Chemosphere 166 (2017) 438–444

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Predicting acute contact toxicity of pesticides in honeybees (*Apis mellifera*) through a k-nearest neighbor model

F. Como [a,*], E. Carnesecchi [b], S. Volani [b], J.L. Dorne [b], J. Richardson [b], A. Bassan [c], M. Pavan [c], E. Benfenati [a]

[a] IRCSS Istituto di Ricerche Farmacologiche Mario Negri, via La Masa 19, 20146 Milano, Italy
[b] European Food Safety Authority, Via Carlo Magno 1A, 43126 Parma, Italy
[c] S-IN Soluzioni Informatiche S.r.l., via G. Ferrari 14, 36100 Vicenza, Italy

HIGHLIGHTS

- A model to predict acute contact toxicity for bees was built for screening pesticides.
- The model developed will address future risk assessments of pesticides of concern.
- The accuracy of k-NN model is good and equal to 65% for the highly toxic compounds.

ARTICLE INFO

Article history:
Received 29 June 2016
Received in revised form
9 September 2016
Accepted 21 September 2016

Handling Editor: Jim Lazorchak

Keywords:
Pesticides
Honey bees
k-NN
In silico models
Acute contact toxicity

Abbreviations:
EFSA
European Food Safety Authority
US-EPA
Unites States Environmental Protection Agency
k-NN
K-Nearest Neighbor
LD50
Dose killing half the test organisms after a specified test duration
MCC
Matthew Correlation Coefficient
OECD
Organization for Economic Co-operation and Development

ABSTRACT

Ecological risk assessment of plant protection products (PPPs) requires an understanding of both the toxicity and the extent of exposure to assess risks for a range of taxa of ecological importance including target and non-target species. Non-target species such as honey bees (*Apis mellifera*), solitary bees and bumble bees are of utmost importance because of their vital ecological services as pollinators of wild plants and crops. To improve risk assessment of PPPs in bee species, computational models predicting the acute and chronic toxicity of a range of PPPs and contaminants can play a major role in providing structural and physico-chemical properties for the prioritisation of compounds of concern and future risk assessments.

Over the last three decades, scientific advisory bodies and the research community have developed toxicological databases and quantitative structure-activity relationship (QSAR) models that are proving invaluable to predict toxicity using historical data and reduce animal testing. This paper describes the development and validation of a k-Nearest Neighbor (k-NN) model using in-house software for the prediction of acute contact toxicity of pesticides on honey bees.

Acute contact toxicity data were collected from different sources for 256 pesticides, which were divided into training and test sets. The k-NN models were validated with good prediction, with an accuracy of 70% for all compounds and of 65% for highly toxic compounds, suggesting that they might reliably predict the toxicity of structurally diverse pesticides and could be used to screen and prioritise new pesticides.

* Corresponding author. Laboratory of Environmental Chemistry and Toxicology, IRCSS-Istituto di Ricerche Farmacologiche Mario Negri, Via Giuseppe La Masa, 19, 20156 Milan, Italy.
E-mail address: francesca.como@marionegri.it (F. Como).

http://dx.doi.org/10.1016/j.chemosphere.2016.09.092

PPDB
The Pesticide Properties Database
PPPs
Plant Protection Products
QSAR
Quantitative Structure Alert Relationship
SMILES
Simplified Molecular Input Line Entry System

1. Introduction

Pesticides as plant protection products (PPPs) have an important role in agriculture through the protection of crops from pests such as insects, plant diseases, fungus and weeds and consequently improve productivity and generally speaking food security. From an environmental perspective, scientific advisory bodies and governmental agencies have developed, over the last three decades, methods and frameworks to assess ecological risks of PPPs for a range of taxa including target and non-target species such as birds, bees and fish. The basic principles of these methods and frameworks follow the classic steps for risk assessment (hazard identification and characterization), exposure assessment combining the hazard (toxicological) dimension and the exposure dimension for risk characterization (EFSA, 2012; Dorne and Fink-Gremmels, 2013).

Non-target species of utmost importance in an environmental risk assessment include honey bees (*Apis mellifera*), solitary bees and bumble bees because of their vital ecological services as pollinators of wild plants and crops.

The honey bee, *Apis mellifera*, is arguably the most important pollinator of agricultural crops and is especially exposed to chemicals when visiting melliferous plants (Klein et al., 2007). Special attention has therefore been paid to their protection, not only for their ecological importance by contributing to the maintenance of wild plant biodiversity, but also for their economic value as honey producers and crop-pollinating agents (EFSA, 2012), in fact, new methods of risk assessment specific to bees have been developed recently; see Sanchez-Bayo and Goka (2014). A large number of research studies have investigated pesticide toxicity on bees, new risk assessment methods for bees and the impact of single and multiple stressors (chemicals, diseases, nutrition, etc.) on honey bee colony losses influence honey bee colony losses, and poisoning incidents (Gallai et al., 2009; EFSA, 2012, 2013, 2014, 2015).

During their foraging flights, while they collect nectar, pollen, plant resins, and ingest the nectar and water, honey bees inadvertently come into contact with a wide array of inorganic and organic xenobiotics, often taking them back to the colony where they may induce lethal and sub-lethal effects.

Pesticide sensitivity in honey bees depends on the intrinsic toxicity of the chemical in relation to its structure and target as well as it toxicokinetics including its persistence in the organism and whether metabolism generate toxic metabolites or detoxifies the parent compound (EFSA, 2012). Many environmental factors can also modulate pesticide sensitivity in bees such as the season (Meled et al., 1998; Decourtye et al., 2003), age (Guez et al., 2001), brood rearing temperature (Medrzycki et al., 2010), and stressors such as mixtures of chemicals and pathogens (Alaux et al., 2010; Aufauvre et al., 2012).

Another aspect of the sensitivity of bee species (including honey bees, solitary bees and bumble bees), has been demonstrated with the recent sequencing of the honey bee genome. In deed, comparison with other insect genomes showed that the honey bee genome is markedly deficient in a number of genes encoding detoxification enzymes, including cytochrome P450 monooxygenases (P450s), glutathione-S-transferases and carboxylesterases compared with other insects (Claudianos et al., 2006; Johnson, 2015).

Toxicity of PPPs in honey bees has been currently assessed using 24 h or 48 h acute toxicity laboratory bioassays to determine the median lethal concentration, LC_{50} (concentration that induces 50 percent death) or lethal dose, LD_{50} (dose that induces 50 percent death) as described in the OECD test guidelines (OECD, 1998) and in European regulation (European Regulation (EU) No. 283/2013; PP 1/170, 2010).

To assess PPPs exposure honey bees for different routes (e.g. contact and oral routes) and exposure scenarios need to be addressed and these will depend on the product formulation and its intrinsic characteristics such as its physico-chemical properties. Such physico-chemical properties contribute to environmental persistence, spatial dispersal and potential contamination of feed sources for honey bees (e.g. nectar, bee bread). Overall, the risk assessment thus combines exposure data for possible routes of exposure such as the administered dose or concentration of a toxicant, the duration of the exposure and toxicity data for the active ingredient.

In the last thirty years, scientific advisory bodies and the research community have developed toxicological databases and computational models such as quantitative structure-activity relationship (QSAR) models that are invaluable to predict hazard (toxicity) using historical data, reduce animal testing and serve as tools for regulatory purposes for both the prioritization of compounds of concern and future risk assessments. These QSAR models for species of ecological importance are often developed using the well-characterised relationships between the physico-chemical properties of chemicals, their persistence and toxicity as well as their global environmental fate (Domine et al., 1992; Devillers and Flatin, 2000). QSAR models have already been applied to the prediction of ecotoxicological endpoints in species of ecological importance including trout, daphnia, quail and bees within the project DEMETRA (Benfenati et al., 2011) and more recently for qualitative and quantitative toxicity prediction in bees using a global quantitative structure-toxicity relationship model (QSTR) (Singh et al., 2014).

To address the needs for QSAR models predicting toxicity of PPPs in honey bees, we developed an in-house software using databases on acute contact data in honey bees from different sources and a k-Nearest Neighbor algorithm (k-NN).

2. Material and methods

2.1. Data

Toxicity data on honey bee acute contact toxicity were collected from different sources. The first one was the DEMETRA project (Benfenati et al., 2011). The second source was the Terrestrial US-