9-16-2019

# The Alfalfa Insecticide Management (AIM) Toolkit: Linking insecticide fate modelling with alfalfa pest and beneficial insect toxicity endpoints

Kimberly J. Hageman
*Utah State University*, kim.hageman@usu.edu

**DATA MANAGEMENT PLAN**

## I. Expected data types

1. Chromatograms and mass-spectrometry data will be generated by GC/MS and LC/MS instrumentation. (Digital)
2. Insecticide concentrations in alfalfa foliage and air will be generated by calibration curves and the software systems in GC/MS and LC/MS instruments. (Digital)
3. Weather data (temperature, wind speed, relative humidity) will be logged with the software system in a small portable meteorological station. (Digital)
4. Alfalfa leaf-air partition coefficients will be calculated from insecticide concentrations measured in plant and air samples. (Digital)
5. Insecticide photodegradation rates on alfalfa foliage will be calculated from insecticide concentrations measured during kinetics studies. (Digital)
6. Toxicology endpoints for alfalfa pests, beneficial insects, and pollinators will be calculated from insect counts associated with laboratory experiments. (Initially non-digital, but will be recorded in spreadsheets as digital data.)
7. Field insect count data be generated during the Wyoming Test Case. (Initially non-digital, but will be recorded in spreadsheets as digital data.)
8. The insecticide fate model will store equations, input data, output data, and predictions. (Digital)

## II. Data format

1. Chromatograms and mass-spectrometry data will be generated in instrument manufacturer specific formats, exported and stored as .csv or .xlsx formats.
2. Insecticide concentrations will be generated in instrument manufacturer specific formats, exported and stored in .xlsx formats.
3. Weather data will be generated in instrument manufacturer specific formats, exported and stored in .xlsx formats.
4. Alfalfa leaf-air partition coefficients will be generated and stored in .xlsx formats.
5. Toxicology endpoints will be transcribed and stored in .xlsx formats.
6. Insect count data will be transcribed and stored in .xlsx formats.
7. The insecticide fate model, inputs, outputs, and predictions will be generated and stored in .xlsx formats and R scripts (.rdata).

## III. Data storage and preservation of access

While data is being generated, we will follow the 3-2-1 Backup Rule by storing three copies of all data, with two backup copies on different storage media, and one of them located offsite. Utah State University uses Box.com for offsite cloud storage while University of Wyoming uses OneDrive. The USDS-ARS lab in Logan, Utah has an offsite dedicated server. Files will be backed up nightly by all systems. Undergraduate and graduate students will be required to provide

copies of their field notebooks, data, processing or modeling scripts, and copies of their theses. Further, the resulting theses and dissertations that result from the current project will contain all analytical data and be will be made available via USU libraries. Electronic versions of spectral data as well as data analysis (e.g. spreadsheets) will be archived as stated above. Physical samples not consumed during experiments (leaves and leaf/air extracts) will be stored under appropriate conditions and properly labeled with identity, date, etc. Physical samples lose integrity over time, but will be stored for at least one year after the conclusion of the project.

After completion of the project, all data from the project will be submitted to USDA Ag Data Commons and DigitalCommons@USU. In addition to graphs, plots and images, data will include data set name, any collaborators to be listed in data citation, dates of data collection/analysis, location information (lat/long), parameter names and units, information required for working with the data, methodologies, error sources/assessment, and references. All laboratory notebooks will remain in the laboratories of the PIs.

**IV. Data sharing and public access**

We will publish our results as rapidly as possible in peer-reviewed journals. To increase accessibility and share our results as we conduct the research, we have requested funds to attend national and international meetings. Data will also be made available after publication both as specified by journal guidelines and as raw instrument files. Details of the results will be included in these publications. Full data sets will be provided in the supporting information of our publications, where possible, with a summary of key results provided in the main text of the article. Full data sets and physical samples will also be made available via requests to other investigators. We anticipate publication in peer reviewed journals to be timely and therefore articles containing all pertinent analyzed data and metadata will be publically available for an indefinite period.

*Policies for re-use and redistribution:* In those journals where we anticipate publishing our results, supporting information is provided to all on an unrestricted, open-access basis. While access to the main text requires a journal subscription, articles will be available to all via direct requests to the authors. Primary data and chromatograms, spreadsheets, and other electronic data will be made available upon request to the investigators after publication of the results in the peer-reviewed literature. Whenever possible, data will be provided in a format such that specialized software is not needed to work with the data, spectra, or images (e.g., .xlsx, .pdf or .jpg formats etc.).

**III. Roles and Responsibilities**

PI Hageman will be ultimately responsible for all data management, with co-PI Jabbour taking over as a contingency plan. Based on the individual involvement in each of the objectives, each co-PI will establish a clear and reproducible protocol for data collection for each objective. These protocols will ensure that appropriate metadata are collected along with raw data and that proper storage and backup protocols are maintained.