Government of Canada | Gouvernement du Canada

National Agri-Environmental Standards Initiative (NAESI)

Synthesis Report No. 7

Environmental Risk-Based Standards for Pesticide Use in Canada

2009

Photos:
Bottom Left- clockwise
Fraser Valley near Abbotsford, B.C.: Wayne Belzer, Pacific Yukon Region, Environment Canada
Crop spraying: Corel CD photo # 95C2840
Elk Creek, BC: Joseph Culp, National Water Research Institute, Environment Canada
Prairie smoke and bee: Emily Wallace, Prairie Northern Region, Environment Canada

This report can be cited as follows:

Mineau, P., T. Dawson, M. Whiteside, C. Morrison, K. Harding, L. Singh, T. Längle, and D.A.R. McQueen. 2009. *Environmental Risk-Based Standards for Pesticide Use in Canada*. National Agri-Environmental Standards Initiative Synthesis Report No. 7. Environment Canada. Gatineau, Quebec. 94 p.

Prepared and published by
Environment Canada
Gatineau, QC

March 2009

ISBN 978-0-662-06037-6
Cat. No. En4-95/2008-MRC

NATIONAL AGRI-ENVIRONMENTAL STANDARDS INITIATIVE

ENVIRONMENTAL RISK-BASED STANDARDS FOR PESTICIDE USE IN CANADA

SYNTHESIS REPORT NO. 7

NOTE TO READERS

The National Agri-Environmental Standards Initiative (NAESI) was a four-year (2004-2008) project between Environment Canada (EC) and Agriculture and Agri-Food Canada (AAFC) and one of many initiatives under AAFC's Agriculture Policy Framework (APF). The goals of NAESI included:

- Establishing non-regulatory national environmental performance standards (with regional application) that support common EC and AAFC goals for the environment;
- Evaluating standards attainable by environmentally-beneficial agricultural production and management practices; and
- Increasing understanding of relationships between agriculture and the environment.

Under NAESI, agri-environmental performance standards (i.e., outcome-based standards) were developed that identify both desired levels of environmental condition and levels considered achievable based on available technology and practice under four thematic areas: Air, Biodiversity, Pesticides, and Water. The development of these performance standards involved both science-based assessments of relative risk and the determination of desired environmental quality.

Outcomes from NAESI contribute to the broader APF goals of improved stewardship by agricultural producers of land, water, air and biodiversity and increased Canadian and international confidence that food from the Canadian agriculture and food sector is being produced in a safe and environmentally sound manner. Development of the standards was led by EC in close partnership with AAFC resulting in science-based products that can be integrated into beneficial agricultural management systems and practices to help reduce environmental risks.

This synthesis report is one of sixteen final deliverables for NAESI and summarizes the research undertaken over the past 4 years. See next page for complete listing of the sixteen final deliverables. Each of the fifteen technical synthesis reports outlines the methodology used to develop the standards, the proposed standards themselves and the scientific rationale behind them, limitations of the proposed standards and gaps in the science, and discusses future considerations with respect to the recommended standards. To ensure scientific integrity these fifteen synthesis reports each underwent an intensive peer-review process involving three unbiased subject matter experts who were not directly affiliated with NAESI or the Government of Canada. An overarching synthesis report accompanies the fifteen technical reports and provides background information about the history of the program as well as a summary of the final results highlighting how these results can inform decision-making.

For additional information regarding this program, please contact:

Environment Canada	Gatineau, QC
National Agri-Environmental Standards	K1A 0H3
Initiative Secretariat	Phone: (819) 997-1029
351 St. Joseph Blvd., 14th floor	Fax: (819) 994-9848

COMPLETE LIST OF FINAL NAESI DELIVERABLES

Bowerman, M., L. Maclean, S. Villeneuve, and E.S. Roberts. 2009. National Agri-Environmental Standards Initiative Overarching Report.

Lillyman, C., K. Buset, G. Beaney, and D. Mullins. 2009. The Development of Regionally Based Performance Standards for Agricultural Ammonia Emissions in the Context of Particulate Matter Formation in Canada.

McPherson, M., C. Nielsen, and K. Proudlock. 2009. The Development of Tier 1 Generalized Habitat-based Standards for Ecozones in Agricultural Regions of Canada.

Neave, E., D. Baldwin, and C. Nielsen. 2009. Tier 2 and 3 Standards – Developing Landscape-specific, Habitat-based Standards Using Multiple Lines of Evidence.

Demers, M.J. and P.B. Jiapizian. 2009. The Development of Ideal Performance Standards for Pesticides.

Rousseau, A.N., P. Lafrance, M-P. Lavigne, S. Savary, B. Konan, and R. Quilbé. 2009. The Development of Agri-Environmental Achievable Performance Standards (APS) for Pesticides at the Watershed Level.

Mineau, P., T. Dawson, M. Whiteside, C. Morrison, K. Harding, L. Singh, T. Längle, and D.A.R. McQueen. 2009. Environmental Risk-Based Standards for Pesticide Use in Canada.

Murphy, C.S., J.P. Mutch, and M.L. Williamson. 2009. The Development of Commodity-Based Ideal Performance Standards for Pesticides Used in Potato Production in Canada.

Williamson, M.L. and N. Glozier. 2009. The Development of Ideal Performance Standards for Pesticide Mixtures to Protect Aquatic Ecosystems: A Prairie Waterbody Case Study.

Ernst, B., R.E. Mickle, S. Howatt, R. Pitbaldo, I. Nichols, T. Rowlandson, and P. MacKinnon. 2009. The Development of Meteorological Standards for Pesticide Application in Canada.

Chambers, P.A., M. Guy, S.S. Dixit, G.A. Benoy, R.B. Brua, J.M. Culp, D. McGoldrick, B.L. Upsdell, and C. Vis. 2009. Nitrogen and Phosphorus Standards to Protect the Ecological Condition of Canadian Streams, Rivers and Coastal Waters.

Guy, M. 2009. Ideal Performance Standards to Prevent Nitrate Toxicity.

Culp, J.M., G.A. Benoy, R.B. Brua, A.B. Sutherland, and P.A. Chambers. 2009. Total Suspended Sediment, Turbidity and Deposited Sediment Standards to Prevent Excessive Sediment Effects in Canadian Streams.

Edge, T.A., I. Droppo, A. El-Shaarawi, V. Gannon, M. Hewitt, R. Kent, I. Khan, W. Koning, D. Lapen, D. Marcogliese, C. Marvin, J. Miller, N. Neumann, R. Phillips, W. Robertson, H. Schrier, I. Shtepani, E. Topp, and E. van Bochove. 2009. An Evaluation of *Escherichia coli* as a Potential Agri-Environmental Waterborne Pathogen Standard.

Baird, D.J., D.L. Peters, R.A. Curry, N. Horrigan, W.A. Monk, and D.E. Tenenbaum. 2009. Establishing Standards and Assessment Criteria for Instream Flow Needs in Agricultural Watersheds in Canada.

Pellerin, P., A. Pietroniro, S. Bélair, V. Fortin, D. Charpentier, B. Bilodeau, I. Doré, J. Töyrä, M. Carrera, B. Davison, B. Toth, and S. Marin. 2009. The Development of Water-Balance Indicators in an Agriculture-Dominated Watershed Using Coupled Hydrometeorological Modeling and Data Assimilation.

KEY FINDINGS

The present report assembles a sufficient empirical database of terrestrial and aquatic field studies to translate laboratory-based toxicity (hazard) indices into an actual probability of impact for most key environmental sectors potentially affected by pesticide use. The passage from a hazard-based assessment of individual pesticides or pesticide use patterns to a risk-based one is a critical step in getting the agricultural community to accept the proposed standards. Only by possessing a reasonable estimate of the actual impacts of pesticides can the risks and benefits of different agricultural interventions, including pesticide use, be weighted against each other.

The setting of standards is a risk-management decision. The standards proposed here are based on a mixture of ecological theory and government mandate, infused with a healthy amount of pragmatic realism borne of the authors' experience with pesticide regulatory systems and agencies. Indeed, where possible, our standards are compared with standard regulatory assessments currently being applied by major regulatory agencies. Whether the standards are considered too stringent or not protective enough, it remains that the present report presents what the authors believe to be the best measurement instruments currently available to gauge the relative environmental impacts of different pesticide treatments. The use of these tools to compare different pesticide treatments transcends the setting of standards and has a clear connection to the area of agricultural indicators and reporting as well as risk reduction.

The table below summarizes the different standards being proposed in this consolidation report.

Summary of risk indices and proposed standards developed under the NAESI program and presented in this report.

Terrestrial risk model	Type of application	Proposed NAESI standard	Impact measures so far below standard that immediate product substitution or effective mitigation is recommended*
Acute Avian	Spray (liquid)	The standard for avian acute toxicity and spray applications is based on the probability that a pesticide will cause detectable mortality in 10% or more of treatments based on a series of empirical models.	The probability that a pesticide will cause detectable mortality in 50% or more of treatments based on a series of empirical models.
	Seed treatment	The standard for avian acute toxicity and seed-treatment applications is based on products having the ability to kill songbirds at the 5% tail of sensitivity distribution with a risk index of 0.1, corresponding to the ingestion of 20 seeds of a preferred seed type or less. This is a provisional standard in the absence of data on seed treatment products of intermediate toxicity.	Products having the ability to kill songbirds at the 5% tail of sensitivity distribution with a single seed of a preferred seed type.
	Granular treatment	The standard for avian acute toxicity and granular applications is based on products having the ability to kill songbirds at the 5% tail of sensitivity and an adjusted risk of 0.1 or higher, in parallel with the proposed seed-treatment standard. This is a provisional standard in the absence of data on granular products of intermediate toxicity.	Products having the ability to kill songbirds at the 5% tail of sensitivity distribution with a single granule.

Terrestrial risk model	Type of application	Proposed NAESI standard	Impact measures so far below standard that immediate product substitution or effective mitigation is recommended*
Avian reproductive	Spray (liquid)	The provisional standard for the chronic (reproductive) toxicity to birds of pesticides applied as spray is based on the product being above environmental levels estimated capable of causing reproductive dysfunction in sensitive 15-g songbirds for one-third or more of the duration of the normal breeding season.	A product being above environmental levels estimated capable of causing reproductive dysfunction in sensitive 15-g songbirds for the duration of the normal breeding season.
	Seed treatment	No standards set	A product capable of causing reproductive dysfunction in sensitive 15-g songbirds with one seed or less per day.
	Granular treatment	No standards set	A product capable of causing reproductive dysfunction in sensitive 15-g songbirds with one granule or less per day.
Small mammal	Spray (liquid)	The standard for small mammals exposed to spray applications is based on the probability that a pesticide will cause a population decline in 10% or more of treatments based on empirical models.	Pesticide applications with a probability of 75% or higher of causing a population decline based on empirical models.
	Seed treatment	The standard for mammalian toxicity and seed-treatment applications is based on products having the ability to kill a small mammal at the 5% tail of sensitivity distribution with a risk index of 0.1, corresponding to the ingestion of 10 seeds of any type. This is a provisional standard in the absence of data on seed treatment products and mammal impacts.	Products having the ability to kill small mammals at the 5% tail of sensitivity distribution with a single seed.

Terrestrial risk model	Type of application	Proposed NAESI standard	Impact measures so far below standard that immediate product substitution or effective mitigation is recommended*
	Granular treatment	The standard for mammalian acute toxicity and granular applications is based on products having the ability to kill a small mammal at the 5% tail of sensitivity at a risk of 0.1 or higher in parallel with the proposed seed-treatment standard. This is a provisional standard in the absence of any data on the impact of granular pesticides on small mammals.	Products having the ability to kill small mammals at the 5% tail of sensitivity distribution with a single granule.
Non-target arthropods (honeybee)	Spray (liquid)	The terrestrial invertebrate standard based on honeybee toxicity will be that applications to areas frequented by natural pollinators should not exceed a hazard ratio of 50 corresponding to a calculated risk score of 0.33. This risk score is computed by comparing the log HR value of the application to a theoretical upper bound of a logHR of 5.	Applications to areas frequented by natural pollinators should not exceed a hazard ratio of 400. This corresponds to a risk score above 0.5 as computed here.
Edaphic invertebrates (earthworms)	Spray (liquid)	The proposed standard for earthworms is that pesticide applications should not be predicted to give rise to a greater than 35% loss of earthworm numbers based on the empirical model developed.	There should not be a greater than 65% loss of earthworm numbers based on the empirical model developed.

Aquatic risk model	Type of application	Proposed NAESI standard	Impact measures so far below standard that immediate product substitution or effective mitigation is recommended*
Crustacea species diversity	Spray (liquid)	To meet the aquatic invertebrate standard, pesticide applications must not result in water concentrations that lead to more than a 20% population decline in the most sensitive model (Copepoda or Cladocera) and must not result in more than 20% of all crustacean species exhibiting significant population declines.	Pesticide applications that result in water concentrations that lead to more than a 50% population decline in the most sensitive model (Copepoda or Cladocera) and in more than 50% of all crustacean species exhibiting significant population declines.
Cladocera abundance	Spray (liquid)		
Copepoda abundance	Spray (liquid)		
Algae abundance	Spray (liquid)	To meet the algal standard for aquatic ecosystems, pesticide applications must not give rise to water concentrations that lead to 20% or more of algal species being significantly affected by treatment.	Pesticide applications must not give rise to water concentrations that lead to a greater than 50% or more of algal species being significantly affected by treatment.
Fish exposure – toxicity ratio	Spray (liquid)	A provisional fish standard is proposed based on the U.S. EPA record of pesticide fish kills. Pesticide active ingredients will be considered to have met the standard if their relative risk to fish (calculated from a risk quotient based on exposure modeling and a fish HC_5 value) is such that all pesticides of equal or lesser hazard are responsible for no more than 10% of all fish kills recorded by the U.S. EPA.	Pesticide active ingredients will be considered for red-listing if their relative risk to fish (calculated from a risk quotient based on exposure modeling and a fish HC_5 value) is such that all pesticides of equal or greater hazard are responsible for more than 50% of all fish kills recorded by the U.S. EPA.

**A three-tier assessment not unlike the universally recognized traffic light system—i.e., green, yellow, red—is being introduced in this report. This labelling scheme will be mentioned repeatedly in the text and used throughout the technical report (Mineau et al., 2008b). "Green-listed" pesticide applications are those that appear to meet an ideal standard of environmental performance for a specific assessed sector. "Yellow-listed" pesticide applications require caution and likely mitigation because they do not appear to meet an ideal standard. "Red-listed" pesticide applications are considered to be so far below standard that product substitution may be the most logical mitigative action. These applications are prime candidates for action under a risk reduction program. This three-level assessment is a binning process. Throughout the text and tables, actual scores, typically expressed as a probability of impact or probability of breaching a very high risk level, are also given for finer tracking of risk*

Having developed these various measures, it is tempting to generate aquatic and terrestrial risk scores for all pesticides registered in Canada. By necessity, products have been compared when used at maximum on-label rate and under standardized conditions, as identified in the various relevant sections of this report. Nevertheless, the comparison should provide an interesting snapshot of existing registered products and help identify those products most in need of replacement or serious mitigation. These tables have been produced as a separate technical report (Mineau *et al*., 2008b). Also, Mineau *et al*. (2008b) shows how comprehensive pesticide-use data, in the form of a recent survey carried out by Statistics Canada for the Pesticide Risk Reduction Program of Agriculture and Agri-Food Canada, can be converted to comprehensive risk scores to assess whether commodity-wide use of pesticides meets the standards proposed herein.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Environment Canada has been tasked with developing environmental standards for implementation under Agriculture and Agri-Food Canada's Agricultural Policy Framework. Setting a standard of acceptability for pesticide use, whether an ideal one or one that is merely currently achievable (Caux and Jiapizian, 2004), implies that we can objectively measure the impact of pesticides on key environmental components. Whereas it may be possible to 'cherry-pick' a few absolute standards of good agricultural practice (e.g., no pesticide application should lead to a large fish or bird kill), many pesticide impacts may be much more subtle or graded in nature. By definition, pesticides carry an inherently high risk to some segments of the environment and choosing the right product often becomes a question of trading-off risk in one environmental component for another.

Measurement-based standards of water quality, in the form of allowable residual concentrations, will be proposed by the National Agri-Environmental Standards Initiative (NAESI) based on an approach that differs from this one (e.g., CEI, 2005). One drawback of measurement-based standards is that not all pesticides are equally well covered by water sampling. In addition, residue sampling is time- and resource-intensive and, more importantly, restricted to a subset of pesticides for which analytical methods and fate characteristics make sampling possible. Also, not all environmental impacts are mediated through movement of the chemical into water. Impacts on terrestrial biota most often occur through exposure pathways that differ from those responsible for the contamination of water bodies. The risk-based standards we propose here are intended to allow a more complete accounting of the environmental footprint of pesticides used on Canadian cropland.

Even though pesticide regulatory systems tend to evaluate pesticides singly based on their own

"merit" (product-specific registration), there have been numerous efforts to devise relative assessment schemes. Depending on the intended use of the resulting metrics, assessment schemes have been called hazard rankings, yardsticks, indices, indicators, screening benchmarks, relative risk rankings, risk assessment tools and more. Levitan (2000) proposed a typology that distinguishes among these approaches depending on whether they are intended for grower decision-support systems for "ecolabelling" purposes or to provide governments and others with an estimate of ecological damage. She distinguished "indicators" from "impact assessment systems." Indicators tend to be succinct summaries and integrations of various trends or highlights that evolve within a framework of policy analysis and risk communication. Impact assessment systems or "tools," on the other hand, retain the ecosystemic perspective along with the depth and complexity that are appropriate to the level of knowledge of a particular environmental component. Levitan's typology has not been uniformly accepted and therefore most of the European literature, for example, refers to "indicators" regardless of the structure, scientific rigour or intended function of the calculated risk index. Risk assessment systems can provide insight not always available from evaluations carried out by pesticide registration authorities. The latter typically consider pesticides singly (product-specific registration procedures), and often make registration decisions under imprecise concepts of risk and benefit.

The standards we propose are based on risk-assessment tools we have developed for different segments of the environment. They were developed after consideration of many other tools and indicators previously advanced (Mineau and Whiteside, 2005), and in light of risk-assessment procedures in the U.S., Canada and the European Union (E.U.). Our proposed assessment tools vary in structure depending on the available information. The standards themselves go one step further than other assessment tools by offering a proposed threshold of environmental

acceptability for each of the measurements. Where the hard data were lacking to adequately set such thresholds, we propose provisional standards. As mentioned above, pesticides are designed to kill either plants, fungi or animals. Strikingly rare are products that have perfect specificity for the pest organism—arguably the ideal standard.

1.1 Overarching principles common to the development of our risk-based standards

A measurement system needs to strike the appropriate balance between complexity and general usability of the proposed tools. It should allow for rapid ranking or prioritization of pesticides used on a commodity, watershed or other designated land unit regardless of whether the product is older and well characterized or newer and just cleared for use. It should, if at all possible, assess risks to valued components of the environment such as terrestrial wildlife (birds, mammals, amphibians and reptiles), beneficial insects (pollinators, predators and parasitoids), soil micro- and mega-fauna, aquatic invertebrates, fish, and aquatic and terrestrial non-target plants. A good measurement system should also take into account not just the hazard of a product based on standard toxicity endpoints, but measure risk by considering the likelihood of exposure, through linkage with fate models or accepted exposure scenarios. Finally, where possible, risk-assessment tools should be calibrated against real-world conditions through an examination of the existing field records. This is the gold standard of risk assessment tools—but one that is rarely if ever met.

A number of general rules apply to the risk-assessment tools we have developed. These rules are in agreement with a review carried out by the E.U. (e.g., the CAPER project – Reus *et al*., 1999; 2002), and also best match the goals of NAESI. Some of these rules have now been incorporated (to varying degrees) into a harmonized European indicator (see the HAIR project at: http://www.rivm.nl/rvs/overige/risbeoor/Modellen/HAIR.jsp.) General points with which we

concur are:

- When possible, include application-specific information such as method and timing of application, formulation and site characteristics.
- Keep individual scores as much as possible rather than combining them and ensure the system is amenable to integration into a farm-specific decision-support system.
- Ideally, make the risk-assessment tools compatible with regulatory evaluation procedures. This favours those risk-assessment tools that are based on ratios of exposure and toxicity. However, we have decided to deviate from such procedures where they are found lagging with respect to the available science.
- Validate where possible. "Validation" is a loaded word and carries a heavy burden of proof. In referring to available field information, we prefer to speak of "calibration" of risk-assessment tools.
- Keep a compromise between complexity and data gaps. Here we need to distinguish between complexities that require additional data from simple computational complexity. With the availability of computer-assisted tools and spreadsheets, computational complexity per se is not an insurmountable problem.
- Based on our needs and a review of the existing science (Mineau and Whiteside, 2005), we added a number of other considerations:
- Information on pesticide use is only to be incorporated as a last step. In part, this is because there is no comprehensive survey of pesticide use in Canada. Indeed, this lack of political will continues to impede Canada's desire to measure the environmental footprint of its agricultural production and places us at a disadvantage with the rest of the developed world. In addition,

risk-assessment measures that are ultimately based on the extent of product use (e.g., water monitoring data) are difficult to use in a farm-specific decision-support system. Our risk-assessment tools will therefore be based on the application of a given product regardless of its popularity.

- Risk should be based on a specific application rate. For ranking purposes, we have tended to use maximum one-time label rates. However, the intent of the tools we present is that they should be used with specific application rates where these are known. Several risk-assessment tools make the implicit assumption that risk increases linearly with the accumulated amount of product used on a land area. It does not consider the assimilative capacity of the receiving ecosystem or that risk may be kept low through a judicious choice of application rate. We will attempt to address the non-linear nature of risk in the tools and standards we develop.
- Whether for measures of toxicity or for exposure estimates, we will use exact values rather than classes. We will favour measures of central tendency over the most conservative values. Mineau and Whiteside (2005) argued that the latter might be appropriate for compound-specific regulatory assessments, but have the perverse effect of distorting comparisons between data-rich and data-poor products. Using the most conservative value (e.g., the lowest available median lethal dose [LD_{50}]) for any given taxon) may give the erroneous impression that newly registered products (typically data-poor) are safer than the older alternatives.
- Where possible, we will use distributions of toxicity data rather than single values on a few standard organisms. The advantage of using species sensitivity distributions (SSDs) has been convincingly argued in a number of recent reviews (e.g., Posthuma *et al*., 2002).

Assuming that a suitable metric of environmental risk can be designed for a given component of

the environment, what should be the standard of acceptability? As alluded above, the ideal standard, which is not likely to be met by many products in current use, is that the compound not have a measurable impact on *any* component of the environment other than the targeted pest. Therefore, setting limits of acceptability and deciding how "clean is clean" will be a compromise by definition. When desirable, limits will be set following ecological principles of sustainability or recovery from perturbations. In many cases, however, limits are not based on strict science but on a more subjective risk-management decision linked to societal and political values. Exercises to set such limits (e.g., Delorme *et al*., 2005 for pesticide registration in Canada) illustrate this. For example, the routine demise of individual invertebrates such as Crustacea tends to be judged acceptable, whereas routine loss of bird or fish life is generally considered below the standard of acceptability, if only because of the legal imperative associated with those impacts (Mineau, 2004b).

For the purpose of presenting our results, we will adopt a three-tier system that parallels the Pesticide Environmental Assessment System (PEAS), developed in the U.S. for the Protected Harvest certification system (http://www.epa.gov/pesp/strategies/2006/ph06.htm). On the basis of predicted risk scores, and analogous to the universally accepted traffic-light system, pesticide applications will be labelled as green, yellow or red. *Red-listed* applications are those thought to be so far below the standard of acceptability that immediate measures are needed to reduce their predicted impact on the environment. *Yellow-listed* applications are considered to be below an ideal standard. These applications are ripe for reconsideration and possible mitigation of effects. *Green-listed* applications will be those meeting the ideal standard (i.e., considered at this point in time to be reasonably benign or capable of causing easily reversible impacts).

For the purpose of this and all of our previous reports in this series, a list of registered pesticides

was established with the assistance of the Pest Management Regulatory Agency (PMRA). It initially included a list of 286 active ingredients currently registered in Canada for commercial, agricultural or restricted use in agriculture, but not applied directly to bodies of water (i.e., no aquatic weed control herbicide or mosquito insecticide). We excluded applications in greenhouses, on ornamental plants, in or around buildings, on machinery, on harvested produce, on livestock, etc. A few active ingredients were excluded because of a complete lack of information. Six fumigants (metam, potassium n-methyldithiocarbamate, methyl isothiocyanate, chloropicrin, 1, 3-dichloropropene and methyl bromide) were also excluded because of uncertainties about how they would fit (or not) into the various models that were developed. This is an area for further work. For example, if the water runoff model used here (Section 7.1) did apply to fumigants, we showed that, in part because of their very high application rates, these products would be amongst those with the highest calculated risk for aquatic systems (Whiteside *et al*., 2006). At least one of the fumigants (1, 3-dichloropropene) has been detected as a surface-water contaminant in areas of use (Merriman *et al*., 1991). Nevertheless, for the present exercise, this gave us a subset of a little over 200 conventional active ingredients believed to represent the mainstay of agricultural uses as of the year 2005.

All pesticides were initially assessed using their highest one-time labelled application rate (Mineau *et al*., 2008b). For some pesticides, the rate of application is extremely variable. Therefore, it is possible that some specific use patterns of some active ingredients may meet our proposed standard, even if the maximum label rate does not. Following an initial ranking and assessment at maximum label rate, we will present a test case (carrot crops nationwide) where actual use information (recently obtained by Agriculture and Agri-Food Canada through the offices of Statistics Canada) is used to compare the relative performance of different growers of

the same commodity (Mineau *et al*., 2008b). If improvements are to be made in reducing the environmental footprint of pesticides by a more judicious choice of products, we believe this is most likely to come from a within-commodity comparison. De Snoo (2003) showed that, when the relative risk (measured through the 'Dutch yardstick') of different growers' operations was tallied, a few of the growers were shown to be responsible for a large proportion of the overall risk "cup" (sensu U.S. EPA) for that specific commodity.

This synthesis report draws, in part, on previous reports produced under this and a related initiative, including Mineau and Whiteside, 2005; Whiteside *et al*., 2006; Mineau *et al*., 2006; Harding *et al*., 2006; Harding *et al*., 2007; Singh, 2007; Mineau, 2007a; Callahan and Mineau, 2007; and Mineau, 2007b.

2 AVIAN ACUTE STANDARDS

2.1 Methodology for choosing a toxicity measure from laboratory-based avian data

Mineau *et al*. (2001) provides a detailed description of the procedures followed to arrive at a standard toxicity value that is representative of all bird species. The median lethal dose (LD_{50}), a statistically derived single oral dose of a compound that causes 50% mortality of the test population will be the measure we use; Mineau *et al*. (1994) have argued against the use of the median lethal concentration (LC_{50}), which derives the concentration of a substance in the diet that is expected to lead to 50% mortality of the test population.

Pesticides are customarily tested on one to three bird species, yet there are an estimated 10 000 species in the world and over 800 species occur in Canada and the U.S. alone. Different strategies have been devised over the years to compare the toxicity of different pesticides to individuals of a given taxon. These strategies were reviewed in Mineau *et al*. (2001). The chosen here is a

distribution-based method, by fitting available toxicity endpoints to a mathematically defined distribution, typically a log normal distribution (Posthuma *et al*., 2002) with two modifications: 1) introduction of a scaling factor for body weight to improve cross-species comparisons of toxicological susceptibility (Mineau *et al*., 1996); and 2) development of a small-sample strategy to deal with chemicals that had insufficient data to derive a distribution, while keeping the derived endpoints compatible with scaled values.

2.2 Methodology for Modeling the Field Record for Spray Applications

Whereas it is customary to have some form of toxicity/exposure ratio (TER) or risk quotient (RQ) at the core of most indicators for avian risk assessment, this assumes that we are able to accurately estimate the exposure received by birds as they frequent pesticide-treated areas. Algorithms that attempt to predict exposure through the ingestion of contaminated foods have been devised and debated over the last decades. However, there is still no adequate way to model other routes of exposure, nor is it possible to determine when non-oral routes may in fact be dominant. The available experimental evidence (Driver *et al*., 1991) suggests that dietary exposure is not the predominant route, at least under some conditions. An alternative is to use empirical (field) evidence to assess safety while bypassing the explicit need to measure exposure. We now have a reasonable sample of field studies on which to base a standard for acute effects in avians. The first analysis of these studies was published by Mineau (2002). Most of the reviewed work followed a similar pattern: insecticides (invariably cholinesterase-inhibiting compounds) were applied to various crops or forests and the impact on birds was measured through a combination of carcass searches and/or cholinesterase-inhibition measurements. Mineau developed an overall risk model that incorporated a toxicity component and application rate, as well as two other predictors built upon several physico-chemical measures, namely molecular

weight and volume, octanol-water partition coefficient, and Henry's law constant. Mineau (2002) concluded that better predictions (i.e., whether mortality would occur) were possible when the physico-chemical properties were considered, and that this was probably related to dermal and other non-dietary routes of exposure.

A re-analysis of these data by Mineau (2007a) showed that the situation was more complicated than initially presented (Mineau, 2002). One notable finding was that the model fit was quite different between direct inhibitors (carbamates and organophosphate insecticides) and indirect inhibitors (i.e., those needing activation to an oxon like phosphorothioates or dithioates). Possible reasons for these apparent differences were proposed but, until more research becomes available, the mechanism responsible for these effects remains unclear. Therefore, there is some uncertainty in extending the models, which incorporate physico-chemical constants, to classes of pesticides other than the cholinesterase-inhibiting pesticides that were modeled. In order to calculate whether pesticide applications exceed the proposed standard, three separate models are used to provide the best estimate of the lethal risk to birds: the direct-inhibitor model, the indirect-inhibitor model, and a model for all non-cholinesterase-inhibiting pesticides. Although the models used here were obtained with the full dataset (i.e., not separated into training and validation sets), the adequacy of the approach was demonstrated through a leave-one-out cross-validation (Mineau and Whiteside, 2006).

The following formulas are based on a logistic fit of available field studies. A score of 1 is defined as a study showing some compound-related mortality, while a score of 0 means that no such mortality was detected. The determination itself carries a degree of uncertainty, which is something that was addressed in a recent E.U. exercise (see Section 2.2.5). The reported probabilities do not address the extent of the mortality (i.e., the proportion of exposed birds that

died). Not all studies were capable of yielding this information and, in addition, the number of dead birds is dependent on the number of birds exposed as much as on the toxicity of the pesticide.

Details on how the various models were obtained and how they were chosen from competing models are given in Mineau (2007a). The principal variable employed in all the predictive models is a combination of toxicity and application rate. For our purposes, we used the same combination as that of Mineau (2002)—namely, the number of HD_5 equivalents per kilogram of body weight per square metre of field area.

2.2.1 The selected model for estimating risk from direct cholinesterase-inhibiting pesticides

In addition to the number of HD_5 equivalents per kilogram (kg) of body weight (bw) per square metre (m^2) of field area, the model developed for direct-acting cholinesterase inhibitors included a measure of skin permeation (Kp), estimated from the octanol/water partition coefficient and molecular weight (MW) of the pesticide. This model performed extremely well, with a classification success rate topping 95% of studies.

$$p = \left(\frac{e^{a+b(TP)+c(\log Kp[SRC])}}{1-e^{a+b(TP)+c(\log Kp[SRC])}} \right)$$

Source: Mineau, 2007a

Where:

a = - 39.000

b = 10.000

c = -7.04591

TP (toxic potential) = log HD_5 equivalents (Mineau *et al*., 2001)/kg bw/m^2 of treated area.

and

log Kp = -2.72 + 0.71 log K_{ow} – 0.0061 MW (from SRC's DERMwin program version 1.43 – Equation 1).

The net result of the Kp predictor is a higher risk associated with compounds of lower lipophicity, possibly because they are released faster from the skin's stratum corneum and other membranes (Mineau, 2007a). The risk of acute intoxication from cholinesterase inhibitors is dependent on C_{max} or peak concentrations of the active ingredient at target sites (typically neuro-muscular junctions); this is to be distinguished from pesticides where "area under the curve" considerations prevail.

2.2.2 The chosen model for estimating risk for indirect-acting cholinesterase-inhibiting pesticides (i.e., needing oxon conversion)

The influence of physico-chemical constants on indirect cholinesterase inhibitors appears to be lesser than for direct inhibitors. Nevertheless, calculation of a "dermal toxicity index" (see details in Mineau, 2007a) did help overall model fit and is used here. Because a measure of dermal toxicity is rarely available in birds, relative dermal-to-oral toxicity is approximated through a combination of molecular weight and vapour pressure (Vp).

$$p = \left(\frac{e^{a+b(TP)+c(DTI)}}{1-e^{a+b(TP)+c(DTI)}} \right)$$

Source: Mineau, 2007a

Where:

a = 4.056710 (for agricultural field applications),

6.986740 (for forest pest applications),

and 3.85394 (for orchard use).

Because our consideration of active ingredients registered in Canada did not differentiate between field and orchard uses (and indeed, some products may be used on both), the agricultural field coefficient was used in the calculations.

b = 4.41475

c = - 3.76727

TP is calculated as above – the number of log HD_5 equivalents/kg bw/m^2 of treated area.

Avian Dermal Toxicity Index (DTI) = 8.514711 -0.337085 MW^0.5 - 0.105831 log Vp

Source: Mineau, 2007a

Where:

Vp is preferentially calculated at 20°C and expressed as mPa. The net effect of the DTI predictor is to lower the predicted risk for compounds of high molecular weight and high volatility.

2.2.3 The chosen model for estimating risk from all other (non-cholinesterase-inhibiting) pesticides

For all other pesticides that are not cholinesterase inhibitors, the model published in Mineau (2002), which relies on toxicity alone, was used. Although this model was developed with a sample of pesticides that share a similar mechanism of toxic action (cholinesterase inhibition), there appears to be no a-priori reason not to extend the applicability of the model to other mechanisms of toxic action—for assessing acute toxicity, at the very least. Some uncertainty does remain with respect to compounds capable of cumulative toxicity (see EFSA, 2007, for a more complete discussion). The current model is likely to underestimate chronic or cumulative impacts. These impacts are more likely to be picked up in the reproductive toxicity index (see Section 3.2).

$$p = \left(\frac{e^{a+b(TP)}}{1 - e^{a+b(TP)}} \right)$$

Equation 1

Where:

a = -3.7917

b = 2.2628.

Again, the algorithm for field crops and pasture was used in preference to those presented in Mineau (2002), and TP is calculated as above.

2.2.4 Comparison of the three avian risk models and model outputs

An analysis of the three risk-assessment models (Table 1) indicates that, for several of the more hazardous pesticides, the choice of model matters very little. For a few active ingredients, incorporating the physico-chemical properties does accentuate the distinction between products that are expected to be lethal to birds and ones that are of lesser concern. Our approach predicts lower risks from methomyl, pirimicarb, carbaryl, and dimethoate than would be predicted by toxicity alone.

Table 1: Predicted probability of mortality according to three empirical models based on the avian field study record.

Active Ingredient (AI)code	AI accepted name	Type	Risk based on direct-inhibitor model – (Mineau, 2007b)	Risk based on indirect-inhibitor model (field sites) – (Mineau, 2007b)	Risk based on toxicity only (Mineau , 2002)
CAF	Carbofuran	Methyl-carbamate insecticide	1.00		0.92
MOM	Methamidophos	Phosphoramidothioate insecticide	1.00		0.58
ACP	Acephate	Phosphoramidothioate insecticide	0.98		0.23
NAL	Naled	Phosphate insecticide	0.96		0.33
MML	Methomyl	Methyl-carbamate insecticide	0.00		0.33
PIR	Pirimicarb	Dimethyl-carbamate insecticide	0.00		0.21
CAB	Carbaryl	Methyl-carbamate insecticide	0.00		0.41
DIA	Diazinon	Phosphorothioate insecticide		1.00	0.97
DUB	Chlorpyrifos	Phosphorothioate insecticide		1.00	0.73
PRT	Phosmet	Phosphorodithioate insecticide		0.98	0.76
GOO	Azinphos-methyl	Phosphorodithioate insecticide		0.89	0.67
DIM	Dimethoate	Phosphorodithioate insecticide		0.13	0.47
MAL	Malathion	Phosphorodithioate insecticide		0.06	0.06

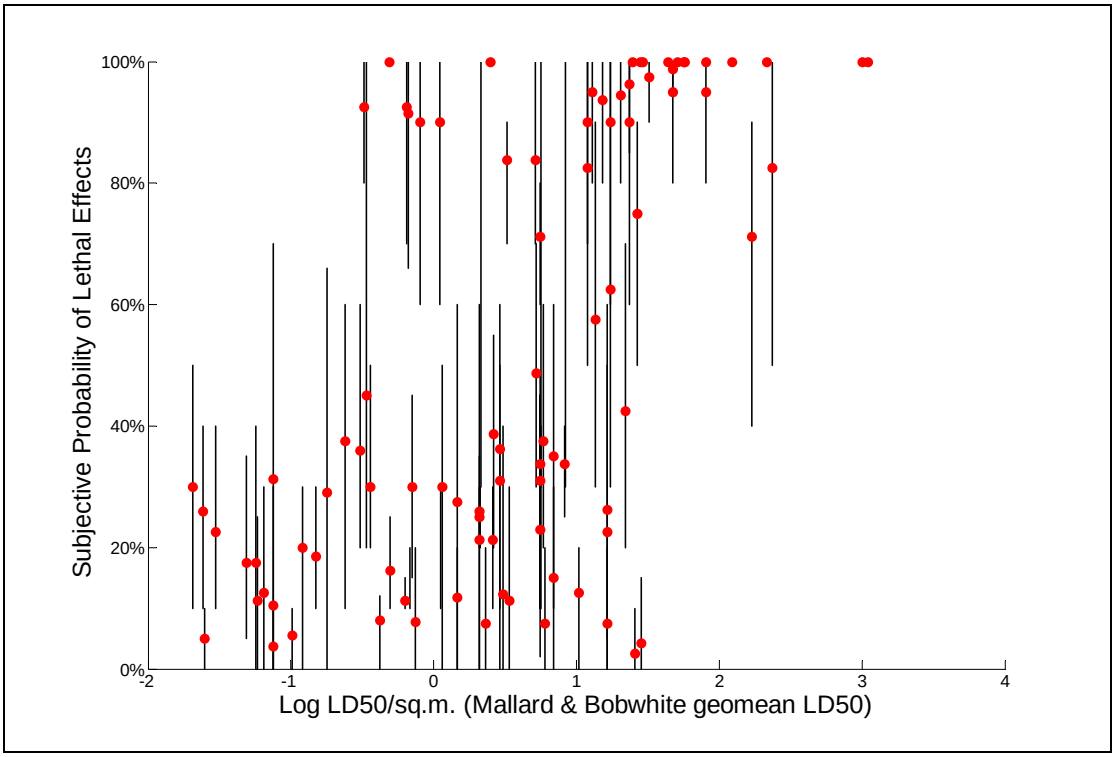
2.2.5 Comparison of field-based risk estimates with current regulatory thinking

As a result of an ongoing exercise with the European Food Safety Authority (EFSA, 2007), the agricultural field studies (i.e., excluding forest spraying studies) contained in the sample used to develop the models in sections 2.2.1. to 2.2.3. were re-assessed by a group of experts. Because results of field studies are not always clear-cut and can be subject to interpretation (e.g., was the observed mortality compound-related?) the expert reviewers were asked to independently score their 'belief' in the extent of the mortality demonstrated by the field study (see EFSA, 2007 for more details). The aim here was to:

- Democratize the ratings by allowing for independent assessment of the evidence by a broader group of individuals, one of whom was nominated by industry. (Initial ratings presented in Mineau (2002) were the work of a team of four Environment Canada reviewers, but these individuals were not working independently of each other).
- Account for uncertainties and flaws in study design that influenced the ability of observers to conclude that there had or had not been compound-related mortality.

Mean belief scores, or the average subjective probabilities that compound-related mortality had taken place, were computed for every field study (Figure 1; EFSA, 2007). The figure also shows the minimum and maximum belief attributed to each study.

Figure 1: Mean belief that compound-related mortality was demonstrated in agricultural field studies reviewed independently by four (sometimes three) reviewers. Vertical bars show the range between the lowest and highest evaluation for each study. (Taken from EFSA, 2007).



Clearly, at least one of the reviewers thought that there was a great deal of uncertainty associated with the results of any one field study. This resulted in part from each individual's experience and, in some cases, the standard of proof they sought in a field study. Keeping this variation in mind, studies were considered to show evidence of mortality if the inter-observer mean belief that there was compound-related mortality in at least one species exceeded 50%. Using this cut-off criterion, results of modeling exercises were found not to differ significantly from the single observer analyses reported in Mineau (2002) and Mineau *et al*. (2006). Also, using this critical value of mean belief to classify studies as positive or negative allowed the number of false negatives and false positives to be computed for different values of a toxicity benchmark (EFSA, 2007). In order to facilitate adoption of easily calculated benchmarks, the toxicity measure chosen

by the EFSA working group was the geometric average of bobwhite and mallard LD_{50} values. These are the two standard test species for which toxicity data are most commonly available. The EFSA working group did not recommend a cut-off value because this is a risk management decision outside of its remit. However, if the intent was to have prevented all cases of avian mortality documented in the field study record, a first tier cut-off in the form of LD_{50}-equivalents per m^2 would have had to be about 0.3 (-0.5 on the log scale). It should be noted that this cut-off is fairly conservative and entails a high proportion (approximately 50%) of false positives. Allowing the risk to increase so that mortality was present in 10% of reviewed studies increases the trigger to approximately 0.5 mallard/bobwhite LD_{50} per m^2 (-0.3 on the log scale).

In addition, our ratings were compared to risk quotients calculated using the current U.S.EPA and Pest Management Regulatory Agency (PMRA) methodology for screening-level assessments. Both jurisdictions attempt to model food intake as the sole source of pesticide exposure. Details of these calculations are provided in Mineau *et al*. (2008b).

2.3 Proposed Standard for Acute Effects of Pesticide Sprays on Birds: Setting a Probability-of-Kill as the Standard

When trying to apply a standard to our modeled probability-of-kill, the choice of a cut-off value is somewhat arbitrary. We would argue, however, that it is less arbitrary than setting a risk quotient threshold at 0.1, 1 or some other round number typically chosen by regulatory agencies. It should be noted that migratory birds in North America are afforded individual protection by statute (Mineau, 2004b). Inspection of predicted probabilities-of-kill for registered compounds (Mineau *et al*., 2008b), as well as consideration of North American and European risk quotients, suggests that a probability of 10% is a reasonable and achievable cut-off. For example, several incidents of avian mortality have been recorded with dimethoate (http://www.abcbirds.org/aims/),

a compound that has a calculated probability of mortality of 0.13 when used at maximum label rate (although, based on toxicity alone, it rates much worse). In contrast, compounds with calculated probabilities lower than 0.10 seldom if ever appear in the kill record.

The standard for avian acute toxicity and spray applications is based on the probability that a pesticide will cause detectable mortality in 10% or more of treatments based on a series of empirical models.

Applications predicted to kill birds 50% of the time or more are probably beyond any degree of possible mitigation. Most of the active ingredients concerned have been associated with mortality incidents; some have a long history of repeated kills around the world. In our estimation, applications that carry a mortality risk of 50% or more should be red-listed—that is, their use should be minimized immediately to reduce impacts on birds.

A quick comparison of our assessment with that of the U.S. EPA or PMRA (Mineau *et al*., 2008b) shows that there would be little argument about applications estimated not to meet our proposed standard. Based on initial levels of concern, U.S. EPA and PMRA procedures are more protective than our standard, but this is normal for a screening-level assessment. It is difficult to justify the continued use of pesticides with risk quotients in the thousands (e.g., diazinon, carbofuran) when the regulatory level of concern is supposed to be 1 or lower—barring overwhelming agricultural benefits from those applications. One notable disagreement between our assessments and those of other regulatory agencies is the insecticide phosmet. This discrepancy was raised by Richards *et al*. (2004). In their assessment, the U.S. EPA used the high toxicity endpoints for the mallard and bobwhite and ignored the much higher, documented sensitivity of small-bodied species to this insecticide. The PMRA also raised the issue of a potentially high risk to small birds (http://www.pmra-arla.gc.ca/english/pdf/pacr/pacr2004-38-

e.pdf). The other important discrepancies are the insecticides that are considered high-risk in U.S. EPA and PMRA screening assessments, but considered to pose minimal risk in our models: methomyl, pirimicarb and carbaryl.

E.U. proposed screening risk quotients in the form of bobwhite/mallard LD_{50}s per m^2 are also presented in Mineau *et al*. (2008b). Again, there appears to be very little difficulty in identifying the pesticide applications that carry the highest risk.

2.3.1 Adjustment of the proposed standard to reflect differential uses: Use pattern adjustment factors

The risk estimates are based on foliar applications and do not consider that some use patterns may mitigate exposure and, therefore, risk. For example, spray applications may be carried out on bare soil and then incorporated through tillage, thereby reducing exposure. In a previous report (Appendix B; Mineau and Whiteside, 2005), we presented an approach developed for the PEAS measurement system (Mineau, 2004a) that might allow incorporation of use-specific conditions through the use of application factors (UPAFs or Use Pattern Application Factors). These factors were obtained after a literature search and in consultation with experts, and are meant to reflect the relative risk of different application conditions. They are meant to be used as multipliers of the final risk scores. Factors proposed for spray applications are reproduced in Table 2 below.

Table 2: Proposed use-application factors for spray applications.

Pre-plant or pre-emergence				Post-emergence		Either
Soil-applied: liquid	**Soil-applied: granular**	**Soil-applied: unspecified**	**Seed treatment**	**Ground foliar applied**	**Soil-applied: liquid**	**Aerial application**
0.5 (surface) 0.1 (sub-surface) 0 (application followed by tarping)	See below	0.5	See below	1	0.5 (surface) 0.1 (sub-surface)	1

Note: Foliar application is set at 1 and other types of application rated in comparison. A factor of 0.5 indicates that an application is expected to carry 50% of the risk of a foliar application.

For example, a 2.76 kg/ha application of 2,4-D received a risk score of 0.16. This means that, based on our models, we believe there is a 16% probability of causing some avian mortality in the case of a foliar application. An application of this type would not meet the proposed standard of 0.1. However, application to bare soil is believed to carry half of the risk or an 8% probability of kill—an acceptable risk based on our proposed standard.

2.4 Methodology for Setting Avian Acute Standards for Granular and Seed Treatment Pesticides

Granular and seed treatment formulations present special challenges because the particles may be sought directly by birds and mammals. The attractiveness of granules depends in large part on their composition (e.g., whether they consist of clay, cellulose, dry corn cob, silica, etc) although many kills. Unfortunately, information on the make-up of granular pesticides is not publicly available, and the granule composition of certain products may change over time. Similarly, the attractiveness of different treated seeds may vary with the seed type and colour, as well as the candidate species. De-husking of seeds and other factors make it difficult to determine exact exposure.

Based on Mineau and Whiteside (2005), an initial recommendation was to treat all such

applications as spray applications and then use UPAFs to bring all risk estimates into a comparable framework. The proposed factors are shown in Tables 3 and 4 below.

Table 3: Use pattern application factors for granular pesticides.

Silica granules	Corn cob (organic) granules	Heat-treated montmorillonite and other non-friable clays, cellulose	Friable granule bases: bentonite and gypsum	Tarping follows granular application
2.0	1.0	0.2	0.1	0

Table 4: Use pattern application factors for seed treatment pesticides.

Rice Millet Sorghum	Spring wheat Corn (maize) Oats	Spring barley	Winter cereals Peas	Rapeseed Mustard Alfalfa	Soybean, field beans Sugar beet Grass
3.0	2.0	1.0	0.4	0.2	0.1

Note: Potato seed pieces were provisionally given a UPAF of 0.1. Although they may indeed be attractive to birds, they are typically buried more carefully than actual seed.

However, as predicted in Mineau and Whiteside (2005), this approach proved difficult to use in practice. With seed treatment chemicals especially, the application rates per hectare (ha) can be very low relative to spray applications, despite the fact that the risk from ingestion of an individual seed may remain high because of the high loading per seed. The low rates of application per ha, when entered in the model developed for spray applications, often fell completely outside of the range of modeled lethal doses per square metre (m^2). We tried an alternative approach, which was to transform concentrations on treated seed into "equivalent spray applications" using the RUD (residue per unit dose) factors favoured by the U.S. EPA and E.U. authorities (EFSA, 2007). RUD values are the levels in parts per million (ppm) expected following an application of one kilogram (kg) of active ingredient (a.i.) per hectare (ha). It should be noted that the RUD principle assumes a linear relationship between application rate and initial

residue levels on food items. U.S. EPA RUDs apply to applications of 1 lb a.i./acre, but they have been converted here for consistency.

Table 5 provides the actual rate of a.i. delivery per ha at typical seeding rates (PMRA, pers. comm.). This is compared to the calculated equivalent spray rates back-calculated from the concentration of a.i. on each seed, and the official average RUD factor used by U.S. EPA and E.U. authorities.

Table 5: Registered seed treatments in Canada for three major crop clusters: corn (maize), cereals and oilseeds.

AI code	AI accepted name	Type of seed treated	Rate in kg a.i/ha of seed treatment at typical seeding rates	Equivalent application rate (kg a.i./ha) with RUD of 25 (proposed E.U. 50th percentile RUD)	Equivalent application rate (kg a.i./ha) with RUD of 7 (U.S. EPA average seed RUD)
CAP	Captan	Corn	0.146	234.0	9.1
COD	Clothianidin	Canola	0.026	127.9	499.5
COD	Clothianidin	Corn	0.039	63.2	246.7
DFZ	Difenoconazole	Canola	0.002	8.3	32.2
DFZ	Difenoconazole	Cereal	0.033	9.6	37.6
DFZ	Difenoconazole	Corn	0.006	9.6	37.6
DIA	Diazinon	Corn	0.009	15.0	58.4
FLD	Fludioxonil	Canola	0.000	2.0	8.0
FLD	Fludioxonil	Cereal	0.007	2.0	8.0
FLD	Fludioxonil	Corn	0.001	2.2	8.7
IMI	Imidacloprid	Canola	0.064	320.0	1250.0
IMI	Imidacloprid	Corn	0.062	99.8	390.0
IPD	Iprodione	Canola	0.024	118.8	464.1
MAN	Maneb	Cereal	0.286	84.9	331.5

Table 5: Registered seed treatments in Canada for three major crop clusters: corn (maize), cereals and oilseeds.

AI code	AI accepted name	Type of seed treated	Rate in kg a.i/ha of seed treatment at typical seeding rates	Equivalent application rate (kg a.i./ha) with RUD of 25 (proposed E.U. 50th percentile RUD)	Equivalent application rate (kg a.i./ha) with RUD of 7 (U.S. EPA average seed RUD)
MCZ	Mancozeb	Corn	0.044	70.4	275.0
MFN	Metalaxyl-m (mefenoxam)	Canola	0.001	5.6	22.1
MFN	Metalaxyl-m (mefenoxam)	Cereal	0.019	5.6	22.1
MFN	Metalaxyl-m (mefenoxam)	Corn	0.004	5.6	22.1
MTA	Metalaxyl	Canola	0.003	14.8	57.8
MTA	Metalaxyl	Cereal	0.050	14.8	57.8
MTA	Metalaxyl	Corn	0.052	83.3	325.5
NXI	Acetamiprid	Canola	0.020	100.8	393.8
TEU	Tebuconazole	Cereal	0.227	67.2	262.6
THE	Thiamethoxam	Canola	0.032	161.5	630.7
THE	Thiamethoxam	Cereal	0.045	13.4	52.2
THE	Thiamethoxam	Corn	0.026	41.1	160.5
THI	Thiram	Canola	0.053	262.6	1026.0
THI	Thiram	Cereal	0.094	27.8	108.4
THI	Thiram	Corn	0.045	72.3	282.3
TLL	Triadimenol	Cereal	0.045	13.4	52.5
TPM	Thiophanate-methyl	Corn	0.018	28.0	109.4
TRT	Triticonazole	Cereal	0.007	2.1	8.4
VIT	Carbathiin	Canola	0.008	40.0	156.2
VIT	Carbathiin	Cereal	0.580	171.9	671.4
VIT	Carbathiin	Corn	0.031	49.7	194.2

As seen in the table above, equivalent spray application rates can be so high as to be outside the range of typical foliar rates of application. We therefore rejected our proposed approach and opted for a standard that more closely paralleled the way regulatory authorities consider granules and seed treatments. The number of particles required to be ingested to reach the adopted criterion was used to rank products and thresholds *provisionally* established from the available field evidence.

2.5 Proposed Standard for Avian Acute Impacts and Seed Treatments

Few if any studies exist that set an absolute guidance of the risk from treated seed. Exposure to seed-treatment chemicals depends on the rate of seed uptake (hence the attractiveness of the treated seed), as well as any seed manipulation such as de-husking. Standard evaluation practices that assume birds obtain an entire day's feed from treated seed are probably over-protective. In order to provide more realistic intake rates for common seed types, we obtained typical one-time seed intakes (i.e., the number of seeds taken in a single observation at a feeding table) for a number of agricultural bird species in the Ottawa area (Smith, 2006). Table 6 below is reproduced from Table 2.3 in Smith (2006).

Table 6: Results of videotaping bird visits at feeding tables summarized by crop and species.

Species	Barley			Corn			Oat			Wheat		
	visits	max.	mean	visits	max.	mean	visits	max.	mean	visits	max.	mean
American Crow	2	10	4									
Brown-headed Cowbird	20	42	10.3				12	11	3.75	13	17	6.85
Blue Jay	2	2	1.5	40	20	5.48	26	27	5	15	14	5.27
Common Grackle	18	13	2.83				14	67	13.29	47	48	9.17
European Starling										1	1	1
House Sparrow	4	2	1.25							7	4	1.71
Mallard				1	92	92						
Mourning Dove	15	37	14.2	2	15	10	4	3	2	4	90	44.75
Red-winged Blackbird	34	33	3.97	10	5	2.2	8.8	71	9.71	63	53	10.22
Savannah Sparrow	16	38	8.25				18	46	10.89	42	23	7.83
Song Sparrow	19	52	17.63				15	15	5.53	32	27	5.13
White-crowned Sparrow	12	33	7.58				2	10	5.5	41	42	9.63

Note: The number of visits, as well as the maximum and mean number of seeds ingested per visit are indicated.

Of note is that no sparrow-sized birds were seen taking corn seed despite repeated documented visits to the seed tables. Based on this observation, the scenario, presented in Mineau *et al*. (2006), for birds of 15-g mass was deemed to be unrealistic. The corn scenario was therefore modified to reflect a 50-g bird, the approximate weight of a red - winged blackbird, which was the smallest species observed taking whole corn seeds. The wheat and canola scenarios were still run with a hypothetical 15-g songbird. Soybeans appeared to be unattractive to the species that

came to the feeding platforms, a confirmation of the low factor given to soy in Table 4. However, larger species expected to be attracted to seeds of that size (e.g., partridge, pheasants, waterfowl), were inadequately sampled.

Seed treatments were ranked by the number of particles required to kill a 15-g bird (or a 50-g bird for corn). In order to reflect the relative appeal of different seed types to birds, the UPAFs shown in Table 4 above were used as multipliers of a product's relative risk.

Results for seed treatment chemicals in the corn, cereal and oilseed clusters are presented in Mineau *et al*. (2008b). For the purpose of assigning an adjustment factor, the cereal cluster product was assumed to refer to wheat. The risk for all products expected to cause mortality with one seed or less was adjusted to reflect a risk of 1, and other products were ranked in comparison. Any seed treatment with an adjusted risk index of 1 (i.e., capable of causing mortality following ingestion of one to two seeds of a preferred type) should, in our opinion, be red-listed. Even if the product is a sensory repellent or capable of leading to learned avoidance, an avoidance response is unlikely to be effective with such a low margin of safety. In the absence of field studies, we propose setting a provisional standard of 0.1 based on our relative risk index. For a small songbird, this corresponds to the ingestion of 20 seeds of a preferred seed type (with an adjustment factor of 2). Because this is well under the maximum meal size recorded for several agricultural species (Table 6), setting a standard based on the likelihood that only 20 seeds will be consumed may therefore be under-protective. This should be a provisional standard until more field-based information is made available on products of intermediate toxicity. A comparison with U.S. EPA and PMRA risk quotients (Mineau *et al*., 2008b) does suggest that this standard may not be stringent enough to ensure protection of avian species.

The standard for avian acute toxicity and seed-treatment applications is based on products having the ability to kill songbirds at the 5% tail of sensitivity distribution with a risk index of 0.1, corresponding to the ingestion of 20 seeds or less of a preferred seed type. This is a provisional standard in the absence of data on seed treatment products of intermediate toxicity.

2.6 Proposed Standard for Avian Acute Impacts and Granular Insecticides

Because specific information on granule mass is not publicly available, we assumed equal mass for all active ingredients and calculated the number of granules to reach the HD_5 for a 15-g songbird. Ratings for past and current granular products were compared to known kill incidents. Mineau *et al*. (2008b) gives ratings for current granular products; ratings for other products no longer registered were calculated but are not shown. On the British Columbia lower mainland, especially, there is a long history of bird kills with granular insecticides used on potato and other root crops. Kills were attributed to the use of fensulfothion, carbofuran, phorate, terbufos and, most recently, chlorpyrifos. This last product is responsible for some waterfowl mortality, although we have yet to see any evidence of secondary poisoning (Environment Canada, unpublished). As calculated above, the three most toxic granular products have the potential to kill a 15-g songbird at the 5% tail of the avian sensitivity distribution with a single granule (before application of any factor). We propose that these products (terbufos, phorate and diazinon) be red-listed. With a single granule being capable of causing a lethal intoxication, we would not expect the exact composition of the granule base or any avoidance response to have much influence on the likelihood of poisoning.

It is thought that some granules, especially those on an organic matrix, may be mistaken for seeds or seed fragments. In keeping with the abovementioned seed standard, we propose a provisional standard be set at a risk index of 0.1, once the indices have been corrected for attractiveness based

on granule composition. Dazomet, a fumigant, is an example of a granule type that has potential to alter the rating. The mode of action of this active ingredient suggests this granule may be very short-lived once exposed to soil moisture, and consequently the product may present minimal risk to birds. If this is shown to be the case, a UPAF of 0.1 would be applied to the current risk estimate of 0.24, bringing it below the proposed 0.1 threshold.

U.S. EPA guidelines for the assessment of granulars are somewhat ambiguous. Two assessment methods are recommended: calculating the number of granules to lethality, as we have done here, and converting granule application rates and estimated fractions left uncovered as the number of LD_{50}s per square foot. A specific threshold is not proposed, and there is also no method to account for the differential attractiveness of granule bases. PMRA guidance on the matter is also lacking. For this reason, our risk scores for granular products were not compared to North American regulatory risk quotients.

The standard for avian acute toxicity and granular applications is based on products having the ability to kill songbirds at the 5% tail of sensitivity and an adjusted risk of 0.1 or higher—in parallel with the proposed seed-treatment standard. This is a provisional standard in the absence of data on granular products of intermediate toxicity.

2.7 Methodology of Risk Scaling and Assessing Multiple Applications

As discussed in Mineau and Whiteside (2005), it was necessary to make a decision about the correct scaling of estimated risks once they were adjusted with a UPAF.

Examples of possible outcomes with a simple multiplication of UPAFs are shown in Table 7. The most hazardous condition (the seeding of rice seed), and a raw risk index of 1, would yield a worst-case corrected risk index of 3.0. At the opposite end of the spectrum, risk from any product could be reduced to naught through measures that exclude bird exposure such as tarping.

Table 7: Examples of possible avian risk scores when the raw risk scores (from 0 to 1) are combined in a scalar fashion with use pattern adjustment factors obtained from expert opinion.

Risk of detectable avian mortality (from sections 3.1.2 and 3.1.3. above)		Examples of application types and applicable adjustment factors					
		Application of rice seed	Seeding of corn or use of silica granules	Foliar treatment	Heat-treated clay granules	Sub-surface liquid	Soil-surface application and tarping
	Use pattern adjustment factor ⟶	3	2	1	0.2	0.1	0.5*0 =0
1		3	2	1	0.2	0.1	0
0.8		2.4	1.6	0.8	0.16	0.08	0
0.5		1.5	1	0.5	0.1	0.05	0
0.2		0.6	0.4	0.2	0.04	0.02	0
0		0	0	0	0	0	0

Although it may be argued that a risk score of 3 represents a level of risk that is higher than a score of 1.5, for example, both applications are predicted to result in avian mortality. Both scenarios would be proposed for red-listing, regardless of the exact threshold chosen for the standard. It should be noted that where spray applications are concerned, the models presented in Section 2.2 do not distinguish on the basis of extent of mortality, but rather on the probability that some mortality should be observed if adequately sought. It is recognized that the extent of mortality depends first and foremost on the number of birds available to be killed. Therefore, it does not seem prudent to attribute significance to the difference between scores of 1 and those over 1. We therefore propose that the risk index should plateau at 1, and hence continue to represent the likelihood of mortality between 0 and 100%. Corresponding risk estimates are shown in Table 8.

Table 8: Examples of possible avian risk scores from Table 7 adjusted to reflect a risk plateau of 1.

Risk of detectable avian mortality – model output from Mineau, 2002 on a scale of 0 to 1.		Examples of application types and applicable adjustment factors					
		Application of rice seed	Seeding of corn or use of silica granules	Foliar treatment	Heat-treated clay granules	Sub-surface liquid	Soil-surface application and tarping
	Use pattern adjustment factor →	3	2	1	0.2	0.1	0.5*0 =0
1		1	1	1	0.2	0.1	0
0.8		1	1	0.8	0.16	0.08	0
0.5		1	1	0.5	0.1	0.05	0
0.2		0.6	0.4	0.2	0.04	0.02	0
0		0	0	0	0	0	0

All risk scores calculated in this exercise will follow a similar structure wherein 1 represents risk that carries a high certainty of mortality and 0 represents an application considered relatively benign. Having all risk scores defined as probabilities of impact suggests a simple method of combining risk from several applications on the same field: through the product of the independent probabilities of non-impact.

For example, if a certain application carries a risk of 0.6, the probability of non-impact is 1—that is, 0.6 or 0.4. If a second application on the same field carries a 0.01 (or very low) probability of risk, the combined probability of impact would be 1- [(1 - 0.01) * (1 - 0.6)] = 0.604. This calculation is continued for the number of applications on the field.

Therefore, in general terms, the combined risk score is:

$$= 1 - [\prod_{k=1}^{n} (1 - Pk)]$$

Equation 2

Where:
P is the risk score (probability from 0 to 1)
n is the total number of applications on the field.

Applications with no probability of impact (i.e., with a score below 0.01 in any of the indices calculated in the report) need not be included in the calculation.

3 AVIAN REPRODUCTION STANDARDS

3.1 Methodology for the Avian Reproduction Standard

Field data on the impact of pesticides on avian reproduction (excluding a few organochlorine insecticides) do not exist. If the potential for harm to avian reproduction is to be scored, we have to rely on laboratory-generated data. Mineau (2005) reviewed in detail the study endpoints typically generated for the purpose of pesticide registration. Unfortunately, current study protocols measure an unnatural and truncated reproductive performance. The species (both indeterminate layers) are induced to lay a much larger clutch than their wild counterparts, and parental influence is restricted to only two aspects of reproductive performance and behaviour: copulation and egg laying. In the wild, there are many aspects of parental performance that may directly affect reproductive success in the period from pair formation to fledging. Quantified or even quantifiable measures of behaviour are totally lacking from the current protocol.

Details of these studies are not publicly available. Therefore, it is not possible to consider the

endpoints separately, as recommended by Mineau (2005) and put into effect by Bennett *et al*. (2005) and Shore *et al*. (2005). However, the U.S. EPA does make single NOAEC (no observable adverse effect concentration) values available in database form for a large number of registered pesticides (one-liner database, Brian Montague, U.S. EPA, pers. comm.). This data is therefore available for a ranking and standard-setting exercise such as this one. Mineau *et al*. (2006) proposed a method to use reproductive NOAECs in the light of the best available science. Readers are referred to Report No. 2-43 for details (Whiteside *et al*., 2006).

The derived risk measure for birds from spray applications is the amount of time that residue levels in food items (insects) remain high enough that the daily chemical intake of our model bird exceeds the reproductive effect threshold. The mallard and bobwhite endpoints are averaged and adjusted to reflect a 15-g insectivorous bird. As recommended by a recent expert panel (Luttik *et al*., 2005), the inter-specific variance in acute toxicity is used as an estimate of the inter-specific variance in chronic toxicity. In the absence of a measured variance, a default value based on all pesticides is used. The reproductive NOAEC is therefore expressed as a critical daily intake for a 15-g bird. Based on an insect diet, this critical intake is converted to a residual level. Values calculated here were modified from those given in Mineau *et al*. (2006) to account for the newly proposed foliar insect residue levels (EFSA, 2007). The 50th percentile RUD of 11.2 ppm was used to replace the more conservative value of 29 ppm. Finally, the foliar DT_{50} (measured or estimated) was used to calculate how long after application residues would exceed this critical residue level (Mineau *et al*., 2008b). Because of the mobility and replacement of insect prey in treated areas, the foliar DT_{50} is thought to be an underestimate of residue loss on and in insect prey. One obvious assumption in this approach is that the bioavailability of residues remains constant over time.

3.2 Proposed Standard for Avian Reproductive Risk Following Spray Applications

The length of a typical breeding season for agricultural songbirds (e.g., egg dates) is approximately 90 days in heavily agricultural areas such as southern Ontario (Peck and James, 1987), and most pairs would be able to produce two clutches during that time. On that basis, the length of the reproductive season was set at 90 days. All applications were expressed as the proportion of the total reproductive season that they were likely to be interfering with avian reproduction. All products expected to be used during the reproductively active time for more than 90 days were given the maximum risk score of 1. All pesticides present for over one-third (33%) of the total reproductive season were provisionally considered below standard. Assuming that breeding was already underway and a nest was close to fledging, an application that made the nesting attempt fail and prevented re-nesting for a full month would likely remove any chance of successful breeding for that season. Products with a risk score of 1 are red-listed on a provisional basis (Mineau *et al*., 2008b). However, as has been pointed out (Mineau, 2005) there are serious extrapolation issues between the standard laboratory reproductive tests and avian reproduction in real life. This, and the lack of field validation, should make us cautious of setting a rigid standard based on avian reproduction. Finally, reproductive NOAECs are missing for a large number of active ingredients, making it difficult to fairly assess all products registered in Canada. The standard should therefore be considered provisional.

The provisional standard for the chronic (reproductive) toxicity to birds of pesticides applied as sprays is based on the product being above environmental levels estimated capable of causing reproductive dysfunction in sensitive 15-g songbirds for one-third or more of the duration of the normal breeding season.

For comparison purposes, both U.S. EPA and PMRA chronic RQs were tabulated (Mineau *et al*., 2008b). The proposed level of concern is an RQ of 1 in both jurisdictions. This is clearly much more protective than the provisional standard we have currently set. About 65% of products for which we obtained data would be considered by PMRA as having RQs of concern, which suggests that the trigger is too protective. However, given that these are all in-use products, it is clear that the trigger does not have any real influence on the registration status of pesticides.

3.3 Proposed Standard for Avian Reproductive Risk Following Granular and Seed Treatment Applications

One difficulty with the approach adopted for spray applications is that calculating the rate of disappearance of treated seed or granules is more complex than the first order loss rates assumed for sprayed residues on surfaces. We therefore reverted to calculating only the number of particles needed to exceed the daily critical intake deemed to be above a reproductive threshold for a 15-g songbird. This was calculated as described in Section 2.4 above, but uses reproductive study endpoints instead of acute toxicity. The risk indices are given in Mineau *et al.* (2008b). No standards were set because of the high uncertainty surrounding the continued availability of seeds or granules after application. It is clear, however, that this accessibility will exceed the availability of equivalent spray applications in some cases. For example, prolonged persistence of granular insecticides in the lower B.C. mainland has resulted in wildlife mortality months after application (Wilson *et al.,* 2002).

For seed treatment, our risk scores were compared to U.S. EPA and PMRA risk quotients (Mineau *et al*., 2008b). Perhaps not surprisingly, given the high loading of active ingredients per granule or seed, a large number of registered products are expected to deliver an exposure level that is above the estimated daily reproductive critical dose in a single particle or less.

Unfortunately, there are no field studies available to validate this high predicted risk. For several of the products, the occurrence of reproductive effects is likely to be a moot point given that a single particle is also likely to be lethal. From a scoring point of view, our proposed method has the drawback of not being able to distinguish between the relative risks of several of the registered seed treatments, because all particles expected to deliver an exposure level that is above the estimated daily reproductive critical dose in a single particle or less were given a score of 1. Setting the maximum risk level at one seed per day should perhaps be revisited in order for the risk scores to be more informative.

4 SMALL MAMMAL STANDARDS

4.1 Methodology for Choosing Available Mammalian Toxicity Data

Mammalian acute, sub-chronic and chronic studies are typically carried out in order to assess the safety of pesticides to humans. Mineau (2005) discusses the use of chronic studies to address wild mammal impacts. Although chronic toxicity-based no observable effects levels (NOELs) are available from public and government sources for most in-use pesticides, the varied nature of the endpoints (e.g., cholinesterase inhibition levels, terata, pup toxicity and developmental problems, etc.) makes it difficult to fairly compare different pesticides. Attempts to relate chronic toxicity endpoints to the available field data failed (Mineau, 2007b). For this reason, our standard was developed with acute toxicity information. Although there is an undeniable correlation between the acute toxicity of chemicals to birds and mammals, there are also notable differences that make it worthwhile to develop a separate standard for mammals (Mineau *et al*., 2006).

Rat LD_{50} values are the most widely available toxicity endpoints for most pesticides. However, a comparison of the predictive power of these endpoints with that of SSD-based values showed that the latter were preferable (Mineau, 2007b). It has been shown (Sample and Arenal, 1999) that

there is no need to scale acute data for body mass in mammals. As was done with birds in Section 2.1, we chose the HD_5 or a value at the 5% tail of the estimated distribution of LD_{50} values available for any one active ingredient. We used the ETX 2.0 software (van Vlaardingen *et al*., 2004), which was developed to calculate the hazardous concentrations and fraction affected based on normally distributed toxicity data. For all datasets with five data points or more, visual inspection of the data was critical. Where the sample was considered normal based on a cumulative probability plot and the Anderson–Darling test, an SSD was generated. If, on the other hand, normality was not met, the small-sample method was used (Aldenberg and Luttik, 2002). This consists of estimating the HD_5 on the basis of a mean LD_{50} and pooled variance estimate of 0.36 (for the $\log_{10}$ LD_{50} values), calculated for a large group of pesticides at large. The median estimate of the HD_5 was calculated in order not to bias for data availability.

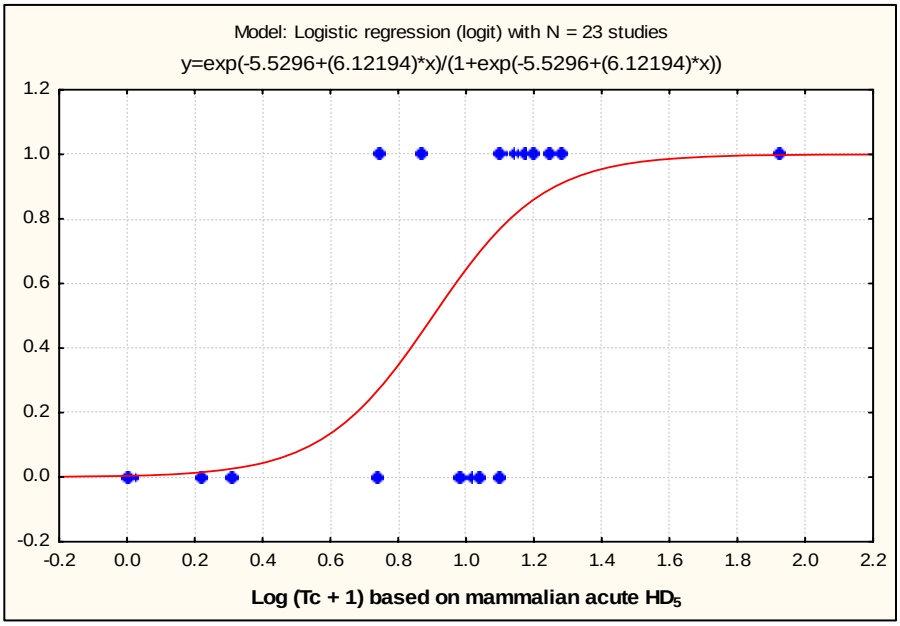
4.2 Methodology for Modeling Small Mammal Population Effects from Available Field Studies Following Spray Applications

The field data available to build an empirical model of pesticide effects is more limited for mammals than it is for birds. Nevertheless, field studies that investigated the impact of pesticides on small mammals were assembled to see whether they could be used to calibrate assessments of either acute or reproductive toxicity. Analyses to that effect were started in Mineau *et al*. (2006) and continued in the context of the European Food Safety Authority (EFSA) working group on the guidance document for the assessment of risks in mammals and birds (Mineau, 2007b). A population response was defined as reductions in some age or sex cohorts, which could indicate mortality, or as changes in reproductive rates (e.g., pregnancy rates, etc.) that indicate a more targeted effect on the reproductive process. Although small mammal populations are able to bounce back quickly from catastrophic mortality events, the impact may have ripple effects on

their consumers.

Acute risk quotients were computed based on previous E.U. guidance (European Commission, 2002; details given in Mineau, 2006), although risk quotients were expressed as North American-styled risk quotients (ETRs [exposure over toxicity] rather than TERs [toxicity over exposure]) and log-transformed. In the original analyses (Mineau *et al*., 2006), the log ETR was referred to as "log exceedance," which referred to the extent to which the chosen effect level was exceeded immediately after pesticide application. The chosen scenario (that of a small 25-g herbivore) was thought to be the best match for the field data. It was based largely on microtine rodents. Based on our analysis showing foliar DT_{50} values to be important in predicting population-level impacts, two other predictors were constructed and tested for their predictive power: the time that residues in the environment remain above a critical level calculated to give one-day exposure levels above the HD_5, as well as the area under the ETR-over-time curve. The best predictor was found to be the time in days that residues remained above critical concentrations (Mineau, 2007b). The plot of this variable against the field study results is presented in Figure 2.

Figure 2: Logistic model showing the probability of small mammal population response on the Y axis (0=no response; 1=significant effect) using, as predictor, a log transform of the time (in days) that residues in the environment are predicted to be present at a level exceeding the 5% acute toxicity threshold determined from a species sensitivity distribution.



4.3 Proposed Standard for Small Mammal Impacts Following a Spray Application

Mineau et al. (2008b) tabulate currently registered pesticides and show how each of the registered active ingredients fared when applied at maximum label rate. As with the acute index for spray applications of birds, we computed current U.S. EPA and PMRA risk quotients. One observation is that there is less agreement between our results and those of the U.S. EPA and PMRA for small mammals than there was for birds. The U.S. EPA and PMRA have different ways of computing RQs, which leads to significant variations in how different applications are assessed. Also, both agencies rely solely on rat data which, counter-intuitively, we have found to be a poor predictor of small mammal toxicity. The PMRA's level of concern (LOC) of 1 corresponds almost exactly to our calculation of a de minimis risk for small mammal populations. This therefore appears reasonable as a screening assessment. The U.S. EPA is less protective in its method of assessment

and LOC. Adopting the PMRA's LOC of 1 (less than a 1% risk based on our estimate) would mean that approximately 37% of assessed products (when applied at maximum label rate) are above the LOC. Based on Figure 2, it is clear that all studies with a critical time of 12 days or longer showed significant population impacts. Keeping in mind that the sample of studies and tested pesticides is very small, there are no false positives at that threshold. The probability of impact based on our model is approximately 75%. Based on their respective methodologies, both U.S. EPA and PMRA would agree that most of these applications pose a very high risk. We propose that all applications scoring above this level (i.e., with a probability of impact of 75% or more) should be red-listed, and action should be taken as soon as possible to minimize their use and reduce mammal impacts. The standard itself should be set at a more protective level to minimize the proportion of pesticide applications that can impact small mammal populations. Where this standard should be set exactly is a subjective decision. The field studies used in the model represent worst-case situations, because they are cases where pastures or old fields are sprayed directly. In many row crops, small mammal activity is probably concentrated around field edges, and their home ranges also include non-crop habitat. Below the predicted probability of impact of 10%, there are no U.S. EPA RQs that exceed the highest U.S. EPA acute level of concern. We propose this be the established standard for small mammals but recognize that it may be considered too protective. However, it is interesting to note that several pesticides deemed to exceed the PMRA level of concern are considered to carry a negligible risk in our assessment: etephon, acetamiprid, chlorothalonil, flufenacet, Napropamide, Mancozeb, diclofop-methyl and fosetyl-al.

The standard for effects in small mammals exposed to spray applications is based on the probability that a pesticide will cause a population decline in 10% or more of treatments based on empirical models.

4.4 Methodology for Considering Small Mammal Impacts of Granular and Seed Treatment Pesticides

As with birds, a different framework is needed to rate the effects of pesticides offered in particulate form on small mammals. Products can be ranked by the number of particles needed to achieve lethality as defined by the mammalian HD_5. However, we cannot assume that seed or granule attractiveness to small mammals will be the same as it is for birds. For example, seeds are not likely to be swallowed whole by small rodents regardless of seed size. For this reason, we are unable to recommend the use of UPAFs for estimating the mammalian risk. Because of uncertainties about the choice of toxicity endpoints by the U.S. EPA or PMRA, the RQs generated by these two agencies have not been computed.

4.5 Proposed Standard for Small Mammals Exposed to Seed Treatment Pesticides

The standard for mammalian toxicity and seed-treatment applications is based on products having the ability to kill a small mammal at the 5% tail of sensitivity distribution with a risk index of 0.1, corresponding to the ingestion of ten seeds of any type. This is a provisional standard in the absence of data on seed treatment products and mammalian impacts.

Mineau et al. (2008b) rank the risk to a 25-g small mammal at the 5% tail of the species sensitivity distribution (SSD) to mammals. In order to anchor the relative risk scores in the same way as was done for birds, we assumed that the worst outcome would be the situation where a single seed is above the lethal dose. No products reached that level of toxicity for small mammals. The standard here was provisionally set at 0.1 as it was in birds. Without any

modification for relative attractiveness (i.e., all UPAFs set at 1), this is equivalent to the consumption of ten seeds.

4.6 Proposed Standard for Small Mammals Exposed to Granular Pesticides

An identical strategy was followed for granular products. Obviously, granules will not be taken up as grit by small mammals. However, some are on an organic base (e.g., corn cob); others use vegetable oil as a carrier and may therefore have some food value and appeal to a foraging small mammal. Molluscicide pellets containing methiocarb have been found to have a high impact on wood mice in the U.K. (Johnson et al., 1991), but these are most likely grain-based. Block and colleagues (1999) documented the exposure of Peromyscus species to terbufos granules (Counter 15G) for up to 15 days after application.

The standard is provisionally set at the same level as for seed treatments (0.1). Only two products (Mineau et al., 2008b) exceed this standard and are toxic enough to be red-listed.

The standard for mammalian acute toxicity and granular applications is based on products having the ability to kill a small mammal at the 5% tail of sensitivity at a risk of 0.1 or higher—in parallel with the proposed seed-treatment standard. This is a provisional standard in the absence of any data on the impact of granular pesticides on small mammals.

5 POLLINATOR STANDARD

In Canada, field tests that consider the impact of pesticides on bees or wild pollinators are seldom carried out as a condition of pesticide registration. Instead, results from laboratory tests are used to trigger product label warnings. For example, the label for Wilson Dimethoate 480 (NuGrow PR Inc.) states: “To reduce injury to bees, restrict application to the period after dark when bees are inside the hives or in the early morning before the bees are foraging.” While these label

recommendations may help reduce the impact of spraying on managed crop pollinators, since hives may be covered or removed prior to application, such recommendations may not be adequate to protect native pollinators in the surrounding habitat. For example, non-*Apis* bees (e.g., bumblebees) are known to have seasonal and diel foraging cycles that differ markedly from those of honeybees (Thompson and Hunt, 1999; Thompson, 2001), and mortality in those species usually goes unnoticed. Of course, different pollinator species may also differ markedly in their sensitivity to different pesticides (Tasei, 2002). Although it would be preferable to have data on a wide range of pollinating species, these data are not routinely available for registered products. Nevertheless, we opted to look at honeybee toxicity and honeybee incident reports as a first-tier surrogate for ranking the risk of pesticides to pollinators in general.

5.1 Methodology for Building a Honeybee Risk Index

Two toxicity tests are typically available for honeybees: oral and contact toxicity. Their ability to explain reported hive incidents were compared (Harding *et al*., 2006; Mineau *et al*., 2008a). Canada has no centralized registry of bee mortality incidents nor does it collect comprehensive pesticide use or sales data; the United Kingdom, however, does both. By making use of a database containing 21 years of honeybee poisoning incidents in the United Kingdom, the Wildlife Incident Investigation Scheme (WIIS), as well as corresponding pesticide-use surveys, we attempted to explain honeybee poisoning incidents in the field using models derived from pesticide-use information, laboratory-generated bee toxicity data, and physico-chemical properties of applied pesticides. Contact toxicity was found to be a slightly better predictor than either oral toxicity or the most protective of either oral or contact toxicity. Although it is clear that oral toxicity needs to be considered in the evaluation of systemic compounds, contact toxicity is the endpoint preferentially used to score registered compounds. Contact bee toxicity data were

assembled from a variety of sources including, principally, Atkins *et al*. (1981), the EPA Pesticide Ecotoxicity Database (http://www.ipmcenters.org/Ecotox/index.cfm), the French AGRITOX database (http://www.dive.afssa.fr/agritox/index.php), INCHEM (http://www.inchem.org/) and *The Pesticide Manual* (Tomlin, 2003). Where there was more than one oral or contact LD_{50} record available, the geometric mean was calculated. A simple index was computed in the form of a hazard ratio (HR) obtained by dividing application rates in grams of active ingredient per hectare (g a.i. /ha) by LD_{50} (µg/bee) to obtain the number of million lethal doses of a pesticide applied per hectare. Our attempts to improve the index through introduction of physico-chemical constants such as foliar persistence did not substantially improve our ability to predict field incidents.

5.2 Methodology for Calibrating the Bee Risk Index with Field Poisoning Incidents

A total of 234 poisoning incidents were found in the WIIS dataset and available for modeling. The WIIS relies on beekeepers and other interested organizations or individuals to report suspected poisoning incidents and submit dead bee samples for analysis. All of the poising incidents used in this study were cases where dead bees were found at the hive. The bee samples are analyzed to rule out non-poisoning incidents (e.g., mite infestations) and to detect any pesticide residues that may have caused bee mortality. Pollen (from pollen baskets on the dead bees) is also analyzed to help determine the crops on which the bees have been foraging. The incidents analyzed to calibrate our risk index included mortality incidents resulting from pesticide applications that may have been improperly timed (e.g., applied to crops in flower or in ways otherwise inconsistent with the label); because our goal was to develop models applicable to wild pollinators, these incidents were retained for the analysis. However, bee mortality incidents

resulting from the abusive use of pesticides (e.g., pesticides being applied to the hive directly) either to kill the bees or to kill bee pests such as the varroa mite (*Varroa jacobsoni*) were excluded from this analysis. The WIIS reports annually and reports are now (1998–current) available for download at http://www.pesticides.gov.uk/environment.asp?id=1861.

Two crops reporting the highest number of incidents were retained for purposes of calibrating the risk index: field beans/peas and rapeseed (canola). The latter was found to carry a higher risk to bees and it was therefore used to help define the potential impact associated with any given pesticide application. This may be because bees are more vulnerable in oilseeds on account of their foraging behaviour or that beekeepers are more likely to associate mortality with pesticide use in that crop. Using the best predictor model (area treated, $HR_{contact}$), we estimated the probability that a bee mortality incident would be reported to the monitoring scheme if the area of treated rapeseed equalled the geometric mean of the 20-year cumulative area treated for each insecticide in the sample of insecticides used in the U.K. during the total survey period (see Harding *et al*., 2006, for details).

5.3 Proposed Pollinator Standard

Although it is difficult from our previous analysis to set strict HR limits that should apply to Canadian agriculture, a few general conclusions can be drawn. There appears to be negligible risk from applications of pesticides with $HR_{contact}$ values below 50. This is a useful validation of the first-tier cut-off value of 50 proposed in the European Commission guidance document on terrestrial ecotoxicology (European Commission, 2002), which was apparently established from unpublished field trials. Beyond an $HR_{contact}$ value of 400, the risk of recording hive mortality incidents is extreme (~ 50% probability) for any pesticide in broad usage. It is clear that the lack of any mortality incident data is no grounds to declare a product safe to bees and the area treated

has an overwhelming influence on predicting whether incidents with any particular insecticide are reported.

Hazard ratios (HR) were generated for the active ingredients used on crops in Canada by taking the maximum application rate and dividing it by the bee contact toxicity value. To set an appropriate standard, the HR cut-off value of 50 proposed by the European Commission (2002) was chosen. All HR values reported above 50 do not meet the bee toxicity standard. To further designate extreme risk, all compounds with an HR value exceeding 400 were flagged as red-list compounds. The standard may be considered by many to be too protective in that, if adopted, it would essentially prevent all possibility of mass mortality of native pollinators. However, as reviewed in Harding *et al*. (2006), pollination is currently in crisis and this carries a real economic and social cost.

Not surprisingly, many insecticides, because of their mode of action, fall under the extreme risk category for bees. When examining calculated HR values in Mineau *et al*. (2008b), it is important to keep these in perspective—the maximum application rate is being used to generate these HR values. Many farmers are aware of the risk to pollinators and attempt to modify their chemical use accordingly. Without actual pesticide use information, it is not possible to get an accurate picture of how real practices impact bees and we can, therefore, only generate a worst-case scenario approach.

$HR_{contact}$ values for all pesticide applications range from 0.01 to more than 80 000. In order to accommodate this very wide range, all values were log-transformed. In order to make the risk scores conform to the 0–1 range adopted for all other indices, log HR values are expressed as a proportion of an absolute worst-risk to pollinators. The insecticide chlorpyrifos, at its maximum application rate, has an HR over 80 000. We arbitrarily set the worst case at 100 000 or 5 on the

log scale. This gives a rating of 0.98 for chlorpyrifos and linearizes all other ratings. The HR value of 400 for red-listed applications corresponds to a risk score of 0.5 on the proportional log scale. The proposed standard of an HR of 50 corresponds to a score of 0.33.

The terrestrial invertebrate standard based on honeybee toxicity will be that applications to areas frequented by natural pollinators should not exceed a hazard ratio of 50 corresponding to a calculated risk score of 0.33. This risk score is computed as a ratio of the log HR value for the application to a theoretical worst case of a log HR of 5.

6 EARTHWORM STANDARD

When pesticides are applied to agricultural land, the soil is inevitably exposed to the compounds. This occurs either through direct deposition of aerosol, incorporation of solids or from being washed off foliage. Soil-dwelling organisms may therefore be exposed to substantial amounts of pesticides.

Earthworms are frequently used as indicator organisms to gauge the effect of pesticides on terrestrial invertebrates. Earthworms are of great importance for soil health and have an enormous impact on the soil and the entire ecosystem. As detritivores feeding on litter, they have a key role in the breakdown of dead plant material. Their activity contributes to the mixing of soil layers and the aeration and water flow in the soil system. Further to this, the digestive activity of earthworms improves the availability of nutrients to plants and other soil organisms.

The role of earthworms in ecosystems can, for instance, be gauged by the dramatic changes these organisms cause in naturally earthworm-free environments, such as temperate and boreal forests in North America. In these ecosystems, the introduction of earthworms can alter the whole soil food-web and above-ground plant community, including the loss of tree species such as sugar maple (Frelich *et al*., 2006). The decline of earthworms in agricultural ecosystems dependent on

their activity would likely be equally dramatic. Adequate protection is therefore essential.

The pesticide risk indicator models (as listed by Mineau and Whiteside, 2005) were analyzed to determine the different approaches used to predict risk to earthworms. The basis for predicting earthworm risk in most models was the estimation of soil pesticide concentrations. However, the level of sophistication of the reviewed systems varied widely, and ranged from a complex leaching model to a simple approximation of soil concentration that assumed standard penetration depths and bulk density. In systems where standard penetration depths were used, the assumed values varied, which resulted in substantial differences in estimated environmental concentrations (EECs). Due to the nature of dose-response relationships, this could mean that, in one case, the concentration would be below the NOEL concentration, whereas the other model would indicate substantial mortality. It was therefore deemed critical to investigate how various physical and chemical parameters influence the actual penetration depth of a given pesticide.

6.1 Methodology for Obtaining Earthworm Toxicity Data

The acute effect of pesticides on earthworms is generally assessed in laboratory tests. A frequent test protocol is the Organisation for Economic Cooperation and Development (OECD) guideline for testing chemicals No. 207 (OECD, 1984). These tests are most commonly conducted with the species *Eisenia fetida* (or *Lumbricus terrestris*). Tests with these species are relatively inexpensive and straightforward to carry out, and a substantial dataset is available for comparison of pesticide toxicity relative to other substances.

Earthworm data were obtained from several sources including, but not restricted to, the U.S.EPA's one-liner database (Brian Montague, pers. comm.), *The Pesticide Manual* (Tomlin, 2003), the French Agritox database (http://www.dive.afssa.fr/agritox/index.php) as well as a comprehensive literature review by Jänsch, Frampton and colleagues carried out under the British

Wefram initiative (see Frampton *et al*., 2006 for details). The Wefram initiative also provided references that were used to identify the key field studies used for model validation. Toxicology data from lab studies were screened for data quality. Generally, all values were included except when there were inconsistencies, such as known deviations in testing conditions (e.g., the test followed OECD guideline 207 protocol, but at 28°C to mimic tropical conditions). If toxicity values included a categorical qualifier (e.g., > 1000 ppm), then those values were only used if no "absolute" values were available. The geometric mean was calculated for each substance across all test data that passed this screening process.

However, a comprehensive review of field and lab studies (Frampton *et al*., 2006; Jänsch *et al*., 2006) indicates that earthworms are not the most sensitive of soil organisms, and furthermore, that the standard test earthworm species (*Eisenia fetida, Lumbricus terrestris*) seem to be less sensitive to pesticides than other (smaller) earthworm species or other soil invertebrates. Nevertheless, due to the limited availability of data for other taxa and the relative importance of earthworms in agricultural ecosystems, we will focus on the pesticide risk to earthworms as a predictor of pesticide impact on overall soil health. It is important to keep in mind the above-noted constraints, which indicate that other soil invertebrates may not be protected by a standard that is based on toxicity to the earthworm species for which the bulk of toxicity data have been generated.

6.2 Methodology for Calibrating Earthworm Mortality Estimates using Field Data.

Approximately 35 published field studies on earthworm mortality following pesticide application were screened for data quality and comparability of conditions. Data points accepted for further processing met the following conditions: (i) a liquid pesticide solution or suspension was sprayed

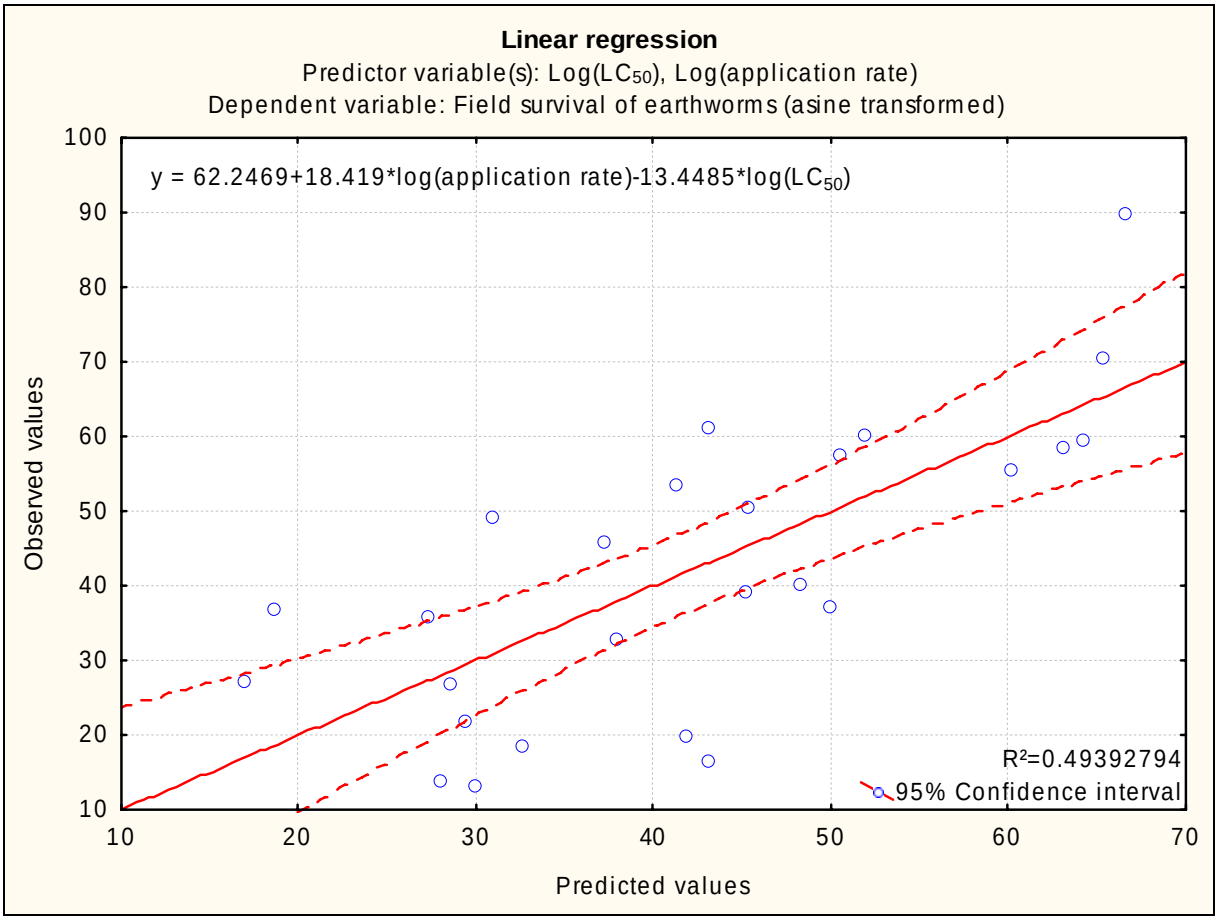
on soil or plant cover; (ii) no soil incorporation techniques such as rotary tilling or others were used; (iii) the time between pesticide application and earthworm counts/mortality assessment was as close as possible to the assessment times used in lab studies (14 to 28 days); (iv) only one application occurred prior to earthworm counts. Earthworm counts conducted more than 100 days post-application were not used, because the influence of factors such as earthworm regeneration, pesticide degradation, and loss to leaching were considered too unpredictable beyond that time period. Although somewhat arbitrary, the cut-off of 100 days was deemed appropriate given the range of generation times for different earthworm species and the degradation rates of pesticides in the soil.

A total of nine studies with one or more trials were retained for further analysis (Edwards *et al*., 1967; Voronova, 1968; Thompson and Sans, 1974; Tomlin and Gore, 1974; Tomlin, 1981; Shires, 1985; Parmelee *et al*., 1990; Potter *et al*., 1990; Kula, 1995; Römbke *et al*., 2004). For each field trial, relevant environmental parameters were extracted for use in the prediction models. Predicted survival and field survival were asine-transformed.

As a first approach, multiple linear regression models were built with the variables noted for each study. Of the variables initially included for the analysis (soil organic matter, bulk density, field capacity, DT_{50}, LC_{50}, application rate), only the LC_{50} and the application rates were found to be useful predictors of earthworm mortality. These two variables were selected using Akaike's Information Criteria (AIC).

The resulting regression model is shown in Figure 3 below.

Figure 3: Performance of a multiple linear regression model using log-transformed LC_{50} and application rates to predict field survival of earthworms (asine-transformed). Observed values were obtained from published field data. See text for more details. The R^2 value is an indication of the proportion of the total variance that is explained by the model.



It is assumed in the current risk assessment carried out by the PMRA that pesticides mix uniformly into the top 15 cm of field soil. However, when actual field mortality was compared to mortality expected following a 15-cm mixing depth, the fit was exceedingly poor. A simple percolation model (Längle *et al.*, in prep.) suggested that 15 cm was an unrealistically deep mixing depth and that the fit to mortality was best with predicted soil concentrations almost exactly 10-fold shallower—or in other words, a mixing depth of approximately 1.5 cm. Other studies published in recent years (e.g., Spurgeon *et al.*, 2003) concur that the assumption of a

homogeneous distribution in the top 15 cm is not realistic.

The development of a percolation model (or use of an existing one such as the Pesticide Root Zone Model [PRZM]) is a promising tool to help more accurately predict the pesticide concentration to which earthworms are exposed. Their application to help predict eartworm toxicity will be investigated further contingent upon resource availability. Ideally, such an exposure model should be coupled to an effect model based on the probit or logit dose-response relationship for each pesticide. Unfortunately, we have found that dose-response slopes are seldom provided for earthworm tests.

Based on the above-noted and for the purpose of this consolidation report, we retained the simpler regression model that uses log-transformed LC_{50} values and application rates, in order to assess the risk to earthworms from the list of registered products. The prediction of earthworm mortality is described in Equation 3. This formula resulted from the back-transformation of asine-transformed survival rates (compare Figure 3) and their expression as the percentage of earthworms killed.

$$Mortality\ [\%] = 100 - \left\langle \sin\left\{ \frac{\left[62.2469 + 18.419 \times \log(\mathrm{r}) - 13.4485 \times \log\left(\mathrm{LC}_{50}\right)\right]}{180} * \pi \right\} \right\rangle^{2} \times 100$$

Equation 3

Where:

r is the application rate in grams per hectare

LC_{50} is the geometric mean of available earthworm toxicity data from laboratory tests.

The predicted earthworm losses following the application of compounds registered in Canada at

the highest permitted rate are tabulated in Mineau *et al*. (2008b).

6.3 Proposed Standard for Earthworm Population Health

While the natural density of earthworms varies widely depending on the ecosystem (Edwards, 2004), it is broadly agreed that a significant reduction in earthworm populations resulting from agricultural practices is not acceptable.

Both the European and Mediterranean Plant Protection Organization (EPPO) standard for the risk assessment of plant protection products as well as the ISO (International Standards Organization) guideline for the determination of field effects on earthworms, deal with the question of what constitutes an acceptable level of effects on earthworms. The recovery of earthworms from a decline in density within a year from exposure plays an important role in both of these documents (ISO 11268-3, 1999; Sheppard *et al*., 1998; EPPO, 2003).

The long-term effect of pesticides on earthworms depends on the acute effect, the reproductive toxicity and the persistence of the substance. This makes it difficult to accurately predict earthworm recovery after pesticide exposure. However, population modelling can lend some insight into this question. According to a model developed for *Lumbricus terrestris*, the recovery of earthworm populations after a severe reduction in density due to a single pesticide application takes a minimum of one year, even for non-persistent substances (Baveco and Roos, 1996). Multiple applications, long persistence times, reproductive toxicity, and synergistic effects can further extend this time or may entirely prevent a recovery to pre-treatment levels.

While research on synergistic negative effects is very limited, it is known that many substances, when present for extended periods due to high persistence or repeated applications, can affect earthworm reproduction at levels far below acute toxicity levels. For instance, lindane and cadmium inhibit reproduction of earthworms at concentrations that are approximately ten times

lower than the acute toxicity threshold (Robidoux *et al*., 1999; Robidoux *et al*., 2000; Lock and Janssen, 2001; Lock, 2002; Robidoux *et al*., 2004).

The available data only allows for an accurate prediction of acute effects, and due to the factors discussed above, there is substantial uncertainty involved in the prediction of earthworm recovery. Taking into consideration these constraints, as well as the risk categories proposed in EPPO Standard PP 3/7 (EPPO, 2003), we propose that a predicted earthworm loss exceeding 65% should be considered severe (red-listing) and require mitigative measures; predicted earthworm losses of less than 35% are deemed to pose a low risk to the ecosystem (green-listing).

The proposed standard for earthworms is that pesticide applications should not be predicted to give rise to a greater than 35% loss of earthworm numbers based on the empirical model developed.

7 AQUATIC RISK-BASED STANDARDS

Setting a risk-based aquatic standard that can be used by growers at the field level requires an estimation of the likely drift and runoff of pesticides from single field treatments. Most water monitoring networks measure residual concentrations downstream from use areas in second- or third-order streams. Our analyses of surveillance data from the San Joaquin Basin in California (Appendix F; Whiteside *et al*., 2006), as well as that of others (e.g., Kreuger and Tornqvist, 1998) have shown that one of the main determinants of the frequency and magnitude of residual detections in aquatic surveillance exercises is the extent of use of any given product in the watershed—something over which individual growers have no control. If a pesticide standard is to be used eventually to help in a certification process, the risk associated with different pesticides has to be assessed at the level of the individual user. The strategy adopted here therefore differs significantly and is complementary to the setting of concentration-based standards in other

NAESI initiatives. The setting of risk-based standards involved the following steps:

- A realistic assessment of aquatic exposure potential following product use;
- The development of laboratory-based toxicity indices that fairly measure the toxic potential of different concentrations of different pesticides;
- A review of mesocosm and pond studies and the construction of models relating toxic potential to effects;
- Estimation of effects based on typical application rates and methods for the list of registered pesticides.

7.1 Methodology for Assessment of Exposure by Modeling Runoff

Our original assessment (Whiteside *et al*., 2006) used the GENeric Estimated Exposure Concentration (GENEEC) program to estimate possible aquatic exposures from runoff and drift. Developed by the U.S. EPA, the GENEEC model provides a simple method to estimate predicted concentrations of pesticides in water at the field edge. It is a truncated version of other well-known models like the Pesticide Root Zone Model (PRZM) and the Exposure Analysis Modeling System or EXAMS, which are commonly used in Canada and the U.S. for regulatory risk assessments, but differs in that it includes a drift component through the U.S. AGDRIFT (Agricultural Drift) model. It is truncated in that it does not consider site-specific attributes such as rainfall, soils, topography or crop. As a result, GENEEC requires few input variables and is easy to use. It is analogous to the E.U.'s Focus Step1-2 models. Input variables are related to the application method (e.g., the rate of application and application directions), as well as pesticide properties (physical and chemical properties and fate variables). Comparing the results of this model with pesticide residue concentrations measured in the San Joaquin Valley in California

(Whiteside *et al*., 2006; Appendix F), we showed that GENEEC gave a reasonable ranking of products, but there was a concern that the estimated pesticide concentrations were high, in keeping with this model's use as a protective screening-level tool.

For a more realistic estimation of runoff, the mass of pesticide moving in surface runoff was estimated using the PRZM pesticide fate model by the group developing the Indicator of the Risk of Water Contamination by Pesticides (IROWC-Pest) under the National Agri-Environmental Health Analysis and Reporting Program (NAHARP) being developed by Agriculture and Agri-Food Canada.

The NAHARP group is using the province of Manitoba as a test case for their approach. One reason for the choice of Manitoba is the availability of comprehensive pesticide-use information under the crop insurance program of the Manitoba Agricultural Services Corporation (MASC). Canada does not collect comprehensive national statistics on pesticide use.

In an ideal world, a grower would be able to adapt a pesticide standard to the relevant site-specific information such as the exact application rate, the type of application and equipment, field size, drainage basin, slope, soil type, crop, immediate weather conditions (e.g., wind speed and direction), soil water content, as well as proximity of any water bodies to the crop being treated. All of this is feasible in the context of the methodology proposed here, provided the appropriate modeling tools are used. For the purpose of this report, the best that could be achieved in the absence of site-specific information is an approximation of what products used at what application rate are most likely to cause aquatic impacts that would be considered below an acceptable standard.

Our first step was to compare results obtained with complete PRZM runs of actual soil landscapes and weather patterns with worst-case results obtained with our earlier GENEEC simulations.

Unfortunately, the comparison is not an exact one in that drift is not considered in the PRZM runs described here. Drift did not play an appreciable role in the ranking of products in Whiteside *et al*. (2006), if only because the chosen scenario involved low boom applications of the pesticide in medium to large droplet sizes typical of an early-season herbicide application to a field crop. Nevertheless, reliance on PRZM or other equivalent runoff models alone is equivalent to assuming that drift issues are of no consequence, or that drift has been adequately mitigated through best management practices—possibly through the use of drift reduction measures and buffers or, more simply, through careful selection of spray times and wind direction or a combination thereof. A consequence of our failure to factor in drift will be the underestimation of risk in the case of high drift applications (e.g., high boom, air blast and aerial applications).

PRZM uses the physical and chemical properties of pesticides as well as soil and climatic data to determine the mass of pesticide moving in surface runoff. On a daily basis, PRZM keeps track of the mass of pesticide remaining after degradation, and the movement of the remaining pesticide

All fate properties used were obtained from sources cited in McQueen *et al*. (2007). The soil-specific information was obtained from the National Soils Database of Agriculture and Agri-Food Canada in SLC (Soil Landscapes of Canada) form. The SLC framework provides site-specific soil data for the major agricultural regions of Canada. This information is needed for the PRZM model. Areas in Manitoba defined under the SLC framework varied in size from 10 000 to one million hectares. Climate information, scaled to the ecodistrict level, was obtained from Environment Canada.

The PRZM simulation was run using a time window of May through July 2004. The year 2004 was chosen because it was the most recent one for which pesticide-use data were available; the May–July window was chosen as a period representing the main months of pesticide application.

The year 2004 was a relatively dry year, but typical of the 2000–2005 period (http://www.msc-smc.ec.gc.ca/contents_e.html). Actual rainfall data were used to allow soil moisture content to be established in the different ecodistricts. Pesticide application was assumed to take place on July 11, 2004, in all SLCs. Runoff was estimated following a single major rainfall event of 150 mm set to occur on July 14. A rainfall event of 150 mm represents the maximum daily rainfall observed in Manitoba from 1975 to 2004. It therefore provides us with a relatively worst-case event, but under realistic soil moisture conditions. The three-day lag between application and major rainfall coincides with the period of time between application and rainfall used by GENEEC simulations. This period of time allows the pesticide time to degrade on the soil so that predicted runoff concentrations are not necessarily worst-case. Under good farming practices, a grower would likely avoid applying a pesticide if heavy rain was imminent.

The fraction of pesticide that moved on July 14 and the volume of surface runoff were extracted from the PRZM output file. The fraction of pesticide moved is defined as the mass of pesticide in the surface runoff on July 14 divided by the mass of the pesticide applied on July 11. In order to perform this calculation, an *average* application rate was obtained for each crop-pesticide combination based on label instructions. The *actual* mass of pesticide in surface runoff on July 14 was then estimated by multiplying the fraction of pesticide expected to move by the mass of pesticide applied on July 11. The mass was based on the fraction of each SLC polygon that was treated with each pesticide in 2004 (from the Manitoba Agricultural Services Corporation [MASC]) and the average application rate. The resulting concentration of runoff water was expressed on a per-treated-hectare basis, allowing for a fair among-pesticide comparison regardless of each product's popularity. Estimates of residue concentrations in puddle water and in the theoretical farm pond were obtained for each active-ingredient–SLC combination. In order

to avoid extreme outliers, the 95% percentile concentration was retained for each active ingredient.

Two different scenarios were run with the pesticide concentration data obtained for each SLC. The first scenario assumed that a small pond would be completely recharged with runoff water. This is referred to as our "puddle scenario." The second and probably more realistic scenario entails dilution of the runoff by having the field drain into the standard U.S. EPA pond, used by that agency for risk-assessment purposes. This standard pond drains 10 ha of field, it has a surface area of 1 ha and a depth of 2 m for a total volume of 20 000 m^3 (U.S. EPA, 2000). This standard pond is the one used in GENEEC as well as in the more complex EXAMS model. Given that the rainfall event we chose to model was a reasonably worst case (highest in a 25-year series), we opted to use an average value (geometric mean) for the amount of runoff flowing into the standard pond.

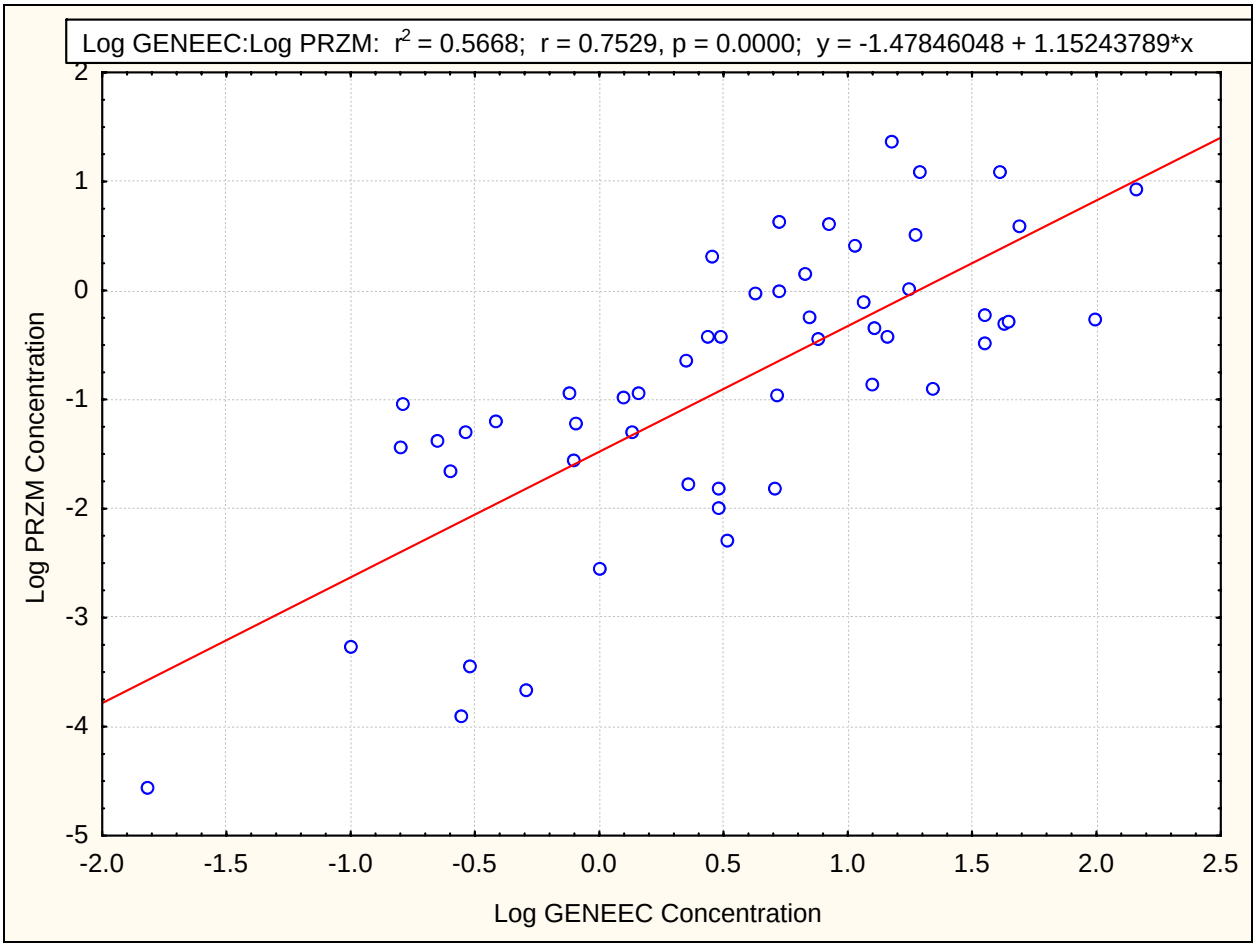
The data obtained from the NAHARP group contained modelled concentrations for 55 pesticides used in Manitoba during the 2004 growing season. GENEEC was used to provide an estimate of water concentration for all 205 compounds registered for agricultural use in Canada (Whiteside *et al*., 2006). In order to compare the two modeled estimates of water concentration, a second run was done with GENEEC on the same 55 compounds, using the same application rates used for the PRZM modeling (average label rates rather than maxima). For the purpose of this comparison, the puddle scenario (concentration in the raw runoff) was compared to GENEEC.

As stated earlier, GENEEC assumes a worst-case scenario and has been shown to predict water concentrations several orders of magnitude above those obtained in monitoring exercises (Whiteside *et al*., 2006). Unfortunately, few monitoring exercises have been carried out at the farm level. However, any dataset could be used in the same way we are using the Manitoba

NAHARP estimates of runoff—to calibrate or refine the worst-case nature of GENEEC or other such programs. Ideally, the comparison should be based on empirical farm-level information so as to avoid calibrating one set of model outputs with another; unfortunately, those farm-level water concentrations are not currently available.

All concentrations were log-transformed. The 95th percentile values for each active ingredient were regressed against the GENEEC-derived values (Figure 4). It is important to remember that during the PRZM-based modelling performed earlier, the date of the application and subsequent rainfall were timed to match the GENEEC program. The fungicide Mancozeb was an outlier (the predicted 95th percentile concentration from PRZM modelling gave a value of 0 µg/L) and eliminated from the regression equation.

Figure 4: Log-transformed PRZM-based concentrations according to our "puddle scenario" vs. log-transformed GENEEC concentrations modified to reflect the same average label application rates.



As predicted, water concentrations estimated in the PRZM puddle scenario were approximately one order of magnitude lower than the GENEEC estimates. The 0.57 r^2 value and range of modeled concentrations was judged adequate to use the regression equation and refine the GENEEC estimates derived for *all* registered pesticides in Whiteside *et al*., (2006). Carrying out this refinement step provides us with an estimate of how a PRZM-based model would have predicted the runoff concentrations of these other pesticides under the same conditions witnessed in Manitoba. Ideally, this procedure should be repeated for other cropland regions of Canada with very different soil types and pluviosity (e.g., Prince Edward Island for extremes of sand content in

soil).

This regression equation for Manitoba conditions is as follows:

95th percentile log PRZM concentration = -1.48 + 1.15 * log GENEEC estimate

7.2 Methodology for Development of Laboratory-Based Toxicity Indices

The construction of these indices was described in detail in Whiteside *et al*. (2006, 2008). Both LC_{50} and EC_{50} (immobilization for animals; growth and population effects in the case of algae) data were assembled from several sources. A series of steps were used to derive HC_5 values for fish, crustaceans, insects and algae. Inter-species variances were calculated and generic values proposed to obtain comparable values where sample size did not permit the use of a standard SSD approach.

7.3 Methodology for the Review of Mesocosm and Pond Studies and the Construction of Models Relating Toxic Potential to Effects

The detailed analysis and modeling of aquatic pond and mesocosm studies were reported in Singh (2007). A wide-ranging literature review was conducted, from which 60 studies were selected representing 184 experiments on 33 pesticides. Only studies with sound experimental designs and with quantification of effects (either explicitly stated or lifted from figures provided) were retained. The aquatic ecosystems selected for modeling were 80 litres or more in volume.

No modeling was done for fungicides alone as there were too few contained in the database. All lotic water regimes (flowing-water systems) were removed from the database prior to analysis, as the endpoints were not directly comparable to those from lentic (still-water) systems). Toxicity and exposure were also bundled together for modeling purposes, as described below. The dependent variable used in modeling was either the *count ratio* or the log-transformed *abundance ratio*. They were defined as follows:

Count ratio = number of affected species within a given taxon (statistically significant changes as reported by the study authors)/total number of species of that taxon within the system.

Log-transformed abundance ratio = log (number of individuals of a given taxon in control/number in treatment).

It should be noted that, in order for affected species to be included in the count ratio, their abundance generally had to drop by approximately 70%, reflecting the temporal and spatial variability of the systems under study. For the abundance ratio, we entered data as reported whether or not declines were considered statistically significant. When abundance was not reported or could not be obtained from the original graphs, it was given a value of 1 if reported as non-significantly different from the controls.

Many regulatory jurisdictions use toxicity and exposure combined for screening assessment tools. These are often referred to as *toxic units* (TUs). Toxic units here were calculated as follows:

TUs = Measured concentration of pesticide in µg/L/geomean LC_{50} or HC_5 also in µg/L.

Because regulatory agencies often use single-species tests (primarily *Daphnia*) to protect aquatic organisms from the impacts of pesticides, we always compared TUs based on *Daphnia* LC_{50} with TUs based on HC_5 values, calculated as described in Whiteside *et al*. (2006). To achieve normality, all of the independent continuous variables (fate, physico-chemical, and system structural properties along with TUs) were log-transformed.

Using Akaike's Information Criteria (AIC), explanatory variables were incorporated into different predictive models based on their significance and then ranked. When possible (e.g., for the crustacean models below but not the algal model), the original dataset was split into a training and a validation set. Details of these validation steps were presented in Singh (2007). To increase the sample size and generate more robust models, original training and validation sets were

combined for the present exercise and best approximating models re-defined. This explains why the models described here differ from those presented in Singh (2007).

7.3.1 Crustacea Models

The datas*et al*lowed for the modelling of two different pesticide groupings: all pesticides combined and insecticides alone. It was not possible to model data for herbicides alone due to under-representation. The presence of fish in some systems was found to lead to a high degree of variation in the study results. In small systems, fish can have a marked effect on the invertebrate community, leading to system instability. The best models were therefore obtained from fish-free systems.

7.3.1.1 Modeling Crustacea Abundance Ratios

Data were available for 14 families of Crustacea representing eight orders (Figure 1 in Singh, 2007). Seventy-one entries were available for modeling. They were drawn from 30 studies covering 22 pesticides. Of 63 possible AIC-ranked model combinations incorporating TUs and uncorrelated variables, 11 were considered to best approximate models based on an AIC difference of less than two. Contrary to expectation, TUs calculated for *Daphnia* were as well represented in the best models as TUs calculated with the HC_5 for all Crustacea combined (HC_5-C). The most predictive model included both $TU_{Daphnia}$ and a hydrolysis rate constant. However, this model only explained 2% of additional variance over a much simpler model based on HC_5-C only.

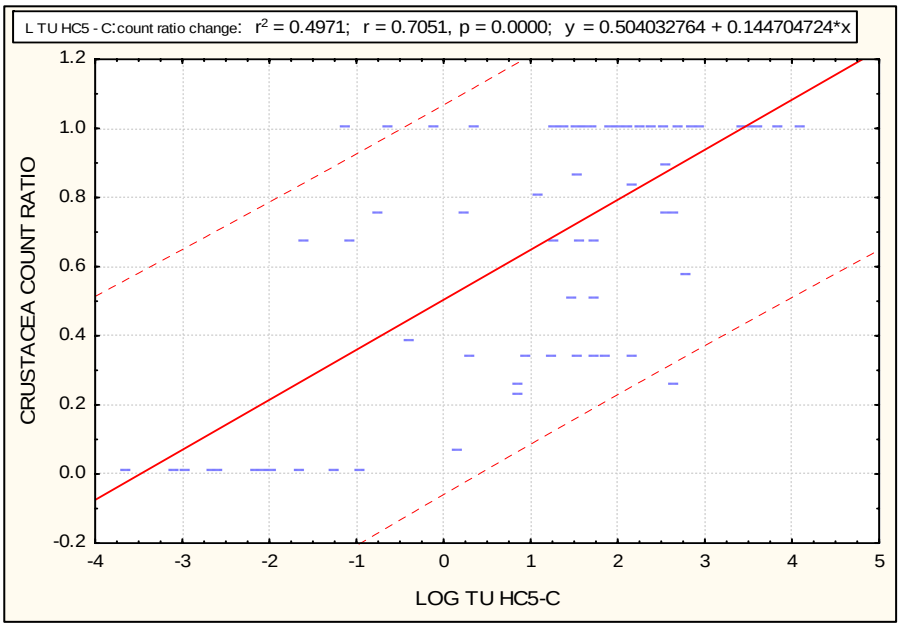
Crustacea count ratio = 0.5040 + (0.1447 * Log TU HC_5-C)

R^2 = 0.49 (N=71, 22 pesticides)

Equation 4

The relationship between the raw TU HC_5-C scores and the species count ratio is shown in Figure 5. We fitted the best linear regression but the data are suggestive of a step function with a threshold value around -1 TU units on the log scale.

Figure 5: Regression between the log number of toxicity units (TUs) calculated with the crustacean HC_5 value and the crustacean species count ratio (proportion of affected species) for a sample of 71 pond and mesocosm studies representing 22 pesticides. 95% prediction intervals are shown.



7.3.1.2 Modeling Copepod Abundance Ratios

The 50 entries available for modeling were collected from 15 studies on 15 pesticides. log abundance ratios (LAR) data were again modeled from all pesticides in fish-free systems.

From the 31 possible AIC-ranked model combinations, all most-plausible models used TU values computed from the crustacean HC_5 rather than the simple *Daphnia*-based equivalents. The best model based on the lowest AIC score was the model containing TU HC_5-C, log K_{ow}, log photolysis half-life, and log aerobic aquatic biotransformation. The second-best model had an almost identical AIC score (difference of 0.03), but only required TU HC_5-C and log K_{ow}.

However, further examination of the data revealed one outlier (gluphosinate ammonium at extreme concentration with a higher impact than expected). The removal of this data point resulted in the model with TU HC_5-C and log K_{ow} having a lower AIC score than the simple model with TU HC_5-C only. The best model of the new data set excluding the outlier was once again the model containing TU HC_5-C, log K_{ow}, log photolysis half-life, and log aerobic aquatic biotransformation. Forward stepwise regression confirmed the statistical significance of all variables except the hydrolysis rate constant with a *p* of 0.08. This model accounted for 59% of total variation. However, the simplest model containing TU HC_5-C only, although clearly less predictive (ΔAIC = 2.97), still accounted for 54% of total variation. Because of the difficulty of obtaining photolysis and aquatic biotransformation rates (these endpoints are considered proprietary in Canada and may not be released by the PMRA unless the manufacturer agrees to the release), and because these variables can be pH- and system-dependent, we opted to model copepod density reductions with TU HC_5-C only.

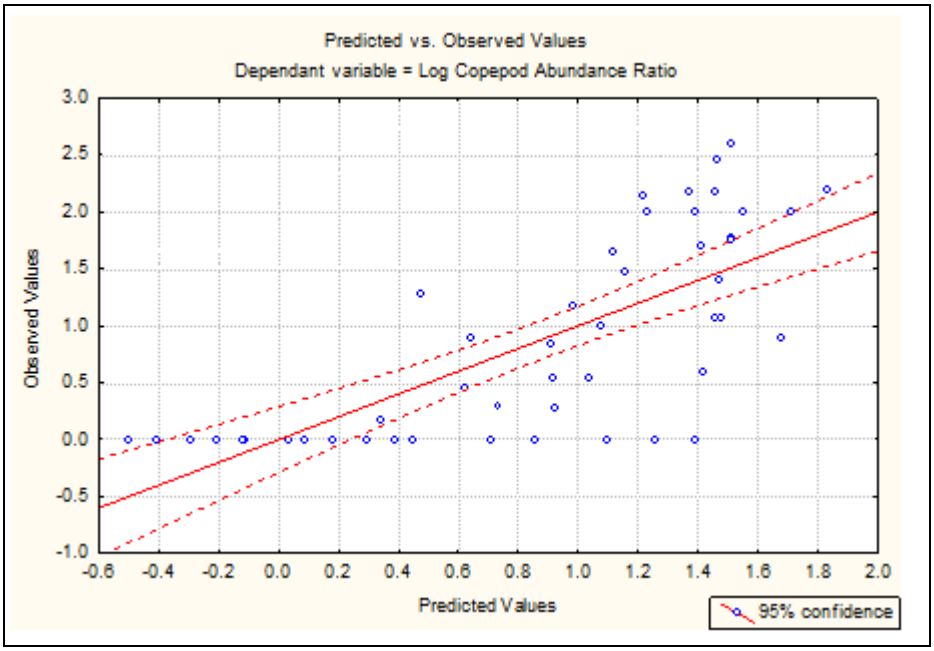
$LAR_{COPEPODA} = 0.6699 + (0.2954 * L\ TU\ HC_5\text{-}C)$

$R^2 = 0.54$ (N=49, 15 pesticides)

Equation 5

Model fit, as well as the 95% confidence interval around the fit, is shown in Figure 6.

Figure 6: Model fit for Equation 5 showing the log abundance ratio for copepods. A value of 1 means a 10-fold reduction in the number of individual copepods in the treated water body.



7.3.1.3 Modeling Cladoceran Abundance Ratios

The dataset consisted of 43 entries from 16 studies reporting on 16 pesticides. Log abundance ratio (LAR) data were modeled from all pesticides combined in fish-free systems.

Of the 63 AIC-ranked model combinations, the best model derived from fish-free systems for all pesticides contained the variables TU-HC_5-C, log K_{ow} and log photolysis half-life. However, a TU variable based on *Daphnia* LC_{50} data proved almost as good. The simpler models with a TU variable alone or with a TU variable and log K_{ow} fared much worse. However, once the same gluphosinate data point was removed from the sample, the relative performance of the different models changed. The best models all incorporated a TU based on *Daphnia* toxicity; the most parsimonious included a *Daphnia* TU value and photolysis rate. This model accounted for 52% of overall variation. A simpler model was a TU based on the HC_5 again, although it represented a loss of 5% of the explained variance (adjusted $r^2 = 0.47$). Because of the difficulty of obtaining

photolysis half-lives for many of the pesticides registered in Canada, the simpler model was retained despite its poorer performance.

$LAR_{CLADOCERA} = 0.7099 + (0.4002 * L\ TU\ HC_5\text{-}C)$

$R^2 = 0.47$ (N=42, 16 pesticides)

Equation 6

7.3.2 Algal models

Algae are considered primary producers and are a very important component of the aquatic ecosystem. Changes in their function or abundance can cause a secondary impact to non-target species, which may rely on them as a food source.

Due to the specific mode of action of different pesticides, it is unlikely that insecticides will have much of a direct effect on algae in aquatic ecosystems unless they have unforeseen herbicidal activity. For this reason, the model selected looked at algal responses for all herbicides combined. All systems (with or without fish) were used for this analysis. The main dependent (effect) variable was a count ratio of species and the main predictor variable was a TU constructed with an algal HC_5 value. Exploratory analyses showed that structural properties of the system were important. The surface area-to-volume ratio was chosen as the best predictor of system structure.

Of the 200 possible AIC-ranked model combinations, the best model contained TU HC_5-A, log photolysis half-life (L WPHL), log aquatic aerobic biotransformation (L AAB) and log total surface area to volume ratio (L TSA/VR). The photolysis half-life was the only variable not statistically significant as determined by forward entry regression.

The best model equation, one that predicts impacts to algae from herbicide use, is listed below with appropriate N and R^2 values. A search for a simpler model with toxicity only or with toxicity and a system-structural variable proved very inferior, with a loss of approximately one-third of

overall explained variance. It should be kept in mind that this is the smallest sample yet on which a model was built.

Count ratio = -0.6618 + (0.3036 * L TSA/VR) + (0.2105 * L TU HC_5-A) + (0.3268 * L AAB)

Algae: R^2 = 0.39 (N=30, 7 herbicides)

Equation 7

Where:

TSA/VR = total surface area-to-volume ratio

AAB = aquatic aerobic biotransformation.

7.4 Proposed aquatic standards for invertebrates and algae

7.4.1 Proposed aquatic invertebrate standard

Estimated water concentrations (based on the conversion from GENEEC estimates to the 95% upper tail of PRZM runs for Manitoba) were calculated for all compounds for both our puddle and pond scenarios (Mineau *et al*., 2008b). For the development of proposed standards, only the results derived from the pond scenario were used, as they are believed to be more realistic—or at least to reflect aquatic systems we should attempt to protect. Combining these estimated concentrations with the appropriate toxicity values (Section 7.2.), the number of expected toxicity units was calculated for all pesticides. Finally, these TU values were entered into the crustacean, cladoceran, copepod or algal models and effect levels—either count ratios of affected species or abundances—were computed (Mineau *et al*., 2008b).

We propose to set the level of acceptability at a 20% loss—whether of total biomass or the proportion of significantly affected species. This impact level was the one suggested by Plafkin *et al*. (1989) for U.S. EPA's rapid bioassessment procedure; others have proposed similar values

(e.g., Suter *et al*., 2000). We believe that applications with impact levels exceeding 50% should be red-listed and slated for immediate replacement and/or mitigation. This is somewhat arbitrary in that our assessment does not allow consideration of recovery of the systems from disturbance. However, disturbances affecting 50% of species or more than 50% of total biomass are likely to be of long duration and/or to have ripple effects on the rest of the aquatic community.

As is the case with other standards proposed in this report, it is important to remember that the impact scores are those predicted when a product is used at its maximum application rate and without the benefit of a buffer strip or other mitigative measure. Ideally, parallel work underway under the NAESI initiative can be used to derive use pattern adjustment factors (UPAFs) for vegetated buffer strips and other best management practices (BMPs). On the other hand, as mentioned earlier, the ratings given here have ignored the potential for aquatic contamination through drift. This is expected to be the dominant exposure pathway under some use conditions.

For three invertebrate indicators (Crustacea species counts, Copepoda and Cladocera abundance), a method was needed to combine the results into a single aquatic invertebrate index. The more conservative of the two abundance ratios was retained. However, in order to be red-listed, both abundance and count ratio variables have to exceed 50%. Conversely, to meet our proposed standard, both abundance and count ratio effects must be maintained below 20%.

To meet the aquatic invertebrate standard, pesticide applications must not result in water concentrations that lead to more than a 20% population decline in the most sensitive model (Copepoda or Cladocera) and must not result in more than 20% of all crustacean species exhibiting significant population declines.

7.4.2 Proposed algal standard

Although algae are key components of freshwater ecosystems, the population response and recovery from a pesticide application is rarely recorded for algae alone. The majority of population response and recovery studies look at the ecosystem in general, including invertebrates. Several studies report the ability of the entire ecosystem,which includes algae and crustacean species (e.g., Plafkin *et al*., 1989) to recover, while others only report the recovery of crustacean species (Roessink *et al*., 2005). Because of the aforementioned cut-off of 20% for whole-community metrics, we decided to use this same value for the algal standard despite the expectation that algal species demonstrate faster recovery from perturbation because of a faster generation time. The disturbance of algal communities can also lead to community shifts and to the appearance of undesirable toxin-producing species. We likewise propose to consider applications leading to significant impacts on 50% or more of the algal species to be red-listed applications.

To meet the algal standard for aquatic ecosystems, pesticide applications must not give rise to water concentrations that lead to 20% or more of algal species being significantly affected by treatment.

7.5 Methodology for modeling the probability of fish kills

The pond and mesocosm studies reviewed in Section 7.3 did not provide an adequate empirical basis for assessing the risk of pesticides to fish. A different approach therefore had to be developed (Whiteside *et al*., 2006).

The 206 active ingredients used on crops in Canada, and for which we have information, were ranked according to the hazard to fish. This was accomplished by dividing the PRZM-corrected GENEEC 96-hour predicted exposure concentration (as described in Section 7.1) by the HC_5 fish toxicity to yield a fish *exposure toxicity ratio* (ETR). Fish kills are a type of ecological incident

that may be more visible than others and may therefore stand a better chance of being reported. The U.S. EPA's EIIS database (Nick Mastrota, pers. comm.) is a compilation of incident reports from pesticide registrants, government agencies and other voluntary submissions from state and federal agencies. The State of California incident data was obtained from the California Department of Fish and Game (Bob Hosea, pers. comm.). Although some incidents in Prince Edward Island were also tabulated in Whiteside *et al*. (2006), Canada does not have a comprehensive fish kill reporting system, thus the use of U.S. data for this analysis. Because the U.S. EPA's EIIS database includes incidents from California, and in order to avoid double-accounting, we used the maximum number of incidents reported for each active ingredient—whether from one database or the other. For the analysis, certain criteria were required. The U.S. EPA's EIIS database classifies incidents as "highly probable," "probable," "possible," "unlikely" and "unrelated" in terms of being caused by the pesticide. Only records with a certainty of highly probable, probable or possible were retained. Likewise, only records where pesticides had a "registered use" or where the use was "undetermined" were kept for analysis (as opposed to pesticides that were misused or spilled). Furthermore, incidents where pesticides were applied to crops or turf were used, as were records when the use type was not reported. The California databases did not contain as highly detailed information as the U.S. EPA's EIIS database, but did require selection of incidents that were reported as "highly probable," "probable" and "possible," as opposed to "unlikely" or "unrelated." Five incidents were attributed to unknown forms of glyphosate; eight were assigned to all copper pesticides combined. For the purpose of scoring individual active ingredients, all glyphosate-based pesticides were assigned the five kills; copper products were all assumed to have caused eight kills each.

We compiled 438 fish kills related to the list of pesticides covered by this ranking exercise. An

examination of the incident data reveals good correspondence between the hazard-based rankings and the reported ecological incidents (Mineau *et al*., 2008b). Azinphos-methyl had the highest number of incidents for any pesticide in the U.S. EPA EIIS database. Following the methodology described above, we assigned it 98 kills; it placed fifth in our ranking of relative hazard. Endosulfan (fourth in our rankings) also had a high number of fish incidents with 58 reported cases. Terbufos, which came eighth in our ranking, had 67 reported incidents while chlorpyrifos, with 26 incidents, came sixth. The two top-ranked pesticides for their toxicity to fish—tefluthrin and phorate—had seven and ten associated incidents respectively. It therefore appears that our rankings are quite predictive of fish kills. Not all top-rated pesticides have been found to give rise to fish kills, however. As outlined in our analysis of U.K. bee kills (Section 5.1 above, or Harding *et al*., 2006), it is likely that an important predictor of the number of kills is the extent of use of the different pesticides, which is not available here. Also, it may be easier to diagnose kills from some classes of pesticides (e.g., cholinesterase inhibitors through a cholinesterase assay) than others. An absence of recorded kills is not very meaningful in itself.

7.6 Proposed fish standard: Defining benchmark pesticides for fish kills

The concept of using benchmark pesticides to define ideal performance standards (IPS) under the NAESI program was suggested by Mineau and Whiteside (2005). This is the obvious approach where full and comprehensive field data to build a predictive model are lacking. A benchmark approach was adopted by Mineau and Duffe (2001) for birds before the models described in Mineau (2002) were developed. In this approach, risk indices associated with specific mortality incidents were used to infer lethal risk with other untested pesticide uses. Similarly, Sheehan *et al*. (1995) developed benchmarks of acceptability for pesticide impacts on prairie slough based on selected studies investigating the level at which the loss of invertebrate biomass would affect

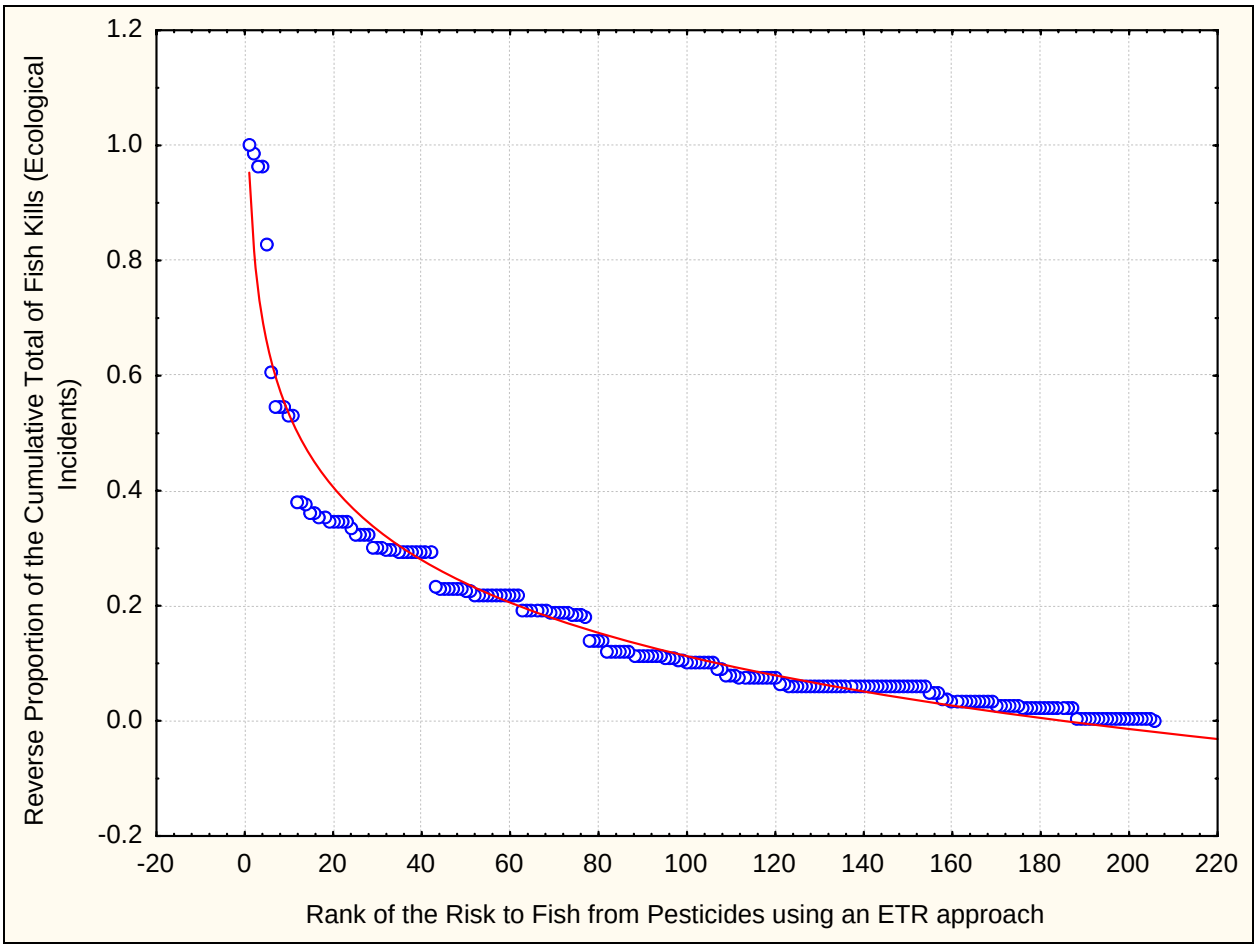
consumers. This approach may not provide for a linear scale of pesticide risk, but it does allow for the selection of an empirically determined level of acceptability rather than an arbitrarily chosen hazard index. Forty-eight out of 206 pesticides have an associated incident or incidents and, despite their heavy concentration in the top ranks of our fish hazard compilation, they are distributed throughout our rankings. The pesticide furthest down the ranked list with a recorded incident was fosetyl-aluminum. It was ranked 205/206 and the incident was related to an application to turf (golf course). However, 80% of recorded fish kills are found in the first 62 ranked compounds; 24 out of those 62 pesticides are associated with incidents. Figure 7 shows a plot of the cumulative proportion of incidents against rank. We propose to use the cumulative proportion of kills as the risk score. Like all previous risk scores and indices, this method has the advantage of scaling from 0 to 1, as well as showing a regular gradation of risk scores for all pesticides for which incidents were recorded. Interpretation of the risk score is also simple. For example, a risk score of 0.2 for pesticide X can be interpreted as follows: all pesticides considered to be of lower risk to fish than pesticide X have, as a group, been responsible for 20% of all fish kills recorded in the U.S.

The curve begins to level out around the 0.10 risk score mark (corresponding to a rank of 100), thus making it an appropriate place to set the standard. Any compounds below this level are associated with relatively few fish kills (45 kills distributed among 107 compounds), and thus can be considered as low risk. With this cut-off value of 0.10, many of the higher-risk pesticides (those associated with substantial numbers of fish kills) fail to meet the standard. About one-third of the pesticides above this mark have been noted to cause fish kills. Both atrazine and metolachlor are major-use compounds; it is therefore not surprising that a higher number of incidents would be recorded with them compared to other active ingredients with a similar or

even higher ranking. A group of the worst 11 pesticides are responsible for 50% of all fish kills. These compounds should be red-listed. We recognize that this is a preliminary step toward validating our risk-based ranking with actual recorded ecological incidents, given that the incident data is only a partial representation of what is happening in the field (because many incidents are not reported or observed). Nevertheless, we believe that the risk indices as defined here could form the basis of workable protection standards, whether ideal or currently achievable. However, one major difference between the fish risk score and all other risk scores presented in this report is that it is not dependent on application rate. Application rates associated with incidents are rarely if ever given and it is therefore impossible to see how closely they correspond to maximum label rates used in our compilation. Because we have no information on the application rates or,more importantly, on the runoff conditions associated with the kills, it is not possible to directly calibrate the ETR values with the kill information. We can attach some confidence to the relative rank of pesticides on the ETR scale, but the ETR value itself is not terribly meaningful. Until more information becomes available, standards will have to be defined by virtue of the relative hazard associated with each active ingredient.

A provisional fish standard is proposed based on the U.S. EPA record of pesticide fish kills. Pesticide active ingredients will be considered to have met the standard if their relative risk to fish (calculated from a risk quotient based on exposure modeling and a fish HC_5 value) is such that all pesticides of the same or lesser hazard are responsible for no more than 10% of all fish kills recorded by the U.S. EPA.

Figure 7: Inverse cumulative proportion of EIIS & California fish kills plotted against the ranked fish hazard index (ETR).



8 LIMITATIONS OF STANDARDS AND GAPS IN SCIENCE

8.1 Avian acute standard

Some of the uncertainties associated with the models have been noted already. In addition, the above assessment does not take into account sublethal impacts that result from pesticide exposure. Impacts that are initially sublethal are likely to increase the risk of avian mortality through a number of physiological mechanisms (Mineau, 2003). Whereas field records are reasonable for compounds applied as liquids, the same sort of analysis has not been performed for granular formulations (although a number of industry studies do exist) or for seed treatments (few

field studies exist). Comparisons of the risks of liquid applications relative to granular applications have always been difficult. The scoring system we have developed does attempt to make this comparison, although it is undoubtedly an aspect that could be improved further. Not helping the situation is the fact that a number of uncertainties exist regarding routes of exposure in the course of spray applications (Mineau, 2007a).

8.2 Avian reproductive standard

The lack of field verification is clearly the most important constraint. More new products fail the avian reproduction assessment screening than other parts of the ecological evaluation of pesticides. By necessity, our approach had to align itself with a regulatory approach. Barring the obtention of field data, the only possible improvements would necessitate using raw study results to separate the various reproductive effects (e.g., distinguishing between parental and embryonic effects), allowing for a population modeling approach. We were not in a position at this point to propose a standard for seed treatment or granular pesticides, arguably two groups of pesticides with a high potential for causing reproductive effects, because of the likelihood of high exposure.

8.3 Small mammal approach

The available toxicological information reinforces the notion that mammals and birds should be assessed separately (Mineau *et al*., 2006). The approach chosen does not provide protection to the individual wild mammal, but uses perturbations in population trajectories as an endpoint. An acknowledged limitation is the sample of compounds for which studies are available. Although they may be adequate to assess the impact of acutely toxic pesticides, they do not allow for a consideration of potential reproductive impacts in the absence of acute toxicity.

It is somewhat paradoxical that the mammalian index is so limited in scope given the amount of mammalian data generated for human safety assessments. However, as described earlier, there are

many questions about the ecological significance of many of the findings that drive human risk assessments (Mineau, 2005). Recent attempts to use information from chronic toxicity tests failed to help explain the limited field record (Mineau, 2007b). The summarized assessments that are commonly tabulated—whether in the form of RfDs (reference doses), ADIs (allowable daily intakes) or PADs (population-adjusted doses)—are based on various enpoints to which a number of safety factors have been applied. This is a protective measure that has evolved over time to ensure the safety of human workers, bystanders and consumers, but it is far from an accurate assessment of the expected harm to a wild mammal entering a treated field. Clearly, more work is required here.

8.4 Acute honeybee standard

Some pesticides (e.g., triazophos and dimethoate) caused a disproportionate number of incidents. In some cases, this may be due to variables not considered in the present model. For example, it was suggested that the high risk of dimethoate was a result of its systemic properties and its concentration in nectar (M.J. Smirle, Agriculture and Agri-Food Canada, pers. comm.).

For honeybees, the crop, as well as the method of spraying (e.g., foliar vs. ground spray), is expected to be a key risk element. The risk index calculated here assumes that exposure to pollinators is likely. Under normal circumstance, many of the products scored in Mineau *et al*. (2008b) would be expected to pose a minimal risk to honeybees merely based on the crop and/or application details. However, it has to be remembered that we are trying to assess the risk to wild pollinators typically foraging in field borders and other nearby non-crop areas potentially exposed to pesticide drift. The availability of toxicological data on non-*Apis* pollinators is needed to ascertain whether our standard has to be more stringent in order to protect species other than honeybees.

The logged risk score described here is different in meaning from a probability of impact, even if it obeys the same 0–1 structure. There may be ramifications to our ability to combine scores from different applications that we have not yet explored.

In a separate report (Harding *et al*., 2006), we attempted to use honeybee toxicity as a predictor of impacts on beneficial insects and spiders. An empirical database of published field studies (SELECTV) was mined and modeling attempted. In all of the derived models, the hazard ratio using honey bee LD_{50} was important in predicting mortality, with the exception of spider mortality. However, it was shown that hazard risk alone did not predict mortality, except for groups which feed on pollen when other prey is scarce. We did not find an acceptable model to predict mortality in all invertebrate species without including categorical factors such as crop type and invertebrate family or feeding guild. The sheer complexity of having to deal with so many different models forced us to abandon our attempts at deriving a simple risk score for toxicity to beneficial or terrestrial invertebrates at large.

8.5 Acute earthworm standard

Multiple applications and synergistic effects between substances were not accounted for in this model (or any other standard presented here). Including these factors would result in higher predicted mortalities.

Also, as noted above, most earthworm species, as well as other soil-dwelling invertebrates, seem on average to be more sensitive to pesticides than the standard test species. The proposed standard is unlikely to be protective enough for these other groups. A correction factor may need to be applied.

8.6 Aquatic standards

A critical requirement for applying risk-based aquatic standards at the field level is to generate realistic water concentrations for the chosen exposure scenario—a pond or other water body at the field edge. The way in which this was done here is not wholly satisfactory. Ideally, we should have enough of a dataset for the various agricultural regions of the country to apply empirically based corrections to our chosen runoff modeling (GENEEC is used here but others are possible). Alternatively, we might consider running more sophisticated modeling (as was done for Manitoba conditions) for all agricultural regions, although the question of calibration/validation will continue to be an issue. Clearly, the unique placement of sensitive water bodies near individual fields is difficult to model exactly. Also, we need to be able to separately account for drift exposure because it will obviously dominate in some instances.

Based on HD_5 values, aquatic insects are frequently more sensitive than Crustacea. Unfortunately, there were not enough pond studies with good insect data to allow for modeling of loss rates—the technique used here to generate the standard. In the same way that extrapolation may be required to protect non-*Apis* pollinators or soil invertebrates other than *Lumbricus* or *Eisenia,* it may prove necessary to apply correction factors to the aquatic standards in order to be protective of non-Crustacea species.

8.7 Limitations shared by most standards

A limitation shared by most of the standards is the inability to account for ecological recovery. Reasons for not including recovery—essentially a lack of information or data—have been discussed where relevant, but this clearly remains an area of further development. Similarly, all of our indices have concentrated on direct and, usually, acute effects. Delayed effects and indirect effects resulting from pesticide impacts are certainly possible (Liess *et al*., 2005), but these data

have not been generated consistently and thus cannot be used yet in an exercise such as this one.

9 FUTURE CONSIDERATIONS AND SUMMARY

Several possible improvements have already been noted in the previous section. The reader should be mindful of the fact that all of the models presented in this report represent somewhat severe application conditions and may not take into consideration farmers' current efforts to reduce their ecological footprint. There are many best management practices (BMP) used by farmers across the country and incorporating more site- and use-specific information into these risk indices is an important next step.

All of the calculations that went into the various impact measurement tools, although complex in appearance, are amenable to a spreadsheet environment. In Mineau *et al*. (2008b), we provide some examples of how the impact scores can be used to quickly characterize pesticide use for a given agricultural commodity. We envision a Web-enabled system that would automate the calculation of risk scores based on pesticide use pattern information, and that would provide growers with immediate feedback on their pest control choices.

In addition, once a standardized and centralized source of pesticide use data becomes available, it should be possible to take regional agricultural snapshots anywhere, regardless of crop, in order to determine overall risk to birds, small mammals, bees, earthworms, fish and aquatic invertebrates. With a powerful tool such as this, product substitution or best management practices can be implemented where and when they are needed to ensure that pesticide use meets a basic standard of environmental sustainability.

The research and analyses we have carried out in the course of the NAESI project have enabled us to develop new tools and approaches to assess the likely environmental impact of pesticide use. These tools, in turn, have allowed us to propose risk management-based standards for several

segments of the terrestrial and aquatic environments. The Table in the Key Findings section summarizes the different standards being proposed.

10 ACKNOWLEDGMENTS

We are grateful for the support of the NAESI initiative that allowed us to perform this work. Elizabeth Roberts and Paul Jiapizian handled the budgetary issues and let us get on with the work.

Peter Delorme, Connie Hart, Vasiliki Flari, Mike Smirle, and Nick Mastrota offered constructive comments on earlier reports or sections of the present report.

The senior author greatly benefited from interactions with other members of the European Food Safety Authority Committee involved in the drafting of the new guidance document for bird and mammal assessments in Europe, notably Dr. Andy Hart.

We are also very grateful to Ian Kennedy, who rewrote GENEEC to allow batch entries.

Greg Malis, Connie Hart, Ian Kennedy and Lizanne Avon (PMRA) and Ed Odenkirchen (U.S. EPA) helped us with the intricacies of the assessment methodology in use by those respective regulatory agencies.

Peter Delorme, Scott Findlay and Mark Forbes were advisors to the student responsible for NAESI Technical Series Report No. 3-31, which was an essential step in modeling the pond and mesocosm studies.

The NAHARP IROC-Pest Group (A.J. Cessna, A. Farenhorst, D.A.R. McQueen and C. Sheedy) kindly agreed to let us hijack Ross McQueen to help with the modeling of water concentrations from PRZM and provided useful comments. We would similarly like to thank the Pest Management Centre of Agriculture Agri-Food Canada, especially Tim MacDonald, for allowing Dr. Tobias Längle to assist us with the new work on the earthworm standard. Tim, Martin

Beaulieu and Sheila Young (Statistics Canada) were also instrumental in our use of the pesticide use survey in carrots.

Brian Collins was a key advisor on all matters mathematical and statistical over the last four years.

Several individuals have contributed to gathering toxicity endpoints over the years. They include Gerard Grolleau, Robert Luttik, Gerhard Joerman, Karen Benbrook and Brian Montague. Without Brian and the U.S. EPA's one-liner database, much of our work would not have been possible. Stephan Jänsch and Geoff Frampton were very generous in sharing the earthworm data compiled in the course of the webfram project. In-house, a large number of summer students including Julia Duinker, Stephen Weber and Omar Mahfoudi were responsible for hunting down and entering data. Alain Baril was also instrumental in this task.

Nick Mastrota (U.S. EPA) and Bob Hosea (State of California) provided the fish kill information; Mark Fletcher and David Garthwaite provided the data necessary for the analysis of U.K. bee kills. Paul Jepson was kind enough to give us access to the SELECTV database.

Chuck Benbrook and Jennifer Curtis provided the PEAs framework and forum for the development of Pesticide Use Adjustment Factors. Several individuals were consulted and offered comments on the avian factors, notably Rick Bennett, Peter Edwards, Niels Elmegaard, David Fischer, Andy Hart, Mike Hooper, Gerhard Joermann, Robert Luttik, Miguel Mora and Joe Sullivan.

Of course, any errors of omission or commission, as well as the opinions expressed herein are the sole responsibility of the authors.

11 REFERENCES

Aldenberg, T. and R. Luttik. 2002. Extrapolation factors for tiny toxicity data sets from species sensitivity distributions with known standard deviation In: L. Posthuma, G.W. Suter II, and T.P. Traas (eds.). *Species Sensitivity Distributions in Ecotoxicology*. Lewis Publishers, New York, USA. pp. 103–118.

Atkins, E.L., D. Kellum, and K.W. Atkins. 1981. Reducing pesticide hazards to honey bees: Mortality prediction techniques and integrated management strategies. Available online at: <http://bees.ucr.edu/reprints/2883.pdf>

Baveco, J. M. and A.M.D. Roos. 1996. Assessing the impact of pesticides on Lumbricid populations: an individual based modelling approach. *The Journal of Applied Ecology*. 33(6): 1451–1468.

Bennett, R.S., I.C. Dewhurst, A. Fairbrother, A.D.M. Hart, M.J. Hooper, A. Leopold, P. Mineau, S.R. Mortensen, R.F. Shore, and T.A. Springer. 2005. A new interpretation of avian and mammalian reproduction toxicity test data in ecological risk assessment. *Ecotoxicology*. 14(8): 801–815.

Block, E.K., T.E. Lacher Jr., L.W. Brewer, G.P. Cobb III, and R.J. Kendall. 1999. Population responses of Peromyscus resident in Iowa cornfields treated with the organophosphorus pesticide counter. *Ecotoxicology*. 8: 189–200.

Callahan, J. and P. Mineau, 2007. An Evaluation of Clinical Sign Data from Avian Acute Oral Toxicity Studies. Guidance document of the Scientific Panel on Plant Protection Products and Their Residues for the Risk Assessment for Birds and Mammals under Council Directive 91/414/EEC, Appendix 10. Draft for public consultation. 9 p. Available online at: <http://www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1178660557918.htm>

Caux, P.-Y. and P. Jiapizian. 2004. Reducing Pesticide Environmental Risks: The National Agri-Environmental Standards Initiative (NAESI). Presentation: Aquatic Toxicity Workshop, Charlottetown, P.E.I., October 24–27, 2004.

CEI (Cantox Environmental Inc.). 2005. *Scoping Assessment of Environmental Performance Standards for in-use Pesticides from Agricultural Sources*. National Agri-Environmental Standards Initiative Technical Series Report No. 1-18. 253 p.

Delorme, P., D. François, C. Hart, V. Hodge, G. Kaminski, C. Kriz1, H. Mulye, R. Sebastien, P. Takacs, and F. Wandelmaier. 2005. *Final Report for the PMRA Workshop: Assessment Endpoints for Environmental Protection*. Unpublished Report for the Pest Management Regulatory Agency. Ottawa, Ont. 63 p.

De Snoo, G.R. 2003. Variations in agricultural practice and environmental care. In: F. den Hond, P. Groenewegen, and N.M. van Straalen (eds). *Pesticides: Problems, Improvements, Alternatives*. Blackwell Science. Oxford, U.K. pp. 100–112.

Driver C.J, M.W. Ligotke, P. Van Voris, B.D. McVeety, B.J. Greenspan, and D.B. Drown. 1991. Routes of uptake and their relative contribution to the toxicologic response of northern bobwhite (*Colinus virginianus*) to an organophosphate pesticide. *Environmental Toxicology and Chemistry.* 10: 21–33.

Edwards, C.A. (ed.). 2004. *Earthworm Ecology*. Second Edition. CRC Press.

Edwards, C.A., A.R. Thompson, and J.R. Lofty. 1967. Changes in soil invertebrate populations caused by some organophosphate insecticides. *Proceedings of the Fourth British Insecticide Fungicide Conference.* 1: 48–55.

EPPO (European and Mediterranean Plant Protection Organization). 2003. *EPPO Standards: Environmental Risk Assessment Scheme for Plant Protection Products*. PP 3/7 (revised). European and Mediterranean Plant Protection Organization Bulletin 33: 195–209.

EFSA (European Food Safety Authority). 2007. *Guidance Document of the Scientific Panel on Plant Protection Products and Their Residues for the Risk Assessment for Birds and Mammals under Council Directive 91/414/EEC.* Draft for public consultation. Available online at: <http://www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1178660557918.htm>

European Commission. 2002. *Guidance Document on Terrestrial Ecotoxicology under Council Directive 91/414/EEC*. European Commission, Health and Consumer Protection Directorate-General (draft – working document). Report Number: SANCO/10329/2002 Rev 2 final. 17 October 2002.

Frampton, G.K., S. Jänsch, J.J. Scott-Fordsmand, J. Rombke, and P.J. Van den Brink. 2006. Effects of pesticides on soil invertebrates in laboratory studies: a review and analysis using species sensitivity distributions. *Environmental Toxicology and Chemistry.* 25(9): 2480–2489.

Frelich, L.E., C.M. Hale, S. Scheu, A.R. Holdsworth, L. Heneghan, P.J. Bohlen, and P.B. Reich. 2006. Earthworm invasion into previously earthworm-free temperate and boreal forests. *Biological Invasions.* 8(6): 1235–1245.

Harding K., P. Mineau, M. Whiteside, M.R. Fletcher, and D. Garthwaite. 2006. *Developing Risk-based Rankings for Pesticides in Support of Standard Development at Environment Canada*. National Agri-Environmental Standards Initiative Technical Series Report No. 2-44. 89 p.

Harding, K., P. Mineau, M. Whiteside, P. Jepson, and T. Dawson. 2007. *Developing Risk-based Rankings for Pesticides in Support of Standard Development at Environment Canada: Predicting Non-target Invertebrate Mortality to Pesticides Using Honeybee Toxicity Values and Physicochemical Properties of the Pesticides*. National Agri- Environmental Standards Initiative Technical Series Report No. 3-30. 69 p.

ISO (International Standards Organization). 1999. *Soil Quality – Effects of Pollutants on Earthworms, Part 3: Guidance on the Determination of Effects in Field Situations*. International Standards Organisation, Geneva.

Jänsch, S., G.K. Frampton, J. Rombke, P.J. Van den Brink, and J.J. Scott-Fordsmand. 2006. Effects of pesticides on soil invertebrates in model ecosystem and field studies: A review and comparison with laboratory toxicity data. *Environmental Toxicology and Chemistry.* 25(9): 2490–2501.

Johnson, I.P., J.R. Flowedew, and R. Hare. 1991. Effects of broadcasting and drilling methiocarb molluscicide pellets on field populations of wood mice, *Apodemus sylvaticus*. *Bulletin of Environmental Contamination and Toxicology.* 46: 84–91.

Kreuger, J. and L. Tornqvist. 1998. Multiple regression analysis of pesticide occurrence in streamflow related to pesticide properties and quantities applied. *Chemosphere.* 37: 189–207.

Kula, H. 1995. Comparison of laboratory and field testing for the assessment of pesticide side effects on earthworms. *Acta Zoologica Fennica.* 196: 338–341.

Levitan, L. 2000. 'How to' and 'why': Assessing the enviro-social impacts of pesticides. *Crop Protection.* 19: 629–636.

Liess, M., C. Brown, G.P. Dohmen, S. Duquesne, A. Hart, F. Heimbach, J. Kreuger, L. Lagadic, S. Maund, W. Reinert, M. Streloke, and J.V. Tarazona (eds.). 2005. *Effects of Pesticides in the Field*. Society of Environmental Toxicology and Chemistry (SETAC) Press, Brussels, Belgium. 136 p.

Lock, K. 2002. The effect of lindane on terrestrial invertebrates. *Archives of Environmental Contamination and Toxicology.* 42(2): 217–221.

Lock, K. and C.R. Janssen. 2001. Cadmium toxicity for terrestrial invertebrates: taking soil parameters affecting bioavailability into account. *Ecotoxicology.* 10(5): 315–322.

Luttik, R., P. Mineau, and W. Roelofs. 2005. A review of interspecies toxicity extrapolation in birds and mammals and a proposal for long-term toxicity data. *Ecotoxicology.* 14(8): 817–832.

McQueen, D.A.R., A. Farenhorst, S. Allaire, and A.J. Cessna. 2007. Automation and evaluation of three pesticide fate models for a national analysis of leaching risk in Canada. *Canadian Journal of Soil Science*. Special Issue 2007. 87(2): 203–212.

Merriman, J.C., J. Struger, and R.S. Szawiola. 1991. Distribution of 1,3-dichloropropene and 1,2-dichloropropane in Big Creek watershed. *Bulletin of Environmental Contamination and*

Toxicology. 47: 572–579.

Mineau, P. 2002. Estimating the probability of bird mortality from pesticide sprays on the basis of the field study record. *Environmental Toxicology and Chemistry.* 24(7): 1497–1506.

Mineau, P. 2003. Avian species. *In*: Plimmer, J.R., D.W. Gammon, and N.N. Ragsdale, eds.,*Encyclopedia of Agrochemicals. Vol. 1*. Wiley Interscience. pp. 129–156.

Mineau, P. 2004a. PEAS-AR: An index of pesticide acute risks to birds. Version 2004.2. Report to Benbrook Consulting dated 24 August 2004 (unpublished). 9 p.

Mineau, P. 2004b. Birds and pesticides: Are pesticide regulatory decisions consistent with the protection afforded migratory bird species under the Migratory Bird Treaty Act? *The William and Mary Environmental Law and Policy Review*. 28(2): 313–338.

Mineau, P. 2005. A review and analysis of study endpoints relevant to the assessment of long-term pesticide toxicity in avian and mammalian wildlife. *Ecotoxicology.* 14(8): 775–799.

Mineau, P. 2007a. *Developing Risk-based Rankings for Pesticides in Support of Standard Development at Environment Canada: Risk-based Approach for Terrestrial Biota Continued: Incorporating Dermal Exposure in Pesticide Risk Assessments for Birds.* National Agri-Environmental Standards Initiative Technical Series Report No. 3-32. 112 p.

Mineau, P. 2007b. Attempts to Validate/Calibrate Mammal Risk Assessments with Field Data. Guidance Document of the Scientific Panel on Plant Protection Products and Their Residues for the Risk Assessment for Birds and Mammals under Council Directive 91/414/EEC, Appendix 18. Draft for public consultation. 16 p. Available online at: <http://www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1178660557918.htm>

Mineau, P., A. Baril, B.T. Collins , J. Duffe, G. Joerman, and R. Luttik. 2001. Reference values for comparing the acute toxicity of pesticides to birds. *Reviews of Environmental Contamination and Toxicology.* 170: 13–74.

Mineau, P., B.T. Collins, and A. Baril. 1996. On the use of scaling factors to improve interspecies

extrapolation of acute toxicity in birds. *Regulatory Toxicology and Pharmacology*. 24: 24–29.

Mineau, P., T. Dawson, M. Whiteside, C. Morrison, K. Harding, L. Singh, T. Längle, and D.A.R. McQueen. 2008b. *Environmental Risk-based Standards for Pesticide Use in Canada: Supporting Tables for the Synthesis Report and Proof of Concept*. National Agri-Environmental Standards Initiative Technical Series Report No. 4-22. 151 p.

Mineau, P. and J. Duffe. 2001. Birds and insecticide toxicoses: A preliminary risk assessment for Israel. In: Shlosberg, A. and O. Bahat (eds.). *The Risks of Toxicoses from Pesticides and Pollutants in Raptors and Wildlife in Israel*. Published by the Society for the Protection of Nature in Israel. Tel Aviv. pp. 30–59.

Mineau, P., K.M. Harding, M. Whiteside, M.R. Fletcher, D. Garthwaite, and L.D. Knopper. 2008a. Using reports of bee mortality in the field to calibrate laboratory-derived pesticide risk indices. *Journal of Environmental Entomology*. 37(2): 546–554.

Mineau, P., B. Jobin, and A. Baril. 1994. *A Critique of the Avian 5-day Dietary Test (LC_{50}) as the Basis of Avian Risk Assessment*. Technical Report No. 215. Canadian Wildlife Service, Headquarters.

Mineau, P., C. Morrison, M. Whiteside, and K. Harding. 2006. *Developing Risk-based Rankings for Pesticides in Support of Standard Development at Environment Canada: Preliminary Terrestrial Rankings*. National Agri-Environmental Standards Initiative Technical Series Report No. 2-43. 92 p.

Mineau, P. and M. Whiteside. 2005. Development of Comparative Environmental Risk Assessment Tools for Pesticides in Support of Standard Development at Environment Canada. National Agri-Environmental Standards Initiative Technical Series Report No. 1-17. 199 p.

Mineau, P. and M. Whiteside. 2006. The lethal risk to birds from insecticide use in the U.S.: A spatial and temporal analysis. *Environmental Toxicology and Chemistry*. 25(5): 1214–1222.

OECD (Organisation for Economic Cooperation and Development). 1984. Guideline 207. Earthworm, Acute Toxicity Tests. Guidelines for Testing Chemicals, Section 2: Effects on Biotic Systems. Organisation for Economic Cooperation and Development. 9 p.

Parmelee, R.W., M.H. Beare, W. Cheng, P.F. Hendrix, S.J. Rider, D.A. Crossley, and D.C. Coleman. 1990. Earthworms and enchytraeids in conventional and no-tillage agroecosystems: A biocide approach to assess their role in organic matter breakdown. *Biology and Fertility of Soils.* 10(1): 1–10.

Peck, G.K. and R.D. James. 1987. Breeding Birds of Ontario. Nidiology and Distribution. Volume 2: Passerines. Royal Ontario Museum. Toronto, Ontario. 387 p.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish. EPA/444/4-89-001. United States Environmental Protection Agency. Washington, D.C.

Posthuma, L., G.W. Suter II, and T.P. Traas. 2002. Species Sensitivity Distributions in Ecotoxicology. Lewis Publishers, Boca Raton, Florida. 587 p.

Potter, D.A., M.C. Buxton, C.T. Redmond, C.G. Patterson, and A.J. Powell. 1990. Toxicity of pesticides to earthworms (*Oligochaeta: Lumbricidae*) and effect on thatch degradation in Kentucky bluegrass turf. *Journal of Economic Entomology.* 83(6): 2362–2369.

Reus, J., P. Leendertse, C. Bockstaller, I. Fomsgaard, V. Gutsche, K. Lewis, C. Nilsson, L. Pussemier, M. Trevistan, H. van der Werf, F. Alfarroba, S. Blümel, J. Isart, D. McGrath, and T. Seppälä. 1999. Comparing Environmental Risk Indicators for Pesticides. Results of the European CAPER Project. Centre for Agriculture and Environment (CML). The Netherlands. 184 p.

Reus, J., P. Leendertse, C. Bockstaller, I. Fomsgaard, V. Gutsche, K. Lewis, C. Nilsson, L. Pussemier, M. Trevistan, H. van der Werf, F. Alfarroba, S. Blümel, J. Isart, D. McGrath, and T. Seppälä. 2002. Comparison and evaluation of eight pesticide environmental risk indicators developed in Europe and recommendations for future use. *Agriculture,*

Ecosystems and Environment. 90: 177–187.

Richards, N.L., P. Mineau, and D.M. Bird. 2004. Exposure of the eastern screech-owl (*Otus asio*) to organophosphorus insecticides and anticoagulant rodenticides in apple orchards of southern Quebec, Canada. In: R.D. Chancellor and B.-U. Meyburg (eds.). *Raptors Worldwide.* World Working Group on Birds of Prey. pp. 389–408.

Robidoux, P.Y., J. Hawari, S. Thiboutot, G. Ampleman, and G.I. Sunahara. 1999. Acute toxicity of 2, 4, 6-trinitrotoluene in earthworm (*Eisenia andrei*). *Ecotoxicology and Environmental Safety.* 44(3): 311–321.

Robidoux, P.Y., C. Svendsen, J. Caumartin, J. Hawari, G. Ampleman, S. Thiboutot, J.M. Weeks, and G.I. Sunahara. 2000. Chronic toxicity of energetic compounds in soil determined using the earthworm (*Eisenia andrei*) reproduction test. *Environmental Toxicology and Chemistry.* 19(7): 1764–1773.

Robidoux, P.Y., G.I. Sunahara, K. Savard, Y. Berthelot, S. Dodard, M. Martel, G. Ping, and J. Hawari. 2004. Acute and chronic toxicity of the new explosive CL-20 to the earthworm (*Eisenia andrei*) exposed to amended natural soils. *Environmental Toxicology and Chemistry.* 23(4): 1026–1034.

Roessink, I., G.H.P. Arts, J.D.M. Belgers, F. Bransen, S.J. Maund, T.C.M. Brock. 2005. Effects of lambda-cyhalothrin in two ditch microcosm systems of different trophic status. *Environmental Toxicology and Chemistry.* 24(7): 1684–1696.

Römbke, J., C.A.M. Van Gestel, S.E. Jones, J. E. Koolhaas, J.M.L. Rodrigues, and T. Moser. 2004. Ring-testing and field-validation of a Terrestrial Model Ecosystem TME – An instrument for testing potentially harmful substances. Effects of carbendazim on earthworms: Terrestrial Model Ecosystems. *Ecotoxicology.* 13: 105–118.

Sample, B.E. and C.A. Arenal. 1999. Allometric models for interspecies extrapolation of wildlife toxicity data. *Bulletin of Environmental Contamination and Toxicology.* 62: 653–663.

Sheehan, P., A. Baril, P. Mineau, and D. Paustenbach. 1995. Predicting the effects of insecticides

on aquatic systems and the waterfowl that use them. In: G.M. Rand (ed.). *Fundamentals of Aquatic Toxicology*. Second Edition. Taylor and Francis. North Palm Beach, Florida. pp. 827–857.

Sheppard, S. Bembridge, J., Holmstrup, M., and L. Posthuma (eds.). 1998. Advances in Earthworm Ecotoxicology: Proceedings from the Second International Workshop on Earthworm Ecotoxicology, April 1997. Amsterdam, the Netherlands. Society for Environmental Toxicology and Contamination, (SETAC) Press, Pensacola, Florida.

Shires, S.W. 1985. A comparison of the effects of cypermethrin, parathion-methyl and DDT on cereal aphids, predatory beetles, earthworms and litter decomposition in spring wheat. *Crop Protection.* 4(2): 177–193.

Shore, R.F., D.R. Crocker, H.R. Akcakaya, R.S. Bennett, P.F. Chapman, M. Clook, M. Crane, I.C. Dewhurst, P.J. Edwards, A. Fairbrother, S. Ferson, D. Fischer, A.D.M. Hart, M. Holmes, M.J. Hooper, M. Lavine, A. Leopold, R. Luttik, P. Mineau, D.R.J. Moore, S.R. Mortenson, D.G. Noble, R.J. O'connor, W. Roelofs, R.M. Sibly, G.C. Smith, M. Spendiff, T. A. Springer, H.M. Thompson, and C. Topping. 2005. Case study part 1: how to calculate appropriate deterministic long-term Toxicity to Exposure Ratios (TERs) for birds and mammals. *Ecotoxicology.* 14(8): 877–893.

Singh, L. 2007. *Developing Risk-based Rankings for Pesticides in Support of Standard Development at Environment Canada: Predicting Effects of Pesticides in Freshwater Aquatic Ecosystems Using Field Data Validation.* National Agri-Environmental Standards Initiative Technical Series Report No. 3-31. 207 p.

Smith, G.K. 2006. Risks to birds from pesticide-treated seed and the possible role of ultraviolet reflection in seed colour preferences and repellent strategies. MSc. Thesis (unpublished). Carleton University. Ottawa, Ontario. 175 p.

Spurgeon, D.J., J.M. Weeks, and C.A.M. Van Gestel. 2003. A summary of eleven years progress in earthworm ecotoxicology. *Pedobiologia.* 47(5/6): 588–606.

Suter II, G.W., R.A. Efroymson, B.E. Sample, and D.S. Jones. 2000. Ecological Risk Assessment

for Contaminated Sites. Lewis Publishers, Boca Raton, Florida.

Tasei, J.N. 2002. Impact of agrochemicals on non-*Apis* bees. In: Devillers, J. and M.H. Pham-Delègue (eds). Honey Bees: Estimating the Environmental Impact of Chemicals. Taylor and Francis. New York, N.Y. pp. 101–131.

Thompson, A.R. and W.W. Sans. 1974. Effects of soil insecticides in southwestern Ontario on non-target invertebrates: Earthworms in pasture. *Environmental Entomology.* 3: 305–308.

Thompson, H.M. 2001. Assessing the exposure and toxicity of pesticides to bumblebees (*Bombus* sp.). *Apidologie.* 32: 305–321.

Thompson, H.M. and L.V. Hunt. 1999. Extrapolating from honeybees to bumblebees in pesticide risk assessment. *Ecotoxicology.* 8: 147–166.

Tomlin, A.D. 1981. Effects on soil fauna of the fungicide, Benomyl, used to control earthworm populations around an airport. *Protection Ecology.* 2: 325–330.

Tomlin, A.D. and F.L. Gore. 1974. Effects of six insecticides and a fungicide on the numbers and biomass of earthworms in pasture. *Bulletin of Environmental Contamination and Toxicology.* 12(4): 487–492.

Tomlin, C.D.S. 2003. The Pesticide Manual. British Crop Protection Council. Alton, Hampshire, U.K.

U.S. EPA (United States Environmental Protection Agency). 2000. Development and Use of GENEEC Version 2.0 for Pesticide Aquatic Ecological Exposure Assessment. Ron Parker (unpublished). United States Environmental Protection Agency.

van Vlaardingen, P.L.A., T.P. Traas, A.M. Wintersen, and T. Aldenberg. 2004. Etx2.0. A Program to Calculate Hazardous Concentrations and Fraction Affected, Based on Normally-distributed Toxicity Data. RIVM report (and software) 601501028/2004. National Institute for Public Health and the Environment, The Netherlands.

Voronova, L.D. 1968. The effect of some pesticides on the soil invertebrate fauna in the south

Taiga zone in the perm region (USSR). *Pedobiologia.* 8: 507–527.

Whiteside, M., P. Mineau, C. Morrison, and K. Harding. 2006. *Developing Risk-based Rankings for Pesticides in Support of Standard Development at Environment Canada: Setting Pesticide Priorities for Aquatic Ideal Performance Standards and Ranking the Aquatic Hazard of in-use Pesticides in Canada*. National Agri-Environmental Standards Initiative Technical Series Report No. 2-42. 173 p.

Whiteside, M., P. Mineau, C. Morrison, and L.D. Knopper. 2008. Comparison of a score-based approach with risk based ranking of in-use agricultural pesticides in Canada to aquatic receptors. *Integrated Environmental Assessment and Management.* 4(2): 215–236.

Wilson, L.K., J.E. Elliott, R.S. Vernon, D.D. Smith, and S.Y. Szeto. 2002. Persistence and retention of active ingredients in four granular cholinesterase-inhibiting insecticides in agricultural soils of the Lower Fraser River Valley, British Columbia, Canada, with implications for wildlife poisoning. *Environmental Toxicology and Chemistry.* 21(2): 260–262.