Environment Canada Environnement Canada

National Agri-Environmental Standards Initiative (NAESI)

Report No. 3-30

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

Technical Series 2007

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:

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.

Prepared and published by
Environment Canada
Gatineau, QC

December 2007

NATIONAL AGRI-ENVIRONMENTAL STANDARDS INITIATIVE TECHNICAL SERIES

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.

REPORT NO. 3-30

NOTE TO READERS

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

- 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) will be established that identify both desired levels of environmental condition and levels considered achievable based on available technology and practice. These standards will be integrated by AAFC into beneficial agricultural management systems and practices to help reduce environmental risks. Additionally, these will provide benefits to the health and supply of water, health of soils, health of air and the atmosphere; and ensure compatibility between biodiversity and agriculture. Standards are being developed in four thematic areas: Air, Biodiversity, Pesticides, and Water. Outcomes from NAESI will contribute to the 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.

The development of agri-environmental performance standards involves science-based assessments of relative risk and the determination of desired environmental quality. As such, the National Agri-Environmental Standards Initiative (NAESI) Technical Series is dedicated to the consolidation and dissemination of the scientific knowledge, information, and tools produced through this program that will be used by Environment Canada as the scientific basis for the development and delivery of environmental performance standards. Reports in the Technical Series are available in the language (English or French) in which they were originally prepared and represent theme-specific deliverables. As the intention of this series is to provide an easily navigable and consolidated means of reporting on NAESI's yearly activities and progress, the detailed findings summarized in this series may, in fact, be published elsewhere, for example, as scientific papers in peer-reviewed journals.

This report provides scientific information to partially fulfill deliverables under the Pesticide Theme of NAESI. This report was written by K. Harding, P. Mineau, M. Whiteside, P. Jepson, and T. Dawson of Environment Canada. The report was edited and formatted by Denise Davy to meet the criteria of the NAESI Technical Series. The information in this document is current as of when the document was originally prepared. For additional information regarding this publication, please contact:

Environment Canada
National Agri-Environmental Standards
Initiative Secretariat
351 St. Joseph Blvd. 8th floor
Gatineau, QC
K1A 0H3
Phone: (819) 997-1029
Fax: (819) 953-0461

NOTE À L'INTENTION DES LECTEURS

L'Initiative nationale d'élaboration de normes agroenvironnementales (INENA) est un projet de quatre ans (2004-2008) mené conjointement par Environnement Canada (EC) et Agriculture et Agroalimentaire Canada (AAC) et l'une des nombreuses initiatives qui s'inscrit dans le Cadre stratégique pour l'agriculture (CSA) d'AAC. Elle a notamment comme objectifs :

- d'établir des normes nationales de rendement environnemental non réglementaires (applicables dans les régions) qui soutiennent les objectifs communs d'EC et d'AAC en ce qui concerne l'environnement;
- d'évaluer des normes qui sont réalisables par des pratiques de production et de gestion agricoles avantageuses pour l'environnement;
- de faire mieux comprendre les liens entre l'agriculture et l'environnement.

Dans le cadre de l'INENA, des normes de rendement agroenvironnementales (c.-à-d. des normes axées sur les résultats) seront établies pour déterminer les niveaux de qualité environnementale souhaités et les niveaux considérés comme réalisables au moyen des meilleures technologies et pratiques disponibles. AAC intégrera ces normes dans des systèmes et pratiques de gestion bénéfiques en agriculture afin d'aider à réduire les risques pour l'environnement. De plus, elles amélioreront l'approvisionnement en eau et la qualité de celle-ci, la qualité des sols et celle de l'air et de l'atmosphère, et assureront la compatibilité entre la biodiversité et l'agriculture. Des normes sont en voie d'être élaborées dans quatre domaines thématiques : l'air, la biodiversité, les pesticides et l'eau. Les résultats de l'INENA contribueront aux objectifs du CSA, soit d'améliorer la gérance des terres, de l'eau, de l'air et de la biodiversité par les producteurs agricoles et d'accroître la confiance du Canada et d'autres pays dans le fait que les aliments produits par les agriculteurs et le secteur de l'alimentation du Canada le sont d'une manière sécuritaire et soucieuse de l'environnement.

L'élaboration de normes de rendement agroenvironnementales comporte des évaluations scientifiques des risques relatifs et la détermination de la qualité environnementale souhaitée. Comme telle, la Série technique de l'INENA vise à regrouper et diffuser les connaissances, les informations et les outils scientifiques qui sont produits grâce à ce programme et dont Environnement Canada se servira comme fondement scientifique afin d'élaborer et de transmettre des normes de rendement environnemental. Les rapports compris dans la Série technique sont disponibles dans la langue (français ou anglais) dans laquelle ils ont été rédigés au départ et constituent des réalisations attendues propres à un thème en particulier. Comme cette série a pour objectif de fournir un moyen intégré et facile à consulter de faire rapport sur les activités et les progrès réalisés durant l'année dans le cadre de l'INENA, les conclusions détaillées qui sont résumées dans la série peuvent, en fait, être publiées ailleurs comme sous forme d'articles scientifiques de journaux soumis à l'évaluation par les pairs.

Le présent rapport fournit des données scientifiques afin de produire en partie les réalisations attendues pour le thème des pesticides dans le cadre de l'INENA. Ce rapport a été rédigé par K. Harding, P. Mineau, M. Whiteside, P. Jepson, et T. Dawson d'Environnement Canada. De plus, il a été révisé et formaté par Denise Davy selon les critères établis pour la Série technique de l'INENA. L'information contenue dans ce document était à jour au moment de sa rédaction. Pour plus de renseignements sur cette publication, veuillez communiquer avec l'organisme suivant :

Secrétariat de l'Initiative nationale d'élaboration de normes agroenvironnementales
Environnement Canada
351, boul. St-Joseph, 8eétage
Gatineau (Québec) K1A 0H3
Téléphone : (819) 997-1029
Télécopieur : (819) 953-0461

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1 INTRODUCTION AND EXECUTIVE SUMMARY

Environment Canada (EC) has been tasked with developing environmental standards for implementation in Agriculture and Agri-Food Canada's Agricultural Policy Framework (AAFC; APF). The Wildlife Toxicology Division of the Wildlife and Landscape Science Directorate of EC's Science and Technology Branch was tasked specifically with developing comparative environmental risk assessment tools for pesticides in support of standard development. This will provide environmentally-oriented advice to AAFC under the APF, allowing for the promotion of reduced risk pest management strategies. Furthermore, standardised pesticide assessment tools will enable EC to objectively assess the environmental impact of alternative pesticide products and prioritize them for research and monitoring.

Agricultural pesticide use is a major threat to beneficial insects; the removal of which can have a severe economical impact on the crop. Without the beneficial insects, secondary pests which are normally a food source for these beneficial insects will greatly increase in number, thus creating a stronger dependence on pesticides. Non-target invertebrates are also important as a food source for birds and other wildlife.

Although the submission of information on non-target invertebrate toxicity is becoming more common, this information is still lacking for most of the pesticides currently in use. The only species which data are routinely available and made public is the European honey bee. Our objective was to test whether honeybee toxicity data could be used to predict impacts on other groups of non-target insects.

In 1988, as a part of a master's thesis, a database called SELCTV was created to combine the published literature on pesticide side-effects on beneficial arthropods. SELECTV contains published articles ranging from 1911 until 1999 with the majority of articles published from 1970 – 1985 (68%). This database was specifically constructed to predict the susceptibility between different arthropod species, including pests, predators, and parasitoids. Within this database, the entries contained both laboratory dosing trials and field application data. For the purposes of our analysis, all laboratory data were excluded from analysis so the models we derived would better reflect field conditions.

Bee toxicity data included oral and contact LD_{50} values. These values were used to derive hazard ratios (HR), which in turn were used for statistical modeling. Physicochemical properties were also obtained to help model field effects.

Using statistical analysis software, models were derived using the different physicochemical properties of the pesticides being examined. The different variables were subjected to numerous tests to determine which were statistically significant and could be used for overall modeling. The different species were broken down into the following divisions for modeling: all insect data combined, pollen feeders, aphid feeders, and generalist feeders.

The results from this modeling show that toxicity based hazard ratios (HR) were important in predicting mortality for most insects but not spiders. However, HR alone did not produce the best predictions. Other factors such as physicochemical properties, crop, or the insect family were needed to explain the variation in the data. From all the different models derived, only the pollen feeding insect model offered reasonable predictions. We believe that the reasonable fit of this model reflected the similarity in the ecological susceptibility and exposure routes to pesticides between honeybees and our pollen-feeding guild. This suggests that ecological niche and exposure routes may be more important than toxicological susceptibility in explaining or predicting impacts on non-target insects. Unfortunately, having knowledge of bee toxicity alone does not allow us to predict impacts on terrestrial invertebrates at large.

2 METHODS AND DATABASES

2.1 Introduction to SELCTV

The majority of the data for this analysis was obtained from the SELCTV database. The SELCTV database contains a worldwide compilation of published literature on pesticide side-effects on natural enemies from 1911-1999 (1078 references in total). The database was initially part of a masters thesis published in 1988, with the majority of articles dating from 1970-1985, however it is fairly out of date as only 12% of the articles were published after 1985. The SELCTV database was constructed to predict susceptibility to pesticides between different arthropod species, including pests, predators and parasitoids. Some terrestrial arthropods are important in the biological control of agronomic pests. For each beneficial arthropod species the SELCTV database includes (when available in the original article) the sex, developmental stage (larvae, nymph, adult), and the type of insect (predator, parasitoid). The database includes information from both laboratory and field pesticide applications, however all laboratory pesticide applications were excluded from this analysis. The database includes observations of invertebrate mortality in 32 different crops; the main ones of interest being apple orchards (46%), alfalfa (13%), cotton (10%), and rice (5%).

The authors' aim in compiling the database was to include the majority of existing articles on beneficial arthropods from several countries around the world including Canada, US, and Europe, to name a few. When available, the formulation name, type, and active ingredient concentration was included in the database

2.1.1 Database description

The SELCTV database contained studies which compared beneficial arthropod presence in a sprayed plot to an unsprayed control plot. Percent mortality or percent loss is then used as a measure of arthropod reduction. Some sources in the database used statistics to determine if the organism loss was significant compared to the control. The authors of the database reported percent loss regardless of statistical significance. It is important to remember that percent loss measurements are based on count data and there is error associated with this measurement due to the difficulty locating organisms in both the control and sprayed plot.

There were occasional problems with multiple entries within the database. Some articles listed mortality on several days such as 1, 3, and 7 days post spray, all of which qualified to be included in the database. For our purposes only the results from the day where mortality was highest were used. A few sources in the SELCTV database gave different mortality values for the same pesticide, application rate, exposure duration, and species observed. In these cases, some of the related articles were obtained which showed the data to be from chemical applications in different years, therefore these multiple entries were retained.

2.1.2 Hymenoptera in SELCTV

The Hymenoptera in the SELCTV database consist primarily of parasitoid wasps and one ant entry (Formicidae). Parasitoid wasps are exposed to pesticides though their hosts and their toxicity responses are very different from that of honey bees. The mortality observations in SELCTV were made either by finding adult wasps in the field or by taking hosts from the field into a lab and observing adult wasp emergence. These two methods were not differentiated within the database and for that reason, hymenoptera were omitted from the analysis.

2.2 Honeybee toxicity data

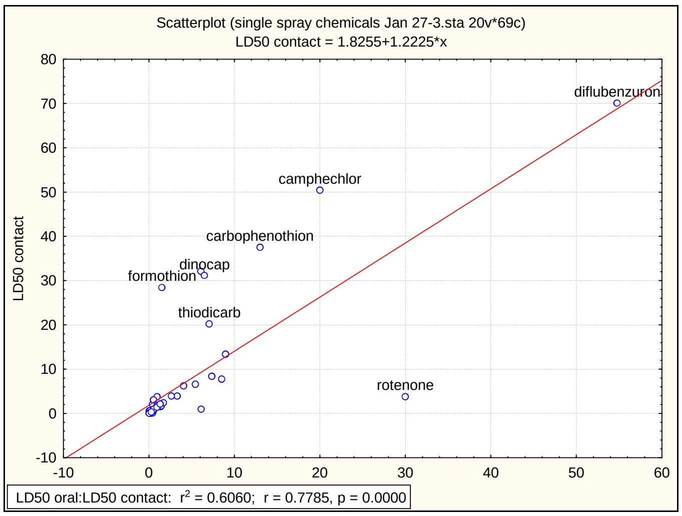
In the course of pesticide registration very few invertebrate toxicity tests are ever required by or submitted to regulatory authorities. The data gaps are more severe in North America than in Europe as more routine testing of beneficial arthropods occurs in European countries. Due to this lack of data, the pesticide toxicity to the majority of invertebrate species in the database remains unknown. More research has been conducted on bees in recent years and therefore honeybee toxicity values are known for the majority of pesticides in the database. The honeybee LD_{50} values therefore were used as an approximation of the toxicity of the pesticides toward other invertebrate species. Honey bees are directly exposed to pesticides as they forage crops. Other beneficial organisms such as Coleoptera and Hemiptera are exposed in the same way.

The honey bee oral and contact LD_{50}s were obtained from the British Crop Protection Council (BCPC) Pesticide Manual, the USEPA one-liner database, Institut National de la Recherche Agronomique (INRA) AGRITOX, publications from the University of California, and other published sources. Where there were multiple LD_{50} values available, the geometric mean was calculated to be used in the analysis (Appendix C).

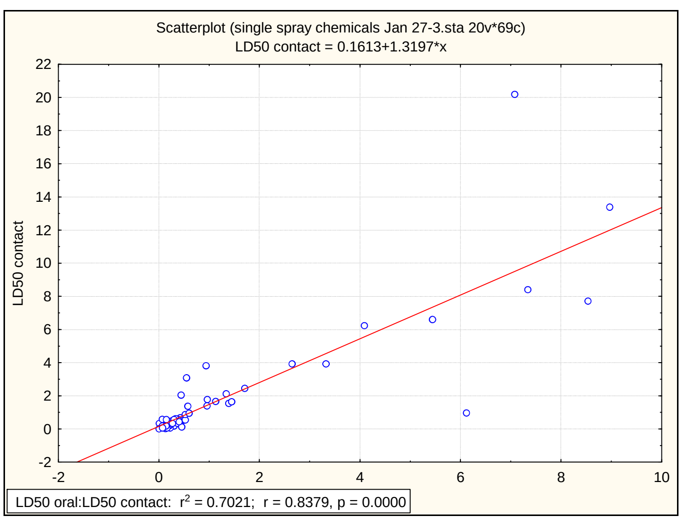
For honeybees, both contact and oral LD_{50}s are typically measured. Oral and contact LD_{50} are correlated (Fig. 1 -- r^2=0.94 $p<0.01$) and because the contact LD_{50} data is more complete, we only used the contact LD_{50} to compute hazard ratios. Furthermore, an analysis of bee poisoning incidents (Harding et al., 2006) shows that contact toxicity is a slightly better predictor of poisoning events. Thiometon and ethirimol were the only chemicals with an oral LD_{50} and no contact LD_{50}; therefore we used the oral LD_{50} for these chemicals in our analysis to avoid data gaps. The application rate along with the LD_{50} value, were used to construct a hazard ratio (HR) which is defined as follows:

$$HR \text{ (million } LD_{50}\text{s/ha)} = \text{application rate (g/ha)/ } LD_{50} \text{ (µg/bee)}$$

Figure 1: a) Correlation between oral and contact LD_{50} values for honeybees for those pesticides represented in the SELCTV database. The identity of the principal outliers is given on the graph. 1b: Correlation between oral and contact LD_{50} values for honeybees for the most toxic products in the SELCTV database.



(a)



(b)

2.3 Physicochemical data

Risk assessments are currently carried out without consideration of important factors such as application rate or physicochemical parameters of the pesticides. These physicochemical parameters can provide valuable information about the chemical in question, such as the environmental fate and persistence, ability to cross barriers such as insect cuticles and so on. For this reason we attempted to include these values in our modeling in order to get a more accurate picture of the activity of a particular pesticide. We obtained K_{oc}, soil and foliar DT_{50} estimates from the USEPA's Groundwater Loading Effects of Agricultural Management Systems (GLEAMS) database, the USDA toxicology database, the Oregon State University pesticide database, and the Pesticide Manual (Tomlin, 2003) in that order. The vapour pressure (vp), log K_{ow}, molecular weight (mw), and water solubility were taken preferentially from the Pesticide Manual (Tomlin, 2003), and then from the other databases in the order listed above.

Of the 89 chemicals included in the analysis, the log K_{ow} was missing for 22 chemicals, the vapour pressure for 8 chemicals, the water solubility for 4 chemicals and the K_{oc} for 15 chemicals. The EPI Suite program from the Syracuse Research Corporation was used to estimate the missing parameters. This program uses the SMILECAS Database, three modeling programs, and the Estimation Programs Interface for Windows (EPIWIN). The programs used were:

KOWWIN - estimates octanol-water partition coefficient

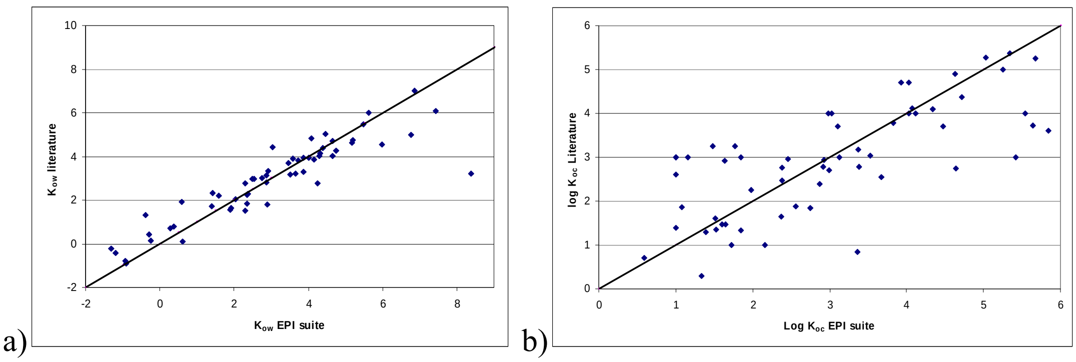
MPBPWIN - estimates melting point, boiling point, and vapour pressure

PCKOCWIN - estimates soil sorption coefficient (K_{oc})

WSKOWWIN - estimates water solubility

In order to test the validity of these programs we compared the log K_{ow} and K_{oc} from EPI suite with log K_{ow} and K_{oc} values from literature included in the program itself. The relationship between log K_{ow} values from both sources as seen in Figure 2a is close to 1:1 with dienochlor as the one obvious outlier (Figure 2a). For the K_{oc} data there is more scatter about the 1:1 line (Figure 2b). The scatter is due to several factors related to the measurement of K_{oc} including the measurement and type of organic matter in the soil. Since the literature and EPI Suite data are correlated, we included both in subsequent analyses.

Figure 2: Relationship between a) K_{ow} and b)K_{oc} from two sources, literature and EPI KOWWIN or PCKOCWIN respectively. Trend line reflects a 1:1 ratio with the points being scattered around it.



2.3.1 Foliar Half-life

The foliar half lives or DT_{50} values are missing for many chemicals in the SELCTV database and it is the least standardized variable in the database. Variance occurs at the field level with foliar DT_{50} values ranging from plant to plant and is also due to weather conditions such as rainfall, humidity, and sunlight intensity. Furthermore, there is a lack of method standardization in the literature. Some sources examine only pesticides on the leaf surface, while others blend fruit and leaves together for examination. We therefore attempted to create a model that would estimate foliar DT_{50} from other more accessible parameters. Details of methods used to develop this model are in Appendix E.

2.4 Standardization of Application Rates

With the variety of sources, different standards for reporting the pesticide application rate were found in the articles. The application rate, in original units, was entered into the database for each pesticide, and then all application rates were standardized by converting them to metric units. Non-convertible units were excluded (i.e. L/tree or % formulation). Articles were retained where it was possible to calculate tank mix concentrations (g or mL per L of spray solution), or rates applied per defined area (g or ml per hectare), and these two measures of application rate were analyzed in separate groups. Also, the specific density of water was used to convert the application rate into either g/ha or g/L.

However, some of the articles listed the application rate as the active ingredient (ai) (for example g ai/ha), while others listed the application rate as a formulation (g/ha). Ideally, all chemicals in the database would have application rates expressed in g ai/ha. If the percent of ai in the formulation was available, it was a simple matter of converting the application rate to g ai/ha. Unfortunately, many entries did not include the percent of ai in the formulation. This occurred because of inconsistencies in the original articles as well as with how data were entered into the database. Therefore in order to standardize the application rate for all entries, we found the range, mean, and standard deviation of the application rates per chemical. With this information, it was possible to create a list of records where the application rate was over two times the standard deviation of the mean application rate. Once these outliers were identified it was then possible to retrieve articles and correct the application rates where possible. After the initial standardization, a second set of outliers was identified from the corrected dataset and relevant articles were again tracked down. Using this method it was possible to uncover approximately 50% of the newly identified outliers from the corrected dataset and therefore ascertain that application rates were exactly as cited and no further formulation information could be obtained from the articles.

We then proceeded with the analysis starting with 497 entries in the final single spray dataset, of which roughly half (49%) had units in g ai/ha. For the other entries, we cannot always be certain that the stated application rates were reported in g ai/ha. Frequency distributions of application rates were constructed for different active ingredients and we attempted to review the source material to clarify different outliers. This does however remain a possible source of error. Active ingredient concentrations are typically in the range of 50% for commercial pesticides; therefore this may have introduced approximately a two-fold error in some of the tabulated rates. When introduced into an effects model, this would tend to underestimate the toxicity of some pesticides.

2.5 Type of Sprays

Many farmers spray their crops several times during a season and many of the articles in the database included multiple spray data. The database did not differentiate between single and multiple sprays per crop. Since the number of pesticide applications will change the observed mortality of beneficial insects, we attempted to separate entries based on the post-spray evaluation time. The post-spray evaluation time ranged from 1-120 days and was occasionally listed as 'variable days.' After reviewing articles from the database associated with 'variable days', it was determined that they fit one of the following patterns: a)one observation was made at the end of the season (for example spider presence in the field after the spray applications were finished); or b) there were observations made repeatedly after the first spray, and the mortality results were based on a pool of all observations. Using this division we created two separate smaller databases: one including only single applications and one containing references that likely had multiple applications. It is rare for a farmer to spray a field more than once a week; therefore we used a 7 day cut off point for the single spray database. The multiple spray database included the 'variable day' entries and any entry that listed the exposure duration as greater than 30 days (including 30, 60, 77 and 120 days post initial exposure).

2.6 Final manipulations of the databases

The final database (summarised in Table 1) contains results of single spray events reported in g/ha from 52 different sources including 49 chemicals (listed in Appendix A), and comprising 497 entries. There are 46 different species divided into 6 orders (Appendix B). Furthermore, there are 20 different crops included in the database, the majority of which are alfalfa (33%), followed by cotton (17%) and rice (13%). There were fewer than 10 entries in the remaining family groups, as listed in Table 2, which was not sufficient for analysis.

Table 1: Distribution of each invertebrate order in each crop in the single spray g/ha database.

Order	Family	alfalfa	cereal grains	cotton	rice	vegetable
Araneae	Lycosidae				3	
Araneae	Micryphantidae				3	
Araneae	Unknown			7	28	
Coleoptera	Coccinellidae	47	16	40		29
Coleoptera	Malachiidae	6		3		
Coleoptera	Staphylinidae			3		
Diptera	Syrphidae	9	1	1		
Hemiptera	Anthocoridae	51		18		
Hemiptera	Belostomatidae				1	
Hemiptera	Lygaeidae	26		15		3
Hemiptera	Miridae				28	
Hemiptera	Nabidae	44		22		
Hemiptera	Reduviidae	5		1		
Hemiptera	Veliidae				28	
Neuroptera	Chrysopidae	26	3	22		
Neuroptera	Hemerobiidae					4
Thysanoptera	Aeolothripidae	4				

Table 2: Families Removed from SELCTV analysis.

Aeolothripidae
Belostomatidae
Hemerobiidae
Lycosidae
Malachiidae
Micryphantidae
Reduviidae
Staphylinidae

The SELCTV database lists % mortality as the final endpoint. According to the sources listed in the database the % mortality represents the % loss of organisms compared to the unsprayed plots. The percent mortality data was not normally distributed and in order to achieve a normal distribution the data were transformed using arc sine square root transformations.

The factors log K_{ow} and molecular weight (mw) were normally distributed, while other factors, including K_{oc}, vapour pressure (vp), soil DT_{50}, foliar DT_{50} and water solubility, were log transformed to achieve normal distributions. Principal component analysis (PCA) is commonly used to reduce the dimensionality of a data set in which there are large numbers of interrelated variables, while attempting to retain as much variation as possible. This reduction is completed by transforming to a new set of variables, also known as the principal components, which are uncorrelated and ordered so that the first few retain most of the variation present in all the original variables. From here, it is reduced to a solution of an eigenvalue-eigenvector problem for a positive-semidefinite symmetric matrix (Jolliffe, 1986). The factors molecular weight (mw), log K_{oc}, log vp, log water solubility and log K_{ow} are inter-correlated as shown in Table 4. These variables were subjected to the PCA which revealed that vapour pressure did not contribute to the first factor, which is evident in Tables 5 and 6. Therefore we only used log K_{ow}, log K_{oc}, and log solubility in the PCA analysis. With these three variables, the eigenvector explained 82%

of the total variance in the data, expressing each variable equally as seen in Tables 7 and 8.

Table 3: Correlations between factors in this analysis. mw= molecular weight, vp= vapour pressure.

	mw	**log vp**	**log soil DT_{50}**	**log solubility**	**log K_{oc}**	**log foliar DT_{50}**	**PCA**
log K_{ow}	.4713	-.1751	.2829	-.6478	.6633	.1619	.8649
	p=.000	p=.150	p=.019	p=.000	p=.000	p=.184	p=0.00
mw		-.3855	.1270	-.4929	.5686	.0464	.5569
		p=.001	p=.298	p=.000	p=.000	p=.705	p=.000
log vp			-.2340	.4838	-.3223	-.0533	-.3721
			p=.053	p=.000	p=.007	p=.664	p=.002
log soil				-.5007	.4138	.3819	.4460
				p=.000	p=.000	p=.001	p=.000
log solub					-.8533	-.2714	-.9308
					p=0.00	p=.024	p=0.00
log K_{oc}						.2550	.9062
						p=.034	p=0.00
log foliar							.2528
							p=.036

Table 4: Contribution of each variable to each of the eigenvalues

	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**
log K_{ow}	0.238849	0.188017	0.572193	0.000942
log vp	0.110002	0.779896	0.081254	0.028848
log solubility	0.336066	0.000968	0.106565	0.556401
log K_{oc}	0.315083	0.031120	0.239988	0.413809

Table 5: Contribution of each variable to each of the Eigenvalues, when log K_{ow}, log K_{oc}, log vp, and log solubility are included in PCA analysis.

	Eigenvalue	% Total variance	Cumulative Eigenvalue	Cumulative %
1	2.649145	66.22863	2.649145	66.2286
2	0.866681	21.66703	3.515826	87.8957
3	0.356845	8.92113	3.872672	96.8168
4	0.127328	3.18321	4.000000	100.0000

Table 6: Contribution of each variable to each of the eigenvalues.

	Factor 1	Factor 2	Factor 3
log K_{ow}	0.2910	0.7080	0.0010
log solubility	0.3523	0.1651	0.4825
log K_{oc}	0.3567	0.1268	0.5165

Table 7: Eigenvalues of correlation matrix and related statistics when log K_{ow}, log K_{oc}, and log solubility are included in PCA analysis.

	Eigenvalue	% Total variance	Cumulative Eigenvalue	Cumulative %
1	2.447213	81.57375	2.447213	81.57
2	0.406390	13.54633	2.853602	95.12
3	0.146398	4.87992	3.000000	100.00

Table 8: Eigenvectors for each variable and each eigenvalue.

	Factor 1	Factor 2	Factor 3
log K_{ow}	0.5394	0.8414	0.0313
log solubility	-0.5936	0.4064	-0.6946
log K_{oc}	0.5972	-0.3562	-0.7187

Despite efforts to minimize data gaps, some chemicals were missing the contact LD_{50}, oral LD_{50}, and/or soil DT_{50}. Most of these chemicals are older experimental pesticides which were not widely used. As a result of the missing data, we removed 19 chemicals from the single spray analysis because they were missing LD_{50} values and an additional 3 chemicals because they were missing soil DT_{50} values. This resulted in 40 entries being removed from the data set. The chemicals which were not included in analysis appear in Table 9.

Table 9: Chemicals removed from the single spray database because of missing information.

Chemical Name	CAS	LD_{50} contact	soil DT_{50}
bromophos	2104963		
cartap	15263533		3
chlordecone	143500		
chlordimeform	6164983		60
dienochlor	2227170		300
dilan	8027007		
dioxathion	78342		
fenobucarb	3766812		10
fenpropathrin	39515418		3
fentrifanil	62441547		
genite	97165		
heptenophos	23560590		1.4
isoprocarb	2631405		
metolcarb	1129415		
schradan	152169		
thiofanox	39196184		
ethiofencarb	29973135	31.6	
TEPP	107493	0.197	
tetradifon	116290	117.2604	

2.7 Model building

We selected models by the best subset method, an iterative method based on maximum likelihood estimation, and Akaike's Information Criterion (AIC) (Burham and Anderson, 2002). The AIC penalizes for the number of independent variables in the model and therefore suggests the most parsimonious solutions. AIC was employed to determine whether physicochemical properties influenced invertebrate mortality. The AIC_c correction factor was used when the sample size was small (less than 200). Akaike weight ratios along with resulting r^2 values (when category factors were not in the model) were used to compare the best models with all likely models. When category factors were included in the model, in order for the model to be accepted, the continuous predictor variables needed to significantly predict mortality without the categorical factor.

2.7.1 Global analysis

The first analysis included all invertebrates and all crops sprayed once with application rates listed in g/ha. Originally, we attempted to model using all the variables present. Subsequent modeling attempts, both by crop and then family can be seen in Appendix D. The best model (model 1) suggests that hazard ratio (HR), vapour pressure, soil DT_{50}, and crop type (entered as a categorical variable) could be combined to predict mortality as demonstrated in Tables 10 and 11. While model 1 contains HR, soil DT_{50}, and family as a categorical variable, overall there is a poor fit in the model which can be seen in Figure 5. Within this model the factor with the largest effect on mortality is crop type, not HR as might be expected.

Table 10: Akaike's Information Criteria (AIC) for all data in single spray g/ha dataset. All predatory species, crops and pesticides with application rate in g/ha included in analysis (N =497).

variable 1	variable 2	variable 3	variable 4	variable 5	variable 6	df	AIC	Δ AIC	Akaike weight	ratio	Chi2	p
log HR	log soil	Family				18	4665.18	0.00	0.17	1.00	112.96	0.000000
log HR	log vp	log soil	log foliar	Family		20	4665.22	0.04	0.17	1.02	116.92	0.000000
log HR	log vp	log soil	Family			19	4665.47	0.29	0.15	1.16	114.67	0.000000

Table 10: Akaike's Information Criteria (AIC) for all data in single spray g/ha dataset. All predatory species, crops and pesticides with application rate in g/ha included in analysis (N =497).

variable 1	variable 2	variable 3	variable 4	variable 5	variable 6	df	AIC	Δ AIC	Akaike weight	ratio	Chi2	p
log HR	log soil	log foliar	Family			19	4665.61	0.43	0.14	1.24	114.53	0.000000
log HR	log vp	log soil	log foliar	PCA	Family	21	4666.82	1.64	0.08	2.27	117.32	0.000000
log HR	log vp	log soil	PCA	Family		20	4667.09	1.91	0.07	2.60	115.05	0.000000
log HR	log soil	PCA	Family			19	4667.09	1.91	0.07	2.61	113.04	0.000000
log HR	log soil	log foliar	PCA	Family		20	4667.54	2.36	0.05	3.26	114.60	0.000000
log HR	log soil	Crop				6	4673.81	8.63	0.00	74.75	80.33	0.000000

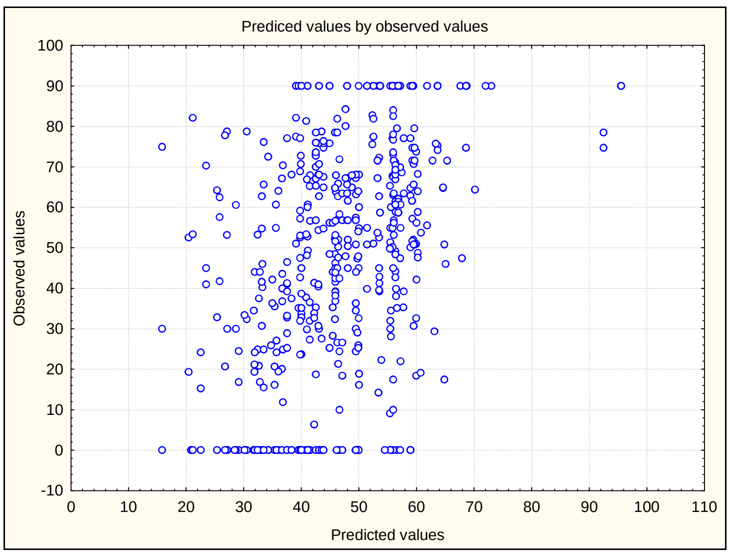
Table 11: Model 1 intercept values for each family for input into the equation:

Asin(sqrt mort)= intercept+8.03*(log HR) –3.06*(log soil DT_{50})

	n	intercept
Lycosidae	3	11.64
Micryphantidae	3	30.16
Unknown	35	14.61
Coccinellidae	132	40.53
Malachiidae	9	45.20
Staphylinidae	3	57.52
Syrphidae	11	52.44
Anthocoridae	69	27.36
Belostomatidae	1	-23.21
Lygaeidae	44	32.12
Miridae	28	26.78
Nabidae	66	29.97
Reduviidae	6	51.06
Veliidae	28	-3.35
Chrysopidae	51	37.87
Hemerobiidae	4	8.30
Aeolothripidae	4	28.48

With a large database such as this one (497 entries and 17 different families), it is possible for a factor to be included as a variable in the best model, yet not have much influence on predicting mortality because of the sheer magnitude of overall variance. Upon analyzing the different models, it became apparent that the strongest effect on mortality was family. Belostomatidae, Veliidae and Hemerobiidae had the lowest mortality, while Reduviidae, Syrphidae, and Staphylinidae had higher mortality. When the other factors in model 1 were held constant using the mean value, mortality ranged from 12% (Veliidae) to 72% (Staphylinidae) which is a 6 fold difference in mortality.

Figure 5: Observed versus predicted values for model 1 (crop type as a categorical factor) with all predatory invertebrates included.



2.7.2 Analysis of separated data

It was shown above that the invertebrate family was an important variable in predicting mortality. Given the difficulty in predicting mortality from the entire dataset, separating the data into logical groups could be advantageous. Related insects are likely to have similar behaviour and morphology and are also likely to have more correlative toxic responses to pesticides than phylogenetically distant groups. Crop type is an equally plausible factor which has influence on mortality, since the effectiveness of chemical applications vary from crop to crop depending on leaf type, plant culture practices, and field micro-habitat.

The family with the most entries in the SELCTV database is Coccinellidae (N=132), as illustrated in Table 11 above. We used data from this family alone to test the effect of HR, and soil DT_{50} in accordance with the best model obtained with the full dataset. The variable with the most influence on mortality was HR. Log HR contact varied from -0.63 to 4.71 and mortality ranged from 30-74%, which is a two fold difference. This suggests that the HR derived from honeybee mortality is somewhat predictive of insect mortality for this taxonomically-unrelated group. Soil half life only caused a 1.2 fold difference in mortality over the range observed here (0.5-3980 days).

Segregation of the dataset into logical groups and subsequent analysis of the data by family led to a separation of the data into four groups of invertebrates. The first category includes spiders, for which the best model (spider model) incorporates crop type and foliar DT_{50}. The second group includes general predators, including all the Hemiptera families (except for Anthocoridae), and within each family there were no models that significantly explained the variability in mortality. Taking the predator family groups as a whole, the best model (predator model) to predict mortality used primarily crop type, while the continuous variables did not predict mortality. The remaining 2 groups of invertebrates included the aphid predators, the Coccinellidae and Neuroptera, and the pollen feeders, the Anthocoridae and Chrysopidae adults. These were the only families for which the best models significantly predicted mortality. These models are discussed in detail in Appendix D.

Separation of the data into smaller crop-specific datasets at times appeared to give promising models (e.g. impacts documented on vegetable crops with Log HR, soil DT_{50} and insect order; r^2=0.45 – Appendix D). However, the need to differentiate among crops reduces the general applicability of such models and their usefulness to arrive at a general ranking of products under NAESI.

2.7.2.1 Pollen Feeders

Adult Anthocoridae and Chrysopidae feed on pollen when other sources of food are scarce. This tends to occur especially after an application when much of the food source has been destroyed by the pesticides. When Anthocoridae and Chrysopidae are combined, the best AIC model (pollinator model), listed in Tables 12 and 13 (log HR and log vp), significantly predicts mortality in these species as seen in Figure 6 (r^2= 0.46 p=0.00041). It is likely that this is the only group for which mortality can be predicted since they have the same food source as honey bees. In the best model for pollen feeders, log HR has a 7 fold effect on predicted mortality and log vapour pressure has a 2 fold effect on mortality. This is the only group for which HR significantly predicts mortality without any other factors.

Table 12: Akaike's Information Criteria (AIC) for all pollen feeders (adult *Orius* spp and adult Chrysopidae (N=28) with application rates in g/ha. Only the top 6 models presented.

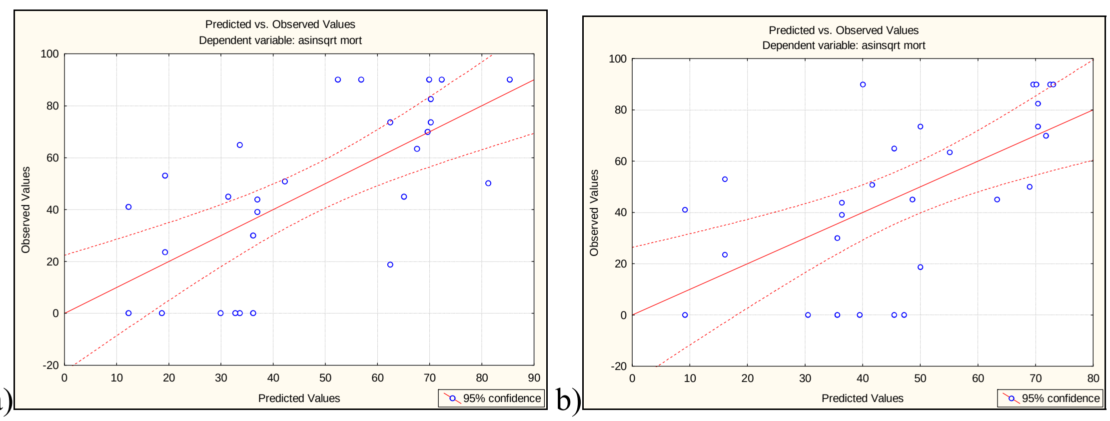
K	variable 1	variable 2	variable 3	df	AIC_c	ΔAIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
4	log HR	log vp		2	263.24	0.00	0.22	1.00	17.48	0.00016
5	log HR	log vp	PCA	3	264.06	0.82	0.15	1.51	19.64	0.00020
3	log HR			1	264.72	1.48	0.10	2.10	13.26	0.00027
5	log HR	log vp	log foliar	3	265.33	2.09	0.08	2.84	18.38	0.00037

5	log HR	log vp	log soil	3	265.45	2.22	0.07	3.03	18.25	0.00039
5	log HR	log vp	Crop	3	266.12	2.88	0.05	4.23	17.59	0.00054

Table 13: Coefficients for the best four models from all pollen feeding insects (adult *Orius* spp and adult Chrysopidae).

Intercept	log HR	log vp	PCA	r^2	p
-54.05	24.09	-8.93		0.46	0.00041
-66.39	26.03	-11.49	-1.70	0.50	0.00065
-6.34	24.03			0.38	0.00051

Figure 6: Predictive capability of the best two models for pollen feeding insects, a) log HR and vp b) log HR alone.

a) 

2.7.2.2 Aphid Predators

There are two groups of insect aphid predators in the database. The first group, Coccinellidae, feed on aphids as both larvae and adults. The second group, most Neuroptera species, also feed on aphids as larvae and adults; however a family within the Neutroptera species, the Chrysopidae, can feed on pollen as adults, especially when food is scarce. Since both Coccinellidae and Neuroptera larvae feed on aphids, they were combined into one group. We also added adult

Micromus tasmaniae (Neuroptera, Hemerobiidae) to the group, which is a species that exclusively feeds on aphids as adults. The AIC analysis suggests that the best model (insect model 1) for predicting mortality for this group contains HR, soil half life, foliar half life, PCA, and crop type (Table 14). This model significantly predicts mortality in aphid predators as illustrated in Table 15 and Figure 7. The best model without crop type (insect model 2) (Δ AIC=6.10) significantly predicts mortality (Figure 8) but not with any confidence (r^2=0.20 p=0.0000005). Comparing these two models showed a 0.1-36% difference in mortality prediction, indicating that crop type greatly influences model predictive capability. In the insect model 1, there is a 2 fold effect of HR and soil DT_{50}, a 1.5 fold effect of foliar DT_{50}, and a 1.3 fold effect of PCA. This indicated that PCA was the least significant predictor of mortality. Therefore we then excluded PCA in favour of the third best model (insect model 3), Δ AIC=0.89). Within insect model 3, each variable had more than 1.5 fold effect on mortality. Without including crop as a variable, HR, soil half life and foliar half life significantly predicted mortality – but again, with very low confidence (r^2=0.16 p<0.000001).

Table 14: Akaike's Information Criteria (AIC) for all aphid feeders (all Coccinellidae and larvae Chrysopidae) (N=176) with application rates in g/ha. Only models with weight ratio less than 6 presented.

df	K	variable 1	variable 2	variable 3	variable 4	variable 5	variable 6	AIC_c	Δ AIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
7	9	log HR	log soil	log foliar	PCA	Crop		1656.21	0.00	0.23	1.00	48.84	0.00000002
8	10	log HR	log vp	log soil	log foliar	PCA	Crop	1656.63	0.42	0.19	1.23	50.66	0.00000003
6	8	log HR	log soil	log foliar	Crop			1656.95	0.74	0.16	1.45	45.87	0.00000003
7	9	log HR	log vp	log soil	log foliar	Crop		1657.10	0.89	0.15	1.56	47.94	0.00000004
6	8	log	log	PCA	Crop			1658.	1.9	0.09	2.6	44.6	0.00000

		HR	soil					13	2		2	9	005
5	7	log HR	log soil	Crop				1658. 52	2.3 1	0.07	3.1 7	42.1 1	0.00000 006
7	9	log HR	log vp	log soil	PCA	Crop		1660. 17	3.9 6	0.03	7.2 4	44.8 8	0.00000 014
6	8	log HR	log vp	log soil	Crop			1660. 42	4.2 1	0.03	8.2 1	42.4 0	0.00000 015
5	7	log HR	log vp	log soil	log foliar	PCA		1662. 31	6.1 0	0.01	21. 08	38.3 2	0.00000 033

Table 15: Parameter estimates of insect model 1 for mortality in aphid feeders. Numbers in brackets are the SE of that coefficient.

Asin(sqrt mortality)= intercept + 9.42 (2.25)*log HR - 11.15 (2.59)*log soil+15.37 (7.50)*log foliar+ 1.29 (0.75)*PCA		
	Intercept	**N**
cotton	34.49	57
alfalfa	39.86	67
cereal grains	17.02	19
vegetable	43.47	33

Table 16: Parameter estimates of the insect model 3 for mortality in aphid feeders. Numbers in brackets are the SE of that coefficient.

Asin(sqrt mortality)= intercept +9.72 (2.27)*log HR – 8.32 (2.03)*log soil + 14.74 (7.56)*log foliar		
	Intercept	**N**
cotton	34.03	57
alfalfa	39.30	67
cereal grains	13.51	19
vegetable	40.79	33

Table 17. Parameter estimates for the insect model 2 (Δ AIC=6.1; r^2=0.20 p<0.000001).

	Intercept	**log HR**	**log vp**	**log soil**	**log foliar**	**PCA**
B	21.86	10.33	-2.92	-13.14	16.98	1.39
SE of B	11.07	2.35	1.67	3.00	8.48	0.75

Figure 7: Predictive capability of the models (a) insect model 3 (b) insect model 1

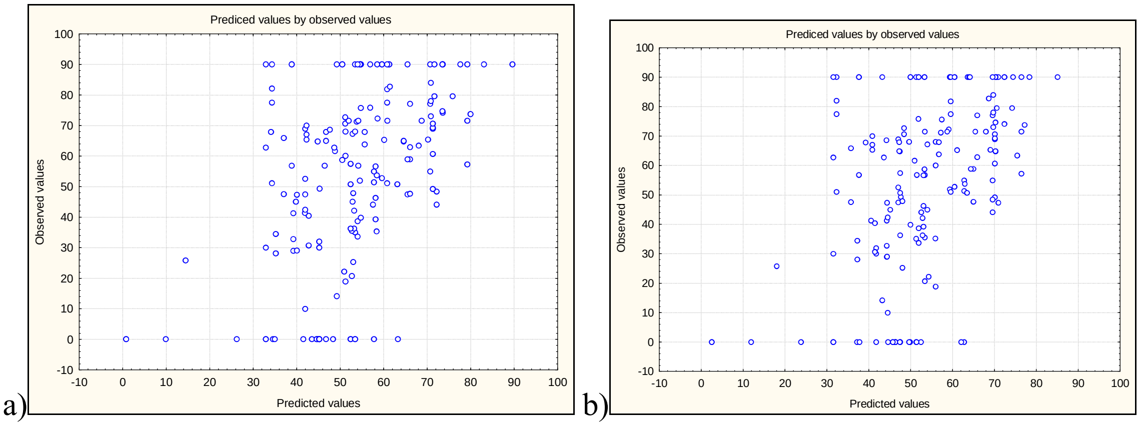
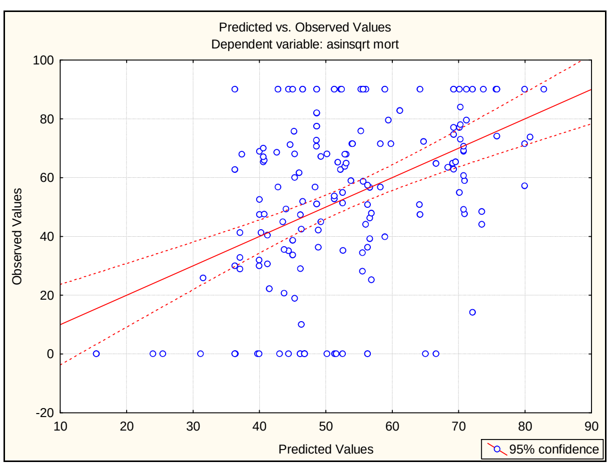


Figure 8: For all insect aphid predators in all crops, predictive capability of insect model 2 (Δ AIC=6.1; r^2=0.20 p<0.000001).



The insect model 1 encompassed all crops, including alfalfa, cotton, vegetable and cereal crops as categorical factors. The smallest crop, cereals, had the lowest observed invertebrate mortality (Table 18). The intercept coefficients for alfalfa, cotton, and vegetables are very similar, but quite different from that of cereal crops. Therefore we decided to remove cereal crops and reanalyse as seen in Table 19. The best model (insect model 4) contained HR, soil and foliar DT_{50} values and was significant but weak (r^2=0.20 p=0.0000003) (Figure 9 and Table 20). There was a two fold effect of HR and soil DT_{50}, and a 1.5 fold effect of foliar DT_{50}. With cereals included in analysis, the same three factors provided a poorer model fit (r^2=0.16 p<0.00001).

Table 18: mean observed mortality in each crop.

Crop	Observed mean % mort	N
cotton	59.8%	57
alfalfa	71.7%	67
cereal grains	30.5%	19
vegetable	74.5%	33

Table 19: Akaike's Information Criteria (AIC) for all aphid feeders (all Coccinellidae and larvae Chrysopidae) when cereals have been removed (N=157) with application rates in g/ha. Only models with weight ratio less than 2 presented.

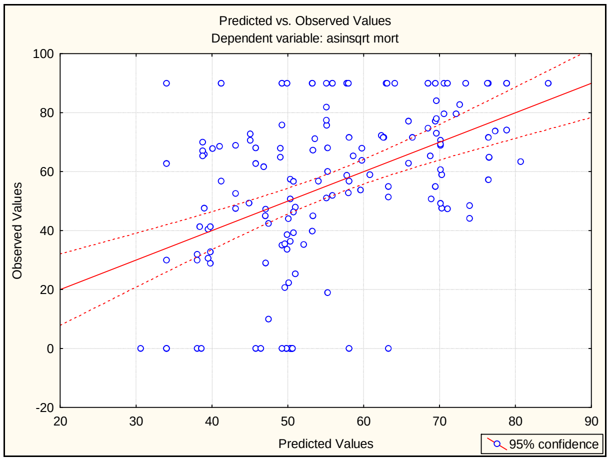
df	K	variable 1	variable 2	variable 3	variable 4	variable 5	AIC_c	Δ AIC_c	Akaike weight	ratio	r^2	p
3	5	log HR	log soil	log foliar			1464.73	0.00	0.18	1.00	0.20	0.0000003
4	6	log HR	log vp	log soil	log foliar		1465.14	0.42	0.15	1.23	0.20	0.0000005
2	4	log HR	log soil				1465.95	1.22	0.10	1.84	0.18	0.0000003
6	8	log HR	log vp	log soil	log foliar	Crop	1466.07	1.34	0.09	1.96		
4	6	log HR	log soil	log foliar	PCA		1466.16	1.43	0.09	2.04		

Table 20: Coefficients in the top three models predicting mortality in aphid feeding insects.

r^2	p	Intercept	log HR	log vp	log soil	log foliar
0.20	0.0000003	34.20	11.48		-8.51	14.10
0.20	0.0000005	21.03	11.66	-2.76	-10.52	17.61

0.18	0.0000003	39.64	11.37		-6.81	

Figure 9: Predictive capability of the insect model 4 predicting mortality in aphid predators.



The final feeding group present in the SELCTV data set consists of generalist predators. These insects feed on anything smaller and slower than they are. The best model (general predator model 1) contained HR, PCA, and family as categorical factors (Table 21). The categorical factor family was more important than crop (Δ AIC= 5.5), and had a great influence on mortality which is visualized in Figure 10. Family is important because it divides insects into categories with similar behaviour, feeding habits and morphology. However, in order to predict mortality in these insects, a better measure of toxicity or a better indication of pesticide uptake is needed.

As with the insect models, any general predator families with less than 10 entries were removed, which left all terrestrial opportunist (or general) predators for analysis. The best model (general predator model 2) included HR and PCA (Table 22), however it did not explain variation in the data seen in Figure 11 (r^2=0.071 p=0.00097, Table 22).

Table 21: Akaike's Information Criteria (AIC) for all general predators (insects only) (N=210) with application rates in g/ha. Only models with weight ratio less than 6 presented.

variable 1	variable 2	variable 3	variable 4	variable 5	variable 6	df	AIC	Δ AIC	Akaike weight	ratio	L. Ratio Chi^2	p
log HR	PCA	Family				9	1965.21	0.00	0.21	1.00	42.49	0.000003
log HR	log foliar	PCA	Family			10	1966.33	1.12	0.12	1.75	43.37	0.000004
log HR	log vp	PCA	Family			10	1966.44	1.24	0.11	1.86	43.25	0.000004
log HR	log soil	PCA	Family			10	1966.96	1.75	0.09	2.40	42.74	0.000006
log HR	log soil	log foliar	PCA	Family		11	1967.35	2.14	0.07	2.92	44.35	0.000006
log HR	log vp	log foliar	PCA	Family		11	1967.59	2.38	0.06	3.28	44.11	0.000007
log HR	log vp	log soil	PCA	Family		11	1968.36	3.16	0.04	4.85	43.33	0.000009
log HR	log foliar	Family				9	1968.86	3.66	0.03	6.22	38.83	0.000012
log HR	log vp	log soil	log foliar	PCA	Family	12	1968.98	3.78	0.03	6.61	44.71	0.000012
log HR	Family					8	1969.00	3.79	0.03	6.65	36.70	0.000013
log HR	log soil	Family				9	1969.81	4.60	0.02	9.97	37.89	0.000018
log HR	log soil	log foliar	Invert Family			10	1970.65	5.44	0.01	15.20	39.05	0.000025
log HR	log vp	PCA	Crop			6	1970.73	5.52	0.01	15.80	30.97	0.000026
log HR	log vp	log foliar	Family			10	1970.84	5.63	0.01	16.70	38.86	0.000027
log HR	log vp	Family				9	1970.91	5.70	0.01	17.32	36.79	0.000029
log HR	log vp	log foliar	PCA	Crop		7	1971.26	6.06	0.01	20.65	32.44	0.000034
log HR	PCA	Crop				5	1971.63	6.42	0.01	24.77	28.07	0.000035
log HR	log vp	log soil	Family			10	1971.76	6.55	0.01	26.41	37.94	0.000039
log HR	log foliar	PCA	Crop			6	1972.17	6.97	0.01	32.57	29.52	0.000048
log HR	PCA					2	1972.50	7.30	0.01	38.39	21.20	0.000025

Figure 10: Predictive capability of a) the general predator model 1 and b) the general predator model 2 (Δ AIC=7.3; r^2=0.095 p=0.00003).

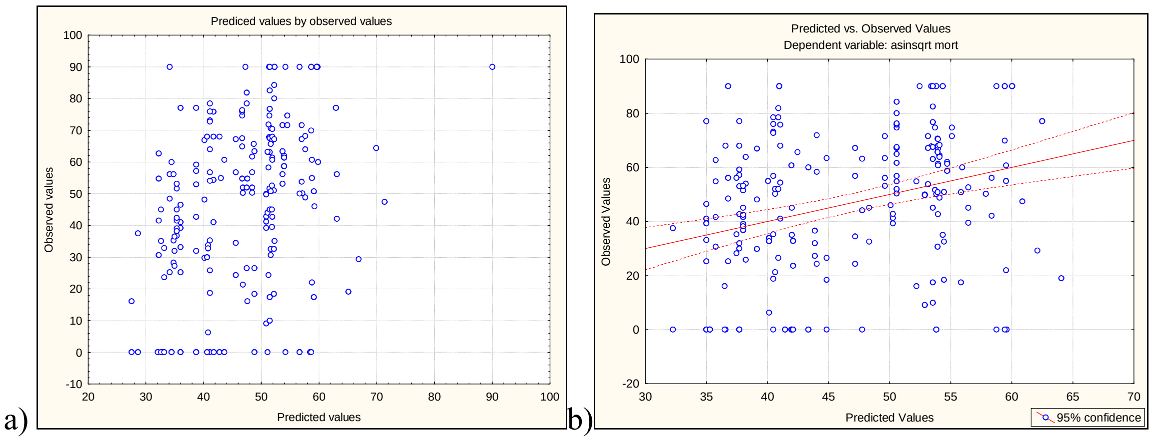


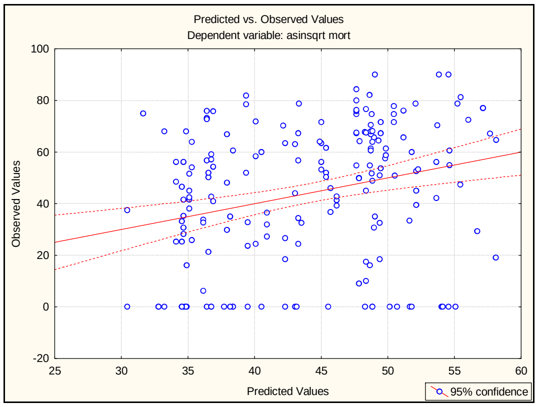
Table 22: Akaike's Information Criteria (AIC) for all general predacious insects (N=190) with application rates in g/ha. Only models with weight ratio less than 2 presented.

	K	variable 1	variable 2	variable 3	AIC_c	Δ AIC_c	Akaike weight	ratio
2	4	log HR	PCA		1778.12	0.00	0.14	1.00
3	5	log HR	log vp	PCA	1779.03	0.91	0.09	1.58
5	7	log HR	PCA	Crop	1779.46	1.34	0.07	1.96
1	3	log HR			1779.97	1.86	0.05	2.53

Table 23: Coefficients in the top three models predicting mortality in aphid feeding insects.

r^2	p	Intercept	log HR	log vp	PCA
0.071	0.00097	27.39	7.66		-1.01
0.077	0.0018	16.76	7.79	-2.12	-1.30

Figure 11: Predictive capability of the general predator model 2 with a sample size larger than 10.



3 CONCLUSIONS

In all of the insect models, the hazard ratio using honey bee LD_{50} was important in predicting mortality, however this did not hold true for the spiders. It was shown that HR alone does not predict mortality, except for groups which feed on pollen when other prey is scarce.

We did not find an acceptable model which predicts mortality in all invertebrate species without including categorical factors such as crop type and invertebrate family. These factors are very important in predicting invertebrate mortality in the field. In this database, often each family includes only one genus, so family type is a good predictor of feeding strategy. We found the best predictors of mortality occurred when the insects were divided into feeding guilds. For the aphid predators, mortality was best predicted by log HR, log soil DT_{50}, and log foliar DT_{50}. For those species which feed on pollen, log HR and log vp gave the best prediction of mortality.

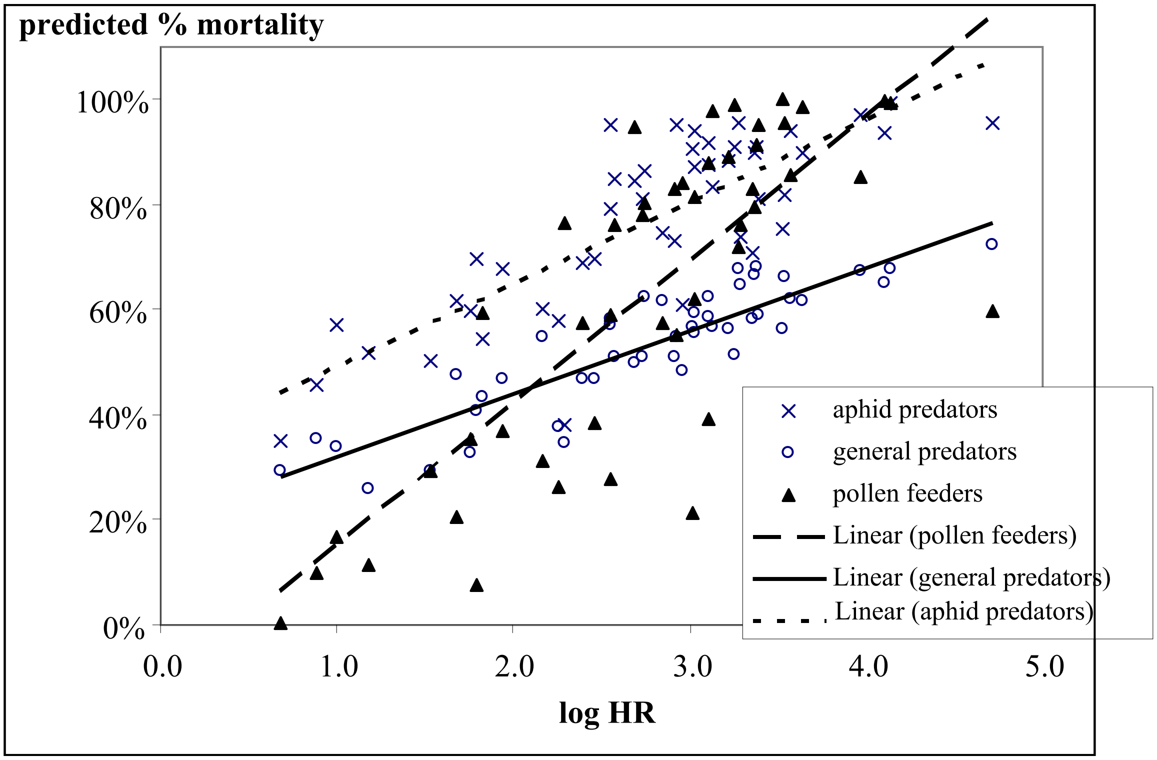
Overall, the best (and possibly the only useful) model was for pollen feeders – a group with a dietary niche not unlike that of honey bees. The models for aphid- feeding insects and general predators had poor predictive power overall; however the model fit slightly better for aphid feeding insects, than for general predators. This suggests that the toxicity values for honey bees can only be used to predict mortality in organisms with a similar feeding strategy. The analysis we carried out on coccinellidae showed that, whereas bee toxicity may be a rough predictor of toxicity to other species, the variation brought about by variation in species' ecology means that we are no further ahead in predicting the field impacts on groups that do not 'behave' as honeybees do. Toxicity values derived for other terrestrial invertebrates might improve model fit but our results suggest that this improvement might be a modest one. This would leave actual field experimentation as the only recourse for assessing field impacts on non-target invertebrates.

Nevertheless, it is instructive to compare average predicted mortalities between the models so as to assess the relative impact of pesticides on the different invertebrate guilds. In order to include as many chemicals as possible in this comparison, we used all chemicals in the database and a geometric mean of the application rates to calculate a common HR for each active ingredient. The best models are summarised in Table 24 and the results plotted in figure 12.

Table 24: Variables and intercepts for the best model for each feeding group.

	Intercept	**log HR**	**log vp**	**log soil**	**log foliar**	**PCA**
aphid predators	34.20	11.48		-8.51	14.10	
general predators	27.39	7.66				-1.01
pollen feeders	-54.05	24.09	-8.93			

Figure 12: Change in predicted mortality with change in log hazard ratio.



It appears that, for any bee-based HR, the impact on aphid predators tends to be higher than the impact on generalist predators. This is likely a result of their relative position on and near the crop. The generalist predators include more ground-dwelling species that may receive some protection from crop interception of the spray. Both groups show approximately the same incremental impact for increases in the HR.

The pollen-feeding group is initially less affected by the other two groups but this situation is reversed at higher hazard ratios. The dose-response relationship for this group appears to be steeper than for the two predatory guilds. We cannot suggest a reason for these results.

4 REFERENCES

Burnham, K.P. and D.R. Anderson. 2002. Model Selection and Multimodel Inference: A Practical Information –Theoretic Approach.

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: Using reports of bee mortality to calibrate laboratory-derived risk indices: An analysis and modeling of 21 years of bee incidents in the UK. National Agri-Environmental Standards Initiative Technical Series Report No. 2-44. 89 p.

Jolliffe, I.T. 1986. Principal Component Analysis. Springer Series in Statistics.

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.

Tomlin, C.D.S. 2003. The Pesticide Manual. BCPC, Alton, Hampshire, UK.

Villa, S., M. Vighi, A. Finizio, and G.B. Serini. 2000. Risk assessment for honeybees from pesticide-exposed pollen. Ecotoxicology. 9: 287-297.

5 APPENDICES

APPENDIX A: Chemicals included in the analysis, and the number of entries in the single spray g/ha data set.

Active Ingredient Name	CAS	LD_{50} contact	N
acephate	30560191	1.352997	5
aldicarb	116063	0.322727	3
azinphos-ethyl	2642719	1.39	3
azinphos-methyl	86500	0.4837	6
camphechlor	8001352	50.4	43
carbaryl	63252	3.911333	9
carbofuran	1563662	0.179105	12
carbophenothion	786196	37.5	7
chlorfenvinphos	470906	3.070831	3
chlorpyrifos	2921882	0.05948	4
DDT	50293	6.596509	44
deltamethrin	52918635	0.014562	3
demeton	8065483	2.440082	32
diazinon	333415	0.318124	7
dichlorvos	62737	0.580323	3
dicrotophos	141662	0.167009	6
dieldrin	60571	0.212424	17
dimethoate	60515	0.164097	20
disulfoton	298044	0.96	4
endosulfan	115297	8.394479	6
endrin	72208	1.768136	14
ethion	563122	4.179414	1
fenthion	55389	0.398271	9
fenvalerate	51630581	0.418185	3
formothion	2540821	28.447	2
heptachlor	76448	0.864	11
lindane	58899	0.334664	3
malathion	121755	0.534175	17
methidathion	950378	0.21742	4
methomyl	16752775	0.672701	4
methoxychlor	72435	23.57	5
mevinphos	26718650	0.176692	16
mexacarbate	315184	0.553	1
monocrotophos	6923224	3.907685	17

Active Ingredient Name	CAS	LD_{50} contact	N
naled	300765	0.528091	3
oxydemeton-methyl	301122	3.801853	4
parathion	56382	0.168392	46
parathion-methyl	298000	0.254134	14
phenthoate	2597037	0.306	3
phosphamidon	13171216	1.634564	10
pirimicarb	23103982	32.06508	11
propoxur	114261	2.120259	8
pyridaphenthion	119120	0.08	3
quinalphos	13593038	0.17	3
tetrachlorvinphos	22248799	1.535057	3
thiometon	640153	0.56	4
triazophos	24017478	0.055	3
trichlorfon	52686	59.8	36
vamidothion	2275232	0.56	2

APPENDIX B: Species and crops included in the analysis of the single spray g/ha dataset.

Invert Order	Invert Family	genus spp	Crop	N
Araneae	Lycosidae	Lycosa pseudoannulata	rice	3
Araneae	Micryphantidae	Oedothorax insecticeps	rice	3
Araneae	Unknown	Unknown spp	cotton	7
Araneae	Unknown	Unknown spp	rice	28
Coleoptera	Coccinellidae	Adalia bipunctata	vegetable	3
Coleoptera	Coccinellidae	Brumus spp	cotton	5
Coleoptera	Coccinellidae	Coccinella 9-nota	cotton	3
Coleoptera	Coccinellidae	Coccinella septempunctata	vegetable	10
Coleoptera	Coccinellidae	Coccinella undecimpunctata	cotton	3
Coleoptera	Coccinellidae	Coccinella undecimpunctata	vegetable	3
Coleoptera	Coccinellidae	Coleomegilla maculata	alfalfa	1
Coleoptera	Coccinellidae	Coleomegilla maculata	cotton	5
Coleoptera	Coccinellidae	Cycloneda sanguinea	cereal grains	5
Coleoptera	Coccinellidae	Hippodamia convergens	alfalfa	28
Coleoptera	Coccinellidae	Hippodamia convergens	cotton	10
Coleoptera	Coccinellidae	Hippodamia convergens	vegetable	9
Coleoptera	Coccinellidae	Hippodamia spp	alfalfa	6
Coleoptera	Coccinellidae	Hippodamia spp	cotton	5
Coleoptera	Coccinellidae	Scymnus spp	cereal grains	5
Coleoptera	Coccinellidae	Scymnus spp	cotton	6
Coleoptera	Coccinellidae	Unknown spp	alfalfa	12
Coleoptera	Coccinellidae	Unknown spp	cereal grains	6
Coleoptera	Coccinellidae	Unknown spp	cotton	3
Coleoptera	Coccinellidae	Unknown spp	vegetable	4
Coleoptera	Malachiidae	Collops spp	alfalfa	4
Coleoptera	Malachiidae	Collops spp	cotton	1
Coleoptera	Malachiidae	Collops vittatus	alfalfa	2
Coleoptera	Malachiidae	Collops vittatus	cotton	2
Coleoptera	Staphylinidae	Paederus alfierii	cotton	3
Diptera	Syrphidae	Scaeva pyrastri	alfalfa	4
Diptera	Syrphidae	Syrphus spp	alfalfa	5
Diptera	Syrphidae	Unknown spp	cereal grains	1
Diptera	Syrphidae	Unknown spp	vegetable	1
Hemiptera	Anthocoridae	Orius insidiosus	cotton	8
Hemiptera	Anthocoridae	Orius spp	alfalfa	27

Invert Order	Invert Family	genus spp	Crop	N
Hemiptera	Anthocoridae	Orius spp	cotton	5
Hemiptera	Anthocoridae	Orius tristicolor	alfalfa	24
Hemiptera	Anthocoridae	Orius tristicolor	cotton	5
Hemiptera	Belostomatidae	Belostoma spp	rice	1
Hemiptera	Lygaeidae	Geocoris pallens	cotton	5
Hemiptera	Lygaeidae	Geocoris punctipes	cotton	5
Hemiptera	Lygaeidae	Geocoris spp	alfalfa	26
Hemiptera	Lygaeidae	Geocoris spp	cotton	5
Hemiptera	Lygaeidae	Geocoris spp	vegetable	3
Hemiptera	Miridae	Cyrtorhinus lividipennis	rice	28
Hemiptera	Nabidae	Nabis alternatus	alfalfa	4
Hemiptera	Nabidae	Nabis alternatus	cotton	3
Hemiptera	Nabidae	Nabis americoferus	cotton	7
Hemiptera	Nabidae	Nabis ferus	alfalfa	23
Hemiptera	Nabidae	Nabis ferus	cotton	3
Hemiptera	Nabidae	Nabis roseipennis	cotton	4
Hemiptera	Nabidae	Nabis spp	alfalfa	17
Hemiptera	Nabidae	Nabis spp	cotton	5
Hemiptera	Reduviidae	Sinea diadema	alfalfa	5
Hemiptera	Reduviidae	Unknown spp	cotton	1
Hemiptera	Veliidae	Microvelia atrolineata	rice	28
Neuroptera	Chrysopidae	Chrysopa spp	alfalfa	26
Neuroptera	Chrysopidae	Chrysopa spp	cotton	16
Neuroptera	Chrysopidae	Chrysopa vulgaris	cotton	6
Neuroptera	Chrysopidae	Unknown spp	cereal grains	3
Neuroptera	Hemerobiidae	Micromus tasmaniae	vegetable	4
Thysanoptera	Aeolothripidae	Aeolothrips fasciatus	alfalfa	4

APPENDIX C: Values of bee toxicity for chemicals used in this analysis. The geometric mean of the available LD_{50} estimates was used to calculate the hazard ratio.

Chemical Name	CAS	geomean LD_{50} (µg/bee)	Study Type	values of LD_{50} (µg/bee)	Source	Reference
acephate	30560191	1.35	Contact	1.20	AGRITOX; REF	ATKINS; PM 2000 (12th)
			Contact	1.20	One liner 2004	WSU 1971
			Contact	1.72	University of California	
aldicarb	116063	0.32	Contact	0.29	AGRITOX	US department of Agriculture
			Contact	0.29	One liner 2004	UCR 1975
			Contact	0.49	University of California	
azinphos-ethyl	2642719	1.39	Contact	1.39	University of California	
azinphos-methyl	86500	0.48	Contact	0.42	One liner 2004	UCR 1976
			Contact	0.42	AGRITOX	ATKINS
			Contact	0.64	University of California	
camphechlor	8001352	50.40	Contact	50.40	One liner 2004	UCR 1973
carbaryl	63252	3.91	Contact	1.00	AGRITOX	PM 96
			Contact	1.00	REF	PM 2000 (12th)
			Contact	1.02	AGRITOX	EHC 153
			Contact	1.02	One liner 2004	UCR 1975
			Contact	1.30	One liner 2004	REF 1978
			Contact	4.05	University of California	
			Contact	100.00	University of California	
carbofuran	1563662	0.18	Contact	0.16	One liner 2004; Agritox	UCR 1975; US department of Agriculture
			Contact	0.24	University of California	
carbophenothion	786196	37.50	Contact	37.50	University of California	
chlorfenvinphos	470906	3.07	Contact	4.10	REF	PM 2000 (12th)

Chemical Name	CAS	geomean LD_{50} (µg/bee)	Study Type	values of LD_{50} (µg/bee)	Source	Reference
			Contact	0.4-4.1	AGRITOX	Shell Chimie
chlorpyrifos	2921882	0.059	Contact	0.010	One liner 2004	REF 1969
			Contact	0.059	AGRITOX	Dictionary of substances and their effects
			Contact	0.059	One liner 2004	REF 1978
			Contact	0.070	AGRITOX; REF	DowElanco; PM 2000 (12th)
			Contact	0.11	One liner 2004	UCR 1976
			Contact	0.15	University of California	
			Contact	0.059-0.07	AGRITOX	Makhteshim (ISRAEL)
DDT	50293	6.60	Contact	3.90	One liner 2004	REF 1968
			Contact	6.40	One liner 2004	WSU 1963
			Contact	11.50	University of California	
deltamethrin	52918635	0.015	Contact	0.0015	One liner 2004	WLI 1991
			Contact	0.0015	REF	Europa reports
			Contact	0.010	REF	Europa reports
deltamethrin	52918635	0.015	Contact	0.051	INCHEM; REF	Stevenson et al. (1978); PM 2000 (12th); Europa reports
			Contact	0.067	One liner 2004	UCR 1976
			Contact	0.124	University of California	
demeton	8065483	2.44	Contact	2.29	University of California	
			Contact	2.60	One liner 2004	UCR 1975
diazinon	333415	0.32	Contact	0.20	One liner 2004	REF 1968
			Contact	0.22	One liner 2004	REF 1964
			Contact	0.37	One liner 2004	UCD 1975
			Contact	0.37	AGRITOX	ATKINS
			Contact	0.54	University of California	

Chemical Name	CAS	geomean LD_{50} (µg/bee)	Study Type	values of LD_{50} (µg/bee)	Source	Reference
dichlorvos	62737	0.58	Contact	0.50	AGRITOX	ATKINS
			Contact	0.50	One liner 2004	UCR 1975
			Contact	0.65	AGRITOX	EHC 79
			Contact	0.71	University of California	
dicrotophos	141662	0.17	Contact	0.08	One liner 2004	REF 1968
			Contact	0.37	University of California	
dieldrin	60571	0.21	Contact	0.14	One liner 2004	UCR 1973
			Contact	0.16	One liner 2004	REF 1968
			Contact	0.43	University of California	
dimethoate	60515	0.16	Contact	0.12	AGRITOX	Agrochemicals Handbook
			Contact	0.16	One liner 2004	HRC 1972
			Contact	0.17	One liner 2004	HRC 1974
			Contact	0.19	One liner 2004	UCR 1975
			Contact	0.32	University of California	
			Contact	0.098-0.12	INCHEM	Stevenson (1968)
			Contact	0.1-0.2	REF	PM 2000 (12th)
disulfoton	298044	0.96	Contact	0.96	One liner 2004	MIL 1992
			Contact	1.11	One liner 2004	MIL 1992
			Contact	4.10	One liner 2004	REF 1968
			Contact	72.70	University of California	
endosulfan	115297	8.39	Contact	4.50	One liner 2004	UCR 1967
			Contact	7.10	INCHEM; One liner 2003	Stevenson et al. (1978); REF 1968
			Contact	7.81	AGRITOX	ATKINS
			Contact	19.90	University of California	
endrin	72208	1.77	Contact	0.65	One liner 2004	REF 1978

Chemical Name	CAS	geomean LD_{50} (µg/bee)	Study Type	values of LD_{50} (µg/bee)	Source	Reference
			Contact	2.02	One liner 2004	UCR 1975
			Contact	4.21	University of California	
ethion	563122	4.18	Contact	0.85	One liner 2004	REF 1973
			Contact	20.55	One liner 2004	UCR 1975
fenthion	55389	0.40	Contact	0.31	One liner 2004	UCR 1975
			Contact	0.52	University of California	
fenvalerate	51630581	0.42	Contact	0.23	REF	PM 2000 (12th)
			Contact	0.23	REF	PM 2000 (12th)
			Contact	0.41	INCHEM	Atkins et al. (1981)
			Contact	0.41	One liner 2004	REF 1981
			Contact	0.41	INCHEM	Atkins et al. (1981)
			Contact	0.41	One liner 2004	REF 1981
			Contact	0.79	University of California	
			Contact	0.79	University of California	
formothion	2540821	28.45	Contact	28.45	REF	PM 2000 (12th)
heptachlor	76448	0.86	Contact	0.86	University of California	
lindane	58899	0.33	Contact	0.20	One liner 2004	REF 1978
			Contact	0.56	One liner 2004	UCR 1975
malathion	121755	0.53	Contact	0.20	One liner 2004	REF 1978
			Contact	0.27	One liner 2004	REF 1968
			Contact	0.71	One liner 2004	UCR 1967
			Contact	0.71	AGRITOX; REF	ATKINS; PM 2000 (12th)
			Contact	1.60	University of California	
methidathion	950378	0.22	Contact	0.24	One liner 2004	UCR 1975
			Contact	0.34	University of California	
methomyl	16752775	0.67	Contact	0.10	REF	PM 2000 (12th)

Chemical Name	CAS	geomean LD_{50} (µg/bee)	Study Type	values of LD_{50} (µg/bee)	Source	Reference
			Contact	0.16	One liner 2004	GAB 2000
			Contact	1.29	INCHEM; Agritox	Atkins et al. (1976); EHC 178
			Contact	1.51	AGRITOX	La Littorale
			Contact	4.42	University of California	
methoxychlor	72435	23.57	Contact	23.57	One liner 2004	UCR 1975
mevinphos	26718650	0.18	Contact	0.03	REF	PM 2000 (12th)
			Contact	0.45	University of California	
mexacarbate	315184	0.55	Contact	0.55	University of California	
monocrotophos	6923224	3.91	Contact	0.51	University of California	
			contact	25-35	REF	PM 2000 (12th)
naled	300765	0.53	Contact	0.48	AGRITOX	ATKINS
			Contact	0.58	University of California	
oxydemeton-methyl	301122	3.80	Contact	0.54	AGRITOX	Bayer France
			Contact	2.15	One liner 2004	UCR 1980
			Contact	3.00	One liner 2004; Agritox	UCR 1975; Atkins
			Contact	9.35	University of California	
			Contact	24.39	One liner 2004	UCR 1980
parathion	56382	0.17	Contact	0.18	One liner 2004; Agritox	UCR 1975; Agrochemicals Handbook
parathion-methyl	298000	0.25	Contact	0.11	One liner 2004	UCR 1981
			Contact	0.17	INCHEM	Dankana et al. (1986)
			Contact	0.20	University of California	
			Contact	0.21	One liner 2004	UCR 1981
			Contact	0.29	AGRITOX	ATKINS
			Contact	0.43	University of California	

Chemical Name	CAS	geomean LD_{50} (µg/bee)	Study Type	values of LD_{50} (µg/bee)	Source	Reference
phenthoate	2597037	0.31	Contact	0.31	REF	PM 2000 (12th)
phosphamidon	13171216	1.63	Contact	1.46	One liner 2004	UCR 1975
			Contact	1.83	University of California	
pirimicarb	23103982	32.07	Contact	18.70	AGRITOX	ATKINS
			Contact	18.72	One liner 2004	UCR 1975
			Contact	52.20	University of California	
			Contact	53.00	REF	PM 2000 (12th)
			Contact	20-50	AGRITOX	Imperial Chemical industries
propoxur	114261	2.12	Contact	1.35	One liner 2004	REF 1969
			Contact	3.33	University of California	
pyridaphenthion	119120	0.080	Contact	0.08	REF	PM 2000 (12th)
quinalphos	13593038	0.17	contact	0.17	REF	PM 2000 (12th)
			contact	0.17	REF	PM 2000 (12th)
tetrachlorvinphos	22248799	1.54	Contact	1.37	One liner 2004	UCR 1975
			Contact	1.72	University of California	
triazophos	24017478	0.055	Contact	0.055	INCHEM	
trichlorfon	52686	59.80	Contact	59.80	One liner 2004; Agritox	UCR 1975; Atkins
			Contact	59.80	One liner 2004; Agritox	UCR 1975; Atkins
vamidothion	2275232	0.56	Contact	0.56	REF	Villa et al 2000

APPENDIX D: Detailed description of exploratory models, having divided the database by crop and by family group.

Section A. Effect of Crop Type.

There are 5 main crop types in the SELCTV database: alfalfa, cotton, rice, vegetables and cereal grains. Vegetable crops include non-described vegetable crops, tomato, cabbage, sugar beets, beans, soybeans, and sorghum. Each crop was analyzed with both order and family included as categorical factors.

The largest crop in the data set is alfalfa, with 218 entries. The best model includes HR, vp, and soil half life as continuous predictors, and family as a categorical variable (Table D1). The continuous predictors also combine to significantly predict mortality, although not much of the variation is explained (r^2=0.12 p=0.00001; Figure D1). In both these models, there is a 2 fold effect of HR on mortality, while soil and vapour pressure factors only have about a 1.2-1.3 fold effect on mortality. The categorical variable family has the most predictive value.

Table D1: Akaike's Information Criteria (AIC) for alfalfa. N =218 All predatory species, and pesticides with application rate in g/ha included in analysis.

variable 1	variable 2	variable 3	variable 4	variable 5	df	AIC	Δ AIC	Akaike weight	ratio	L. Ratio Chi^2	p<
log HR	log vp	log soil	Family		11	2019.59	0.00	0.15	1.00	54.37	0.0000001
log HR	Family				9	2019.67	0.08	0.14	1.04	50.29	0.0000001
log HR	log vp	Family			10	2019.92	0.32	0.13	1.18	52.04	0.0000001
log HR	log vp	PCA	Family		11	2020.76	1.17	0.08	1.79	53.20	0.0000001
log HR	log soil	Family			10	2020.84	1.25	0.08	1.87	51.12	0.0000001
log HR	PCA	Family			10	2021.40	1.81	0.06	2.48	50.55	0.0000001

Table D1: Akaike's Information Criteria (AIC) for alfalfa. N =218 All predatory species, and pesticides with application rate in g/ha included in analysis.

variable 1	variable 2	variable 3	variable 4	variable 5	df	AIC	Δ AIC	Akaike weight	ratio	L. Ratio Chi^2	p<
log HR	log vp	log soil	log foliar	Family	12	2021.41	1.82	0.06	2.48	54.55	0.0000001
log HR	log vp	log soil	PCA	Family	12	2021.56	1.97	0.06	2.67	54.40	0.0000001
log HR	log foliar	Family			10	2021.60	2.01	0.05	2.73	50.36	0.0000001

Equation D1: Model parameters for the best model in alfalfa.

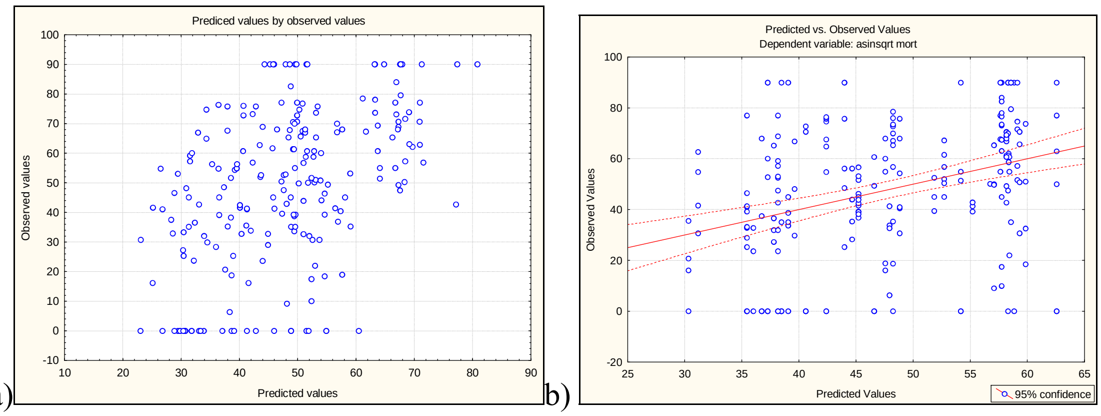
Asinsqrt(mortality)=intercept +9.24*(log HR) – 3.04*(log vp) – 2.66*(log soil DT_{50})

	Intercept
Coccinellidae	24.34
Malachiidae	34.18
Syrphidae	26.62
Anthocoridae	6.23
Lygaeidae	9.84
Nabidae	8.28
Reduviidae	28.71
Chrysopidae	20.73
Aeolothripidae	9.63

Equation D2: Predictive capability of HR, vp, and soil half life in alfalfa crops

Asinsqrt(mortality)=11.58 +9.46*(log HR) – 3.58*(log vp) – 2.65*(log soil DT_{50})

Figure D1: Predicted vs observed mortality in alfalfa. a) Best model with family as a categorical factor Equation A1, b) Best model with only HR, and soil DT_{50}, Equation A2.



The smallest crop represented in the SELCTV database is cereal grains (N=20). The best model for this crop contains log soil DT_{50} and PCA as listed in Table D2. The r^2 is higher (0.31) but this is likely spurious given the very small sample size. Soil half life has a three fold effect on mortality, while PCA has a 6 fold effect on mortality. All of the organisms in this group are aphid feeders (except for the one Syrphidae).

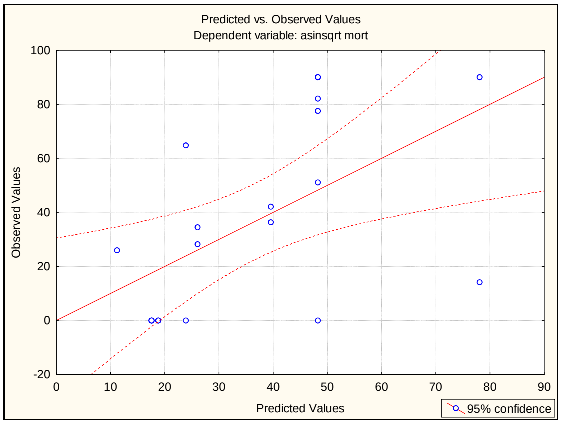
Equation D3: The best model in cereal grains:

Asinsqrt(mortality)=59.03 - 19.67*(log soil DT_{50})+ 8.36*PCA

Table D2: Akaike's Information Criteria (AIC) for cereals. N =20 All predatory species, and pesticides with application rate in g/ha included in analysis.

df	K	variable 1	variable 2	AIC_c	Δ AIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
2	4	log soil	PCA	199.40	0.00	0.15	1.00	7.32	0.026
1	3	PCA		199.92	0.52	0.12	1.30	3.63	0.057
1	3	log soil		200.61	1.21	0.08	1.83	2.94	0.086
2	4	log vp	PCA	201.36	1.96	0.06	2.67	5.35	0.069
2	4	log HR	PCA	201.44	2.04	0.06	2.77	5.28	0.071
2	4	Family		201.79	2.39	0.05	3.31	4.92	0.085

Figure D2: Predicted vs observed mortality in cereal crops.



Cotton crops include data on 5 orders, 10 families, and 26 species of invertebrates. In spite of this variation, the categorical factors order and family do not occur in the top models (Table A3). In cotton crops, the best model (HR and PCA) did not significantly predict mortality (r^2=0.076 p=0.006; Figure A3).

Equation D4: The best model in cotton crops:

Asin(sqrt) mortality in cotton crops= 36.60+ 7.67*(log HR) - 1.14*(PCA)

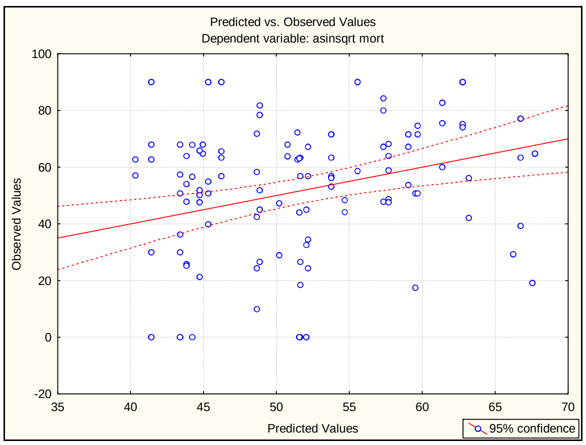
Table D3: Akaike's Information Criteria (AIC) for cotton crops (N =131). All predatory species, and pesticides with application rate in g/ha included in analysis. Only models with weight ratio less than 5 presented.

df	K	variable 1	variable 2	variable 3	AIC_c	Δ AIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
2	4	log HR	PCA		1291.90	0.00	0.16	1.00	8.83	0.012
2	4	log HR	log soil		1293.09	1.19	0.09	1.82	7.64	0.022
3	5	log HR	log soil	PCA	1293.73	1.83	0.06	2.50	9.15	0.027
1	3	log HR			1293.77	1.87	0.06	2.55	4.84	0.028

Table D3: Akaike's Information Criteria (AIC) for cotton crops (N =131). All predatory species, and pesticides with application rate in g/ha included in analysis. Only models with weight ratio less than 5 presented.

3	5	log HR	log vp	PCA	1293.90	2.00	0.06	2.72	8.99	0.029
3	5	log HR	log soil	log foliar	1293.93	2.03	0.06	2.76	8.96	0.030
3	5	log HR	log foliar	PCA	1294.06	2.15	0.05	2.94	8.83	0.032
1	3	PCA			1294.55	2.64	0.04	3.75	4.06	0.044

Figure D3: Predicted versus observed mortality in cotton crops.



Rice crops, unlike cotton, grass and vegetables, are often grown submerged under water. Data represent 3 orders including Spiders (Lycosidae, Micryphantidae and unknown families; 36% of total), Coleoptera (Dytiscidae, and Hydrophilidae; 4% of total), and Hemiptera (Miridae, and Veliidae; 60%). The best model includes foliar DT_{50} and family as a categorical variable (Table D4). This model does not explain variation in mortality as seen in Figure D4.

Table D4: Akaike's Information Criteria (AIC) for rice crops (N =90). All predatory species and pesticides with application rate in g/ha included in analysis.

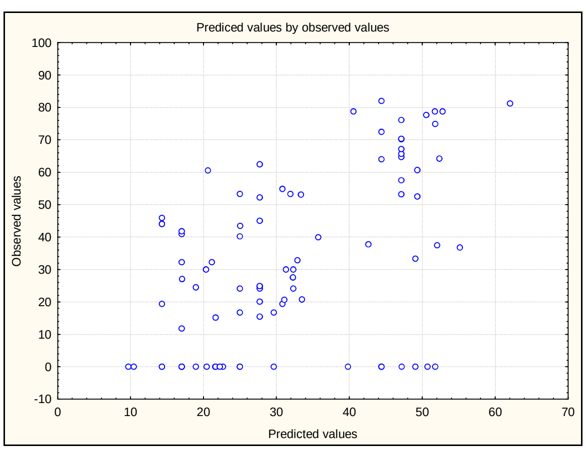
K	variable 1	variable 2	variable 3	df	AIC_c	ΔAIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
7	log foliar	Family		5	832.86	0.00	0.17	1.00	25.78	0.00010
6	Family			4	833.62	0.76	0.12	1.46	22.67	0.00015
7	log vp	Family		5	834.24	1.38	0.09	1.99	24.40	0.00018
8	log vp	log foliar	Family	6	834.72	1.86	0.07	2.54	26.33	0.00019

Equation D5: Model parameters for the best model in rice crops.

Asinsqrt(mortality)=intercept +15.36*(log foliar DT_{50})

	Intercept
Lycosidae	26.17
Micryphantidae	44.69
Unknown	20.36
Miridae	39.80
Veliidae	9.68

Figure D4: Predicted versus observed mortality in rice crops.



The only good model prediction occurred for vegetable crops, which included 4 orders, 4 families, and 10 species of invertebrates. The best model included family as a categorical variable, and HR and soil DT_{50} as factors (Table D5). The majority of the species present are in the family Coccinellidae (78%), which shows that the categorical factor family corrects for only a few entries. The variables soil DT_{50} and HR significantly predicted mortality (r^2=0.45 p=0.00003) without the addition of family (Figure D5). In both models, there is a three fold effect of HR, and a 2.6 fold effect of soil DT_{50}.

Table D5: Akaike's Information Criteria (AIC) for vegetable crops (N =37). All predatory species and pesticides with application rate in g/ha included in analysis. Only models with weight ratio less than 10 presented.

K	variable 1	variable 2	variable 3	variable 4	df	AIC_c	ΔAIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
7	log HR	log soil	Order		5	332.05	0.00	0.26	1.00	41.58	0.0000001
7	log HR	log soil	Family		5	332.05	0.00	0.26	1.00	41.58	0.0000001
8	log HR	log vp	log soil	Family	6	334.76	2.71	0.07	3.88	42.15	0.0000002
8	log HR	log vp	log soil	Order	6	334.76	2.71	0.07	3.88	42.15	0.0000002
8	log HR	log soil	PCA	Family	6	335.28	3.23	0.05	5.03	41.63	0.0000002
8	log HR	log soil	PCA	Order	6	335.28	3.23	0.05	5.03	41.63	0.0000002
8	log HR	log soil	log foliar	Family	6	335.33	3.28	0.05	5.15	41.58	0.0000002
8	log HR	log soil	log foliar	Order	6	335.33	3.28	0.05	5.15	41.58	0.0000002

Equation D6 Model parameters for the best model in vegetable crops.

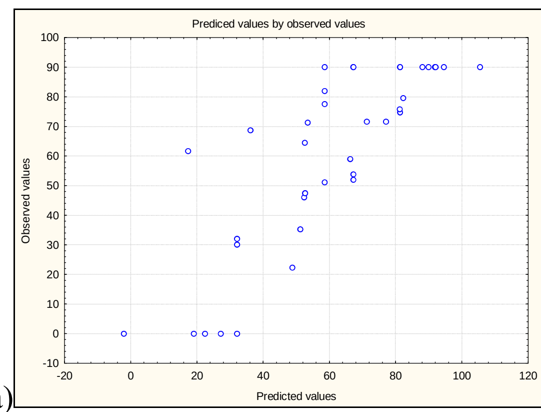
$$\text{Asinsqrt(mortality)} = \text{intercept} + 15.53 * (\log HR) - 14.78 * (\log \text{soil } DT_{50})$$

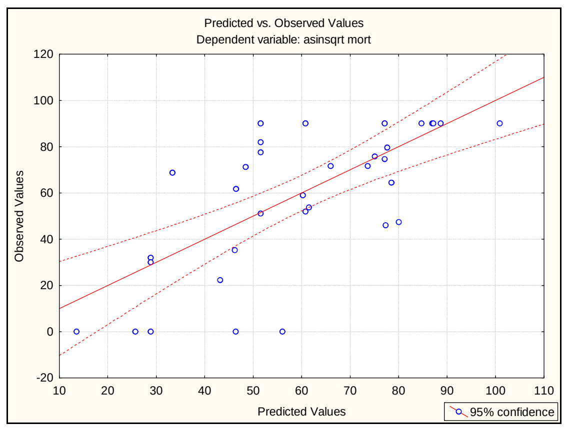
Order	N	Intercept
Coleoptera, Coccinellidae	29	50.3
Diptera, Syrphidae	1	81.7
Hemiptera, Lygaeidae	3	19.9
Neuroptera, Hemerobiidae	4	14.9

Equation D7 Model parameters for the best model in vegetable crops.

Asinsqrt(mortality)=38.83 + 17.47*(log HR) – 13.75*(log soil DT_{50})

Figure D5: Predicted versus observed mortality in vegetable crops. a) Best model with a categorical factor of family (Equation D6). b) Best model with only HR, and soil DT_{50} (Equation A7).

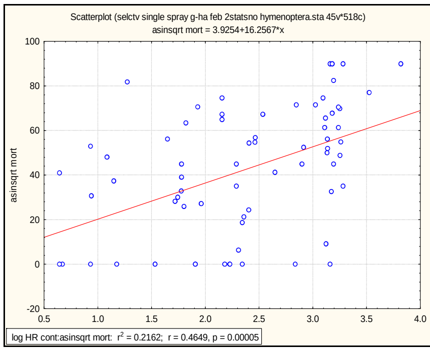
a) 

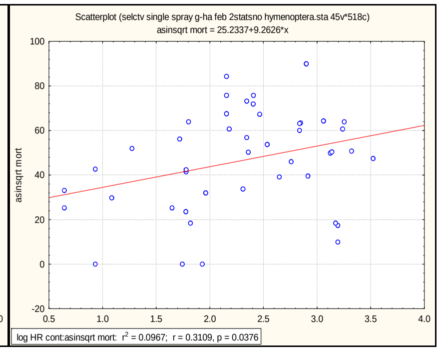
b) 

Section B. Effect of Invertebrate Order.

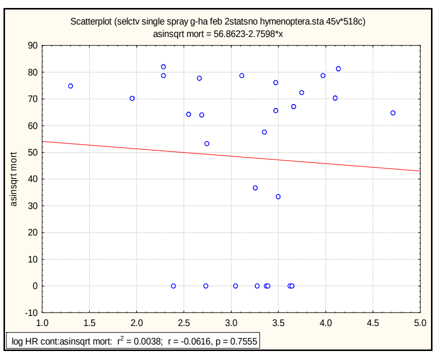
Insects related to honey bees are likely to have similar morphology, and similar toxic responses to pesticides than phylogenetically distant groups. Honey bees are insects with complete metamorphosis, meaning that they have a larval stage and a pupa stage. Other insects (such as Hemiptera) have incomplete metamorphosis, which means that the nymphs grow simply by shedding skin until the adult life stage is reached. Organisms with complete metamorphosis have stronger relationship between HR and mortality, than those with gradual metamorphosis (Figure B1).

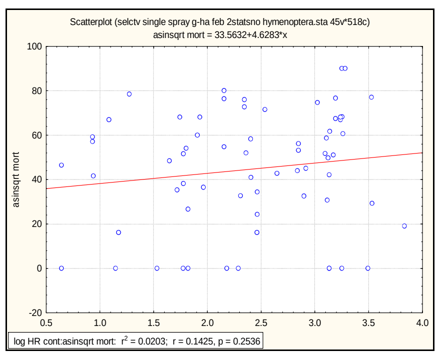
Figure D6. Correlation between HR calculated with honey bee LD_{50} and mortality in five Hemiptera families.



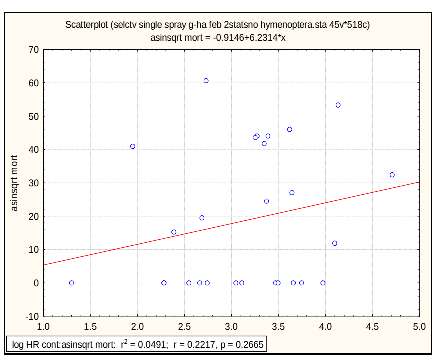


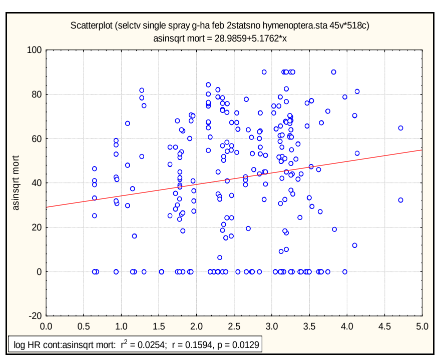
Anthocoridae Lygaeidae



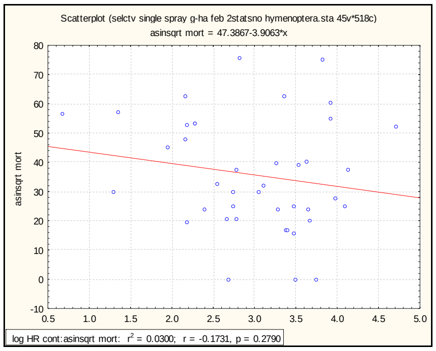


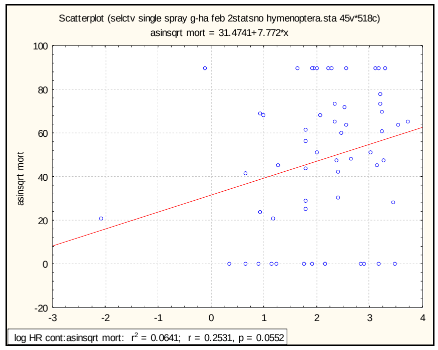
Miridae Nabidae





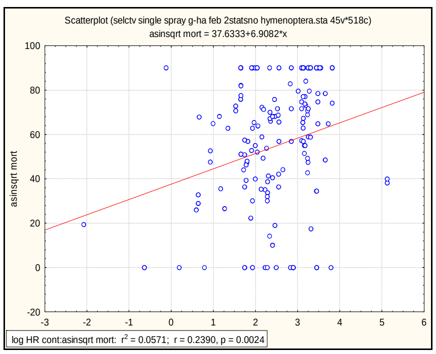
Veliidae All Hemiptera

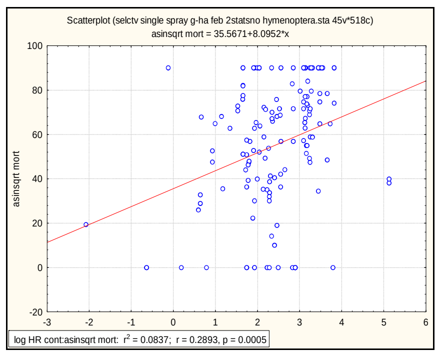




Araneae

Neuroptera





All Coleoptera

Coccinellidae

SELCTV contains one non-insect order which is spiders (Arachnida, Araneae); although some mites are included in the multiple spray data. For the majority of entries however the family has not been identified. SELCTV contains four references to spider mortality, where one study identified two families, and the others did not identify families or species. These studies only examined spider mortality in cotton and rice crops. Of the 10 likely models (weight ratio <10), only two contain HR, which indicates that spiders do not have a similar toxic response to honey bees as seen in Table D6. The best model contains crop type and log soil DT_{50} as seen in Equation D8. Both soil half life and crop are equally important in the model; they both have a two fold effect on mortality.

Table D6: Akaike's Information Criteria (AIC) for all Araneae (N=41) with application rates in g/ha. Only models with weight ratio less than 10 presented.

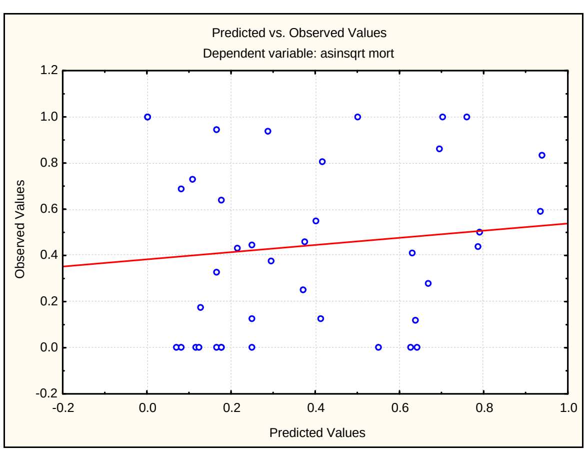
K	variable 1	variable 2	variable 3	df	AIC_c	$\Delta\ AIC_c$	Akaike weight	ratio	L. Ratio Chi^2	p
4	log soil	Crop		2	346.15	0.00	0.22	1.00	18.73	0.000086
3	Crop			1	346.92	0.77	0.15	1.47	15.50	0.000083
5	log soil	PCA	Crop	3	348.52	2.38	0.07	3.28	18.96	0.000279
5	log vp	log soil	Crop	3	348.67	2.53	0.06	3.54	18.81	0.000300
5	log HR	log soil	Crop	3	348.75	2.60	0.06	3.67	18.74	0.000310
5	log soil	log foliar	Crop	3	348.75	2.60	0.06	3.67	18.73	0.000310
4	log vp	Crop		2	349.22	3.07	0.05	4.64	15.66	0.000397
4	log foliar	Crop		2	349.33	3.18	0.04	4.90	15.55	0.000420
4	log HR	Crop		2	349.36	3.21	0.04	4.98	15.52	0.000426
4	PCA	Crop		2	349.37	3.23	0.04	5.02	15.51	0.000429

Equation D8 Model parameters for the best model for all Araneae.

Asinsqrt(mortality)= Intercept + – 8.26*(log soil DT_{50})

	Intercept
rice	20.19
cotton	49.72

Figure D7: Best model in Araneae



The only large group in SELCTV with incomplete metamorphosis is Hemiptera, or true bugs. These insects tend to be generalist predators, and search the leaves and soil for prey. In this order, the best models for Miridae, Nabidae, and Veliidae did not contain HR. The best model for Miridae only contains soil DT_{50} (Table D7) and does not significantly explain mortality (r^2=0.12, p=0.078, Figure D8). The best model for Veliidae only contains PCA, and does not explain the variation in mortality (r^2=0.21, p=0.014, Figure D9). The lack of model fit is likely due to their rice crop habitat. Perhaps, even though these groups are not found immersed in water, factors related to chemical persistence in water and aquatic organism toxicity are better predictors.

Table D7: Akaike's Information Criteria (AIC) for all Miridae (N=28) with application rates in g/ha. Only models with weight ratio less than 2 presented.

K	variable 1	variable 2	df	AIC_c	Δ AIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
3	log soil		1	275.89	0.00	0.18	1.00	3.42	0.064
4	log soil	log foliar	2	276.55	0.67	0.13	1.40	5.49	0.064

Table D8: Akaike's Information Criteria (AIC) for all Veliidae (N=27) with application rates in g/ha. Only models with weight ratio less than 2 presented.

K	variable 1	variable 2	df	AIC_c	Δ AIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
3	PCA		1	238.15	0.00	0.22	1.00	7.02	0.0081
4	log soil	PCA	2	239.25	1.10	0.13	1.73	8.69	0.013
4	log vp	PCA	2	239.88	1.73	0.09	2.37	8.06	0.018

Equation D9: Model parameters for the best model for Miridae:

Asinsqrt(mortality)= 73.56 + – 19.42*(log soil DT_{50})

Equation D10: Model parameters for the best model for Veliidae:

Asinsqrt(mortality)= 12.11 + 3.43*(PCA)

Equation D11: Model parameters for the best model for Nabidae:

Asinsqrt(mortality)= 0.44 + – 9.67*(log soil DT_{50})+ 2.86*(PCA)

Figure D8: Predictive capacity of the best model in Miridae

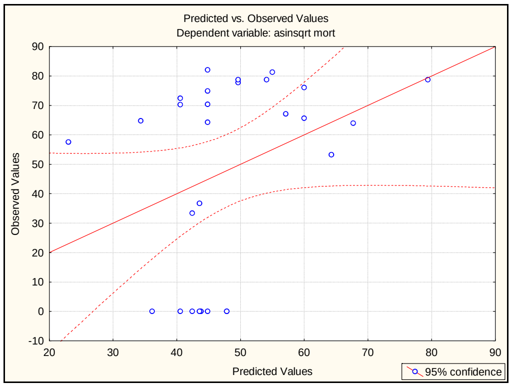
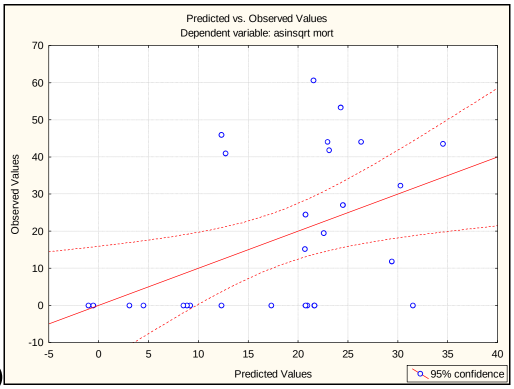
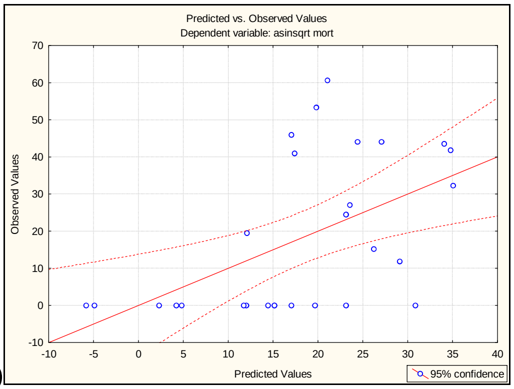


Figure D9: Predictive capacity of the best two models for Veliidae. a) Equation D8 b) Equation D11.

a) 

b) 

The Nabidae in SELCTV are represented by one genus, *Nabis* spp. These are general predators, eating anything smaller than them that they can catch, including many pest species. However, they are not widely recognised as beneficial, since they also feed on other beneficial insects. The best models contained single variables (Table D9), and did not significantly predict mortality (Figure D10).

Table D9: Akaike's Information Criteria (AIC) for all Nabidae (N=66) with application rates in g/ha. Only models with weight ratio less than 2 presented.

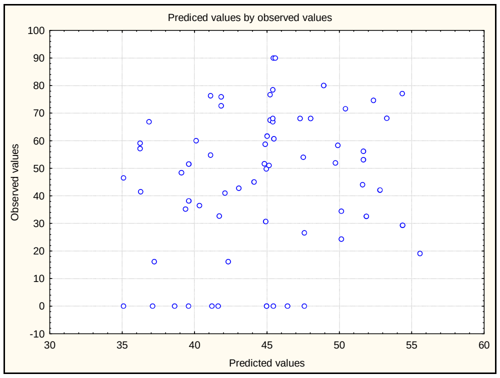
K	variable 1	variable 2	df	AIC_c	ΔAIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
3	Crop		1	620.41	0.00	0.08	1.00	1.69	0.19
3	log HR		1	620.74	0.33	0.07	1.18	1.35	0.24
3	log foliar		1	621.54	1.13	0.05	1.76	0.56	0.46
4	log HR	Crop	2	621.68	1.27	0.04	1.89	2.68	0.26
3	log vp		1	621.73	1.32	0.04	1.93	0.37	0.54

Equation D12 Model parameters for the best model for Nabidae:

Asinsqrt(mortality)= Intercept + 3.97*(log HR)

	Intercept
cotton	3.91
alfalfa	-3.91

Figure D10: Predictive capacity of HR and crop (Equation B5) for Nabidae.



Most Hemiptera are detrimental to crops since many species feed on the juice of plants. Many pesticides have therefore been developed to target these particular pests. However, there are two genera of Hemiptera widely recognised as beneficial, *Geocoris* spp (Hemiptera Lygaeidae) and *Orius* spp (Hemiptera Anthocoridae). Both are the only genera in their family within the data sets. *Geocoris* spp feed on eggs and small larvae or nymphs of most Lepidoptera and Hemiptera, and also on all life stages of whiteflies, mites and aphids. The best model for *Geocoris* spp contains log HR, PCA, and log soil DT_{50} (Table B6). However, the model does not clearly explain the variation in mortality as seen in Figure B6 (r^2=0.21 p=0.0022). In this model all variables have a similar effect on mortality, as there is a 2 fold effect of HR and soil half life, and a 1.6 fold effect of PCA.

Table D10: Akaike's Information Criteria (AIC) for *Geocoris* spp (Hemiptera Lygaeidae) (N=44) with application rates in g/ha.

df	K	variable 1	variable 2	Variable 3	variable 4	AIC_c	Δ AIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
3	5	log HR	log soil	PCA		394.11	0.00	0.16	1.00	10.44	0.015
2	4	log HR	log soil			395.00	0.89	0.10	1.56	6.99	0.030
1	3	log HR				395.70	1.59	0.07	2.22	3.87	0.049

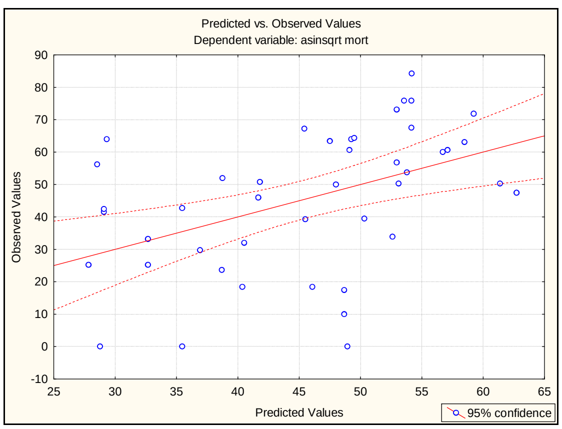
Table D10: Akaike's Information Criteria (AIC) for *Geocoris* spp (Hemiptera Lygaeidae) (N=44) with application rates in g/ha.

df	K	variable 1	variable 2	Variable 3	variable 4	AIC_c	ΔAIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
4	6	log HR	log soil	log foliar	PCA	396.23	2.12	0.05	2.89	11.01	0.026
1	3	log soil				396.30	2.19	0.05	2.99	3.27	0.071

Equation D13 Model parameters for the best model for Lygaeidae:

Asinsqrt(mortality)= 13.72 + 3.97*(log HR) + 8.87*(log soil DT_{50})+ 1.77*(PCA)

Figure D11: Predictive capability of the best model for *Geocoris* spp (Hemiptera Lygaeidae).



Orius spp consume thrips, spider mites, insect eggs, aphids, and small caterpillars. Additionally, adults feed on pollen or plant juices when prey is scarce, which occurs especially after a pesticide application. This makes their post-spray exposure pattern similar to honey bees, since pesticide applications tend to reduce the numbers of prey. The best model for all *Orius* spp (adults, larvae and those with non specified life stage) contains log HR and PCA (table B6). This model has the best mortality predictive capability of any model in Hemiptera order (r^2=0.39 p=0.0000004).

Hazard ratio has the largest effect on mortality (a 5 fold effect); while soil half life and PCA have a 1.7 and 1.8 fold effect respectively.

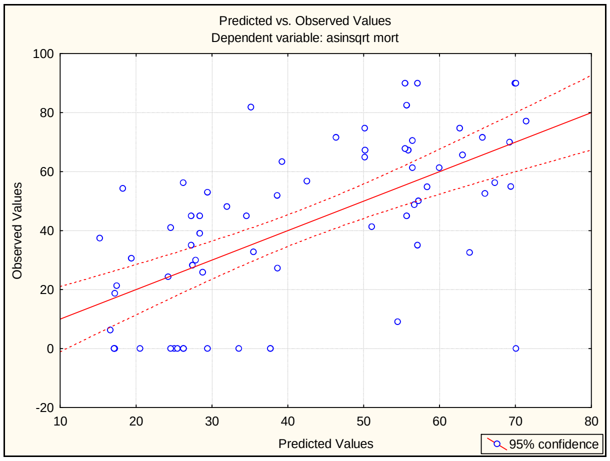
Table D11: Akaike's Information Criteria (AIC) for *Orius* spp (Hemiptera Anthocoridae) (N=69) with application rates in g/ha. Only models with Δ AIC_c less than 2 presented.

df	K	variable 1	Variable 2	variable 3	variable 4	variable 5	AIC_c	Δ AIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
3	5	log HR	log soil	PCA			630.83	0.00	0.11	1.00	34.37	0.0000002
2	4	log HR	PCA				630.97	0.13	0.10	1.07	31.92	0.0000001
4	6	log HR	log soil	PCA	Crop		631.35	0.52	0.08	1.30	36.26	0.0000003
4	6	log HR	log vp	log soil	PCA		631.37	0.54	0.08	1.31	36.24	0.0000003
3	5	log HR	log foliar	PCA			632.07	1.23	0.06	1.85	33.14	0.0000003
5	7	log HR	log vp	log soil	PCA	Crop	632.12	1.29	0.06	1.90	37.97	0.0000004
3	5	log HR	log soil	Crop			632.21	1.38	0.05	1.99	33.00	0.0000003
3	5	log HR	PCA	Crop			632.22	1.38	0.05	2.00	32.99	0.0000003
3	5	log HR	log vp	PCA			632.29	1.45	0.05	2.07	32.92	0.0000003
2	4	log HR	log soil				632.74	1.90	0.04	2.59	30.14	0.0000003

Equation D14 Model parameters for the best model for Anthocoridae:

Asinsqrt(mortality)= 15.19 + 16.76*(log HR) - 5.60*(log soil DT_{50}) – 1.93*(PCA)

Figure B12: Predictive capability of the best model for *Orius* spp (Hemiptera Anthocoridae).



The Neuroptera are voracious aphid predators, although some adults can feed on pollen as well. The biggest family represented in the data set is Chrysopidae, a group which feeds on pollen as adults. The best model includes HR, PCA, and crop as a categorical factor, thought this is closely followed by a model with only HR and PCA shown in Table B7. The second best model does not significantly explain the variation in mortality as illustrated in Figure B8 (r^2=0.19 p=0.0063). The slightly better fit for the model with crop as a category appears to be due to the correction with the cereal grains.

Table D12: Akaike's Information Criteria (AIC) for Chrysopidae (N=51) with application rates in g/ha. Only models with Δ AICc less than 2 presented.

df	K	variable 1	variable 2	Variable 3	variable 4	variable 5	AIC_c	Δ AIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
4	6	log HR	PCA	Crop			490.40	0.00	0.08	1.00	16.00	0.0030
2	4	log HR	PCA				490.61	0.21	0.07	1.11	10.75	0.0046

Table D12: Akaike's Information Criteria (AIC) for Chrysopidae (N=51) with application rates in g/ha. Only models with Δ AICc less than 2 presented.

df	K	variable 1	variable 2	Variable 3	variable 4	variable 5	AIC_c	ΔAIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
4	6	log HR	log soil	log foliar	PCA		491.18	0.78	0.06	1.48	15.22	0.0043
5	7	log HR	log foliar	PCA	Crop		491.40	1.00	0.05	1.65	17.70	0.0033
3	5	log HR	log foliar	PCA			491.84	1.44	0.04	2.06	11.99	0.0074
4	6	log HR	log foliar	Crop			491.84	1.44	0.04	2.06	14.56	0.0057
1	3	PCA					491.84	1.44	0.04	2.06	7.16	0.0075
6	8	log HR	log soil	log foliar	PCA	Crop	491.90	1.50	0.04	2.11	20.03	0.0027
3	5	log HR	log soil	PCA			492.08	1.68	0.04	2.32	11.75	0.0083
3	5	log HR	Crop				492.16	1.77	0.03	2.42	11.66	0.0086
3	5	log soil	log foliar	PCA			492.29	1.89	0.03	2.57	11.54	0.0091
3	5	PCA	Crop				492.29	1.89	0.03	2.57	11.54	0.0091

Equation D15 Model parameters for the best model for Chrysopidae:

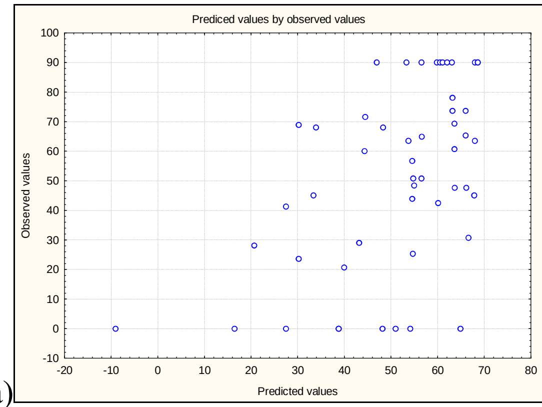
Asinsqrt(mortality)= Intercept + 9.50*(log HR)+ 2.04*(PCA)

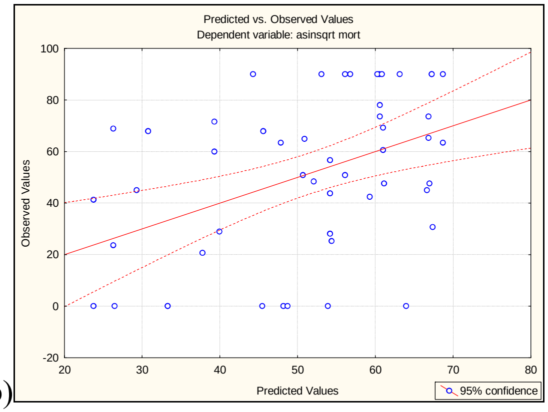
	Intercept
cotton	12.62
alfalfa	12.68
cereal	-12.14

Equation D16 Model parameters for the second best model for Chrysopidae:

Asinsqrt(mortality)= 23.64 + 8.83*(log HR)+ 2.55*(PCA)

Figure D13. Predictive capacity of the best two models for Neuroptera. a) Equation B8 b) Equation D16.

a) 

b) 

The final family is Coccinellidae, a group of aphid feeders. They feed on aphids during larval and adult life stages. The best model contains all factors except for PCA, shown in Table B8. The model without crop does not significantly explain variation in mortality (r^2= 0.23 p=0.0000007 Figure 15). In both models, HR, vapour pressure, and sand soil half life have a 2-3 fold effect on mortality, while foliar half life has only a 1.5 fold effect.

Table D13: Akaike's Information Criteria (AIC) for Coccinellidae (N=132) with application rates in g/ha. Only models with Δ AIC_c less than 2 presented. The hanging table presents the best models without crop as a categorical factor.

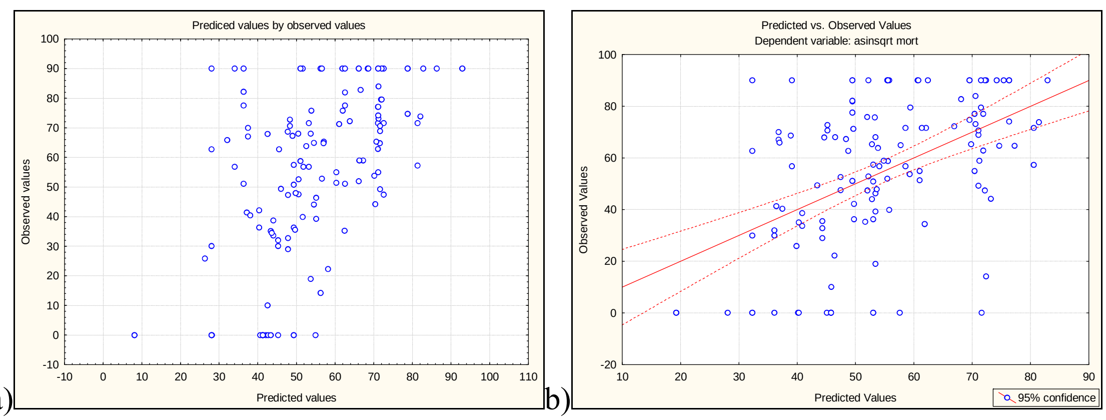
df	K	variable 1	variable 2	Variable 3	Variable 4	variable 5	AIC_c	Δ AIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
7	9	log HR	log vp	log soil	log foliar	Crop	1227.52	0.00	0.22	1.00	47.82	0.00000004
5	7	log HR	log soil	Crop			1227.73	0.21	0.20	1.11	43.05	0.00000004
6	8	log HR	log soil	log foliar	Crop		1227.89	0.37	0.18	1.20	45.15	0.00000004
6	8	log HR	log vp	log soil	Crop		1229.16	1.64	0.10	2.27	43.88	0.00000008

Table D13: Akaike's Information Criteria (AIC) for Coccinellidae (N=132) with application rates in g/ha. Only models with Δ AIC_c less than 2 presented. The hanging table presents the best models without crop as a categorical factor.

df	K	variable 1	variable 2	Variable 3	Variable 4	variable 5	AIC_c	Δ AIC_c	Akaike weight	ratio	L. Ratio Chi^2	p
4	6	log HR	log vp	log soil	log foliar		1233.22	5.70	0.01	17.32	35.32	0.00000040
2	4	log HR	log soil				1233.43	5.91	0.01	19.23	30.75	0.00000021

Figure D14: Predictive capacity of the best two models for Coccinellidae. a) Equation D18 b) Equation D17.

a)



Equation D17 – Model parameters for the first best model for Coccinellidae:

Asinsqrt mort = Intercept + 9.45*(log HR) – 2.84*(log vp) -12.04*(log soil) + 17.84*(log foliar)

	Intercept
Vegetable	37.52
Cotton	23.63
Alfalfa	29.15
Cereal	11.34

Equation D18 – Model parameters for the second best models for Coccinellidae :

Asinsqrt mort = Intercept + 8.98*(log HR) – 8.31*(log soil)

	Intercept
Vegetable	56.44
Cotton	43.59
Alfalfa	48.49
Cereal	30.02

APPENDIX E: Predicting foliar half life

The foliar DT_{50} value was available for 134 of the 207 pesticides ranked here. Most of the values were obtained from the USDA Natural Resources Conservation Service (NRCS) in the form of their Pesticide Properties Database (PPD). The foliar DT_{50} variable was used previously in the evaluation of terrestrial risks to vertebrates (Mineau et al., 2006). Unfortunately, this is probably the least standardised variable to be collected on pesticide active ingredients. Variation undoubtedly occurs at the field level from plant to plant, insect to insect and also because of weather effects (rainfall, humidity, sunlight intensity etc.). Furthermore, there is a lack of method standardization in the literature, so that some sources examine only pesticides on the leaf surface, while others blend fruit or leaves for examination. We used USDA estimates where available and also attempted to create a model that would estimate foliar DT_{50} from other more accessible parameters.

To develop this model, foliar half life values for all chemicals in the Gleams and USDA 2005 PPD databases were used. The complete dataset of Gleams and USDA information was divided in two; chemicals used in Canada, and those not used in Canada. The chemicals used in Canada were included in a training set and all other chemicals were used to validate the model. Some foliar half-lives in the GLEAMS and USDA databases were marked as 'estimated values', and these were not included in our analysis. Subsequent analysis showed that most of these values were clear outliers in the models we developed, suggesting that they had been poorly estimated. For each chemical, the values for the octanol water partition coefficient (K_{ow}), the organic carbon soil sorption coefficient (K_{oc}), and soil DT_{50} were also obtained from the Gleams and USDA databases, as well as from a proprietary database of company data from the Pest Management Regulatory Agency (PMRA) and the Pesticide Manual. Any chemical missing one of the above values was removed from the analysis, which left a total of 123 pesticides with complete data.

In our dataset, the foliar DT_{50} ranged from 0.5-30 days. Some of the values appeared to be over-represented and were therefore suspect; there were 23 chemicals (18%) with a foliar DT_{50} value of 3 days and 25 pesticides (20%) with a value of 5 days. There were also 19 pesticides with a reported DT_{50} of 30 days (15% of total); perhaps this value is more a maximum rather than an actual determined value. Even though we suspect that some of these values are approximations, they were all used in the analysis.

In some studies soil DT_{50} is used to approximate foliar DT_{50} (Villa *et al.,* 2000), and in our data we found that there was a strong correlation between foliar DT_{50} and soil DT_{50} ($r = 0.62$ $p<0.01$; table 1). The variables log water solubility and log K_{ow} were also correlated ($r = -0.74$ $p<0.001$). The log normalized water solubility was correlated with molecular weight and log K_{oc} (both $r = -0.47$ $p< 0.01$). The other variables were not strongly correlated ($r<0.30$).

We used AIC (Akaike information criteria) (Burnham and Anderson, 2002) to identify the most plausible models and relevant variables to select. Small sample size corrections (AIC_c) were calculated in order to keep the models as parsimonious as possible. In keeping with the results of Villa and colleagues (2000), our analysis found that the best models all included log soil DT_{50} (Table 2). The best model contained log soil DT_{50}, log water solubility and log vapour pressure, followed by a model with log soil DT_{50}, log K_{ow} and log vapour pressure. The weight ratio of 1.86 between these two models indicated that they are equally acceptable. The similarity is not surprising, given the high inverse correlation between log water solubility and log K_{ow}. The best model is highly significant and moderately predictive ($R^2 = 0.43$ $p<<0.00001$, Figure 1). There are some outliers, some of which proved to be older chemicals not in use today.

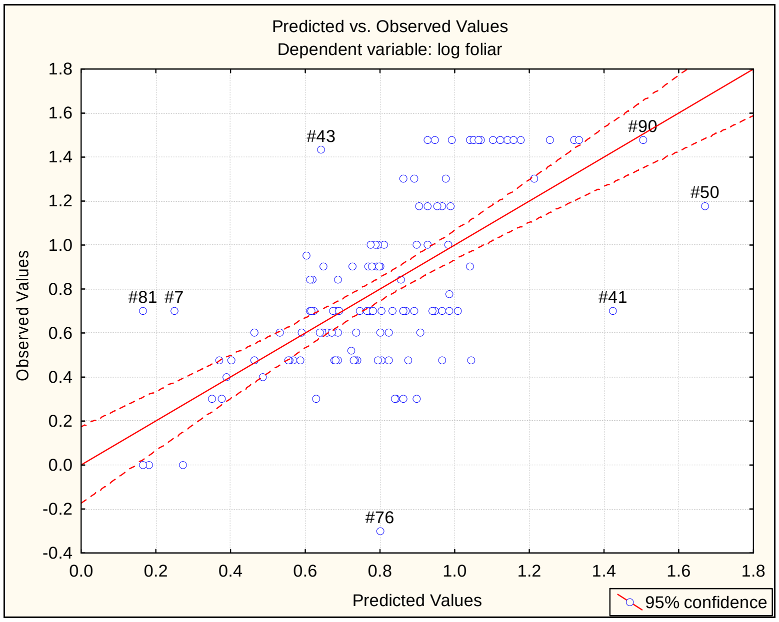
Table E1: Correlations between variables used in analysis.

	Log K_{ow}	mw	log foliar	log soil	log solubility	log K_{oc}	log vp
Mw	.2900	1.0000	.0102	.0491	-.4663	.3071	-.2775
	p=.001	p= ---	P=.911	p=.589	p=.000	p=.001	p=.002
log foliar	-.0731	.0102	1.0000	.6295	.0531	.1230	-.1891
	p=.422	p=.911	P= ---	p=.000	p=.560	p=.175	p=.036
log soil	.0916	.0491	.6295	1.0000	-.1073	.2668	-.0791
	p=.313	p=.589	P=.000	p= ---	p=.238	p=.003	p=.385
log solubility	-.7419	-.4663	.0531	-.1073	1.0000	-.4733	.0392
	p=0.00	p=.000	P=.560	p=.238	p= ---	p=.000	p=.667
log K_{oc}	.3144	.3071	.1230	.2668	-.4733	1.0000	-.1426
	p=.000	p=.001	P=.175	p=.003	p=.000	p= ---	p=.116
log vp	.1117	-.2775	-.1891	-.0791	.0392	-.1426	1.0000
	p=.219	p=.002	P=.036	p=.385	p=.667	p=.116	p= ---

Table E2: Top models found by using AIC analysis N=123. molecular weight = mw vapour pressure=vp

factor 1	factor 2	factor 3	factor 4	df	AICc	Δ AICc	Akaike weight	Ratio	R2	p<<
log soil	log solubility	log vp		3	60.01	0.00	0.177	1.00	0.43	0.000001
log soil	log Kow	log vp		3	61.24	1.24	0.095	1.86	0.43	0.000001
log soil	log vp			2	61.45	1.44	0.086	2.06	0.41	0.000001
log soil	log solubility	log vp	log Koc	4	62.16	2.16	0.060	2.94	0.42	0.000001
log soil	log solubility			2	62.20	2.19	0.059	2.99	0.40	0.000001
log soil	log solubility	log vp	mw	4	62.20	2.20	0.059	3.00	0.42	0.000001
log soil	log Kow			2	62.41	2.41	0.053	3.33		
log soil	log vp	log Koc		3	62.46	2.46	0.052	3.41		
log soil	log vp	mw		3	63.01	3.01	0.039	4.49		
log soil	log vp	log K_{ow}	log Koc	4	63.05	3.04	0.039	4.58		
log soil				1	63.33	3.32	0.034	5.26	0.40	0.000001

Figure E1: Model fit for the best AIC model: log soil DT_{50}, log water solubility and log vapour pressure (R^2= 0.43 p<<0.00001). Some of the worst outliers are identified: 7 Anilazine; 41 Dieldrin; 50 Endrin: 43 Diflubenzuron; 76 Methomyl; 90 Naled.



The equation for the best model was:

Log foliar DT_{50}= -0.024 +0.41* log soil DT_{50} + 0.023* log solubility - 0.031*log vp

Despite the higher AIC_c score, the inclusion of water solubility and vapour pressure in the model only added 3% to the explained variance. Indeed, for the 5 pesticides for which we used the model to predict foliar DT_{50}, we obtained the same answer whether we used the full model or soil DT_{50}, once the answer was rounded to the nearest day.

The validation set of data included 129 chemicals not used in Canada. The best model from table 2 was used to predict foliar half life in the validation set. There was a significant relationship between calculated and observed foliar half life although only 27% of overall variance was explained by the model (r^2=0.27 p<0.000001; see Figure 2).

Figure E2: Observed log foliar DT_{50} vs. calculated log foliar DT_{50} of the validation set, for chemicals not used in Canada. Some of the most obvious outliers are identified.