced Water, J. P. Ray and F. R. Engelhart, eds. Plenum Press, New York, Pp. 187-198.

Hauer, F. R. and V. H. Resh. 2006. Macroinvertebrates. In Methods in Stream Ecology, F. R. Hauer and G. A. Lamberti eds. Academic Press, New York, P. 446.

Hellawell, J.M. 1986. Biological indicators of freshwater pollution and environmental management. Elsevier, London. 546 p.

Hose, G.C. and P.J. Van den Brink. 2004. Confirming the species sensitivity distribution concept for Endosulfan using laboratory, mesocosm, and field data. Archives of Environmental Contamination and Toxicology 47: 511-520.

Ingersoll, C.G., N. Wang, J.M.R. Hayward, J.R. Jones, S.B. Jones, and D.S. Ireland. 2005. A field assessment of long-term laboratory sediment toxicity tests with the amphipod *Hyallela azteca*. Environmental Toxicology and Chemistry 24: 2853-2870.

Johnson, B.M., L.E. Kanagy, J.H. Rodgers, and J.W. Castle. 2008. Chemical, physical, and risk characterization of natural gas storage produced waters. Water Air and Soil Pollution 191:33-54.

Kater, B.J., J.F. Postma, M. Dubbledam, and J.T.H.J. Prins. 2001. Comparison of laboratory and in situ sediment bioassays using *Corphium volutator*. Environmental Toxicology and Chemistry 20: 1291-1295.

Kharaka, Y. K., J. J. Thordsen, E. Kakouros, and W. N. Herkelrath. 2005. Impacts of petroleum production on ground and surface waters: results from the Osage-Skiatook petroleum environmental research A site, Osage county, Oklahoma. Environmental Geosciences 12: 127-138.

Kharaka, Y.K., E. Kakouros, J.J. Thordsen, G. Ambats, and M.M. Abbott. 2007. Fate and groundwater impacts of produced water releases at OSPER "B" site, Osage County, Oklahoma. Applied Geochemistry 22: 2164-2176.

Marshall, N.A. and P.E. Bailey. 2004. Impact of secondary salinisation on freshwater ecosystems: effects of contrasting, experimental, short-term releases of saline wastewater on macroinvertebrates in a lowland stream. Marine and Freshwater Research 55: 509-523.

Merrit, R. W. and K. W. Cummins. 1996. An introduction to the aquatic insects of North America, 3rd ed. Kendall/Hunt Publishing Company, Dubuque.

Mount, D., Gulley, D., Hockett, J., Garrison, T., and Evans, J., 1997. Statistical models to predict the toxicity of major ions to *Ceriodaphnia dubia*, *Daphnia magna* and *Pimephales promelas* (fathead minnows). Environ. Toxicol. Chem. 16, 2009-2019.

Olive J. H., Jackson J. L., Keller D., Wetzel P. 1992. Effects of oil-field brines on biological integrity of 2 tributaries of the Little Muskingum Ruver, Southeastern Ohio. Ohio Journal of Science 92: 139-146.

O'Neil, P. E. S. C. Harris, M. F. Mettee, H. R. Isaacson, and J. M. Evans. 1992a. Biological fate and effect of coalbed methane produced waters discharged into streams of the Warrior Basin, Alabama. In Produced Water, J. P. Ray and F. R. Engelhart, eds. Plenum Press, New York, Pp. 315-327.

O'Neil, P. E., H. R. Isaacson, and J. M. Evans. 1992b. Surface discharge of coalbed methane produced waters in the Warrior Basin of Alabama, the Cedar Cove model. In Produced Water, J. P. Ray and F. R. Engelhart, eds. Plenum Press, New York, Pp. 329-341.

Otton, J. K., R. A. Zielinski, B. D. Smith, M. M. Abbott, and B. D. Keeland. 2005. Environmental impacts of oil production on soil, bedrock, and vegetation at the U.S. Geological Survey Osage-Skiatook Petroleum Environmental Research site A, Osage county, Oklahoma. Environmental Geosciences, 12: 73-87.

Peeters, E. M., J.P. Gardeniers, and A.A. Koelmans. 2000. Contribution of trace metals in structuring in situ macroinvertebrate community composition along a salinity gradient. Environmental Toxicology and Chemistry 19: 1002-1010.

Rosenberg, D.M. and V.H. Resh. 1993. Introduction to freshwater biomonitoring and benthic macroinvertebrates. *In* Freshwater Biomonitoring and Benthic Macroinvertebrates, D.M. Rosenberg and V.H. Resh, eds. Chapman and Hall, New York, pp. 1-9.

Sasson-Brickson, G. and G.A. Burton, Jr. 1991. In situ and laboratory sediment toxicity testing with *Ceriodaphnia dubia*. Environmental Toxicology and Chemistry 10: 201-207.

Schroer, A.F.W., J.D.M. Belgers, T.C.M. Brock, A.M. Matser, S.J. Maund, and P.J. Van den Brink. 2004. Comparison of laboratory single species and field population-level effects of the pyrethroid insecticide λ-Cyhalothrin on freshwater invertebrates. Archives of Environmental Contamination and Toxicology 46:324-335.

Short, T.M., J.A. Black, and W.J. Birge. 1991. Ecology of a saline stream: community responses to spatial gradients of environmental conditions. Hydrobiolgia 226: 167-178.

Smith, D. G. 2001. Pennak's freshwater invertebrates of the United States, 4th ed. John Wiley and Sons, New York.

Soucek, D. J., T. S. Schmidt, and D. S. Cherry. 2001. In situ studies with Asian clams (*Corbicula fluminea*) detect acid mine drainage and nutrient inputs in low-order streams. Canadian Journal of Fisheries and Aquatic Sciences, 58: 602-608.

Tietge, J., Hockett, J., and Evans, J., 1997. Major ion toxicity of six produced waters to three freshwater species: application of ion toxicity models and TIE procedures. Environ. Toxicol. Chem. 16, 2002-2008.

Tucker, K.A. and G.A. Burton, Jr. 1999. Assessment of nonpoint-source runoff in a stream using *in situ* and laboratory approaches. Environmental Toxicology and Chemistry 18: 2797-2803.

U.S. Environmental Protection Agency. 2000. Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates. EPA/600/R-99/064. Duluth, MN.

U.S. Environmental Protection Agency. 2002. Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms. EPA/821/R-02/012. Washington, DC.

Winston, RB. 2000. Graphical User Interface for MODFLOW, Version 4: U.S. Geological Survey Open-File Report 00-315, 27 p.

Zielinski, R.A., W.N. Herkelrath, and J.K. Otton. 2007. Composition of pore water in lake sediments, research site “B”, Osage County, Oklahoma: Implications for lake water quality and benthic organisms. Applied Geochemistry 22: 2177-2192.

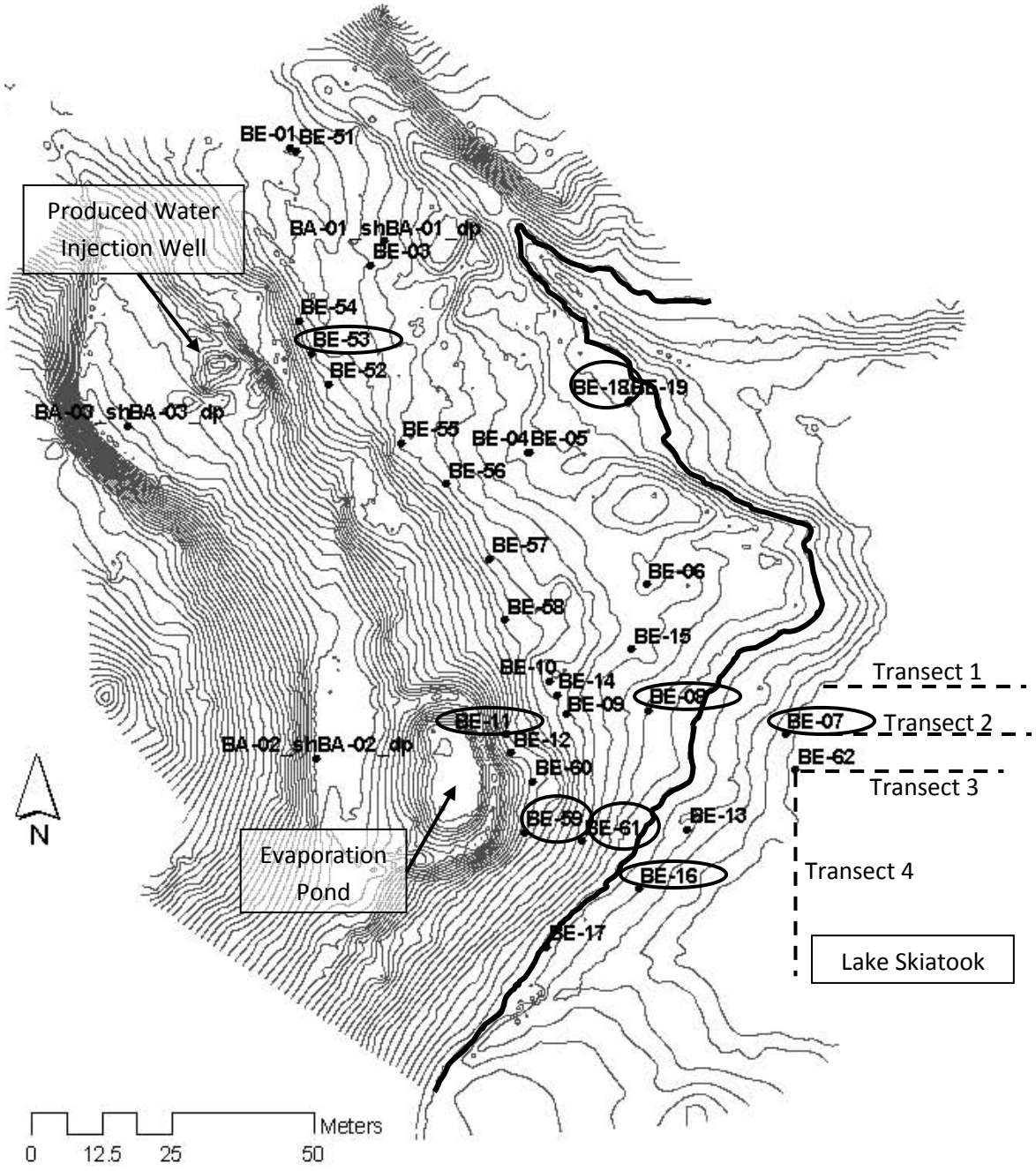

Figure 1. Map of the site influenced by produced water on Lake Skiatook, Oklahoma showing elevation contours (each line represents a 20cm change in elevation increasing from the lake or east side), groundwater well locations, sources of produced water and lake benthic transects. Circled groundwater well labels indicate wells selected for bimonthly toxicity monitoring. The conservation pool level for Skiatook at 217.5 m is indicated by the thick black contour line.

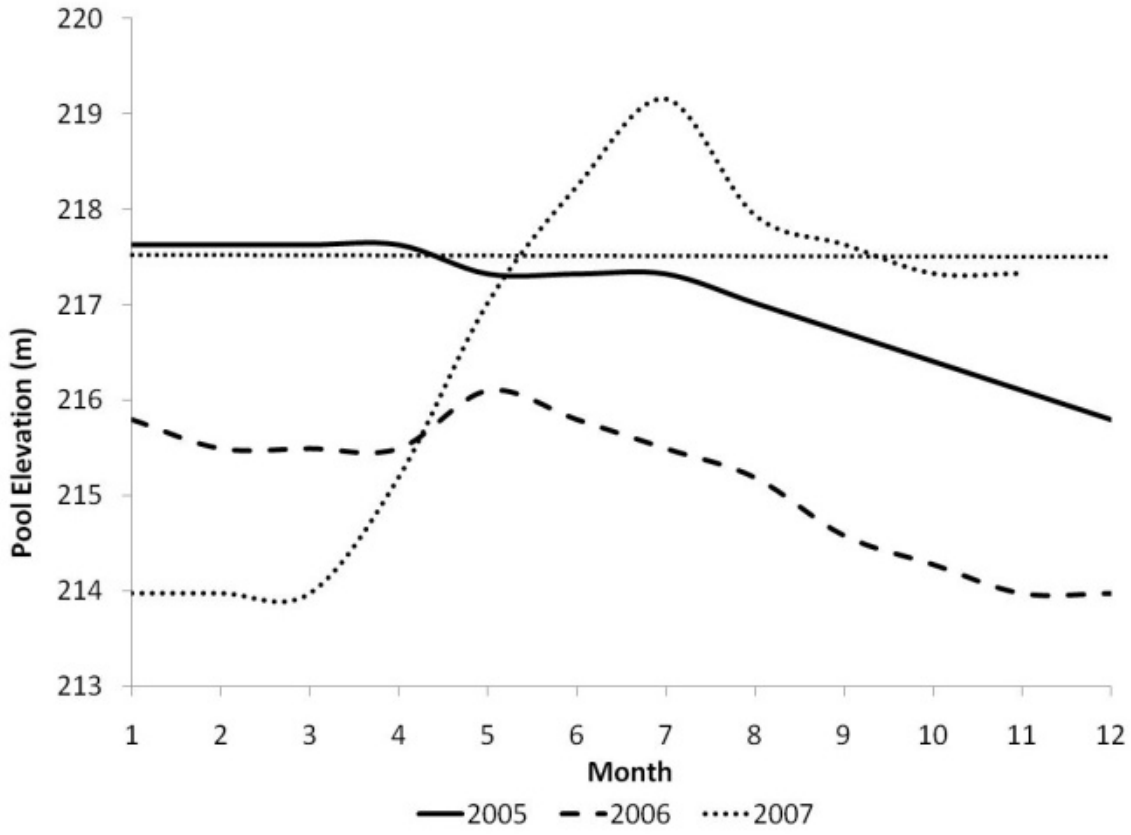

Figure 2. Hydrograph of Skiatook Lake water levels during the study period. The conservation pool elevation is indicated by a horizontal dashed black line.

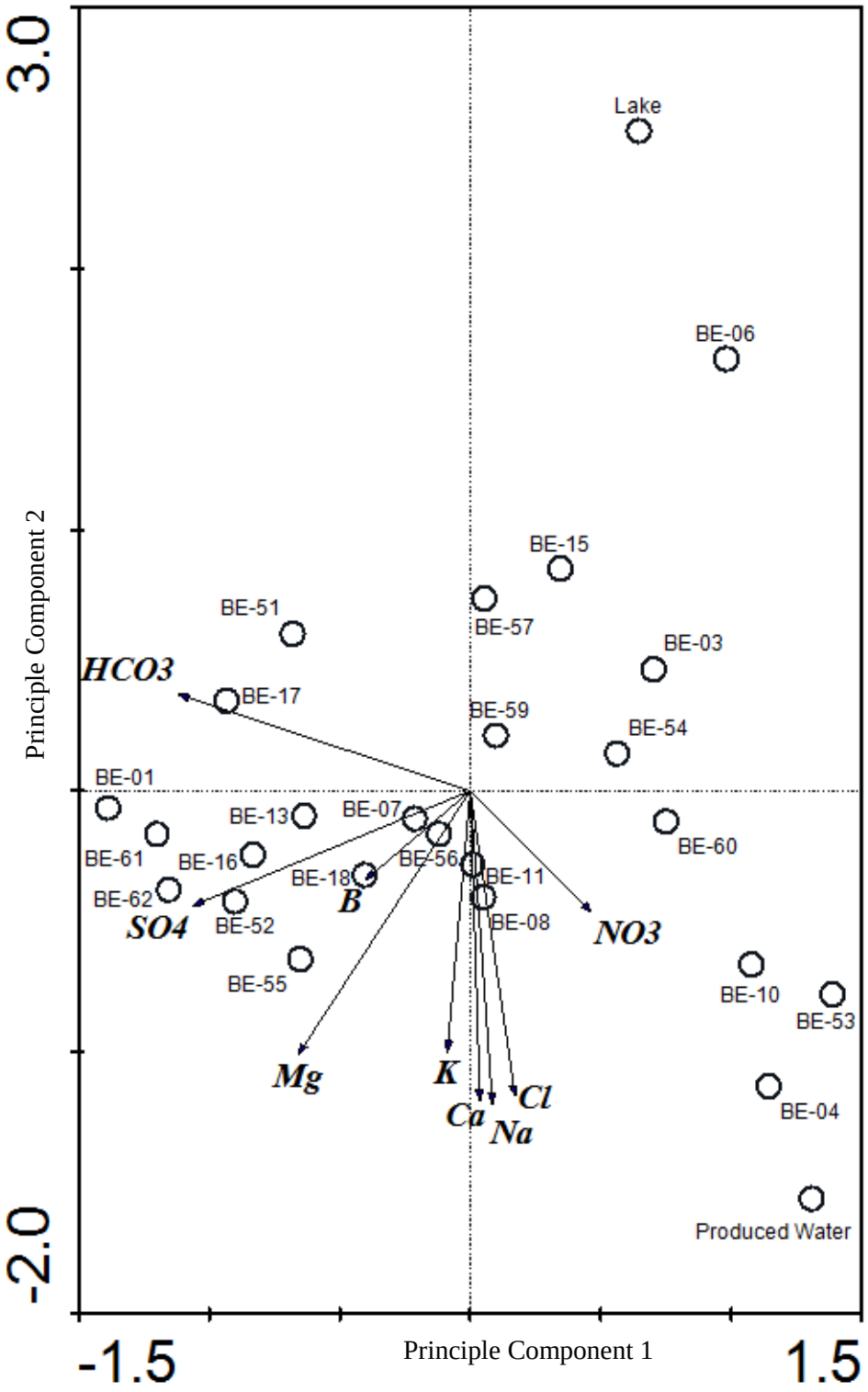

Figure 3. Principle Component Analysis biplot of study site groundwater wells and ion profiles. Labeled points are groundwater wells and vectors are measured parameters.

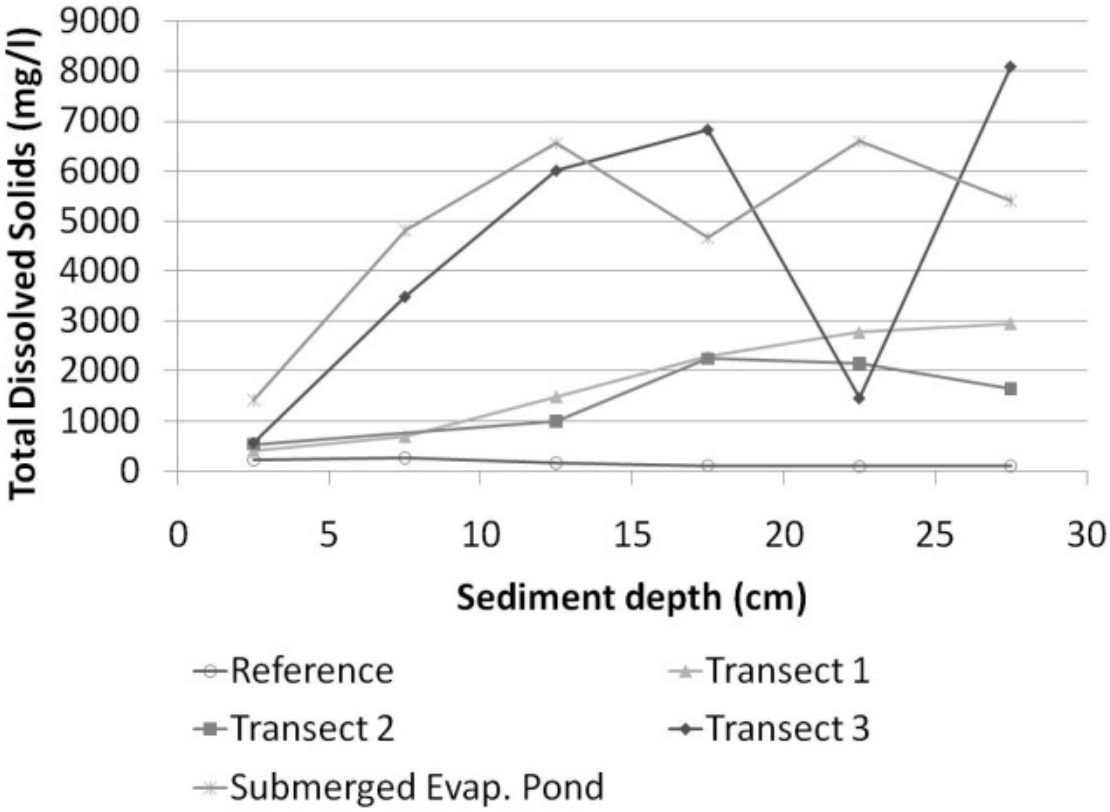

Figure 4. Plot of total dissolved solids for lake sediment porewater at the impact site. Increased total dissolved solids indicate brine presence in porewater samples.

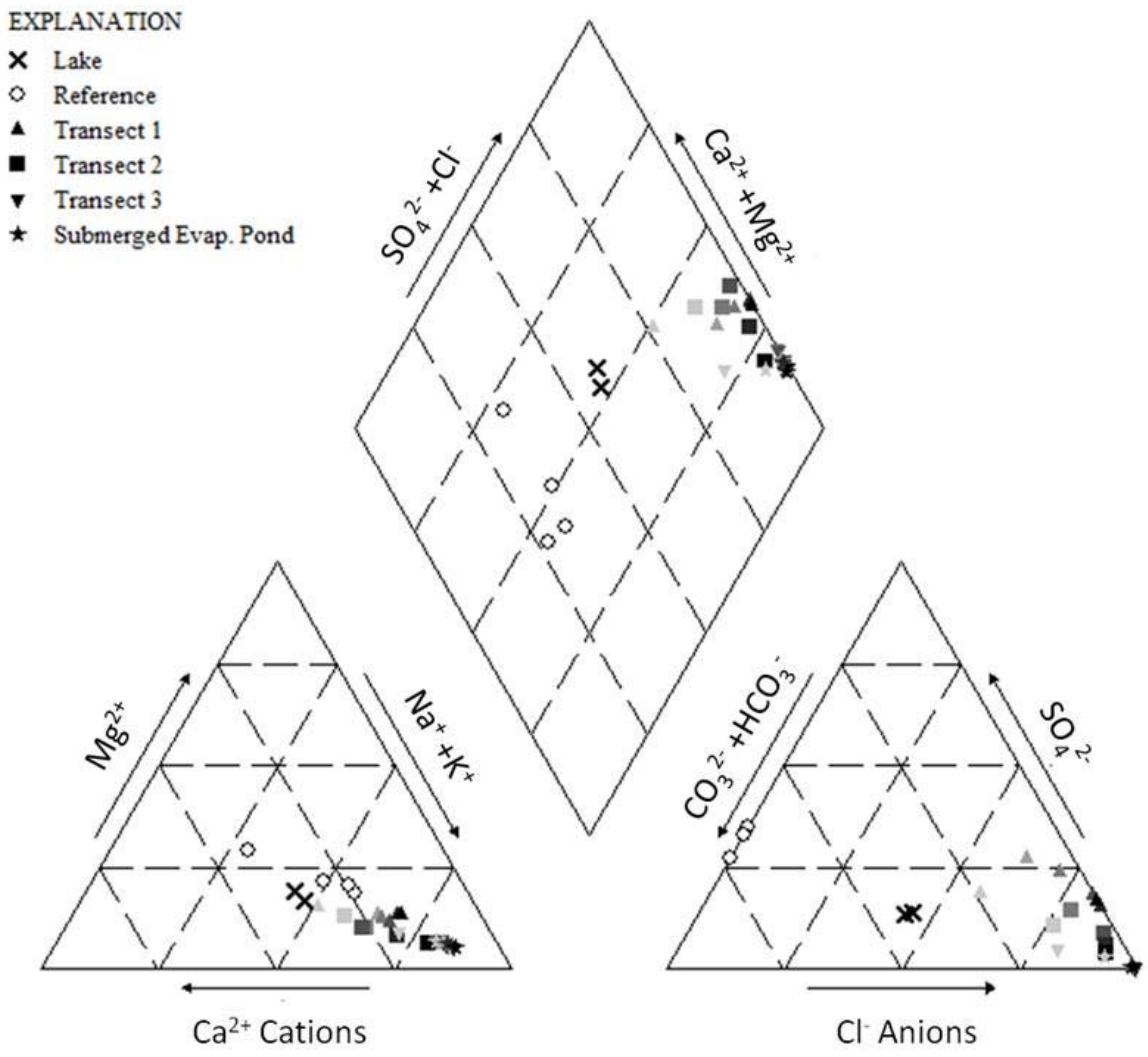

Figure 5. Piper diagram for Skiatook lake sediment porewater from the impact site. The Piper diagram plots the percentage of meq/l of cations (left triangle) and anions (right triangle). Cations and anions are projected onto the center diamond to simultaneously show all major ion meg/l percentages. Darker symbols (light gray or shallow depth to black or deeper depth) indicate increased sediment depth of porewater samples.

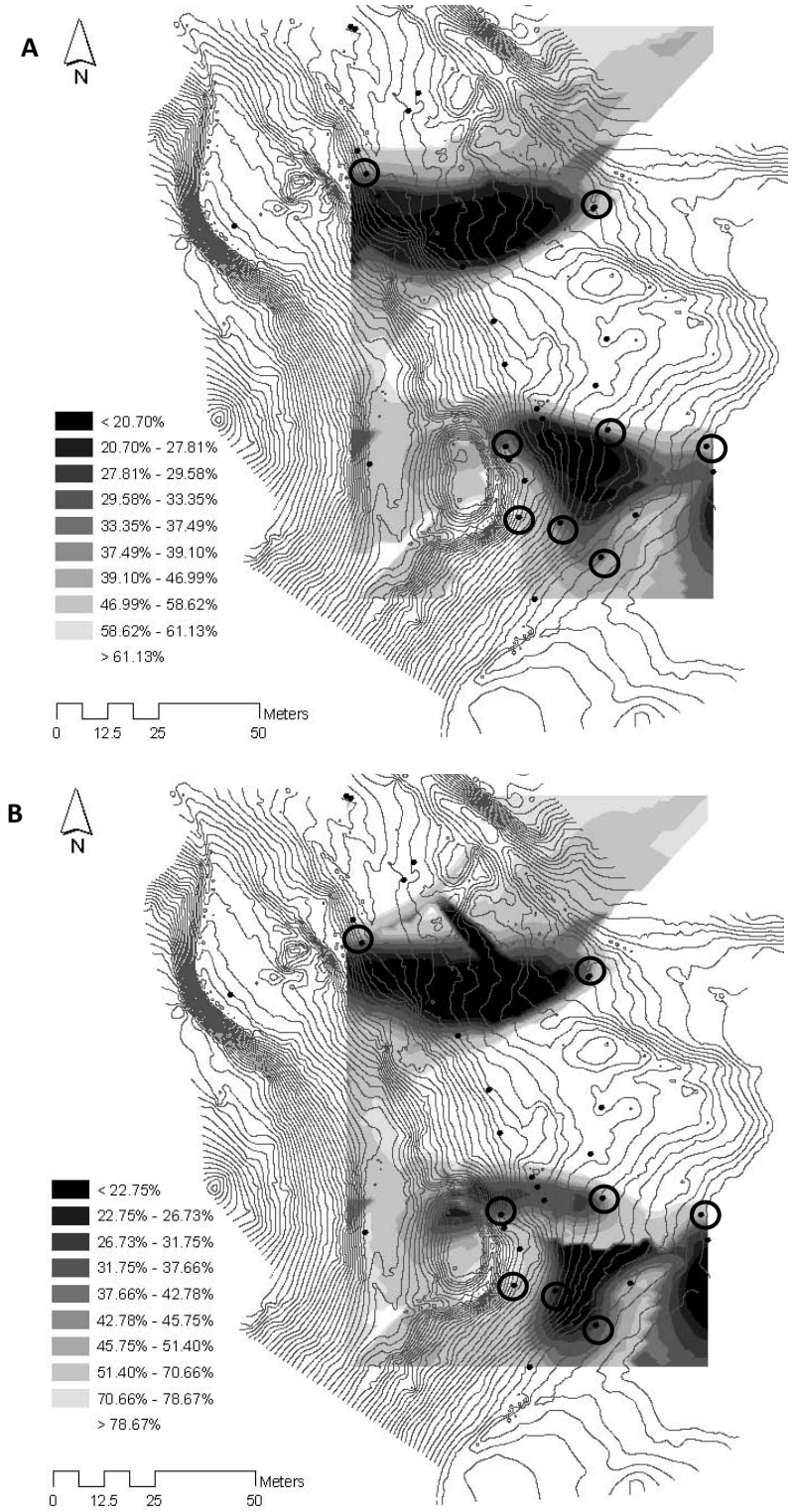

Figure 6. Groundwater brine contamination at the impact sit shown as predicted major ion toxicity to *Ceriodaphnia dubia* (A) and fathead minnows (B). Darker shading indicates increased predicted toxicity and toxicity is expressed as the percent concentration of groundwater that corresponds to the 48h LC50. Black dots indicate groundwater wells and circled wells were wells monitored bimonthly for acute toxicity.

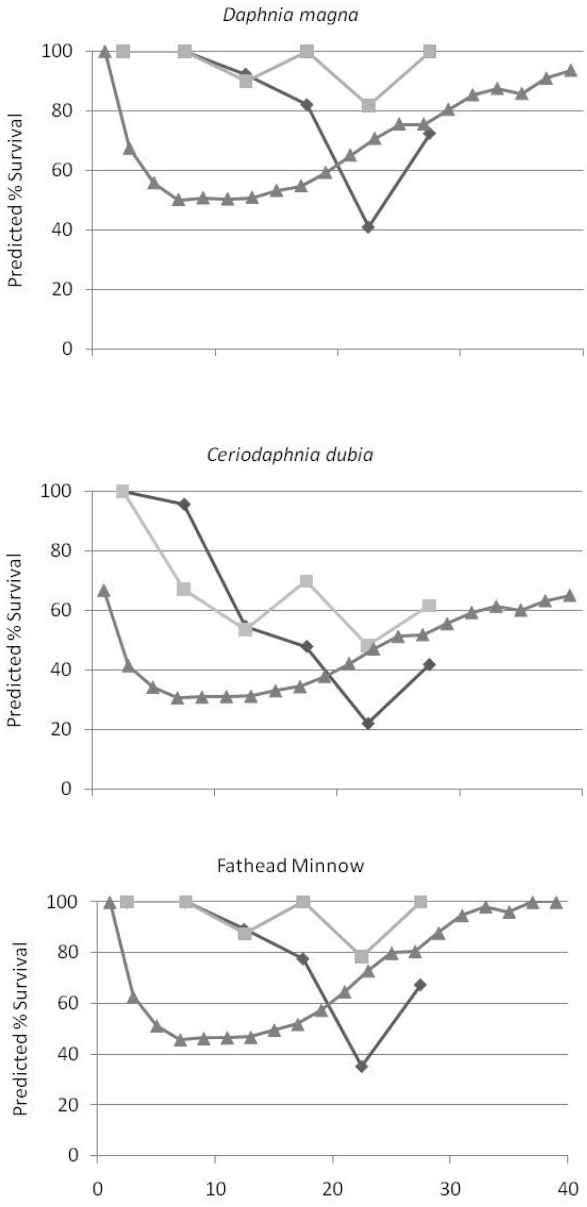

Figure 7. Predicted Lake Skiatook, OK sediment porewater toxicity using measured major ion concentrations to bioassay organisms. Sampling locations where toxicity was predicted were Transect 3 (diamonds), the submerged evaporation pond (squares), and the conductivity hot spot evaluated by a previous study (triangles, see Zielinski et al. 2007).

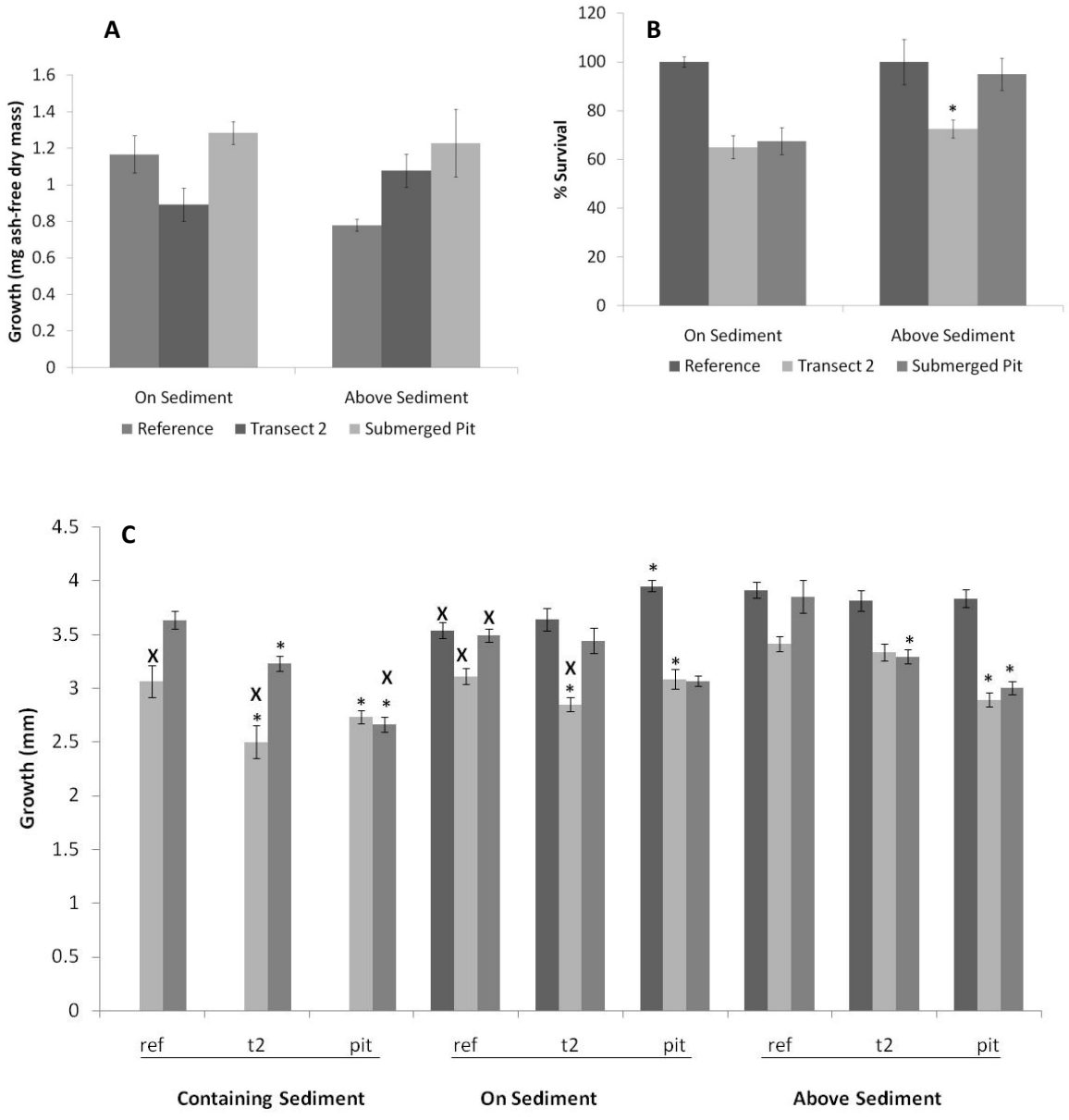

Figure 8. Results of *in situ* bioassays for midge and amphipod experiments for midge growth A), percent survival (B), and amphipod mean growth (C). Error bars are standard error of the mean. An asterisk (*) indicates a significant difference from reference site values for each sediment treatment at α=0.05. An X indicates a significant difference between the containing sediment or on sediment treatment from the water column (above sediment) treatment. Labels along x-axis indicate placement of exposure chambers where ref is the reference site, t2 is along transect 2, and pit is the submerged evaporation pond.

Table 1. Growth rates (mm/day) and standard error (in parentheses) of Asian clams from *in situ* growth experiments testing for the effects of produced water contamination. Rates are corrected for length of experiment (30 to 35 days) and initial size. No data is reported for Transect 2 during 2006 because low water levels prevented experiments at that location.

Year	Year Mean	Reference	Transect 2	Submerged Evap. Pond
2005	0.0337 (0.0015)	0.0513 (0.0024)	0.0367 (0.0024)	0.0387 (0.0019)
2006	0.0270 (0.0019)	0.0217 (0.0023)	-	0.0087 (0.0018)
2007	0.0380 (0.0019)	0.0297 (0.0048)	0.0257 (0.0029)	0.0280 (0.0018)
Total		0.0350 (0.0021)	0.0287 (0.0020)	0.0263 (0.0018)

Table 2. Summary of macroinvertebrate samples from Lake Skiatook benthos. Taxa richness is the total number of taxa per sample across years. Diversity is Shannon-Wiener diversity of each sample. No data is reported for Transect 2 during Fall because low water levels prevented macroinvertebrate sampling at that location.

Location	Season	N	Mean number of Taxa	Mean Taxa Richness	Diversity (H)
Reference	Spring	10	55.7	6.9	1.24
Transect 1	Spring	5	57.2	9.2	1.43
	Fall	3	187	9.0	1.41
Transect 2	Spring	5	59.0	10.0	1.55
Transect 3	Spring	3	77.7	10.7	1.34
	Fall	3	131	9.3	1.62
Evap. Pond	Spring	3	72.0	8.0	1.05
	Fall	3	68.0	9.3	1.82

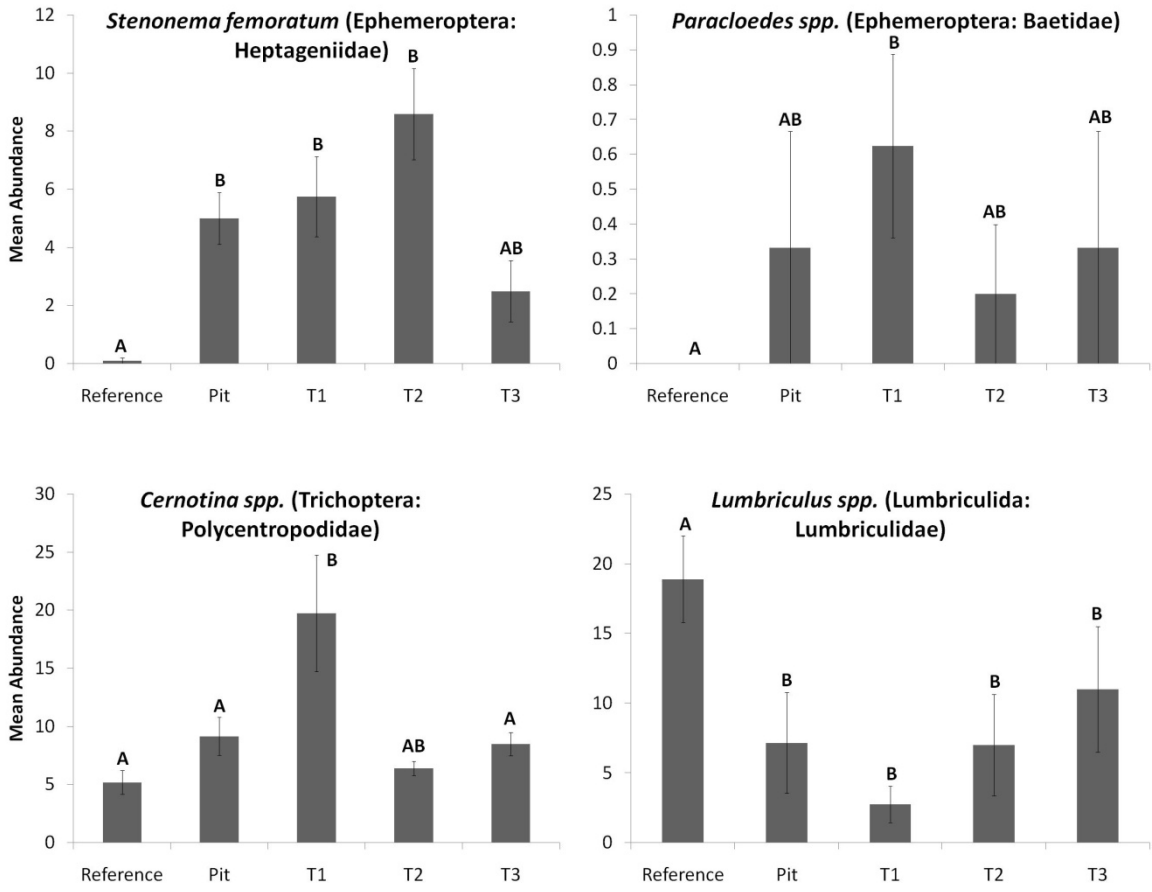

Figure 9. Mean abundance of macroinvertebrates from each sampling location (Reference = Reference site, Pit = submerged evaporation pond, T1 = transect 1, T2 = along transect 2, and T3 = transect 3) where significant ($\alpha = 0.05$) differences were detected. Significance groupings within each taxon are denoted by letters. Error bars are standard error of the mean.

CHAPTER IV

ASSESSMENT OF THE JOINT EFFECTS OF TWO HERBICIDES AND RESOURCE LIMITATION ON ALGAE USING MIXTURE MODELING APPROACHES

Abstract

Aquatic organisms frequently encounter stressors from numerous sources, in many different combinations, and at different relative effects levels. This study expanded on current mixture modeling and analysis techniques in order to assess the effects of mixtures containing chemical and nonchemical stressors, any number of stressors at any relative potency, and for any combination of similar and dissimilar modes of action. The protocols presented here allow assessment of concentration addition, independent action, synergism, antagonism, dose level-dependent toxicity, and dose ratio-dependent toxicity. These techniques were first grounded in existing mixture modeling protocols by analyzing literature datasets and comparing results between techniques. These techniques were then used to assess the joint action of atrazine, norflurazon, and resource limitation (light limitation). These anthropogenic and natural stressors affected the study organism, *Pseudokirchneriella subcapitata*, via similar and dissimilar modes of action. Mixture effects of the tertiary mixture were predicted within a factor of two of the observed data using a mixture model that incorporated both concentration addition and independent action.

Introduction

Aquatic organisms rarely experience single environmental stressors. Toxicants, multiple toxicants, or unfavorable physical and ecological conditions may all occur simultaneously, confounding environmental assessments. Pesticides are particularly problematic because they are intentionally applied to the environment in a logistically un-testable number of combinations resulting in very little knowledge about their combined effects (Lydy et al. 2004). Therefore, predictive models are necessary to adequately estimate risk.

Concentration addition (CA) and independent action (IA) are the most commonly employed predictive methods for determining toxicant combined effects. The toxicological theory supporting concentration addition is the assumption that all toxicants act similarly or are dilutions of each other (Altenburger et al. 2000). Independent action assumes toxicants interact with organisms in completely different and independent manners (Backhaus et al. 2000). Each of these models is predictive – in that the models were developed based on individual tests prior to evaluating the experimental units that had more than one variable (i.e., chemical stressor) applied to it. Through comparison of the experimentally estimated values to that of the expected values based on the models, predictability of each model is determined providing insight into the toxicological joint action. A metric for comparing the fit of multiple predictive models is the model deviance ratio (MDR, Belden et al. 2007). This ratio is calculated by dividing the concentration predicted to cause toxicity (e.g., median lethal concentration) by the concentration observed to cause toxicity. A value of one indicates exact agreement between the predictive model and data, while values within a factor of two (between 0.5 and 2.0) are considered to indicate adequate agreement (Belden et al. 2007). Pesticides with similar modes of action (MOA) tend to be modeled best by the CA model (Altenburger et al. 2003; Belden et al. 2007). However the applicability of this to two similar MOA stressors, of which one is not a pesticide, is not known Testing of the predictive procedures is frequently conducted using equipotent mixtures,

meaning that a single mixture is produced with the concentration of all toxicants at a ratio of their effective concentrations (Altenburger et al. 2000; Altenburger et al. 2003; Belden and Lydy 2006). However, the probability that any two stressors would co-occur at equipotent or equitoxic levels, though incalculable, is likely extremely low.

This study utilized three stressors: atrazine, norflurazon, and resource limitation (i.e., light limitation). Atrazine causes toxicity by interrupting photosystem II, which decreases a plant's ability to derive energy from sunlight and causes oxidative cellular damage when the energy being processed through photosystem II is released into the cell as a reactive oxygen species (Shimabukuro and Swanson 1969). Norflurazon causes toxicity by inhibiting the synthesis of carotenoids, removing cellular protections from naturally occurring reactive oxygen species (Ahrens 1994). Light is the energy source for both the algal cells and the reactive oxygen species. Decreasing light exposure may both limit the energy available to algal cells for population growth and lessen the relative effect of herbicides with light dependent modes of action.

The goal of this study was to assess the joint action of an environmental stressor with similarly acting anthropogenic chemical stressors and evaluate how well models grounded in ecotoxicological theory account for these combined effects. The specific objectives were to 1) modify concentration addition (CA) and independent action (IA) approaches for mixtures of any dimension where toxicity is not evenly divided between toxic components, 2) compare the methods described here with other methods for non-equitoxic mixtures, specifically those of Jonker et al. (2005), and 3) examine the joint action of atrazine, norflurazon and resource limitation, stressors with dissimilar modes of action, but potentially similar physiological effects.

Methods

Models

In order to assess the joint action of multiple stressors, several methods commonly or potentially employed for this purpose will be compared. Concentration addition and IA are the most commonly employed predictive methods for determining toxicant combined effects (Figure 1). Mathematically, CA is often represented by using summed standardized individual concentrations (Toxic Units), which can be calculated as follows:

$$ECx_{mix} = \left(\sum_{i=1}^{n} \frac{p_i}{ECx_i}\right)^{-1} \quad (1)$$

where ECx_{mix} is the total mixture concentration that causes *x* effect, p_i indicates the component *i* mixture proportion, and ECx_i indicates the component *i* concentration that would cause *x* effect (Altenburger et al. 2000). Mixture effects from dissimilarly acting compounds using IA can be calculated with the following equation:

$$E(c_{mix}) = 1 - \prod_{i=1}^{n}[1 - E(c_i)] \quad (2)$$

where $E(c_{mix})$ is the total mixture effect and $E(c_i)$ is the expected component *i* effect (Backhaus et al. 2000).

Both of these models assume that all mixture components occur at concentrations where all components would, if tested independently, cause the same measured effect or are equitoxic. It is important to note that CA essentially sums the standardized component concentrations before determining the dose response relationship, whereas IA combines the expected effects from each individual dose response in a probabilistic manner (Figure 1).

The toxicity unit (TU) is a way to standardize mixture components based on potency and is simply the part of the CA equation being summed for each component. The individual TUs of all components in an equitoxic mixture would be equal. Jonker et al. (2005) described a method for describing the TU proportion for mixture components as follows:

$$z_i = \frac{TU_i}{TU_{sum}} \quad (3)$$

where z_i is the TU proportion, TU_i is the toxicity units of the mixture component of interest, and TU_{sum} is the sum of all mixture toxicity units. Using this approach, any component in an equitoxic mixture would have a z_i equal to the inverse of the total number of mixture components (n^{-1}). However, as the mixture departs from equitoxicity, at least one component will have a z_i greater than n^{-1}, and at least one of the remaining components will have a proportion less than the equitoxic levels. This change in proportional toxicity contribution of mixture components likely changes the final mixture dose response from that derived using CA. Because the mixture has changed to be dominated by one component, then the dose response curve of that mixture should approach the response of the dominant component (Figure 1).

Concentration addition for an equitoxic mixture can be thought of as the average dose response for all mixture components at the TUs for the total mixture. When the mixture departs from equitoxicity, that final dose response curve should reflect the increased toxicity contribution of the dominant component. Concentration addition could then be modified as a weighted average of the individual dose response curves, where the weight is z_i. Mathematically, this would be expressed as follows:

$$E(c_{mix}) = \sum_{i=1}^{n} z_i \cdot E(TU_{sum})_i \quad (4)$$

where $E(c_{mix})$ is the expected response and $E(TU_{sum})_i$ is the expected response of each component for the sum of all toxicity units.

Independent action does not need to be weighted because it is a probabilistic response combination and thus responses are directly calculated with the model.

Organisms and Chemical Analyses

Atrazine certified at 99.5% purity was obtained from Chem Service (West Chester, PA, USA) and norflurazon certified at 98.6% purity was obtained from Sigma Aldrich (St. Louis, MO, USA). All other chemicals used for culture mediums were of reagent grade. Algae (*Pseudokirchneriella subcapitata*) for testing were obtained from Aquatic Biosystems, Inc. (Fort Collins, CO, USA).

Experiments were conducted in algal culture medium (US EPA 2002). Chlorophyll was measured daily in algal experiments using a Trilogy Laboratory Fluorometer (Turner Designs, Sunnyvale, CA, USA). Measurements were taken in vitro and measured chlorophyll–a in Relative Fluorescence Units (RFU). All fluorescence measurements for each experimental run were standardized to percent control (24h light and no herbicide) for that experiment.

Experimental Design and Testing Protocol

Chronic bioassays used the algae *Pseudokirchneriella subcapitata* (4d, USEPA 2002). Algal tests varied daily illumination to limit resource availability. Based on preliminary toxicity tests, atrazine treatment levels were set at 0, 85.5, 171, 342, and 684 μg/l, norflurazon treatment levels were set at 0, 2, 4, 10, 20, and 40 μg/l, and resource treatments were 0h (control), 2.5h, 5h, 10h, 15h, and 20h of darkness per day. An acetone carrier was used to integrate atrazine and norflurazon in the culture medium, however, final acetone carrier concentration did not exceed 0.1 ml/L in any treatment and solvent controls were employed.

Atrazine and norflurazon were tested in combination with resource limitation such that all pesticide concentrations were evaluated at each level of resource limitation. Equipotent mixtures were included in all stressor combinations. However, treatment combinations were also chosen to represent the widest possible variation in dose levels and dose ratio (the ratio between toxicants) which was achieved using a complete block design and resulted in only 45% of mixture treatment combinations approximating equipotency. Each treatment combination

contained four replicates with 100µl stock culture per 3 ml replicate with three test runs per experiment. The daily measured algal endpoint was chlorophyll-a fluorescence.

Data and Analyses

This study also utilized data analyzed by other authors using a mixture modeling technique proposed by Jonker et al. (2005) in order to demonstrate the data analysis technique presented here. Their technique allows the analysis of non-equitoxic mixtures and specifically tests for CA, IA, synergism, antagonism, dose level-dependent toxicity, and ratio dependent toxicity. Response data means for *Caenorhabditis elegans* to mixtures of copper and zinc and for *Folsomia candida* to mixtures of cadmium and zinc were originally published by Jonker et al. (2004) and Jonker et al. (2005), respectively. Response data for *Lumbricus rubellus* to mixtures of fluoranthene and desiccation stresses were simulated based on mean, standard error, and sample data presented in Long et al. (2009). The data were simulated to demonstrate how the techniques proposed here work with full data sets, instead of just treatment means. These data are only presented here to ground the currently proposed technique in existing mixture toxicity practice.

Model Deviation Ratios (MDRs) are used to compare mixture toxicity model predictions with observed bioassay endpoints. MDRs are calculated by dividing the predicted EC50 by the observed EC50 (Belden et al. 2007). Ratios that fall within a factor of two (between 0.5 and 2.0) are considered to show adequate agreement between predicted and observed values. Belden et al. (2007) found that for pesticide mixture experiments where concentration addition was expected, greater than 90% had MDRs within a factor of 2. This degree of deviation is only slightly larger than the magnitude of the uncertainty usually associated with toxicity testing. For instance, the U.S. EPA (2002) reported the coefficient of variation for intra-laboratory and inter-laboratory bioassays ranged from 8% to 41% and 13% to 38.5%, respectively, or within a factor of 1.7.

Equipotent mixture combinations were compared to predictions based on CA, IA, and a combination model that included both CA and IA (tertiary mixture only) following Altenburger et al. (2000). These combinations were graphed with model predictions using a logistic regression and estimated specific effective concentrations (5, 10, 15, 50, 85, 90, 95% effect; following Belden and Lydy 2006) with 95% confidence intervals.

This study also evaluated the predictive ability of mixture modeling techniques at any level of toxicant mixture concentrations. As such, model predictions were assessed using a simple ratio as follows:

$$MeDR = \frac{E(TU)_{predicted}}{E(TU)_{observed}} \quad (5)$$

where MeDR is a modified MDR called the Model effects Deviation Ratio, $E(TU)_{predicted}$ is the predicted effect based on a mixture model at some TU level, and $E(TU)_{observed}$ is the observed effect at that mixture treatment level. Pointwise comparisons based on MeDRs provide a statistically powerful way to assess mixture modeling predictions across the entire range of experimental units.

Jonker et al. (2005) included explicit statistical tests in a spreadsheet environment for quantifying the dynamics of the organismal response to a binary toxicant mixture. Specifically, tests were done to evaluate if CA or IA best fit the data, if there was synergism or antagonism, and if toxicity responses changed from predictions based on the dose applied (dose dependent toxicity) or the ratio of the toxicants applied (ratio dependent toxicity). Here, model fit and toxicity dynamics are examined using two plots of the data (dependent vs independent): MeDRs vs sum TU and MeDRs vs toxicant TU ratios.

To determine whether CA or IA best fit the data, a box containing MeDRs between 0.5 and 2.0 (a factor of two around one) is placed on the graph. MeDRs from the model best

contained within a factor of two indicate the best fitting model. This comparison can also be achieved using summary statistics and comparing the mean MeDR between models.

Synergism refers to a toxicological interaction where the measured effects from the exposure of an organism to a toxicant mixture results in negative effects greater than those predicted by CA or IA. Antagonism is the opposite where the measured negative effects are lower than predicted. MeDRs with values consistently above one indicates a synergistic relationship, while those below one indicate an antagonistic relationship. This can be assessed by visual inspection of the MeDR plots, comparing means, or using the above mentioned z-test to test if standardized MeDRs are greater than zero (synergism) or less than zero (antagonism).

Dose-dependent and ratio-dependent deviation from models can both be tested using simple linear regressions on one of the two MeDR plots. Regressing MeDRs on the sum toxicity unit graph provides a test for dose-dependent deviation. The same type of regression on the TU ratio graph provides a test for ratio dependent deviation. A significant slope (when regressed using a statistical analysis program) on the sum TU or TU ratio graph indicates significant dose-dependent or ratio-dependent toxicity, respectively.

Statistical analyses were performed using SAS version 9.1.3 (SAS Institute, Cary, NC, USA). Generalized linear mixed models, including dose response relationships and simple linear regressions, were analyzed using Proc GLIMMIX, while other models, such as those requiring a hormetic response, utilized Proc NLIN. Dose response relationships for atrazine, norflurazon, and resource limitation (light) were derived using Brain-Cousens logistic regression (Schaebenberger et al. 1999), log-linear regression, and logistic regression, respectively. Predictions from mixture models, summary statistics, confidence intervals, and derived single compound dose responses were analyzed in Microsoft Excel (Microsoft Corporation, Redmond, WA, USA).

Results

Demonstration of MeDR approach using literature derived values

Measured toxicity for *C. elegans* exposed to a mixture of zinc and copper began to depart from mixture toxicity predictions at approximately four toxicity units (Figure 2). Overall, the mean MeDRs for these data were 1.54 for CA and 1.06 for IA. Simple linear regressions fit to the models indicate significant positive relationships for CA (P = 0.0319) but not IA (P = 0.2209) predictions. Due to the lack of variability in mixture TU ratios for this data, no reliable trends regarding ratio dependent toxicity could be discerned.

Measured toxicity for *F. candida* exposed to zinc and cadmium began to depart from mixture toxicity predictions between two and three toxicity units (Figure 2). Overall, the mean MeDRs for these data were 0.57 for CA and 0.99 for IA. Simple linear regressions fit to the models indicate a significant relationship for CA (negative relationship, P = 0.0482) but not IA (P = 0.0773). Both models showed a significant decreasing trend with increasing relative amounts of cadmium (P = 0.0053 for CA, P = 0.0133 for IA).

Measured toxicity for *L. rubellus* exposed to fluoranthene and dessication departed from mixture toxicity predictions at higher TU's (Figure 2). The mean MeDRs for these data were 1.51 for CA and 1.38 for IA. Simple linear regressions indicated no relationship for CA (P = 0.0792) or for IA (P = 0.0792). Though no trend for CA (P = 0.0510) was observed, IA (P = 0.0311) MeDRs exhibited significant positive trends based on toxicant ratios.

Assessment of Herbicide-Herbicide and Herbicide-Resource Limitation Effects Using Equipotent Mixtures

Using only the equipotent combinations (Figure 3), atrazine and resource limitation most likely mix in an independent (IA) manner, though the MDRs for these relationship place both

models within a factor of two of observed data (Table 1). Norflurazon and resource limitation plots exhibited (Figure 3) less toxicity than predicted by IA or CA, though IA predictions were closer to observed (Table 1). Atrazine and norflurazon mixtures were plotted closest to IA (Figure 3) and the IA model had an MDR closer to one (Table 1). However, norflurazon mixed with atrazine and resource limitation in binary experiments in a manner that suggested less observed toxicity than predicted by any of the three models (Figure 3). Model MDRs (Table 1) also indicated that predicted EC50s for CA, IA and the combination model were more than a factor of two more toxic than observed.

Assessment of Herbicide-Herbicide and Herbicide-Resource Limitation Effects Using MeDR Approach

Binary mixtures of atrazine and light limitation using *P. subcapitata* remained close to CA predictions up to 2..0 TUs (Figure 4). Mean MeDRs for these data were 0.62 for CA and 2.12 for IA. Simple linear regressions fit to CA indicate a negative relationship for MeDRs as TU's increase ($P < 0.0001$). IA predictions departed from observed values at 2.0 TUs and had a significant positive slope ($P < 0.0001$). No relationship was observed for IA predictions ($P = 0.4074$) but a significant negative relationship was found for CA predictions ($P < 0.0001$) based on the ratio of atrazine to resource limitation stress.

Binary mixtures of norflurazon and resource limitation were modeled similarly using both CA and IA (Figure 4). Mean MeDRs for these data were 0.71 for CA and 0.89 for IA. Simple linear regressions fit to CA and IA based on TUs indicate a significant negative slope for both models (CA, $P < 0.0001$; IA $P = 0.0002$). CA did not exhibit any ratio dependent toxicity ($P = 0.0796$), while a slight negative relationship for IA ($P = 0.0232$) existed based on TU ratios.

Binary mixtures of atrazine and norflurazon were modeled similarly for both models up to 2.50 TUs where both models predictions diverged from one (Figure 4). Mean MeDRs for

these data were 0.54 for CA and 0.86 for IA. Both models significantly diverged from agreement with observed values with a negative relationship for both CA ($P < 0.0001$) and for IA ($P = 0.0207$). No relationship based on TU ratios for CA ($P = 0.1216$) or IA ($P = 0.2524$) was found.

The tertiary mixture containing atrazine, norflurazon, and resource limitation was modeled using CA for all components, IA for all components, and using a mixing model that combined atrazine and resource limitation with CA and then combined that combination with norflurazon using IA (Figure 5). Mean MeDRs for these data were 0.40 for CA, 0.97 for IA, and 0.24 for the combination model. Two of three models exhibited a significant negative relationship with increasing TUs ($P < 0.0001$ for CA, $P < 0.0001$ for IA, and $P = 0.0256$ for the combination model). Few significant trends based on toxicant ratios for atrazine ($P = 0.6451$ for CA, $P = 0.6887$ for IA, and $P = 0.6545$ for the combination model) and norflurazon ($P = 0.2544$ for CA, $P = 0.0661$ for IA, and $P = 0.0125$ for the combination model) were found. However, all three models exhibited significant positive trends ($P = 0.0240$ for CA, $P = 0.0130$ for IA, and $P < 0.0001$ for the combination model) when plotted against resource limitation (light) ratios.

Discussion

Method comparison

Using our methods the results for the *C. elegans* data were similar to that reported in Jonker et al. (2004). Our examination found that observed toxicity increased less with increasing TUs than predicted by either model indicating a synergistic response. The original study concluded that synergism occurred and these results support that conclusion. Description of any ratio dependent effects for thesedata was not reliable because regressions would have only described the relationship between two points: approximate TU ratios of 0.5 (twice as much zinc as copper) and 4.0 (four times as much copper as zinc).

The results for *F. candida*, using our methods were also similar to those of the original study (Jonker et al. 2005). On average, CA and IA model predictions fell within a factor of two of the observed data, suggesting that, like the original analysis, CA provided adequate fit for the observed responses and unlike the original analysis, that IA may provide a better fit. Deviations from model predictions did occur at elevated total TUs, indicating dose level dependent antagonism with the CA model only. Concentration addition antagonism was only slight because, despite that deviation, average MeDRs were still within a factor of two, and none exceeded an order of magnitude difference. Our analysis also found, like the original analysis, that dose ratio-dependent deviation from CA and IA occurred. Both studies indicate that toxicity decreases when toxicant ratios switch from being dominated by zinc to being dominated by cadmium. Both analyses showed that this deviation was an antagonistic response with increasing cadmium.

The previous analysis of *L. rubellus* responses only considered IA as a possible mixture model between fluoranthene and drought effects (Long et al. 2009). The results presented here indicate that IA fit the observed data similarly to CA as evidenced by MeDRs within a factor of two and no significant dose level dependent toxicity deviations. Though our analysis did indicate slight synergism for toxicant ratios with increased fluoranthene based on IA, CA only deviated from a factor of two at the highest relative fluoranthene concentration and this deviation was insignificant. The primary difference between our analysis and the original analysis was that CA was considered here. By including CA as a potential predictor model, we were able to demonstrate that the observed data was modeled similarly with CA and IA. CA predictions did not exhibit any dose ratio-dependent effects and, despite a slightly higher MeDR, better explained the mixture effects of fluoranthene and decreased soil moisture.

Our analyses of previously published datasets indicate many similarities between the two analytical methods. Differences in results between methods were few, minor in relevance (such

as the slight CA antagonism in the *F. candida* study), or the result of the original study's authors not considering both theoretical models. The latter point provides further support for the statements by Jonker et al. (2005) that researchers should "consider both concepts as equally valid alternatives." Our methods do not disprove the isobole and multiple linear regression mixture analysis method provided in Jonker et al. (2005). Instead, our methods provide a simpler and more parsimonious protocol to assess the many facets of mixture toxicity. By treating each aspect of mixture interaction as a separate question, these protocols allow model fit, dose level dependent toxicity, and dose-ratio dependent toxicity to be evaluated with summary statistics, scatter-plots, t-tests, and simple linear regressions.

Natural and anthropogenic stressors

Atrazine mixed with the natural stressor of resource limitation was best modeled by CA, a result supported by both the equipotent and full mixture analyses. Additionally, because high light levels coupled with high atrazine concentrations did not result in increased or synergistic toxicity, reactive oxygen species produced by uncoupling photosystem II play a minor role compared to resource limitation in inducing toxicity in this system. In fact, glucose addition can counteract the effects of atrazine on the algae *Chlorella vulgaris* without any atrazine induced cellular abnormalities (Ashton et al. 1966). Increased relative atrazine concentrations (dose ratio dependent toxicity) decreased MeDRs. This antagonism is likely a result of high light levels (high resource levels) overcoming the atrazine-dependent resource limitation and a lack of oxidative effects. A study by Cedergreen et al. (2005) found that for terbutylazine, a triazine herbicide similar to atrazine, that five of seven aquatic macrophyte species had decreased EC50s (increased toxicity) when exposed to 60% less irradiance during the 16h light period of two-week growth experiments. The previous study only demonstrated that resource limitation can contribute to toxicity, as found here, but did not quantify a wide range of stressor effects.

Norflurazon mixed with the natural stressor was intermediate to and modeled similarly by both models, based on mean MeDRs and MeDRs across the range of TU ratios. However, IA better predicted toxicity across the range of dose-levels and was slightly better at modeling this toxicant mixture. Because IA alone, without antagonism, best predicted these mixture effects, it is likely that the amount of photonic energy required to produce toxic levels of reactive oxygen species is smaller than the light treatments used here. Additionally, these data revealed that growth inhibition due to light limitation has a greater effect on the algal cells than decreased toxicity due to the reduction in energy available to produce reactive oxygen species or that the pesticide concentration, not the available energy, determines the amount of oxidative damage.

Overall, atrazine and norflurazon mixtures were modeled similarly with CA or IA, based on all metrics. However, the equipotent modeling indicated that the observed response was more similar to CA than IA. Both toxicants have both similar and dissimilar modes of action. Atrazine (Shimabukuro and Swanson 1969), like norflurazon (Ahrens 1994), increases the exposure of a cell to light fueled reactive oxygen species. Unlike norflurazon, atrazine also limits the ability of cells to generate usable energy from light. Because pesticides with similar modes of action tend to be modeled best with CA, then CA likely best models the shared mode of action (Belden et al. 2007). However, because a dissimilar mode of action is also present in this binary mixture, IA likely best models that aspect of toxicity. This mixture was modeled equally well by both models and the observed results were intermediate to both models because both models theoretically modeled different aspects of the same mixture.

Three mixture modeling combinations were used to model the tertiary mixture. Atrazine and resource limitation were best modeled by CA. The other two binary experiments suggested that both CA and IA would model the mixture similarly. The tertiary mixture modeling combinations were selected to reflect these results and were: CA only, IA only, and a combination model (atrazine and resource limitation combined using CA and that combination

then combined with norflurazon using IA). Equipotent combinations were most similar to the combination model, though fit was indistinguishable from the combination model and CA below one TU. The three models only differed when compared using overall accuracy or MeDRs, with the combination model as the only model with predicted values within a factor of two of observed values. All models over-predicted toxicity (antagonism) when resource limitation dominated the mixture. This antagonism was caused by the complex interaction of two resource limiting stressors with two oxidative damaging stressors, all of which were dependent on light levels.

Only one other study to date has used a natural stressor modeled with a chemical stressor in a mixture modeling framework. The study by Long et al. (2009) assumed that the natural stressor (drought) and the PAH fluoranthene acted independently when in a mixture without actually testing that assumption. This study is the first to demonstrate a natural stressor interacting with chemical stressors in both additive (CA) and independent (IA) manners. Though the stressors used in this study were chosen for their similar modes of action, other modes of action are known, but additive effects still adequately described toxic effects. Though drought-stress and PAH-induced narcosis may seem to be dissimilar modes of action, fluoranthene can be more toxic to the aquatic sediment dwelling midge, *Chironomus dilutus*, than would be expected based on narcosis alone (Schuler et al. 2006). Additionally, that increased toxicity was attributed to a fluoranthene metabolite with an unknown mode of action.

In addition to biotransformations of compounds that can change modes of action, organisms can be exposed to many stressors that have both anthropogenic and natural origins. For example, major ion salts can enter the aquatic environment and cause toxicity, but the source of that salt toxicity can be from natural sources like saltwater intrusion (Horrigan et al. 2007) or anthropogenic sources like oil field produced water (Fisher et al. 2010). Anthropogenic stressors can induce a chain of events that result in negative effects from traditionally natural stressors. For example, agricultural nutrient runoff, an anthropogenic stressor, reduces dissolved oxygen,

typically a natural stressor, in the Gulf of Mexico resulting in negative effects or toxicity to resident marine biota (Dodds 2006). Thus, natural stressors should be evaluated and modeled in the same way as anthropogenic chemical stressors. Standardizing the assessment of all stressors would allow hazard assessments to incorporate all types of stressors using specific effective concentrations like the LC50. As an example, eutrophic stream systems could be ranked using hazard quotients based on the oxygen saturation in those streams compared to the oxygen saturation LC50 for a sensitive or economically relevant local fish.

Conclusions

The mixture modeling and analysis protocol presented here, though based on the work of Jonker et al. (2005), differs significantly from those methods in three key ways. First, this protocol treats each possible deviation pattern as its own separate research question with its own separate analysis. The graphically based nature of these analyses simplifies both the understanding and communication of the generated results. Second, while MS Excel and SAS were used to generate and analyze data, the protocols presented here are simple and robust enough to be adapted to any statistical analysis program or regression capable spreadsheet program, alone or in combination. This allows users to generate results using software with which they are already familiar and forgoes the need to introduce and interpret a new suite of analytical variables. Finally, the methods presented here allow the scientist greater control over modeling and data analysis process. In this method, the user can, like the previous method, test for CA or IA combinations of toxicants or both at the same time (unlike the previous method) for mixtures containing more than two components. Those mixtures with more than two components can have some toxicants combined using CA and then combine those CA combinations with any number of IA combinations, regardless of relative constituent potency.

This study demonstrated that the derived techniques presented here adequately modeled and analyzed mixture toxicity. While equipotent mixture comparisons agreed with full mixture analyses, the full dynamic of toxic response was only distinguishable using the full, non-equipotent mixture results. These results suggest that mode of action is vital to predicting how two stressors will jointly affect an organism. The modes of action of stressors, not sources, determine how an organism will react when exposed to a mixture of stressors.

Acknowledgements

We thank J. Crawford for assistance in the laboratory. This manuscript was improved using valuable comments from K. Baum and D. Shoup.

Literature Cited

Ahrens, W.H., ed. 1994. Herbicide Handbook, seventh edition. Weed Science Society of America, Champaign, IL, USA.

Altenburger, R., T. Backhaus, W. Boedeker, M. Faust, M. Scholze, and L.H. Grimme. 2000. Predictability of the toxicity of multiple chemical mixtures to *Vibrio fischeri*: mixtures composed of similarly acting chemicals. Environmental Toxicology and Chemistry 19:2341-2347.

Altenburger R, Nendza M, Schuurmann G. 2003. Mixture toxicity and its modeling by quantitative structure–activity relationships. Environmental Toxicology and Chemistry 22:1900–1915.

Ashton, F.M., T. Bisalputra, and E.B. Risley. 1966. Effect of atrazine on *Chlorella vulgaris*. American Journal of Botany 53: 217-219.

Backhaus, T., R. Altenburger, W. Boedeker, M. Faust, M. Scholze, and L.H. Grimme. 2000. Predictability of the toxicity of a multiple mixture of dissimilarly acting chemicals to *Vibrio fischeri*. Environmental Toxicology and Chemistry 19:2348-2356.

Belden, J.B. and M.J. Lydy. 2006. Joint toxicity of chlorpyrifos and esfenvalerate to fathead minnows and midge larvae. Environmental Toxicology and Chemistry 25:623-629.

Belden, J.B., R.G. Gilliom, and M.J. Lydy. 2007. How well can we predict the toxicity of pesticide mixtures to aquatic life? Integrated Environmental Assessment and Management 3:364-372.

Cedergreen, N. and J.C. Streibig. 2005. The toxicity of herbicides to non-target aquatic plants and algae: assessment of predictive factors and hazard. Pest Management Science 61:1152-1160.

Dodds, W.K. 2006. Nutrients and the "dead zone": the link between nutrient ratios and dissolved oxygen in the Gulf of Mexico. Frontiers in Ecology and the Environment 4:211-217.

Fisher, J.C., J.B. Belden, and J.R. Bidwell. 2010. Can site-specific heuristic toxicity models predict the toxicity of produced water? Chemosphere 80:542-547.

Horrigan, N., J.E. Dunlop, B.J. Kefford, and F. Zavahir. 2007. Acute toxicity largely reflects the salinity sensitivity of stream macroinvertebrates derived using field distributions. Marine and Freshwater Research 58: 178-186.

Jonker, M.J., C. Svendsen, J.J.M Bedaux, M. Bongers, and J.E. Kammenga. 2005. Significance testing of synergistic/antagonistic, dose level-dependent, or dose ratio-dependent effects in mixture dose-response analysis. Environmental Toxicology and Chemistry 24: 2701-2713.

Jonker, M.J., R.A.J.C. Sweijen, and J.E. Kammenga. 2004. Toxicity of simple mixtures to the nematode *Caenorhabditis elegans* in relation to soil sorption. Environmental Toxicology and Chemistry 23: 480-488.

Kuwabara, K., A. Nakamura, and T. Kashimoto. 1980. Effect of petroleum oil, pesticides, PCBs and other environmental contaminants on the hatchability of *Artemia salina* dry eggs. Bulletin of Environmental Contamination and Toxicology 25:69-74.

Lam, G.K.Y. 1989. Analysis of interaction for mixtures of agents using the linear isobole. Bulletin of Mathematical Biology 51:293-309.

Long, S.M., F. Reichenberg, L.J. Lister, P.K. Hankard, J. Townsend, P. Mayer, J. Wright, M. Holmstrup, C. Svendsen, and D.J. Spurgeon. 2009. Combined chemical (fluoranthene) and drought effects on *Lumbricus rubellus* demonstrate the applicability of the independent action model for multiple stressor assessment. Environmental Toxicology and Chemistry 28: 629-636.

Lydy, M., J. Belden, C. Wheelock, B. Hammock, and D. Denton. 2004. Challenges in regulating pesticide mixtures. Ecology and Society 9: 1.

Pelletier, E., P. Sargian, J. Payet, and S. Demers. 2006. Ecotoxicological effects of combined UVB and organic contaminants in coastal waters: a review. Photochemistry and Photobiology 82:981-993.

Schaebenberger, O., B.E. Tharp, J.J. Kells, and D. Penner. 1999. Statistical tests for hormesis and effective dosages in herbicide dose response. Agronomy Journal 91: 713-721.

Shimabukuro, R.H. and H.R. Swanson. 1969. Atrazine metabolism, selectivity, and mode of action. Journal of Agricultural and Food Chemistry 17:199-205.

Schuler, L.J., P.F. Landrum, and M.J. Lydy. 2006. Comparative toxicity of fluoranthene and pentachlorobenzene to three freshwater invertebrates. Environmental Toxicology and Chemistry 25:985-994.

U.S. Environmental Protection Agency. 2002. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. EPA/821/R-02/013. Washington, DC.

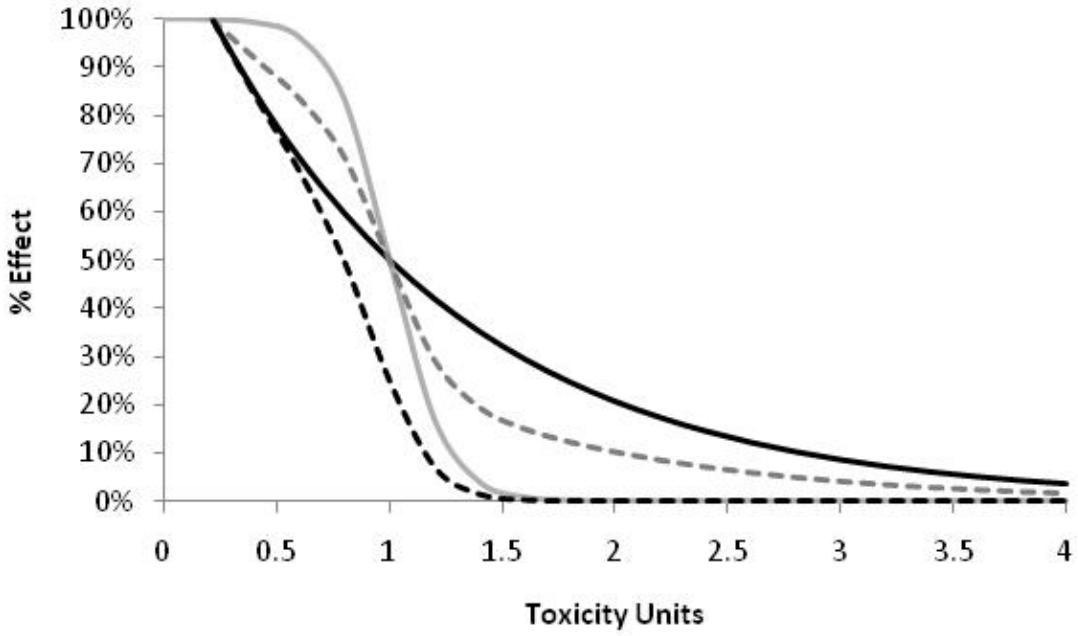

Figure 1. Dose response curves for a theoretical binary mixture. The two solid lines are the individual component dose response curves. The dashed lines are equitoxic mixture effect predictions based on CA (gray line) and IA (black line).

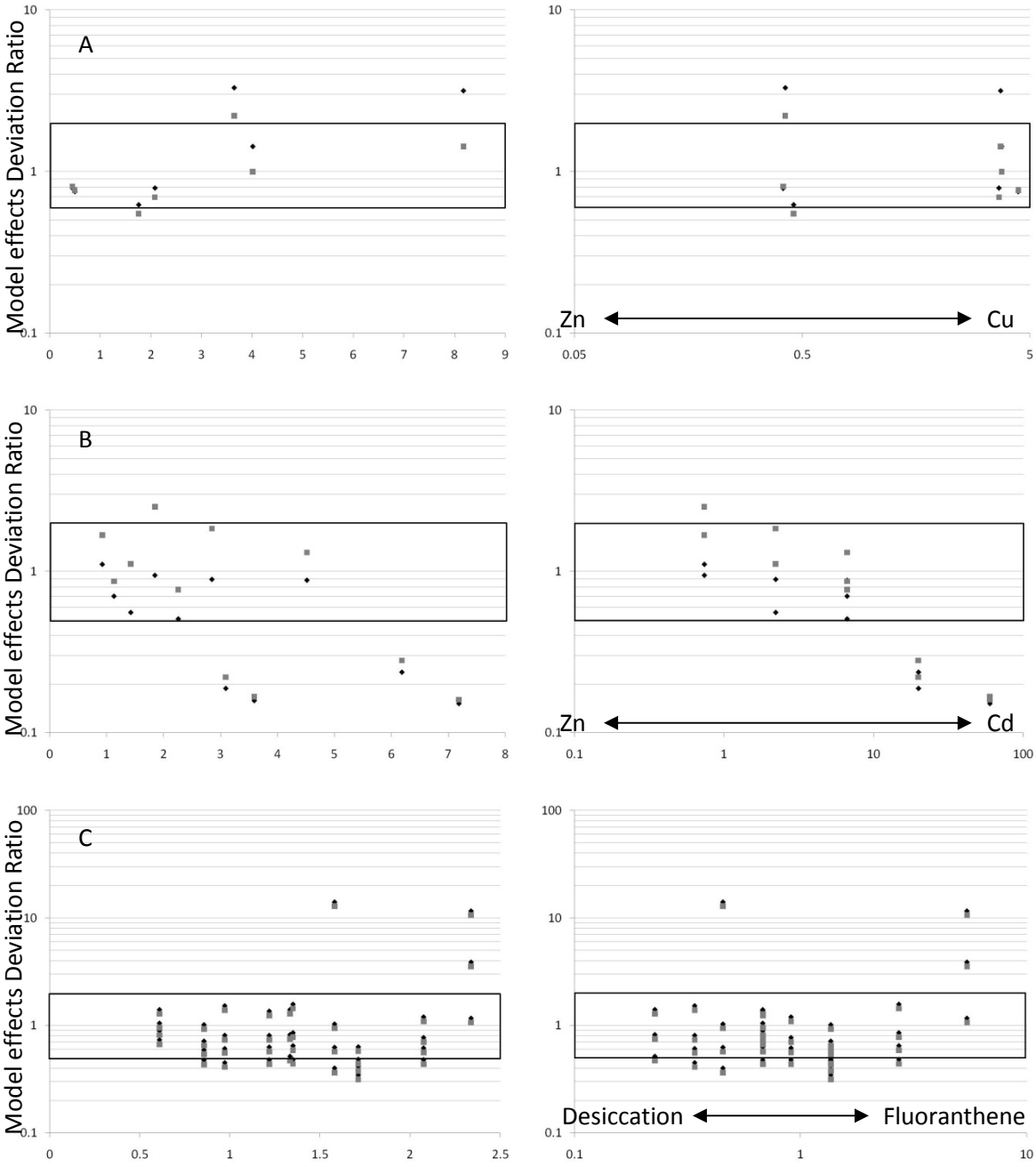

Figure 2. Mixture model fit plots for *C. elegans* (Cu and Znrow A, Jonker et al. 2004), *F. candida* (Cd and Zn, row B, Jonker et al. 2005), and *L. rubellus* (fluoranthene and desiccation, row C, Long et al. 2009). Model fit (model effects deviation ratios) are plotted against total toxicity (left column) and against the ratio of toxicants in the mixtures (right row). MeDRs are black diamonds for concentration addition and gray squares for independent action. MeDRs within a factor of two (0.5 to 2.0) are enclosed in black boxes. Deviations above a factor of two indicate synergism, while those below indicate antagonism.

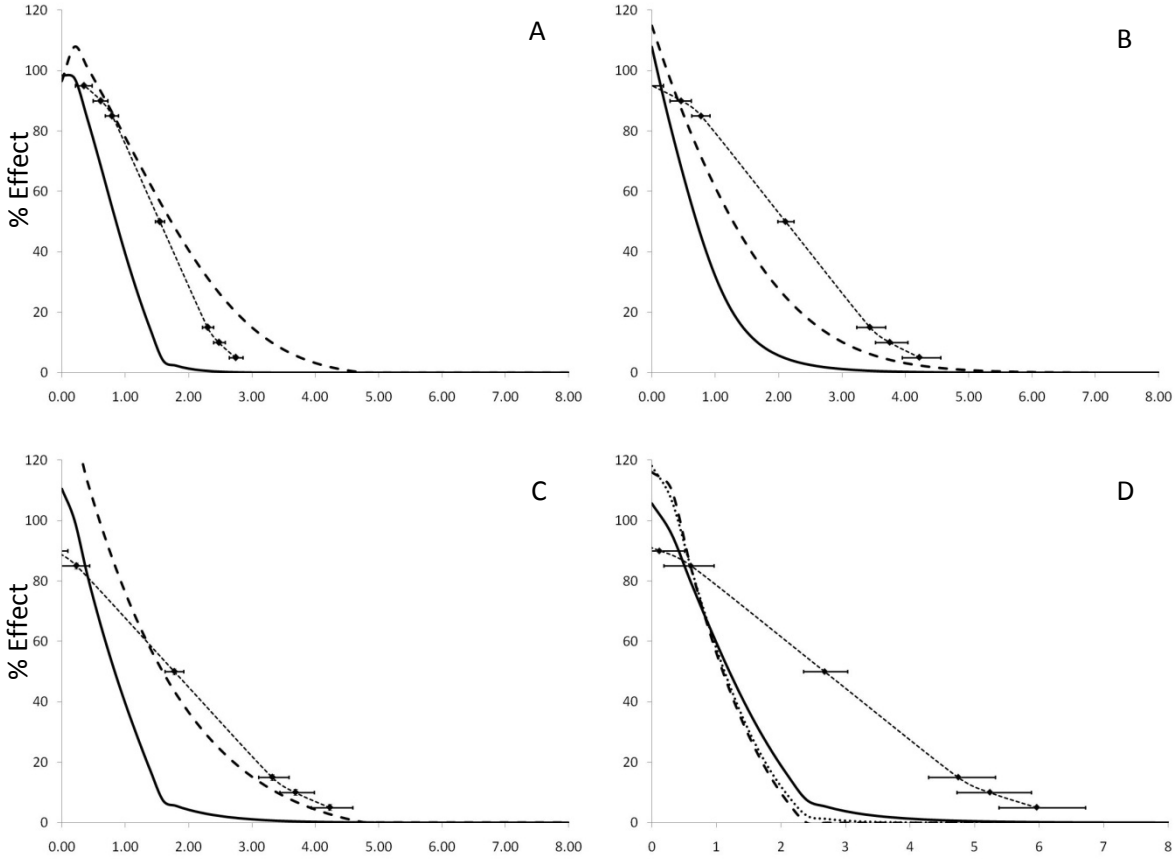

Figure 3. Equipotent mixture toxicity for atrazine and resource limitation (graph A), norflurazon and resource limitation (graph B), atrazine and norflurazon (graph C), and all three toxicants (graph D). Mixture modeling predictions are represented by a solid line for CA, a large dashed line for IA, and a dotted line for the model using CA for atrazine and resource limitation and IA for norflurazon with the CA model. Fitted data (small dashed line) are percent control algal growth. Points along the fitted data line are specific effective concentrations (5, 10, 15, 50, 85, 90, and 95% effect) with 95% confidence interval error bars.

Table 1. Comparison of MDR, based on EC50s, and mean MeDR values for each equipotent stressor combination. Standard error of the mean MeDRs are presented in parentheses next to each MeDR. The combination model using both CA and IA is abbreviated as Comb.

	MDR			MeDR		
Mixture	CA	IA	Comb.	CA	IA	Comb.
Atrazine and Resource Limitation	0.55	1.12	-	0.62 (0.03)	2.12 (0.15)	-
Norflurazon and Resource Limitation	0.33	0.61	-	0.71 (0.07)	0.89 (0.07)	-
Atrazine and Norflurazon	0.47	0.90	-	0.54 (0.04)	0.86 (0.06)	-
All three stressors	0.45	0.41	0.42	0.40 (0.06)	0.98 (0.13)	0.24 (0.04)

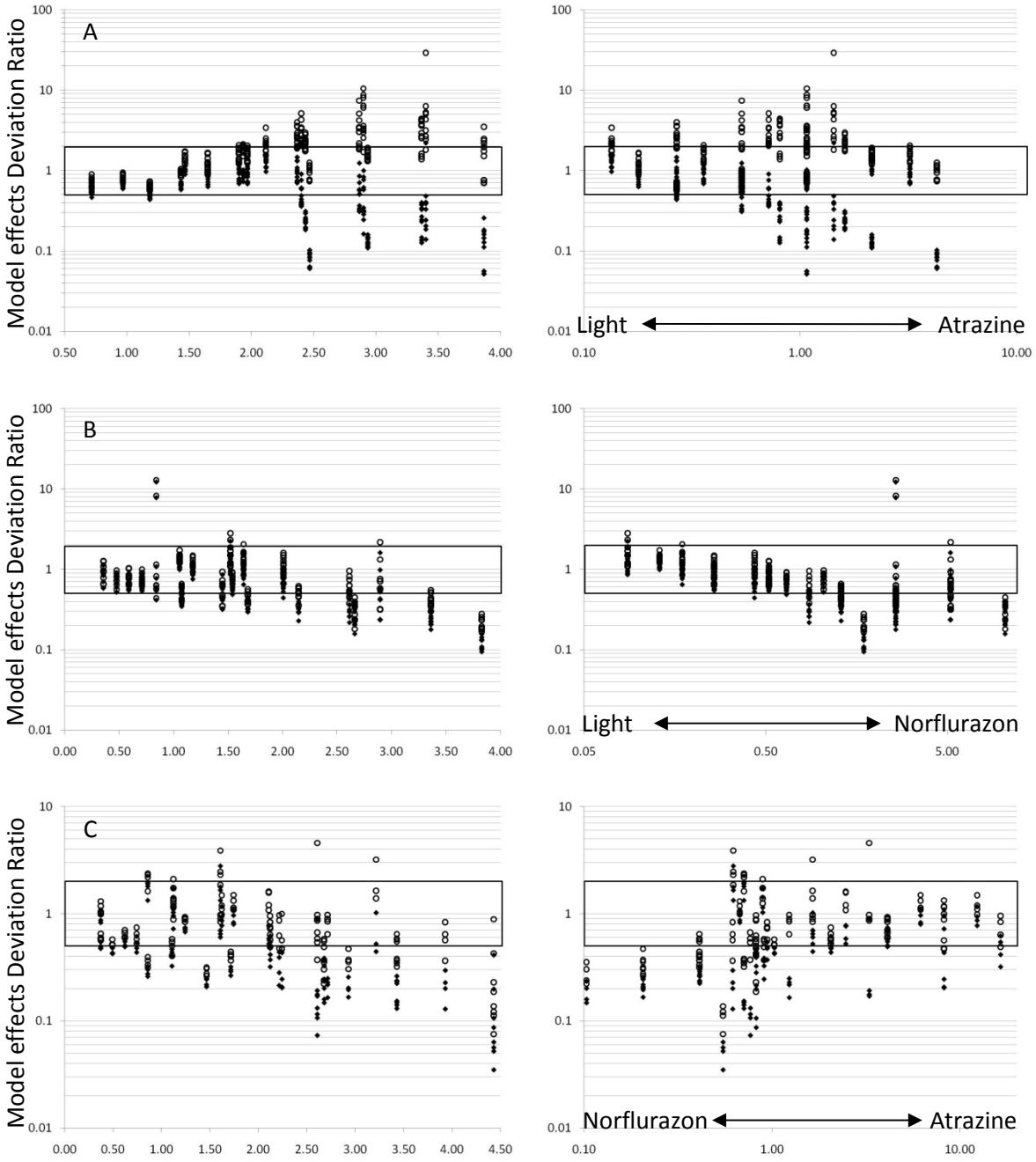

Figure 4. Mixture model fit plots for atrazine and light (row A), norflurazon and light (row B), and atrazine and norflurazon (row C). Model fit (model effects deviation ratios) are plotted against total toxicity (left column) and against the ratio of toxicants in the mixtures (right column). MeDRs are black diamonds for concentration addition and open circles for independent action. MeDRs within a factor of two (0.5 to 2.0) are enclosed in black boxes. Deviations above a factor of two indicate synergism, while those below indicate antagonism.

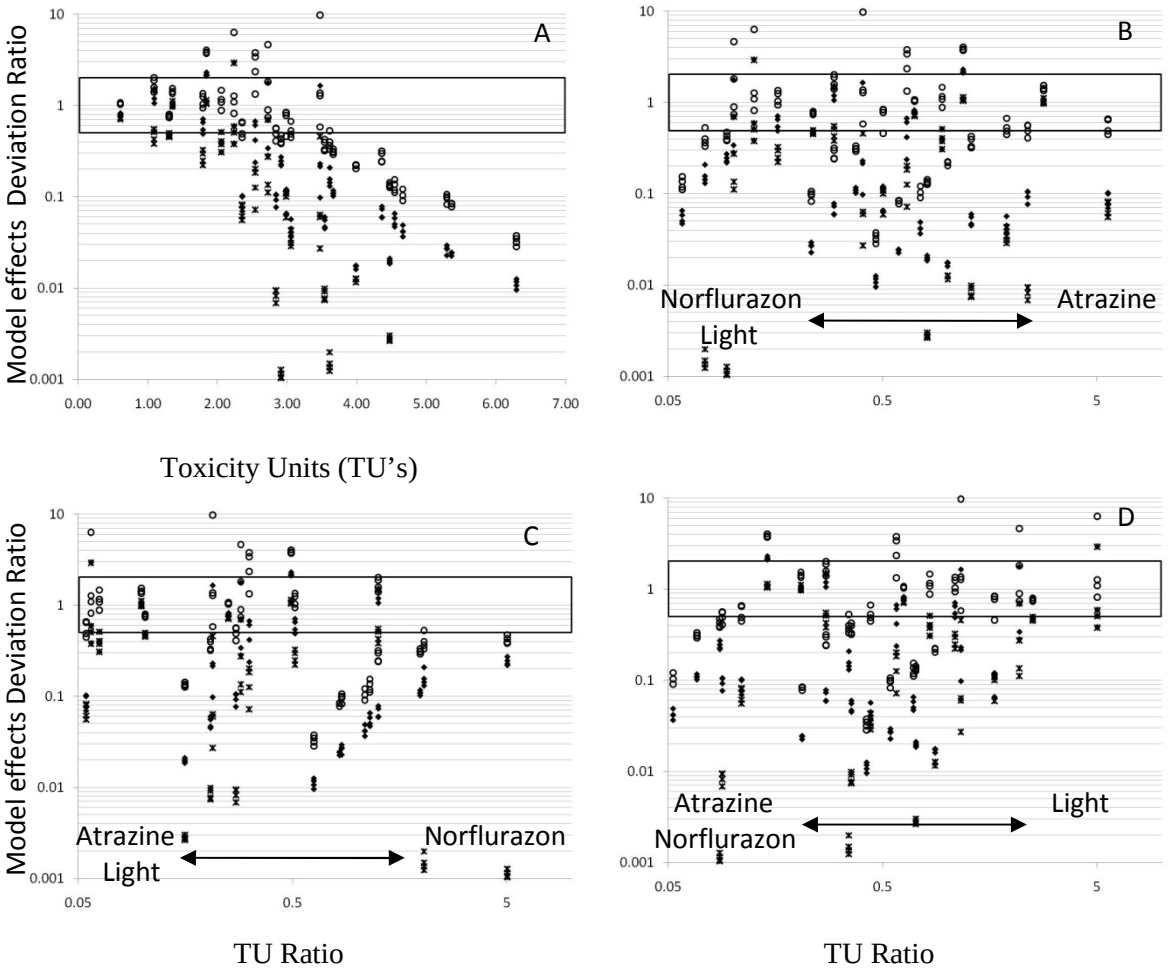

Figure 5. Mixture model fit plots for atrazine, norflurazon, and light. Model fit (model effects deviation ratios) are plotted against total toxicity (graph A) and against the ratio of toxicants in the mixtures (graphs B, C, and D). MeDRs are black diamonds for concentration addition, open circles for independent action, and black asterisks for the model incorporating both CA and IA mixing. MeDRs within a factor of two (0.5 to 2.0) are enclosed in black boxes.

VITA

Jonathan C. Fisher

Candidate for the Degree of

Doctor of Philosophy

Thesis: LABORATORY TOXICITY AND FIELD EFFECTS OF A COMPLEX MIXTURE: OIL-FIELD PRODUCED WATER

Major Field: Zoology

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Zoology at Oklahoma State University, Stillwater, Oklahoma in July, 2010.

Completed the requirements for the Master of Science in Fisheries at Louisiana State University, Baton Rouge, LA in 2005.

Completed the requirements for the Bachelor of Science in Agriculture at Louisiana State University, Baton Rouge, LA in 2003.

Experience:

Assistant Professor of Biology, Aquatic Toxicology, Northeastern State University, Tahlequah, OK, August 2010.

Professional Memberships:

Society for Environmental Toxicology and Chemistry

Name: Jonathan C. Fisher Date of Degree: July, 2010

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: LABORATORY TOXICITY AND FIELD EFFECTS OF A COMPLEX MIXTURE: OIL-FIELD PRODUCED WATER

Pages in Study: 87 Candidate for the Degree of Doctor of Philosophy

Major Field: Zoology

Abstract

This dissertation investigated how organisms in the field and the laboratory responded to complex mixtures or combinations of stressors. Organisms are continually exposed to natural and anthropogenic stressors in the environment. The first two chapters of this study examined the potential for and measured aquatic effects of a near-shore petroleum operation along Skiatook Lake, Osage County, Oklahoma, USA. The field based part of this study showed that produced water from a near-shore petroleum production operation travelling through the groundwater into the lake was sufficiently diluted by the lake to result in few observable biotic effects. For the laboratory based part of this study, an empirically derived model of major ion toxicity was combined with other toxicity assessments to account for the observed toxicity in field-collected produced water and produced-water contaminated groundwater. The accuracy and precision of the ion toxicity model, calculated using model deviation ratios (MDR) and simple linear regressions, was determined for fathead minnows, *Ceriodaphnia dubia*, and *Daphnia magna*. Those results indicated that predicted major ion toxicity, combined with other measured parameters, could accurately and precisely account for observed responses in test organisms to field collected samples. The last chapter of this study expanded on current mixture modeling and analysis techniques in order to assess the effects of mixtures containing chemical and nonchemical stressors, any number of stressors at any relative potency, and for any combination of similar and dissimilar modes of action. New techniques were derived and then used to assess the joint action of atrazine, norflurazon, and resource limitation (light limitation). Mixture effects of a tertiary mixture all three stressors were predicted within a factor of two of the observed data using a mixture model that incorporated both concentration addition and independent action.

ADVISER'S APPROVAL: Dr. Jason B. Belden