Literature review and analysis

SENSITIVITY OF WILD PLANT AND CROP SPECIES

IN CONTEXT OF 1107/2009

Author:

Heino Christl
on behalf of the SETAC workshop on “Non-target terrestrial plants”

Title:

Comparative assessment of the sensitivity of wild plant and crop species to plant protection products and their active substances, evaluated in laboratory and field tests, published data and regulatory (unpublished) studies, in context of Regulation 1107/2009 and the upcoming new Terrestrial Guidance document

An initiative of the SETAC tripartite workshop on Terrestrial plants

held in Wageningen, April 1-3, 2014 (and extended after the second workshop)

Data Requirement

Regulation (EC) No 1107/2009, SANCO/10329/2002 rev 2 final (17 October 2002)

Realisation:

tier3 solutions GmbH
Kolberger Strasse 61-63
D-51381 Leverkusen
Germany

tier3 report number:

<B14037 - B15060>

Date of Finalisation:

2017-05-19
(yyyy-mm-dd)

Sponsors:

Responsibilities

Sponsors

European Crop Protection Association
ADAMA Agan Ltd
BASF SE
Bayer CropScience
Dow AgroSciences
DuPont Crop Protection
FMC Corporation
Monsanto Company

Realisation

Heino Christl
tier3 solutions GmbH
Kolberger Strasse 61-63
D-51381 Leverkusen
Germany

Signature and Approval

This literature review was written to the current state of scientific knowledge.

2017-05-19	Heino Christl
Date (yyyy-mm-dd)	Dr. Heino Christl Team leader Risk Assessment tier3 solutions GmbH Kolberger Strasse 61-63 D-51381 Leverkusen Germany Phone + 49 (0) 2171 55997 14

Table of Contents

1 SUMMARY **5**

2 INTRODUCTION **7**

3 MATERIAL AND METHODS **9**

3.1 INITIAL STEPS 9
3.2 SELECTION CRITERIA 14
3.3 PRINCIPLE OF COMPARISON 18
3.4 AUXILIARY ANALYSES 23
3.5 MULTIPLE REGRESSION 23
3.6 ADDITIONAL STATISTICAL ANALYSIS (JOHN W. GREEN) 25
3.7 ABBREVIATIONS FREQUENTLY USED 25

4 RESULTS **27**

4.1 GENERAL OUTCOME AND INITIAL CONSIDERATIONS 27
4.2 SUMMARY OF RESULTS 28
4.3 MULTIPLE REGRESSION ANALYSIS 33
4.4 FURTHER DETAILS OF THE COMPARISON OF CROPS' AND WILD PLANT SPECIES' SENSITIVITY (AUXILIARY ANALYSES) 41

5 DISCUSSION **49**

5.1 RETRIEVAL OF ENDPOINTS FROM THE LITERATURE 49
5.2 QUALITY OF LITERATURE DATA 49
5.3 TEST CONDITIONS 51
5.4 WHICH MEASURE OF SENSITIVITY (LOWEST OR AVERAGE ENDPOINTS) TO BE USED 53
5.5 OVERALL QUOTIENT APPROACH 54
5.6 HETEROGENEITY OF ENDPOINTS AND SELECTION OF FINAL DATASET 54
5.7 DIFFERENCES IN SENSITIVITY BETWEEN LAB AND FIELD TEST SYSTEMS 56
5.8 BACKGROUND INFORMATION ON THE TWO MOST EXTREME DATASETS. 56
5.9 MULTIPLE REGRESSION ANALYSIS 58
5.10 PROBLEMS CAUSED BY CENSORED ENDPOINTS 60
5.11 GERMINABILITY OF CROP AND WILD SPECIES 62
5.12 FURTHER ASPECTS POTENTIALLY RELEVANT FOR THE PROTECTION GOAL BUT NOT ASSESSED IN THIS REVIEW 63
5.13 DEFICIENCIES 65
5.14 OUTLOOK 65

6 CONCLUSION **67**

7 REFERENCES **68**

8 APPENDIX 1 - BRIEF SUMMARIES OF THE MOST RELEVANT PAPERS AND THEIR USE FOR THE PRESENT REVIEW **80**

9 APPENDIX 2 – LIST OF ACTIVE SUBSTANCES **94**

10 APPENDIX 3 – LIST OF SPECIES FOR WHICH NUMERIC ENDPOINTS WERE FOUND IN THE LITERATURE OR IN GLP DATA SUBMITTED FOR REGISTRATION **96**

11 APPENDIX 4 – LIST OF ALL AVAILABLE DATA SETS **103**

12 APPENDIX 5 - LISTS OF TABLES AND FIGURES **108**

13 APPENDIX 6 - INDIVIDUAL DATA SETS PRESENTED AS FIGURES AND SUMMARIZED IN TABLES 120

14 APPENDIX 6 - ADDITIONAL STATISTICAL ANALYSIS OF CROP AND WILD PLANTS (MULTIPLE REGRESSION) Z. GAO ET AL. 168

TODOS 168

INTRODUCTION 168

RESULTS 168

MDD CALCULATION 198
MIXED EFFECT MODELS 204

CONCLUSION 205

DISCUSSION 207

REFERENCES 208

USEFUL LINKS 208

15 APPENDIX 7 - ADDITIONAL STATISTICAL ANALYSIS OF CROP AND WILD PLANTS (COMPARISONS OF DISTRIBUTIONS – JOHN W. GREEN) 209

1 Summary

One of the recommendations of the first SETAC workshop on “Non-target terrestrial plants” was to perform a literature review to compare the sensitivity of terrestrial plant species (crop species and wild species). Published literature and unpublished data generated in the course of the registration of PPPs was searched for this information. The hypothesis propagated by some authors and rejected by others is that wild plant species could be more sensitive to plant protection products than the standard test species - which generally, though not exclusively, are crop species.

Potential differences between greenhouse single species tests and field tests were considered. Variability due to the choice of different endpoint types was minimized by comparing assessment endpoints as reported in published literature and those from standardized tests. The scope of work was adapted and extended during the second SETAC workshop on “Non-target terrestrial plants”, Wageningen 2015.

The overall finding was that based on biomass-based ER10, ER25 and ER50 vegetative vigour endpoints (the largest fraction of data) there were no consistent differences in sensitivity between wild plant species and crop species[1]. The two groups of species were found to be similarly sensitive, and although there were single instances with significant correlations between wild or crop species and modes of action were found, these differences occurred in both directions and were balanced, i.e. there was no trend for any of the two to be more sensitive than the other. MDD-analysis and multivariate regression analysis of factorial modified data sets indicated that for the dataset of this size and heterogeneity, differences between crop and wild species would have been detected as statistically significantly different if they had differed by a factor of 1.5 or more.

It can thus be concluded that for the taxonomic groups for which data were available there is no consistent difference between crop species and wild plant species. Testing 6 to 10 crop species as model organisms in standard toxicity tests is by far more extensive than standard testing in any other ecotox area and covers a range of uncertainties versus testing only one representative organism as usually done. It seems thus to be a pragmatic approach as some basic requirements of the testing guideline can be fulfilled by cultivated species and not by wild species such as the requirement for at least 70% germination within a short timeframe to allow for all plants to be at the same growth stage during application. Based on the subset

[1] Data of seedling emergence studies (seedling emergence, survival, shoot height, vegetative vigour, biomass), and other endpoints from vegetative-vigour-like studies (such as survival, measured shoot height etc. were available for some active substances and were initially assessed as well. They did not show a fundamentally different pattern, but added further uncertainty; and were too few to be analysed on their own. For the final analysis of this report it was decided to focus on numeric vegetative vigour biomass data, the largest subset of data. However, in an additional statistical assessment also censored values were included, and the main figure is also presented as a variant including seedling emergence data and endpoints based on other measured parameters, such as shoot height or survival.

of endpoints with most data[2] there appears to be no reason to include more wild species in standard ecotoxicity testing or to add an additional uncertainty factor for studies where only crop species have been tested.

[2] While the assessment of other endpoints (listed in the previous footnote) was not presented in detail in this paper, the assessment of ER10, ER25 and ER50 endpoints of vegetative vigour (the largest fraction) is considered to allow also the more categorical conclusion given here.

2 Introduction

Terrestrial non-target plants are one of the organism groups regularly tested in the process of approval of plant protection products (PPP). For practical reasons generally crop species are tested as representative for the different plant morphotypes and systematic groups. Seeds of crop species are readily available throughout the year and germinate reliably. This allows having consistent plant material and hence also contributes to consistent and reproducible results. In April 2014 a SETAC workshop on “Non-target terrestrial plants” was organised in Wageningen (The Netherlands). One of the recommendations of the workshop was to perform a literature review investigating whether wild plant species[3] might be generally more sensitive than crop species, and hence whether the testing requirements for plant protection products are protective of non-target terrestrial plants (Boutin et al. 1995, FIFRA / Levis et al 2001, Olszyk et al. 2008, Schmitz et al. 2013). The results and analyses of such literature review are presented in this report.

This assessment was performed considering the SETAC tripartite principle (with members of business, Government and academia contributing). Therefore the selection of the search criteria for this review as well as the review of this document were set and conducted by representatives from business, Government and academia.

Data seen as evidence of supporting the research hypothesis that wild plant species might be significantly more sensitive (see e.g. Davy et al. 2001, EFSA 2014) were generally based on only few species and few active substances (e.g. Strandberg et al. 2012). Also field data and lab/greenhouse data and different endpoint types were sometimes pooled, disregarding that wild plants and crop species were not equally present in all groups (which however is also a weak point in the database assessed here). Furthermore often a variety of non-standard sublethal experimental endpoints was assessed in field studies, without being able to differentiate precisely between actual differences in sensitivity and confounding factors such as testing conditions and test design. Other authors who did test crop- and non-crop species side-by-side under equivalent test conditions generally could not confirm any general trend in terms of different sensitivity between crop and wild plant species (e.g. Boutin et al. 2012, Carpenter and Boutin, 2010, Clark et al. 2004, Egan et al. 2014a, McKelvey et al. 2002, Strandberg et al. 2012, White and Boutin, 2007); so according to their observations the null-hypothesis (‘there is no indication for an intrinsic difference in sensitivity between crop and wild plant species’) must be maintained.

In this paper it was therefore tried to consider all available published data in which ERx-experimental endpoints were recorded, and to merge these data with the experimental

[3] In this review the term “wild plant species” was used in a wide sense, basically any species that had been tested on PPPs and was no crop species. We did not differentiate if the tested species was particularly relevant in terms of likelihood to be exposed (typical for field margins), belonging to a community considered to be vulnerable, being considered as an agricultural pest or weed. It must also be considered that any such differentiation would vary between regions, with the same species being an integral part of field margin communities in one region and an invasive species in another. Virtually all of these tested non-crop species were annual, biennial or perennial herbaceous plants.

endpoints of standard regulatory studies (most of them unpublished). These endpoints were provided by members of the ECPA and CLI NTP-group, and in some cases there were also publicly available data e.g. from DARs, EFSA conclusion reports and the official European Lists of Endpoints. Experimental endpoints provided to the US-EPA were included when they were cited in published papers[4]. All these endpoints were included in the database.

The ultimate goal was a comparison based on standard test parameters (comparing similar assessment endpoints as closely as possible), trying to find evidence for differences in sensitivity between the two groups, and if yes, which of the two groups is more sensitive; the null-hypothesis being that there is no difference in sensitivity between crop species and wild plant species.

[4] No US-EPA database was accessible to be used directly as source.

3 Material and methods

3.1 Initial steps

3.1.1 Literature search

In a first step known published literature to the topic was searched for ERx values ("species – test-substance combinations") and these included in the database.

In a second step a formal literature search was performed on CAB Abstracts, BIOSIS Previews, Web of Science Core Collection, Current Contents Connect, Food Science and Technology Abstracts (FSTA), Medline, Chinese Science Citation Index, and SciELO. Duplicate references across databases were removed. The search strategy was:

#1 ts=((nontarget* or "non target*" or native) near/5 (plant or plants or vegetation) near/5 (crop or crops))

#2 ts=((noncrop* or "non crop" or wildflower* or "wild species") near/5 (crop or crops))

#3 ts=(resist* or sensitiv* or toleran* or phytotox* or ecotox* or susceptib*)

(#1 OR #2) AND #3 n data. TSall , which was then d ##

This very broad search resulted in 1418 unique citations which were then individually reviewed (based on title and abstract), and scored on a star system with 5 stars being those references expected to be most relevant and no star for those not relevant. Of the references returned by the search criteria, 22 references were ranked as most relevant and included citations by authors known to have relevant work (e.g. Boutin, Pfleeger, Olszyk, etc.) or for example had titles or abstracts that included the terms "appraisal of herbicide bio-efficacy...", "tolerance of ..", "sensitivity of selected wild and crop...", etc. A total of 44 references received a rating of 4 stars with titles or abstracts that mentioned terms such as "Effects of XXX on non-treated plants...", "Effects of herbicides on species diversity...", "Effects of XXX on Siberian Elm...", "XXX for control of...", " Assessment of the resistance of some tree species...", etc. A total of 39 references received a rating of 3 stars with titles or abstracts that included terms related to efficacy, or weed control or suppression, control of vegetation on roadsides or shelterbelts, or control of "aquatic weeds". A total of 23 references received a rating of 2 stars and included references to irrigation, risks from release of GMO, biological control of weeds, application techniques for aquatic weed control, etc. References that did not include PPP's (e.g. heavy metals, ozone, salinity, etc.), studies that compared effects of plant pests (insect and fungal), compared application methods, evaluated plant resistance (typically at multiples of the maximum application rates), or efficacy studies with only one or two application rates, studies on genetically modified plants, or conference proceedings announcing new products were assigned no stars. Among the references scored with 4 stars only two references were found to contain relevant experimental endpoints[5], and among those scored 3 not a single one appeared to cite ERx values of individual “species – test-substance combinations” and hence to be useful for the present work. It was hence felt that the used rating system was appropriate to identify the relevant publications for the literature search. Relevant references were obtained and reviewed in detail, extracting any reported ERx values. Initially a total of 29 papers were found to report ERx values, or information in a way that – at least partially – could be translated into ERx values (some only approximately). Several of other papers that appeared to be promising in terms of title and abstract only reported the results of their meta-analysis but did not share the actual experimental endpoints they used. The final database includes endpoints of 54 published papers (plus a few from unpublished studies by different authors cited in the former papers), plus data from confidential GLP-studies.

The experimental endpoints from all references identified from the published literature were collated in an Excel database, and numbers of endpoints per active substance and formulation were checked. This data base contained already some crop species endpoints; although most of the data points were for wild plant species. After the publication of the EFSA Scientific Opinion in July 2014, a comparison of the cited literature was conducted, and also vegetative endpoints of studies collected in context of the assessment of reproductive endpoints were included in the database. In a last step data collected for a different paper assessing differences in sensitivity between vegetative and reproductive endpoints (Christl 2017) were assessed and additional vegetative endpoints were also included in the database. Ultimately the two databases were merged into one.

3.1.2 Access to confidential data sets

In a fourth step substances were identified that were likely to allow an individual assessment, i.e. those of which more than 6 wild plant species endpoints were available. For those the members of CLI plant group and ECPA were asked to provide the confidential experimental endpoints of their regulatory studies (GLP, standard guideline, i.e. OECD 227 or OPPTS 850.4250 Vegetative vigour, - OCSPP 850.4150 - Vegetative Vigour (June 2012)). The companies provided summaries of reports or lists of the endpoints, which were also incorporated in the data base. Sometimes data of just one formulation or even only of the technical material were available. In several cases product data of only those formulations that had been tested when the DAR had been generated were available, but not for newer formulations (see discussion). In others data for several formulations were provided.

3.1.3 Modes of action and anonymisation

[5] To avoid confusion in the following sections we generally differentiate between ‘experimental endpoint’, e.g. an ER25 determined in a specific study for a specific substance and plant species, and in ‘species’ endpoint’; which may be either identical with an experimental endpoint (if the latter is the only one) or be the geometric mean of several independent experimental endpoints for the same active substance / test type / species combination, further details see further down, page 13.

In a fifth step the active substances were classified according to their mode of action. This was primarily done to anonymise[6] the different active substances and formulations. Most company data come from confidential studies that fall under data protection laws that limit use in product registrations, therefore anonymisation was required. As a consequence the data sets were merged by mode of action. Different classification schemes for modes of action are available, see e.g. Martin / Ontario Ministry of Agriculture & Food 2014, Ross M.A. and Jordan T N 1999, Menne & Köcher 2007, Schmidt, R. 1999 and WSSA Herbicide Handbook, Weed Science Society of America (1994 and Supplement 1998). A list of the different modes of action (MoAs) considered in this analysis MoAs is given in Table 1. Within each Mode of action-group active substances were numbered. The full list of actives considered in the analysis is provided in Appendix 2 – List of active substances.

Table 1 Modes of action (MoA) of the active substances for which terrestrial plant species endpoints were available. Modes of actions not listed were either not available (no non-target-plant endpoints) or merged with other modes of action.

Code	Mode-of-action	n
AASI	Amino Acid Synthesis Inhibitors	7
SGI	Seedling Growth Inhibitors	5
GW	Growth Regulators	12
PHI	Photosynthetic Inhibitors	10
LSI	Lipid Synthesis Inhibitors	3
CMD	Cell Membrane Disrupters	4
ACI	Acetyl CoA inhibition	2
ICD	Inhibition of cell division	3
OTH	Other (lumped unique or unknown MoA)	7

Please note that this initial list includes any active substance for which a wild-plant endpoint was available (94 substances in total). However, for several of these substances this was only true for one or two plant species, or only censored values were available (see Paragraph “Censored endpoints”, p.17). The number of substances where an individual comparison between crop and wild plant species sensitivity was possible is roughly 50% of the figures above (see results section).

In addition to anonymizing names all ERx endpoints were normalized (based on the geometric mean of all plant endpoints of a particular compound). So instead of expressing endpoints as [g a.s./ha] data were normalized by dividing them through the overall geometric

[6] Many of the crop data come from confidential company studies which are data protected.

mean of all endpoints included for a particular active substance[7]. The anonymisation of the active substances and their corresponding endpoints was a fundamental requirement for two reasons: (1) companies have confidentiality issues if they display data that has been generated for post Annex-1 which has not yet received data protection and is considered a competitive advantage; (2) antitrust laws so as not to compare access to markets or the consequences to individual products. To be on the safe side only the methods and strategies are presented, but no specific actives or products.

Then endpoints were classified by type (measured variable, "x" of ERx, field or lab data etc., for criteria details see further down). At this step it became apparent that the data were very heterogeneous, e.g. for one given active substance only dicot data would be available, or only ER25 values based on shoot height, whereas for another one only ER50 endpoints based on biomass (fresh weight or dry weight) were available. Often endpoints of wild species would stem almost exclusively from field tests, whereas the majority of crop species endpoints would come from lab/greenhouse experiments. Furthermore it was noted that there were many cases of multiple testing, i.e. the same species had been tested repeatedly on a given active substance (by different authors, on different formulations, in different test designs etc.).

3.1.4 Combining several endpoints of the same species

Hence an option was implemented to combine multiple experimental endpoints of a given "species – endpoint test-substance combination" (e.g. several ER25 vegetative vigour biomass endpoints of the selected a.s. reported from lab tests on *Stellaria media*). In line with European requirements, the geometric mean of all these experimental endpoints was used. Obviously we did not combine different endpoint types such as ER25 and ER50, also shoot height or biomass data were kept separate. Field- and greenhouse data were not combined either at this step (otherwise their potential influence on variance could not have been assessed). If a given species had been tested both in the field and in standard lab test design, geometric means would be generated only within these two categories, ultimately resulting in a lab/greenhouse species endpoint and in a field test species endpoint for a given "species – test-substance combination", see Figure 1 as an example.

---

[7] The steps were as follows: The dataset of a substance to be displayed was retrieved by selecting criteria (e.g. substance X; lab and field studies; wild and crop species; monocots and dicots, only veg. vig. studies: only biomass experimental endpoints (wet–or dry weight), only numeric ER25 values). The resulting ER25 experimental endpoints (as active substance/test-type/species/endpoint combination) were checked for double entries (several tests of the same combination) and if there were such, the species' geometric mean endpoint was calculated. Of the resulting 'species endpoints' (one per active substance/test-type/species combination, see e.g. Figure 1 as an example) an overall geometric mean was calculated, by which all endpoints were divided (normalisation). So each selection results in a different overall geometric mean. For the calculation of quotients (crop/wild) this normalisation has no effect, the quotient crop species/wild species is the same before and after normalisation. In the charts (Figures 1, 2 and 11 ff.) the element that is affected by this normalisation is the scaling of the ordinate, the vertical grid line crossing the abscissa at '1' depicts the neutral point; data sets left of this are more sensitive and those right of it less sensitive than the geometric mean of the current selection. (If a geometric mean were calculated from the normalised endpoints (wild & crop) of any selection, it would be zero.)

Only experimental endpoints of the same measurement type were combined, i.e. shoot length with shoot length, survival with survival and biomass with biomass data. However, biomass data based on dry weight or on wet weight were combined for two reasons: (1) often only the measured parameter “biomass“ was recorded, but no details were given if these were based on fresh weight or on dry weight; (2) differentiating these would have left the evaluator with many different categories each with very few observations.

In several datasets with cases of multiple testing n is larger if individual experimental endpoints of a given crop-active substance combination are processed, than if geometric means per species (species’ endpoints) are processed. In the latter case n equals the number of species tested in a given test design (e.g. lab or field) and providing the chosen numeric endpoint (e.g. ER25 biomass).

The retrieval module of the database allowed to assess by species, i.e. based on the geometric means[8] here termed ‘species endpoints’, or based on the individual experimental endpoints (by experimental endpoint).

At this stage it was decided to focus on

- Vegetative-vigour-like studies, i.e. post-emergence exposure, foliar application
- Biomass data (fresh weight or dry weight)[9]
- ER25 and ER50 endpoints (only for these enough data were available both from lab and from field experiments); later we included also ER10.
- Distinguish between design options, either lab/greenhouse data (according to standard testing guidelines) or field test and multi species test data (details see section 3.2 Selection criteria).
- In case of lab tests to focus on vegetative vigour data. Several seedling emergence studies were included in the database and assessed in early versions, but were not included in the final assessment. Also there is only little seedling emergence data on wild plant species.

These selections were made to make the most out of the database in terms of endpoints that could be pooled together without adding too many potential confounding factors.

One instance of striking inconsistency in the reported and published data was identified at this stage. A typo in the final paper is the most likely explanation. In Egan et al. 2014 there was just one single data point, an ER25 of 0.4 g/ha that was a factor 50 lower than the next lowest experimental endpoint. Coincidently the very species had been tested by the same

[8] Alternatively also the median was used as central measure of each group. While in some cases it differs from the geometric mean by more than 20%, deviations are in both directions and the overall outcome proved to be same whatever central measure is assessed.
[9] Including some scoring systems that were considered to be based mainly on vegetative performance, with biomass the main parameter contributing to it, though biomass was not measured as such

authors twice in succeeding years, the endpoint obtained for that species in the first year indicated average sensitivity close to the overall geometric mean of all species (448 g a.s./ha). Also the former endpoint was by a factor of exactly 1120 higher than the low outlier (1120 is the conversion factor between lb/a and g/ha). It is suspected that there was just a typing error in the final table, possibly due to confused units (e.g. g a.s./ha or kg a.s./ha or lb/a). In this paper there were further cases where deviations between ER25-values obtained in the two succeeding years[10] exceeded two orders of magnitude. We excluded only the one value that clearly stood out (the deviation exceeded three orders of magnitude). However, any choice has only little effect on the overall medians and no effect on the minima.

Another problematic case was not as clear. The authors (Reuters &Siemoneit-Gast 2007) had compared endpoints of single-species tests and multi-species test designs. Their assessment dates were 2, 4 and 6 weeks post application in both test designs; the latter was considered to be an indicator for recovery of the plants. Standard vegetative vigour tests are however evaluated 3 weeks after application. In order to get test conditions as close as possible, here geometric means were calculated from the ERx-values calculated for the week 2 and week 4 assessment. However, there were single instances where no numeric endpoint was available at any of these two data points, but only smaller-than or greater-than values. In a few of these the evaluation after 6 weeks however yielded a numeric endpoint. In these cases the latter were included in the database, considering that a possibly slightly biased numeric value is still better than no value at all.

As a rule only vegetative vigour studies were evaluated (and all foliar applied field tests), as there were no corresponding field data available for seedling emergence glasshouse data. Therefore seedling emergence data were not used in the ultimate overall evaluation.

3.2 Selection criteria

3.2.1 Test substance:

As there may be pronounced differences between different formulations of the same active substance, the original aim was to compare formulation by formulation. The data proved however to be too heterogeneous for this approach. For some formulations a number of wild species experimental endpoints were available but only from field data, in other cases the formulation was not specified in the publication (this applies in particular to publications that did meta-analysis based on other data bases; here only the active substance was given). Hence ultimately all data of a given active substance were expressed as [g a.s./ha] and

[10] Egan et al 2014 obtained in the year 2011 overall lower endpoints than those obtained in 2010 (factor of 4) ; there were however also several cases where it was the other way round; e.g. for *Elymus hystrix* the ED25 obtained for Dicamba in 2010 was 11 g/ha, whereas in 2011 it was 311 g/ha (deviation by a factor of 28). Such inconsistencies indicating poor reproducibility appear to be common with testing NTPs, in particular when testing wild species in the field.

assessed together, in case of multiple experimental endpoints for one given species calculating the geometric mean.

3.2.2 Environmental conditions

The environmental conditions during testing may play an important role in the measurement values and can affect reproducibility of experimental endpoints, in particular if conditions are outside of the normal range required for the individual species or if the plants have not been properly maintained. Laboratory-greenhouse tests with controlled environmental conditions, many of them defined in the guidance documents are less problematic than the varying conditions in published field tests. For example in at least one case, plants grown in pots outside of the greenhouse (e.g. reported as semi-field) were too large for the pot size, did not receive adequate water, and suffered from desiccation before termination of the study. Many published datasets did not provide any details on test criteria, and those where details were given, they were not consistent (e.g. sometimes temperature and humidity was reported but not day length and/or irradiation).

3.2.3 Test design (Lab/field) – effects on sensitivity?

A fundamental parameter that was always reported was if tests had been performed in the lab/greenhouse or in the field. Therefore, data were filtered based on this parameter and – where the number of endpoints allowed - either only lab endpoints, only field endpoints[11] or all datasets were assessed. The majority of field test data were on wild species, whereas the majority of endpoints for crop species was obtained from lab data. For many test substances virtually all crop endpoints were lab and all but one or two wild species' endpoints from field tests. Accordingly, lab and field data had to be combined for comparison of wild and crop species.

There were 18 substances where pairs of endpoints of the same species were available both from lab and from field studies, four of these with both ER25 and ER50 endpoints, the others only one of the two, so a total n of 22 endpoint-substance combinations and a total n of 54 cases where matching endpoints were available both from lab and from field test. These were assessed separately for any differences between lab/greenhouse and field tests, again applying the quotient approach see 3.3.3. (Further details and results see 4.4.2) The overall result was calculated both as a standard geometric mean and as a weighted geometric mean. Here the number of common species (tested on a particular active substance both in the lab/greenhouse and in the field) was the basis of weighing (for details of weighing see 3.3.3; .numeric outcome see results section, point 4.4.2).

[11] The various test designs reported in the field were lumped for the final evaluation in just two categories, standard single species lab– or green-house tests and non-standard semi-field and field tests. The category "field tests" thus includes a few non-standard tests performed in the greenhouse but where artificial communities of several species had been tested. These multi-species assemblages – being non-standard-tests – were considered to be more similar to field tests than to the other greenhouse tests and therefore were evaluated with these.

3.2.4 Compliance with test guidelines and GLP

Another fundamental parameter generally available was if the test was performed under GLP. This was the rule for unpublished company data from the lab but hardly ever the case for published data, which generally reported field experiment results. Similarly compliance with test guidelines was generally the case with lab data provided by the companies, but obviously not with field test data (for which little guidance exists). Both parameters were included in the database, but neither were used as selection criteria. This is of course not ideal, as any guideline-compliant GLP data has to meet validity criteria, whereas for the non-GLP-field data no validity criteria are even defined. Often little is known of the actual performance of tested species that are considered in published papers.

Reliability of publications was evaluated in parallel to inclusion into the data base by assigning reliability indices (see e.g. Klimisch et al. 1997). Main points considered and implemented as a three-category-score were whether the test substance was unequivocally reported, whether the exposure route was described, whether information regarding dosing was sufficient (field rates, treatment levels), whether the observation time was indicated, whether the number of organisms / of replicates was reported, whether the test design was described in sufficient detail, and whether there were any inconsistencies that could not be explained. Further points that are relevant for regulatory studies were not considered here, as including these would inevitable have resulted in rejection of the majority of data: Aspects such as minimum germination rate, density, defined growing medium, reported fertilization, analytical verification, defined growth performance etc. were generally not considered. The resulting three-category-score system was simply ‘yes’, ‘partim’ or ‘no’. (1,2,3) and allowed an approximate grading of the reliability. However, some of the papers considered as relevant in the EFSA Scientific Opinion should have been regarded as less reliable based on our criteria, we lacked the information to assess it thoroughly (unpublished etc.) but it would have been inappropriate not to consider that very data which was the basis for EFSA’s conclusions.

Overall we included and considered all data that allowed the derivation of rate-based endpoints with some certainty, i.e. only highly subjective criteria such as fruit colour or marketability were not used, or data where the actual doses could not be translated into an area-based unit such as g/ha.

3.2.5 Systematic position of species

Some authors tested pairs of closely related species. Therefore we considered to compare closely related wild species and crop species also here. However, there were only few cases where closely related species were tested in the same test design i.e. both in the lab or both in the field. We did not deem it appropriate to compare endpoints in a paired design when the test designs were very different (i.e. glass house or field). Therefore results of paired test designs - performed by other authors - are cited and discussed individually; also these data were included in the database and contribute to our numeric evaluation, which however does not allow comparison at lower taxonomic level (other than what the authors did themselves).

However, data allowed assessing most compounds differentiating between monocotyledons and dicotyledons.

3.2.6 Traits of species

Tested species largely belonged to grasses or shrubs, i.e. annual, biennial or perennial herbaceous plants, both crops and wild species. Also a few tree species were tested. To utilize this information in addition to the species level, assessments were performed at higher levels, defining the plant family, the class (monocot/dicot) and crop or wild species as potential explanatory variables.

3.2.7 Endpoints

In this review we focussed on vegetative vigour endpoints, for apparent reasons there is only little data on effects on seedling emergence of wild plant species. Many wild species do not germinate readily but rather form a seed bank (as an adaptation to deal with environmental uncertainty (see e.g. Gardarin &Colbach 2014, Grime et al. 2007, Strassburger, E. 1998, Thompson, 1987, 2000, Rees 1994), so there would not be a defined control germination rate and tremendous noise in any seedling emergence data of these wild plant species. In contrast crop varieties have been selected over centuries to germinate instantaneously and readily; without this prerequisite a validity criterion such as ≥70% seedling emergence (OECD 208, 227), could normally not be met. This may be the main reason why published data testing wild plant species seem to focus exclusively on effects on vegetative vigour of young or mature plants, and sometimes on e.g. reproductive endpoints, usually in combination with foliar applications, but neither on any application on bare ground that would control germination (seedling emergence study type), nor on the seedling emergence endpoint.

According to European Data Requirements both study types (OECD 208, and 227, seedling emergence and vegetative vigour, respectively) are required, therefore any non-crop species proposed for testing must fulfil the germination requirement, or valid testing of that species including seedling emergence will hardly be possible.

3.2.8 Censored endpoints

A multitude of endpoints was listed as “greater than the highest test rate”. These are not strictly numeric but define a range that is only defined at one side. (“less-than” or “greater-than”). It is problematic to include them into numeric evaluations e.g. SSD, in particular when the censored values are not the lowest or highest, but are framed by higher or lower numeric values, details see discussion. However, it appeared that many of the censored values were not the only endpoint for a particular substance-species combination (i.e. often there was also a numeric endpoint for the same species), hence omitting censored values from the evaluation did not reduce the data base substantially. Therefore, while the censored values were included in the database and are listed as (n) in brackets, it was decided not to include them in any numeric evaluation; Details see discussion.

In contrast, in the additional statistical analysis performed by John W Green (Appendix 7) explicitly included censored values.

3.3 Principle of comparison

The fundamental approach in this paper was to compare like with like, i.e. endpoints selected based on the criteria described above with each other, and to calculate a quotient, which is the simplest approach to indicate quantitatively if the endpoints of a particular active substance differed between wild plant species and crop species. In this report the results based on biomass, ER10, ER25 and ER50 endpoints from vegetative vigour studies are presented, as this was the largest subset of data by a long way. John W Green considered in his additional statistical analysis (Appendix 7) initially also seedling emergence data.

To visualize the approach a comparatively small data set was generated (from data of two AASI – compounds, here termed AASI xx). We did not use any real dataset for this purpose for confidentiality reasons.

3.3.1 Assessment by experimental endpoint or by species' endpoint

In cases where one species was tested multiple times in the lab or in the field, the default was to use the geometric mean of all numeric experimental endpoints[12] of a species for the comparison. Lab and field data were however kept separate (see below) so if a particular species had been tested both in the lab and in the field, and if lab and field data are displayed together, two species' endpoints of that particular species will appear in the data set. In the example overleaf this is the case for *Silene nutans* and *Galium mollugo*. If in this example analysis was based just on lab data, the subset of wild species data would have two values less and the quotients would shift slightly.

In contrast if not assessed by species, but by experimental endpoint (i.e. any numeric endpoint is used as a data point), the overall "n" would be higher, but species tested twice would also be double accounted. This approach may still be useful if to compare the minimum values of each group, see example AASI 3 further down, and discussion.

3.3.2 Visualisation of sensitivities of crop species and wild plant species

These datasets may be visualized either by means of an SSD-chart, or by means of a box plot, see Figure 1 and Figure 2 for the comparison. The two charts display the same data;

[12] Of a given type, e.g. ER25 biomass endpoint from a veg.-vig.-like study. We did not merge e.g. ER25 with ER50 endpoints in one analysis, neither seedling emergence or vegetative vigour endpoints etc. Only in the revised statistical analysis one run was performed combining ER25 and ER50 ("pooled") data sets; however in this case the endpoint type (ER25 or ER50) was included as predictor, so its effect could be separated out.

which can be checked e.g. by comparing the positions of the lowest and the highest endpoints in Figure 1 with the error bars in Figure 2.

To compare the overall sensitivity of the two groups, either the lowest endpoints of the two groups are compared (i.e. the two most sensitive species), or a central value. The boxplots in Figure 2 display the median as the central line within the box, which as the middle score would also be an appropriate central measure for the data (which are approximately log-normal distributed, hence displayed on a log-scale). Furthermore the overall geometric mean was calculated for each data set, displayed in the box plot as diamond (rhombus sign). The geometric mean is another suitable central point for these data clouds, and in the European registration process the preferred way to combine several endpoints on a concentration scale (either from several tests on the same species, or from several species into one overall endpoint that serves as basis for the risk assessment (e.g. EFSA 2005, 2013).

Either of the two central measures of the data set would be suitable for the comparison exercise here, but here we are choosing the geometric mean as the central point of each data set, as it is closer to the approach applied in the European registration process.

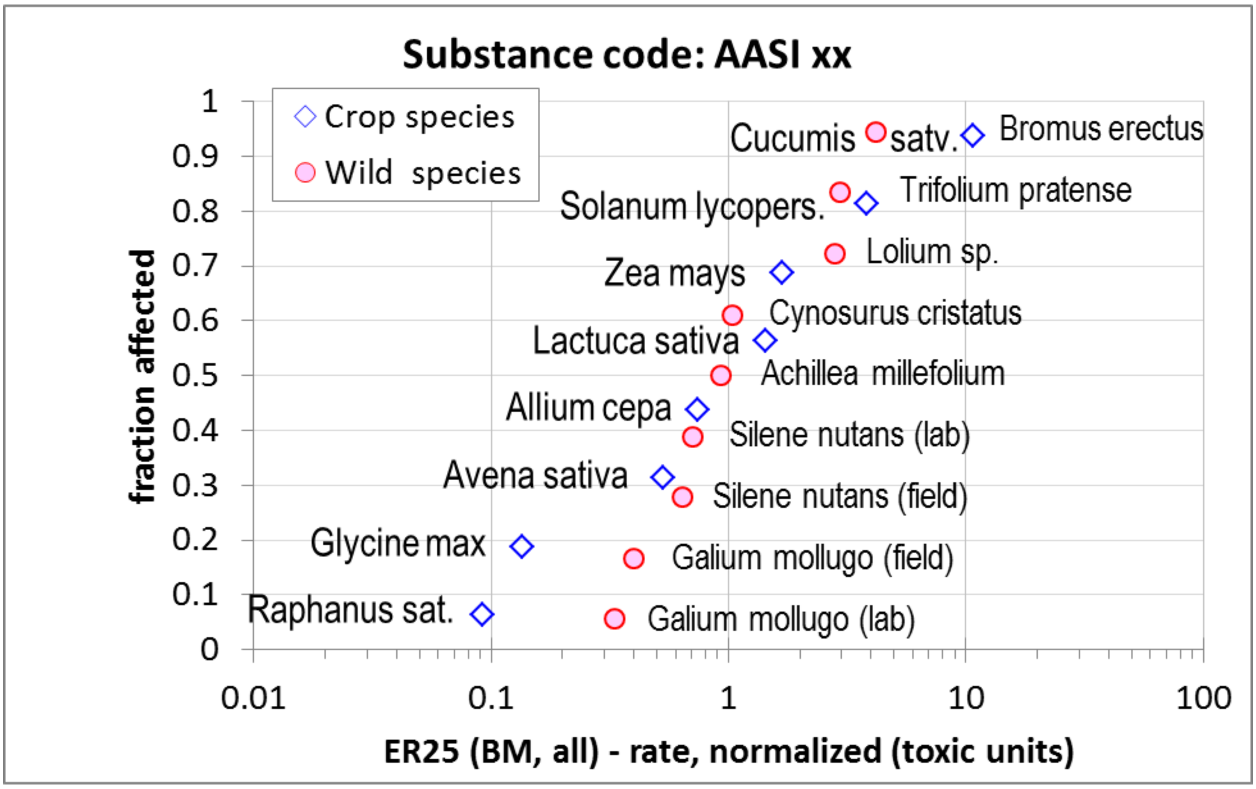


Figure 1 **Example plot visualising the approach from the individual species' endpoints via SSD. Distribution of ER25 endpoints (biomass; greenhouse and field endpoints combined) of wild plant and crop species. Abscissa indicating normalized field rates (endpoints divided by the geometric mean of all data). Species sorted by their sensitivity. Further explanations see text.**

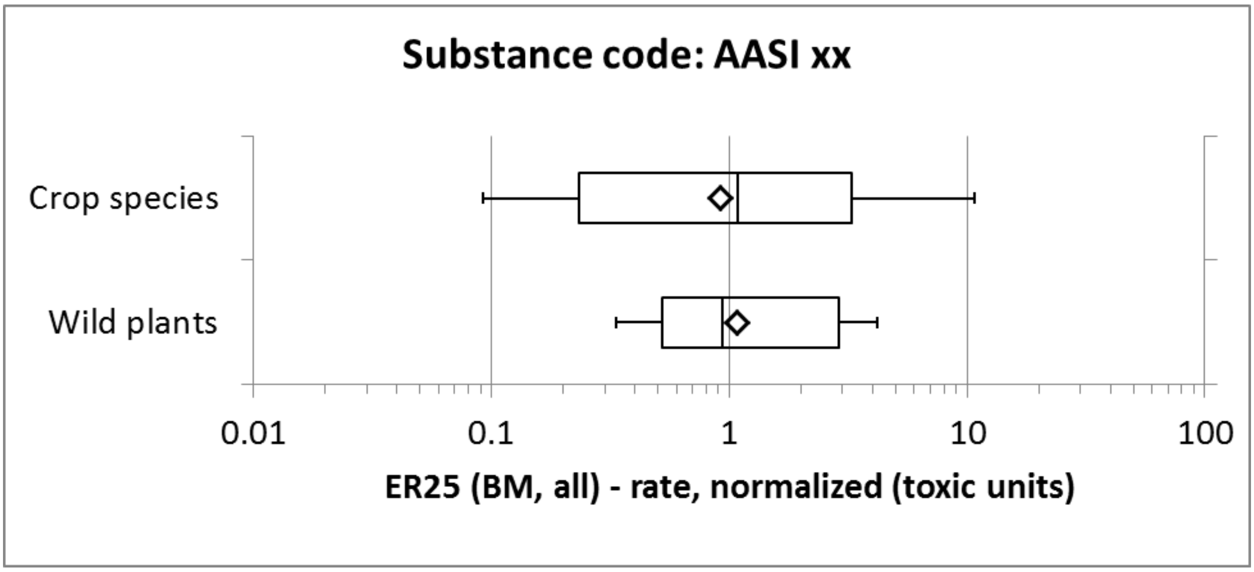


Figure 2: **Example plot visualising the approach from the individual species' endpoints via SSD to the box plots used throughout the paper. Distribution of ER25 endpoints (biomass) of wild plant and crop species. The rhombus marks the geometric mean of data points...**

3.3.3 Comparison of sensitivity using quotients

The quotients of the minimum pairs or the central pairs are displayed as quotients, always dividing the overall crop endpoint by the overall wild species endpoint (using either the lowest or central points of each group). In cases where crop species were more sensitive than wild species (i.e. crop endpoints lower than wild species' endpoints), the resulting quotient is less than one, if wild species were more sensitive than crop species the quotient is greater than one. In addition any quotient larger than 5 (wild species' endpoints >5 times lower than crop endpoints) or smaller than 0.2 (i.e. crop endpoints >5 times lower than wild species' endpoints) is printed in bold.

In the example overleaf wild species would appear to be slightly more sensitive only if based on the median values of the endpoints. Based on minima or on geometric means wild species appear to be less sensitive than crop species, see Table 2.

Table 2: AASI xx: Lowest endpoint and Geometric mean (alternatively median) of wild plant species and crop species, and resulting quotients.
<u>Quotients (crop/wild):</u> Values < 1 indicate that crop endpoints were lower, i.e. crops more sensitive than wild species, and *quotients > 1 that wild plant species' endpoints were lower*, i.e. wild species more sensitive than crop species. Any quotient smaller than '0.2' or greater than '5' is printed in bold.

Group		N	Minimum	Quotient	GeoMean	Quotient	Median	Quotient
All (monocots + dicots)	Crop	8 (species)	0.092	0.28	0.916	0.85	1.090	*1.17*
	Wild	9 (species)	0.333		1.081		0.931	

These calculations were applied to all datasets where each ≥ 3 endpoints of wild plant and crop species where available for a given data selection, i.e. endpoint - active substance test type combination. See also Figures 11 to 56. In addition, where data allowed also the systematic subgroups 'monocots' and 'dicots' (s.l.) were differentiated within each group (wild plants or crop plants).

In Appendix 6 detailed data are presented for each anonymized active substance. Some seedling emergence studies were also available, and these also provide biomass – and sometimes shoot height – endpoints. As discussed previously, the default selection was only to use endpoints from vegetative vigour studies due to the paucity of data on seedling emergence with wild plant species, and with the absence of data from field tests.

Quotients were calculated either based on lowest endpoints = most sensitive species (i.e. dividing the lowest crop species' endpoint by the lowest wild species' endpoint), and based on average sensitivity (i.e. dividing the geometric mean of all crop species' endpoints by the geometric mean of all wild species' endpoints).

Further down an overall table with all quotients is presented (basically combining the individual tables of the Section 'Tables and Figures' in Appendix 6, p. 120); based on biomass data (vegetative vigour studies or field study endpoints) as this is the most comprehensive dataset. Comparisons based on ER10, ER25 and ER50 endpoints were combined, see Chapter 4.4.1.

From these quotients an overall quotient was calculated. Two approaches were applied:

a) overall standard average (i.e. geometric mean of all quotients) - no weighting

b) overall weighted geometric mean quotient.

In the standard average approach any quotient of any substance-endpoint combination bears the same weight, disregarding that some quotients are based on more data points than others, results see penultimate row in Table 3. Calculation as a weighted average (geometric mean) considers the number of endpoints entering the calculation of a quotient. The weight of each individual active substance is based on the lower of the two ‘n’ of the two groups wild plants / crop species) that were compared (‘n’ labelled ‘lower n”), i.e. if 15 wild species endpoints but only 6 crop species endpoints were available, ‘n’ used in weighting is 6. The weighted geometric mean was thus calculated as the product of all true weighted quotients (each of which was raised to the power the total ‘n’ (sum of the two comparison groups). The overall product then was raised to the power of the reciprocal of the sum of all ‘n’s (again the lower number of species of the two groups that were compared)[13]. Results are presented in the last row of Table 3.

[13] Formulas for the weighted geometric mean (first variant was applied)

$$\bar{x} = \left(\prod_{i=1}^{n} x_i^{w_i}\right)^{1/\sum_{i=1}^{n} w_i} = \exp\left(\frac{\sum_{i=1}^{n} w_i \ln x_i}{\sum_{i=1}^{n} w_i}\right)$$ with

$x_{i\ (1\text{-}n)}$ the values and $w_{i\ (1\text{-}n)}$ the weights (here: lower number of species per groups for the individual a.s.). (Wikipedia 2014), http://en.wikipedia.org/wiki/Weighted_geometric_mean

3.4 Auxiliary analyses

Supplementary evaluation of the available data aimed to verify (or not verify) whether there was a fundamental difference between quotients based on ER10, ER25 or based on ER50 endpoints (always comparing like with like), and the hypotheses that endpoints from laboratory/greenhouse are generally different (higher or lower) than field-test endpoints.

Active substances with data for all four combinations, i.e. ‘wild ER25, and crop ER25; and wild ER50, and crop ER50, respectively, were collated, and quotients (dividing the crop species’ minimum by the wild species minimum, or the crop central point by the wild central point, respectively. As the data sets for ER25 and ER50 consisted of different species, there is considerable additional scatter of data. Still the resulting quotients are listed in a table and plotted. In addition the multiple regression analysis allowed to extract information regarding the representativeness of ER25, ER50, and pooled assessment of both together, respectively.

In addition we checked whether outcomes from lab and field studies were fundamentally different. Plant tests were performed either in the lab/greenhouse or in the field. There is contradicting information as to whether lab-greenhouse test systems or field test systems generate the lower endpoints. In this evaluation results of both test systems had to be assessed together, as otherwise a large proportion of available data could not have been utilized. Initially it was only considered that any underlying differences in sensitivity between greenhouse and field test would affect both to wild and to crop species and therefore – while adding noise to the data set – overall cancel each other out. In the course of the work it became however apparent that a) the available data was not equally distributed, but there were more wild plant endpoints from field tests than from lab tests, and b) the available data seemed suitable also to assess numerically if there was any trend in terms of differences in sensitivity between lab and field tests.

The database was therefore searched for pairs of endpoints tested both in the lab and in the field. To rule out any bias from differences between species, only those cases were included where numeric ER10, ER25 or ER50 endpoints based on biomass measurement after foliar applications (vegetative-vigour study design) were available. Of these data pairs again quotients were calculated and presented in a table, and overall average quotients calculated, details for the weighting procedure see previous page. This assessment was not repeated with ER10 data as further information was available with other approaches i.e. the regression approach (see next chapter) the additional statistical analysis comparing distributions including censored values, see Appendix 7, and a comparison of lab and field endpoints again by means of a paired quotient approach.

3.5 Multiple Regression

In addition to the quotient approach described above, an ANOVA analysis was performed based on original (non-normalized, but log(e)-transformed) data, again aiming to falsify the null-hypothesis that there are no significant differences in sensitivity to plant protection products between wild plant species and crop species.

Furthermore a factorial model was applied on the log-transformed biomass-data.

Crop species and wild species categories were treated as a factor in a linear regression model where the endpoints (ER25 or ER50 in the log scale) were modelled as a function of several factors including CW.fin, MoA, Lab/Field, ClassM/D (explanation see further down). The results were summarized in Table 1.

Starting with only ER25 data (model1, “ER25”), first the full model was fitted including all the interaction terms between crop and wild species (CW), mode of action (MoA), laboratory/glasshouse versus field test (Lab-Field), taxonomic group, i.e. monocot or dicot s.l. (ClassM.D). Second a best model with smallest Akaike information criterion (AIC) was selected using a stepwise procedure. As no justification for any biologically defined interaction terms was available, this purely mathematically optimized full model must be interpreted with caution. The standard model (aka reduced model) with no interaction terms is therefore considered to be the biologically more meaningful option, and the overall evaluation was based on the standard model without interactions.

Next a power analysis was applied, calculating MDDs (minimum detectable difference) for different combinations of predictors. The result (expressed as percent) indicates which level of difference could be detected as being significantly different. However, if the MDD calculation is based on each combination of factor predictors (as it is usually calculated), then the limited sample size and large variance in this specific combination would result in some very high MDD for several specific combinations, and we will not actually assess the data’s general difference but only the difference of specific sub-category, i.e. specific combinations of predictors. The outcome of this analysis was therefore inconclusive, which was considered to be a fundamental methodical problem. For our goal the ‘wild group’ and ‘crop group’ data need to be treated as a whole to perceive how a consistent negative difference would look like (see discussion).

Therefore, a different approach was finally applied, in which artificially modified (“manipulated”) data sets were submitted to the same factorial analysis as the original data. The factorial modification was that all crop endpoints were rescaled (by multiplying them by 1.5 or 2) whereas the wild endpoints remained unchanged. The concept behind this manipulation is that it should reveal if any consistent difference between crop and wild species would have been detectable with data exactly as heterogeneous, clustered and variable as the original set of data analysed here. The factorial manipulation of the original data (multiplication by a factor of 2) equals the addition of 0.69 of the natural log-transformed data points. In other words, a factorial coefficient of 0.69 determined for a particular predictor translates into a deviation by a factor of 2 on the original scale.

While this factorial manipulation changes the overall variance of the data, the variance and standard deviation of the log-transformed data (that were used for the factorial analysis) within each group is not affected by the manipulation. The approach should therefore give a realistic indication whether the observed heterogeneity and variability of data would have prevented detecting any differences of this magnitude.

3.6 Additional statistical analysis (John W. Green)

In addition distributions were compared applying MLE-methods including censored endpoints to relevant subsets of data. in this process individual censored ERx values are replaced by adjusted values based on the distribution obtained from MLE methods. Conceptually, a right-censored value was replaced by the 90th percentile of the fitted distribution and a left-censored value was replaced by the 10th percentile of the distribution.

The resulting adjusted datasets underwent ANOVA analysis. Detail see Appendix 7.

3.7 Abbreviations frequently used

AI-factor Proportion of variance capturing the influence of the 'Active Ingredient' (AI) on the endpoints (two-way ANOVA)

Average Arithmetic mean

BM Biomass, a measured variable (weight of plant material above ground) used to calculate ERx endpoints, either based on wet weight = fresh weight (WW) or dry weight (DW)

D Dicotyledonous species (sensu lato)[14]

DAR Draft assessment report

DW dry weight (above ground)

ECPA European Crop Protection Association

EFSA European Food Safety Authority

ERx Effect rate, i.e. a treatment rate (often expressed in [g a.s./ha]) at which a certain degree x of inhibition was observed, e.g. ER25, ER50. Some papers also list EC50 (mostly incorrectly as the unit is still a rate) or IC50 (Inhibition concentration) which – considering the rate unit – are again incorrectly used and here interpreted as synonyms.

GeoMean Geometric mean, a central measure of a sample that is better for non-linear distributions than the arithmetic mean

[14] The term 'dicotyledons' stands for a paraphyletic group; so effectively covers eudicots or tricolpates and magnoliids (the latter were however not present in the database, so basically 'D' i.e. 'dicots' means 'eudicots' in this report)

Median	Central point of a distribution, with 50% of values above and 50% below, another central estimate that is more robust against skewed distributions than the arithmetic mean
Max.	Maximum, here the highest endpoint of the selected groups (e.g. monocotyledonous crop species), not used for overall assessment
Min.	Minimum, here the lowest endpoint of the selected groups (e.g. monocotyledonous crop species
M	Monocotyledonous species
SE	Seedling emergence (e.g. studies acc. to OECD 208), application of soil prior to germination and emergence of plant seedlings
SH	Shoot height, a measured variable used to calculate ERx endpoints, sometimes also termed shoot length
Surv.	Survival of plants, measured variable also used to calculate ERx endpoints
USEPA	United States Environmental Protection Agency
VV	Vegetative vigour (e.g. studies acc. to OECD 227) usually folia application
WW	wet weight = fresh weight (above ground)

4 Results

4.1 General outcome and initial considerations

The database consists of 2873 data sets, i.e. substance-species-test-design combinations. Many of these contain just one endpoint (e.g. an ER25), others several, often generated within the same study. However, only a subset of these consisted of numeric endpoints; ca 30.0 % of endpoints were censored[15] (discussion see paragraph “Censored endpoints”, p.17). A total of 525 ER10, 1760 ER25 and 2062 ER50 values were available, but only 308, 1152 and 1286 of them were numeric, respectively[16]. Furthermore there were many cases where the same species had been tested twice or even more frequently. For the assessment by species all matching endpoints of a given species were combined into an overall geometric mean per species (see material and methods). This resulted in 139 ER10, 584 ER25 and 632 ER50 numeric species’ endpoints (by species, endpoint and test system). Including censored endpoints there were 183 ER10, 825 ER25 and 887 ER50 species’ endpoints (by species, endpoint and test system). Most of them were based on measurement of biomass in vegetative vigour studies or field studies comparable to vegetative vigour studies[17] (foliar application of fully emerged plants at generally early growth stages – in the field growth stages were more variable though).

For the final evaluation the largest homogenous data set was used; endpoints from vegetative vigour studies, ER10, ER25 or ER50 (only comparing like-with-like), but combining lab/greenhouse single species and field studies/multi species data, based on the measured variable “biomass” (but not differentiating between wet weight and dry weight). This resulted in a total of 504 crop (84 ER10, 168 ER25 and 252 ER50) and 747 wild species endpoints (96 ER10, 275 ER25 and 376 ER50 endpoints respectively).

4.1.1 Selection of data sets

Initially we considered that generally a minimum of three species with numeric endpoints for either crop or wild species should be available to attempt an individual comparison of wild species’ and crop species’ sensitivity[18]. In addition a weighted geometric mean approach was applied, in which all data sets irrespective of the number of numeric endpoints could be considered. The active substances that could be assessed numerically at least with one subset of endpoint (e.g. ER25) and corresponding quotients are listed in Table 3. In the

[15] 30.4% crop endpoints and 29.6% wild endpoints.
[16] Of all 2541 experimental endpoints, 30.0% were censored; 30.4% crop endpoints and 29.6% wild endpoints. So if ignoring censored endpoints should have introduced any bias, it would be introduced to both groups similarly.
[17] including some scoring systems considered to reflect vegetative vigour
[18] Only for the additional evaluation on similarity or dissimilarity between quotients based on ER25 or on ER50 endpoints (in Chapter 4.4.1) we deviated from this principle. To have some additional cases where ER25 and ER50 values of the same test substance could be compared, we included also datasets were one of the four data groups had less than 3 numeric endpoints. See in Table 6.

overall assessments (multiple regression (Appendix 6) or comparison of MLE-distributions, (Appendix 6), all data are included, irrespective of their n.

4.2 Summary of results

A total of 2872 data sets were entered (species – test combinations) with 525 ER10, 1760 ER25 and 2062 ER50 endpoints[19], plus 77 ER05, 87 ER75 and 50 ER90 endpoints. Considering only numeric endpoints and merging multiple endpoints of equivalent substance-species-test system combinations (multiple testing of same species) resulted in 139 ER10, 584 ER25 and 632 ER50 numeric species' endpoints.

The data derived as summarized above and detailed in Appendix 6, is combined below into one overall table, and the resulting quotients (only of total assessments, no differentiation between monocots and dicots, which are detailed in Appendix 6) are assessed for any consistent difference between wild plants and crops (quotients above or below 1), and by which value. All data sets are included that allowed a comparison of vegetative vigour biomass[20] data like-with like. There were borderline cases with low 'n' that are not displayed individually in Appendix 6. However, the evaluation weighted by the number of species is not significantly affected by the decision whether or not to include minor data sets, as the overall outcome is dominated by the data sets for which many endpoints exist. Therefore ultimately all datasets based on vegetative vigour biomass-like measurements[21] that allowed calculation of quotients were included in the overall assessment overleaf.

Quotients were calculated (a) based on the most sensitive species of each group and (b) on the average sensitivity of each group, always dividing the crop species endpoint by the wild species endpoint. Established on the most sensitive species, 30 quotients were above 1 indicating that wild plant species were more sensitive than crop species (the endpoint of the latter is higher); and 26 quotients below 1 indicating that the crops were more sensitive than the tested wild species. Based on the average sensitivity (geometric means) also 30 quotients were above 1, and 26 quotients below 1. Based on these quotients overall estimates were calculated, as a standard average (i.e. geometric mean) - no weighting - and weighted average quotients were calculated from all quotients based on most sensitive species, and from all quotients based on average sensitivity. In the standard average approach any quotient of any substance-endpoint combination bears the same weight, disregarding that some quotients are based on more data points than others, results see penultimate row in the table overleaf. Calculation as a weighted average (geometric mean) considers the number of endpoints entering the calculation of a quotient, see materials and methods.

[19] disregarding endpoints given an a not-area-related unit (µmol),these could not be converted.
[20] Biomass as pooled variables (fw or dw)
[21] including some scoring systems considered to reflect vegetative vigour in a way that allowed derivation of ERx- estimates.

Table 3: Summary of all sets where a quotient could be calculated considering ER10, ER25 and ER50 endpoints from VV studies / field studies based on biomass data. Quotients calculated from lowest endpoints (minima), and from the geometric means. A quotient x greater than 1 indicates that wild species were more sensitive (by factor x) than crop species, quotients below 1 indicate the opposite. Quotients above 5 or below 0.2 are printed in bold, those above 5 (indicating that wild species were more sensitive than crops) are underlined. Resulting overall quotients between crop species and wild plant species are shown at the bottom; (a) as overall geometric mean of all quotients (not weighted), (b), as weighted geometric mean (weighting based on the lower 'n' of each pair), (c) medians of all quotients.

Code Substance-effect level	n crop	n wild	n total	Quotients (x) based on minimum	geomean	P-value (ln())*	more sensitive
AASI01 (ER25)	8	8	16	0.712	3.08	p = 0.161	
AASI01 (ER50)	7	9	16	0.903	4.00	**p = 0.048**	wild
AASI02 (ER25)	1	5	6	<u>29.7</u>	1.97	n.d.	
AASI02 (ER50)	10	4	14	0.204	0.196	p = 0.132	
AASI03 (ER10)	13	9	22	1.22	1.29	p = 0.727	
AASI03 (ER25)	16	85	101	1.01	1.22	p = 0.492	
AASI03 (ER50)	14	62	76	3.09	3.03	**p = 0.0002**	wild
AASI04 (ER25)	2	15	17	<u>82.8</u>	<u>14.1</u>	n.d.	
AASI04 (ER50)	14	10	24	2.06	0.99	p = 0.9892	
AASI05 (ER25)	12	9	21	0.0511	0.233	**p = 0.0334**	(crop…)
AASI06 (ER25)	6	7	13	0.867	0.566	p = 0.2782	
AASI07 (ER25)	9	8	17	0.016	0.557	p = 0.3997	
AASI14 (ER10)	2	1	3	0.862	0.885	n.d.	
AASI14 (ER25)	3	1	4	0.347	0.606	n.d.	
AASI14 (ER50)	4	2	6	0.11	0.465	n.d.	
AASI15 (ER25)	2	3	5	1.07	0.555	n.d.	
AASI17 (ER10)	10	4	14	4.1	1.73	p = 0.7327	
AASI17 (ER25)	10	1	11	0.0116	0.0858	n.d.	
AASI17 (ER50)	11	26	37	<u>8.57</u>	<u>5.56</u>	**p = 0.0004**	**wild**
ACI2 (ER50)	1	9	10	4.5	0.634	n.d.	
ACI3 (ER50)	2	1	3	0.457	1.53	n.d.	
CMD01 (ER50)	9	29	38	1.15	0.807	p = 0.5447	
GW01 (ER10)	8	2	10	0.0216	0.208	n.d.	
GW01 (ER25)	10	17	27	0.353	0.836	p = 0.8261	
GW01 (ER50)	12	27	39	3.14	<u>5.83</u>	**p = 0.0265**	**wild**
GW03 (ER10)	1	11	12	1.96	0.69	n.d.	
GW03 (ER25)	1	11	12	2.97	1.09	n.d.	
GW03 (ER50)	7	13	20	1.31	1.3	p = 0.3602	
GW05 (ER10)	1	7	8	<u>43.0</u>	1.15	n.d.	
GW05 (ER50)	17	8	25	0.239	1.37	p = 0.4191	
GW06 (ER25)	11	14	25	0.205	0.309	p = 0.1833	
GW07 (ER50)	2	1	3	0.06	0.363	n.d.	

Code Substance-effect level	n crop	n wild	n total	Quotients (x) based on		P-value (ln())*	more sensitive
				minimum	geomean		
GW09 (ER25)	9	16	25	4.73	1.77	p = 0.2843	
GW10 (ER25)	7	3	10	4.3	1.68	p = 0.7837	
GW10 (ER50)	7	14	21	2.49	1.2	p = 0.7545	
GW11 (ER25)	8	7	15	5.03	0.79	p = 0.8154	
GW13 mix (ER50)	12	8	20	0.398	0.276	p = 0.2071	
LSI1 (ER25)	2	2	4	1.22	1.1	n.d.	
LSI1 (ER50)	8	6	14	0.8	1.45	p = 0.3835	
LSI2 (ER50)	5	3	8	0.104	0.191	p = 0.1298	
OTH06 (ER25)	10	14	24	2.75	6.65	**p = 0.0033**	**wild**
PHI01 (ER25)	11	2	13	0.153	0.243	n.d.	
PHI01 (ER50)	10	3	13	0.172	0.234	p = 0.2076	
PHI02 (ER10)	9	1	10	0.0006	0.0462	n.d.	
PHI02 (ER25)	15	38	53	1.49	1.31	p = 0.5719	
PHI02 (ER50)	7	14	21	0.489	0.572	p = 0.538	
PHI03 (ER50)	6	15	21	2.74	2.45	**p = 0.0158**	wild
PHI04 (ER50)	8	31	39	0.0881	0.23	**p = 0.041**	(crop…)
PHI07 (ER25)	5	6	11	0.979	1.21	p = 0.8292	
PHI09 (ER50)	3	2	5	0.0333	0.314	n.d.	
SGI2 (ER10)	9	1	10	2.08	30.1	n.d.	
SGI2 (ER25)	9	1	10	2.75	35.8	n.d.	
SGI2 (ER50)	9	16	25	301	20.2	**p = 0.0004**	**wild**
SGI3 (ER25)	13	13	26	1.28	1.69	p = 0.1326	
SGI3 (ER50)	1	2	3	3.33	2.54	n.d.	
SGI4 (ER50)	2	2	4	3.85	7.86	n.d.	
Overall (numeric) n	421	639		minimum	geomean		
Average quotients				0.849	1.097		
Weighted average quotients				1.092	1.292		
Median				1.110	1.125		

* t-test comparing logarithmic (ln)-transformed endpoints of wild and crop species, two-sided, p = 0.05

Based on this subset of dataset (only substances with both crop and wild plant endpoints, no seedling emergence endpoints, no shoot height or survival-based endpoints etc., including field and lab data, only numeric endpoints), all calculations resulted in figures very close to 1 (ranging from 0.85 to 1.29), indicating that overall there was no apparent difference in sensitivity between wild plant species and crop species.

In addition the ln-transformed endpoints of crops and of wild species were assessed individually for any significant differences (two-sided, p = 0.05). There were 34 cases (out of 56) where at least three species with numeric endpoints were available per group. Based on two-sided t-test based on log-endpoints, only in 9 of 34 cases significant differences between crops and wild species were detected. Of these, wild species were more sensitive than crops in 7 cases and crop species were more sensitive than wild species in 2 cases. Based on these vegetative biomass data the overall quotient (geometric mean of the individual quotients) based on minima was 0.85 and the one based on the geometric mean was 1.10 (n = 56, Figure 3 A)

If also seedling emergence data are included, and other endpoints (such as those based on shoot height or survival etc.), there are a total of 64 cases, and the overall quotient based on minima was 0.96 and the one based on the geometric mean was 1.16 (n = 64, excluding censored endpoints) (details not presented in report).

Including also censored endpoints with a correction factor f = 2 increases the number of cases to 76, (Figure 3 B). The resulting overall quotient based on minima was 0.97 and the one based on the geometric mean was 0.87 (n = 76) (details not presented in report).

The figures display quotients plotted on a log-scale; sorted by the quotients magnitude. The individual cases (combinations of substances, and endpoints) were kept together, so the distance between the two data rows (triangles and rhombi) indicate how the two quotients calculated per data set[22] do scatter (Figures 3 A and B).

[22] quotient either based on the central values of the two groups, or on the minima of the groups (most sensitive species)

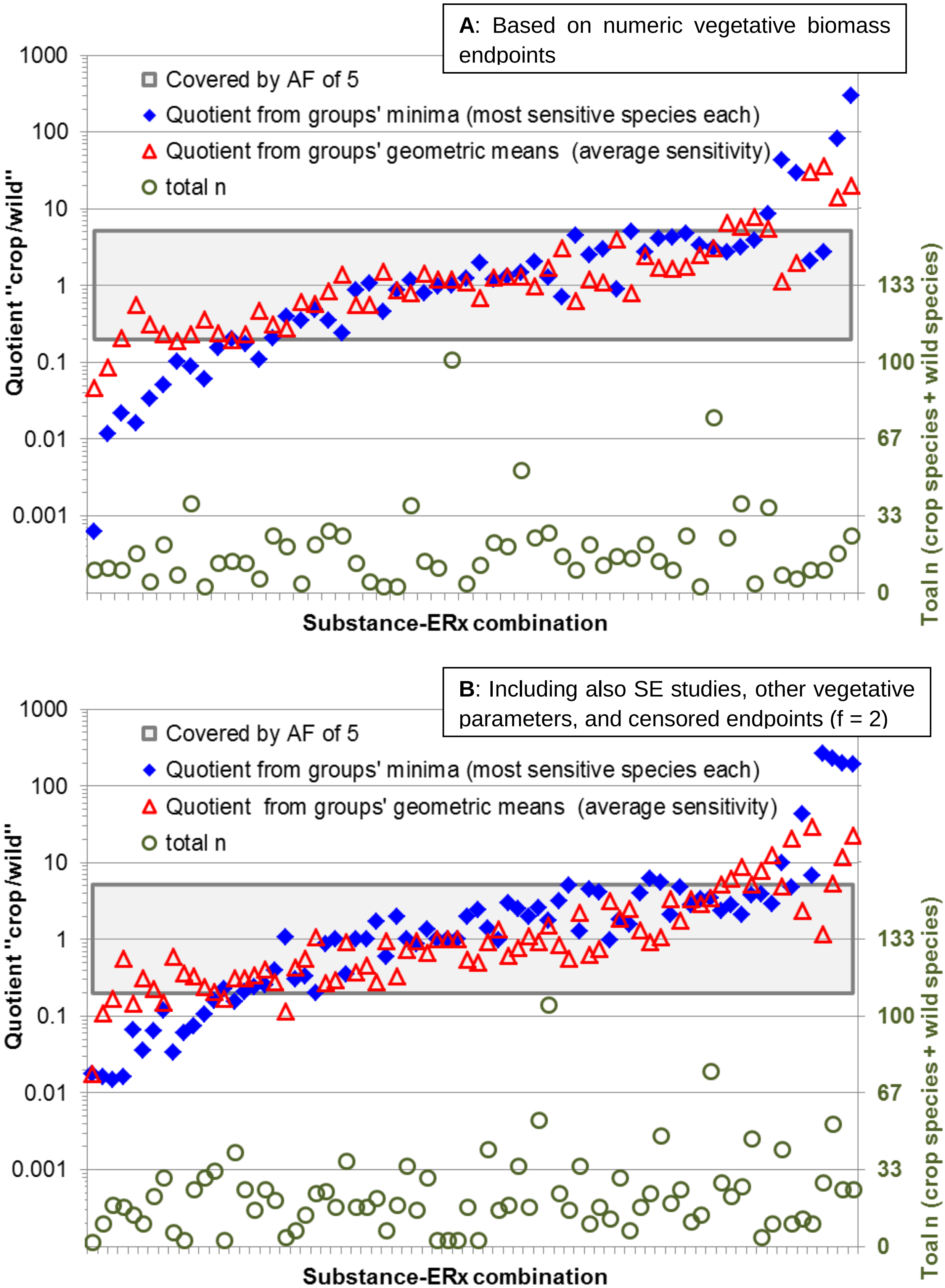


Figure 3: **Upper figure A based on numeric vegetative biomass endpoints, lower figure B also including seedling emergence studies, other parameters such as shoot**

height or survival, and considered censored values with a correction factor of 2. Quotients (triangles and rhombi) of the individual cases (combinations of substances and endpoints) on a log scale, sorted by average quotients (of minima- and geometric means-quotients) in ascending order. Cases above 1 indicate that wild species were more sensitive than crop species, cases below 1 that crop species were more sensitive; either based on their central value (geometric mean) = triangles, or on the groups' minima (comparison of the most sensitive species of each group) = rhombi.. The green circles at the bottom indicate the number of species (lower n of the compared groups, secondary ordinate).

In these charts it is apparent that there are as many low quotients (below 1) as above 1, confirming the results presented in the Table 3.

The circles at the bottom indicate the number of species (lower n of the compared groups). The trend-line illustrates that the datasets at the far ends – where large differences between wild and crop species sensitivity were observed –generally were the datasets with a lower n and hence lower reliability.

In addition it is apparent that quotients based on the minima i.e. the two most sensitive species of each dataset (blue rhombi) show larger deviations from the central value (which is close to one) than the quotients based on the average sensitivity of the groups (red triangles). Based on the initial assessment (considering only numeric endpoints, Figure 3A) 38 (of 56) quotients based on minima were within the rectangle indicating the area covered by an assessment factor of 5, but 44 (of 56) quotients based on average sensitivities. 6 quotients based on minima and 8 quotients based on average sensitivity were greater than 5; 12 quotients based on minima but only 4 of those based on average sensitivity were smaller than 0.2., i.e. outside the area covered by an assessment factor of 5.

The overall outcome is thus a strong support that by and large there is no consistent difference in sensitivity between crop and wild plant species (at least for the systematic groups for which data were available), <u>if matching endpoints are assessed i.e. only like with like is compared.</u>

4.3 Multiple regression analysis

4.3.1 Results of the different methods applied

In addition to the quotient approach described further up, a multiple regression analysis was performed.

To assess the influence of confounding parameters a factorial model was fitted including further categorical predictors, i.e. Mode of action, Lab or field test, monocot or dicot, and (in case of pooled analysis) ER10, ER25 or ER50, details see Material and Methods. The models were run four times, 1. with just ER10, 2. with ER25, 3. with ER50, and 4. pooling all three sets of data (while introducing the kind of endpoint 'ERx' as another categorical predictor, results see bottom of Table 4. Two models were run: a standard, i.e. reduced (no interactions) and a full model (i.e. iterative optimized model, defining interactions on a purely numerical basis), details see materials and methods and Appendix 6, p. 168. Many of the interactions defined by the latter were not supported by any data sets though (i.e. there were no tests with the corresponding combinations of predictors). The model with interactions is assuming there is a different relationship between crop and wild in each MoA, L/F, Class. As the numerically optimized 'best' model indicated different relationships for different MoA, this was considered to be another indicator against the hypothesis of a fundamental and ubiquitous difference in sensitivity between crop and wild plant species. The assumption made was that there should be a common trend and no interactions. Hence we focussed on the standard model, results of which are presented here; summarized in Table 4.

Based on the **ER10** endpoints, the standard model found wild species to have on average significantly *lower* ER10 values (factorial coefficient = -0.37, which equals a factor of 0.69 on the original scale[23] compared to the crop species (negative coefficient, $p < 0.05$). Different modes of action had different ER10. and all had significantly higher ER10 values than the baseline, which is AASI, among which particularly potent herbicides are found (i.e. herbicides with low efficient field rates and hence also particularly low endpoints). Class M (monocots) overall had higher ER10 than class D (dicots), and lab endpoints were slightly higher than the field endpoints but not significantly when based on the standard model (no interaction). Based on the 'best' mode with interactions the canonical coefficient of 2.32 indicated that wild endpoints were significantly *higher* than crop endpoints, also endpoints from lab studies were significantly higher than those of field studies, but the outcomes based on models with interactions must be interpreted with care, as the interactions may be just mathematical artefacts.

Based on the **ER25** endpoints, the standard model found wild species to have on average significantly *higher* ER25 values (factorial coefficient = +0.65, which equals a factor of 1.92 on the original scale[24] compared to the crop species (positive quotient; $p < 0.05$), and again highly significant based on the model with interactions (Best model). Here no significant differences were detected between lab and field data, while class M (monocots) overall had higher ER25 than class D (dicots). All modes of action had significantly higher ER25 than the baseline, AASI; see previous paragraph. Based on the 'best' mode with interactions the canonical coefficient of 2.242 indicated again that wild endpoints were significantly *higher* than crop ER25 endpoints. Lab endpoints were slightly lower but not significantly different from those of field tests.

Based on the **ER50** data and the standard model (no interactions) the canonical coefficient was -0.01 equal to a factor of 0.99 on the original scale, so crop and wild endpoints did not differ. Again all modes of action had significantly higher ER50 than the baseline, AASI, lab endpoints were lower than field endpoints when based on the standard model, but higher

[23] Back-transformed from the natural logarithm of -0.37 i.e. wild ER10 were on average a factor of 0.69 the crop ER10, or 1.45 times lower than crop endpoints
[24] Back-transformed from the natural logarithm of 0.65 i.e. wild ER25 were on average 1.92 times higher than crop specie endpoints.

when based on the model with interactions, and also the outcome comparing monocots and dicots was inconsistent. Based on the 'best' mode with interactions the canonical coefficient of 1.25 indicated again that wild endpoints were again *higher* than crop ER50 endpoints, but not significantly.

As differences between crop species and wild species were identified to be significantly positive for ER25 but significantly negative for ER10 and ER50, a further approach was added, in which ER10, ER25 and ER50 data were **pooled**, assuming a common difference between crop species and wild species irrespective of the reported endpoint. In this approach the effect level x (of the ERx) was considered as another predictor.

Both **pooled** models (standard and with interaction) found endpoints of wild species to be significantly *higher* than those of the crop species, i.e. less sensitive than the latter. Again all modes of action had significantly higher endpoints than the baseline, Class M and Gymn had significantly higher endpoints compared to class D; lab and field endpoints were not significantly different based on the standard model. Also ER25 and ER50 endpoints were significantly different from the baseline ER10; ER25 being somewhat higher and ER50 distinctly higher than ER10, as to be expected. Details see Table 4.

Table 4: Summary of fitted factorial models (no interaction) to ER10, ER25, ER50, and pooled data. Positive coefficient signs indicate that the predictor at the right end of the predictor code was on average higher, negative quotient signs that it was lower. E.g. 'CW.finW': ER10, '-0.37': From the groups C (crops) and W (wild) endpoints of the latter (W) were lower than the former, i.e. wild species were more sensitive than crop species, whereas based on ER25 it was the other way round (coefficient +0.65) wild plant endpoints higher than crop endpoints. Standard errors of the coefficient estimation inside parenthesis.

Predictor code	ER10	ER25	ER50	Pooled
(Intercept)	0.45 (0.5)	2.00 (0.23)***	2.25 (0.22)***	1.15 (0.19)***
CvWW	-0.37 (0.43)	0.65 (0.15)***	-0.01 (0.14)	0.32 (0.10)**
MoA.CODEACI	5.27 (1.50)***	0.62 (0.44)	0.03 (0.35)	0.38 (0.28)
MoA.CODECMD	5.59 (0.89)***	1.90 (0.45)***	2.65 (0.21)***	2.58 (0.19)***
MoA.CODEGW	2.23 (0.45)***	1.68 (0.20)***	2.29 (0.18)***	1.93 (0.13)***
MoA.CODEOTH	4.94 (2.07)*	0.53 (0.38)	4.31 (0.78)***	1.39 (0.33)***
MoA.CODEPHI	2.74 (0.69)***	2.62 (0.21)***	3.16 (0.21)***	2.86 (0.15)***
MoA.CODESGI	3.38 (0.68)***	4.28 (0.31)***	3.76 (0.31)***	3.93 (0.21)***
LabvFieldi	0.17 (0.73)	(no data)	0.18 (0.39)	-0.33 (0.32)
LabvFieldL	0.55 (0.50)	-0.35 (0.19)	-0.27 (0.19)	-0.15 (0.13)
Class.MvDM	1.38 (0.40)***	0.48 (0.16)**	0.74 (0.15)***	0.66 (0.10)***
MoA.CODEICD		5.75 (2.38)*	5.18 (0.70)***	5.23 (0.71)***
MoA.CODELSI		2.60 (0.48)***	3.74 (0.40)***	3.18 (0.32)***
Class.MvDGymn			2.48 (1.04)*	2.49 (1.10)*
Effect.level.ERxER25				0.68 (0.16)***
Effect.level.ERxER50				1.01 (0.16)***
R^2	0.26	0.24	0.32	0.27
Adj. R^2	0.23	0.24	0.31	0.27
Num. obs.	308	1148	1281	2737
RMSE	2.85	2.38	2.29	2.42

In case of the additional model (no interaction) the intercept is the estimated value of the response variable for the first modalities of each factor under the assumption of additivity.

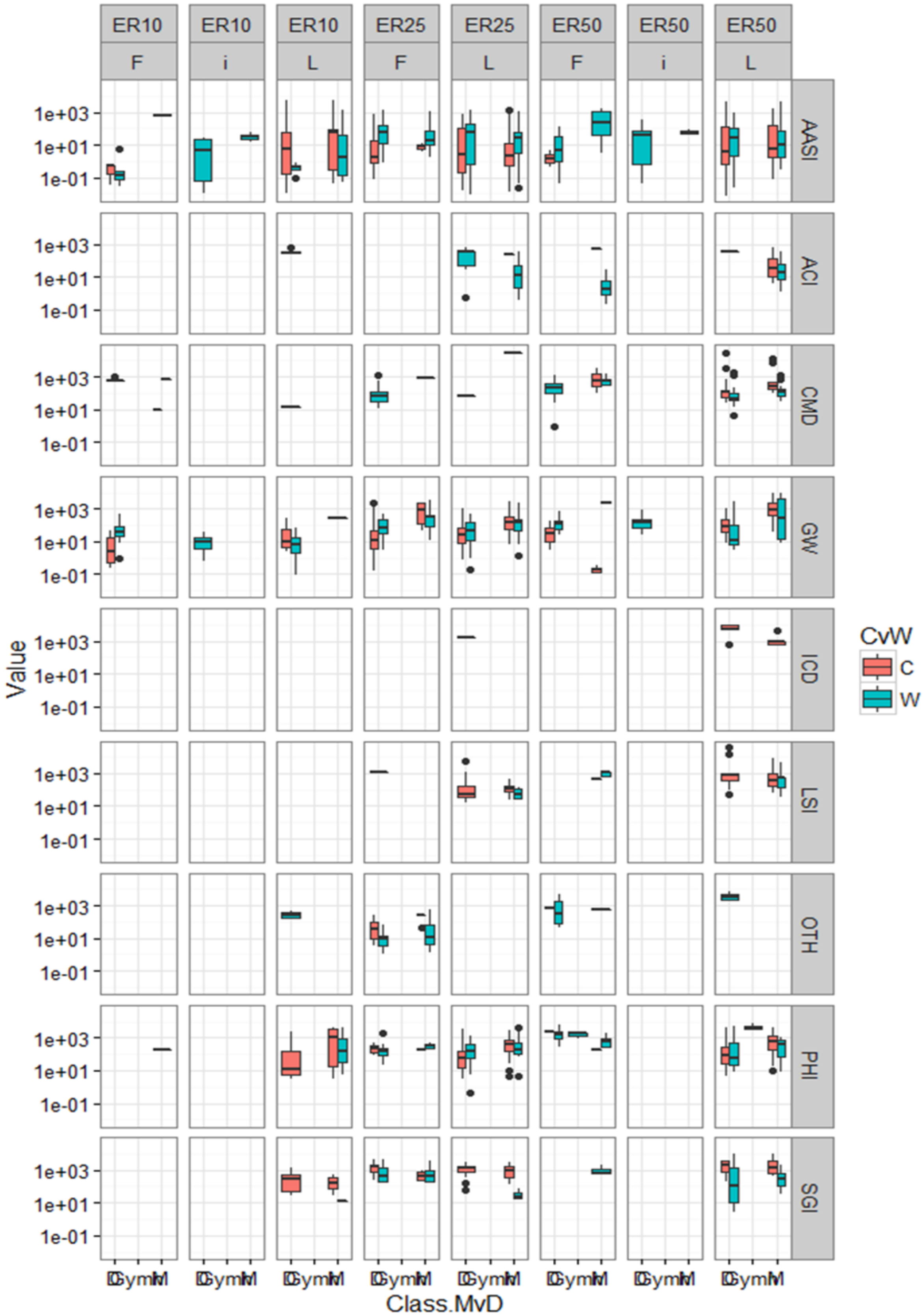


Figure 4: **Boxplot of the endpoints (log10 scale) for different Mode of Action (MoA), Class (M,D or Gymn), and Lab/Field categories.**

The boxplots in Figure 4 illustrate that the data were clustered, with a lot of data for some MoA and little or none for others, but also they show that overall no clear differences between crop and wild species were detected.

4.3.2 Assessment of statistical power

Finally it was attempted to calculate some measure of power of the analysis. The normal MDD approach considers each combination of different predictor levels and calculates an MDD based on the confidence intervals for the predicted/estimated endpoint value for both crop and wild species in this specific combination of different predictor levels. This approach thus takes the sample size and variance in this specific group into consideration, hence MDDs for different combinations of predictors vary greatly. Thus the MDDs for some predictor combinations were reasonably low (35 – 64%), but in other cases very high (for example, there is no observed data in this combination of predictor levels and a prediction based on this model would have a very large CI). However, as a first indication of power it was observed that based on Pooled ER10, ER25 and ER50 data the MDD (expressed as canonical coefficient of 'CvWW' was) 1.28 or 0.32 (with and without interactions, respectively) indicating that (back-transformed from the natural log) a difference of as little as factor 1.38[25] could result in significant deviations, but the prediction interval was wide. For the reasons given above this MDD approach is only of limited value to determine the power of the test to detect an overall consistent difference.

A different approach was to repeat the assessment with a purposely modified ("manipulated") database. The entire dataset (ER10, ER25 and ER50) was modified in such a way that all crop species endpoints were increased by factors of 1.5 or 2.0, while the wild species' endpoints were left unchanged. While this factorial modification affects the total variance, but the relative distance between the different crop species' endpoints remains the same, and so does the variance of the log-transformed crop species' data, which had been used for the factorial analysis above.

This modified database was submitted to the same regression analysis as the original data[26]. The outcome (presented in Table 5) thus may be directly compared with the outcome based on the original data with the pooled model (rightmost column of Table 4, repeated as first data column pair in Table 5).

Based on the fact that in the original data wild endpoints were higher than crop endpoints, and the expectation that based on MDD estimates two groups differing by a factor of 1.4 or more should be detected as statistically significantly different, a manipulation of one group by a factor of 1.5 should make the significant differences of the original data vanish, and a

[25] in other words, if crop endpoints are set to 100%, wild endpoints being on average around 70% or 140% would have been detected as significantly different by the fitted factorial model applied here.
[26] In addition versions were calculated using Heteroscedasticity Consistent Standard Errors, based on Long & Ervin (2000) and MacKinnon & White (1985). With this approach the standard error for the predictor crop-wild were slightly larger compared to the original results (so p-values were even larger); standard errors of a few other predictors were lower. However, in terms of differences in sensitivity between crop and wild plant species the outcome was the same.

manipulation by a factor of 2 should revert the observed relationship between the crop and the wild species' endpoints, i.e. after the modification wild plant's endpoints should be detectable as significantly lower than those of the crop species.

Table 5: Summary of two runs with modified data, fitted factorial model (with interactions or no interaction), where based on the original data all crop endpoints had been increased by a factor of 1.5 or 2. Positive coefficient signs indicate positive deviations, negative signs negative deviations compared to the corresponding baseline, e.g. that that wild species were more sensitive than crops species. The figures are based on the log-transformed data (natural logarithm). Inside parentheses the standard error of the coefficient estimation.

Predictor code	Original Data		Modified 1 F = 1.5		Modified 1 F = 2.0	
	Pooled best	Pooled no interaction	Pooled best	Pooled no interaction	Pooled best	Pooled no interaction
(Intercept)	0.04 (0.38)	1.15*** (0.19)	0.44 (0.38)	1.56*** (0.19)	0.73 (0.38)	1.85*** (0.19)
Effect.level.ERxER25	0.56*** (0.16)	0.68*** (0.16)	0.56*** (0.16)	0.68*** (0.16)	0.56*** (0.16)	0.68*** (0.16)
Effect.level.ERxER50	1.07*** (0.15)	1.01*** (0.16)	1.07*** (0.15)	1.01*** (0.16)	1.07*** (0.15)	1.01*** (0.16)
CvWW	1.28** (0.45)	**0.32**** (0.10)	0.88 (0.45)	**-0.09** (0.10)	0.59 (0.45)	**-0.38*****(0.10)
MoA.CODEACI	5.15* (2.52)	0.38 (0.28)	5.15* (2.52)	0.38 (0.28)	5.15* (2.52)	0.38 (0.28)
MoA.CODECMD	2.67 (1.56)	2.58*** (0.19)	2.67 (1.56)	2.58*** (0.19)	2.67 (1.56)	2.58*** (0.19)
MoA.CODEGW	1.98*** (0.44)	1.93*** (0.13)	1.98*** (0.44)	1.93*** (0.13)	1.98*** (0.44)	1.93*** (0.13)
MoA.CODEICD	6.23*** (0.96)	5.23*** (0.71)	6.23*** (0.96)	5.23*** (0.71)	6.23*** (0.96)	5.23*** (0.71)
MoA.CODELSI	5.24 (3.49)	3.18*** (0.32)	5.24 (3.49)	3.18*** (0.32)	5.24 (3.49)	3.18*** (0.32)
MoA.CODEOTH	3.21*** (0.65)	1.39*** (0.33)	3.21*** (0.65)	1.39*** (0.33)	3.21*** (0.65)	1.39*** (0.33)
MoA.CODEPHI	4.98*** (1.0)	2.86*** (0.15)	4.98*** (1.0)	2.86*** (0.15)	4.98*** (1.0)	2.86*** (0.15)
MoA.CODESGI	6.24*** (0.7)	3.93*** (0.21)	6.24*** (0.7)	3.93*** (0.21)	6.24*** (0.7)	3.93*** (0.21)
LabvFieldi	-0.5 (0.48)	-0.33 (0.32)	-0.5 (0.48)	-0.33 (0.32)	-0.5 (0.48)	-0.33 (0.32)
LabvFieldL	0.93* (0.38)	-0.15 (0.13)	0.93* (0.38)	-0.15 (0.13)	0.93* (0.38)	-0.15 (0.13)
Class.MvDGymn	0.5 (1.4)	2.49* (1.1)	0.5 (1.4)	2.49* (1.1)	0.5 (1.4)	2.49* (1.1)
Class.MvDM	2.53*** (0.52)	0.66*** (0.1)	2.53*** (0.52)	0.66*** (0.1)	2.53*** (0.52)	0.66*** (0.1)
(interactions…)	not displayed		not displayed		not displayed	
R^2	0.33	0.27	0.33	0.27	0.34	0.27
Adj. R^2	0.32	0.27	0.32	0.27	0.32	0.27
Num. obs.	2737	2737	2737	2737	2737	2737
RMSE	2.34	2.42	2.34	2.42	2.34	2.42

***p < 0.001, **p < 0.01, *p < 0.05

With the modified data most coefficients are identical to those obtained by the pooled analysis of ER10, ER25 and ER50 data (as expected), except for the predictor 'CvW.W' crop/wild. The outcome of the runs with modified data was thus as to be expected.

With the updated database (compared to an earlier version of this report) wild endpoints were by and large higher than the crop endpoints (baseline). Based on the original data, differences between crops and wild planta are significant, but in that direction that wild endpoints were higher, thus crops were slightly more sensitive than wild plants. Coefficients of 1.28 and 0.32 for the "best" and the "no interaction" model are equivalent to average differences by a factor of 3.60 and 1.38 respectively. If crop endpoints increased by a factor of 1.5 (which corresponds to a coefficient of 0.449 based on the natural log-transformed data) were no longer significantly different from the wild plants' endpoints. Coefficients of 0.88 and -0.09 for the “best” and the “no interaction” model equal factors of 2.41 and 0.91 respectively, so in the modified data groups were no longer significantly different. Only if crop endpoints were increased by a factor of 2.0 (which corresponds to a coefficient of 0.69 based on the natural log-transformed data), then the standard model (no interactions) detects significant differences between the two groups in the ‘critical direction’ (i.e. wild plants more sensitive than crops); coefficients of 0.59 and -0.38 for the “best” and the “no interaction” model equal factors of 1.80 and 0.68 , the latter would be matched with high significance ($P < 0.001$, see rightmost column of Table 5). Considering that based on the original data (pooled, standard model) the coefficient crop-wild had been 0.32 (see leftmost column of Table 5), calculating 0.32 – 0.69 should be -0.37. The value found for the crop-wild predictor was -0.38, see rightmost column of Table 5, which is attributable to rounding errors.

Based on back-transformed ratios, 2.41 and 0.91 are indeed a factor of *ca* 1.5 lower than 3.6 and 1.38 (from the original data) and 1.80 and 0.68 are indeed a factor of *ca* 2 lower than the former. These checks confirm that the models did calculate what they were supposed to do. The other significant differences between explanatory variables were the same or very near to those determined from the original data (for simplicity now solely focussing on the standard model (no interactions, right column each): ER25 and ER50 endpoints were significantly higher than ER10 endpoints (defined as the baseline), all modes of action (except for ACI) had distinctly higher endpoints than the baseline mode of action ‘AASI’; in terms of lab- or field tests (the latter defined as baseline) the lab endpoints were higher than endpoints of field tests according to the model with interactions, but lower if based on the standard model. Intermediate test designs (e.g. plants grown in the field but tested in the greenhouse, or the other way round) produced lower endpoints than field tests but not significantly different, monocots had significantly and gymnosperms slightly higher endpoints than dicots, as to be expected.

To conclude, if wild endpoints had been by a factor of 1.5 (reciprocal of 0.69) lower than the crop endpoints, this would have been detected by the standard model as highly significant. The extended original dataset found wild endpoints to be slightly higher than crop endpoints, and the two groups were found to be significantly different in the direction not expected; i.e. crop species were by and large more sensitive than wild plant species.

The test with the modified data above thus confirms that the fitted factorial model applied is sufficiently powerful to detect an intrinsic difference in sensitivity between wild and crop species even from data as heterogeneous and variable as found here. If the endpoints of the first group are different by a factor of ca 1.5 or more from the second, the deviations are detected as statistically significant. With the updated database significant differences between crop and wild species were detected, but indicating that overall wild species were slightly less sensitive than crop species.

4.4 Further details of the comparison of crops' and wild plant species' sensitivity (Auxiliary analyses)

4.4.1 Comparison of outcomes based on ER10, ER25 and ER50 endpoints

In the main assessment it became evident that no fundamental differences between effect levels, i.e. ER10, ER25 and ER50 endpoints appeared to exist in terms of representing relative sensitivity of crop vs. wild species. In the original dataset there were 4 substances where n of both wild and crop species was large enough and ER25 and ER50 endpoints were available to directly compare quotients obtained from ER25 and ER50 endpoints (AASI 1, AASI 3, GW 01 and PHI 02), see Table 6. Further five substances allow comparison of quotients from ER25 and ER50 while each one of the four data sets violates the minimum "n" criterion of 3, i.e. "n" = 2. These five substances are AASI 4, GW 05, LSI 1, PHI 01 and SGI 3 and their results are also listed in Table 6. There were not sufficient substances with n high enough to compare ER10 and ER50 as well, and it was not attempted to analyse these. The results of the initial analysis (based on ER25 and ER50 endpoints) are presented below.

Table 6: Comparison of outcomes based on ER25 (VV BM) and ER50 (VV BM).

	Quotients based on				Comments
	Most sensitive species		Geometric means		
Substance	ER25	ER50	ER25	ER50	
AASI 1	0.71	1.30	2.86	2.93	
AASI 3	1.21	1.62	1.72	2.40	
AASI 4	*0.072*	*0.013*	0.27	0.23	ER25 wild: n = 2
GW 01	*0.35*	*3.13*	*0.87*	*5.57*	
GW 05	*0.050*	*0.22*	*0.17*	*1.07*	ER25 wild: n = 2
LSI 1	2.83	2.70	1.67	1.77	ER25 wild: n = 2 ER25 crops: n = 2
PHI 01	0.14	0.19	0.27	0.25	ER25 wild: n = 2
PHI 02	1.49	0.54	1.03	0.62	
SGI 3	1.81	3.33	2.08	1.72	ER50 crops: n = 2
Geometric mean	0.47	0.64	0.83	1.18	
25-75%ile	(0.34 to 1.9)	(0.49 to 2.65)	(0.11 to 2.3)	(0.21 to 2.3)	
	Mean deviation between ER25- and ER50-based quotients				
Mean dev.	1.35		1.42		
	Overall mean quotient between crop-wild species				
		ER25	ER50		
Overall mean quotient		0.63	0.87		
25-75%ile		(0.24 to 1.74)	(0.24 to 2.75)		
Mean dev. overall		1.39			

Quotients were based both on the lowest endpoint (most sensitive species of each group) and the central point (geometric mean of each group). The sensitivity of wild species and crop species to these 9 substances (18 cases) either based on ER25 or on ER50 endpoints generally did not vary by more than a factor of 3 with five exceptions[27]; these pairs of quotients are printed in italics in the table above. On average ER25/ER50 quotients (based on species sensitivity quotients) ranged between 0.47 and 1.18. Overall, quotients based on comparison of ER25 values were slightly lower than quotients based on ER50 values, i.e. if only ER25 had been considered, crop species would appear to be slightly less sensitive than wild species compared to an assessment solely based on ER50 endpoints. However, the differences were marginal (on average by a factor 1.39), and deviations occurred in both directions. Thus there was neither trend nor evidence for quotients based on ER25 values being any different from quotients based on ER50 values.

Considering the wide range of quotients and deviations in both directions overall there appears to be no reason **not** to combine quotients from ER25 and from ER50 comparisons and to assess them together. Hence quotients based on ER25 were pooled with quotients from ER50 endpoints to draw overall conclusions regarding the relative sensitivity between crops and non-crops.

The quotients are also displayed in Figure 5.

[27] AASI 4 if based on most sensitive, and GW 01 and GW 05 both either based on most sensitive species or on average (geometric mean) sensitivity each.

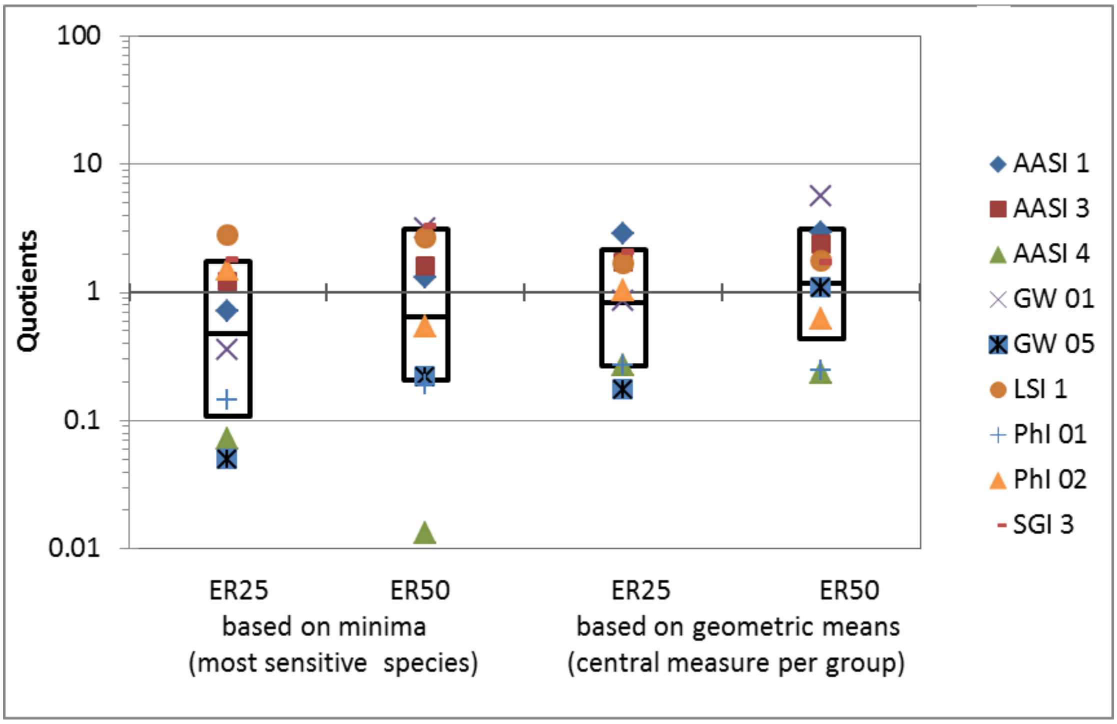


Figure 5: Differences in sensitivity between wild and crop species based on ER25 or on ER50 endpoints. The point signatures visualise the quotients (above 1 indicating wild species being more sensitive than crop species, below 1 crop species being the more sensitive group). Each set is either based on the minima, i.e. the most sensitive species of each group, or on the central points of each group (geometric mean of the group's endpoints). The boxes indicate 25%ile and 75%ile, with the central line displaying the geometric mean as central point[28]. For numeric values see Table 6; further explanations see text.

Furthermore considering that there are several substances where either only ER25 endpoints or ER50 endpoints were available in sufficient numbers it was decided that all quotients from assessing ER25 endpoints and those from assessing ER50 should be combined.

The multiple statistical approaches (based on all data, not only the data pairs assessed further up) confirmed that ER10 values were overall lower than the ER50 values (as to be expected). Based on the standard model, no interactions, the average coefficient was 0.68 (±0.16) on the logarithmic scale and the coefficient between ER10 and ER25 was 1.01 (±0.16) respectively) see 4.3.1, Table 4. The repetition of the analysis based either solely on ER10, ER25 or on ER50 estimates revealed that based on the available data sets there were some differences in terms of sensitivity of wild plants and of crops; Based on ER25 values wild species appeared to be more sensitive than crop species, based on ER10 and ER50 values it was the other way round; details see 4.3.1. The degree of deviation was

[28] In standard boxplots the median would be the central value to be displayed. As here the geometric mean was used in all tables, introducing another central point would probably rather confuse than contribute to clarity. Also generally there were no pronounced differences between median and geometric means, hence resulting quotients were similar.

+0.65 for ER25, -0.37 based on ER10 -0.01 based on ER50 and +0.32 based on the pooled model (all on the log scale), respectively. These canonical coefficients translate into factors of 1.9 based on ER25, 0.69 based on ER10, 0.99 based on ER50, and 1.38 based on the pooled model, respectively. These factors indicate the factorial difference at the original dose response scale. All these differences are within inter-lab variability and balanced around 1.

4.4.2 Comparison of outcomes based on lab/greenhouse studies or on field studies

There were 27 substances where pairs of endpoints of the same species were available both from lab and from field studies, often paired data were available for several species, so a total n of 79 endpoint-substance combinations were available, and sometimes more than one quotient calculable due to presence of lab and field endpoints at the same effect level, hence a total of 81 data pairs (excluding one extremely high and on extremely low outlier pair). The pairs of endpoints for individual species/test substance combinations were combined either by active substance and then assessed weighting them by n (number of species per active substance) or alternatively assessed as individual pairs for any differences in sensitivity between lab/greenhouse and field tests. The overall result was again calculated both as a standard geometric mean and as a weighted geometric mean (if assessed by active substance). Here the number of common species (tested on a particular active substance both in the lab/greenhouse and in the field) was the basis of weighting (for details of weighting see 3.3.3). It must be considered that only a small proportion of species had been tested both in the field and in the lab, so that any finding must be interpreted cautiously, however there is other data focussing on potential differences between test designs (greenhouse/lab or field, see discussion). Of the 22 cases[29] for which pairs of tests were found, (data displayed in Table 7 13 indicated field tests to be more sensitive than lab-greenhouse test systems, i.e. the latter species endpoint was higher than the matching species endpoint of a field test. However, 8 cases indicated the opposite, i.e. lab/greenhouse test systems were more sensitive than field test systems, and one provided identical endpoints. On average field test systems were as sensitive as lab/greenhouse test systems (differing by a factor of 1.03). Details are presented in Table 7. See also discussion Point 5.3 on p. 51.

[29] 18 active substances were available with numeric endpoints from both lab and field tests with at least one identical species. In four instances both ER25 and ER50 values were available that could be compared, hence the total number of 22 cases. A total of only 32 species-test substance combinations could be assessed, so on average these individual comparisons were based on just 1.5 species per test substance. The outcome must therefore be interpreted cautiously.

Table 7: Comparison of endpoints of species/active substance/endpoint combinations tested in the laboratory/greenhouse and in semi-field/field test systems (Option A = listed by active substance). Numbers in brackets give the range per substance/endpoint combination. If just one figure, there had been only one species with both endpoints (i.e. just one quotient).

Code	ERx	n common species	Quotient lab/field	
AASI 1	ER25	2	0.87	(0.83 - 0.91)
AASI 1	ER50	4	0.77	(0.32 - 1.29)
AASI 3	ER25	12	0.70	(0.14 - 3.18)
AASI 3	ER50	6	1.05	(0.52 - 1.81)
ACI 2	ER50	4	5.41	(1.19 - 11.24)
CMD 1	ER50	3	0.22	(0.114 - 0.37)
CMD 2	ER50	1	2.06	(2.06)
GW 05	ER50	1	11.3	(11.33)
GW 08	ER50	1	1.55	(1.55)
LSI 2	ER50	1	1.00	(1.00
OTH 3	ER50	1	1.25	(1.25)
OTH 7	ER50	1	1.51	(1.51)
PHI 02	ER25	6	0.81	(0.45 - 1.28)
PHI 02	ER50	2	1.86	(1.07 - 3.22)
PHI 04	ER50	2	0.71	(0.6 - 0.84)
PHI 08	ER50	1	3.26	(3.26)
PHI 09	ER50	1	0.26	(0.26)
PHI 10	ER50	1	1.70	(1.7)
SGI 3	ER25	1	1.10	(1.1)
SGI 3	ER50	1	2.18	(2.18)
SGI 4	ER50	1	1.08	(1.08)
SGI 5	ER50	1	0.88	(0.88)
min			0.114	
max			11.3	
Simple quotient			**1.253**	
Overall weighted quotient			**1.032**	

Based on the assessment of individual species (n = 54), the largest deviations ranged from 0.114 to 11.3 i.e. in one instance the lab endpoint was almost a factor of 8.76 lower than the field endpoint, the other extreme was a lab endpoint by a factor of 11.3 higher than the corresponding field endpoint. 25%iles and 75%iles were 0.70 and 1.39, respectively, again indicating symmetric scatter of quotients around 1. Also the calculation by individual species pairs (Option B) resulted in an overall quotient of 1.03, i.e. very close to 1, (the simple overall mean was 1.25). While it must be considered that this subset of data pairs was relatively small and variation within it was high, it is obvious that variation between e.g. larger systematic groups (monocots/dicots) was much larger than any variation caused by different test systems.

While the assessment just presented was not repeated including ER10, the following quotient approach was performed with all data and effect levels. The database was searched for substance-species combinations where the pairs of the same ERx were available from lab tests and from field tests, 81 data pairs were available[30], 57 wild species and 24 crop species, of which quotients were calculated, here dividing the lab endpoint by the field endpoint.

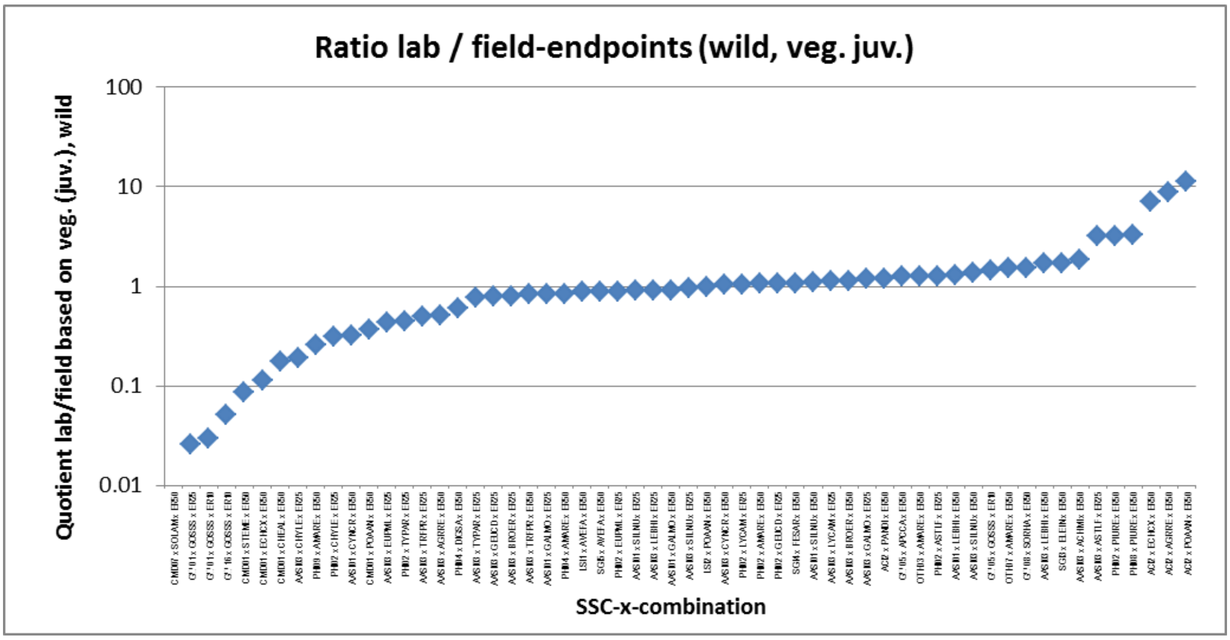


Figure 6: Differences between lab and field endpoints (wild species only), again expressed as quotients. Quotients below 1 indicate that the lab endpoint was lower than the field endpoint, greater 1 that the field endpoint was lower, hence more sensitive; further explanations see text.

Based on these quotients the overall geometric mean and median quotients were 0.90 and 1.06, respectively (n = 81). The quotients lab/field are thus distributed fairly symmetrically around 1, hence do not indicate that one endpoint type (lab or field) was generally lower than the other.

This data may also be split up into crop and wild plant species, and charts based on these fractions are displayed as Figures 6 and 7.

Figure 6 is based only on wild endpoint pairs (i.e. wild species-substance combinations of the same effect level for which lab and field endpoints are available). Based on these, the overall geometric mean an median quotients were 0.80 and 0.95, respectively (n = 57).

[30] not counting two outliers, one extremely high and one extremely low quotient that are regarded to be due to data errors.

Considering only on pairs of crop species' endpoints (Figure 7) the overall geometric mean an median quotients were 1.23 and 1.11, (n = 24).

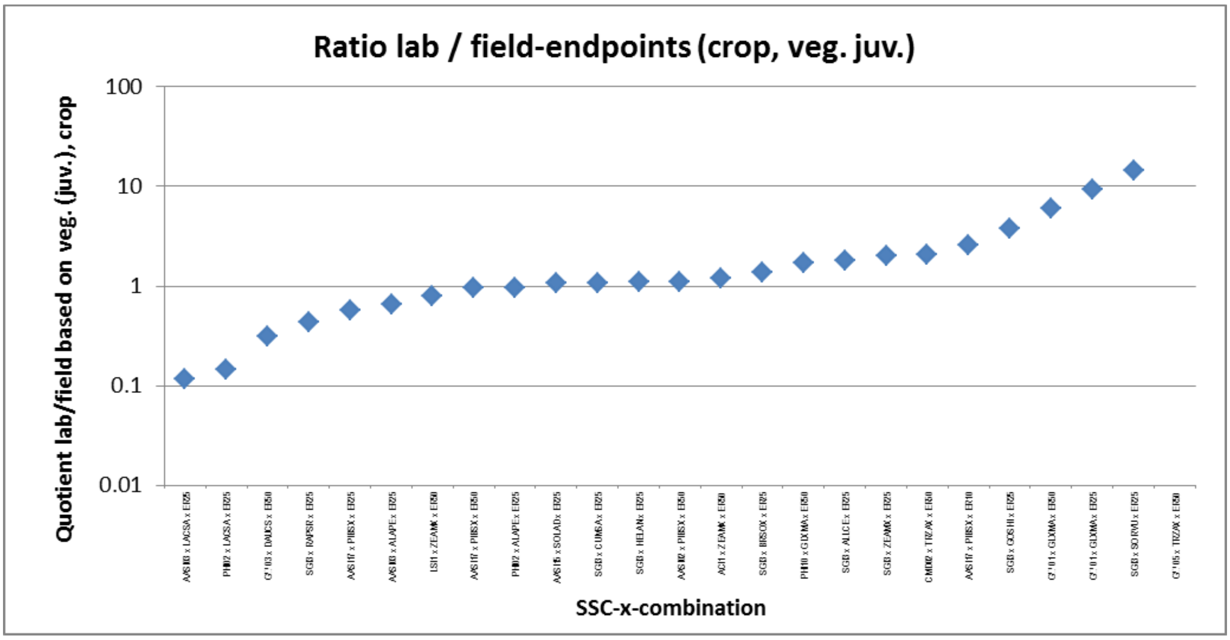


Figure 7: Differences between lab and field endpoints (crop species only), again expressed as quotients. Quotients below 1 indicate that the lab endpoint was lower than the field endpoint, greater 1 that the field endpoint was lower, hence more sensitive; further explanations see text.

Overall quotients were found to be close to 1 in both subsets, which allows two conclusions: First, endpoints from field tests were quite similar to endpoints from lab tests in this database; and second, the main question, whether crop and wild species differ in sensitivity, is unlikely to be markedly affected by any bias due to varying frequencies of lab-data or field-data in the different subgroups (particularly crop or wild)..

Finally the multiple regression approach (based on all data, not just on the pairs assessed above) also addressed the question whether lab and field data were systematically different. The evaluations based on different effect levels resulted in inconsistent findings. Based on ER10 endpoints, lab-test values were slightly higher than the field endpoints but not significantly when based on the standard model[31]. Based on ER25 and ER50 lab endpoints were slightly lower but again not significantly different from those of field tests, and based on all effect levels (pooled standard model) lab and field endpoints were not significantly different either, confirming that in this database, and based on all numeric data including ER10, differences between lab and field endpoints were slight. A canonical coefficient

[31] Based on the model with interactions, the canonical coefficients was 0.99 and indicated just significant differences, with lab endpoints slightly higher than field endpoints, but it must be considered that the interactions were defined solely by the algorithm, as no mechanisms for potential interactions are known, so these may well be just a mathematical artefact.

of -0.15 on the logarithmic scale translates into a factor of 0.86 on the original scale, so lab endpoints were on average only slightly lower than matching endpoints of field studies. These differences are very minor compared to the considerable random noise of data caused e.g. by inter-lab variability.

Overall no consistent and significant difference in sensitivity between lab/greenhouse and field test systems was detected in this database, irrespective of whether based on comparisons of distributions or based on matching datapairs of individual substance-species combinations.

5 Discussion

5.1 Retrieval of endpoints from the literature

In the process of notification of plant protection products these are tested for effects on non-target terrestrial plants. Generally crop species are tested as representative for the different plant morphotypes and systematic groups. While a few wild species are also regularly tested (e.g. *Lolium perenne*), generally crop species are preferred for practical reasons; crop seeds from known sources are readily available throughout the year, they germinate reliably, which is a validity criterion according to the OECD guidelines, and thus enable test facilities to generate consistent and reproducible results.

However, currently there is a debate if this selection of crop species as representatives for flora of Europe potentially present in the agricultural landscape is adequate. There is data that suggest that wild species could be more sensitive than the tested crop species, and therefore risk assessment based on the latter might not be protective. (e.g. UBA 2009 in EFSA 2014), and it is therefore suggested to test wild herbs frequently found adjacent to fields. Other scientists however feel that testing 6, 8 or even 10 plant species (numbers cited in Boutin et al., 2012, EFSA 2010, 2014, Regulation (EC) No 1107/2009, SANCO/10329/2002 rev 2 final(17 October 2002) that are often herbaceous weed species (and therefore r-strategy species[32]) might also not be sufficient to represent the flora of Europe that is potentially present in the agricultural landscape (Boutin et al., 2012, EFSA 2010, 2014) so suggest to test even more species, among these also representatives of woody species. However, these proposals seem not to take into account practical implications, such as number of tests, validity of tests (e.g. germination criterion according to OECD 208), the duration of any testing program, and costs. While some of these suggestions are considered in the most current opinion paper of EFSA (2014) on the science behind the risk assessment of terrestrial non-target plants, the first and foremost question seems to be if there is indeed a need to change the plant testing strategy, based on the question whether wild plant species do show a higher intrinsic sensitivity compared to standard test species, or if the current testing scheme can remain unchanged.

A related question not considered in the present paper is which are the protection goals for non-target terrestrial plants, and - based on this – how much effect is considered acceptable at different distances to the crop-field. These protection goals are yet to be clarified.

5.2 Quality of literature data

Potential reasons for the widespread perception that wild species were more sensitive than crop species:

[32] r-strategy species have a short life cycle, a high reproductive rate and a high recolonisation potential as opposed to K-strategy species.

- Comparison of different measured experimental endpoints - standardized survival / shoot height / biomass endpoints generally tested on crop species, non-standard phytotoxicity symptoms or reproductive endpoints often tested on wild species.
- Comparison of different test designs - standardized lab/greenhouse tests generally tested on crop species, non-standard community and field studies mostly tested on wild species.
- Use of old and less efficient formulations in old regulatory studies (as cited in DAR and hence publicly available) or even just endpoints from the technical material, which is often the only crop endpoints available to authors from academia as the regulatory studies performed under GLP fall under data protection and normally are not published; modern and optimized formulations generally tested on wild species in most of the more recent tests.

The last point was discussed by White & Boutin 2007. They observed "[…] glyphosate elicited a significantly less toxic response than Round-Up Original (Table 2)." Although technical material was applied at higher rates than the formulated product, there were several plant species where the technical material did not affect growth significantly at any test rate, whereas the product did (White & Boutin 2007, Fig. 1). Similar correlations may be expected for most active substances, in particular those that have been on the market for some time, not only for glyphosate. From own experience with herbicides there is often a striking pattern of newer formulations generating lower endpoints than older formulations or the technical material itself. Considering that formulations are developed to be most efficacious at low rates, it seems to be coherent that the more time has been invested in refining formulations, the more potent they become. This is particularly important when assessing effects of herbicides on non-target plant species, as the non-target plants are often very closely related to the target plants, the weeds to be controlled. Hence if endpoints obtained from technical material or first-generation products are related to effects caused by modern optimized formulations, the potentially significant differences between the two might often just be caused by differences in efficacy between technical material/old formulations and modern products.

In studies where wild and crop species were tested on the same formulations in parallel, there were generally no clear differences in sensitivity between crop species and wild plant species (e.g. Clark et al. 2004, Egan et al. 2014a, Strandberg et al 2012). The latter conclude "The sensitivity of non-target plants to herbicides measured as survival and biomass does not vary significantly from the sensitivity of crop species" (Strandberg et al. 2012).

The difference between these conclusions and other authors who state that there were relevant differences in sensitivity between wild plant and crop species (e.g. US-EPA, Francois T, D. 2001) can probably be explained by the circumstance that the former compared like with like, the latter did not. Test designs varied (lab/greenhouse tests or field studies), test materials varied (old / new formulations), and endpoint types varied (standardized survival / shoot height / biomass endpoints) versus non-standard phytotoxicity symptoms or reproductive endpoints. This is in line with observations discussed by Strandberg et al 2012, who stress that the influence of test conditions on the observed sensitivity in tests may have been underestimated in the past, and they propose to use any data base only when test conditions were also recorded, as in absence of such information wrong or misleading conclusions on species sensitivity may be drawn (Strandberg et al. 2012). So this is indeed a weak point of the current data. However, the auxiliary analysis 4.4.2 and the multiple regression analysis 4.3.1 indicated that by and large lab and field tests resulted in similar endpoints (based on endpoints of species tested in both test systems). This indicates that although some bias could have been introduced by pooling the data, this is unlikely to have changed the overall outcome. See also 5.3

5.3 Test conditions

Another issue is that potential effects of culturing conditions on sensitivity are difficult to quantify. Allison et al. (2013) investigated effects on a particular parameter, soil organic matter content, and found differences but not a clear pattern. The results suggest that nitrophilous and non-nitrophilous plants might require different culturing conditions for most realistic results. This has however to be balanced with the complexity of testing requirements and the lack of data for most wild plant species. Also Bidelspach et al. (2008) considered different soil types and propagate using standard artificial soil, which would reduce one potential source of variability between tests. Any parameter affecting growth may also affect endpoint i.e. the perceived sensitivity. Bidelspach et al 2008 found strong influences of soil type in particular on growth of control plants, whereas growth of exposed plants was less dependent on the soil type. If ERx values had been calculated (which the authors did not, also only two treatment levels were tested) the values would have been greatly affected by the control performance, which in some soil-species combination could vary by more than factor 2 (Bidelspach et al. 2008).

Another point causing additional noise in the data is the less harmonized growth stage in field tests. In lab/greenhouse standard guideline tests plants are applied at defined growth stages, in the field studies most endpoints were assessed following a foliar application of fully emerged plants at generally early growth stages, i.e. overall they are similar to vegetative vigour studies. However, growth stages vary greatly in field studies, and often are not even reported in the reviewing papers or only qualitatively described. The lack of harmonized growth stages in field studies is considered to add further noise to the data. We considered that by and large the comparison between vegetative vigour study endpoints and field endpoints is the most appropriate approach (in particular considering that in both approaches exposure is via post-emergence foliar application). In cases where only few numeric vegetative vigour endpoints were available for a given group, outcomes must however be interpreted with caution.

In lab- and greenhouse tests exposure is standardised as far as possible. Still the composition and intensity of irradiation may affect the persistence of the acting active substance and its metabolites. Depending on the exact composition of light (including intensity of UV-radiation which may cause photolysis), effective exposure will vary even in different lab- and greenhouse test systems; however these details are generally not reported, and if they were, it would still be extremely difficult to assess quantitatively the actual foliar exposure for every individual case. Next, growth of the plants results in a dilution of residues on foliage. Plants that continue growing during the test will reduce their effective foliar exposure compared to those with little growth. Last but not least, weathering may reduce exposure substantially in the field; which however will vary between different formulations and active substances, with some of them more others less mobile, and between the actual weather conditions at the days after application. All these parameters do affect the actual exposure of plants and thus their sensitivity. However, it appears not to be feasible to quantify them all, and to put the results in a causal chain.

Other parameters that might affect growth and hence also sensitivity, but that also add realism to the studies, are e.g. soil type, nutrient content, temperature, light supply and circadian rhythm. Optima for different species do differ, but this is rarely implemented when testing uncommon species. Also this cannot be fulfilled for several species at the same time in the same container or field, which could be seen as a major disadvantage of testing communities of different species in the lab, in particular if the results are then interpreted as quantifying intrinsic toxicity of a species.

Growth conditions can be controlled in order to get consistent results, in particular for commonly used test species (i.e. largely crops). Satisfying growth conditions are ensured by following the guidelines (e.g. OECD 208, OECD 227). Hence the protected company data used in this data base is assumed to satisfy most requirements defined in the test guidelines.

In contrast, for tests on wild species validity criteria are often ignored. Some criteria may even never be applied, as wild plant species follow other reproduction strategies than crops (e.g. formation of a seed bank, see point 3.2.7 on p. 17). It must therefore be expected that tests on wild plants will often be performed under varying conditions, consequently there is another source of noise in the data, which is expected to further have widened the range of endpoints obtained for wild plant species.

For suitability of wild species see also Pallett et al. 2007. The tests of White and Boutin, 2007; Carpenter and Boutin, 2010, and Pallett et al. 2007 who refer to a ring test on various wild plant species have been interpreted in different ways. In their Scientific Opinion (EFSA 2014) the EFSA Panel considers that numerous phytotoxicity studies have successfully been conducted using non-crop plants […] [and] that using non-crop species in greenhouse testing was straightforward (EFSA 2014). In contrast the authors themselves (e.g. Pallett et al. 2007) were quite cautious to suggest wild species as standard test species, as several issues occurred, in particular differences between seed sources, low and inconsistent emergence (despite the recommended pre-treatment of seeds), unacceptably large variability in biomass. Also it was admitted that in most cases germination characteristics were not tested (EFSA 2014).

Overall the authors perceive a lack of fundamental research on which factors affecting plant performance increase and which decrease sensitivity to toxic substances or other adverse agents. While it is generally accepted that rapidly growing plants are more sensitive than those growing at a slow rates (e.g. OECD 208, 227, Mayer et al. 1998, Smith et al 2000) it is not clear (and there may not even be an ubiquitous answer) up to which extent “hardening” e.g. of seedlings prior to transferring them in the field may increases their resilience or in other cases may weaken them and make them less resilient to additional stressors such as xenobiotics. Higher plants are very complex test systems, and the interactions between these different aspects appear not to be fully understood yet.

Obviously the multitude of factors influencing sensitivity also varies between lab and field studies, without them being quantifiable – at least not for this data set. More work is needed to assess in which way different counteracting parameters may influence exposure and growth of plants in different test systems, multispecies test systems growth and in turn sensitivity of plants in different test systems, which is further complicated when considering e.g. multispecies lab- and greenhouse or field test systems and resulting competition (see e.g. Damgaard et al. 2008, Reuter & Siemoneit-Gast 2007). All these parameters together affect the sensitivity of a given test species to a given herbicide.

So test conditions do affect exposure and plant growth in many ways, which in turn affects the plants' sensitivity. However, we consider that overall the added variability will have affected endpoints in both directions, and also will have affected both crop and wild plant species. Therefore it is considered unlikely that a systematic bias could have been introduced by pooling different test designs in this database.

5.4 Which measure of sensitivity (lowest or average endpoints) to be used

The regulatory relevant endpoints in the European risk assessment are primarily the lowest endpoints of all tested species, hence it is obvious that a comparison between crop and wild species based on minima is relevant (e.g. SANCO/10329/2002, US-EPA (2004), EFSA 2014). However, minima and maxima are more dependent on the number of tested species. In particular in case of low numbers the presence or absence of one extreme value will determine the outcome of the comparison.

In contrast, the central points of any distribution, e.g. median, arithmetic mean or – in this case most appropriate – the geometric mean, are less influenced by any extreme values (e.g. Sokal & Rohlf 1995). While we included all endpoint values including the outstanding extreme ones in the numeric assessment[33] the fundamental question whether any group is generally more sensitive than the other can most robustly be answered by comparing the central points of the two distributions, not just the minima. Therefore from a purely scientific viewpoint comparison of central points should be preferred. To address both the fundamental question and the aspect of greatest regulatory relevance the quotient assessment was performed in parallel twice, first based on central points of the distributions (geometric means), then on the minima (i.e. lowest endpoints, viz. the most sensitive species of each group).

In parallel different ANOVA assessments (s.l.) were performed, including multiple regression analysis, for which these considerations do not apply as the actual numeric species' endpoints were assessed, not just the central points or minima of any distribution.

[33] Except for one single outlier value that was apparently wrong

5.5 Overall quotient approach

The overall quotients used in the conclusion (based on the individual quotients of all datasets) may be calculated either as a standard geometric mean (no weighing) or weighing the individual quotients by the number of endpoints forming it. The latter approach is considered to better reflect the actual status, the simple average quotient was only included for completeness; also it can be verified easily. However to avoid deviations due to datasets with just one or two endpoints per crop, it was considered here only to use quotients calculated from at least three numeric endpoints per group (≥ 3 wild and ≥ 3 crop species).

The standard weighted approach simply considers ‘n’ per group. However the values weighted here are quotients, formed out of two groups an ‘n’ each. For this special case of a weighted approach there are different options to calculate it, either to weigh by the total number of endpoints (n_{wild} + n_{crop}) and second to weigh by the lower ‘n’ of the two. The first approach is straightforward but disregards that a quotient based on 3 crop and 17 wild species is not as robust as one based on 10 crop and 10 wild species (‘n’ = 20 in both cases). The second approach considers this difference in predictive power, but disregards that the quotient based on 3 crop and 17 wild species is still not as weak as the one based on 3 crop and 3 wild species (17 endpoints would make the group estimates (minimum, geometric mean) more robust for the wild species than the group estimates of the crops), but the lowest ‘n’ is three in both cases. As the overall power is more affected by the weakest element than by the total ‘n’, the approach basing the weighing on the lower n of the two groups was considered to be the best choice and thus was applied here.

Initially it was considered that quotients should only be calculated from cases with at least three numeric endpoints (three species per plant groups (wild or crop species). However, there were borderline cases with low ‘n’ that were deemed useful for comparison of ER25 and ER50-derived quotients (without these the number of cases for this supplementary evaluation would have been very low), so it was attempted to include these as well in the main analysis, with only little effect on the overall outcome. Apparently deviations existed in both directions and cancelled each other out. Also the evaluation weighted by the number of species (see previous paragraph) is not significantly affected by the decision whether or not to include minor data sets, as the overall outcome is dominated by the data sets for which many endpoints exist. Finally including all possible data sets prevents any concern about the threshold set (minimum n = 3). Therefore ultimately all datasets based on vegetative vigour / biomass measurements that allowed calculation of quotients were included in the overall assessment.

5.6 Heterogeneity of endpoints and selection of final dataset

There is considerable heterogeneity in experimental endpoints being reported to describe effects on non-target plants (e.g. Fletcher 1985, Strandberg et al 2012). There is the effect level, viz. ER05, ER25, ER50…, but any of these may be obtained from a greenhouse or a field test, a seedling emergence study or a vegetative vigour study, and within these be based on different measured variables, e.g. on survival, shoot height, on biomass (wet weight or dry weight), on phytotoxicity (various definitions including reproductive endpoints[34]) or different scoring systems. Some studies did not follow the rate-response design but tested just one test rate and ranked species by the effect level observed on this sole tested rate.

The different endpoints reported in the literature were, unfortunately, not evenly distributed as there were not many seedling emergence data from wild plant species, but numerous from crops. Crop endpoints were regularly based on shoot height, wild species' endpoints rarely. There were more wild species endpoints from field tests or intermediate test designs and fewer from lab/greenhouse tests. While in many reports (in particular from regulatory studies) various parameters were measured and turned into several endpoints from one experiment, other data reported often just one particular endpoint type, often even without clearly specifying the basis of the endpoint (e.g. wet weight or dry weight).

In this evaluation it was considered paramount to compare like with like. However, if this approach had been applied rigorously, only few dataset would have allowed comparison of wild and crop species endpoints; therefore compromises had to be made, balancing the desire to compare only true pairs of experiments with the target to include as much of the available data as possible.

The resulting selection for the final evaluation was to combine field study endpoints with greenhouse endpoints[35], to include ER10, ER25 and ER50 values (always comparing like with like), based on biomass measurements (the largest fraction) but not differentiating between wet weight or dry weight measurements)[36]. Seedling emergence data were ultimately not included in the main assessment for two reasons: there is only little seedling emergence data for wild species, and for most species/active substance combinations that provided seedling emergence data also vegetative vigour data were available. Analysis solely based on greenhouse data or only on field data would have been possible, but substantially reduce the number of comparisons possible. Exclusion of shoot height endpoints did not make much difference, as there were only few wild species endpoints based on measurement of shoot height, and for those available generally there were also biomass data available for the same species, often generated within the same study. Also we checked if there were many cases where the lowest endpoint was not based on biomass. There were only a few cases where the SH-endpoint was the lowest, but the difference was generally less than a factor 1.7. The one instance with the largest deviation was a crop species; its shoot-height based endpoint was by a factor of 5.1 lower than the one based on biomass. If not evaluating shoot-height based endpoints, in this particular case the crops will appear to be less sensitive than if including shoot height, i.e. for our comparison it is a conservative approach.

---

[34] Effects of different endpoints including reproductive endpoints are discussed in a separate paper.
[35] It is appreciated that this adds a factor of uncertainty but see 4.4.2 and discussion 5.7.
[36] Generally in a given study biomass endpoints were either based on wet weight or on dry weight, but rarely both variants were presented. In the few cases where both were available, the lowest endpoint of the two was used.

Anyway, considering comments on earlier versions we did include also an alternative approach, considering also seedling emergence data and other measured parameters such as shoot height or survival. Also the additional statistical analysis performed by John Green (Appendix 7) included all endpoints included in the database.

As indicated, for the main evaluation the largest homogenous data set was used; ER10, ER25 and ER50 endpoints from vegetative vigour-type studies, mostly based on the measured variable “biomass”[37], sometimes on scoring systems considered to be a close approximation, but not on survival, shoot height, reproductive endpoints or similar. This largest homogenous dataset utilizes approximately 70% of all available numeric ER25 and ER50 species endpoints, which puts the proportion of not evaluated endpoints into perspective.

5.7 Differences in sensitivity between lab and field test systems

Many authors argue that sensitivity of laboratory and greenhouse test systems and field test systems differs (e.g. Boutin 1995, Dalton & Boutin 2010, Reuter Siemoneit-Gast et al. 2007) so that risk assessment based on a dataset solely from laboratory tests would need an extra safety factor to ensure protection of non-target plants in the field (e.g. Dalton & Boutin 2010). However there is also evidence that sensitivity of plants tested in the lab may be very similar to that of plants tested in field systems (e.g. Fletcher et al, 1990, Egan et al. 2014b, Strandberg et al. 2012) if the same endpoints are being assessed. The assessment presented here, comparing pairs of plant species tested on a given active substance both in the lab and in the field support the latter. Generally endpoints were very similar no matter if tested in the greenhouse or in a field situation. Exceptions did occur but any deviations were found to spread symmetrically in both directions (as many cases in which field tests generated the lower endpoints as those where lab tests generated the lower endpoints). While the dataset is small and neither their systematic position of the available species nor the active substances for which pairs of test endpoints existed are necessarily representative, the finding is in line with other published data where equivalent endpoints were assessed. Due to the paucity of data this outcome should not be regarded as a final conclusion as to whether there may be differences in sensitivity between lab/greenhouse test systems or field test systems or not. However, based on the data evaluated here an assessment factor just to cover extrapolation from the lab to the field situation seems to be dispensable.

5.8 Background information on the two most extreme datasets.

In the following section the background of the two most extreme datasets (in terms of differences in sensitivity between wild species and crops species) is presented.

AASI 2 – ER50: This data set is of an old herbicide, 16 wild species endpoints and 42 crop endpoints were available. However, 13 wild species were greater-than endpoints, so only 3

[37] biomass not differentiating between wet weight and dry weight

numeric endpoints could be used in the assessment, and these were again of just one study in which just one rate has been tested; the effect level happened to be around 50% for these three species. So this old study is not very reliable, and also if further rates had been tested and more numeric endpoints been generated, these would have further raised the average sensitivity of the wild species. In contrast only three of the 42 crop endpoints were greater-than values; the remaining 39 endpoints were obtained from only ten species, (i.e. on average 4 endpoints per species), resulting in 10 numeric endpoints, calculated as geometric means of the individual endpoints (per species). Thus, three numeric endpoints of wild species from an old and not very robust study were compared with 10 extremely robust crop endpoints. Overall in this instance the wild species proved to be insensitive, but based on the details above it is apparent that this quotient must be interpreted very cautiously, and it is by no means suitable to support any hypothesis that such large differences between wild and crop species' sensitivity could be an abundant and often occurring real phenomenon.

The other extreme was the data set SGI 2 - ER50 where the deviation was the other way round. Another old active substance; 15 wild species endpoints were reported in published papers (plus one from a regulatory study), with endpoints covering a huge range (rates spread over more than three orders of magnitude). The crop species data from several regulatory studies were much more consistent (9 numeric values). In the published paper there are no details reported that could explain the large scatter of wild species endpoints, however it must be considered that the paper used endpoints both from own greenhouse experiments and from an old EPA-database not further specified. In the latter only the lowest endpoint each had been used to calculate the ER50, without specifying which measured parameter actually lead to the numeric endpoint. Obviously this adds quite some uncertainty to these endpoints. In addition the regulatory endpoints were generated only with technical material[38], whereas the published endpoints tested effects of a formulation. As formulations tend to be much more efficient than the technical material as such, this could be another factor contributing to the large differences between crop endpoints and wild endpoints observed in this instance.

So overall there are a number of aspects that could have biased the comparison in these cases, and there are more of such cases. One option would have been to exclude such data sets. However, few cases were clear cut, so this would always have included a judgemental factor, and would have attracted criticism. There were many cases in a grey zone, and it would have been difficult to draw a line between data sets still considered acceptable for the comparison exercise and others being potentially biased to such an extent that they should not be included in the overall assessment.

Therefore it was considered to be prudent to include also these problematic cases, expecting that by and large these would cancel each other out. The overall result appears to confirm this expectation.

[38] The author tried to obtain also formulation data for crops, Most formulation data available is however for mixtures, other studies reported only ER25 endpoints, so could not be used to expand this dataset either.

5.9 Multiple regression analysis

Pre-tests indicated that there was heterogeneous variance and uneven sample size at each factor level combinations in the multiple-way ANOVA model, which can affect the homogeneity of variance assumption. ANOVA is considered to be robust to moderate departures from this assumption. However the departure needs to stay smaller when the sample sizes are very different. According to Keppel (1991), there is not a good rule of thumb for the point at which unequal sample sizes make heterogeneity of variance a problem. A practical rule of thumb is that the largest group variance can be up to four times the smallest without posing strong problems.

There may be pronounced issues with unequal sample sizes in factorial ANOVA, if the sample sizes are confounded in the two (or more) factors. For example, in a two-way ANOVA, the two independent variables (factors) could be age (young vs. old) and marital status (married vs. not). If there are twice as many young people as old and the young group has a much larger percentage of singles than the older group, the effect of marital status cannot be distinguished from the effect of age.

A similar case may be here where more wild species have been tested in the field than crop species. However, as apparently there were neither pronounced effects of the test design (field test endpoints and lab endpoints did not differ conspicuously) neither differences between crop species and wild plant species, this fundamental problem – while applicable here – will not have changed the outcome. This would have been different if any relevant deviations had been found.

It must be considered that some ER10, ER25 and ER50 for the same substance are closely related, in particular those derived from the same experiment / dose-response curve; these are not independent from each other. However most of them are, as of by far most experiments only one set of endpoints, either ER25 or ER50, was reported. The fraction of data sets with ER25 and ER50 from the same study was small (a worst-case assessment considering also greater-than cases found 35% of species-substance combinations with both ER25 and ER50 values, and 24% of species-substance combinations with both ER10 and ER50 values), but only a part of these actually delivered numeric ER10 and ER50 or ER25 and ER50 from the same study, as the figures above also considers censored values. Of these, a considerable part are just one of several experimental endpoints that together form the species' endpoint, and considering that matching species ER25-endpoints and ER50 endpoints will be combinations of different experimental endpoints, as some studies only generated ER25 and others ER50 endpoints, there is only a partial overlap, and in terms of independence these ER10, ER25 and ER50 species' endpoints are somewhere in-between. Still, pooled analysis is not "completely proper" since an - albeit small - part of the data sets was not independent. The potential bias introduced by this pooling is however considered to be outweighed by the larger n and thus higher power.

Factorial fitted models:

Again, the heterogeneous composition of the data set is not ideal for determining a general trend/generic conclusion (see Figure 4). Instead of an automatic (purely numeric) model selection, interaction terms could also be defined by the scientist if any causal and biologically relevant relations were known. However, in absence of further information, and also after initial assessment of the data we decided to focus on the outcome of the standard model. (The results of the model with interactions were similar though.) Also the model with interactions assumes different relationships between crop and wild in each MoA, L/F, taxonomic class etc., whereas our fundamental assumption was that if there is an intrinsic difference in sensitivity between wild and crop plants, it should be similar for any active substance, so there should be a general trend. If not, it would not be a fundamental difference in sensitivity between the two groups.

An even more robust comparison would also have been to assess ER10, ER25 and ER50 pairs or triplets derived from the same dose-response curve, ER10, ER25 and ER50 groups derived from both lab and field experiments using the same substance, and above all more experimental data testing wild and crop species tested in parallel in the same test systems on the same active substances/formulations and calculating the same endpoint variants. In general, comparison can be done at best if based on comparable pairs or groups with a balanced sample size instead of pooling all data together, as potential antagonistic effects could theoretically cancel each other out. We did this for the subset of data that could be paired, but do not consider these subsets of data to be sufficiently representative for the whole.

However, these were the data available, and the data set is much larger than those data sets on which previous conclusions regarding differences or lack of differences in sensitivity between wild species and crop species had been based. It is expected that if there had been any consistent difference in sensitivity between wild plant species and crop species, this should have been detectable even from a data set as heterogeneous as this one.

To further assess the power of this analysis, and taking into account the fundamental problems of calculating an overall MDD, (not just MDD-estimates for individual predictor combinations), modified datasets were assessed in addition. The calculation of modified data sets resulted in outcomes fully in line with the findings applying the MDD approach. The MDD approach also confirmed the differences between the ER10, ER25 and the ER50 datasets (which based on paired datasets were fundamentally equivalent in terms of their ability to detect differences between crop and wild species, see 4.4.1, p. 41), which however varied by chance, so that based on ER25 values crops appeared to be more sensitive than wild species, while based on ER10 and ER50 values the difference in sensitivity was the other way round. Considering both the MDD estimates based on the datasets available and the outcome of analysis of modified data sets, it is apparent that if wild species endpoints had generally been lower by a factor of around 2 or more than the wild species' endpoints, the difference would have been detected as statistically significant, no matter which subset of data (only ER10, ER25 or ER50) had been used. Based on the pooled analysis and considering the observed distribution and scatter of the data, any intrinsic difference in sensitivity between crop species and wild species as low as a factor of 1.5 would have been detected as statistically significantly different.

5.10 Problems caused by censored endpoints

Censored endpoints ("less-than" or "greater-than") are not strictly numeric but define a range that is only defined at one side. It is problematic to include them into numeric evaluations e.g. SSD, in particular when the censored values are not the lowest or highest, but are framed by higher or lower numeric values (e.g. Maltby et al. 2010, Brock et al. 2011, Christl 2013, EFSA 2013, Giddings et al. 2010, Giddings 2011). As the actual effect threshold level is not known, the exact position of these censored endpoints in a distribution, e.g. visualized as a SSD curve, is not either. If censored values were included as numeric endpoints anyway, this would involve a judgemental decision i.e. between which numeric values to position the censored value. This positioning would however affect the outcome, i.e. the overall centre of the distribution as well as its tails.

Sometimes authors evaluate censored endpoints numerically anyway[39], but generally without discussing the matter in depth. Some guidance is given in EFSA 2013, where as a rule it is proposed not to use censored values in e.g. an SSD. It is proposed to deviate from this rule only if censored values are outside the range of already available values, e.g. a lower-than value is lower than the lowest toxicity endpoint (EFSA 2013), so including it would expand the range despite its qualitative character.

In terms of censored values in this project, and considering data sets here, there are fundamentally three different data situations, depending on the presence or absence of numeric endpoints of the same substance-species-combination, or at least of the same active-substance-taxon group.

I. Censored values that were the <u>only data for a given active-substance-taxon group</u> (e.g. wild species 'Substance X'). If these could have been included somehow (e.g. implementing any Tobit model, see Amemiya 1984, or using Kon Kam King 2014), the number of calculable quotients would have gone up, (but their reliability would have gone down).

II. Censored values that were <u>one of several of a given active-substance-taxon group</u>, but the only data for a given active-substance-species combination, so other wild species endpoints for that substance did exist, and were numeric. If we could have included the non-numeric ones here, the number of calculable quotients would have remained the same, only the number of species contributing to that quotient would have gone up (and the quotient itself could have changed to some extent, but see below).

III. Censored values that only <u>complement one or several numeric endpoints</u> from other studies of the same substance-species-combination. Including these would not affect number of quotients and also not the number of species contributing to the quotient of this a.s., but only the number of studies contributing to that specific substance-species-combination, and the geometric mean of the substance-species combination could change slightly (the numeric values would still contribute to it).

---

[39] It is even suggested again in a very recent EFSA (2015) opinion paper on NTAs to treat censored values as point values, which is considered to be questionable; but see also DuBoudin et al. 2004.

For II. and III. it also matters whether any censored value is inside the range of numeric endpoints or outside, and if outside, whether “at the expected side” of the range. Depending on this the censored value may just qualitatively confirm the dose-response or sensitivity distribution specified by the numeric endpoints, or distinctly modify it as it is outside of the range and in the direction that would extend the range of numeric endpoints.

To conclude: Censored values are not just censored values, but depending on the data situations – with presence or absence of supporting numeric endpoints, censored endpoints would have to be treated differently. To decide which belong to which group would require reassessing all “active-substance-test-system-endpoint-species combinations”. It is by no means trivial to consider ‘greater-thans’ and ‘less-thans’ quantitatively, which is why most regulators are happy to discard them (unless them being a less-than lower than all numeric endpoints, in which case the regulator still would not consider the less-than endpoint as such, but would request numeric data to be generated).

It is appreciated that there are currently new proposals to make better use of censored data in SSDs, and a user interface was developed by Kon Kam King et al. (2014). However, the authors discuss that there are many ways to censor data and [that there was] no trivial way to choose between them as well (Kon Kam King et al. 2014). Anyway, the tool they developed might be suitable to make use of further data in future projects, e.g. by calculating an individual SSD for each a.s.-experimental endpoint-type-group (crop/wild), and to use the HC50 as central estimator and the HC5 or HC10 as estimator for the most sensitive species.

However, when evaluation of this dataset started Kon Kam King’s tool was not yet available, and was not implemented in this assessment. However we ran some assessments including greater-than-values in the other project, comparing vegetative and reproductive endpoints (Christl 2017, in preparation). A simplistic inclusion of censored endpoints was also implemented in this project, considering any censored values with a correction factor of 2, i.e. greater-than endpoints were doubled and less-than endpoints halved, see Figure 3 B. We did not expect that including any censored experimental endpoints would have changed the outcome significantly; and a comparison of Figures 3A and 3B supports that the overall outcome regarding sensitivity of crop and wild plant species was the same irrespective of inclusion or exclusion of censored endpoints.

Ultimately it was therefore decided to present numeric assessment only based on numeric endpoints, no matter which direction and in which group[40] they occurred.

Statistical analysis of distributions considering censored values

The additional statistical analysis performed by John W. Green addressed the concern of some reviewers that with ignoring censored values a part of the database’s information may not be utilized. While the conclusions based on this assessment differ in detail from the outcome based on the multiple regression approach (based on numeric endpoints only) also

[40] Of all experimental endpoints, 30.0% were censored; 30.4% of crop endpoints and 29.6% of wild endpoints. So if ignoring censored endpoints should have introduced any bias, it would be introduced to both groups similarly.

the additional statistical analysis did not find consistent differences in sensitivity between crop species and wild plant species.

5.11 Germinability of crop and wild species

The authors of EFSA (2014) appreciate that the crop species selected for inclusion in phytotoxicity testing are generally species with consistent and reliable rates of germination. However they also claim that non-crop species had been tested successfully (OECD, 2006a, b; US EPA, 2012a), and that many species had shown to be easy to manipulate and to yield uniform germination.

Numerous phytotoxicity studies have successfully been conducted using non-crop plants (Boutin et al., 2000, 2004; Riemens et al., 2008, 2009; Carpenter and Boutin, 2010; Strandberg et al., 2012; Carpenter et al., 2013). However, in most cases, germination characteristics were not tested (e.g. EFSA 2014, citations above). White et al. 2009 found that of 29 wild species 23 (almost 80%) reached 70 % germination. However, this was rather the exception than the rule. Cole (1993) tested germination of crop- and non-crop plants. All crop species but only six (of 22) non-crop species reached the 70 % threshold. The authors of EFSA (2014) believe that the conditions provided to non-crops for germination were inadequate; e.g. most non-crop species need light to germinate and sometimes some stratification. Pallet et al. (2007) tested the emergence of six non-crop species, and except for *Ipomoea hederacea* (L.), little germination occurred. The authors of EFSA (2014) observed that sowing depths ranged from 2 mm to 10 mm and suggest that this could have caused the low germination rates.

Anyway, non-crop species are sometimes being tested in a regulatory context, e.g. if there is indication of particular sensitivity of particular species (e.g. from efficacy test) or when the lowest endpoints obtained with technical material were from non-crop species. In the latter case it is at least attempted to perform also all formulation tests with these species. However testing such non-standard species is often unsuccessful, in particular the seedling emergence study is extremely problematic, despite advances in knowledge. Even with vegetative vigour studies performed on wild species, frequently only the third or fourth run results in formally valid study results under GLP (own experience, pers. communication with German Study Directors of GLP-laboratories); there was also one case where all attempts to generate seedling emergence data with a particular wild species were futile. Testing wild species for regulatory purposes is not impossible, but challenging for the performing laboratories and limited to some species with good and relatively homogenous germination (e.g. some strains of *Arabidopsis*). It is however not helpful for standard ecotoxicity testing if seeds germinate within a range of up to 6 weeks and seedlings are mixed with already mature plants. These problems are considered to be caused by specific traits of most annual wild herbs as follows.

Fundamentally annual crop species and wild annual herb species should follow the same reproduction strategies. This is however only true for the wild forms of crop species that served as source for the agricultural varieties. Most of the wild forms share with wild annual herbs that by means of seed dormancy or delayed germination they form a seed bank. These traits are seen as an adaptation to deal with environmental uncertainty (see e.g. Gardarin & Colbach 2014, Grime et al. 2007, Strassburger, E. 1998, Thompson, 1987, 2000, Rees 1994). Any trait that prevents instant germination is however unwanted for crop species, so has been selected away over many centuries and now is very rare in all culture varieties. This of course is a fundamental difference to reproduction strategies of wild plant species. There are a few exceptions among wild plant species. E.g. *Arabidopsis thaliana* is a special case, as there is extensive natural variation in the trait of delayed germination (e.g. Koornneef et al. 2004, Leubner 2015), i.e. some varieties do not produce many dormant seeds, so are similar to agricultural crops in this respect. Such varieties would offer themselves to be tested routinely[41]. However, these appear to be only a small fraction of common wild herbaceous species, and it remains to be assessed if these would by any means be more representative of the exposed non-target plants than the crop species currently used for standard glasshouse testing. Most wild plant varieties of this reproductive type do produce dormant seeds and exhibit delayed germination, which is considered to be a major disadvantage for routine testing.

While numerous crop endpoints were tested in seedling emergence studies, only few wild species endpoints are available. In the data assessed here (including both published papers and regulatory studies, both lab/greenhouse and field studies) there were 33 crop ER10, 139 crop ER25 and 242 crop ER50 endpoints (including censored endpoints) from seedling emergence studies, but only 3 ER10, 45 ER25 and 42 ER50 of wild species, respectively.

When collating the data for this review it was also observed that there is hardly any field data on seedling emergence / germination, whereas vegetative vigour endpoints are frequently measured also in the field. The authors suspect that one of the reasons is methodical obstacles (see previous paragraphs), the other that there would not be any reliable and representative control-plot situation. Even if seed densities were known, there was no reliable control germination rate that could serve as a reference rate to base any comparison with treated groups on.

5.12 Further aspects potentially relevant for the protection goal but not assessed in this review

There is only little information on sensitivity of wooden species in the literature. EFSA (2014) cite a few, including Strandberg et al. (2012), and “several studies conducted in Canada and Denmark have shown that there is no significant difference between the sensitivity of short- and long-lived species in terms of intrinsic sensitivity (Boutin et al., 2004; White et al., 2007; Carpenter and Boutin, 2010; Boutin et al., 2012; Carpenter et al., 2013)” taken from EFSA (2014). However, based on their longevity and spatial requirements, it can be assumed that only a fraction of the population, or even only a part of an individual, will be exposed to

[41] Note that handling of A. thaliana is still challenging because the seeds are tiny. Errors may easily occur by losing single seeds or erroneously seeding more than intended, then once seeds touch the soil they may no longer be detectable in the pots, i.e. double checking is aggravated. Testing such species under GLP is thus problematic.

herbicides drifting from a treated area. In addition, the longevity increases the potential for recovery over time. Populations of such species are probably much more dependent on other factors, e.g. habitat availability and density of larger herbivorous.

Even less is known about effects on cryptogam species, e.g. ferns or mosses. Considering the paucity of data, in this review no attempt was made to cover ferns, mosses. liverworts, hornworts, horsetails, and lichens. Any conclusion from this review thus only covers higher herbaceous plants and is in line with the decision of the SETAC Workshop on terrestrial plants held in Wageningen in April 2014.

Differences in sensitivity between different crop varieties and different wild plant species ecotypes are also discussed in EFSA (2014). The authors state that disparity in herbicide susceptibility among crop cultivars and wild species ecotypes had been confirmed in a number of studies, but no conclusions could be drawn on how to select appropriate ecotypes, varieties or cultivars. However, the current evaluation takes data from many different laboratories and countries into account. It can therefore be assumed that there were different varieties (ecotypes, cultivars) used in the many different testing locations. Thus in a certain way this question is included in the overall assessment in the current document without assessing all the different species varieties in detail.

While there may be difference in sensitivity between different ecotypes or cultivars (e.g. Boutin et al. (2010)), these authors found that seasonal fluctuations had a larger impact on sensitivity than differences in terms of ecotypes of the tested species (Boutin et al. 2010). Other than that we are not aware of any systematic testing of different ecotypes or cultivars in parallel (under otherwise identical test conditions), which would be the only way to rank cultivars in terms of sensitivity (to any particular a.s). In reality test varieties are chosen by the test facilities mainly based on practicability, i.e. cultivars for which seeds can be obtained reliably and by different suppliers, and that in previous experiments have proven to produce valid data. I.e. they must germinate readily, grow well under test conditions, and show little variation in growth; anything different would increase the risk of having to repeat studies.

Most of the quotients between wild and crop species were within the factor of 5 that is used as assessment factor in the RA (e.g. SANCO/10329/2002). Considering the wide interval of test rates in which endpoints were observed (sometimes exceeding three orders of magnitude) it is not surprising that there were also some cases where the quotient was larger than 5 or smaller than 0.2. Only in 4 (out of 56) the differences between crop and wild species endpoints[42] differed by a factor of greater than 5 *and* the two groups were also significantly different (see rightmost column in Table 3, bold entries). Of these four cases only one was clearly above 10, the other three were only marginally outside of the range covered by an assessment factor of 5 (factor 5.6 to 6.5). So we consider that even though there were some deviations beyond the factor of 5 they do not indicate that there is any fundamental difference between crops' and wild species' sensitivity that would justify additional testing requirements.

[42] based on the central point of the groups, i,e, the geometric mean of the endpoints.

Last but not least it is noted here that requirements for terrestrial non-target plant to test 6 to 10 species is exceptional in the ecotoxicological risk assessment. For any other organism group only endpoints of one to two standard species are required, and only for one further group (non-target arthropods) further species testing might be triggered by the results of the initial tests (e.g. SANCO/3268/2001, SANCO/10329/2002, EFSA 2013, US-Senate 2012). Hence within the entire Ecotoxicological Section the group with the most robust data situation is terrestrial non-target plants; the risk assessment is based on at least 36 endpoints (seedling emergence and vegetative vigour, six to ten species each, minimum observations in each study are biomass, survival, phytotoxicity rating). This allows a very detailed characterisation of the toxicity of a product to terrestrial plants and should therefore be accounted for in the risk assessment procedures.

5.13 Deficiencies

The largest drawback of the assessment is the heterogeneity of data, see discussion in 5.6. The ideal dataset would consist of an array with paired tests, only one parameter changed at a time. The actual dataset was clustered, with pronounced aggregations of data points for certain predictors, and few or even no data points for other combinations of predictors.

There is additional uncertainty due to having assessed lab/greenhouse and field data simultaneously. However, direct comparison of endpoints obtained from both revealed that differences in sensitivity due to the different test systems were less pronounced than expected. While it would be good to generate and assess further data to confirm these findings, based on the available endpoints no significant differences between the two were found, despite marked differences in exposure growth conditions, growth rates etc.

Availability of different endpoint types: Seedling emergence or vegetative vigour biomass endpoint; either of seedling emergence or of vegetative vigour studies; shoot height, root development, reproductive endpoints etc.

As there were e.g. considerably fewer data sets from seedling emergence studies available, and hardly any of these testing wild species, calculating quotients based on seedling emergence endpoints alone was possible only for few active substances.

However, considering that, if based on one endpoint (vegetative vigour, biomass; the endpoint for which by far most data were available), there were no consistent differences in sensitivity, it is regarded extremely unlikely that based on a different endpoint the outcome would be fundamentally different. While we agree that the vegetative vigour endpoint is not always the lowest, these exceptions are expected to occur in both groups (crop and wild) so are very likely to cancel each other out.

5.14 Outlook

In this study we assessed and tried to disprove the null-hypothesis that there are no inherent differences in sensitivity between wild plant species crop species. A fundamental problem in the concept of hypothesis testing is however that a negative cannot be proven; “one can never prove the nonexistence of something” (Wouters 2014). Eighty years ago Fisher (1935) already pointed out while defining the term “Null-hypothesis” that “it should be noted that the null hypothesis is never proved or established, but is possibly disproved, in the course of experimentation. Every experiment may be said to exit only in order to give the facts a chance of disproving the null-hypothesis.” (Fisher 1935/1971).

We considered that increasing the data base should increase the chance of the facts to disprove the null-hypothesis. So if - despite the extent of data - the null-hypothesis could not be disproved, it must be regarded unlikely that any further data should prove otherwise, even though we are aware of the fact that the heterogeneity of the data assessed here is not ideal. Still, based on this large database, there is no compelling evidence for crop and wild species endpoints being fundamentally different.

There is data that was generated explicitly to address similar questions – albeit on a smaller scale and for much fewer active substances / fewer species, but with the advantage of optimised experimental designs, generating homogeneous data sets, in paired designs etc. (e.g. Allison et al. 2013, Boutin et al. (2010), Boutin et al. 2012, Carpenter and Boutin, 2010, Clark et al. 2004, Egan et al. 2014a, Strandberg et al. 2012, White and Boutin, 2007). Claims that there were significant differences in sensitivity between wild plant species and crop species, with the former being more sensitive, seem to be based on few and not necessarily representative reviews. (Boutin and Rogers 2000, Davy et al. 2001, EFSA 2014, US-EPA database, Schmitz et al. 2013b[43]) in which even different endpoint types were treated as equivalent. We therefore conclude that – considering both published data and regulatory data generated under GLP and assessed by regulatory authorities – there is no sufficiently strong indication for any inherent difference in sensitivity between wild plant species crop species. Furthermore considering the practical difficulties encountered when testing wild species in regulatory studies, the slim chance that this assessment might be proven wrong at some stage seems not to justify the considerable additional effort for testing wild plant species instead of the current standard test species.

[43] In EFSA 2014 the unpublished report of Schmitz et al. 2013b is claimed to be available at www. Umwelt Bundesamt.de, however it could not be retrieved, and also was not made available as yet.

6 Conclusion

In this study published data and confidential data provided by the Chemical industry were combined to test – comparing like with like – the hypothesis that there may be consistent differences in sensitivity between wild non-target plant species and crop species.

The overall finding was that there were no consistent differences in sensitivity between wild plant species and crop species. There was however a trend that based on ER10, ER25 and ER50 (endpoints of vegetative vigour-like lab- and field studies) crop species were overall slightly more sensitive than wild species, but the difference was insignificant. Exceptions do exist, but there was no correlation between these exceptions of the rule and certain modes of action. Also the power of this evaluation was assessed. Based on the dataset above, considering its heterogeneity and variability it was demonstrated that any difference between the two groups would have been detected as significant if the two group's endpoints had deviated on average by a factor of 1.5 or more.

It can thus be concluded that for the taxonomic groups for which data were available there is no consistent difference between crop species and wild plant species. Testing crop species as model organisms and surrogates in standard toxicity tests seem thus to be a conservative approach, and there seems to be little reason for including further wild species in standard ecotoxicity testing.

7 References

(Published papers, the regulatory studies under data protection are not listed, neither are regulatory monographs, DARs or Conclusion Reports.)

Amemiya, T. 1984. Tobit models: A survey. Journal of Econometrics 24 (1–2): 3–61.

Al-Khatib K, Claassen MM, Stahlman PW, Geier PW, Regehr DL, Duncan SR, Heer WF. 2003. Grain sorghum response to simulated drift from glufosinate, glyphosate, imazethapyr, and sethoxydim. Weed Technology 17:261-265.

Al-Khatib K, Tamhane A. 1999. Dry Pea (Pisum sativum L.) Response to Low Rates of Selected Foliar- and Soil-Applied Sulfonylurea and Growth Regulator Herbicides. Weed Technology 13 (4): 753-758

Allison, J. E.; Boutin, C.; Carpenter, D. 2013. Influence of soil organic matter on the sensitivity of selected wild and crop species to common herbicides. Ecotoxicology 22(8): 1289-1302 ISSN: 0963-9292

Ahrens 1994 & Hatzios 1998 WSSA Herbicide Handbook, Weed Science Society of America, 1994 and Supplement 1998.

Anderson, R. L. 1990. Tolerance of safflower (Carthamus tinctorius), corn (Zea mays), and proso millet (Panicum miliaceum) to clomazone. Weed Technology 4 (3): 606-611. ISSN: 0890-037X

Anonymous 2001 Guidance Document on Aquatic Ecotoxicology in the context of the Directive 91/414/EEC. Sanco/3268/2001 rev.4 (final) 17 October 2002

Anonymous 2002 Guidance Document on Terrestrial Ecotoxicology Under Council Directive 91/414/EEC. SANCO/10329/2002 rev 2 final, 17 October 2002

Anonymous (2014 ff) Herbicide Resistance Action Committee (HRAC). (Herbicide resistance management) http://www.hracglobal.com/

Ashton, F. M.; Crafts, A. S. 1973 Mode of action of herbicides. Wiley, New York

Bewick, T. A., Stall. M., Kostewicz S. R., Smith K. 1991. Alternatives for control of paraquat tolerant American black nightshade (Solanum americanum). Weed Technology 5: 61-65

Bhatti M. A., Al-Khatib K., Felsot A.S. 1995. Effects of simulated Chlorsulfuron drift on fruit yield and quality of sweet cherries (Prunus aviuml.). Environmental Toxicology and Chemestry 14 (3): 537-544.

Bhatti M. A., Al-Khatib K., Parker R. 1996. Wine Grape (Vitis vinifera) Response to Repeated Exposure of Selected Sulfonylurea Herbicides and 2,4-D. Weed Technology 10 (4): 951-956

Bidelspach C, Olszyk, D, Pfleeger, T. 2008 Can Artificial Soil be Used in the Vegetative Vigor Test for US Pesticide Registration? Integrated Environmental Assessment and Management 4 (4) pp. 409–415.

Birnie, J. E. 1984. A preliminary study on the effect of some agricultural herbicides on a range of field margin flora. : Technical Report, AFRC Weed Research Organization . (1984 (79) :24pp. (

Bilz M, Kell SP, Maxted N, Lansdown RV 2011. European Red List of Vascular Plants. Luxembourg: Publications Office of the European Union, 144 pp. http://bookshop.europa.eu or IUCN Publications Services, www.iucn.org/publications

Blake, Robin J.; Westbury, Duncan B.; Woodcock, Ben A.; Sutton, Peter; Potts, Simon G. 2012. Investigating the phytotoxicity of the graminicide fluazifop-P-butyl against native UK wildflower species. : Pest Management Science . (2012) : 68 (3) :412-421. : 1526-498X

Blakeley-Smith M. 2007. Performance of Willamette Valley native plants following herbicide exposure. Master of Science thesis submitted to Oregon State University.

Boatman ND, Crocker DR, Hart A, Roelofs W, Smith RM Holland JM, Lutman PJW, Brown VK Ramsay A. 2004. Development of a scheme for the assessment of risks to wider biodiversity arising from the use of pesticides. PS2403 Final report to Defra Centre for Agro-Ecological Research, University of Reading

Boutin C, Aya KL, Carpenter D, Thomas PJ, Rowland O. 2012. Phytotoxicity testing for herbicide regulation: Shortcomings in relation to biodiversity and ecosystem services in agrarian systems. Science of the Total Environment 415:79-92.

Boutin C, Aya KL, Carpenter D, Thomas PJ, Rowland O. 2012. Phytotoxicity testing for herbicide regulation: Shortcomings in relation to biodiversity and ecosystem services in agrarian systems. Science of the Total Environment 415:79-92.

Boutin C, Elmegaard N, Kjaer C. 2004. Toxicity testing of fifteen non-crop plant species with six herbicides in a greenhouse experiment: Implications for risk assessment. Ecotoxicology 13:349-369.

Boutin C, Freemark KE, Keddy, CJ 1995 Overview and rationale for developing regulatory guidelines for nontarget plant testing with chemical pesticides Environmental Toxicology and Chemistry 14 (9) pp. 1465-1475,

Boutin C, Rogers CA. 2000. Pattern of sensitivity of plant species to various herbicides - An analysis with two databases. Ecotoxicology 9:255-271.

Boutin C, White AL, Carpenter D. 2010 (acc. 2009). Measuring variability in phytotoxicity testing using crop and wild plant species. Environmental Toxicology and Chemistry 29:327-337.

Boutin C, Strandberg B, Carpenter D, Mathiassen SK and Thomas P, 2014. Herbicide impact on non-target plant reproduction: what are the toxicological and ecological implications? Environmental Pollution, 185, 295–306.

Boutin, C. Strandberg B. Carpenter D. Mathiassen S K. Thomassen P.J. 2014 Herbicide impact on non-target plant reproduction: What are the toxicological and ecological implications? Env. Poll. 185, p. 295–306. DOI: 10.1016/j.envpol.2013.10.009

Brock, T.C.M.; Arts, G.H.P.; Maltby, L.; Van den Brink, P.J. (2006): Aquatic risks of pesticides, ecological protection goals and common aims in European Union Legislation. SETAC Environmental Assessment and Management, Vol. 2 (4) p. 20-46.

Brock, TCM.; Alix, A. Brown, CD. Capri, E. Gottesbüren, BFF. Heimbach, F., Lythgo, CM. Schulz, R. / Streloke, M. (2009/2010) Linking Aquatic Exposure and Effects Risk Assessment of Pesticides. SETAC / CRC Press (Taylor & Francis group, 410 pp.)

Brock TCM, Arts GHP, Ten Hulscher TEM, De Jong FMW, Luttik R, Roex EWM, Smit CE and Van Vliet PJM, 2011. Aquatic effect assessment for plant protection products: a Dutch proposal that addresses the requirements of the Plant Protection Product Regulation and Water Framework Directive. Alterra Report No 2235

Carpenter G & Boutin C. 2010 Sublethal effects of the herbicide glufosinate ammonium on crops and wild plants: short-term effects compared to vegetative recovery and plant reproduction. Ecotoxicology (2010) 19:1322–1336

Carpenter G & Boutin C, Allison J. E. 2013. Effects of Chlorimuron ethyl on terrestrial and wetland plants: Levels of, and time to recovery following sublethal exposure. Environmental pollution 172: 275-282.

Christl H. 2013 Aquatic macrophytes - does their sensitivity depend on their systematic position (monocotyledons/dicotyledons) SETAC Europe 23rd Annual Meeting in Glasgow, (12-16 May, 2013) 339

Christl H. 2017 Literature review and analysis: Sensitivity of vegetative and reproductive terrestrial plant endpoints in context of 1107/2009, performed on behalf of the SETAC workshop on “Non-target terrestrial plants”

Clark J Ortego L S, Fairbrother A. 2004 Sources of variability in plant toxicity testing. Chemosphere 57 (2004) 1599–1612

Cole JFH, Canning L and Brown RA, 1993. Rationale for the choice of species in the regulatory testing of the effects of pesticides on terrestrial non-target plants. Proceedings Brighton Crop Protection Conference—Weeds, Brighton, UK, 22–25 November, 151–156.

Dalton RL, Boutin C. 2010. Comparison of the effects of glyphosate and atrazine herbicides on non-target plants grown single and in microcosms. Environmental Toxicology and Chemistry 29:2304-2315.

Damgaard, C, Mathiassen, SK Kudsk P 2008 Modeling effects of herbicide drift on the competitive interactions between weeds. Environmental Toxicology and Chemistry 27 (6): 1302-1308.

Davy M. Petrie,R. Smrchek, J. Kuchniki, T. Francois, D. 2001 Scientific Advisory Panel Briefing - Proposal to update NTP testing under NAFTA. Joint presentation by Health Canada – PMRA, and US-EPA- Office of Prevention, Pesticides and Toxic Substances.

Das T.K. 2001. Towards better appraisal of herbicide bio-efficacy. Indian Journal of Agricultural science 71 (10): 676-678

De Snoo GR, Tamis WLM, van de Brink PJ, van der Poll RJ. 2005. Non target plant field study: effects of glufosinate-ammonium on off crop vegetation. Leiden University, Institute of Environmental Sciences (CML), Leiden.

DuBoudin, C., P. Ciffroy, H. Magaud 2004. Effects of Data Manipulation and Statistical Methods on Species Sensitivity Distributions, Environmental Toxicology and Chemistry 23, 489-499.

EFSA 2005 Opinion of the Scientific Panel on Plant health, Plant protection products and their Residues on a request from EFSA related to the assessment of the acute and chronic risk to aquatic organisms with regard to the possibility of lowering the uncertainty factor if additional species were tested. (Question N° EFSA-Q-2005-042), adopted on 14 December 2005. The EFSA Journal (2005) 301, 1-45

EFSA 2010. Scientific Opinion on the development of specific protection goal options for environmental risk assessment of pesticides, in particular in relation to the revision of the Guidance Documents on Aquatic and Terrestrial Ecotoxicology (SANCO/3268/2001 and SANCO/10329/2002). EFSA Journal 2010;8(10):1821, 55 pp. doi:10.2903/j.efsa.2010.1821

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2013. Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. EFSA Journal 2013;11(7):3290, 268 pp. doi:10.2903/j.efsa.2013.3290.

EFSA 2014. Scientific Opinion addressing the state of the science on risk assessment of plant protection products for non-target terrestrial plants. EFSA Journal 2014;12(7):3800, 163 pp. doi:10.2903/j.efsa.2014.3800

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues) , 2015. Scientific Opinion addressing the state of the science on risk assessment of plant protection products for non-target arthropods. EFSA Journal 2015;13(2): 3996, 212 pp. doi:10.2903/j.efsa.2015.3996.

Egan JF, Bohnenblust E, Gosleec S, Mortensena D, Tooker J 2014a. Herbicide drift can affect plant and arthropod communities. Agriculture, Ecosystems and Environment 185 (2014) 77– 87.

Egan JF, Graham IM, Mortensen DA 2014b A comparison of the herbicide tolerances of rare and common plants. Environmental Toxicology and Chemistry, Vol. 33, No. 3, pp. 696–702.

Everitt D., Keeling J. W. 2009. Cotton growth and yield response to stimulated 2,4-D and Dicamba drift. Weed Technology 2009 23: 503-506

FIFRA (Eds. Lewis, P. Matsumura, F. Odiott, O. Roberts, S.M.) 2001 Review of Non-Target Plant Toxicity Tests Under the North American Free Trade Agreement (NAFTA) SAP Report No. 2001-08, 51 pp.

Fisher, R.A. (1935) /1971 (9th ed.) The Design of Experiments. (Oliver and Boyd, Edinburgh, UK, 1935) / Hafner Publishing Company New York 1971.

Fletcher JS, Johnson FL, McFarlane JC. 1990. Influence of greenhouse versus field testing and taxonomic differences on plant sensitivity to chemical treatment. Environmental Toxicology and Chemistry 9:769-776.

Fletcher JS, Muhitch MJ, Vann DR, McFarlane JC, Benenati FE. 1985. Phytotox database evaluation of surrogate plant species recommended by the U.S. environmental protection agency and the organization for economic cooperation and development. Environmental Toxicology and Chemistry 4:523-532.

Fletcher J S Pfleeger T G Ratsch H C Hayes R 1996 Potential impact of low levels of chlorsulfuron and other herbicides on growth and yield of non-target plants. Environmental Toxicology and Chemistry, 15: 1189–1196.

Francois T, D. 2001 Scientific Advisory Panel Briefing - Proposal to update NTP testing under NAFTA. Joint presentation by Health Canada – PMRA, and US-EPA- Office of Prevention, Pesticides and Toxic Substances.

Gardarin A. & Colbach N. 2014 How much of seed dormancy in weeds can be related to seed traits? European Weed Research Society 55, 14–25.

Gealy D. R., Boerboom C. M., Ogg A. G., Jr. 1995. Growth and yield of Pea (Pisum sativum L.) and Lentil (Lens culinaris L.) sprayed with low rates of Sulfonylurea and Phenoxy herbicides. Weed Science 43 (4): 640-647.

Geisthardt, M. (2012): Effekte von Herbiziden auf phytophage Insekten am Beispiel der Kohleule Mamestra brassicae. Diplomarbeit, Institut für Umweltwissenschaften. Universität Koblenz-Landau. Campus Landau.

Giddings JM. 2011, The Relative Sensitivity of Macrophyte and Algal Species to Herbicides and Fungicides: An Analysis Using Species Sensitivity Distributions; for the Species Sensitivity Distribution Working Group (SSD WG), functioning under the umbrella of the Aquatic Macrophyte Ecotoxicology Group (AMEG). Society of Environmental Toxicology and Chemistry (SETAC).

Giddings, J. Loutseti S. Arts, G. Cedergreen, N. Christl H. Davies , J. Dobbs M. Hanson, M. Hommen U. Honegger J. Manson P. Meregalli G. Weyman G. (2010) Relative

Sensitivity of Macrophyte & Algal Species to Herbicides & Fungicides using Species Sensitivity Distributions. SETAC Sevilla 2010 MOPC2-4.

Gilreath JP, Chase CA. 2001. Crop injury from sublethal rates of herbicide. I. Tomato. HortScience 36:669-673.

Gilreath, JP, Chase, CA, Locascio, SJ (2001). Crop injury from sublethal rates of herbicide. II. Cucumber. HortScience 36(4):674-676

Gilreath, JP, Chase, CA, Locascio, SJ (2001). Crop injury from sublethal rates of herbicide. III. Pepper. HortScience 36(4):677-681

Griffin J. L, Bauerle M. J, Stephenson D. O., III , Miller D. K., Boudreaux J. M. 2013. Soybean Response to Dicamba Applied at Vegetative and Reproductive Growth Stages. Weed Technology 27 (4): 696-703.

Grime, J.P. , Hodgson, J G., Hunt, R. 2007 Comparative plant ecology: a functional approach to common British species. Castlepoint Press, 2007 – 748 pp.

Hahn M, Lenhardt PP, Brühl, CA 2014, Characterization of Field Margins in Intensified Agro-Ecosystems—Why Narrow Margins Should Matter in Terrestrial Pesticide Risk Assessment and Management. Integrated Environmental Assessment and Management — Volume, Number (accepted) pp. 1–7

Helsel, DT 2012 Statistics for censored environmental data using minitab and r., John Wiley & Sons, Inc. 322 p.

Hemphill D. D., Jr., Montgomery M. L. 1981. Response of Vegetable Crops to Sublethal Application of 2,4-D. Weed Science 29 (6): 632-635.

Johnson E. N. 2015. Herbicides: Past, present, and future. Presentation based on paper: Holm F. A, Johnson E. N. 2009. The history of Herbicide use for weed management on the Prairies. Prairie Soils and Crops 2:1-10. [http://www.prairiesoilsandcrops.ca]

Keppel, G. (1991) (3rd ed.). Design and Analysis: A Researcher's Handbook. Englewood Cliffs, NJ: Prentice Hall

Kahn M., Donald W. W. 1992. Sulfonylurea Herbicides Reduce Survival and Seed Production of Green and Yellow Foxtails (Setaria spp.). Weed Technology 6 (2): 284-290.

Kjær C. 1994. Sublethal effects of chlorsulfuron on black bindweed (Polygonum convolvulus L.). Weed Research 34: 453-459.

Kjær C, FIFRA M, Erlandsen M. 2006b. Metsulfuron spray drift reduces fruit yield of hawthorn (Crataegus monogyna L.). Science of the Total Environment 356:228-234.

Kleijn D, Snoeijing GIJ. 1997. Field boundary vegetation and the effects of agrochemical drift: botanical change caused by low levels of herbicide and fertilizer. Journal of Applied Ecology 34:1413-1425.

Klimisch HJ, Andreae M. and Tillmann U (1997): A Systematic Approach for Evaluating the Quality of Experimental Toxicological and Ecotoxicological Data. Regulatory Toxicology and Pharmacology Vol 25, pp 1–5, doi:10.1006/rtph.1996.1076, PMID 9056496

Koch H, Weißer P, Strub O. 2004. Comparison of dose response of pesticide spray deposits versus drift deposits. Landesanstalt für Pflanzenbau und Pflanzenschutz, Mainz, 9 pp.

Kon Kam King, G, P Veber, S. Charles, and ML Delignette-Muller 2014. MOSAIC_SSD: a new web-tool for the Species Sensitivity Distribution, allowing to include censored data by maximum Likelihood. Environmental Toxicology and Chemistry 33: 2133–2139.

Koornneef M, Alonso-Blanco C, Vreugdenhil D. (2004) Naturally occurring genetic variation in Arabidopsis thaliana. Annu Rev Plant Biol. 2004;55:141-72.

Leubner G (2015). The Seed Biology Place - Website Gerhard Leubner Lab; Royal Holloway, University of London – http://www.seedbiology.de

Long J. S., Ervin L. H. (2000), Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. The American Statistician, 54, 217–224.

MacKinnon J. G., White H. (1985), Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties. Journal of Econometrics 29, 305–325.

Marrs RH, Frost AJ, Plant RA. 1991a. Effects of herbicide spray drift on selected species of nature conservation interest - the effects of plant-age and surrounding vegetation structure. Environmental Pollution 69:223-235.

Marrs RH, Frost AJ, Plant RA. 1991b. Effect of mecoprop drift on some plant-species of conservation interest when grown in standardized mixtures in microcosms. Environmental Pollution 73:25-42.

Marrs RH, Frost AJ, Plant RA, Lunnis P. 1993. Determination of buffer zones to protect seedlings of nontarget plants from effects of glyphosate spray drift. Agriculture Ecosystems & Environment 45:283-293.

Marrs RH, Williams CT, Frost AJ, Plant RA. 1989. Assessment of the effects of herbicide spray drift on a range of plant species of conservation interest. Environmental pollution (Barking, Essex : 1987) 59:71-86.

Maltby, L. Arnold, D. Arts, G. Davies, J. Heimbach, F. Pickl, C. Poulsen, V. 2010 Aquatic Macrophyte Risk Assessment for Pesticides. SETAC/CRC Press Taylour & Francis Group, Boca Raton, London New York, 135 pp.

Martin H. / Ontario Ministry of Agriculture & Food 2014 Herbicide Mode of Action Categories http://www.omafra.gov.on.ca/english/crops/facts/00-061.htm

Mathiassen SK (unpublished data), endpoints cited in EFSA 2014, no source given.

Mayer, B. Frickmann, J. Christensen, E.R. Nyholm, N. 1998 Influence of growth conditions on the results obtained in algal toxicity tests. Environmental Toxicology and Chemistry, Vol. 17, No. 6, pp. 1091–1098, 1998

McKelvey RA, Wright JP, Honegger JL. 2002. A comparison of crop and non-crop plants as sensitive indicator species for regulatory testing. Pest Management Science 58:1161-1174.

Menne H & Köcher H HRAC Classification of Herbicides and Resistance Development. In Modern Crop Protection Compounds. Edited by W. Krämer & U. Schirmer 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, p. 5 – 26.

Narens, L. (1981). A general theory of ratio scalability with remarks about the measurement-theoretic concept of meaningfulness. Theory and Decision, 13, 1–70

Newman MC, Ownby DR, Mezin LCA, Powell DC, Christensen TRL, Lerberg SB, Anderson BA. 2000. Applying species-sensitivity distributions in ecological risk assessment: Assumptions of distribution type and sufficient numbers of species. Environmental Toxicology and Chemistry 19:508-515.

Obrigawitch TT, Cook G, Wetherington J. 1998. Assessment of effects on non-target plants from sulfonylurea herbicides using field approaches. Pesticide Science 52:199-217.

OECD (Organisation for Economic Co-operation and Development), 2006a. Terrestrial plants test: seedling emergence and seedling growth test (No. 208). OECD Guidelines for Testing Chemicals. OECD, Paris, France.

OECD (Organisation for Economic Co-operation and Development), 2006b. Terrestrial plants test: seedling growth test (No. 208) and vegetative vigour test (No. 227). OECD Guidelines for Testing Chemicals. OECD, Paris, France.

Olszyk D, Blakeley-Smith M, Pfleeger T, Lee EH, Plocher M. 2013. Effects of low levels of herbicides on prairie species of the Willamette valley, Oregon. Environmental Toxicology and Chemistry 32 (11): 2542-2551.

Olszyk D, Pfleeger T, Lee EH, Burdick C, King G, Plocher M, Kern J. 2008. Selecting and evaluating native plants for region-specific phytotoxicity testing. Integrated environmental assessment and management 4:105-117.

Olszyk D, Pfleeger T, Lee EH, Plocher M. 2009. Pea (Pisum sativum) seed production as an assay for reproductive effects due to herbicides. Environmental Toxicology and Chemistry 28 (9): 1920-1929.

Olszyk, D.; Pfleeger, T.; Lee, E. H.; Plocher, M. 2010. Phytotoxicity assay for seed production using Brassica rapa L. : Integrated Environmental Assessment and Management (IEAM) . (2010) : 6 (4) :725-734. ISSN: 1551-3777

Olszyk DM, Burdick CA, Pfleeger TG, Lee EH, Watrud LS. 2004. Assessing the Risks to Non-Target Terrestrial Plants from Herbicides. Journal of Agricultural Meteorology 60:221-242.

Pallett K, Cole J, Oberwalder C, Porch J. 2007. Performance of potential non-crop or wild species under OECD 208 testing guideline study conditions for terrestrial non-target plants. Pest Management Science 63:134-140.

Pfleeger T, Blakeley-Smith M, King G, Henry Lee E, Plocher M, Olszyk D. 2012. The effects of glyphosate and aminopyralid on a multi-species plant field trial. Ecotoxicology:1-17.

Pfleeger T, Olszyk D, Lee EH and Plocher M 2011 Comparing effects of low levels of herbicides on greenhouse and field-grown potatoes (Solanum tuberosum L.) soybeans (Glycine max L.) and peas (Pisum sativum L.). Environmental Toxicology and Chemistry 30 (2), pp. 455–468.

Riemens MM, Dueck T, Kempenaar C. 2008. Predicting sublethal effects of herbicides on terrestrial non-crop plant species in the field from greenhouse data. Environmental Pollution 155:141-149.

Rees, M. 1994 Delayed germination of seeds: A look at the effects of adult longevity, the timing of reproduction, and population age structure. The American Naturalist, 144 (1) 43-64.

Reuter S. & Siemoneit-Gast, S. 2007 Entwicklung einer weiterführenden Methode zur Bewertung des Risikos für terrestrische Pflanzen durch Exposition mit Pflanzenschutzmitteln und ihren Wirkstoffen. Abschlussbericht (=Development of an advanced method to assess the risk to terrestrial plants from exposure to PPPs and their active substances – final report). UFOPLAN 202 67 43 (prepared for the German Federal Environment Agency).

Riemens MM, Dueck T, Kempenaar C. 2008. Predicting sublethal effects of herbicides on terrestrial non-crop plant species in the field from greenhouse data. Environmental Pollution 155:141-149.

Riemens M.M., Dueck T., Kempenaar C., Lotz A. P., Kropff M. J. J. 2009. Sublethal effects of herbicides on the biomass and seed production of terrestrial non-crop plant species, influenced by environment, development stage and assessment date. Environmental Pollution 159: 2306-2313.

Romanowski RR. 1980. Simulated drift studies with herbicides on field-grown tomato. HortScience 15: 793-794.

Ross MA Childs DJ 2000 Herbicide Mode-Of-Action Summary. Cooperative Extension Service Purdue University West Lafayette, Indianapolis https://www.extension.purdue.edu/extmedia/ws/ws-23-w.html

Schmidt, R. 1999 Classification of Herbicides According to Mode of Action, April 1999

Schmitz J, Schäfer K, Brühl CA. 2013a. Agrochemicals in field margins - Assessing the impacts of insecticides, herbicides and fertilizer on the common buttercup Ranunculus acris. Environmental Toxicology and Chemistry, Vol. 32, pp. 1124–1131.

Schmitz J, Stahlschmidt P and Brühl CA, 2013b. Protection of terrestrial non-target plant species in the regulation of environmental risks to pesticides. Report Gz. Z6 – 93104-4/10.Main Report. FKZ 360 03 053. (Gutachten erstellt im Auftrag des Umweltbundesamt, Dessau (see reference UBA 2012)), unpublished report. Acc. to EFSA 2014 available at www.umweltbundesamt.de (*Report could not be retrieved by the authors as yet*)

Schmitz J, Schäfer K, Brühl CA. 2014a. Agrochemicals in field margins - An experimental field study to assess the impacts of pesticides and fertilizers on a natural plant community. Agriculture, Ecosystems and Environment 193 (2014) 60–69.

Schmitz J, Schäfer K, Brühl CA. 2014b. Agrochemicals in field margins - Field evaluation of plant reproduction effects. Agriculture, Ecosystems and Environment 189 (2014) 82–91.

Sciumbato AS., Chandler JM, Senseman SA, Bovey RW, Smith KL. 2004. Determining Exposure to Auxin-Like Herbicides. I. Quantifying Injury to Cotton and Soybean. Weed Technology 18: 1125-1134.

Siemoneit-Gast, S.; Reuter, S.; Kubiak, R.; Höllrigl-Rosta, A. 2008. Development of an extended method for assessing the risk to terrestrial plants exposed to plant protection products and their active ingredients. In: Re, A. A. M. del; Capri, E.; Fragoulis, G.; Trevisan, M. (eds.): Environmental fate and ecological effects of pesticides 2007 649-656.

Smith JL, Burritt D.J. & Bannister, P 2000 Shoot Dry Weight, Chlorophyll and UV-B-absorbing Compounds as Indicators of a Plant's Sensitivity to UV-B Radiation. Annals of Botany 86: 1057-1063

Snoo GR de, Tamis WLM & Brink PJ van de 2005 Non target plant field study effects of glufosinate-ammonium on off crop vegetation. Institute of Environmental Sciences (CML), Leiden University.

Sokal RR, Rohlf FJ 1995 The Principles and Practice of Statistics in Biological Research. W. H. Freeman, 887 pp.

Spaunhorst D. J., Bradley K. W. 2013 . Influence of Dicamba and Dicamba plus Glyphosate Combinations on the Control of Glyphosate-Resistant Waterhemp (Amaranthus rudis). Weed Technology 27: 675-681.

Strandberg B, Mathiassen SK, Bruus M, Kjær C, Damgaard C, Andersen HV, Bossi R, Løfstrøm1 P, Larsen SE, Bak J, Kudsk P. 2012. Effects of herbicides on non-target plants: How do effects in standard plant test relate to effects in natural habitats? Danish Environmental Protection Agency.

Strandberg, B., Sørensen, P. B., Damgaard, C., Bruus, M., Strandberg, M., Navntoft, S., Nielsen, K. E. 2013 Indikatorer for biodiversitetsforbedringer i marknære småbiotoper ved etablering af sprøjtefri randzoner. Bekæmpelsesmiddelforskning fra Miljøstyrelsen nr. 149, 2013. Miljøstyrelsen/Miljøministeriet ISBN 978-87-92903-90-7

Straßburger, E. (1998): Handbook of Botany (Lehrbuch der Botanik). 34 ed., Gustav Fischer, Stuttgart, Jena, New York

Thompson, K. (1987) Seeds and seed banks. New Phytologist (Supplement) 106:23-34.

Thompson K, (2000) The Functional Ecology of soil seed banks. In: Fenner, M. (ed) 2000: Seeds: The Ecology of Regeneration in Plant Communities; Chapter 9. CAB International Publishing, Oxon - New York, 415 pp.

US-EPA (2004) Overview of the Ecological Risk Assessment Process (Environmental risks, regulated under FIFRA) 23 Jan 2004. http://www.epa.gov/espp/consultation/ecorisk-overview.pdf

US EPA (United States Environmental Protection Agency), 2012d. Background and special considerations—tests with terrestrial aquatic plants, cyanobacteria, and terrestrial soil-core microcosms. OCSPP 850.4000. EPA 712-C-013, Washington, DC.

US EPA 2012 Ecological Effects Test Guidelines: Seedling Emergence and Seedling Growth (OCSPP 850.4100), Early Seedling Growth Toxicity Test (OCSPP 850.4230), Vegetative Vigor (OCSPP 850.4150), Background and Special Considerations – Tests with Terrestrial, Aquatic Plants, Cyanobacteria, and Terrestrial Soil-core Microcosms (OCSPP 850.4000)

US-Senate (2012) Federal Insecticide, Fungicide, and Rodenticide Act (FIRFA) 7 U.S.C. §136 et seq. (1996 ff.) www2.epa.gov/laws-regulations/summary-federal-insecticide-fungicide-and-rodenticide-act

Vielhauer B, Siemoneit-Gast S, Velten G & Bruehl C 2009 SETAC GLB Poster. Universität Koblenz-Landau, Landau.

Vielhauer B. 2010. Auswirkungen von Pflanzenschutzmitteln auf terrestrische Nicht-Ziel-Pflanzen Sensitivitätsvergleich von nah verwandten Kultur- und Wildpflanzenarten gegenüber einem Breitbandherbizid. Diploma Thesis, Koblenz-Landau, Landau.

Wall D. A. 1994. Potato (Solanum tuberosum) response to stimulated drift of Dicamba, Clopyralid, and Tribenuron. Weed Science 42: 110-114.

West T M (1988) The activity and post-emergence selectivity of some recently developed herbicides: imazethapyr, BAS 51800H, DPX-L5300, triasulfuron, DPX-A7881. Technical Report IACR – Long Ashton Research Station, (104), pp 58.

West T M (1994) The pre- and post-emergence activity and selectivity of the herbicide amidosulfuron (HOE 075032). Technical Report IACR – Long Ashton Research Station, (110), pp 25.

West, T.M. (1988) The activity and post-emergence selectivity of some recently developed herbicides: imazethapyr, BAS 51800H, DPX-L5300, triasulfuron, DPX-A7881. Technical Report IACR – Long Ashton Research Station, (104), pp 58.

West, T.M. (1994) The pre- and post-emergence activity and selectivity of the herbicide amidosulfuron (HOE 075032).

White AL, Boutin C. 2007. Herbicidal effects on nontarget vegetation: Investigating the limitations of current pesticide registration guidelines. Environmental Toxicology and Chemistry 26: 2634-2643.

White, A. L.; Boutin, C.; Dalton, R. L.; Henkelman, B.; Carpenter, D.. (2009) Germination requirements for 29 terrestrial and wetland wild plant species appropriate for phytotoxicity testing. : Pest Management Science 2009: 65 (1) : 19-26. ISSN: 1526-498X

Wingender & Weddeling (2002) Die Bedeutung der landwirtschaftlichen Nutzung für die Vielfalt wildlebender Tiere und Pflanzen in Deutschland.

Wouters L 2014 Statistical Thinking and Smart Experimental Design. VIK, Vlaams Instituut Voor Biotechnologie, 99 pp.

WSSA 2010 ff Summary of Herbicide Mechanism of Action According to the Weed Science Society of America (WSSA): http://wssa.net/wssa/weed/

8 Appendix 1 - Brief summaries of the most relevant papers and their use for the present review

Below the papers considered in this assessment are briefly summarized, mainly to make transparent which papers contributed considerably to the database and which did not – while still being considered as relevant.

Allison et al. (2013) assessed the influence of soil organic matter on the sensitivity of selected wild and crop species to common herbicides. They tested a number of crop- and non-crop species, differentiating between species of high or low nitrogen affinity. Also they varied the organic matter (OM) in soil (low – standard 3% = 1.5% OC, and high = 9% OM = 4.5 %OC). They present the resulting ER25-values which were included in the data base generated for the present paper, but no ER50 values. According to the authors the observed variation within the results indicated that herbicide-species combinations may respond differently to changes in available nutrient levels. However there was no clear pattern, so no such statement as “Increased OM content increases sensitivity” or “Species with low nitrogen affinity are less affected by organic matter content” was possible. While the authors suggest that “regulatory guidelines may need to be adjusted to allow testing under soil nutrient conditions that are more reflective of natural environments near agricultural areas” (Allison et al. 2013), their own results would not indicate into which direction test conditions had to be shifted to make them more meaningful, and for which species group.

According to **Bilz et al. (2011)** exposure to pesticides appears to play only a minor role for the distribution of threatened European vascular plants (endangered species, see European Red List of Vascular Plants). This is in contrast to the perception of other scientists who recommend to expand NTP test requirements (see e.g. EFSA 2014) . In Bilz’ et al. 2011 conclusions and recommendations exposure to pesticides and pollution are not listed among the main threats to vascular plants. On the contrary the listed threats are intensified livestock farming, recreational activities, wild plant collection, urban development, ecosystem modifications, problematic native species and invasive alien species. The threat “pollution” which includes exposure to plant protection products appears on position 8 (of 18) when based on the number of non-threatened species, and on place 13 (of 18) when based on the number of threatened species. The threat ”Intensive arable farming[44]” is listed on position 10 (of 18) when based on the number of non-threatened species, and on place 15 (of 18) when based on number of threatened species (Bilz et al. 2011, Fig. 8).
This evaluation puts any concerns regarding adverse effects on vascular plants solely due to exposure to agrochemicals into context. Bilz et al. (2011) also stress that intensive arable farming (and pollution caused by it) is a multifactorial complex, and PPPs are only one factor of this complex.

[44] Adverse effects on endangered wild plant species from intensive arable farming is by no means limited to effects due to exposure to pesticides, but encompasses changes in habitat, mechanical disturbance, fertilisation and other procedures required according to good agricultural practice, but affecting wild herbs and weeds.

Birnie (1984) presented initial results from a test series with 17 wild herb and crop species kept in pots. Some were raised from seeds, others collected from field margins. The plants were sprayed by means of a teejet boomsprayer and received just one treatment dose, the recommended field rate, of seven commonly used herbicides. Effects were assessed semi-quantitatively (scoring system) so obviously no ERx values are available.
To use the data anyway, the scores (0 = "dead" to 9 = "indistinguishable from control") were interpreted as follows. Scores 3 to 4 (4: "slight inhibition of growth" to 3: "gross inhibition of growth") were considered to approximate an ER50, and the scores 5 to 6 (6: "slight growth differences, e.g. wilting, chlorosis", 5: "obvious growth defects, e.g. epinasty[45]") were considered to approximate an ER25. Scores higher or lower than these thresholds were interpreted as censored endpoints, higher-than or lower-than the test rate. In doing so at least a part of the info could be included as numeric values in the database, and pronounced sensitivities of wild herbs should be detectable, which might help the task of the present paper, i.e. to evaluate if there is a general trend between crop species and wild herb species in terms of sensitivity to herbicides. Birnie (1984) then concluded that effects were most pronounced on the target weeds (as to be expected), but that in a few cases herbicides were also effective against grass weeds not yet on the label. On the other hand none of the herbicides controlled sterile brome selectively, which according to the authors poses a major weed problem.

Blakeley-Smith M. (2007) investigated effects of herbicide spraydrift on native wild plants in the United States, selecting a natural reserve where plants are unlikely to ever having been exposed to herbicides. He observed relatively low endpoints. While this is not a published paper, it is online available, was supervised by experts known in their field of work, so these low wild endpoints were also included in our database.

Boatman et al. (2004) contributed to the development of a risk assessment scheme for wider biodiversity suitable for use in a regulatory context, elements of which have been considered in the current risk assessment schemes, including case studies and definition of protection goals. They discuss endpoints from crop species but protection of non-crop species, and provided different extrapolation approaches, depending on whether crop endpoints are seen as representative for non-crop endpoints or if not. There is however no statement if there are differences in sensitivity between crop species and wild species or if not. Effect levels at a given rate were provided and have been incorporated in the data base, but as no dose-response test design has been employed, most resulting endpoints in our database are censored (greater-than or less-than). We decided to consider cases where effect levels ranged between 15 and 35% as numeric ER25, and those that ranged between 40% and 60% effect compared to the controls as numeric ER50, respectively. The authors also discussed effects on reproduction and considered based on other papers' conclusions - that a biomass endpoint measured 2-4 weeks after treatment was unlikely to achieve a reliable estimate of seed production (Boatman et al. 1988); hence they advertise long-term testing, but also appreciate problems in terms of practicability.

[45] asymmetric growth of leaf blades, resulting in curling

Boutin & Rogers (2010) evaluated several databases but do not share the individual endpoints. They argue that due to the limitations of the database it is difficult to draw a definitive conclusion as to the number and type of species to be tested in a risk assessment scheme. Nonetheless it appears that grasses tend to respond in a similar way to various chemicals, thus the number of grasses tested can be minimized relative to broad-leaved species. According to the authors more broad-leaved species should be tested than the number currently requested in the U.S EPA guidelines, since with plants, as opposed to animals, an extensive database could be generated despite the extensive work required.

Boutin et al. (2004) present data of a greenhouse experiment where 15 non-crop plant species were sprayed with 6 herbicides. The dataset termed the "dataset (called thereafter Danish/Canadian" was then compared with data submitted to the US-EPA for registration. Unfortunately the latter included only the most sensitive endpoint (see Boutin et al. 2004, page 355). Quantitative endpoints were reported, e.g. dry weight, fresh weight, height, etc. but as it is not recorded which of these provided the lowest endpoint there is inevitably a deviation from the principle to compare only like with like (e.g. biomass DW with biomass DW). Reported endpoints were included in the database anyway, but results must therefore be interpreted with some caution.

Boutin et al. (2010) opine that realistically, only a few species can be used to represent the hundreds of species that have to be protected. The study revealed that some variation in sensitivity to herbicides existed among ecotypes of different plant species and that conclusions regarding the phytotoxicity of any given herbicide may differ depending on the ecotypes chosen for inclusion in risk assessment.
Both crop and wild plant species responded quite variably when they were tested in different seasons as well as when they were tested in a greenhouse or in growth chambers. Abiotic factors, such as temperature and light have to be taken into account in phytotoxicity testing even in greenhouses. The calculated effective doses indicated that seasonal fluctuations have a larger impact on their sensitivity than differences in terms of ecotypes of the tested species. So according to the results obtained it was considered useful to base the test duration after application on environmental variables such as hours of sunlight and/or hours of temperature above a certain threshold rather than just to define a fix number of days.). IC25 endpoints were reported and were incorporated in our database.

Boutin et al. (2012) combined data from different experiments in this paper, closing important knowledge gaps, i.e. how woody plants' sensitivity relates to that of herbaceous plants, effects on ferns, compared vegetative and reproductive endpoints, and investigated species sensitivity based on ecological traits i.e. live span, and whether crop or wild species. Endpoints were collated from tables or approximately determined from figures. While the level of detail reported for the individual experiments is limited (so Klimisch-score was not applicable), the endpoints were included in the database anyway, but results might be affected by unknown aspects.

Carpenter & Boutin (2010) assessed the ability of plant species to recover (biomass and reproduction) when exposed and tested at the juvenile stage (routine regulatory testing), comparing crop and wild species and using the herbicide glufosinate ammonium**.** Ten crop species (4 monocots + 6 dicots) and 10 wild species (4 monocots + 6 dicots) were tested under greenhouse conditions and IC50 values were provided and included in the present assessment.

Clark et al. (2004) investigated several databases with plant effect data for plant protection products, applying a correlative empirical approach to develop general principles for extrapolating results of standard bioassays to non-standard species or test conditions. The sensitivity of agricultural plants versus other species was investigated as well as the similarity of effects seen at different taxonomic levels. Generally they report only meta-data but not the actual endpoints of species and active-substance – combinations. The few ER50 endpoints that were reported (crop- and non-crop species) were included in our data base. The authors observed that genus and family taxonomic groupings generally show similar responses among species (irrespective of whether it is a crop- or a wild species), but less similarity was found when members of the same orders and classes were compared. Comparatively larger differences were detected between endpoints from field and greenhouse studies. Overall the authors concluded that “agricultural species are not consistently more or less sensitive to the herbicides tested than non-crop species” (Clark et al. 2004).

Dalton & Boutin (2010) tested single plants and assemblages of wetland- and “terrestrial” species in four different test container systems (termed “single species” (standard lab test), “greenhouse microcosms”, “long-term microcosm” and “outdoor microcosms”) on atrazine and glyphosate formulations. They found differences in sensitivity depending on the container used for testing, partly significantly different. However the differences they observed varied generally by a factor of 2 to 3 which is well within inter-lab variability. Also, although the authors advertise the multi-species tests as being superior and a more sensitive test systems, from the presented charts the data actually appear to be inconclusive. In 4 of 12 sets the single-species lab test generated the lowest endpoint of the series, in further 4 cases it was one of the two lower values, and in only two instances the single species lab test resulted in the highest endpoint (Dalton & Boutin 2010, Fig. 2). In fact there appears to be no unequivocal trend that would allow stating which test systems resulted in lower endpoints. Unfortunately the individual IC25 endpoints (by species) were not reported in the paper but only results of meta-analysis. However, one figure displays endpoints of six species from different test variants. As a surrogate and approximation the column chart Figure 3 “Comparison of 25% inhibition concentrations” was read and visually transformed into numeric figures, which of course adds considerable uncertainty (assumed to be around ±15 g/ha). However large deviations were clearly visible. These data were included in the database and put in relation with other endpoints reported for the tested formulations.

Egan et al. (2014 a) exposed replicated plots to low doses of dicamba or of 2,4 D designed to simulate herbicide drift and monitored changes in plant and arthropod communities. No ER25 or ER50 are reported, hence the results of the study could not be included in the data base for the overall assessment. The authors conclude that variability across sites and taxonomic groups makes it difficult to offer general conclusions about the risks of dicamba drift to plant and arthropod biodiversity.

Egan et al. (2014 b) state “If herbicides are a significant driver of changes in the diversity and composition of plant communities across agricultural landscapes, then we would expect rare species to show consistently lower tolerances to nonlethal exposures than common species with similar life histories. Contrary to this hypothesis, our bioassay experiments revealed few significant differences in herbicide tolerances between the paired rare and common species” (Egan et al. 2014a). If the ED50 values of common and rare species are compared across all test substances, the average endpoint of the common species is even distinctly lower than the one of the rare species, though differences are not significant due to the variability of data. Also when based on individual active substances only in case of dicamba the average ED50 of the rare species is lower than the one of common species (by a factor of <2), for the other two active substances it is the other way round). The authors conclude that “results do suggest a challenge to the viewpoint that herbicides are a primary driver of plant biodiversity decline” Egan et al. (2014a). ER25 and ER50 endpoints were reported and were incorporated in our database.

Fletcher et al. (1985) analysed an US data base [46] focussing on plant species recommended for testing either by the US-EPA or by OECD. They observed that “in plant science research […] “ [often no ER 25 or ER50 is recorded], “published studies dealing with the responses of vascular plants to applied chemicals lack any degree of uniformity in experimental design, data analysis and/or quantification of results. Hence considerable data [of the database] […] could not be used in the present study.” (Fletcher et al. 1985)”. The same is true for most Fletcher data to be used in this study. Data can however be used in other contexts, e.g. considering the author’s own comparative approaches, the estimation of the level of phytotoxicity anticipated for wild versus crop species. The data displayed in Fletcher’s own paper however, although aiming to standardize towards an endpoint comparable to an ER50, fail to explain how the two units given “µM or kg/ha” relate to each other, as the latter is a rate (amount per area), the former an amount without any specified denominator (plant, square metre, hectare, litre application fluid?). Ultimately only data expressed as rate were included in the database. It should be noted that Fletcher et al. (1985) only list crop species endpoints, and furthermore also there are more reliable available data available by now for the assessed active substances, hence this unresolved unit “µM” is unlikely to have affected the overall outcome of this review.

Fletcher et al. 1990, again evaluating the US-EPA’s data base, focus here on potential influence of either greenhouse or field conditions on the sensitivity of terrestrial plants, and how this influence of test condition relates to differences in sensitivity due to taxonomic differences. The authors discovered that compared to the large differences in sensitivity related to the taxonomic position the test conditions (greenhouse or field) played only a minor role. The authors list quite a number of ER50 endpoints for active substance-species combinations. These were included in the data base assessed in the present paper.

Geisthard (2012) investigated effects of herbicides on phytophagous insects, effects on the plants were only a side aspect of his work. The actual thesis was not available to me, but endpoints were listed in Schmitz et al. 2015. As Schmitz’ endpoints were included in our

[46] Database termed PHYTOTOX, originally developed at the University of Oklahoma

database (unless already extracted directly from the papers), we included also Geisthard's endpoints.

Gilreath I Hortscience (2001) sprayed four different doses of glyphosate in three independent tomato plots at three reproductive growth stages. Plants were rated for foliar injury, and the number of open flowers and fruits per plant were counted. Since no ER25 or ER50 are reported, the data presented in Gilreath I Hortscience (2001) cannot be included in the data base for the overall assessment, but may be discussed on their own.

Gilreath II Hortscience (2001) evaluated the growth and yield development of cucumber in different stages after application in field of sublethal rates of 2,4-D. 2 experiments were carried out; the first assessed pre-bloom applications of 2,4-D and observed that these resulted in an increase in the foliar epinasty and a reduction on plant vigour, early yield and early fruit enlargement as rates of 2,4-D increased from 0 to 112 g a.s./ha. The application rates were applied in a logarithmic distribution: 0, 0.11, 1.12, 11.2 and 112 g a.s./ha. Main effects of 2,4-D in length and fresh weight were presented at the same table together with the yield (early and total). An Excel tool was used in the logarithmic calculation, providing an approximate ER10 value in g a.s./ha.

Gilreath III Hortscience (2001) conducted three field experiments to determine the effects of sublethal rates of dicamba and 2,4-D on pepper growth and yield. Dicamba and 2,4-D were applied as single and multiple application at various concentrations and at different stages of development. Dicamba was found to induce more foliar injury than 2,4-D. Plant vigour was rated applying a scale from 0 to 10, where 0 represented dead plants and 10 no effect compared to controls. Since no further information about scale composition was included in the report, the calculation of endpoints such as ER25 or ER50 was not possible and the findings data could not be included in the data base.

Hahn et al. (2014) characterized the size of field margins at two locations and investigated the prevailing plant community. As this investigation did not quantify any effects of defined pesticide exposure (no effect study) no endpoints were generated that could have been assessed in our review. The authors recorded the width of field margins using digital orthophotos and geographical information systems and reported the proportion of field margins narrower than three metres. They found pronounced differences between the two sites, but concluded that narrow grassy field margins can represent a large part of the available seminatural habitats adjoining agricultural sites. Also they consider many narrow margins not to be relevant for risk management (due to lack of unaffected areas), they question the current practice of pesticide risk assessment and management on a larger scale, and they propose better to protect field margins in Germany and other European countries.

Hemphill & Montomery 1981 assessed effects of 2,4-D on crops in field tests, differentiationg between vegetative end reproductive endpoints, which were included in the database. Ultimately the endpoints were not used in this project, where we focussed on vegetative endpoints from juvenile plants, as the authors seem to have tested vegetative effects on older plants, and reproductive endpoints.

Johnson 2015 cited endpoints of *Euphorbia esula* in his recent paper, investigating plants 14 days and 42 days after application. While many details were not reported, the (few) endpoints were tentatively included in the database, but only those measured after 14 days were assessed in this project.

Kjaer et al. (2006) carried out a study to investigate whether spray drift of metsulfuron has a potential to negatively affect hawthorn hedgerows near agricultural fields. For this purpose four doses of metsulfuron ranging from 5% to 40% of the field dose were sprayed on trees in seven different hawthorn hedgerows.
The results were presented by a linear scale relating the biomass (g) to a metsulfuron deposition (µg/cm^2). This approach unfortunately did not allow extracting relevant endpoints such as ER50 or ER25 for this data base.

Kleijin et al. (2004).The authors tested the hypothesis that standard fertilizer application could provoke an increase in total biomass production in combination with a decrease in species numbers. Their results supported the well-documented theory that an increase in nutrient resources in vegetation will lead to an increase in competition for light (Bobbink 1991; Tilman 1993). In both experiments such correlations were found.
In contrast, the herbicide applications did not provoke similarly uniform effects or effects of a similar magnitude. In the first experiment no herbicide effects were observed, whereas in the second one a reduction in the abundance was perceived.
The author presents the results comparing the presence of a determined species (% plots) over time (years), and observed a pronounced decline. Based on the results presented in the report it is however not possible to estimate any endpoints relevant for our assessment, such as ER25 or ER50.

In **Koch, Weißer, Strub (2004)** drift deposits on wheat plants were quantified and corresponding effects of paraquat assessed in order to describe drift dose response relations. Field experiment were performed assessing paraquat drift doses and the response of field-grown wheat, and laboratory experiments testing spray dose response on glass house grown wheat in the laboratory were carried out in parallel.
Results were based on foliar chlorosis (termed efficiency) related to the drift/spray deposition (ng/cm^2), but it was not possible to extract relevant endpoints to be included in the data base.

Marrs et al. (1989) selected 5 herbicides on the basis of the four risk/usage categories devised by Williams *et al*.(1987) viz. (1) high use/high risk = MCPA and mecoprop; (2) high use/moderate risk = asulam and glyphosate; and (3) low use/high risk = chlorsulfuron. Three series of experimental sprayings were done between 1987 and 1988. The effects of spray drift were tested at several distances downwind and the results were expressed as “safety distances” where no lethal effects, no damaging effects and no suppression of flowering occurred. No numerical data in terms of toxicity were given, hence no data could be included in the assessment.

Marrs et al. (1991a) tested effects of spray drift of three herbicides (glyphosate, MCPA mecoprop) on five plant species, varying distance to sprayer, height of surrounding vegetation, age of exposed plant but not implementing any dose-response design, and consequently not calculating ERx values that could have been incorporated in the database.

Bothe adverse effects and stimulating effects were observed, varying between plant and exposure level. Younger plants tended to be more affected than older plants, which was interpreted in terms to potential effects on population dynamics and depauperating communities at field margins.

Marrs et al. (1991b) demonstrated that damage in the most exposed area (i.e. 0 - 4 m zone downwind of the sprayer) varied greatly between herbicide doses, species, plant age and the structure of the surrounding vegetation. The experiment evaluated the effects of mecoprop in 60 artificially created microcosms, basically containers with 8 different species (*Digitalis purpurea*, Filipendula ulmaria, *Galium mollugo*, *Hypericum hirsutum*, *Lychnis floscuculi*, *Primula veris*, *Ranunculus acris* and *Stachys sylvatica*.
The species showed a reduced performance after mecoprop application, but only NOEC and LOEC were reported, or safety distances which cannot be translated into ERx endpoints either. Thus none of these endpoints could be included in the database.

McKelvey et al. (2000) compared directly effects of 11 herbicides on crop- and wild plant species, comparing the ER25 endpoint; differentiating between pre-emergence exposed seedling-emergence endpoints and foliar applied vegetative vigour endpoints. In this assessment wild species with mean ER25 values within 1 SD of the most sensitive crop ER25 minus 1 SD were classified as being of equivalent sensitivity. Hence only wild species ER25 endpoints falling below the most sensitive crop ER25 minus 1 SD were classified as more sensitive than the most sensitive crop species (McKelvey et al. 2000) and vice versa. Based on eleven substances evaluated they concluded that overall crop species' sensitivity was likely to be representative of non-crop herbaceous species sensitivity. The actual endpoints were not listed as numeric values but could be estimated from figures; these estimates were included in the database.

Newman et al. (2000) performed meta-analysis of NOEC, EC50 and LC50 data by analysing them with species-sensitivity distribution methods. Examples how to determine hazardous concentrations to 5% (HC5), to 10% (HC10), and to 20% (HC20) were presented as linear graphics under results. However as these distributions only displayed meta data but not the actual endpoints for individual species, the results could not be included in the data base but must be assessed on their own.

Obrigawitch et al. (1998) examined the relationship of short-term plant response measurements to plant productivity measurements such as yield or quality. They discuss if short-term plant response measurements have a practical degree of accuracy and precision that is appropriate for hazard assessment of sulfonylureas on non-target plants. This review provides an overview of research quantifying plant-growth effects resulting in crop or plant productivity losses as a result of exposure to sulfonylurea herbicides. The lowest test dose causing significant effect (LOEC) was calculated and presented in this review. While the data on their own were considered to be relevant, again no transformation into ERx values was possible.

Olszyk et al (2008) evaluated a methodology to determine risks to terrestrial native plant species from potential herbicide drift, focusing on 1) selection of native species for testing, 2) growth of these species, and 3) variability in herbicide response among native species and compared with crop plants.

5 native species and 5 crops were treated with sulfometuron methyl, resulting in distinct reductions e.g. in shoot dry weight. EC25 values were given for shoot dry weight, plant height and phytotoxicity (injury); these endpoints have been included in the data base.

In this study, **Olszyk et al. 2009** determined whether a short–growing season plant can indicate potential effects of herbicides on seed production but also measuring plant height or shoot biomass and visible injury. these vegetative endpoints were included in the database. While there were surprising inconsistencies between the ER25 endpoints reported by the authors themselves, and the new ER10 and ER50 estimates (based on raw data) calculated by EFSA, we applied a compromise by including both EFSA's endpoints and those originally reported by the authors, and assigned the same experiment number to it, which resulted in either the geometric mean to be used, or the lowermost endpoints, depending on the approach.

Olszyk 2010 reports results from tests conducted to determine whether a plant species with a short life cycle, such as *Brassica rapa L*., can be used to indicate potential effects on seed production of herbicides applied at relatively low levels (i.e. low fractions of the field application rates [FAR]). The effects of_0.1_FAR of aminopyralid, cloransulam, glyphosate, primisulfuron, or sulfometuron applied 14 d after emergence (DAE), were evaluated for *B. rapa* grown in mineral soil in pots under greenhouse conditions. Reproductive endpoints were also reported, in this assessment however only vegetative endpoints such as shoot dry weight were assessed.

Olszyk et al. 2010 determined whether young potato plants can be used as an assay to indicate potential effects of pesticides on asexual reproduction. His endpoints were relevant for the assessment of reproductive endpoints, and vegetative endpoints of mature plants, but no endpoints from juvenile plants were reported in this study, hence none contributed to this assessment.

Olszyk et al. 2013 tested the effect of glyphosate, tribenuron and fluazifop to 17 non-crop plant species from Oregon`s Willamette Valley. A dose-response test under greenhouse conditions was performed for each pair of test organism and herbicide. Growth rate and shoot dry weight was determined after 14 days of exposure. Results were presented as IC25 values which were transformed into field rates and included to the database.

In **Pfleeger et al. 2008** field trials were conducted to determine if potato (*Solanum tubersum* L.) vegetative growth and tuber yield and quality were affected by herbicides at below recommended field rates. Again these endpoints were relevant for the assessment of reproductive endpoints, and vegetative endpoints of mature plants, but no endpoints from juvenile plants were reported in this study, hence none contributed to this assessment.

Pfleeger et al. 2011 investigated effects of two herbicides on green-house and field-grown potatoes, soybean and peas, including reproductive endpoints. In 2002, plants from all three species were exposed to sulfometuron-methyl at concentrations of 0 (carrier control), 0.00056, 0.0032, 0.018, 0.1 and 1.0 FARs (max. recommended **f**ield **a**pplication **r**ate)s, in 2003, potato plants were exposed to glyphosate, bromoxynil, MCPA ([4-chloro-2-methylphenoxy] acetic acid), and sulfometuron-methyl at concentrations of 0, 0.00056, 0.0032, 0.018 and 0.1 FARs, and effects on various vegetative and reproductive endpoint recorded and ER25 calculated. MCPA and bromoxynil are stated not to have had significant effects on the plants species at the tested rates (data were not shown). Differences between locations were pronounced, and also there were pronounced differences between years (sulfometuron-methyl data tested in two succeeding years are reported). Differences in sensitivity between green-house and field-grown plants were overall found to been only minor, whereas reproductive endpoints were in many cases found to be lower than vegetative endpoints. Several vegetative and reproductive endpoints (EC_{25}) were extracted from the original report and incorporated in the data base. However endpoints must be treated with care, as some of them proved to be grossly extrapolated (up to a factor of 10 in case of sulfometuron-methyl and up to a factor of 40 in case of glyphosate); a three-parameter-Weibull-algorithm and the PROC NLIN procedure in SAS had been applied indiscriminately to all data, no matter whether the observed effect levels covered the calculated ER25 effect (i.e. whether the ER25 endpoint estimates were inside or outside of the tested range of treatment levels). These numeric endpoints thus had to be reassessed as they may be misleading, suggesting a certainty that was definitely not given. The authors could not provide the source data, but stated that this Weibull-function had been applied also to other papers (Olszyk, pers. comm.). As a pragmatic way out, extrapolated values were included as reported (as numeric value) but only when they did not exceed the highest rate tested by more than a factor of two. Larger extrapolations were included as censored values (less-than or greater-than values).

Pfleeger et al. (2012) evaluated effects of glyphosate and aminopyralid by means of a multi-species plant field trial. Three native Oregon plant species were grown together with an introduced species. The experiment was replicated at two locations with glyphosate applied at 0, 0.01 (8.3 g/ha), 0.1 (83.2 g/ha), and 0.2 (166.4 g/ha) and at the FAR (Field Application Rate) of 832 g/ha acid equivalent), and repeated for 3 years Tests with aminopyralid applied at 0, 0.037 (4.6 g/ha), 0.136 (16.7 g/ha), and 0.5 (61.5 g/ha were performed in two consecutive years.
Results were presented graphically. Variation of height and volume of plants are shown but it was not possible to extract numerical information (such as ERx values) from them.

In **Pfleeger et al. 2014**, toxicology tests were conducted on potatoes, peas, and soybeans grown in a native soil in pots in the greenhouse and were compared to plants grown outside under natural environmental conditions to determine toxicological differences between environments. The herbicides bromoxynil, glyphosate, MCPA ([4-chloro-2-methylphenoxy] acetic acid), and sulfometuron-methyl were applied at below field application rates to potato plants at two developmental stages. Peas and soybeans were exposed to sulfometuron-methyl at similar rates at three developmental stages. Again these endpoints were relevant for the assessment of vegetative endpoints of mature plants, but no endpoints from juvenile plants were reported in this study, hence none contributed to this assessment.

Reuter & Siemoneit (1987) performed a series of tests on wild plant species, either in single species tests, or arranged in “artificial communities”, here termed “terrestrial microcosms”. Evaluation of fresh weight was based on OECD 227 (vegetative vigour) effects of a broadband herbicide and a selective herbicide were tested both in potted single-species test according to OECD 227 and in the terrestrial microcosms, each comprising of 6 species. These had been selected to ensure presence of different families, different growth types

(habitus) but also based on practicability (growth period sufficiently long and not too different growth patterns, quantitative harvesting feasible, seeds commercially available etc. (several further species had been tested for suitability in non-dosed test systems)). The dosed tests were ultimately performed with *Trifolium pratense*, *Bromus erectus*, *Cynosurus cristatus*, *Galium mollugo*, *Leontodon hispidus*, and *Silene nutans*. The observation period was extended in order to assess any recovery potential. Hence three sets of endpoints were obtained per test, after 14, 28 and 42 days of observation. ER25, ER50 and ER75 endpoints were reported for the individual species in the single species tests, for individual species in the microcosms' plant communities, and the biomass of the total plant community was also assessed quantitatively.
The endpoints expressed as mL PPP/ha were transformed into g a.s./ha, considering the following details: Roundup® Ultra (Monsanto) containing glyphosate at 360 g pure a.s. per L formulation (486 g of the isopropylamin-salt L-1), and in case of sulfosulfuron the formulation content of 800 g/L. While the authors consider that their artificial community approach is a valuable alternative option, they focussed on the recovery potential of affected plants while competing each other. In our study only to compare potential differences in sensitivity between crop species and wild plants species, only effects after 14 and 28 days were included, as any growth patterns related to recovery would have distorted the outcome.

Riemens et al. (2008) compared different endpoints, i.e. aboveground biomass, seed production, seed germination and recovery of different species grown in the greenhouse and in the field, seeds of *Chenopodium album*, *Stellaria media*, *Poa annua*, and *Echinochloa crus-galli* after application of glufosinate ammonium. Dosages were 0, 0.04, 0.2, 0.4, 2, and 4 L Finale per ha (which corresponded with doses of 0, 6, 30, 60, 300 and 600 g glufosinate ammonium per ha). ED50 values on fresh weight are presented in the report and therefore included in the assessment.

Riemens et al. (2009) compared greenhouse test endpoints with field test endpoints, and also single species tests with a "Mesocosm" setup in which several test species were combined in one test system. Based on just one herbicide tested they concluded that they found some agreement between greenhouse- and field test data, however suggested that a translation of results would be appropriate; and they prefer multi-species tests over single-species tests, considering that the most important parameters were composition of the species in the test system, mode of action, development stages of the species (in the test and on the field margin to be protected) and the chosen endpoint (biomass or reproductive endpoint) ER50 values based on fresh weight were presented in the report and included in this analysis.

Schmitz et al. (2013) To assess the effects of the agrochemical applications on *Ranunculus acris*, plant community assessments were carried out and a photo-documentation of the flowering intensity was performed over two years. In addition, the authors conducted a monitoring survey of *R. acris* in field margins where herbicide were expected to cause a sublethal effect i.e. flower intensity was reduced by 85%. Results were based on measurements of plant density, hence again no numeric endpoints relevant for our assessment were available.

Schmitz et al. (2014a) investigated effects of combined fertilizer, herbicide, and insecticide exposure on the reproduction of four species inputs on four wild plant species (*Ranunculus acris, Lathyrus pratensis, Vicia sepium, Rumex acetosa*) over a period of three years. Differences between controls and treated plots were observed, correlated both to fertilizer application (25% of the field rate) and to the pesticide applications (at 30% of the field rate). Both herbicide and fertilizer exposure had significant effects on the reproductive performance of three of the four species, only *R. acetosa* was not affected. Effects on frequency and biomass were recorded; however as the study did not follow a dose-response design, no ERx-endpoints were determined that could have been included in this data base, but the results are considered in the discussion.

Schmitz et al. (2014b) performed a 3-year field study in a randomized block design investigating individual and combined effects of fertilizer, herbicide, and insecticide inputs on plant communities of field margins. They observed differences between controls and treated plots both correlated to fertilizer application (25% of the field rate) which appeared to promote plants with a high nutrient uptake, and to the pesticide applications (at 30% of the field rate). Frequencies of small and subordinate species were found to decrease and number of species and diversity decreasing compared to the control plots. The authors conclude that repeated annual application of agrochemicals and fertilizers at these rates cause long-term shifts in the plant community. Effects on frequency and biomass were recorded; however as the study did not follow a dose-response design, no ERx-endpoints were determined, so data could not be assessed in this data base, but are considered in the discussion.

Siemoneit-Gast et al. (2007) developed an extended method for assessing the risk to terrestrial non-target plants from pesticides by combining several species into one test system, aiming to study also effects of competition on their resilience to exposure to pesticides. They advertise their test systems as a useful intermediate step between testing in the lab and in the field, but also discuss restrictions, feasibility problems and practical problems. No numeric endpoints are reported in this paper. However, the report forming the basis of their publication (Reuter & Siemoneit-Gast 2007) listed numeric endpoints, which were incorporated in the database (see further up)

Snoo et al (2005) report results from a large scale field experiment over two years, in which glufosinate-ammonium drift rates (expressed as % of target field rate) were applied to edge-of-field communities (pioneer species, i.e. no crop species tested). The target field rate was 800 g a.s./ha, and tested drift rates were thus 16, 32, 128, 256 and 512 g a.s./ha. Effect levels were assessed in terms of overall phytotoxic effect and % coverage. However only the overall analysis is presented, not effect levels of individual plant species, except for Figures 4.3 and 4.4. where example effect plots are given for a few species. While these indicate that *Rumex acetosa*, *Trifolium pratensis* and *Cerastium fontanum* were among the most sensitive species, no individual ERx were available to be incorporated in the database. The overall finding of de Snoo et al, 2005 was that significant short-term effects of glufosinate-ammonium on edge-of field plants could be demonstrated at levels as low as 16 to 32 g a.s./ha (2 and 4% of the target field rate) 10 days after application. At later assessment dates effects on the community were only visible at rates as high as 240 g a.s./ha (30%) of the field rate or higher. One year after spraying no effects could be detected even at the highest rate.

Spaunhorst & Bradley 2013 assessed effects of single herbicides and combinations on glyphosate-resistant waterhemp (*Amaranthus*). The authors followed a rate-response design and reported % control. Also they tested how effects varied depending on the size of the plants at application. Inclusion of these endpoints was borderline, as resistance in weeds could affect the outcome of this assessment (being not worst-case). However, endpoints based on rates of the substance used to break resistance was considered to be relevant, and was thus included in the database.

Strandberg et al. (2012). The project was carried out in order to investigate the effects of herbicides on plants found in natural and semi-natural habitats within the agricultural land such as hedgerows or field borders. 3 herbicides (metsulfuron-methyl, mecoprop-p and glyphosate) were assessed in detail, and results, were presented as ED50, which were included in the data base. The authors conclude that the crop species tested on glyphosate, metsulfuron-methyl and mecoprop-P in general were not less sensitive to herbicides than wild non-target species when dose-response experiments were run under the same conditions. They stress influence of test conditions on outcome may have been underestimated in the past, and they propose to use any data base only when test conditions were also recorded, as in absence of such information wrong or misleading conclusions on species sensitivity may be drawn.

Strandberg et al. 2013 listed a few vegetative endpoints and gave effect curves for reproductive endpoints as well (Fig. 3.1.) The latter were calculated approximately via interpolation, the former could be included directly in the database and were assessed in this project.

Vielhauer (2010) compared closely related crop ans wild plants in this diploma thesis Unfortunately the thesis as such was not made available to us, but some endpoints were presented in a poster (entry below) and other endpoints cited by Schmitz et al. 2015. As Schmitz' endpoints were included in our database (unless already extracted directly from the source papers), and we included also Vielhauer's endpoints as far as available.

Vielhauer & Bruehl (2009) was a SETAC presentation (poster) where the authors assessed the differences in crops sensitivity against broad-spectrum herbicide between closely related crop species and wild species. The author based their conclusion on a very small dataset, which however has been included in our data base.

White & Boutin (2007). Several crops and wild plant species were grown under greenhouse conditions following standard protocol for phytotoxicity testing. Plants were sprayed with five different herbicides (Round-Up original, Glyphosate, Aatrex liquid 480, Pursuit and MCPA Amine 500) at the four- to six-leaf stage, and biomass was recorded at 28 d after spray. The authors concluded that current regulatory protocols were likely to underestimate herbicide phytotoxicity if testing does not include data for the complete tank-mix formulation. The IC25 values were included in the database.

Wingender & Weddeling (2010) provide another recent literature research (on behalf of the German Federal Ministry of Food and Agriculture) investigating effects of agricultural uses on ecological diversity of wild plants and animals. Exposure to pesticides was ***not*** found to be the fundamental cause determining plant diversity in and around agricultural areas (Wingender & Weddeling 2010).

9 Appendix 2 – List of active substances

Below is the list of all active substances for which published plant endpoints were available. For some of them only crop endpoints, for others only wild species endpoints were available; also this list includes cases where most endpoints were censored. This is the reason for the list of substances that allowed numeric assessment was shorter; see Tables 3 and 10.3 and 10.

Table 8: List of all active substances for which published plant endpoints were available. For some of them only crop endpoints, for others only wild species endpoints were available; also these include cases with only censored endpoints.

Active substance	
2,4-D	Glyphosate
2,4-D amine	Imazapyr
2,4-DB	Imazaquin
Acifluorfen	Imazethapyr
Alachlor	Iodosulfuron-methyl-natrium
Aminocyclopyrachlor	Isoproturon
Aminopyralid	K-815910
Amitrole	Lactofen
Atrazine	Linuron
Barban	MCPA
Bentazon	MCPA Amine
Bromoxynil	Mecoprop
Chloramben	Mecoprop-P
Chlorimuron	Mesosulfuron-methyl
Chloroxuron	metazachlor
chlorpropham	Metolachlor
Chlorsulfuron	Metribuzin
Clodinafop-propargyl	Metsulfuron-methyl
Clofop-methyl	MSMA
Clomazone	Nitrofen
Clopyralid	Oxyfluorfen
Cloransulam	paraquat
Cloransulam-methyl	Pendimethalin
Dalapon	Picloram
DCPA	primisulfuron
Dicamba	Prometryne
Diclofop-methyl	Pyridate
Dinoseb	Pyridyloxy A
Diphenamid	Pyridyloxy B
Diquat	Simazine
EPTC	sulfometuron
Ethofumesate	Sulfosulfuron

Active substance	
fenoxaprop-P-ethyl	Tepraloxydim
Flamprop-M-isopropyl	terbacil
Flazasulfuron	thifensulfuron
Fluazifop	Triallate
Fluroxypyr	Triasulfuron
Glufosinate	Tribenuron-methyl
Glufosinate-ammonium	Trifluralin

Mixtures: (included as “a.s.- category” if matching crop and wild endpoints available)
2,4-D + mecoprop + dicamba
2,4-D+glyphosate
Amitrole + ammonium thiocyanate
Carfentrazone-ethyl + Isoprotutron
Chlorsulfuron + Metsulfuron
Chlortoluron + MCPA
Ioxynil + bromoxynil
Paraquat + Diquat
Paraquat + tridiphane
mesosulfuron-methyl + iodosulfuron-methylnatrium + mefenpyr-diethyl)
Thifensulfuron + Metsulfuron 680 + 68
Thifensulfuron + Tribenuron
Thifensulfuron + Tribenuron + 2.4-D

Non-herbicides with plant endpoints (not included, no matching crop and wild endpoints available)
wetting agents only (blank formulation)
adjuvant only (MCDS)
(Lambda-cyhalothrin)
(Dichlobenil)

10 Appendix 3 – List of species for which numeric endpoints were found in the literature or in GLP data submitted for registration

Below is the list of all species for which published plant endpoints were included in the data base. For some of them only published endpoints, for others only data from regulatory studies were available. Again this list includes cases where most endpoints were censored.

Table 9: List of all plant species (wild and crop) with vegetative endpoints from experiments on young plants (NOER, NOEC or ERx ICx or ECx).

Species	BAYER CODE	Family	W/C(wild/crop/ introduced)	Cases incl. cens.
Abutilon theophrasti	ABUTH	Malvaceae	W/C	4
Acer negundo	ACRNE	Aceraceae	W	1
Acer rubrum	ACRRB	Aceraceae	W	1
Achillea millefolium	ACHMI	Asteraceae	W	12
Agrostis stolonifera	AGSST	Poaceae	W	13
Alliaria petiolata	ALAPE	Brassicaceae	C	22
Allium cepa	ALLCE	Amaryllidaceae	C	136
Alopecurus myosuroides	ALOMY	Poaceae	W	11
Amaranthus palmeri	AMAPA	Amaranthaceae	I	1
Amaranthus retroflexus	AMARE	Amaranthaceae	W	11
Amaranthus sp.	AMASS	Amaranthaceae	W	8
Ambrosia artemisiifolia	AMBEL	Asteraceae	W	1
Anagallis arvensis	ANGAR	Primulaceae	W	9
Andropogon gerardii	ANOGE	Poaceae	W	8
Anthriscus sylvestris	ANRSY	Apiaceae	W	6
Apocynum cannabinum	APCCA	Apocynaceae	W	2
Arctotheca calendula	AROCA	Asteraceae	W	1
Arrhenatherum elatius	ARREL	Poaceae	W	14
Asclepias incarnata	ASCIN	Asclepiadaceae	W	2
Asclepias syriaca	ASCSY	Asclepiadaceae	W	12
Asclepias tuberosa	ASCTU	Asclepiadaceae	W	6
Avena fatua	AVEFA	Poaceae	W	15
Avena sativa	AVESA	Poaceae	C	114
Bellis perennis	BELPE	Asteraceae	W	13
Beta vulgaris	BEAVX	Chenopodiaceae	C	74
Bidens cernua	BIDCE	Asteraceae	W	12
Bidens frondosa	BIDFR	Asteraceae	W	11
Brachiaria fasciculata	PANFA	Poaceae	W	1
Brachydopodium sylvaticum	BRCSI	Poaceae	W	14
Brassica juncea	BRSJU	Brassicaceae	C	1

Species	BAYER CODE	Family	W/C(wild/crop/ introduced)	Cases incl. cens.
Brassica napus	BRSNN	Brassicaceae	C	62
Brassica oleracea	BRSOX	Brassicaceae	C	61
Brassica oleracea var. capitata	BRSOL	Brassicaceae	C	2
Brassica rapa	BRSRR	Brassicaceae	C	11
Bromus carinatus	BROCN	Poaceae	W	5
Bromus erectus	BROER	Poaceae	W	17
Bromus sterilis	BROST	Poaceae	W	16
Bupleurum rotundifolium	BUPRO	Apiaceae	W	2
Campanula americana	CMPAM	Campanulaceae	W	1
Campanula rotundifolia	CMPRO	Campanulaceae	W	3
Capsella bursa pastoris	CAPBP	Brassicaceae	W	14
Carthamus tinctorius L.	CAUTI	Asteraceae	C	6
Centaurea cyanus	CENCY	Asteraceae	W	18
Centaurea jacea	CENJA	Asteraceae	W	3
Cerastium fontanum	CERVU	Caryophyllaceae	W	7
Chenopodium album	CHEAL	Chenopodiaceae	W	21
Chondrosum gracile	BOBGR	Poaceae	W	5
Cichorium intybus	CICIN	Asteraceae	W	3
Clarkia amoena	CKAAM	Onagraceae	W	5
Cleome serrulata	CLESE	Capparaceae	W	2
Clinopodium vulgare	STIVU	Lamiaceae	W	6
Collinsia grandiflora	CLCGF	Scrophulariaceae	W	4
Crepis biennis	CVPBI	Asteraceae	W	3
Cucumis sativus	CUMSA	Cucurbitaceae	C	99
Cynodon dactylon	CYNDA	Poaceae	W	12
Cynosurus cristatus	CYNCR	Poaceae	W	9
Cynosurus echinatus	CYXEC	Poaceae	W	5
Cyperus esculentus	CYPES	Cyperaceae	C	2
Cyperus rotundus	CYPBI	Cyperaceae	W	12
Dactylis glomerata	DACGL	Poaceae	W	14
Danthonia californica	DANCF	Poaceae	W	6
Daucus carota ssp. carota	DAUCA	Apiaceae	W	3
Daucus carota ssp. sativus	DAUCS	Apiaceae	C	55
Digitalis purpurea	DIKPU	Plantaginaceae	W	13
Digitaria sanguinalis	DIGSA	Poaceae	W	11
Echinochloa colona	ECHCO	Poaceae	W	1
Echinochloa crus-galli	ECHCX	Poaceae	W	40
Eleusine indica	ELEIN	Poaceae	W	2
elymus canadensis	ELYCA	Poaceae	W	8
Elymus hystrix	ELYHX	Poaceae	W	12
Elymus lanceolatus	AGRDA	Poaceae	W	5

Species	BAYER CODE	Family	W/C(wild/crop/ introduced)	Cases incl. cens.
Elymus repens	AGRRE	Poaceae	W	24
Elymus riparius	ELYRX	Poaceae	W	12
Elymus trachycaulus	AGRTR	Poaceae	W	5
Elymus virginicus	ELYVI	Poaceae	W	4
Emex australis	EMESP	Polygonaceae	W	1
Eriophyllum lanatum	ERHLA	Asteraceae	W	5
Eupatorium maculatum	EUPML	Asteraceae	W	9
Euphorbia heterophylla	EPHHL	Euphorbiaceae	W	6
fagopyrum esculentum	FAGES	Polygonaceae	C	7
Fallopia convolvulus	POLCO	Polygonaceae	W	28
Festuca arundinacea	FESAR	Poaceae	W	7
Festuca roemeri	FESRO	Poaceae	W	2
Festuca rubra	FESRU	Poaceae	W	14
Fragaria ananassa	FRAAN	Rosaceae	C	5
Fragaria virginiana	FRAVI	Rosaceae	W	5
Frangula alnus	RHAFR	Rhamnaceae	W	1
Galium aparine	GALAP	Rubiaceae	W	18
Galium mollugo	GALMO	Rubiaceae	W	14
Galium verum	GALVE	Rubiaceae	W	3
Geranium dissectum	GERDI	Geraniaceae	W	3
Geranium molle	GERMO	Geraniaceae	W	18
Geranium robertianum	GERRO	Geraniaceae	W	18
Geum canadense	GEUCD	Rosaceae	W	12
Gilia capitata	GILCA	Polemoniaceae	W	5
Glebionis segetum	CHYSE	Asteraceae	W	2
Glyceria striata	GLYST	Poaceae	W	5
Glycine max	GLXMA	Fabaceae	C	150
Gossypium hirsutum	GOSHI	Malvaceae	C	33
Gossypium sp.	GOSSS	Malvaceae	W	72
helianthus annuus	HELAN	Asteraceae	C	39
Helianthus tuberosus	HELST	Asteraceae	W	6
Heracleum sphondylium	HERSP	Apiaceae	W	8
Hieracium pilosella	HIEPI	Asteraceae	W	3
Hordeum vulgare	HORVX	Poaceae	C	8
Hypericum perforatum	HYPPE	Hypericaceae	W	7
Hypochaeris radicata	HRYRA	Asteraceae	W	6
Ilex glabra	ILEGL	Aquifoliaceae	W	1
Inula helenium	INUHE	Asteraceae	W	11
Ipomoea hederacea	IPOHE	Convolvulaceae	W	6
Ipomoea lacunosa	IPOLA	Convolvulaceae	W	1
Ipomoea purpurea	IPOHT	Convolvulaceae	W	2

Species	BAYER CODE	Family	W/C(wild/crop/ introduced)	Cases incl. cens.
Ipomoea quamoclit	IPOQU	Convolvulaceae	W	1
Ipomoea wrightii	IPOWR	Convolvulaceae	W	1
Juncus tenuis	IUNTE	Juncaceae	W	4
Koeleria macrantha	KOLMA	Poaceae	W	6
Lactuca canadensis	LACCA	Asteraceae	W	1
Lactuca sativa	LACSA	Asteraceae	C	98
Lactuca serriola	LACSE	Asteraceae	W	6
Lamium purpureum	LAMPU	Lamiaceae	W	2
Lens culinaris	LENCU	Fabaceae	C	30
Leontodon hispidus	LEBHI	Asteraceae	W	11
Leonurus cardiaca	LECCA	Lamiaceae	W	9
Lepidium sativum	LEPSA	Brassicaceae	C	0
Leptochloa mucronata	LEFFI	Poaceae	W	1
Leucanthemum vulgare	CHYLE	Asteraceae	W	33
Linaria canadensis	LINCA	Scrophulariaceae	W	5
Linum usitatissimum	LIUUT	Linaceae	C	1
Lobelia inflata	LOBIN	Campanulaceae	W	5
Lolium multiflorum	LOLMG	Poaceae	W	1
Lolium perenne	LOLPE	Poaceae	W	39
Lolium sp.	LOLSS	Poaceae	W	55
Lotus corniculatus	LOTCO	Fabaceae	W	3
Lupinus albicaulis	LUPAC	Fabaceae	W	4
Lycopersicon esculentum	LYPES	Solanaceae	C	135
Lycopus americanus	LYCAM	Lamiaceae	W	17
Madia elegans	MADEL	Asteraceae	W	4
Malva moschata	MALMO	Malvaceae	W	3
Medicago lupulina	MEDLU	Fabaceae	W	3
Medicago sativa	MEDSA	Fabaceae	W/C	3
Melilotus officinalis	MEUOF	Fabaceae	W	5
Mentha spicata	MENSP	Lamiaceae	W/C	6
Mollugo verticillata	MOLVE	Caryophyllaceae	W	1
Myosotis arvensis	MYOAR	Boraginaceae	W	14
Nepeta cataria	NEPCA	Lamiaceae	W	6
Neslia paniculata	NEAPA	Brassicaceae	W	2
Nicotiana rustica	NIORU	Solanaceae	W	1
Pachysandra terminalis	PCHTE	Buxaceae	W	1
Panicum clandestinum	PANCL	Poaceae	W	3
Panicum miliaceum	PANDI	Poaceae	W/C	17
Panicum virgatum	PANVI	Poaceae	W	8
Papaver argemone	PAPAR	Papaveraceae	W	4
Papaver rhoeas	PAPRH	Papaveraceae	W	12

Species	BAYER CODE	Family	W/C(wild/crop/ introduced)	Cases incl. cens.
Pastinaca sativa	PAVSA	Apiaceae	C	2
Pennisetum glaucum	PESGL	Poaceae	W	108
Persicaria hydropiper	POLHY	Polygonaceae	W	1
Persicaria lapathifolia	POLLA	Polygonaceae	W	11
Persicaria pensylvanica	POLPY	Polygonaceae	W	5
Phalaris aquatica	PHATU	Poaceae	W	1
Phalaris arundinacea	TYPAR	Poaceae	W	10
Phaseolus vulgaris	PHSVX	Fabaceae	C	5
Phaseolus vulgaris	PHSVX	Fabaceae	C	2
Phytolacca americana	PHTAM	Phytolaccaceae	W	5
Picris hieracioides	PICHI	Asteraceae	W	3
Pinus resinosa	PIURE	Pinaceae	W	5
Pisum sativum	PIBSX	Fabaceae	C	140
Plantago lanceolata	PLALA	Plantaginaceae	W	22
Poa annua	POAAN	Poaceae	W	21
Poa compressa	POACO	Poaceae	W	5
Poa palustris	POAPA	Poaceae	W	6
Poa pratensis	POAPR	Poaceae	W	14
Poa trivialis	POATR	Poaceae	W	16
Populus grandidentata	POPGR	Salicaceae	W	1
Portulaca oleracea	POROL	Portulacaceae	W	1
Potentilla gracilis	PTLGR	Rosaceae	W	5
Potentilla recta	PTLRC	Rosaceae	W	9
Prunella vulgaris	PRUVU	Lamiaceae	W	18
Prunus avium	PRNAJ	Rosaceae	C	48
Quercus palustris	QUEPA	Fagaceae	W	1
Quercus rubra	QUERU	Fagaceae	W	1
Ranunculus acris	RANAC	Ranunculaceae	W	8
Ranunculus occidentalis	RANOC	Ranunculaceae	W	4
Ranunculus repens	RANAC	Ranunculaceae	W	8
Raphanus raphanistrum	RAPRA	Brassicaceae	W	2
Raphanus sativus	RAPSR	Brassicaceae	C	78
Rapistrum rugosum	RASRL	Brassicaceae	W	4
Rhamnus cathartica	RHACT	Rhamnaceae	W	1
Ricinus communis	RIICO	Euphorbiaceae	W	1
Rosa wichuraiana	ROSWI	Rosaceae	W	1
Rudbeckia hirta	RUDHI	Asteraceae	W	19
Rumex acetosa	RUMAC	Polygonaceae	W	8
Rumex crispus	RUMCR	Polygonaceae	W	12
Sanguisorba occidentalis	SANOC	Rosaceae	W	5
Scandix pecten-veneris	SCABR	Apiaceae	W	4

Species	BAYER CODE	Family	W/C(wild/crop/ introduced)	Cases incl. cens.
Scorzoneroides autumnalis	LEBAU	Asteraceae	W	3
Senecio vulgaris	SENVU	Asteraceae	W	2
Senna obtusifolia	CASOB	Caesalpinioideae	W	1
Senna occidentalis	CASOC	Caesalpinioideae	W	1
Setaria faberi	SETFA	Poaceae	W	7
Setaria italica	SETIT	Poaceae	C	2
Setaria viridis	SETVI	Poaceae	W	48
Silene latifolia	MELAL	Caryophyllaceae	W	3
Silene noctiflora	MELNO	Caryophyllaceae	W	18
Silene nutans	SILNU	Caryophyllaceae	W	12
Silene vulgaris	SILVU	Caryophyllaceae	W	15
Sinapis arvensis	SINAR	Brassicaceae	W	10
Sisymbrium officinale	SSYOF	Brassicaceae	W	14
Solanum americanum	SOLAM	Solanaceae	W	194
Solanum dulcamara	SOLDU	Solanaceae	W	7
Solanum nigrum	SOLNI	Solanaceae	W	5
Solanum tuberosum	SOLAD	Solanaceae	C	24
Solidago canadensis	SOOCA	Asteraceae	W	18
Sorghum bicolor	SORVU	Poaceae	C	34
Sorghum halepense	SORHA	Poaceae	W	2
Spergula arvensis	SPRAR	Caryophyllaceae	W	1
Stellaria media	STEME	Caryophyllaceae	W	11
Symphyotrichum lateriflorum	ASTLF	Asteraceae	W	12
Syringa vulgaris	SYRVU	Oleaceae	W	1
Taraxacum officinale	TAROF	Asteraceae	W	1
Tridens flavus	TRSFL	Poaceae	W	2
Trifolium dubium	TRFDU	Fabaceae	W	3
Trifolium pratense	TRFPR	Fabaceae	W	30
Trifolium repens	TRFRE	Fabaceae	W	6
Trifolium subterraneum	TRFSU	Fabaceae	W	1
Tripleurospermum inodorum	MATIN	Asteraceae	W	9
Tripleurospermum perforatum	MATIN	Asteraceae	W	2
Triticum aestivum	TRZAX	Poaceae	C	68
Ulmus americana	ULMAM	Ulmaceae	W	1
Urtica dioica	URTDI	Urticaceae	W	6
Verbena hastata	VEBHA	Verbenaceae	W	13
Verbena urticifolia	VEBUR	Verbenaceae	W	12
Veronica americana	VERAM	Plantaginaceae	W	5
Veronica arvensis	VERAR	Plantaginaceae	W	2
Veronica persica	VERPE	Plantaginaceae	W	16
Vicia americana	VICAM	Fabaceae	W	2

Species	BAYER CODE	Family	W/C(wild/crop/ introduced)	Cases incl. cens.
Vicia faba	VICFX	Fabaceae	C	5
Vigna radiata	PHSAU	Fabaceae	C	2
Viola tricolor	VIOTR	Violaceae	W	3
Xanthium strumarium	XANST	Asteraceae	W	6
Zea mays	ZEAMX	Poaceae	C	150

11 Appendix 4 – List of all available data sets

The data listed in Table 3 on p. 29 is the “positive“ subset of a larger table that shows the available endpoints of all substances, not only those above the criterion “n > 3“. The table below (continued overleaf) lists the numbers of ER10, ER25 and ER50 endpoints (biomass, vegetative vigour, lab or field) of all substances, and also indicates where differentiating between monocots and dicots could be informative. To avoid confusion, other selections such as endpoints based on shoot height or survival, or such based on seedling emergence tests are not listed below. Numbers including censored endpoints are listed in separate rows.

Table 10: List of substances and numbers of numeric endpoints differentiating by monocots/dicots and wild and crop species. The column labelled “endpoints” indicates the endpoint’s effect level and -type (here only vegetative of juvenile plants, biomass, and the figures the resulting number of species*type (lab or field) with numeric endpoints or with any endpoint (second row each).

Subst. code	Endpoint	‘n’ based on	Dicots		Monocots		All		Total	Comments
		Spec*type L-F &	wild	crop	wild	crop	wild	crop	all	
AASI 1	ER25 (VV BM)	numerics only	5	5	3	3	8	8	17	Also ER05 and ER75 endpoints
		with cens., f=2	6	6	4	3	10	9	20	
	ER50 (VV BM)	numerics only	6	4	3	3	9	7	17	
		with cens., f=2	8	6	5	3	13	9	23	
AASI 2	ER10 (VV BM)	numerics only	0	0	1	0	1	0	1	Also SH endpoints
		with cens., f=2	0	0	1	0	1	0	1	
	ER25 (VV BM)	numerics only	1	1	4	0	5	1	6	
		with cens., f=2	6	1	10	0	16	1	17	
	ER50 (VV BM)	numerics only	3	6	1	4	4	10	14	
		with cens., f=2	6	7	11	5	17	12	29	
AASI 3	ER10 (VV BM)	numerics only	4	9	5	4	9	13	22	Also ER05, ER75, SH and SE endpoints
		with cens., f=2	4	9	5	4	9	13	22	
	ER25 (VV BM)	numerics only	62	12	23	4	85	16	102	
		with cens., f=2	62	12	24	4	86	16	103	
	ER50 (VV BM)	numerics only	49	10	13	4	62	14	77	
		with cens., f=2	49	10	13	4	62	14	77	
AASI 4	ER10 (VV BM)	numerics only	6	0	0	0	6	0	6	Also SH endpoints
		with cens., f=2	6	0	0	0	6	0	6	
	ER25 (VV BM)	numerics only	11	2	4	0	15	2	17	
		with cens., f=2	24	10	13	4	37	14	51	
	ER50 (VV BM)	numerics only	10	10	0	4	10	14	24	
		with cens., f=2	18	13	7	6	25	19	44	
AASI 5	ER10 (VV BM)	numerics only	7	0	2	0	9	0	9	
		with cens., f=2	7	0	2	0	9	0	9	
	ER25 (VV BM)	numerics only	7	8	2	4	9	12	21	
		with cens., f=2	7	8	3	4	10	12	22	
	ER50 (VV BM)	numerics only	9	0	2	0	11	0	11	
		with cens., f=2	11	0	5	0	16	0	16	
AASI 6	ER25 (VV BM)	numerics only	4	3	3	3	7	6	13	
		with cens., f=2	5	5	3	3	8	8	16	
AASI 7	ER25 (VV BM)	numerics only	2	6	6	3	8	9	17	
		with cens., f=2	2	6	6	3	8	9	17	

AASI14	ER10 (VV BM)	numerics only	0	1	1	1	1	2	3	
		with cens., f=2	1	2	1	2	2	4	6	
	ER25 (VV BM)	numerics only	0	1	1	2	1	3	4	
		with cens., f=2	1	2	1	2	2	4	6	
	ER50 (VV BM)	numerics only	1	2	1	2	2	4	6	
		with cens., f=2	1	2	1	2	2	4	6	
AASI15	ER25 (VV BM)	numerics only	2	1	1	1	3	2	5	
		with cens., f=2	2	3	2	2	4	5	9	
AASI16	ER50 (VV BM)	numerics only	0	1	0	0	0	1	1	
		with cens., f=2	0	1	0	0	0	1	1	
AASI17	ER10 (VV BM)	numerics only	3	6	1	4	4	10	14	
		with cens., f=2	3	6	1	4	4	10	14	
	ER25 (VV BM)	numerics only	0	6	1	4	1	10	11	
		with cens., f=2	0	6	1	4	1	10	11	
	ER50 (VV BM)	numerics only	24	6	2	5	26	11	37	
		with cens., f=2	26	6	2	5	28	11	39	
ACI 1	ER50 (VV BM)	numerics only	0	0	0	2	0	2	2	
		with cens., f=2	0	0	0	2	0	2	2	
ACI 2	ER50 (VV BM)	numerics only	0	0	9	1	9	1	10	
		with cens., f=2	0	0	9	1	9	1	10	
ACI 3	ER50 (VV BM)	numerics only	0	0	1	2	1	2	3	
		with cens., f=2	10	7	3	3	13	10	23	
CMD 1	ER10 (VV BM)	numerics only	0	0	1	0	1	0	1	(found further wild sp. data, overall no change)
		with cens., f=2	0	0	2	0	2	0	2	
	ER25 (VV BM)	numerics only	8	0	1	0	9	0	9	
		with cens., f=2	8	0	2	0	10	0	10	
	ER50 (VV BM)	numerics only	18	6	11	3	29	9	38	
		with cens., f=2	18	6	11	3	29	9	38	
CMD 2	ER50 (VV BM)	numerics only	0	0	0	2	0	2	2	
		with cens., f=2	0	0	0	2	0	2	2	
CMD 3	ER50 (VV BM)	numerics only	0	2	0	0	0	2	2	
		with cens., f=2	0	2	0	0	0	2	2	
CMD 4	ER25 (VV BM)	numerics only	0	0	0	1	0	1	1	
		with cens., f=2	0	0	0	1	0	1	1	
	ER50 (VV BM)	numerics only	0	2	0	2	0	4	4	
		with cens., f=2	0	2	0	3	0	5	5	
CMD 6	ER50 (VV BM)	numerics only	1	0	0	0	1	0	1	
		with cens., f=2	1	0	0	0	1	0	1	
CMD 7	ER50 (VV BM)	numerics only	1	0	0	0	1	0	1	
		with cens., f=2	1	0	0	0	1	0	1	
GW 01	ER10 (VV BM)	numerics only	1	5	1	3	2	8	10	Also ER05 endpoints, also SH endpoints; further endpoints expressed in µmol#
		with cens., f=2	1	7	1	3	2	10	12	
	ER25 (VV BM)	numerics only	13	8	4	2	17	10	27	
		with cens., f=2	13	9	6	3	19	12	31	
	ER50 (VV BM)	numerics only	24	9	3	3	27	12	39	
		with cens., f=2	24	10	4	4	28	14	42	
GW 02	ER25 (VV BM)	numerics only	0	0	0	0	0	0	0	
		with cens., f=2	6	1	10	0	16	1	17	
	ER50 (VV BM)	numerics only	0	0	1	0	1	0	1	
		with cens., f=2	6	1	10	0	16	1	17	
GW 03	ER10 (VV BM)	numerics only	11	1	0	0	11	1	12	11 wild and 4 crop greater-than values
		with cens., f=2	17	1	0	0	17	1	18	
	ER25 (VV BM)	numerics only	11	1	0	0	11	1	12	
		with cens., f=2	17	1	0	0	17	1	18	

	ER50 (VV BM)	numerics only	13	7	0	0	13	7	20	
		with cens., f=2	19	7	1	3	20	10	30	
GW 04	ER25 (VV BM)	numerics only	0	0	0	0	0	0	0	11 wild and 4 crop greater-than val.; 2 smaller-than values.
		with cens., f=2	6	1	10	0	16	1	17	
	ER50 (VV BM)	numerics only	0	0	1	0	1	0	1	
		with cens., f=2	6	1	10	0	16	1	17	
GW 05	ER10 (VV BM)	numerics only	7	1	0	0	7	1	8	Also ER05 endpoints; further crop endpoints in [µmol]#
		with cens., f=2	7	1	0	0	7	1	8	
	ER25 (VV BM)	numerics only	0	8	0	4	0	12	12	
		with cens., f=2	0	8	1	5	1	13	14	
	ER50 (VV BM)	numerics only	8	13	0	4	8	17	25	
		with cens., f=2	8	13	1	5	9	18	27	
GW 06	ER25 (VV BM)	numerics only	7	7	7	4	14	11	25	
		with cens., f=2	7	7	7	4	14	11	25	
GW 07	ER50 (VV BM)	numerics only	1	1	0	1	1	2	3	Further crop endpoints in [µmol]#
		with cens., f=2	1	1	0	1	1	2	3	
GW 08	ER50 (VV BM)	numerics only	0	0	2	0	2	0	2	Further crop endpoints in [µmol]#
		with cens., f=2	0	0	2	0	2	0	2	
GW 09	ER25 (VV BM)	numerics only	9	5	7	4	16	9	25	
		with cens., f=2	9	5	7	4	16	9	25	
GW 10	ER10 (VV BM)	numerics only	6	0	0	0	6	0	6	
		with cens., f=2	6	0	0	0	6	0	6	
	ER25 (VV BM)	numerics only	2	5	1	2	3	7	10	
		with cens., f=2	2	5	1	3	3	8	11	
	ER50 (VV BM)	numerics only	14	5	0	2	14	7	21	
		with cens., f=2	14	5	1	3	15	8	23	
GW 11	ER25 (VV BM)	numerics only	4	5	3	3	7	8	15	
		with cens., f=2	5	5	3	3	8	8	16	
GW 13mix	ER50 (VV BM)	numerics only	7	9	1	3	8	12	20	Further endpoints given only as µmol#
		with cens., f=2	7	9	1	3	8	12	20	
ICD 1	ER50 (VV BM)	numerics only	0	4	0	5	0	9	9	Further endpoints given only as µmol#
		with cens., f=2	0	4	0	5	0	9	9	
ICD 2	ER50 (VV BM)	numerics only	0	0	0	0	0	0	0	
		with cens., f=2	0	1	0	0	0	1	1	
ICD 3	ER50 (VV BM)	numerics only	0	1	0	1	0	2	2	Further crop endpoints in [µmol]#
		with cens., f=2	0	1	0	1	0	2	2	
LSI 1	ER25 (VV BM)	numerics only	1	1	1	1	2	2	4	
		with cens., f=2	6	7	11	3	17	10	27	
	ER50 (VV BM)	numerics only	0	3	6	5	6	8	14	Also SH endpoints. Further endpoints expressed in µmol#
		with cens., f=2	6	8	13	6	19	14	33	
LSI 2	ER25 (VV BM)	numerics only	0	5	0	3	0	8	8	Also SH endpoints; further endpoints expressed only in µmol#
		with cens., f=2	0	6	0	4	0	10	10	
	ER50 (VV BM)	numerics only	0	4	3	1	3	5	8	
		with cens., f=2	0	7	3	4	3	11	14	
LSI 3	ER50 (VV BM)	numerics only	0	1	0	1	0	2	2	Further endpoints expressed in µmol#
		with cens., f=2	0	1	0	1	0	2	2	
OTH 2	ER25 (VV BM)	numerics only	0	0	2	0	2	0	2	
		with cens., f=2	6	1	10	0	16	1	17	
	ER50 (VV BM)	numerics only	0	0	2	0	2	0	2	
		with cens., f=2	6	1	10	0	16	1	17	

OTH 3	ER50 (VV BM)	numerics only	2	0	0	0	2	0	2	
		with cens., f=2	2	0	0	0	2	0	2	
OTH 6	ER25 (VV BM)	numerics only	7	7	7	3	14	10	24	
		with cens., f=2	7	7	7	3	14	10	24	
OTH 7	ER50 (VV BM)	numerics only	2	0	0	0	2	0	2	
		with cens., f=2	2	0	0	0	2	0	2	
OTH 8	ER25 (VV BM)	numerics only	0	0	0	0	0	0	0	14 wild endpoints greater than, 2 wild and 1 crop endpoints less than.
		with cens., f=2	6	1	10	0	16	1	17	
	ER50 (VV BM)	numerics only	0	1	0	0	0	1	1	
		with cens., f=2	6	1	10	0	16	1	17	
PHI 01	ER25 (VV BM)	numerics only	1	7	1	4	2	11	13	Also ER05 endpoints and SH endpoints
		with cens., f=2	6	9	11	4	17	13	30	
	ER50 (VV BM)	numerics only	1	6	2	4	3	10	13	
		with cens., f=2	6	9	11	4	17	13	30	
PHI 02	ER10 (VV BM)	numerics only	0	0	0	0	0	0	0	Also ER05 endpoints; Further endpoints expressed in µmol#
		with cens., f=2	0	6	1	3	1	9	10	
	ER25 (VV BM)	numerics only	31	12	7	3	38	15	53	
		with cens., f=2	31	12	7	4	38	16	54	
	ER50 (VV BM)	numerics only	10	5	2	2	14	7	21	
		with cens., f=2	10	5	2	3	14	8	22	
PHI 03	ER25 (VV BM)	numerics only	0	4	0	1	0	5	5	Also SH endpoints. Further endpoints expressed in µmol#
		with cens., f=2	0	4	0	3	0	7	7	
	ER50 (VV BM)	numerics only	14	6	1	0	15	6	21	
		with cens., f=2	14	7	2	3	16	10	26	
PHI 04	ER25 (VV BM)	numerics only	0	5	0	3	0	8	8	
		with cens., f=2	0	6	1	3	1	9	10	
	ER50 (VV BM)	numerics only	20	5	10	3	31	8	40	
		with cens., f=2	20	6	11	3	32	9	42	
PHI 06	ER50 (VV BM)	numerics only	0	3	0	1	0	4	4	only crop endpoints
		with cens., f=2	0	3	0	1	0	4	4	
PHI 07	ER25 (VV BM)	numerics only	5	4	1	1	6	5	11	only crop endpoints
		with cens., f=2	5	5	3	3	8	8	16	
	ER50 (VV BM)	numerics only	0	2	0	0	0	2	3	
		with cens., f=2	0	7	0	3	0	10	11	
PHI 08	ER50 (VV BM)	numerics only	0	0	0	0	2	0	2	little wild and no crop data, further ER50 endpoints expressed in µmol#
		with cens., f=2	0	0	0	0	2	0	2	
PHI 09	ER50 (VV BM)	numerics only	2	2	0	1	2	3	5	Further endpoints expressed in µmol#
		with cens., f=2	2	3	0	1	2	4	6	
PHI 10	ER50 (VV BM)	numerics only	0	2	0	0	0	2	2	only crop endpoints
		with cens., f=2	0	2	0	0	0	2	2	
PHI 12	ER25 (VV BM)	numerics only	0	0	0	1	0	1	1	only crop endpoints
		with cens., f=2	0	0	0	1	0	1	1	
	ER50 (VV BM)	numerics only	0	0	0	1	0	1	1	only crop endpoints
		with cens., f=2	0	0	0	1	0	1	1	
SGI 1	ER25 (VV BM)	numerics only	0	0	0	1	0	1	1	Numeric endpoints only for wild sp. CROP endpoints only for survival. Also crop SE endpoints but either SH

		with cens., f=2	0	4	0	2	0	6	6	or BM endp. listed. All crop VV BM greater-thans (old substance, no longer marketed)
	ER50 (VV BM)	numerics only	14	0	1	0	15	0	15	Numeric endpoints only for wild sp. CROP endpoints only for survival. Also crop SE endpoints but either SH or BM endp. listed. All crop VV BM greater-thans (old substance, no longer marketed)
		with cens., f=2	14	4	1	2	15	6	21	
SGI 2*	ER10 (VV BM)	numerics only	0	6	1	3	1	9	10	ER10 and ER25 only one wild species endpoint each
		with cens., f=2	0	6	1	3	1	9	10	
	ER25 (VV BM)	numerics only	0	6	1	3	1	9	10	
		with cens., f=2	0	6	1	3	1	9	10	
	ER50 (VV BM)	numerics only	14	6	2	3	16	9	25	
		with cens., f=2	14	6	2	3	16	9	25	
SGI 3	ER25 (VV BM)	numerics only	6	8	7	5	13	13	26	11 wild and 1 crop ER50 endpoints greater-than values; also SH endpoints; further endpoints only expressed in µmol#
		with cens., f=2	7	12	8	8	15	20	35	
	ER50 (VV BM)	numerics only	0	1	2	0	2	1	3	
		with cens., f=2	1	6	3	4	4	10	14	
SGI 4	ER50 (VV BM)	numerics only	0	1	2	1	2	2	4	Further endpoints expressed in µmol#
		with cens., f=2	0	1	2	1	2	2	4	
SGI 5	ER50 (VV BM)	numerics only	0	0	2	0	2	0	2	no crop and just two wild sp. endpoints
		with cens., f=2	0	0	2	0	2	0	2	

Notes:

- Endpoints in the lower line are censored endpoints.

* Crop endpoints only from a.s. (tested on technical material), wild species endpoints from field tests performed with commercial formulations. However, even if this certainly not representative dataset is included, the overall outcome is unchanged.

[#] Endpoints that are given only as [µmol] (no reference to any quantitative measure such as area) cannot be transformed into field rates (not even in original paper, Fletcher et al. 1985)

12 Appendix 5 - Lists of tables and figures

Table 1 Modes of action (MoA) of the active substances for which terrestrial plant species endpoints were available. Modes of actions not listed were either not available (no non-target-plant endpoints) or merged with other modes of action. 11

Table 2: AASI xx: Lowest endpoint and Geometric mean (alternatively median) of wild plant species and crop species, and resulting quotients. Quotients (crop/wild): Values < 1 indicate that crop endpoints were lower, i.e. crops more sensitive than wild species, and *quotients > 1 that wild plant species' endpoints were lower*, i.e. wild species more sensitive than crop species. Any quotient smaller than '0.2' or greater than '5' is printed in bold. 21

Table 3: Summary of all sets where a quotient could be calculated considering ER10, ER25 and ER50 endpoints from VV studies / field studies based on biomass data. Quotients calculated from lowest endpoints (minima), and from the geometric means. A quotient x greater than 1 indicates that wild species were more sensitive (by factor x) than crop species, quotients below 1 indicate the opposite. Quotients above 5 or below 0.2 are printed in bold, those above 5 (indicating that wild species were more sensitive than crops) are underlined. Resulting overall quotients between crop species and wild plant species are shown at the bottom; (a) as overall geometric mean of all quotients (not weighted), (b), as weighted geometric mean (weighting based on the lower 'n' of each pair), (c) medians of all quotients. 29

Table 4: Summary of fitted factorial models (no interaction) to ER10, ER25, ER50, and pooled data. Positive coefficient signs indicate that the predictor at the right end of the predictor code was on average higher, negative quotient signs that it was lower. E.g. 'CW.finW': ER10, '-0.37': From the groups C (crops) and W (wild) endpoints of the latter (W) were lower than the former, i.e. wild species were more sensitive than crop species, whereas based on ER25 it was the other way round (coefficient +0.65) wild plant endpoints higher than crop endpoints. Standard errors of the coefficient estimation inside parenthesis. 36

Table 5: Summary of two runs with modified data, fitted factorial model (with interactions or no interaction), where based on the original data all crop endpoints had been increased by a factor of 1.5 or 2. Positive coefficient signs indicate positive deviations, negative signs negative deviations compared to the corresponding baseline, e.g. that that wild species were more sensitive than crops species. The figures are based on the log-transformed data (natural logarithm). Inside parentheses the standard error of the coefficient estimation. 39

Table 6: Comparison of outcomes based on ER25 (VV BM) and ER50 (VV BM). 41

Table 7: Comparison of endpoints of species/active substance/endpoint combinations tested in the laboratory/greenhouse and in semi-field/field test systems (Option A = listed by active substance). Numbers in brackets give the range per substance/endpoint combination. If just one figure, there had been only one species with both endpoints (i.e. just one quotient). 45

Table 8: List of all active substances for which published plant endpoints were available. For some of them only crop endpoints, for others only wild species endpoints were available; also these include cases with only censored endpoints. 94

Table 9: List of all plant species (wild and crop) with vegetative endpoints from experiments on young plants (NOER, NOEC or ERx ICx or ECx). 96

Table 10: List of substances and numbers of numeric endpoints differentiating by monocots/dicots and wild and crop species. The column labelled “endpoints” indicates the endpoint’s effect level and -type (here only vegetative of juvenile plants, biomass, and the figures the resulting number of species*type (lab or field) with numeric endpoints or with any endpoint (second row each). 103

Table 11: AASI 1: Standardized and anonymized Geometric mean and lowest endpoint of wild plant species and crop species, differentiated by systematic group (Monocot/dicot). Numeric ER25 (biomass) endpoints from lab and field tests. 121

Table 12: AASI 1: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above. 122

Table 13: AASI02: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above. 123

Table 14: AASI03a: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above. 125

Table 15: AASI03a: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above. 126

Table 16: AASI3b: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above. 127

Table 17: AASI03(all): Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER10 endpoints. Parameter selection see within table, further explanations see text and figure above. 128

Table 18: AASI03(all): Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above. 129

Table 19: AASI3(all): Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above. 130

Table 20: AASI04: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 131

Table 21: AASI04: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 132

Table 22: AASI 5: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 133

Table 23: AASI 6: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 134

Table 24: AASI 7: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 135

Table 25: AASI14: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 136

Table 26: AASI15: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 137

Table 27: AASI17: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER10 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 138

Table 28: AASI17: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 139

Table 29: AASI17: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 140

Table 30: CMD 1: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 141

Table 31: GW 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER10 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 142

Table 32: GW 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER10 endpoints. Parameter selection see within table, further explanations see text and figure above. 143

Table 33: GW 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above. 144

Table 34: GW 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above. 145

Table 35: GW03: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above. 146

Table 36: GW 05: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above. 147

Table 37: GW 06: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above. 148

Table 38: GW 09: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above. 149

Table 39: GW 10: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above. 150

Table 40: GW 10: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above. 151

Table 41: GW 11: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above. 152

Table 42: GW 13: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above. 153

Table 43: LSI 1: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above. 154

Table 44: LSI 1: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 155

Table 45: LSI 2: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 156

Table 46: OTH 6: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 157

Table 47: PHI 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 158

Table 48: PHI 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 159

Table 49: PHI 02: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 160

Table 50: PHI 02: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 161

Table 51: PHI03: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 162

Table 52: PHI 04: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 163

Table 53: PHI 07: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 164

Table 54: PHI 09: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 165

Table 55: SGI2: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 166

Table 56: SGI 3: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above....................................... 167

Figures:

Figure 1 Example plot visualising the approach from the individual species' endpoints via SSD. Distribution of ER25 endpoints (biomass; greenhouse and field endpoints combined) of wild plant and crop species. Abscissa indicating normalized field rates (endpoints divided by the geometric mean of all data). Species sorted by their sensitivity. Further explanations see text. 20

Figure 2: Example plot visualising the approach from the individual species' endpoints via SSD to the box plots used throughout the paper. Distribution of ER25 endpoints (biomass) of wild plant and crop species. The rhombus marks the geometric mean of data points... 20

Figure 3: Upper figure A based on numeric vegetative biomass endpoints, lower figure B also including seedling emergence studies, other parameters such as shoot height or survival, and considered censored values with a correction factor of 2. Quotients (triangles and rhombi) of the individual cases (combinations of substances and endpoints) on a log scale, sorted by average quotients (of minima- and geometric means-quotients) in ascending order. Cases above 1 indicate that wild species were more sensitive than crop species, cases below 1 that crop species were more sensitive; either based on their central value (geometric mean) = triangles, or on the groups' minima (comparison of the most sensitive species of each group) = rhombi.. The green circles at the bottom indicate the number of species (lower n of the compared groups, secondary ordinate)... 32

Figure 4: Boxplot of the endpoints (log10 scale) for different Mode of Action (MoA), Class (M,D or Gymn), and Lab/Field categories.. 37

Figure 5: Differences in sensitivity between wild and crop species based on ER25 or on ER50 endpoints. The point signatures visualise the quotients (above 1 indicating wild species being more sensitive than crop species, below 1 crop species being the more sensitive group). Each set is either based on the minima, i.e. the most sensitive species of each group, or on the central points of each group (geometric mean of the group's endpoints). The boxes indicate 25%ile and 75%ile, with the central line displaying the geometric mean as central point. For numeric values see Table 6; further explanations see text. 43

Figure 6: Differences between lab and field endpoints (wild species only), again expressed as quotients. Quotients below 1 indicate that the lab endpoint was lower than the field endpoint, greater 1 that the field endpoint was lower, hence more sensitive; further explanations see text. 46

Figure 7: Differences between lab and field endpoints (crop species only), again expressed as quotients. Quotients below 1 indicate that the lab endpoint was lower than the

field endpoint, greater 1 that the field endpoint was lower, hence more sensitive; further explanations see text. 47

Figure 11: Substance AASI01 - Distribution of ER25 endpoints (biomass) of wild plant and crop species. Presentation by species – in cases of multiple testing the endpoint used is geometric mean per species. Endpoints from lab and field tests. The rhombus marks the geometric mean of data points. 121

Figure 12: Substance AASI 1 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 122

Figure 13: Substance AASI02 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 123

Figure 14: Substance AASI03a - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 125

Figure 15: Substance AASI03a - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 126

Figure 16: Substance AASI03b - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 127

Figure 17: Substance AASI03(all) - Distribution of numeric ER10 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 128

Figure 18: Substance AASI03(all) - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 129

Figure 19: Substance AASI3(all) - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 130

Figure 20: Substance AASI04 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 131

Figure 21: Substance AASI04 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 132

Figure 22: Substance AASI 5 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 133

Figure 23: Substance AASI 6 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 134

Figure 24: Substance AASI 7 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 135

Figure 25: Substance AASI14 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 136

Figure 26: Substance AASI15 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 137

Figure 27: Substance AASI17 - Distribution of numeric ER10 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 138

Figure 28: Substance AASI17 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 139

Figure 29: Substance AASI17 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 140

Figure 30: Substance CMD 1 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 141

Figure 31: Substance GW 01 - Distribution of numeric ER10 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 142

Figure 32: Substance GW 01 - Distribution of numeric ER10 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 143

Figure 33: Substance GW 01 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 144

Figure 34: Substance GW 01 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 145

Figure 35: Substance GW03 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 146

Figure 36: Substance GW 05 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 147

Figure 37: Substance GW 06 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 148

Figure 38: Substance GW 09 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 149

Figure 39: Substance GW 10 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 150

Figure 40: Substance GW 10 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 151

Figure 41: Substance GW 11 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 152

Figure 42: Substance GW13 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 153

Figure 43: Substance LSI 1 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 154

Figure 44: Substance LSI 1 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 155

Figure 45: Substance LSI 2 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 156

Figure 46: Substance OTH 6 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 157

Figure 47: Substance PHI 01 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 158

Figure 48: Substance PHI 01- Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 159

Figure 49: Substance PHI 02 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 160

Figure 50: Substance PHI 02 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 161

Figure 51: Substance PHI03 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 162

Figure 52: Substance PHI 04 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 163

Figure 53: Substance PHI 07 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 164

Figure 54: Substance PHI 09 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 165

Figure 55: Substance SGI2 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 166

Figure 56: Substance SGI 3 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians. 167

13 Appendix 6 - Individual data sets presented as figures and summarized in tables

On the following pages individual comparison of datasets by substance and type of endpoint are depicted in order to visualize the (lack of) differences in sensitivity between wild plants and crop species for individual active substances. To give the numbers of species already in the figure, the following abbreviations were used:

- m^c = monocot crop species
- m^w = monocot wild species
- d^c = dicot crops
- d^w = dicot wild species

Below ER10, ER25 and ER50 endpoints are displayed, from vegetative vigour lab studies or similar field studies with foliar application; biomass data (lumping however fresh-weight and dry-weight based evaluations[47]).

Quotients around 1 indicate that there was no pronounced difference in sensitivity between crop species and wild plant species, quotients > 1 indicate that wild plant endpoints were lower than crop endpoints, and quotients < 1 that crop endpoints were lower than wild plant endpoints. For the overall availability of data and endpoints see Table 3, this is also where occurrences of censored endpoints are listed. Datasets with fewer than 2 numeric endpoints in a group were not displayed.

[47] While it is appreciated that wet weight and dry weight-based endpoints may sometimes differ considerably (hence separate evaluation would be desirable), often the papers do not indicate on which of the two measurements their biomass-derived endpoint are based. Also differentiating between wet weight and dry weight-based endpoints would have reduced the numbers of endpoints and hence the number of substances for which sufficient endpoints for an individual assessment were available.

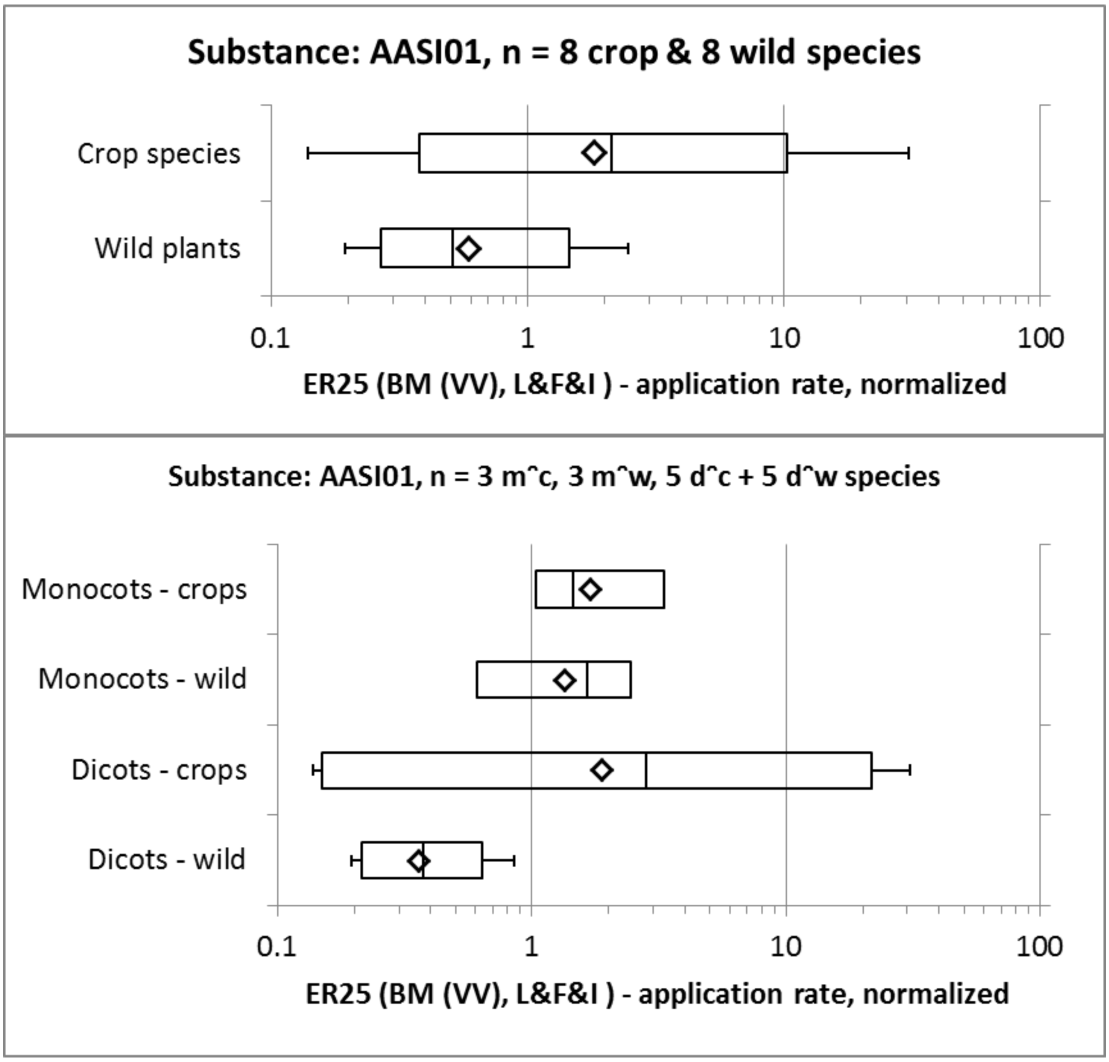


Figure 11: Substance AASI01 - Distribution of <u>ER25</u> endpoints (<u>biomass</u>) of wild plant and crop species. Presentation by species – in cases of multiple testing the endpoint used is geometric mean per species. Endpoints from <u>lab and field tests</u>. The rhombus marks the geometric mean of data points.

Table 11: AASI 1: Standardized and anonymized Geometric mean and lowest endpoint of wild plant species and crop species, differentiated by systematic group (Monocot/dicot). Numeric ER25 (biomass) endpoints from lab and field tests.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI01**	Mono-cots	Crop	3 (species)	1.0362	*1.7*	1.705	*1.3*
Endpoint	ER25		Wild	3 (species)	0.6053		1.347	
Measured	BM	Dicots	Crop	5 (species)	0.1386	0.71	1.887	**5.2**
Lab/field	L&F&I		Wild	5 (species)	0.1948		0.360	
SE, VV or all	VV	All	Crop	8 (species)	0.1386	0.71	1.817	*3.1*
Censored val.	numeric		Wild	8 (species)	0.1948		0.590	

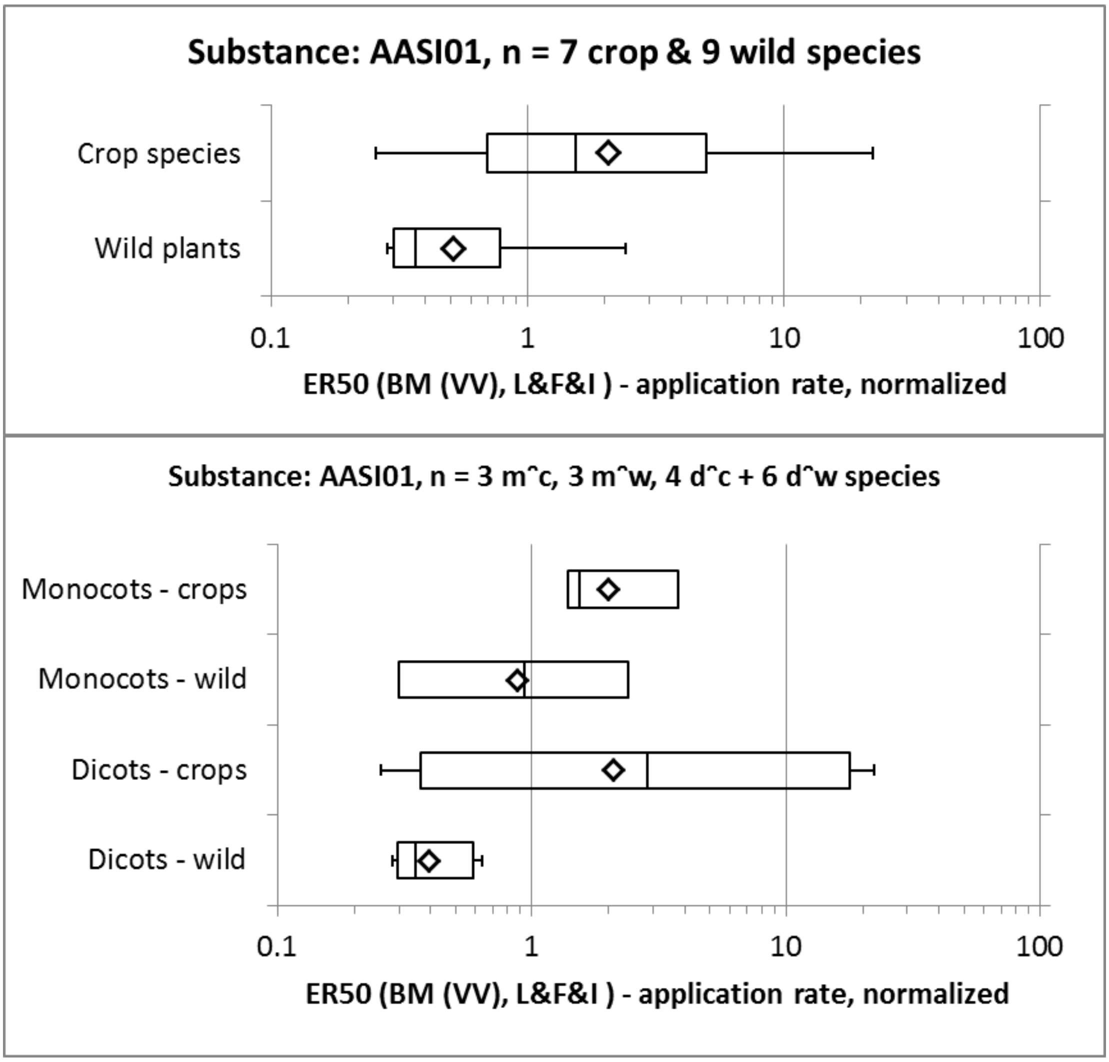


Figure 12: Substance AASI 1 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 12: AASI 1: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI01**	Mono-cots	Crop	3 (species)	1.3869	*4.6*	2.003	*2.3*
Endpoint	ER50		Wild	3 (species)	0.3016		0.877	
Measured	BM	Dicots	Crop	4 (species)	0.2556	0.9	2.106	**5.3**
Lab/field	L&F&I		Wild	6 (species)	0.2831		0.395	
SE, VV or all	VV	All	Crop	7 (species)	0.2556	0.9	2.061	*4*
Censored val.	numeric		Wild	9 (species)	0.2831		0.515	

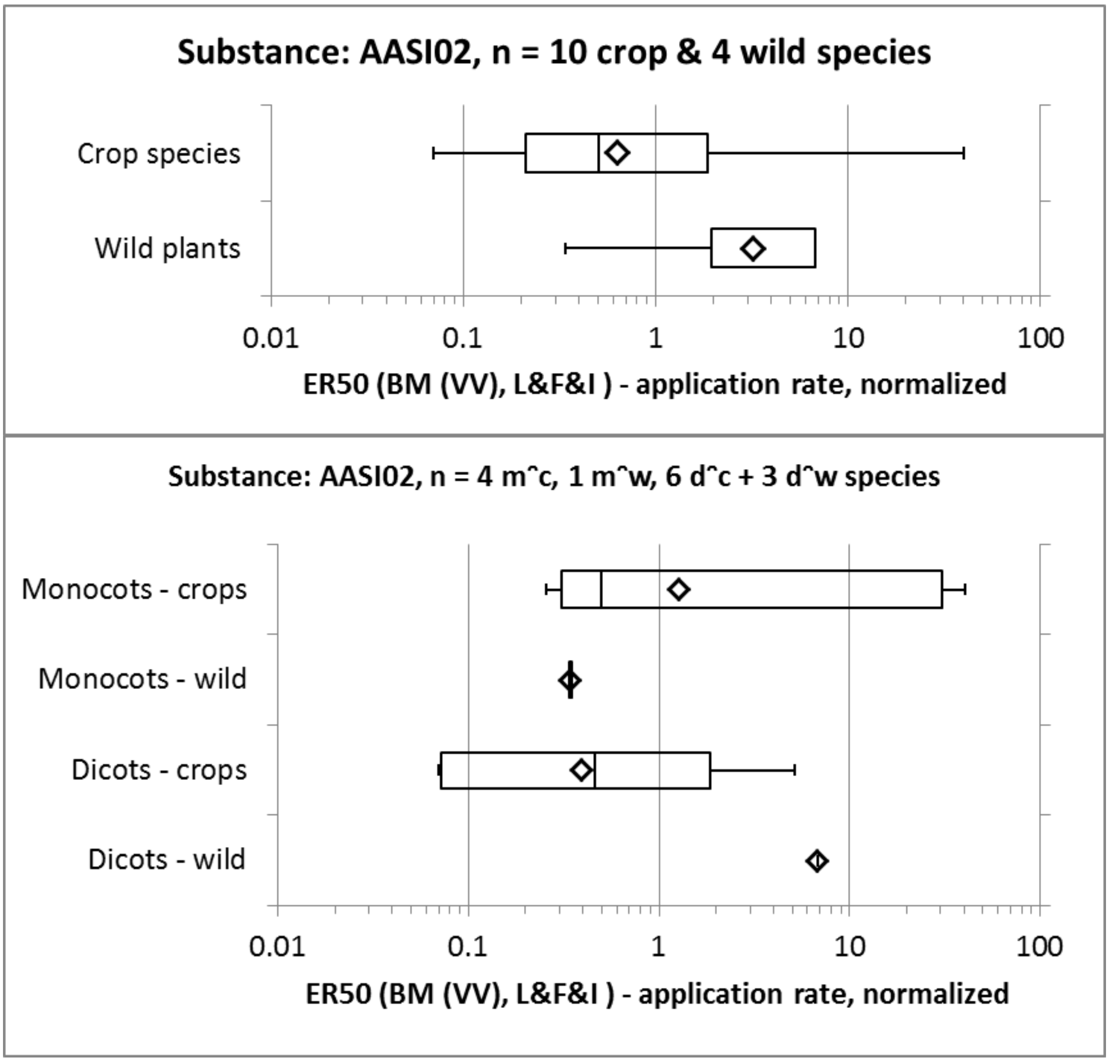


Figure 13: Substance AASI02 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 13: AASI02: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI02**	Mono-cots	Crop	4 (species)	0.2561	0.75	1.268	*3.7*
Endpoint	ER50		Wild	1 (species)	0.3413		0.341	
Measured	BM	Dicots	Crop	6 (species)	0.0697	**0.01**	0.392	**0.058**
Lab/field	L&F&I		Wild	3 (species)	6.7717		6.772	
SE, VV or all	VV	All	Crop	10 (species)	0.0697	0.2	0.627	0.2
Censored val.	numeric		Wild	4 (species)	0.3413		3.209	

Example interpretation:

For substance AASI 2 only ER50 data were available, and most of the wild species data were censored (12 species greater-than x g/ha and one less than x g/ha, three with effect levels around 50% were included as numeric ER50 = x g/ha. For crop species there were numeric endpoints for 10 species. Here quotients were negative and very large, hence crop species appear to be more sensitive than wild species by several orders of magnitude. While numeric endpoints could be determined only for three wild species, the finding that all but one of the censored endpoints were greater than values confirms that based on the published endpoints wild species were indeed considerably less sensitive to AASI 2 than crop species..

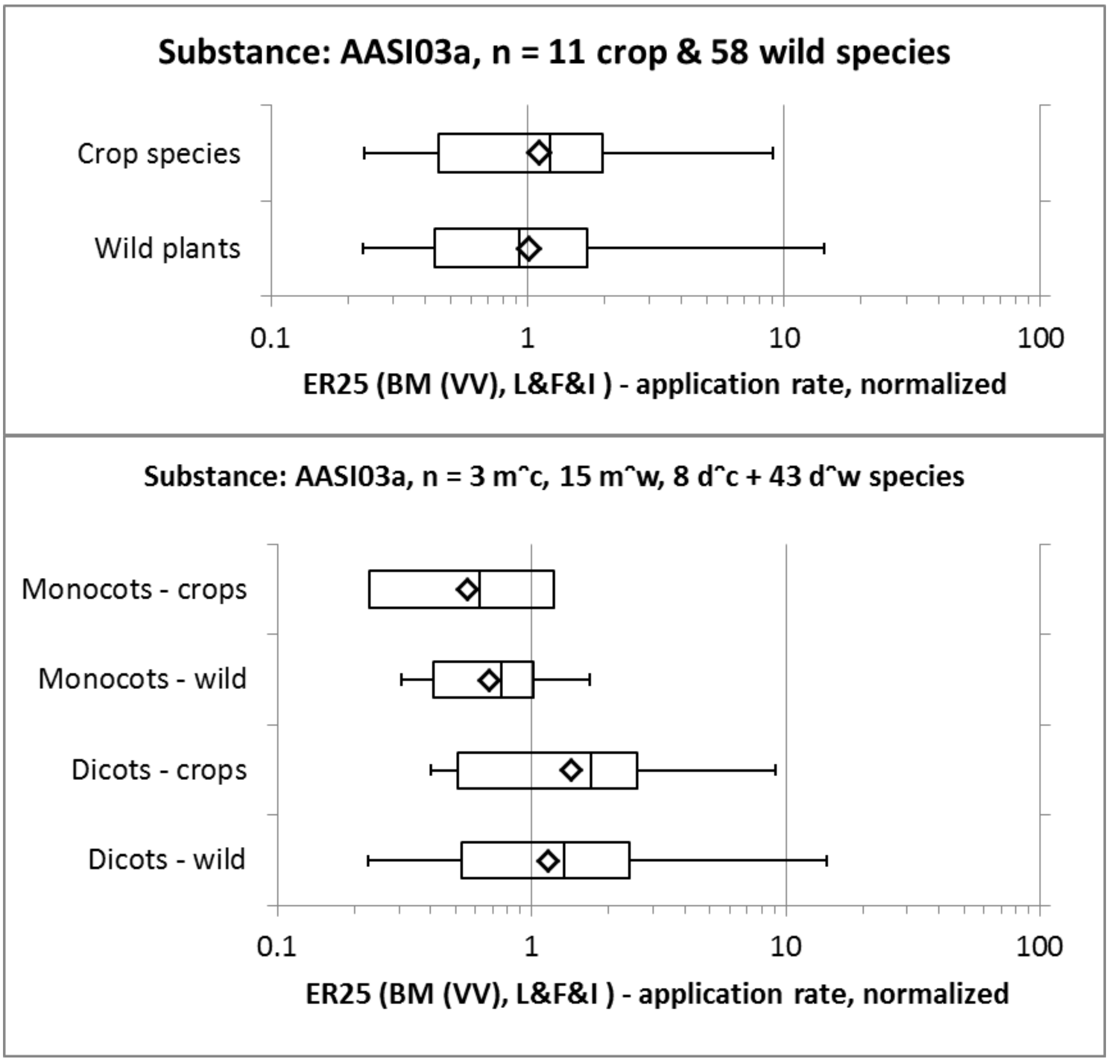


Figure 14: Substance AASI03a - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 14: AASI03a: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI3a**	Mono-cots	Crop	3 (species)	0.2294	0.75	0.559	0.82
Endpoint	ER25		Wild	15 (species)	0.3062		0.677	
Measured	BM	Dicots	Crop	8 (species)	0.4014	*1.8*	1.435	*1.2*
Lab/field	L&F&I		Wild	43 (species)	0.2271		1.155	
SE, VV or all	VV	All	Crop	11 (species)	0.2294	1.0	1.109	*1.1*
Censored val.	numeric		Wild	58 (species)	0.2271		1.006	

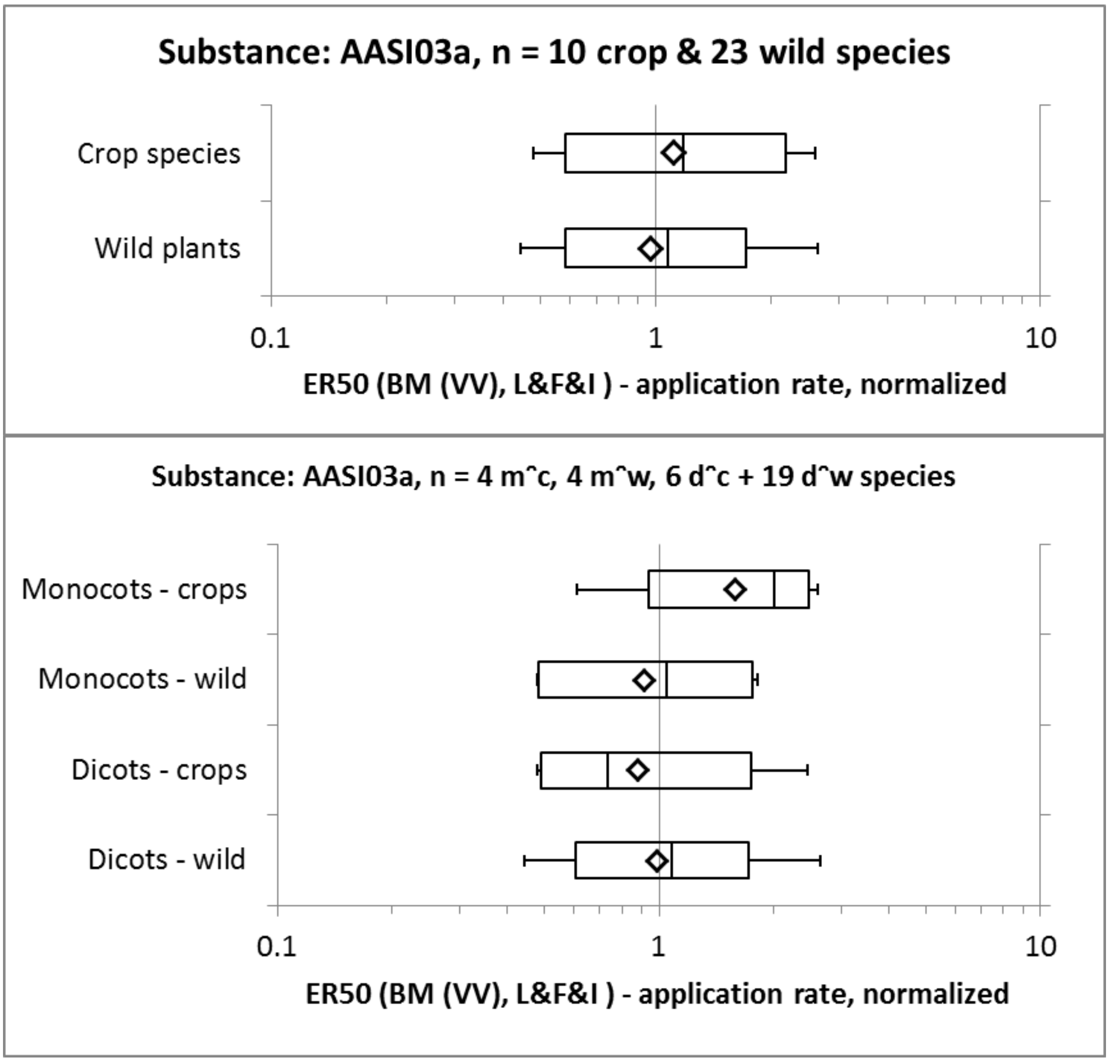


Figure 15: Substance AASI03a - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 15: AASI03a: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI03a**	Mono-cots	Crop	4 (species)	0.6101	*1.3*	1.589	*1.7*
Endpoint	ER50		Wild	4 (species)	0.4791		0.912	
Measured	BM	Dicots	Crop	6 (species)	0.4791	*1.1*	0.881	0.89
Lab/field	L&F&I		Wild	19 (species)	0.4448		0.985	
SE, VV or all	VV	All	Crop	10 (species)	0.4791	*1.1*	1.115	*1.1*
Censored val.	numeric		Wild	23 (species)	0.4448		0.971	

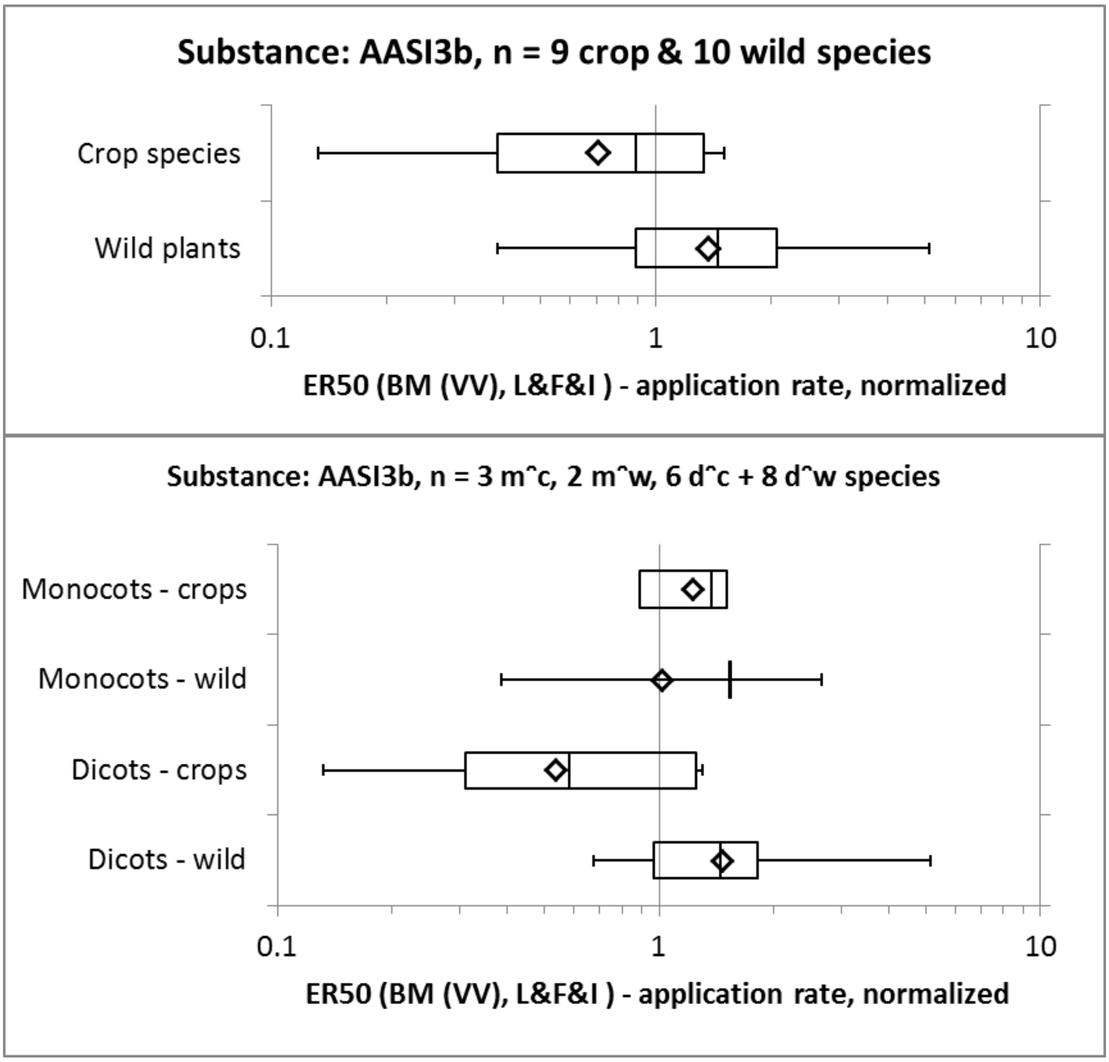


Figure 16: Substance AASI03b - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 16: AASI3b: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI3b**	Mono-cots	Crop	3 (species)	0.8908	*2.3*	1.226	*1.2*
Endpoint	ER50		Wild	2 (species)	0.3873		1.018	
Measured	BM	Dicots	Crop	6 (species)	0.1322	0.2	0.538	0.37
Lab/field	L&F&I		Wild	8 (species)	0.6731		1.468	
SE, VV or all	VV	All	Crop	9 (species)	0.1322	0.34	0.708	0.52
Censored val.	numeric		Wild	10 (species)	0.3873		1.365	

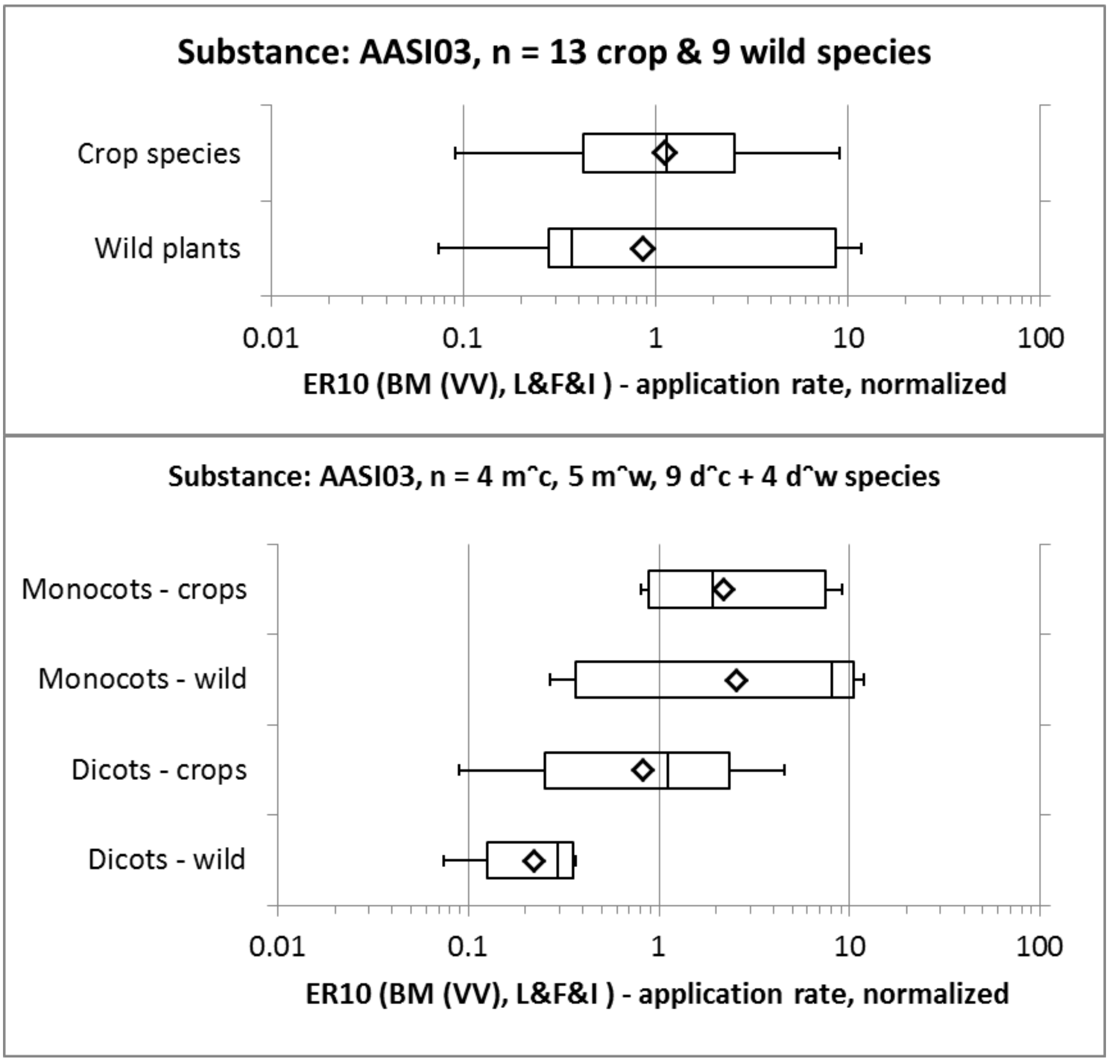


Figure 17: Substance AASI03(all) - Distribution of numeric ER10 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 17: AASI03(all): Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER10 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI03**	Mono-cots	Crop	4 (species)	0.8064	*3*	2.181	0.86
Endpoint	ER10		Wild	5 (species)	0.2683		2.543	
Measured	BM	Dicots	Crop	9 (species)	0.0899	*1.2*	0.823	*3.7*
Lab/field	L&F&I		Wild	4 (species)	0.0740		0.221	
SE, VV or all	VV	All	Crop	13 (species)	0.0899	*1.2*	1.111	*1.3*
Censored val.	numeric		Wild	9 (species)	0.0740		0.859	

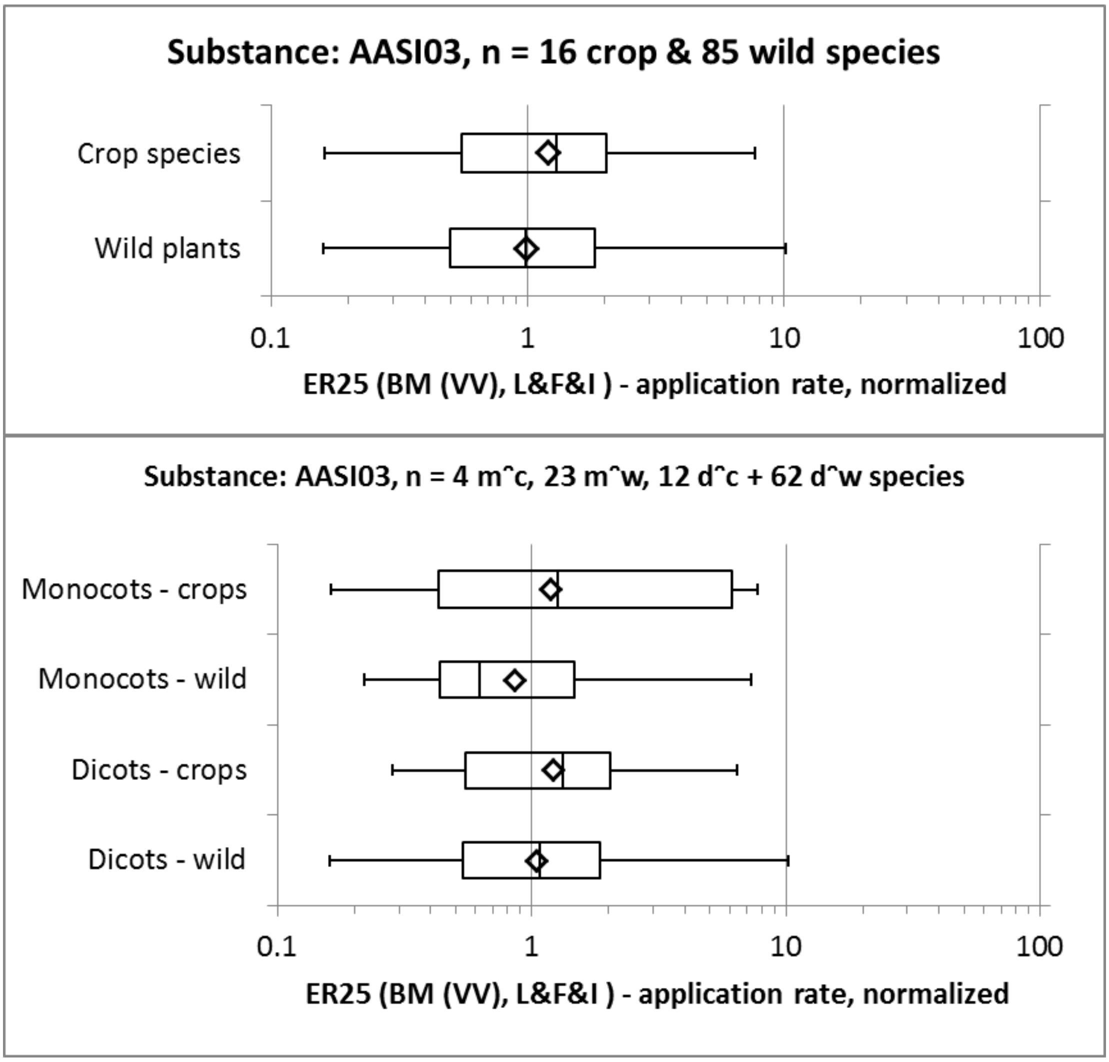


Figure 18: Substance AASI03(all) - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 18: AASI03(all): Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI03**	Mono-cots	Crop	4 (species)	0.1618	0.74	1.185	*1.4*
Endpoint	ER25		Wild	23 (species)	0.2184		0.853	
Measured	BM	Dicots	Crop	12 (species)	0.2831	*1.8*	1.213	*1.2*
Lab/field	L&F&I		Wild	62 (species)	0.1601		1.041	
SE, VV or all	VV	All	Crop	16 (species)	0.1618	1	1.206	*1.2*
Censored val.	numeric		Wild	85 (species)	0.1601		0.987	

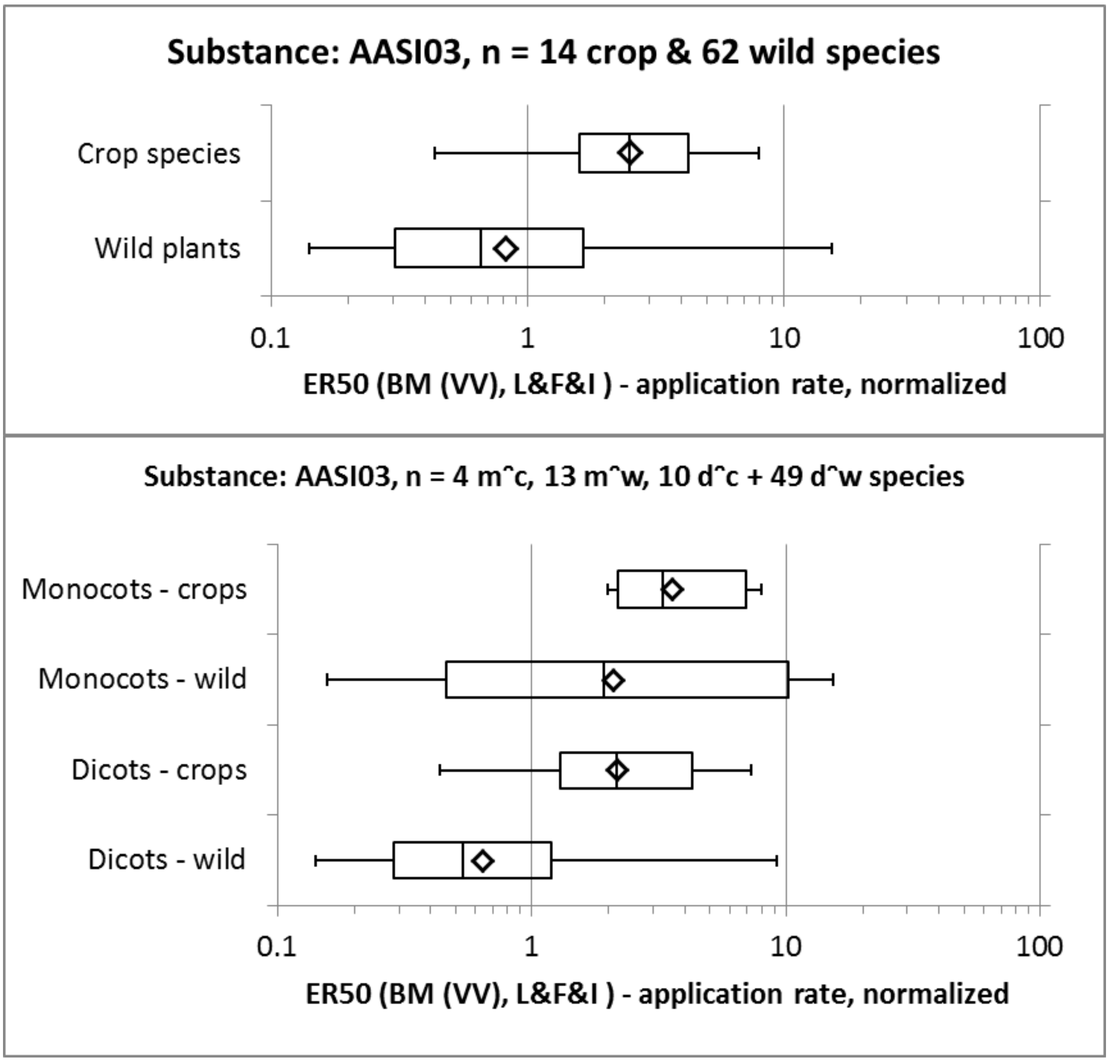


Figure 19: Substance AASI3(all) - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 19: AASI3(all): Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI03**	Mono-cots	Crop	4 (species)	1.9856	**13**	3.590	*1.7*
Endpoint	ER50		Wild	13 (species)	0.1562		2.090	
Measured	BM	Dicots	Crop	10 (species)	0.4367	*3.1*	2.166	*3.4*
Lab/field	L&F&I		Wild	49 (species)	0.1413		0.644	
SE, VV or all	VV	All	Crop	14 (species)	0.4367	*3.1*	2.502	*3*
Censored val.	numeric		Wild	62 (species)	0.1413		0.825	

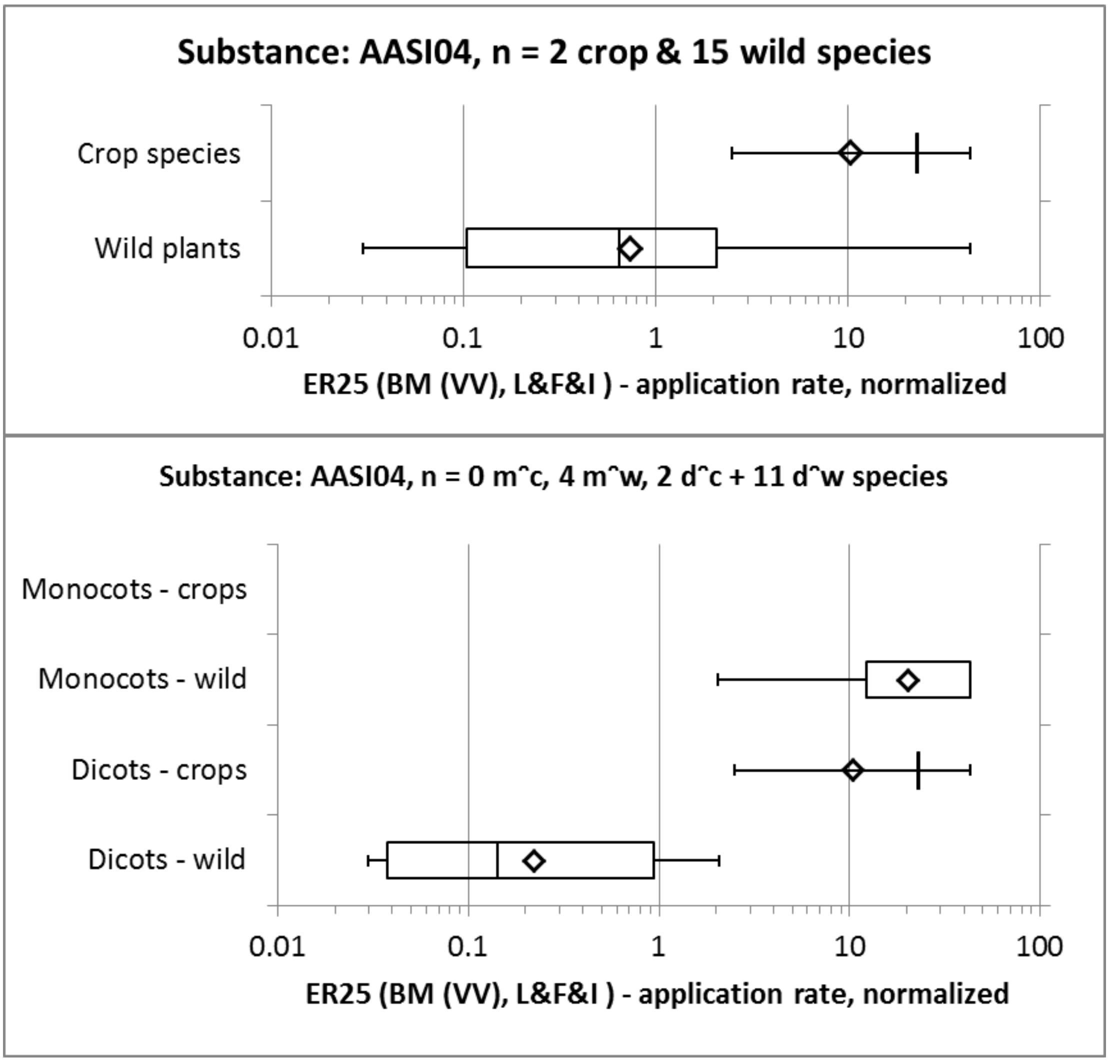
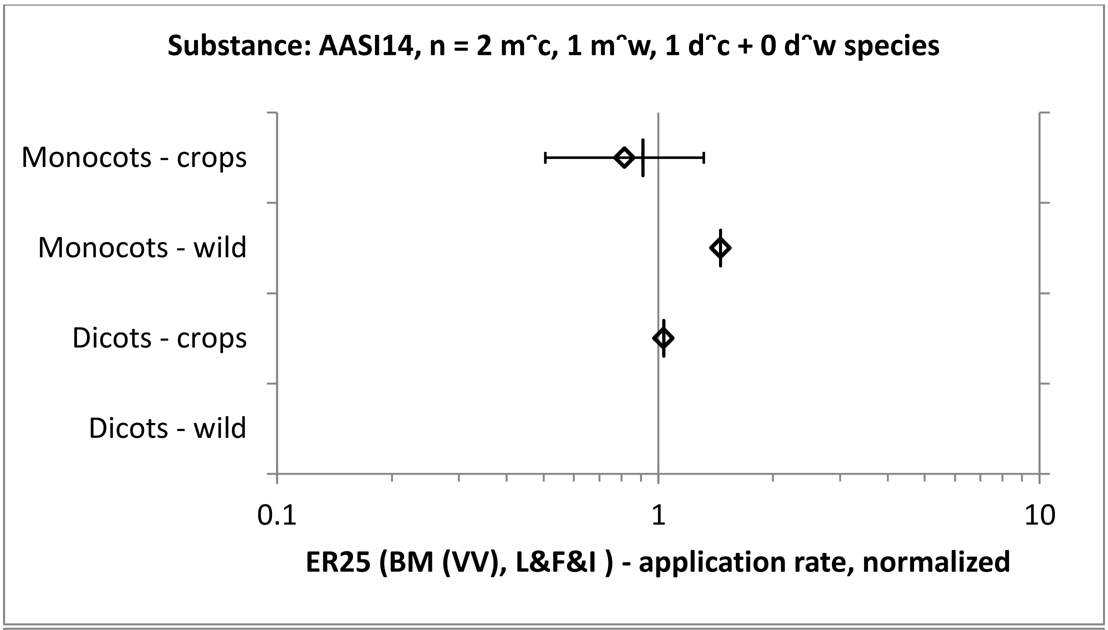


Figure 20: Substance AASI04 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 20: AASI04: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	AASI04	Mono-cots	Crop	0 (species)	#NUM!	#NUM!	#NUM!	#NUM!
Endpoint	ER25		Wild	4 (species)	2.0294		20.111	
Measured	BM	Dicots	Crop	2 (species)	2.4784	83	10.347	47
Lab/field	L&F&I		Wild	11 (species)	0.0299		0.220	
SE, VV or all	VV	All	Crop	2 (species)	2.4784	83	10.347	14
Censored val.	numeric		Wild	15 (species)	0.0299		0.732	



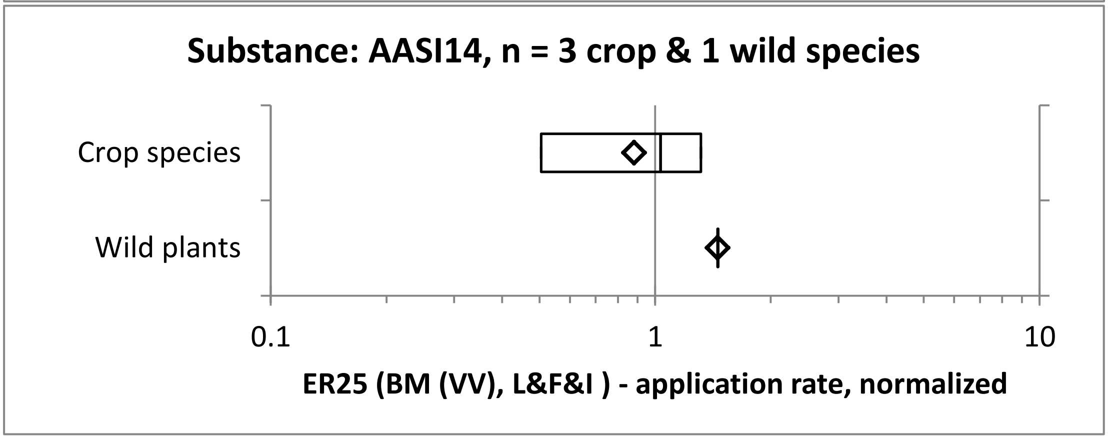


Figure 21: Substance AASI04 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 21: AASI04: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	AASI04	Mono-cots	Crop	4 (species)	1.4088	#NUM!	1.782	#NUM!
Endpoint	ER50		Wild	0 (species)	#NUM!		#NUM!	
Measured	BM	Dicots	Crop	10 (species)	0.0900	2.1	0.789	0.78
Lab/field	L&F&I		Wild	10 (species)	0.0436		1.006	
SE, VV or all	VV	All	Crop	14 (species)	0.0900	2.1	0.996	0.99
Censored val.	numeric		Wild	10 (species)	0.0436		1.006	

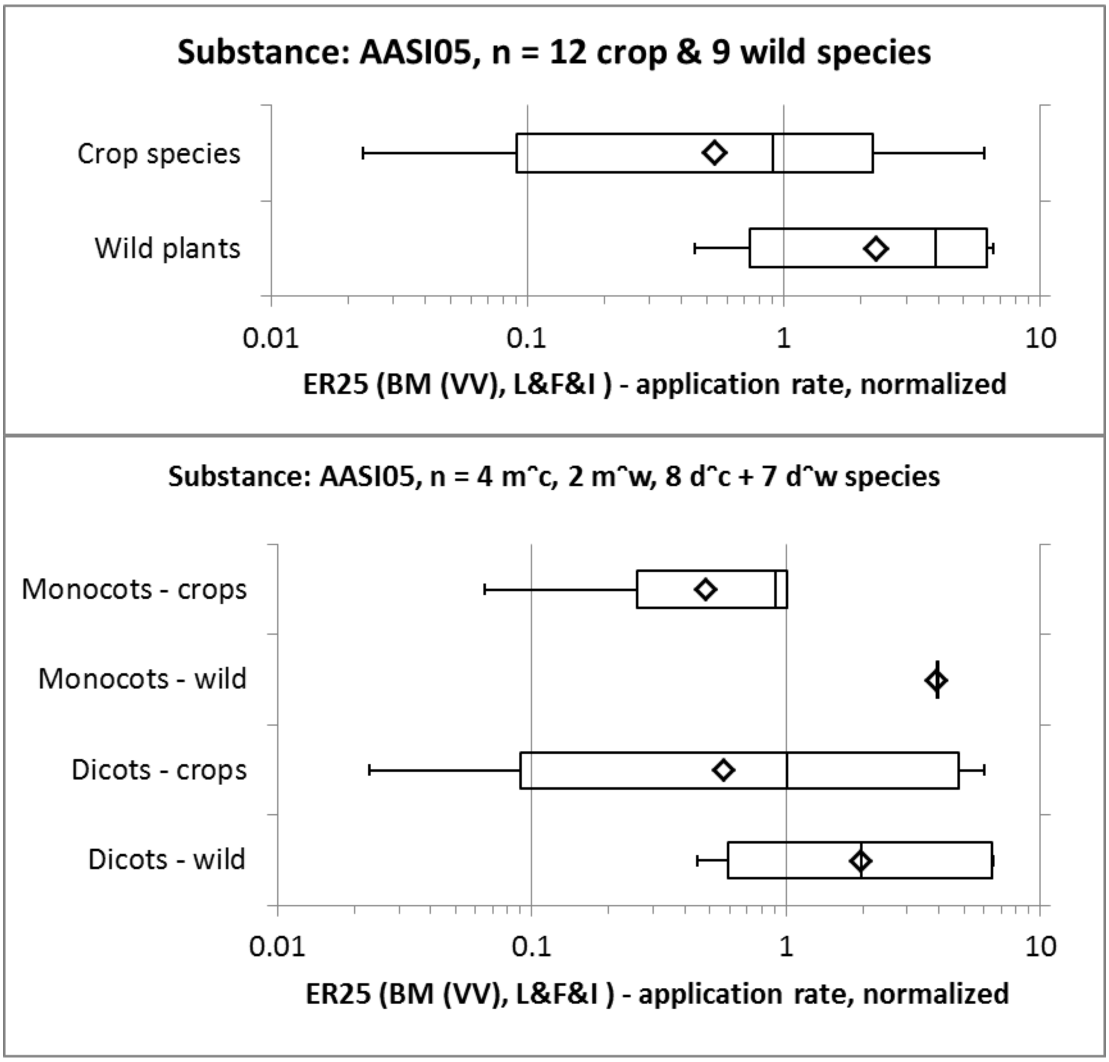


Figure 22: Substance AASI 5 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 22: AASI 5: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI05**	Mono-cots	Crop	4 (species)	0.0654	**0.017**	0.482	**0.12**
Endpoint	ER25		Wild	2 (species)	3.9064		3.913	
Measured	BM	Dicots	Crop	8 (species)	0.0229	**0.051**	0.565	0.29
Lab/field	L&F&I		Wild	7 (species)	0.4483		1.972	
SE, VV or all	VV	All	Crop	12 (species)	0.0229	**0.051**	0.536	0.23
Censored val.	numeric		Wild	9 (species)	0.4483		2.296	

(ER50: No data for crop species)

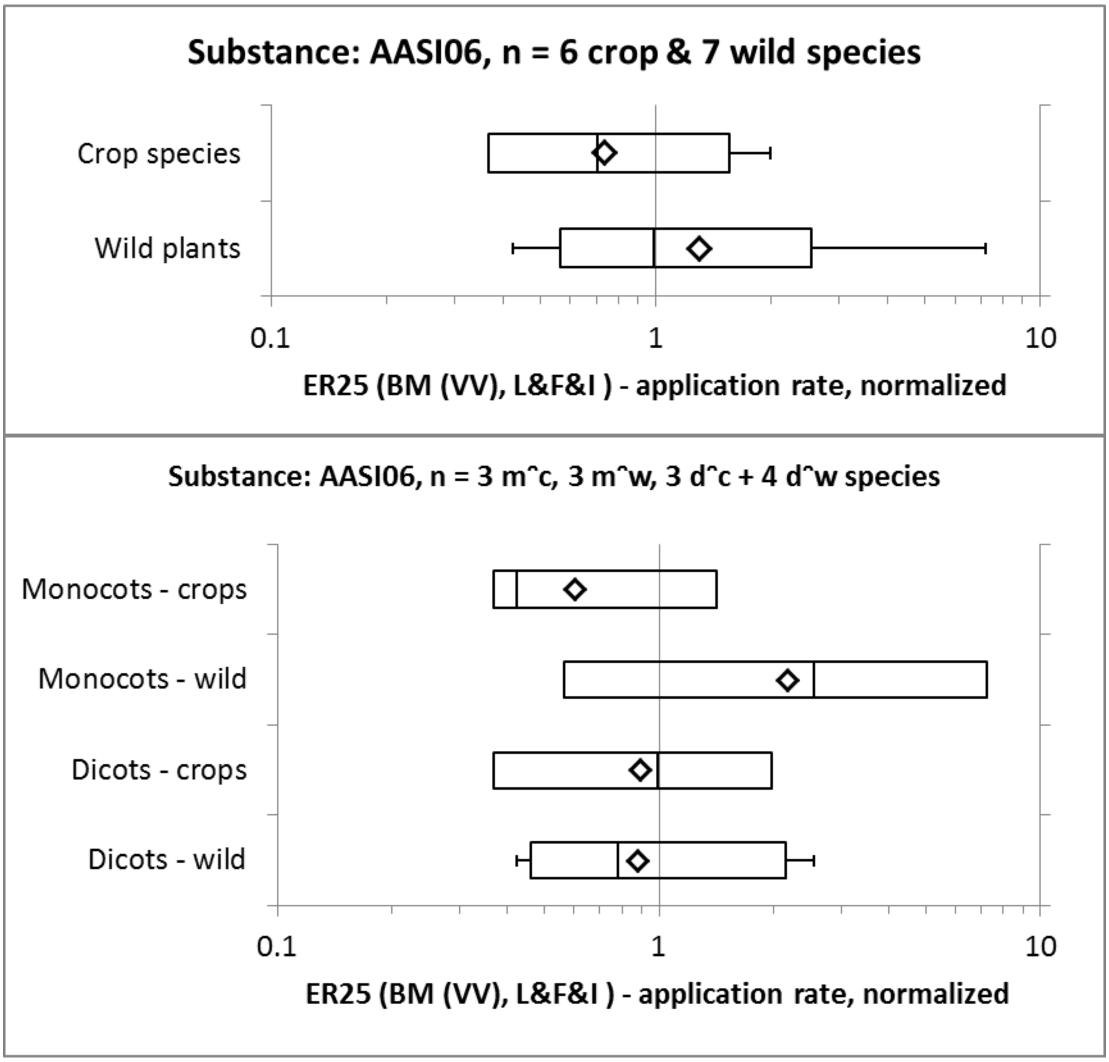


Figure 23: Substance AASI 6 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 23: AASI 6: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI06**	Mono-cots	Crop	3 (species)	0.3677	0.65	0.604	0.28
Endpoint	ER25		Wild	3 (species)	0.5657		2.182	
Measured	BM	Dicots	Crop	3 (species)	0.3677	0.87	0.897	1
Lab/field	L&F&I		Wild	4 (species)	0.4243		0.882	
SE, VV or all	VV	All	Crop	6 (species)	0.3677	0.87	0.736	0.57
Censored val.	numeric		Wild	7 (species)	0.4243		1.300	

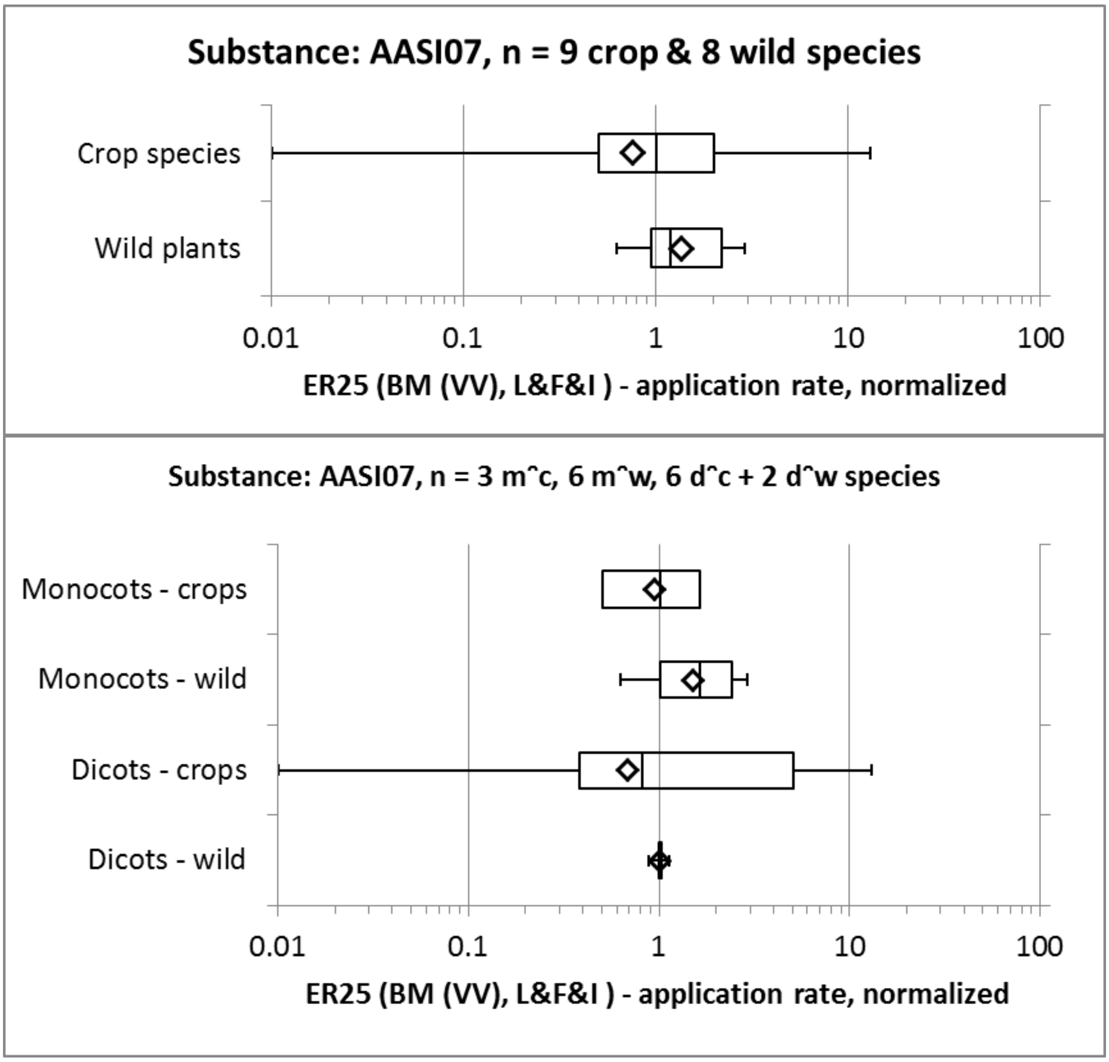


Figure 24: Substance AASI 7 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 24: AASI 7: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI07**	Mono-cots	Crop	3 (species)	0.5039	0.8	0.940	0.62
Endpoint	ER25		Wild	6 (species)	0.6298		1.511	
Measured	BM	Dicots	Crop	6 (species)	0.0101	**0.011**	0.682	0.68
Lab/field	L&F&I		Wild	2 (species)	0.8818		1.000	
SE, VV or all	VV	All	Crop	9 (species)	0.0101	**0.016**	0.759	0.56
Censored val.	numeric		Wild	8 (species)	0.6298		1.363	

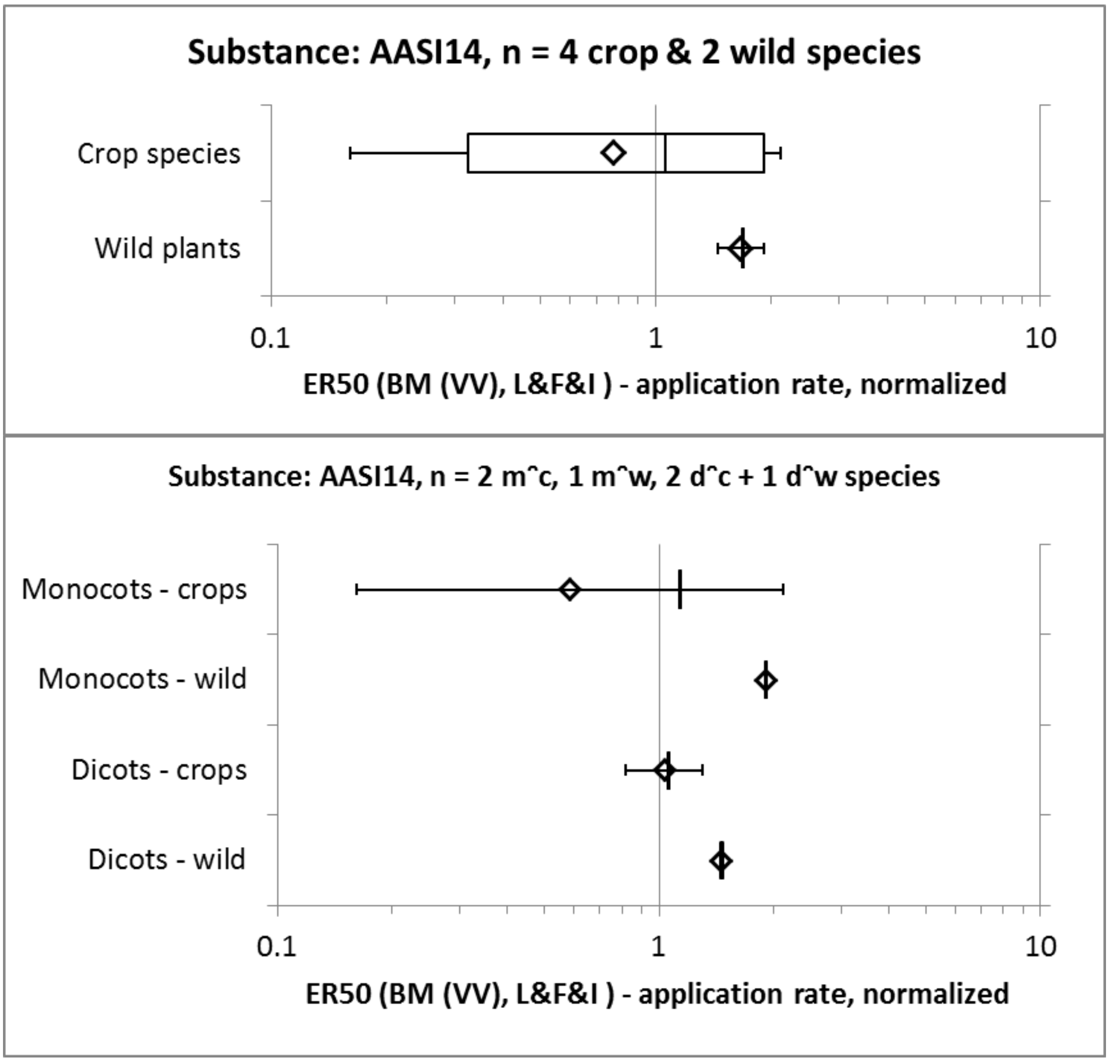


Figure 25: Substance AASI14 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 25: AASI14: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI14**	Mono-cots	Crop	2 (species)	0.1607	**0.084**	0.582	0.31
Endpoint	ER50		Wild	1 (species)	1.9060		1.906	
Measured	BM	Dicots	Crop	2 (species)	0.8197	0.56	1.031	0.71
Lab/field	L&F&I		Wild	1 (species)	1.4550		1.455	
SE, VV or all	VV	All	Crop	4 (species)	0.1607	**0.11**	0.775	0.47
Censored val.	numeric		Wild	2 (species)	1.4550		1.665	

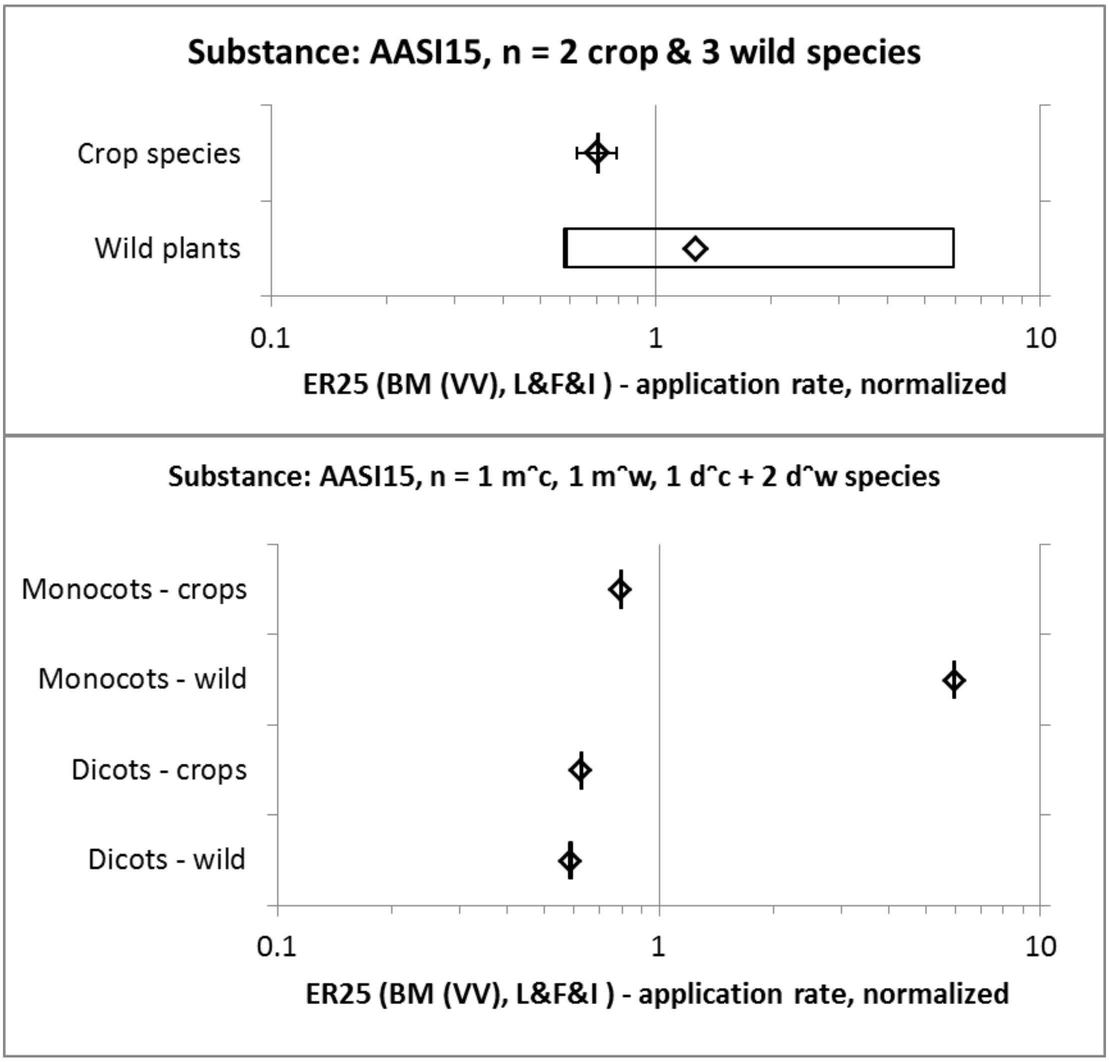


Figure 26: Substance AASI15 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 26: AASI15: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI15**	Mono-cots	Crop	1 (species)	0.7924	**0.13**	0.792	**0.13**
Endpoint	ER25		Wild	1 (species)	5.9427		5.943	
Measured	BM	Dicots	Crop	1 (species)	0.6226	*1.1*	0.623	*1.1*
Lab/field	L&F&I		Wild	2 (species)	0.5800		0.584	
SE, VV or all	VV	All	Crop	2 (species)	0.6226	*1.1*	0.702	0.55
Censored val.	numeric		Wild	3 (species)	0.5800		1.266	

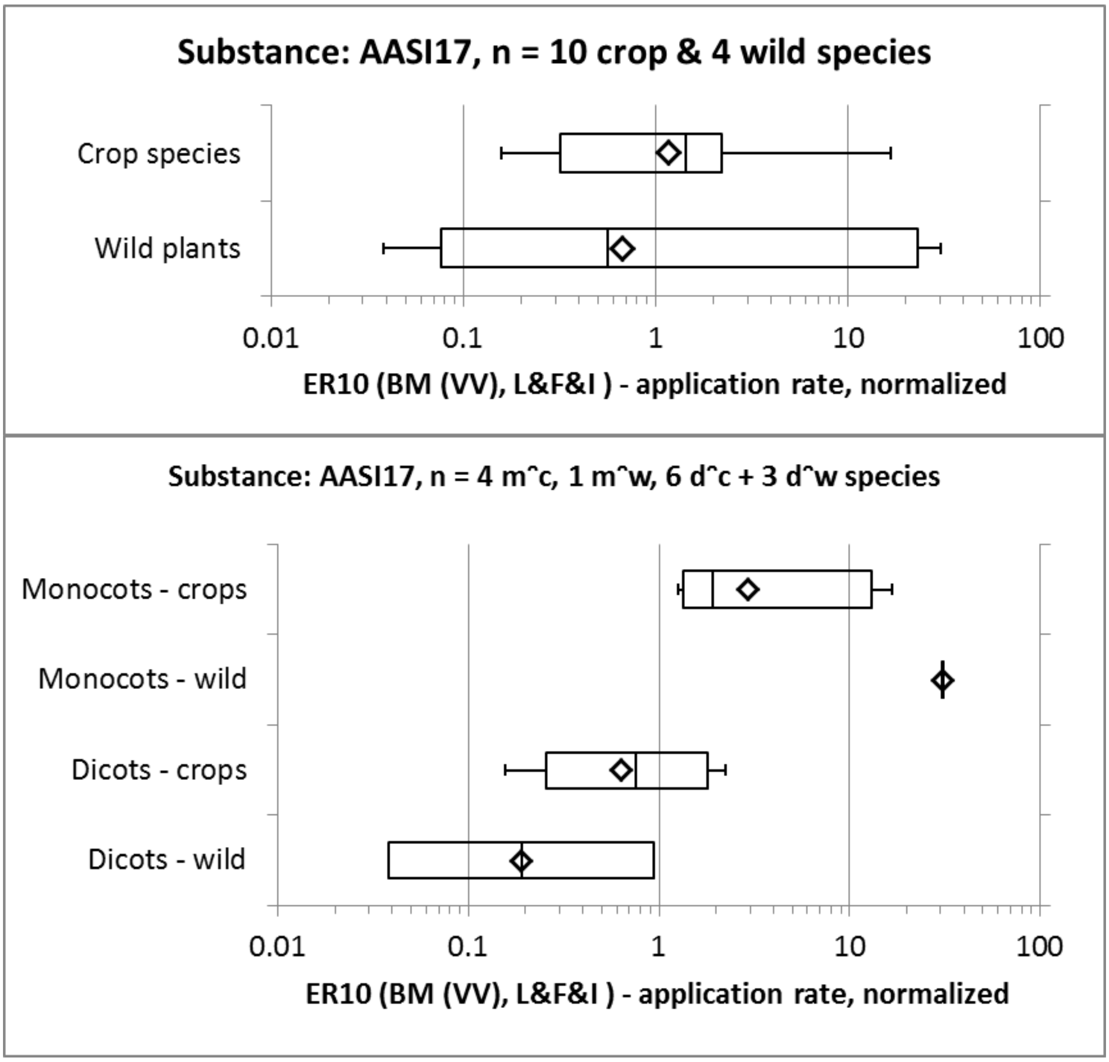


Figure 27: Substance AASI17 - Distribution of numeric ER10 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 27: AASI17: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER10 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI17**	Mono-cots	Crop	4 (species)	1.2524	**0.041**	2.937	**0.096**
Endpoint	ER10		Wild	1 (species)	30.5190		30.519	
Measured	BM	Dicots	Crop	6 (species)	0.1565	*4.1*	0.634	*3.3*
Lab/field	L&F&I		Wild	3 (species)	0.0381		0.189	
SE, VV or all	VV	All	Crop	10 (species)	0.1565	*4.1*	1.170	*1.7*
Censored val.	numeric		Wild	4 (species)	0.0381		0.675	

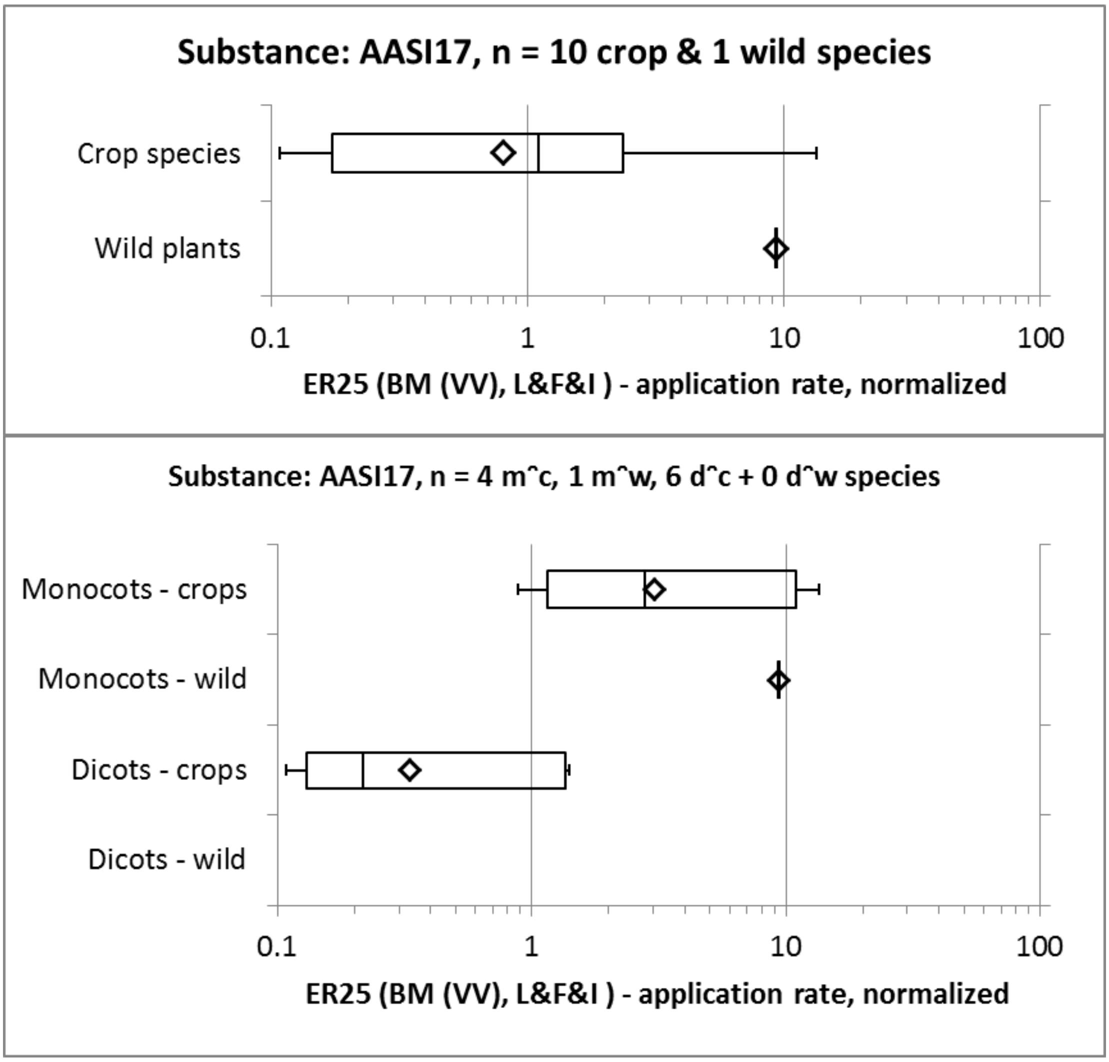


Figure 28: Substance AASI17 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 28: AASI17: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI17**	Mono-cots	Crop	4 (species)	0.8801	**0.094**	3.021	0.32
Endpoint	ER25		Wild	1 (species)	9.3241		9.324	
Measured	BM	Dicots	Crop	6 (species)	0.1084	#NUM!	0.330	#NUM!
Lab/field	L&F&I		Wild	0 (species)	#NUM!		#NUM!	
SE, VV or all	VV	All	Crop	10 (species)	0.1084	**0.012**	0.800	**0.086**
Censored val.	numeric		Wild	1 (species)	9.3241		9.324	

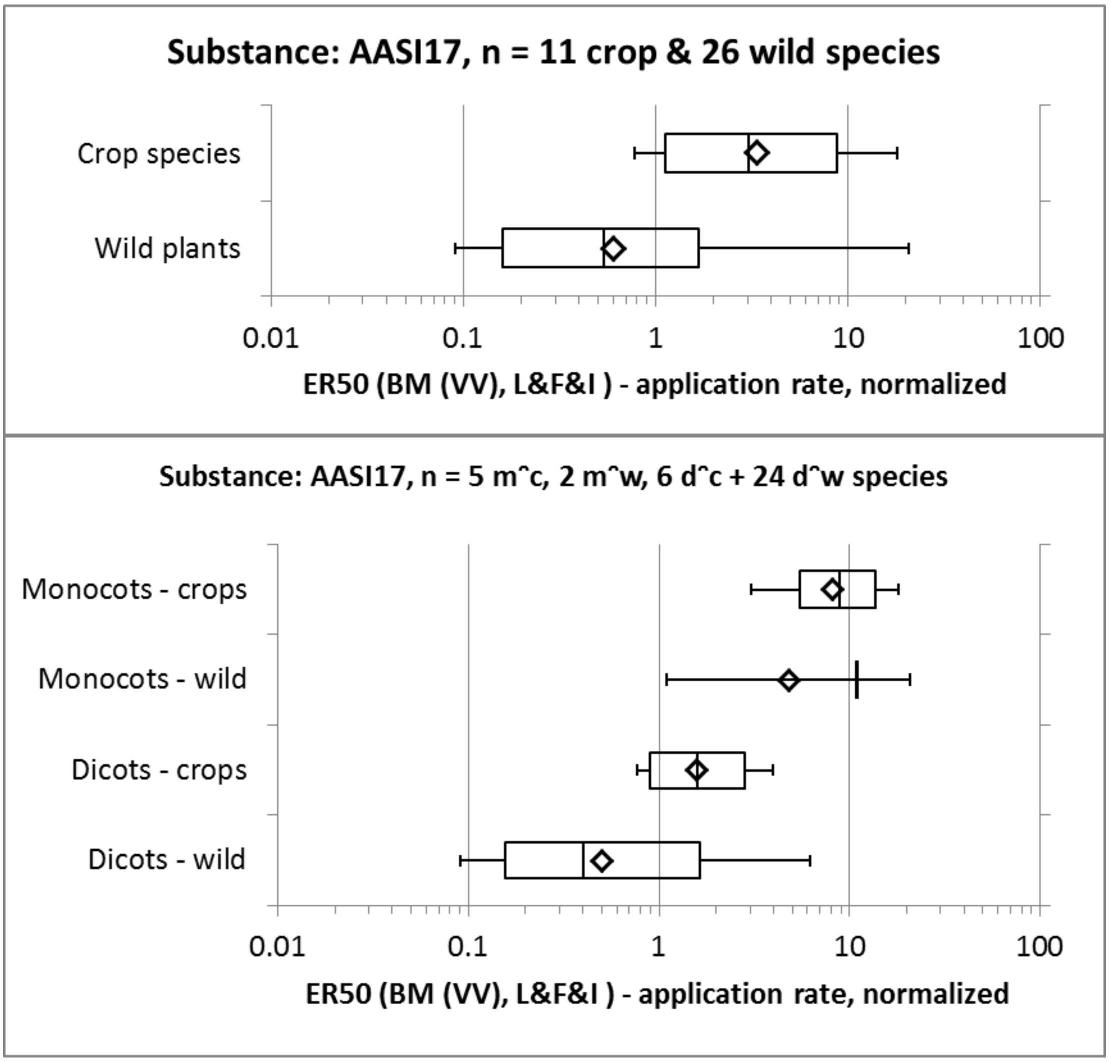


Figure 29: Substance AASI17 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 29: AASI17: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**AASI17**	Mono-cots	Crop	5 (species)	3.0451	*2.8*	8.141	*1.7*
Endpoint	ER50		Wild	2 (species)	1.1051		4.772	
Measured	BM	Dicots	Crop	6 (species)	0.7742	**8.6**	1.589	*3.1*
Lab/field	L&F&I		Wild	24 (species)	0.0903		0.505	
SE, VV or all	VV	All	Crop	11 (species)	0.7742	**8.6**	3.340	**5.6**
Censored val.	numeric		Wild	26 (species)	0.0903		0.600	

ACI1: only one wild & one crop with numeric values quotient close to 1

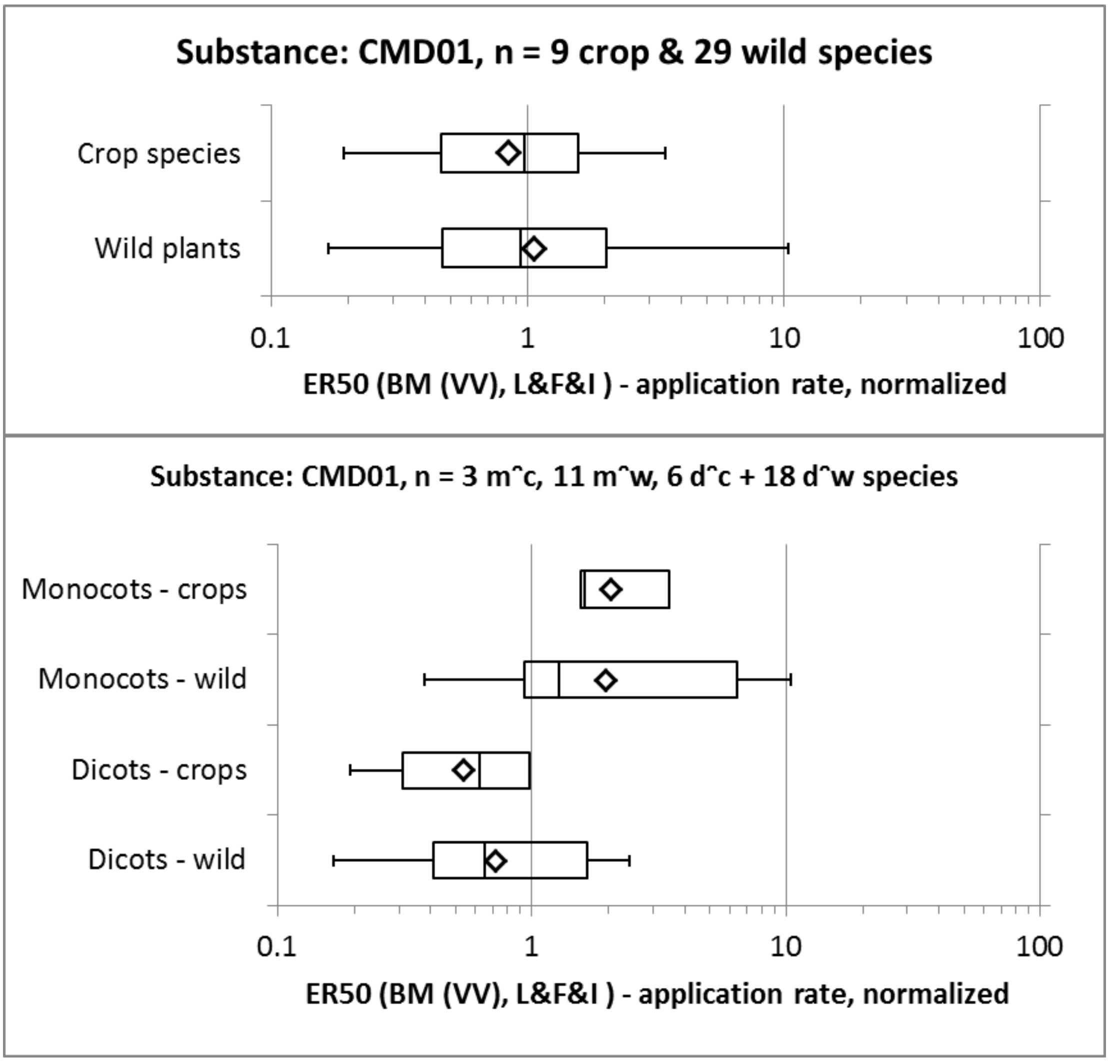


Figure 30: Substance CMD 1 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 30: CMD 1: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**CMD01**	Mono-cots	Crop	3 (species)	1.5617	*4.1*	2.066	1
Endpoint	ER50		Wild	11 (species)	0.3808		1.968	
Measured	BM	Dicots	Crop	6 (species)	0.1931	*1.2*	0.544	0.76
Lab/field	L&F&I		Wild	18 (species)	0.1675		0.718	
SE, VV or all	VV	All	Crop	9 (species)	0.1931	*1.2*	0.849	0.81
Censored val.	numeric		Wild	29 (species)	0.1675		1.052	

Further endpoints; no crop-ER25 available

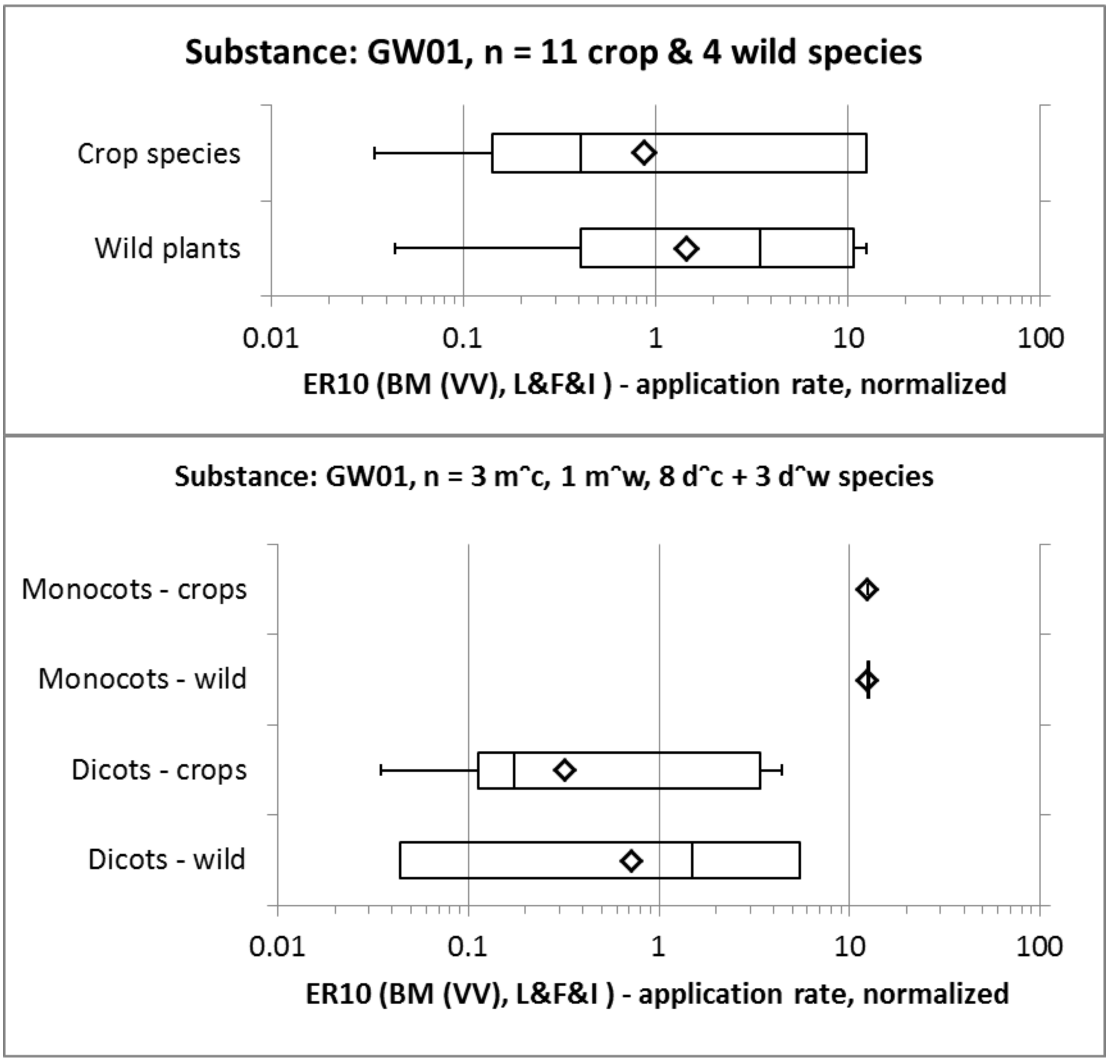


Figure 31: Substance GW 01 - Distribution of numeric ER10 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 31: GW 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER10 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW01**	Mono-cots	Crop	3 (species)	12.3880	1	12.388	1
Endpoint	ER10		Wild	1 (species)	12.3880		12.388	
Measured	BM	Dicots	Crop	8 (species)	0.0346	0.79	0.323	0.45
Lab/field	L&F&I		Wild	3 (species)	0.0440		0.713	
SE, VV or all	VV	All	Crop	11 (species)	0.0346	0.79	0.872	0.6
Censored val.	numeric		Wild	4 (species)	0.0440		1.456	

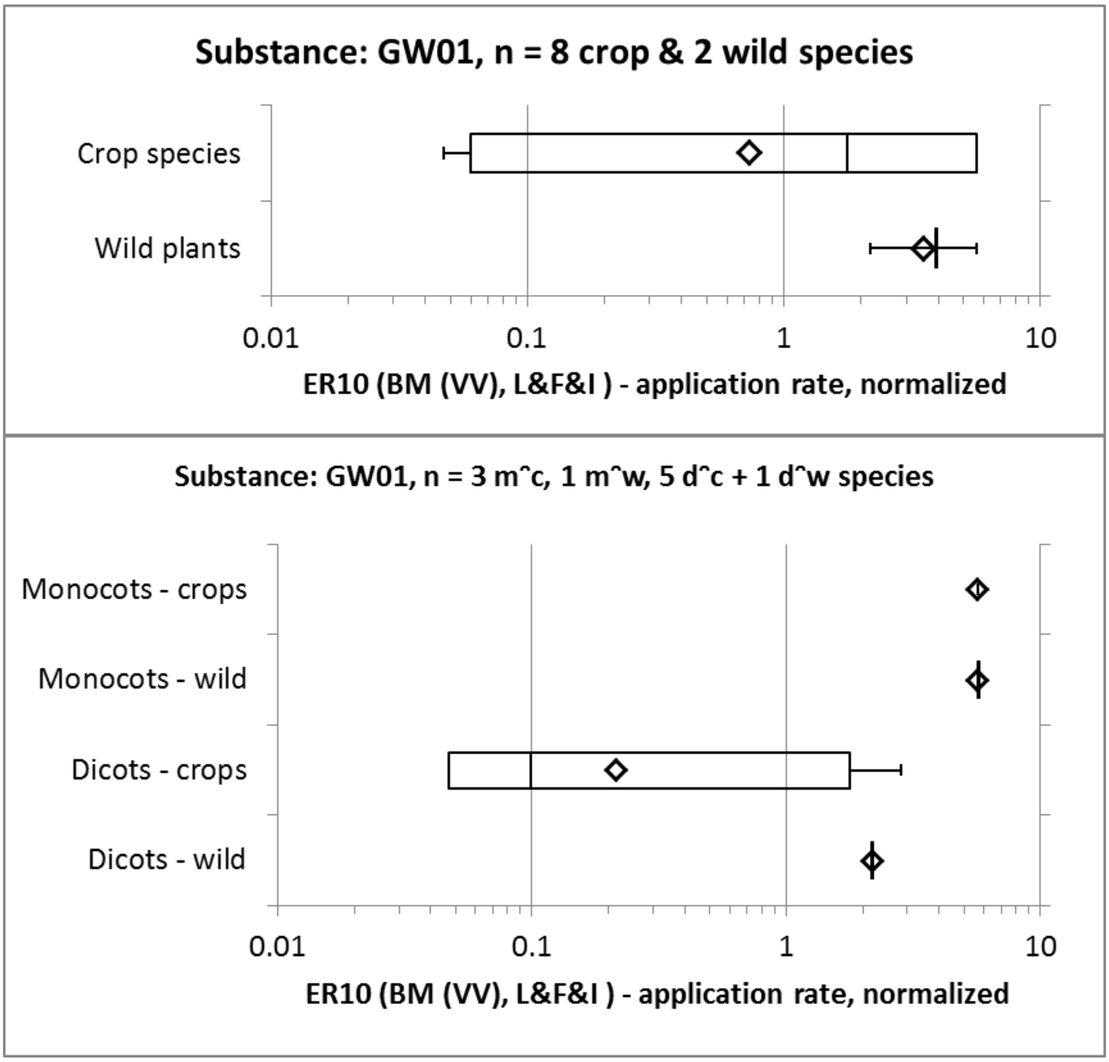


Figure 32: Substance GW 01 - Distribution of numeric ER10 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 32: GW 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER10 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW01**	Mono-cots	Crop	3 (species)	5.6789	1	5.679	1
Endpoint	ER10		Wild	1 (species)	5.6789		5.679	
Measured	BM	Dicots	Crop	5 (species)	0.0471	**0.022**	0.213	**0.098**
Lab/field	L&F&I		Wild	1 (species)	2.1795		2.180	
SE, VV or all	VV	All	Crop	8 (species)	0.0471	**0.022**	0.730	0.21
Censored val.	numeric		Wild	2 (species)	2.1795		3.518	

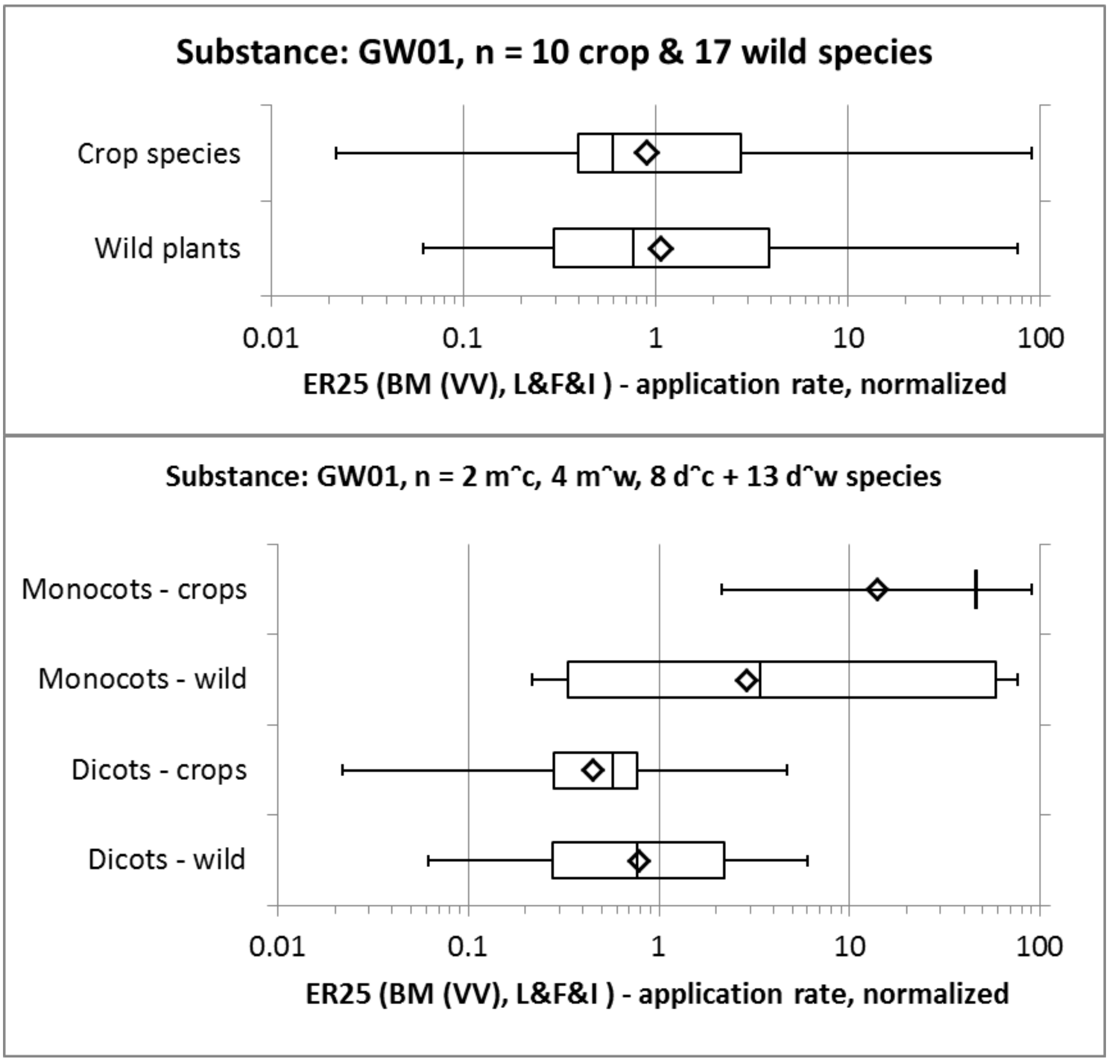


Figure 33: Substance GW 01 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 33: GW 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW01**	Mono-cots	Crop	2 (species)	2.1410	**9.9**	13.865	*4.8*
Endpoint	ER25		Wild	4 (species)	0.2157		2.865	
Measured	BM	Dicots	Crop	8 (species)	0.0218	0.35	0.450	0.57
Lab/field	L&F&I		Wild	13 (species)	0.0616		0.789	
SE, VV or all	VV	All	Crop	10 (species)	0.0218	0.35	0.894	0.84
Censored val.	numeric		Wild	17 (species)	0.0616		1.068	

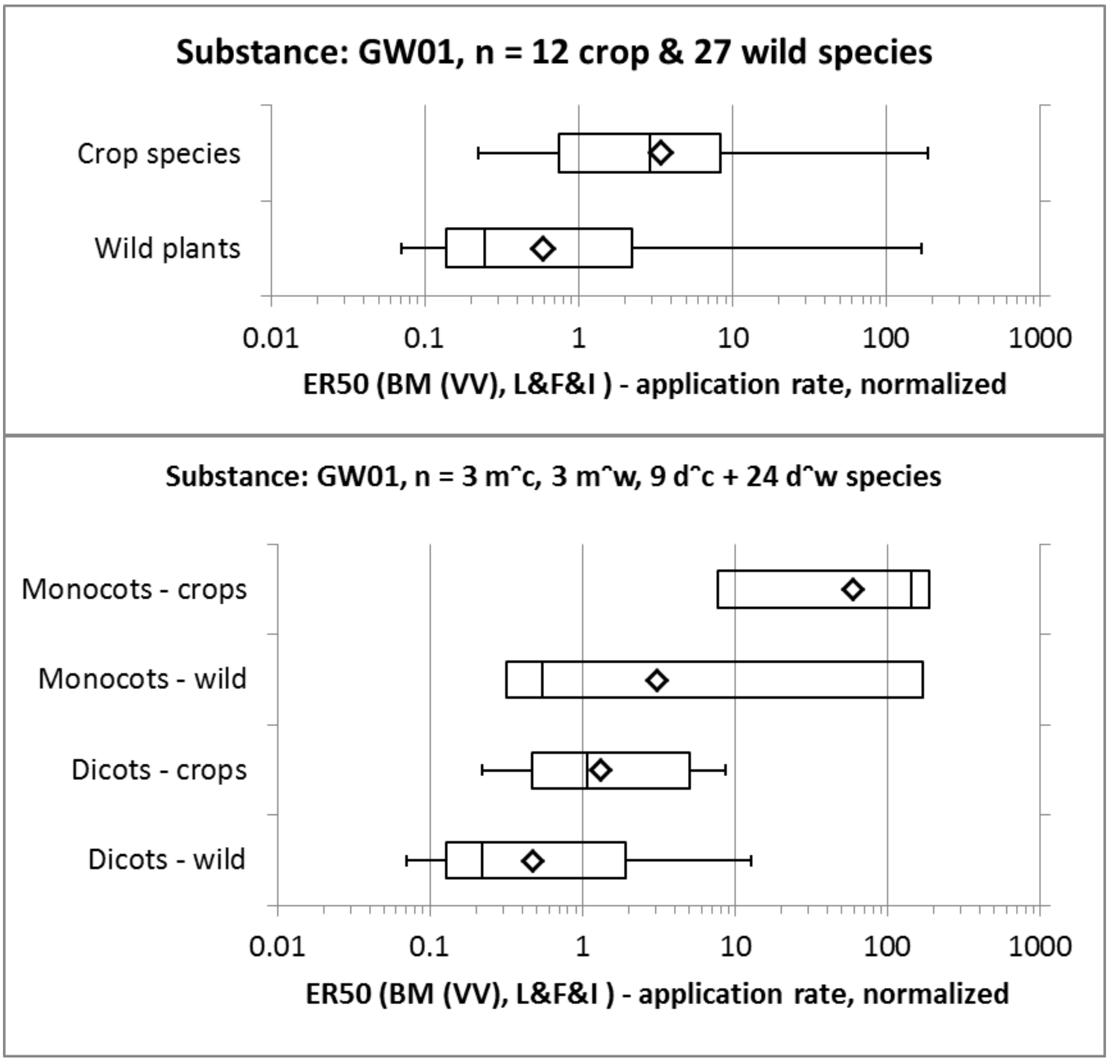


Figure 34: Substance GW 01 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 34: GW 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW01**	Mono-cots	Crop	3 (species)	7.6268	**24**	58.675	**19**
Endpoint	ER50		Wild	3 (species)	0.3190		3.092	
Measured	BM	Dicots	Crop	9 (species)	0.2211	*3.1*	1.311	*2.8*
Lab/field	L&F&I		Wild	24 (species)	0.0704		0.472	
SE, VV or all	VV	All	Crop	12 (species)	0.2211	*3.1*	3.390	**5.8**
Censored val.	numeric		Wild	27 (species)	0.0704		0.581	

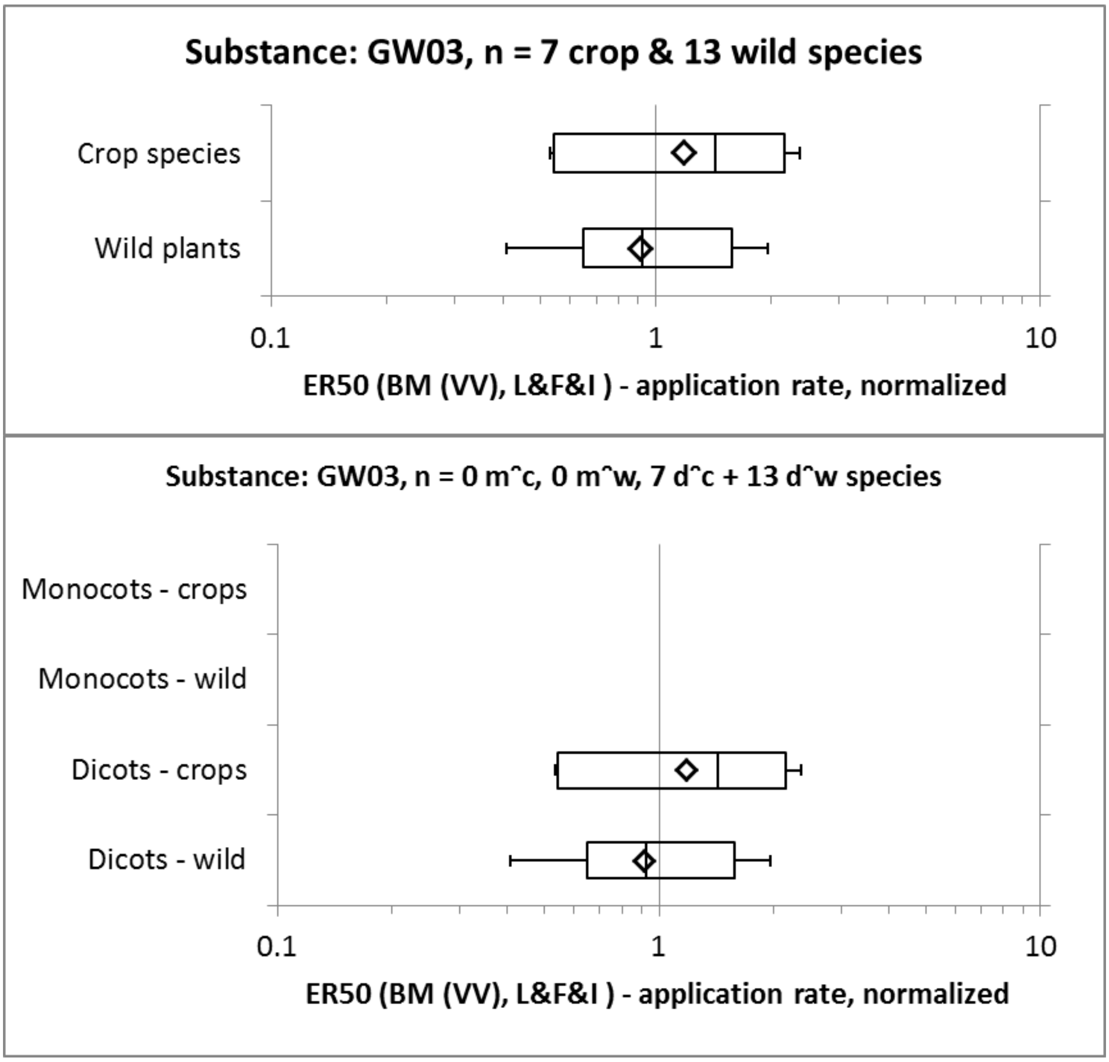


Figure 35: Substance GW03 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 35: GW03: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW03**	Mono-cots	Crop	0 (species)	#NUM!	#NUM!	#NUM!	#NUM!
Endpoint	ER50		Wild	0 (species)	#NUM!		#NUM!	
Measured	BM	Dicots	Crop	7 (species)	0.5325	*1.3*	1.185	*1.3*
Lab/field	L&F&I		Wild	13 (species)	0.4075		0.913	
SE, VV or all	VV	All	Crop	7 (species)	0.5325	*1.3*	1.185	*1.3*
Censored val.	numeric		Wild	13 (species)	0.4075		0.913	

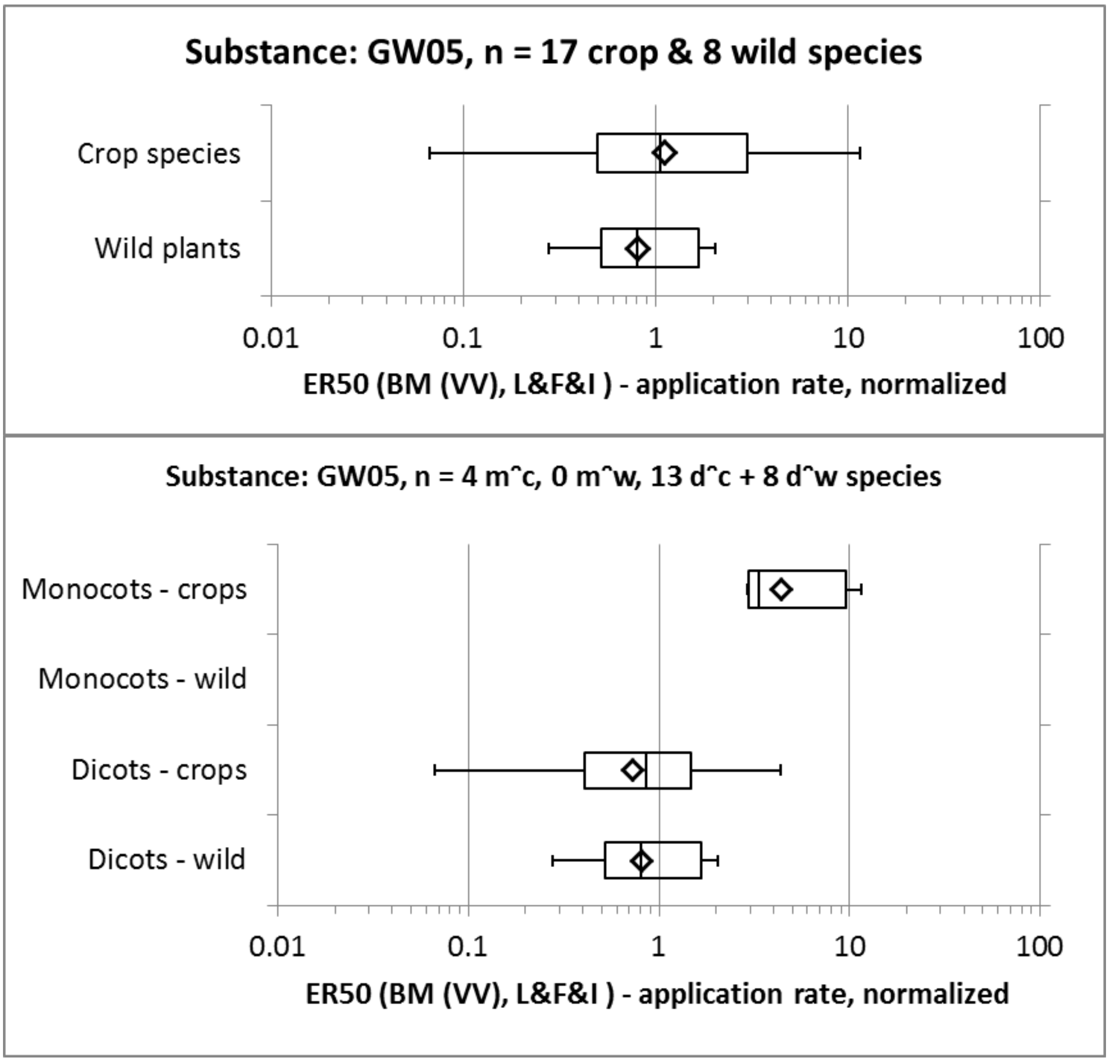


Figure 36: Substance GW 05 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 36: GW 05: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW05**	Mono-cots	Crop	4 (species)	2.9039	#NUM!	4.395	#NUM!
Endpoint	ER50		Wild	0 (species)	#NUM!		#NUM!	
Measured	BM	Dicots	Crop	13 (species)	0.0665	0.24	0.724	0.9
Lab/field	L&F&I		Wild	8 (species)	0.2780		0.805	
SE, VV or all	VV	All	Crop	17 (species)	0.0665	0.24	1.107	*1.4*
Censored val.	numeric		Wild	8 (species)	0.2780		0.805	

.

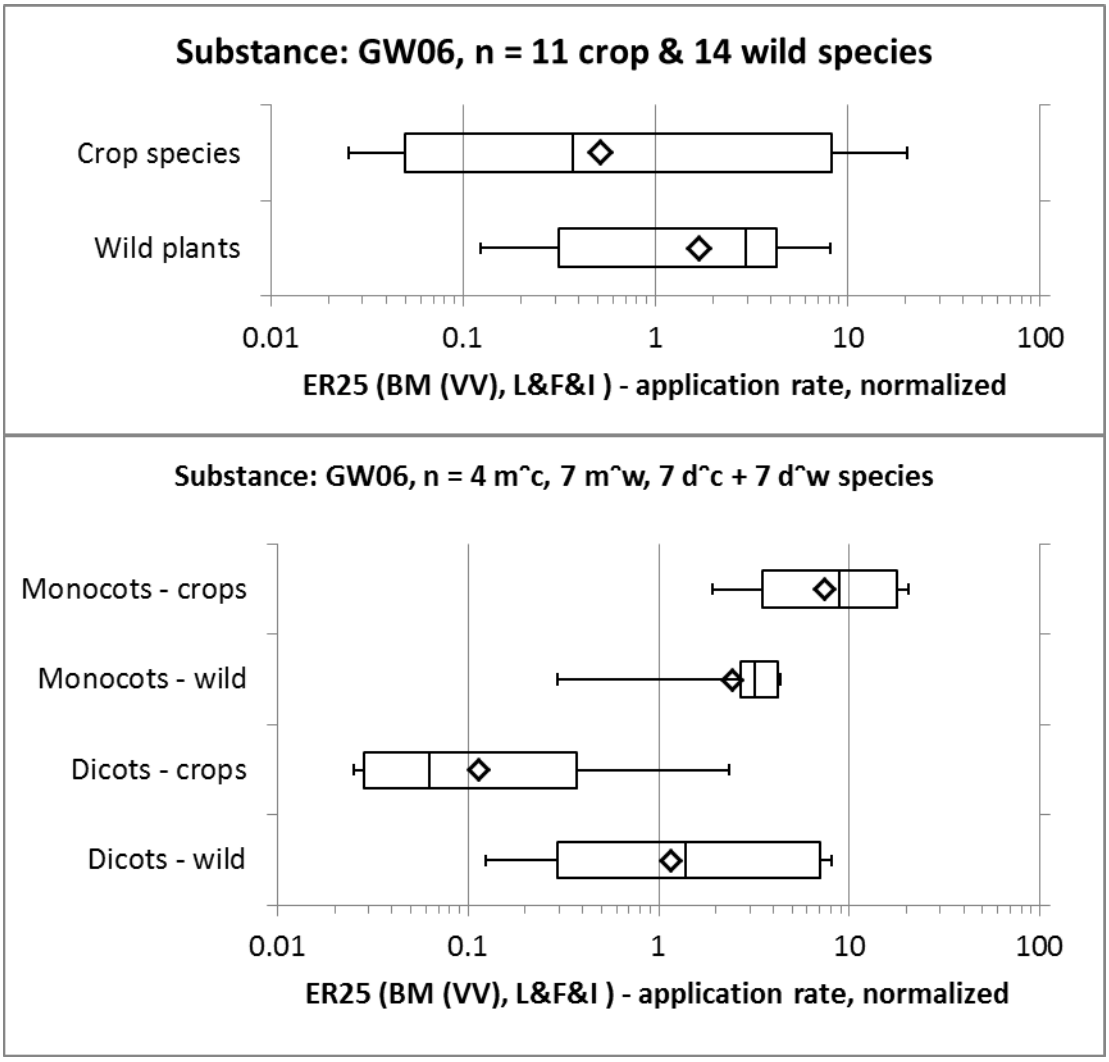


Figure 37: Substance GW 06 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 37: GW 06: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW06**	Mono-cots	Crop	4 (species)	1.9014	**6.4**	7.425	*3.1*
Endpoint	ER25		Wild	7 (species)	0.2951		2.424	
Measured	BM	Dicots	Crop	7 (species)	0.0252	0.2	0.113	**0.097**
Lab/field	L&F&I		Wild	7 (species)	0.1229		1.160	
SE, VV or all	VV	All	Crop	11 (species)	0.0252	0.2	0.518	0.31
Censored val.	numeric		Wild	14 (species)	0.1229		1.677	

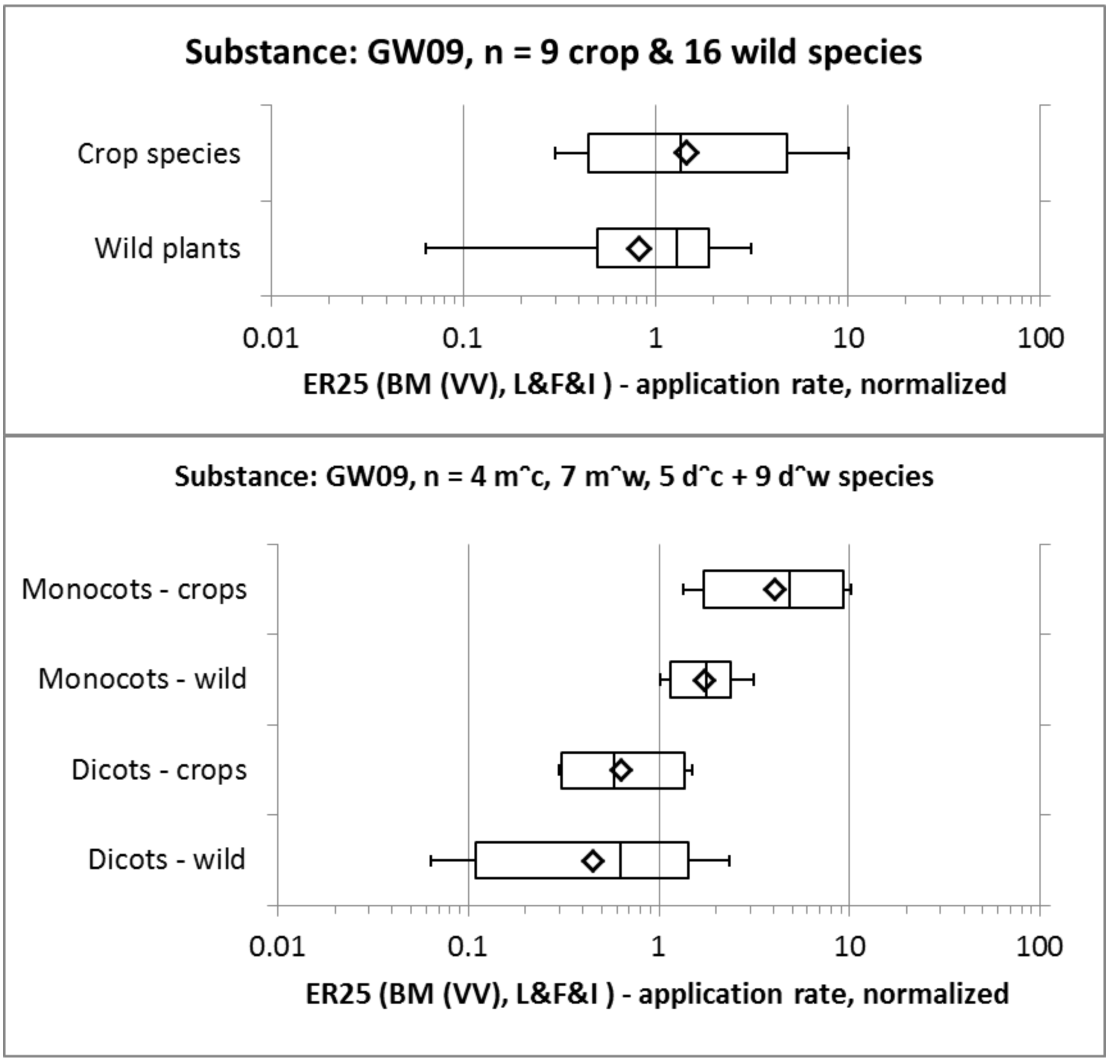


Figure 38: Substance GW 09 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 38: GW 09: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW09**	Mono-cots	Crop	4 (species)	1.3491	*1.3*	4.022	*2.3*
Endpoint	ER25		Wild	7 (species)	1.0208		1.737	
Measured	BM	Dicots	Crop	5 (species)	0.2992	*4.7*	0.633	*1.4*
Lab/field	L&F&I		Wild	9 (species)	0.0633		0.452	
SE, VV or all	VV	All	Crop	9 (species)	0.2992	*4.7*	1.440	*1.8*
Censored val.	numeric		Wild	16 (species)	0.0633		0.814	

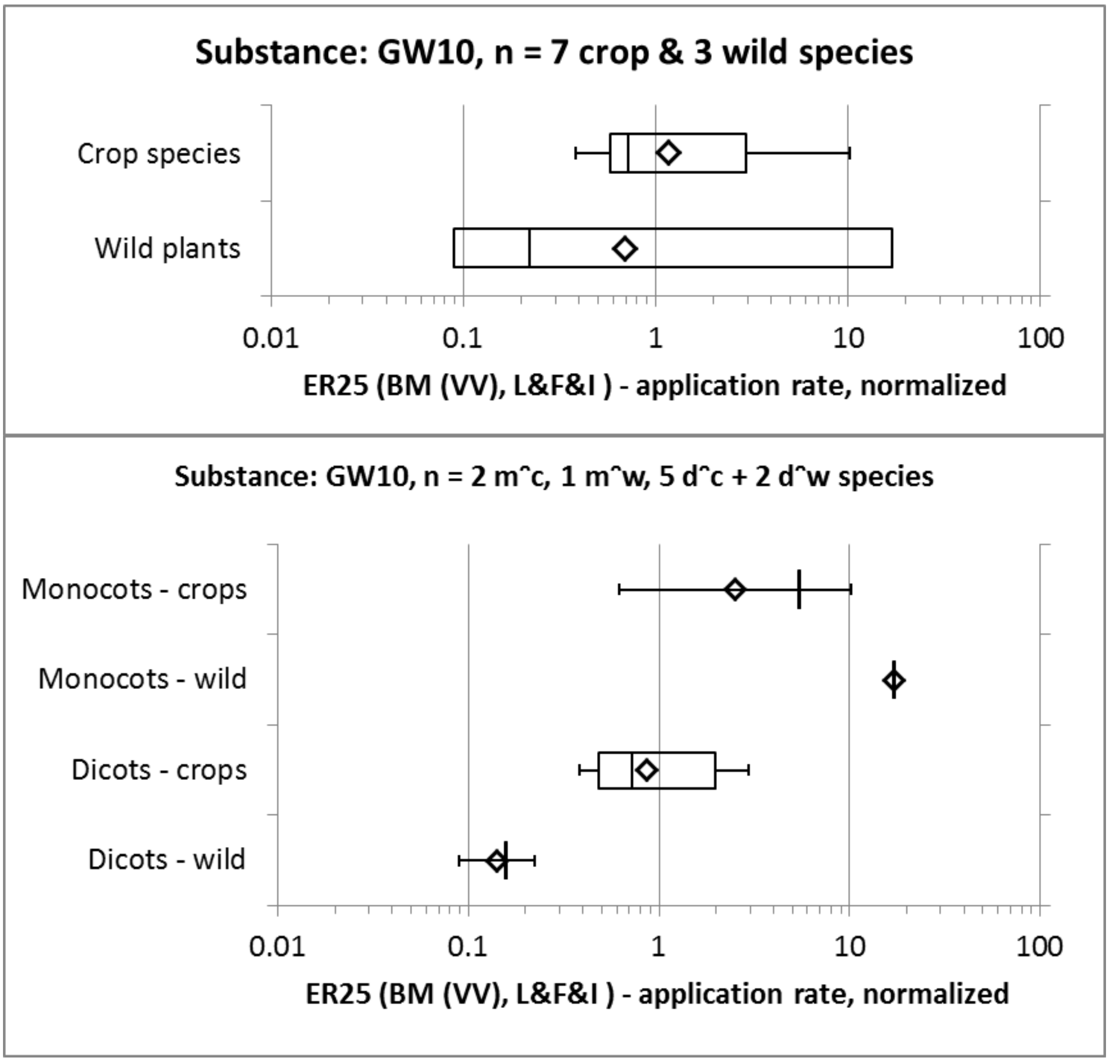


Figure 39: Substance GW 10 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 39: GW 10: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW10**	Mono-cots	Crop	2 (species)	0.6223	**0.036**	2.522	**0.15**
Endpoint	ER25		Wild	1 (species)	17.0688		17.069	
Measured	BM	Dicots	Crop	5 (species)	0.3823	*4.3*	0.858	**6.1**
Lab/field	L&F&I		Wild	2 (species)	0.0889		0.141	
SE, VV or all	VV	All	Crop	7 (species)	0.3823	*4.3*	1.168	*1.7*
Censored val.	numeric		Wild	3 (species)	0.0889		0.696	

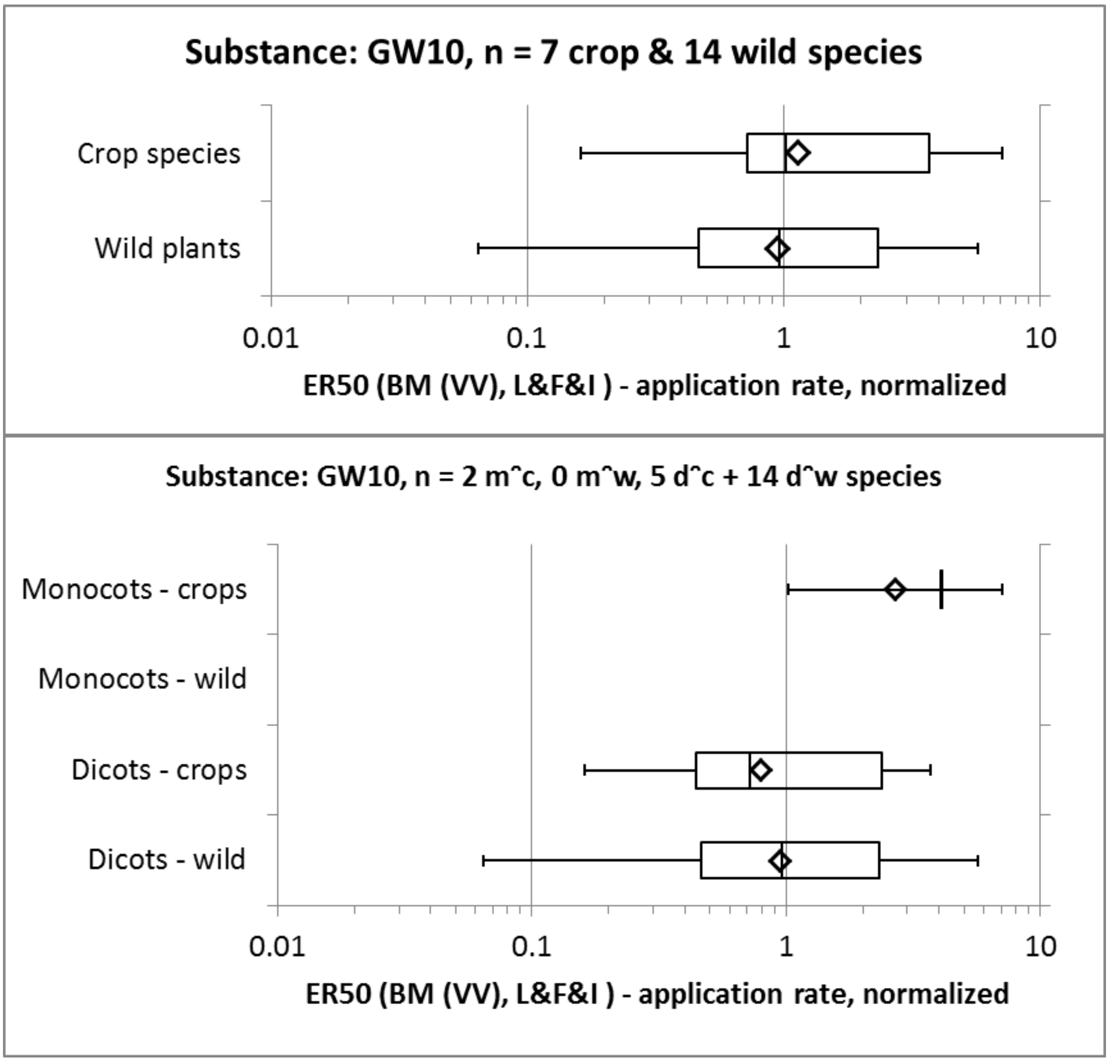


Figure 40: Substance GW 10 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 40: GW 10: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW10**	Mono-cots	Crop	2 (species)	1.0172	#NUM!	2.690	#NUM!
Endpoint	ER50		Wild	0 (species)	#NUM!		#NUM!	
Measured	BM	Dicots	Crop	5 (species)	0.1606	*2.5*	0.798	0.85
Lab/field	L&F&I		Wild	14 (species)	0.0646		0.941	
SE, VV or all	VV	All	Crop	7 (species)	0.1606	*2.5*	1.129	*1.2*
Censored val.	numeric		Wild	14 (species)	0.0646		0.941	

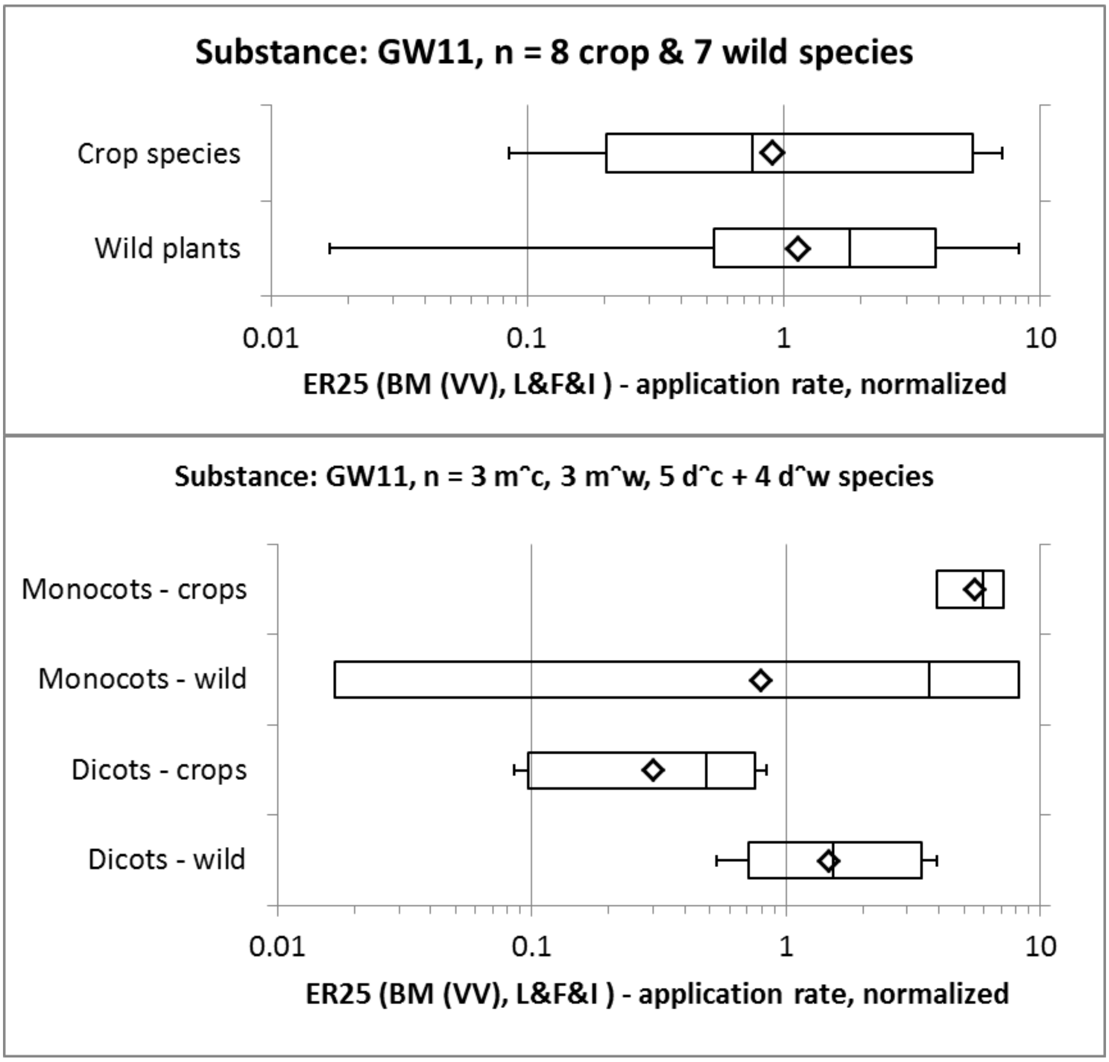


Figure 41: Substance GW 11 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 41: GW 11: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW11**	Mono-cots	Crop	3 (species)	3.9366	**234**	5.503	**6.9**
Endpoint	ER25		Wild	3 (species)	0.0168		0.798	
Measured	BM	Dicots	Crop	5 (species)	0.0848	**0.16**	0.302	0.2
Lab/field	L&F&I		Wild	4 (species)	0.5330		1.476	
SE, VV or all	VV	All	Crop	8 (species)	0.0848	*5*	0.896	0.79
Censored val.	numeric		Wild	7 (species)	0.0168		1.134	

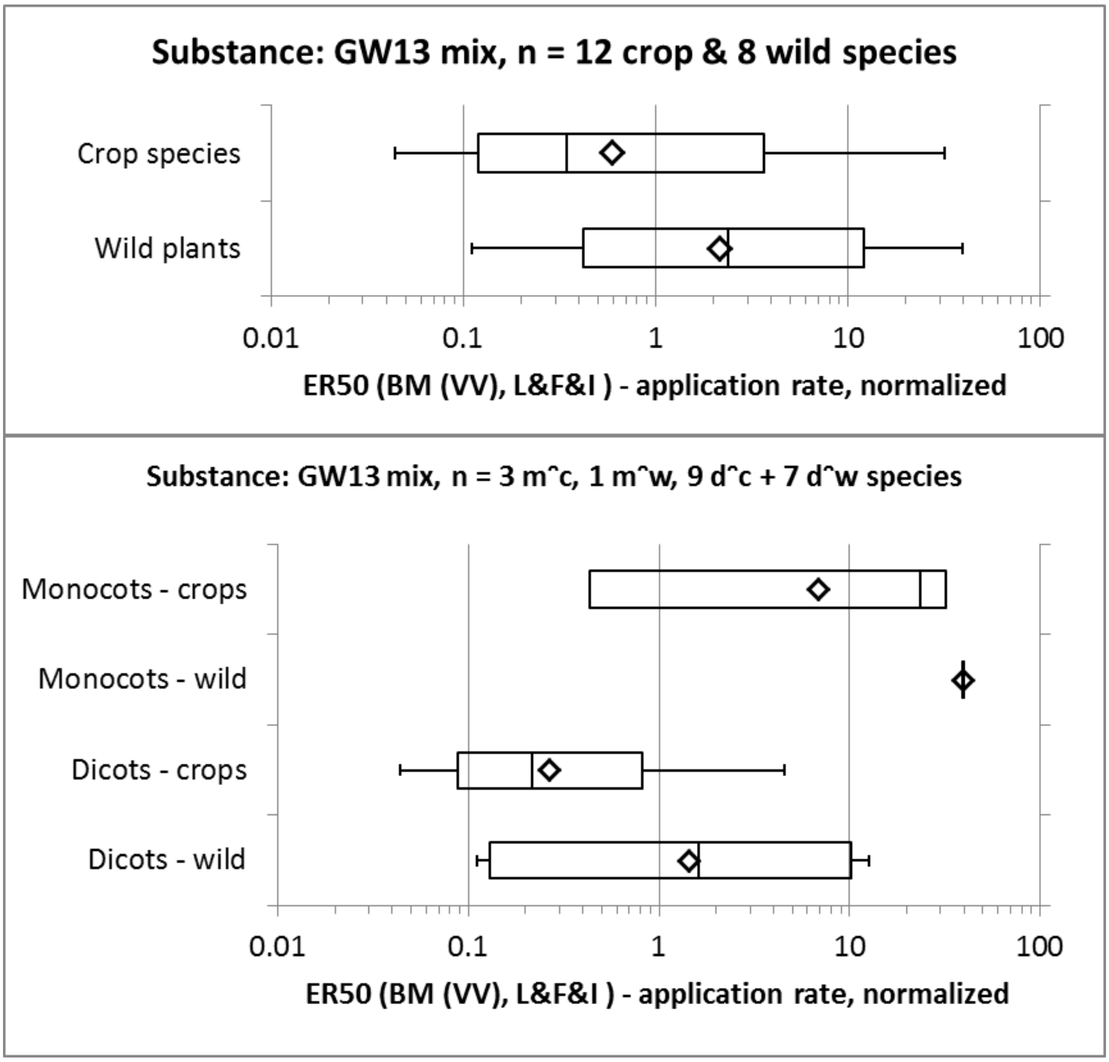


Figure 42: Substance GW13 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 42: GW 13: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**GW13 mix**	Mono-cots	Crop	3 (species)	0.4363	**0.011**	6.900	**0.175**
Endpoint	ER50		Wild	1 (species)	39.3379		39.338	
Measured	BM	Dicots	Crop	9 (species)	0.0441	0.4	0.264	**0.185**
Lab/field	L&F&I		Wild	7 (species)	0.1109		1.431	
SE, VV or all	VV	All	Crop	12 (species)	0.0441	0.4	0.598	0.28
Censored val.	numeric		Wild	8 (species)	0.1109		2.165	

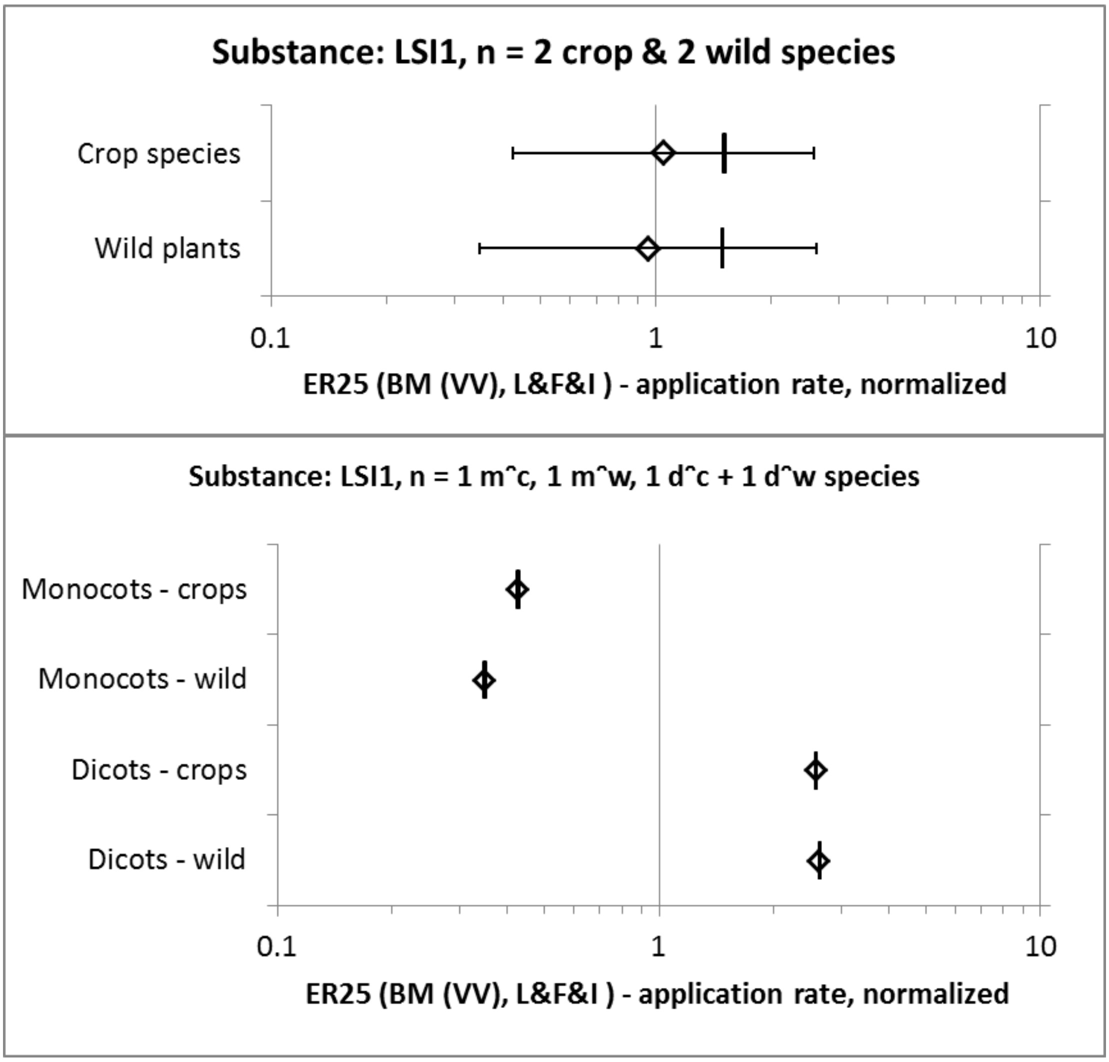


Figure 43: Substance LSI 1 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 43: LSI 1: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**LSI1**	Mono-cots	Crop	1 (species)	0.4254	*1.2*	0.425	*1.2*
Endpoint	ER25		Wild	1 (species)	0.3479		0.348	
Measured	BM	Dicots	Crop	1 (species)	2.5777	0.98	2.578	0.98
Lab/field	L&F&I		Wild	1 (species)	2.6213		2.621	
SE, VV or all	VV	All	Crop	2 (species)	0.4254	*1.2*	1.047	*1.1*
Censored val.	numeric		Wild	2 (species)	0.3479		0.955	

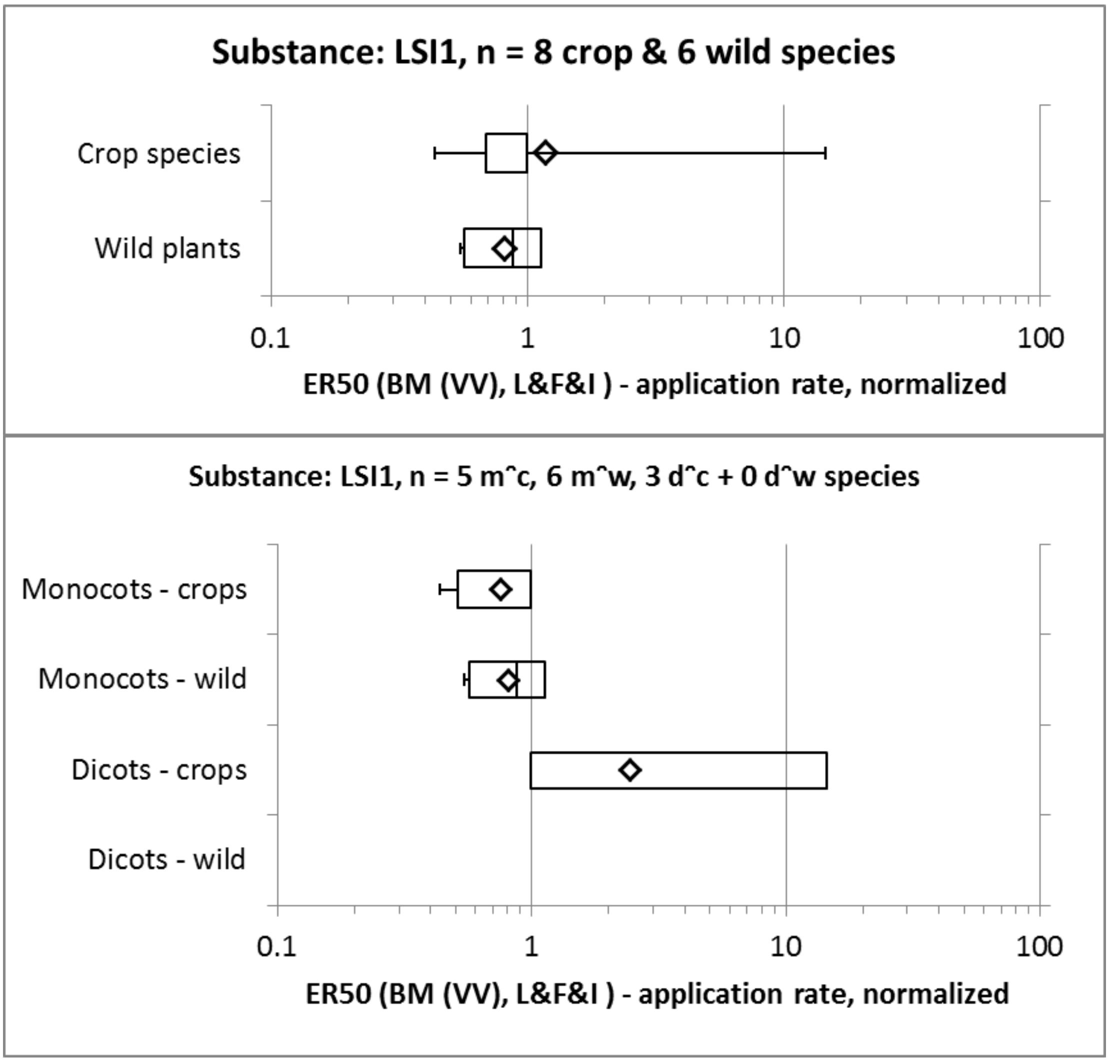


Figure 44: Substance LSI 1 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 44: LSI 1: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**LSI1**	Mono-cots	Crop	5 (species)	0.4359	0.8	0.758	0.94
Endpoint	ER50		Wild	6 (species)	0.5449		0.809	
Measured	BM	Dicots	Crop	3 (species)	0.9907	#NUM!	2.422	#NUM!
Lab/field	L&F&I		Wild	0 (species)	#NUM!		#NUM!	
SE, VV or all	VV	All	Crop	8 (species)	0.4359	0.8	1.172	*1.4*
Censored val.	numeric		Wild	6 (species)	0.5449		0.809	

LSI2: ER25: Only crop data; ER50 previously with typo (unit g/ha vs kg/ha)

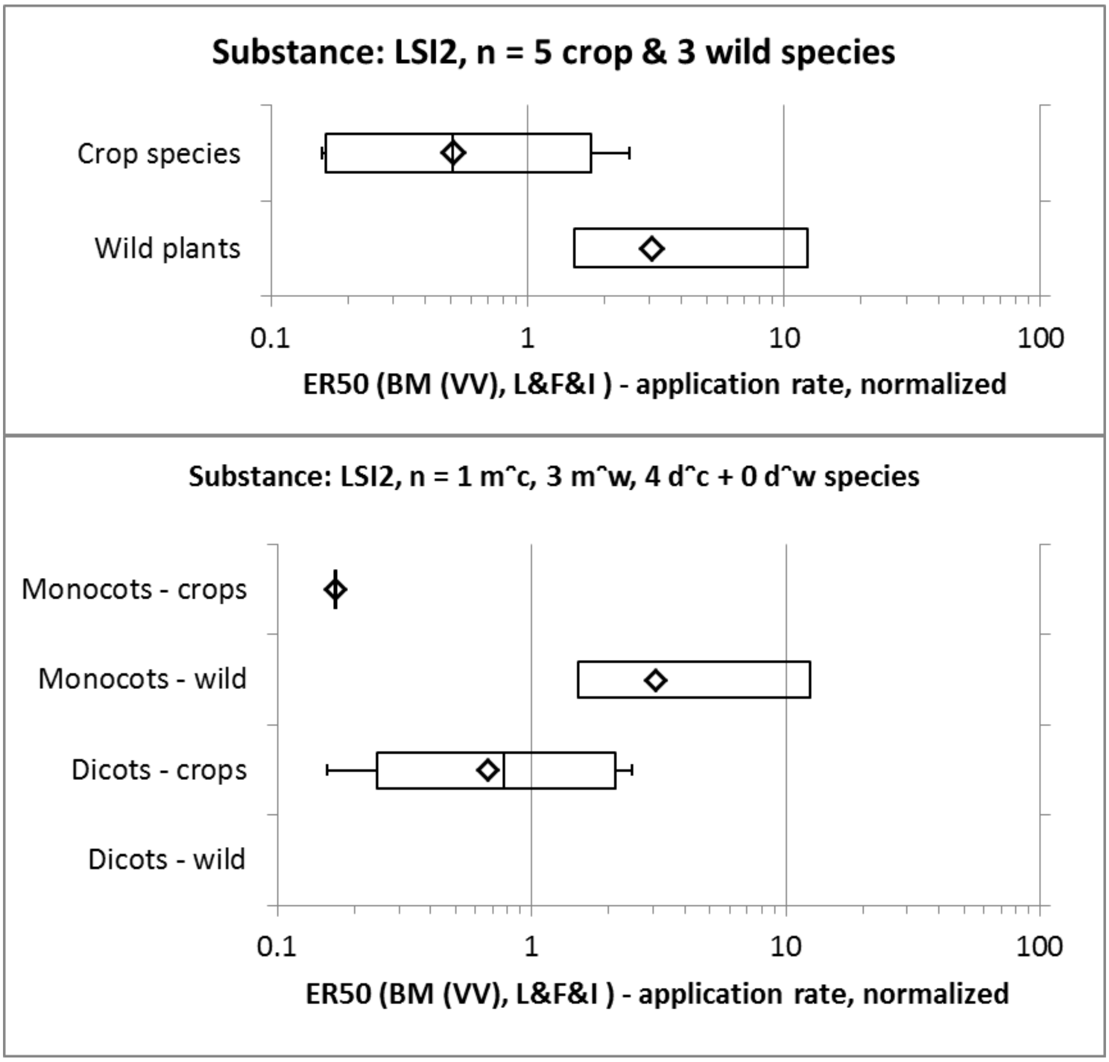


Figure 45: Substance LSI 2 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 45: LSI 2: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**LSI2**	Mono-cots	Crop	1 (species)	0.1549	**0.11**	0.155	**0.055**
Endpoint	ER50		Wild	3 (species)	1.3969		2.815	
Measured	BM	Dicots	Crop	4 (species)	0.1449	#NUM!	0.733	#NUM!
Lab/field	L&F&I		Wild	0 (species)	#NUM!		#NUM!	
SE, VV or all	VV	All	Crop	5 (species)	0.1449	**0.1**	0.537	**0.19**
Censored val.	numeric		Wild	3 (species)	1.3969		2.815	

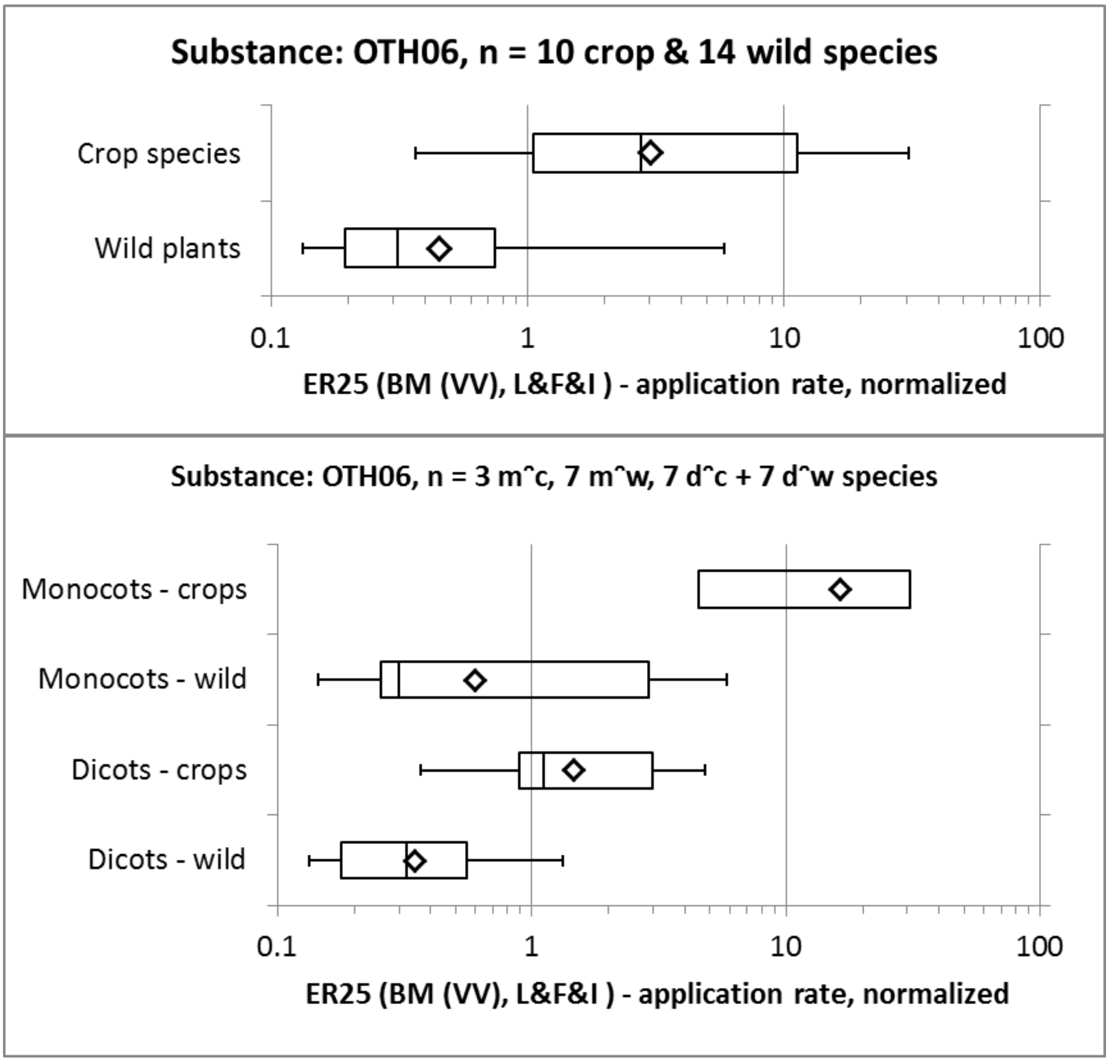


Figure 46: Substance OTH 6 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 46: OTH 6: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**OTH06**	Mono-cots	Crop	3 (species)	4.5499	**32**	16.300	**27**
Endpoint	ER25		Wild	7 (species)	0.1443		0.597	
Measured	BM	Dicots	Crop	7 (species)	0.3662	*2.8*	1.467	*4.2*
Lab/field	L&F&I		Wild	7 (species)	0.1332		0.345	
SE, VV or all	VV	All	Crop	10 (species)	0.3662	*2.8*	3.021	**6.7**
Censored val.	numeric		Wild	14 (species)	0.1332		0.454	

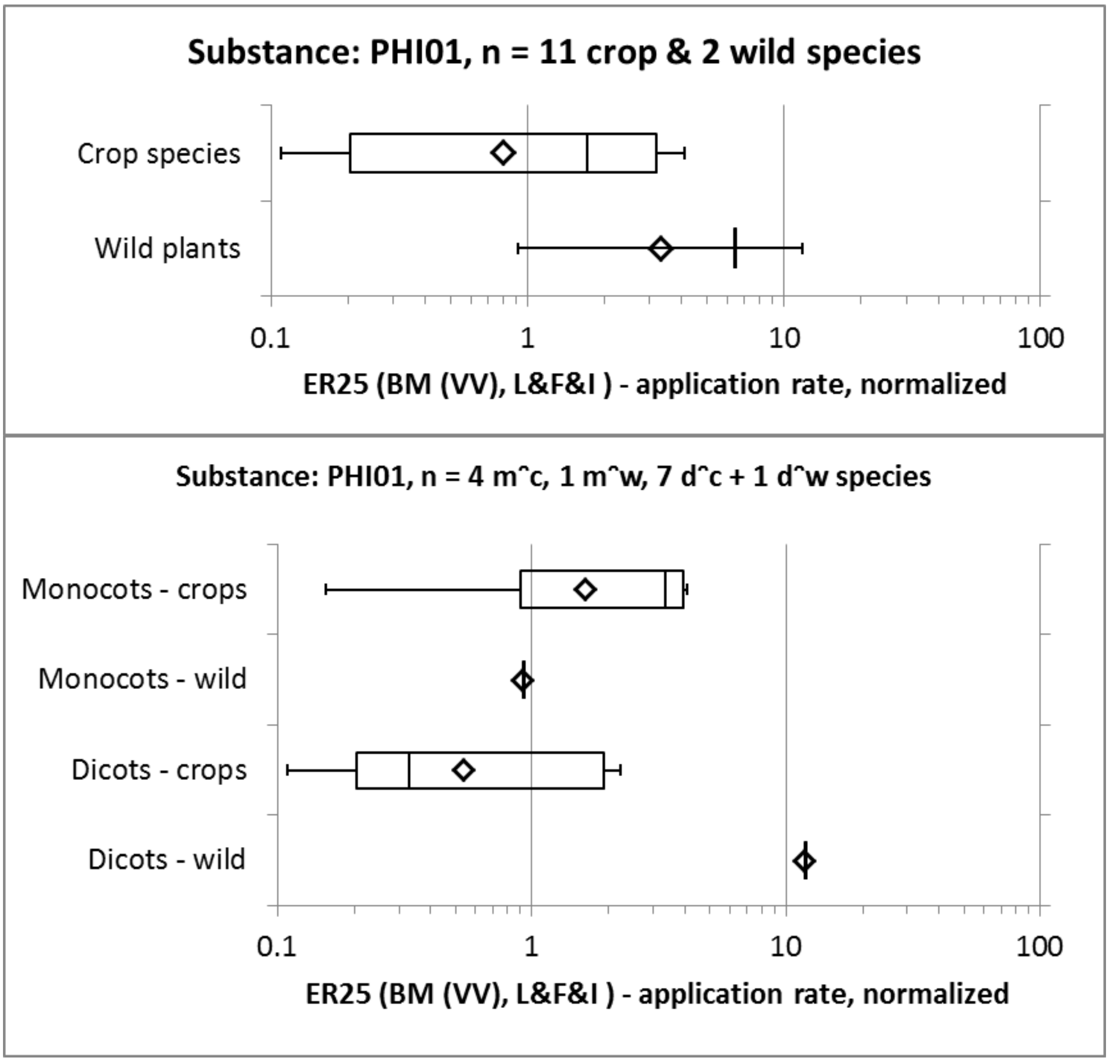


Figure 47: Substance PHI 01 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 47: PHI 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**PHI01**	Mono-cots	Crop	4 (species)	0.1750	0.22	1.572	*1.9*
Endpoint	ER25		Wild	1 (species)	0.8140		0.814	
Measured	BM	Dicots	Crop	7 (species)	0.1243	**0.009**	0.549	**0.041**
Lab/field	L&F&I		Wild	1 (species)	13.4291		13.429	
SE, VV or all	VV	All	Crop	11 (species)	0.1243	**0.15**	0.805	0.24
Censored val.	numeric		Wild	2 (species)	0.8140		3.306	

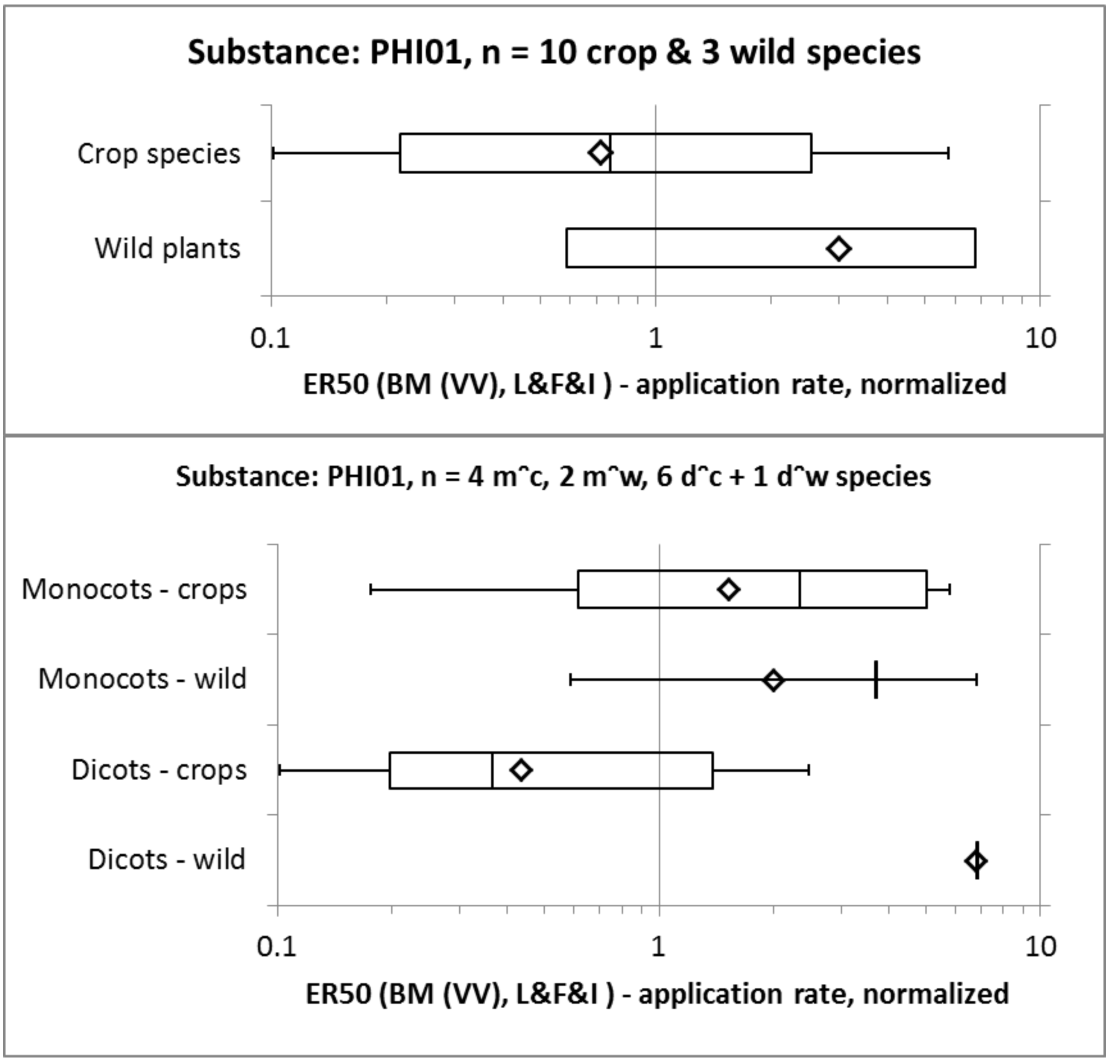


Figure 48: Substance PHI 01- Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 48: PHI 01: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**PHI01**	Mono-cots	Crop	4 (species)	0.1789	0.3	1.551	0.76
Endpoint	ER50		Wild	2 (species)	0.5983		2.033	
Measured	BM	Dicots	Crop	6 (species)	0.1029	**0.015**	0.427	**0.062**
Lab/field	L&F&I		Wild	1 (species)	6.9069		6.907	
SE, VV or all	VV	All	Crop	10 (species)	0.1029	**0.17**	0.715	0.23
Censored val.	numeric		Wild	3 (species)	0.5983		3.056	

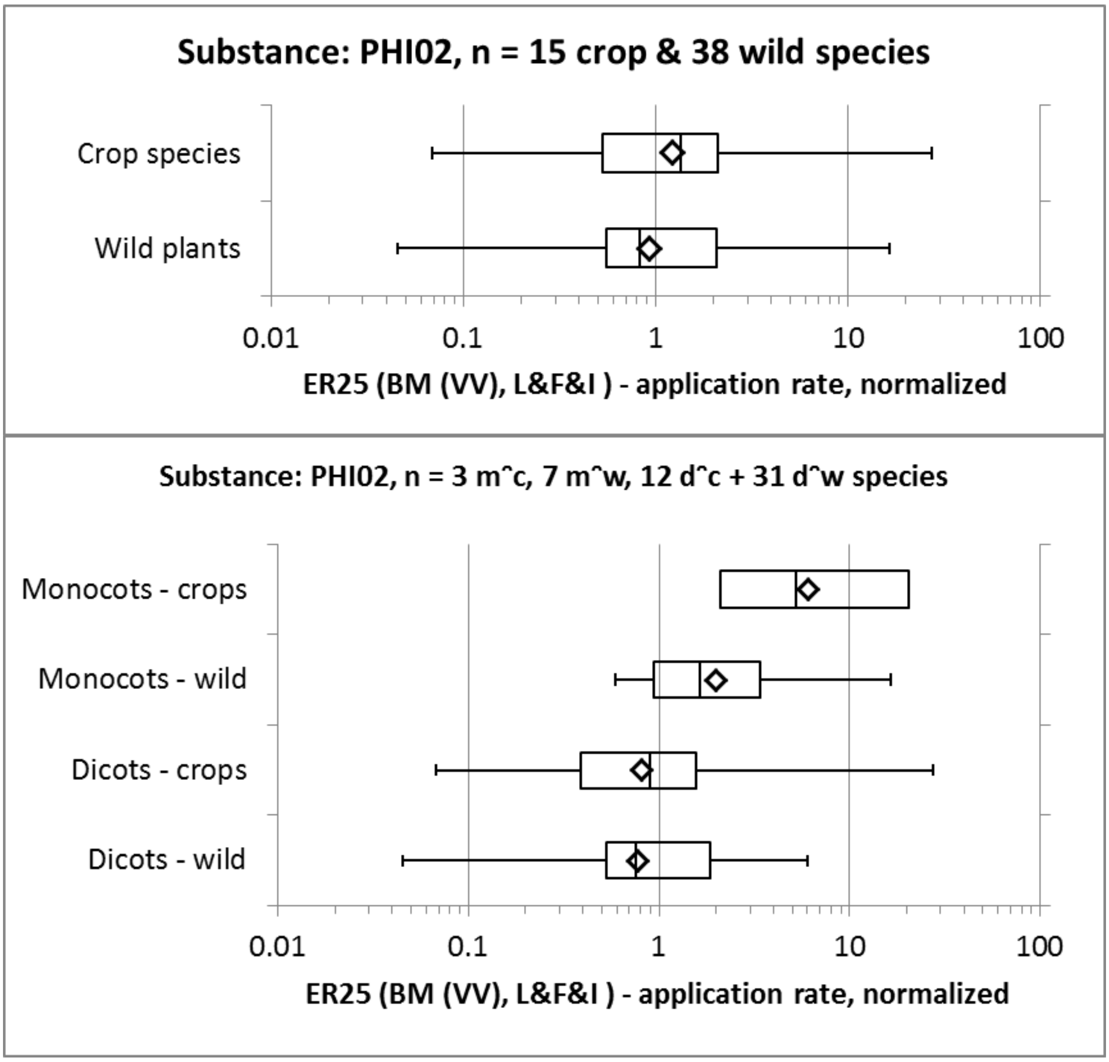


Figure 49: Substance PHI 02 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 49: PHI 02: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**PHI02**	Mono-cots	Crop	3 (species)	2.0915	*3.5*	6.061	*3.0*
Endpoint	ER25		Wild	7 (species)	0.5934		2.000	
Measured	BM	Dicots	Crop	12 (species)	0.0682	*1.5*	0.811	1.0
Lab/field	L&F&I		Wild	31 (species)	0.0456		0.779	
SE, VV or all	VV	All	Crop	15 (species)	0.0682	*1.5*	1.213	*1.3*
Censored val.	numeric		Wild	38 (species)	0.0456		0.927	

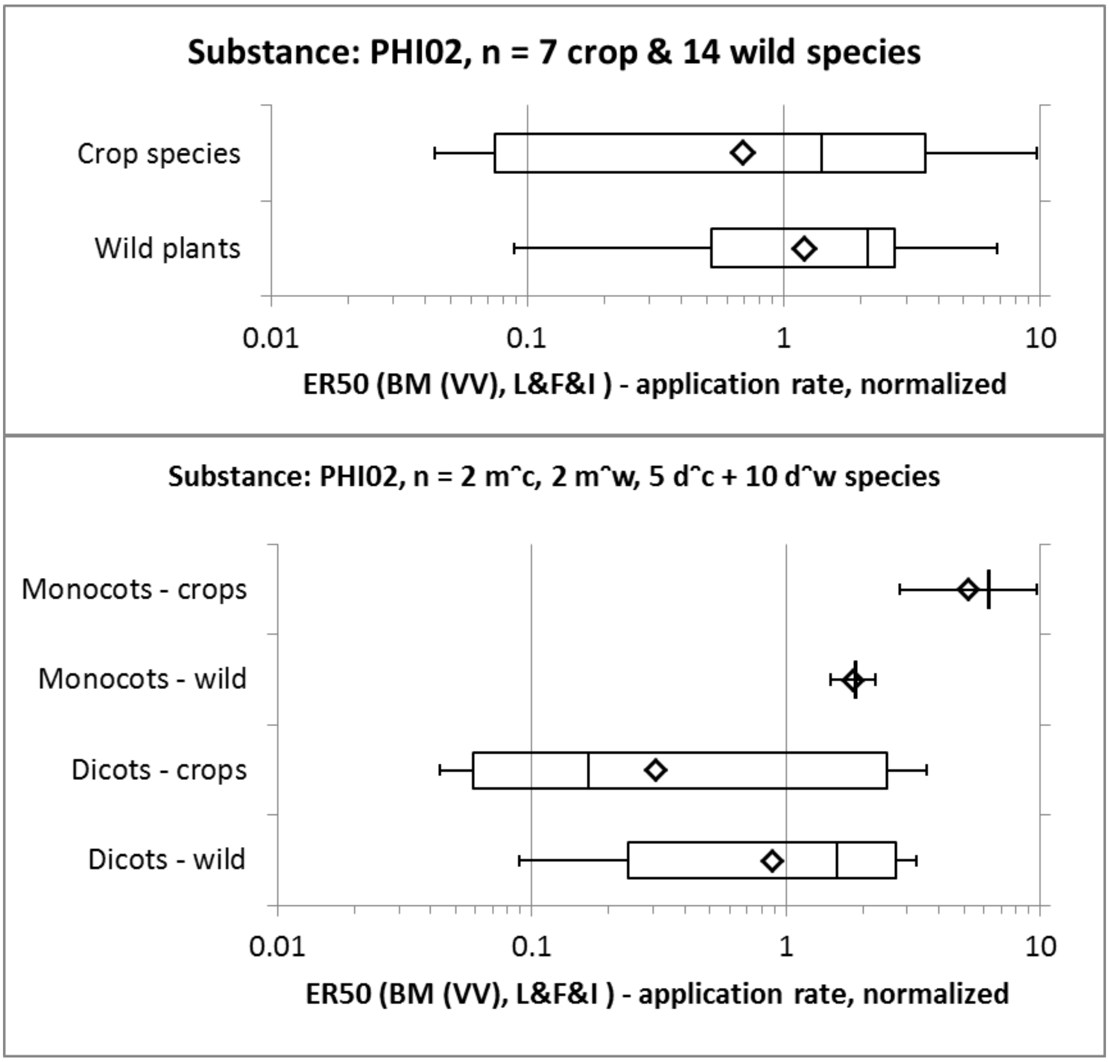


Figure 50: Substance PHI 02 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 50: PHI 02: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**PHI02**	Mono-cots	Crop	2 (species)	2.8146	*1.9*	5.231	*2.9*
Endpoint	ER50		Wild	2 (species)	1.4884		1.829	
Measured	BM	Dicots	Crop	5 (species)	0.0435	0.49	0.307	0.35
Lab/field	L&F&I		Wild	10 (species)	0.0890		0.880	
SE, VV or all	VV	All	Crop	7 (species)	0.0435	0.49	0.689	0.57
Censored val.	numeric		Wild	14 (species)	0.0890		1.204	

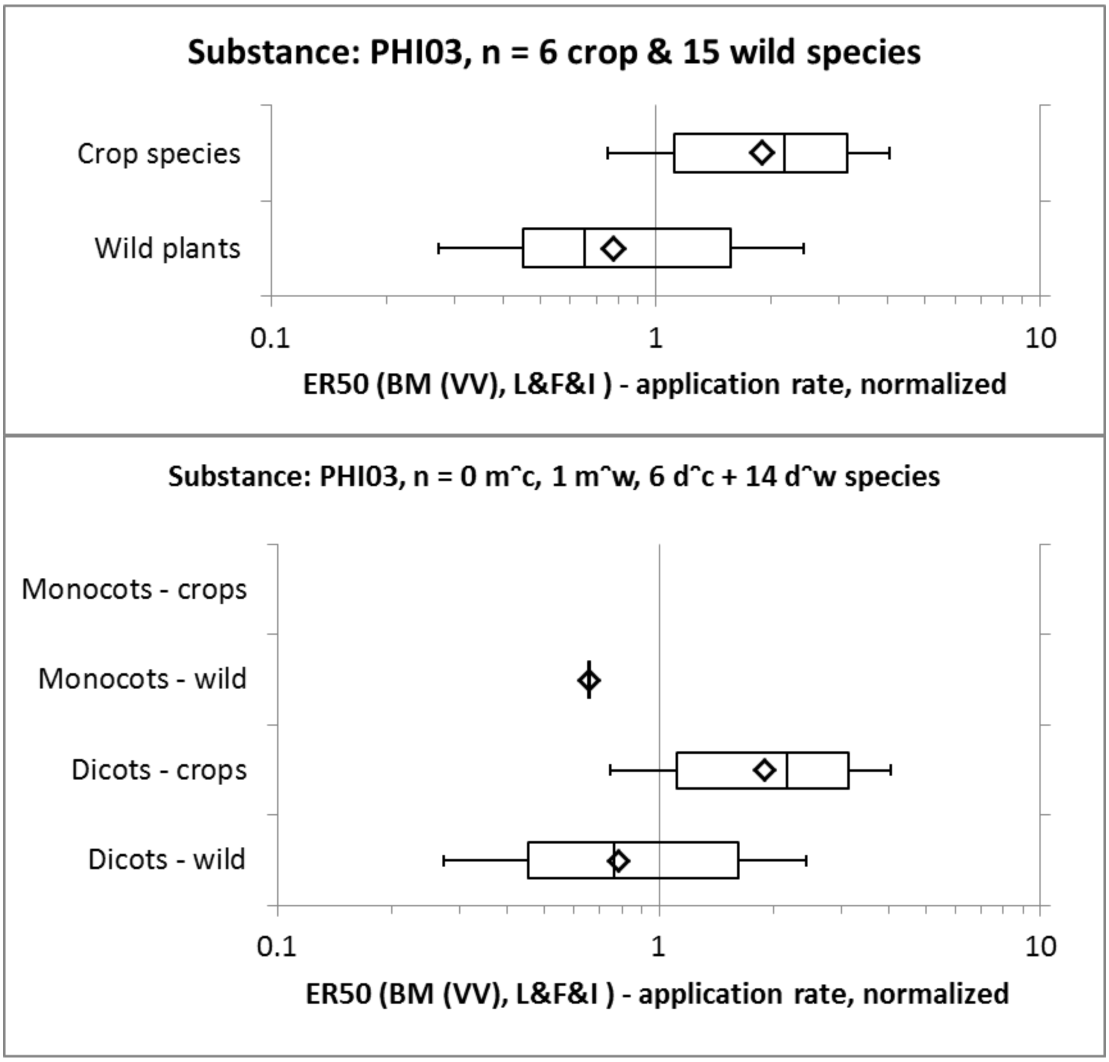


Figure 51: Substance PHI03 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 51: PHI03: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**PHI03**	Mono-cots	Crop	0 (species)	#NUM!	#NUM!	#NUM!	#NUM!
Endpoint	ER50		Wild	1 (species)	0.6542		0.654	
Measured	BM	Dicots	Crop	6 (species)	0.7477	*2.7*	1.897	*2.4*
Lab/field	L&F&I		Wild	14 (species)	0.2729		0.783	
SE, VV or all	VV	All	Crop	6 (species)	0.7477	*2.7*	1.897	*2.4*
Censored val.	numeric		Wild	15 (species)	0.2729		0.774	

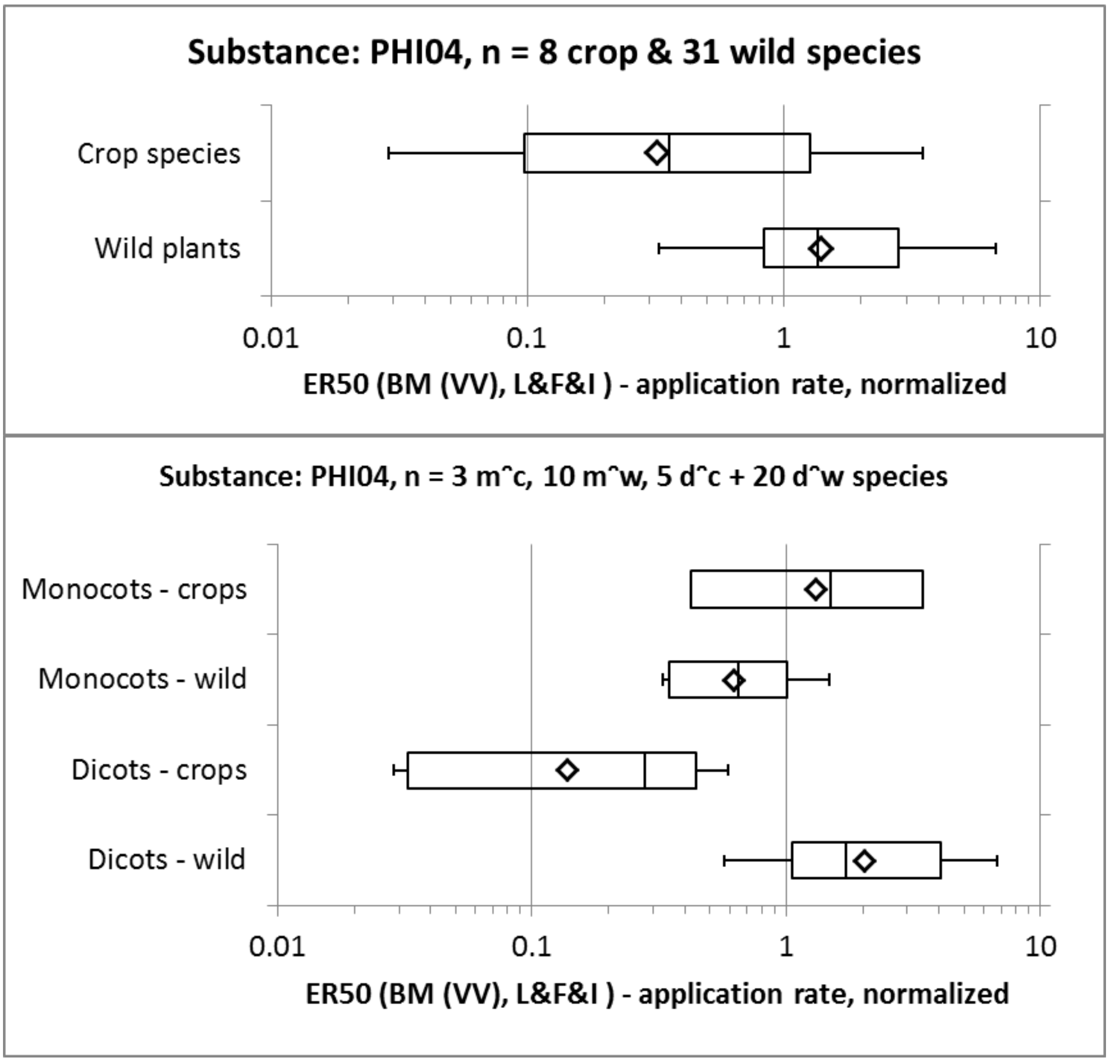


Figure 52: Substance PHI 04 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 52: PHI 04: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**PHI04**	Mono-cots	Crop	3 (species)	0.4248	*1.3*	1.302	*2.1*
Endpoint	ER50		Wild	10 (species)	0.3255		0.622	
Measured	BM	Dicots	Crop	5 (species)	0.0287	**0.05**	0.138	**0.068**
Lab/field	L&F&I		Wild	20 (species)	0.5738		2.034	
SE, VV or all	VV	All	Crop	8 (species)	0.0287	**0.088**	0.320	0.23
Censored val.	numeric		Wild	31 (species)	0.3255		1.391	

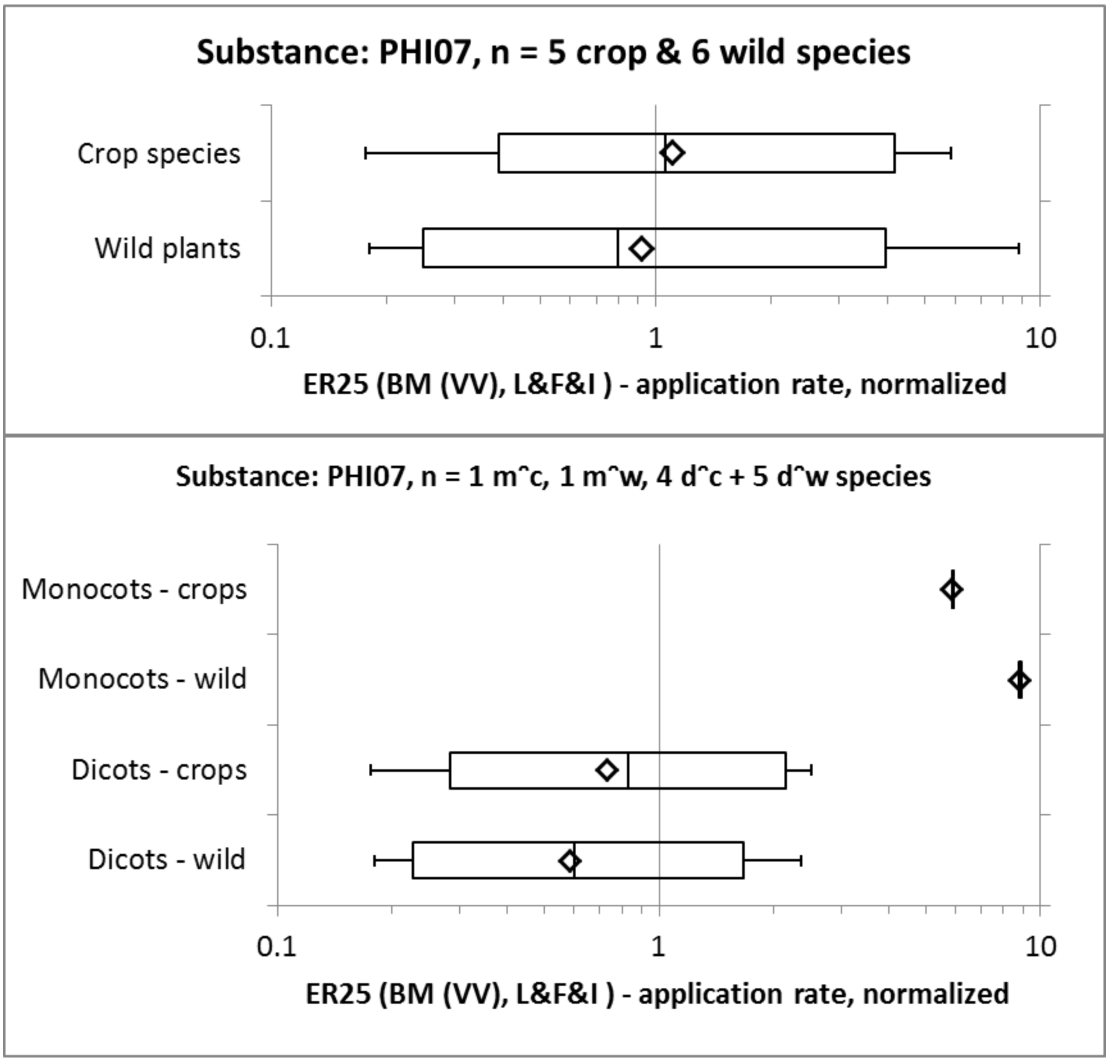


Figure 53: Substance PHI 07 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 53: PHI 07: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**PHI07**	Mono-cots	Crop	1 (species)	5.8633	0.66	5.863	0.66
Endpoint	ER25		Wild	1 (species)	8.8198		8.820	
Measured	BM	Dicots	Crop	4 (species)	0.1759	0.98	0.730	*1.2*
Lab/field	L&F&I		Wild	5 (species)	0.1798		0.584	
SE, VV or all	VV	All	Crop	5 (species)	0.1759	0.98	1.107	*1.2*
Censored val.	numeric		Wild	6 (species)	0.1798		0.919	

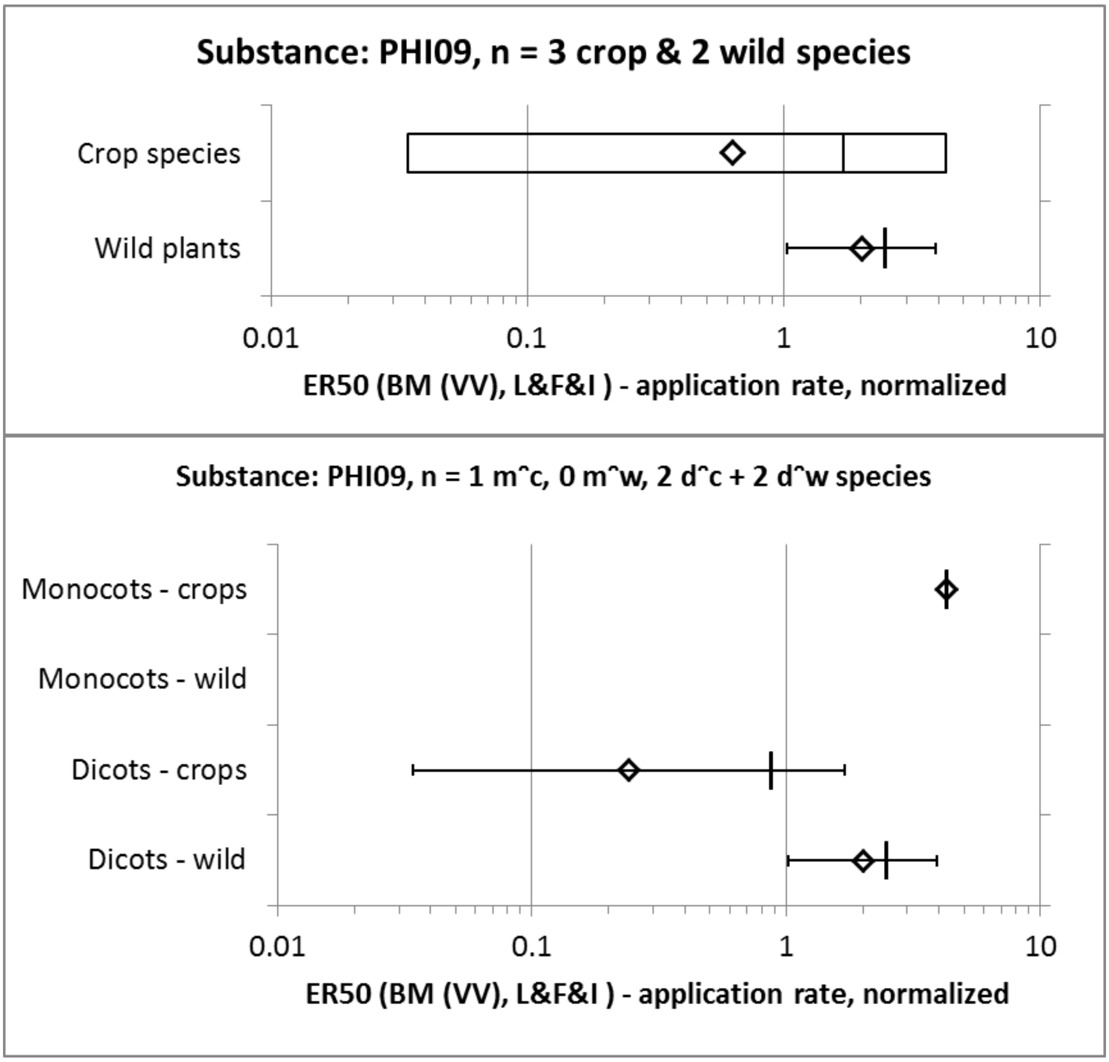


Figure 54: Substance PHI 09 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 54: PHI 09: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**PHI09**	Mono-cots	Crop	1 (species)	4.2675	#NUM!	4.267	#NUM!
Endpoint	ER50		Wild	0 (species)	#NUM!		#NUM!	
Measured	BM	Dicots	Crop	2 (species)	0.0341	**0.033**	0.241	**0.12**
Lab/field	L&F&I		Wild	2 (species)	1.0242		2.005	
SE, W or all	W	All	Crop	3 (species)	0.0341	**0.033**	0.629	0.31
Censored val.	numeric		Wild	2 (species)	1.0242		2.005	

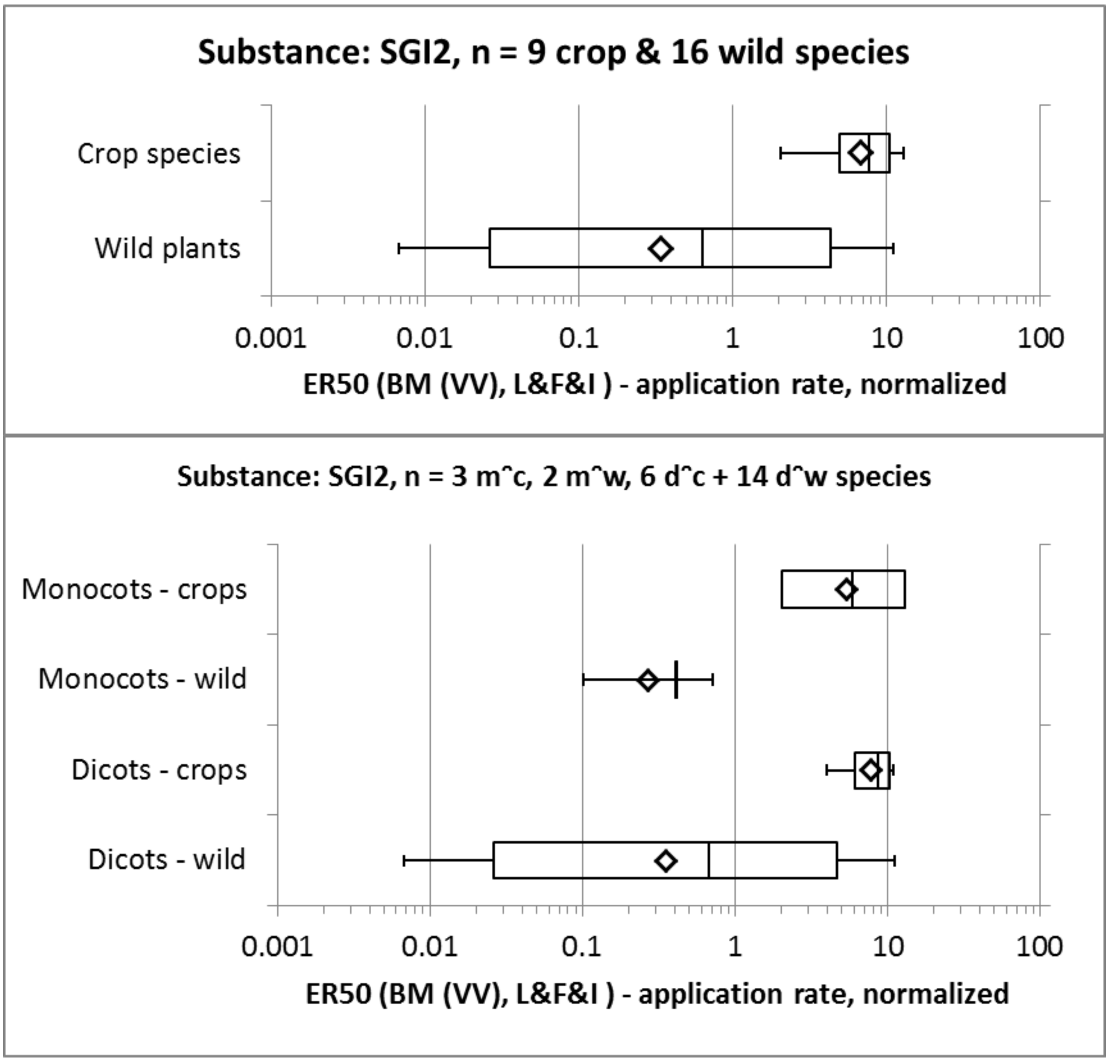


Figure 55: Substance SGI2 - Distribution of numeric ER50 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 55: SGI2: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER50 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**SGI2**	Mono-cots	Crop	3 (species)	2.0316	**20**	5.356	**20**
Endpoint	ER50		Wild	2 (species)	0.1016		0.269	
Measured	BM	Dicots	Crop	6 (species)	4.0016	**592**	7.757	**22**
Lab/field	L&F&I		Wild	14 (species)	0.0068		0.350	
SE, VV or all	VV	All	Crop	9 (species)	2.0316	**301**	6.856	**20**
Censored val.	numeric		Wild	16 (species)	0.0068		0.339	

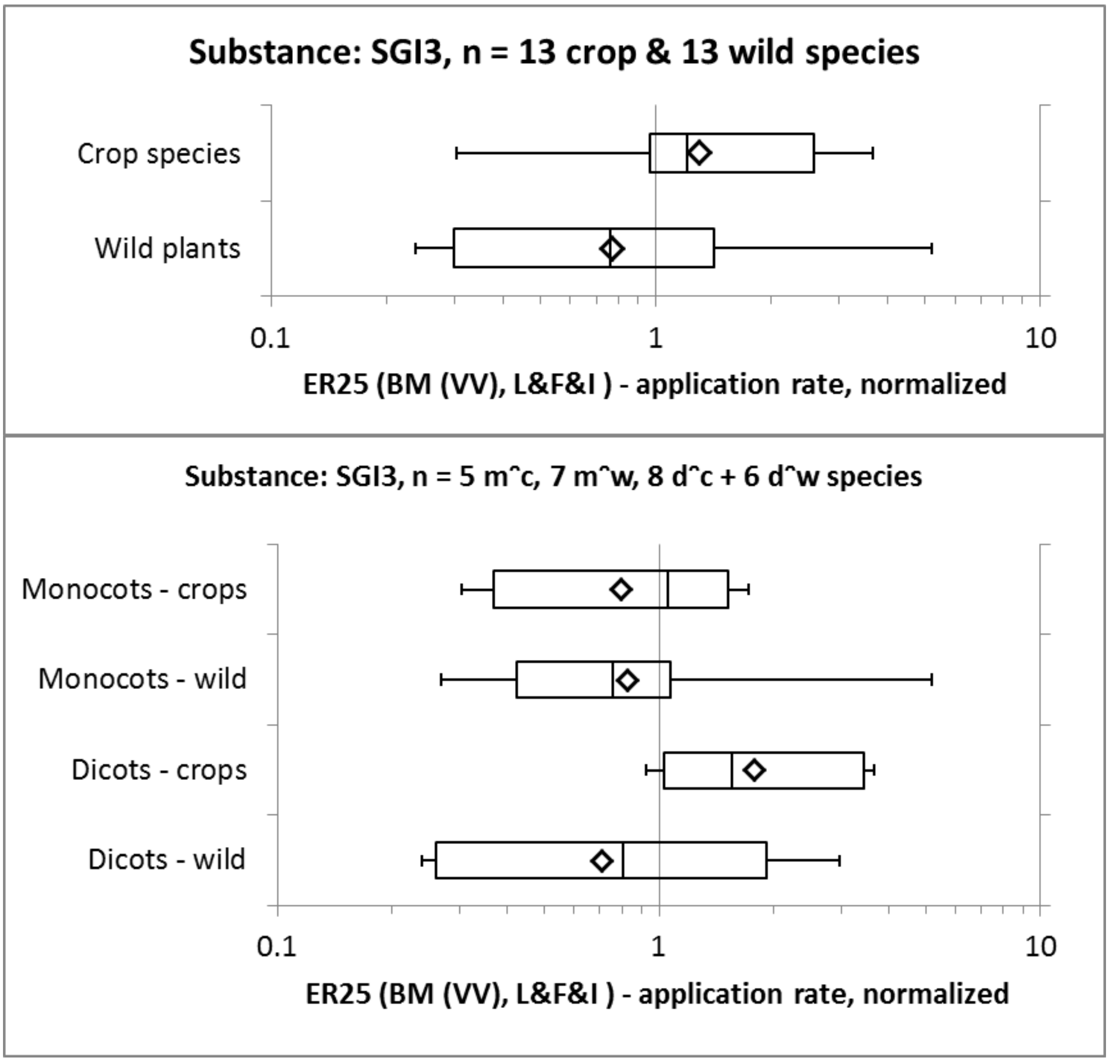


Figure 56: Substance SGI 3 - Distribution of numeric ER25 endpoints (veg. vigour, biomass) of wild plant and crop species, greenhouse- and field data; endpoints by species (any multiple endpoints combined to one geometric mean per species and test design (lab/greenhouse and multispecies/field test endpoints). The rhomb symbols mark the geometric means, the central lines in the boxes the medians.

Table 56: SGI 3: Lowest endpoints, geometric means, and resulting quotients between crop species and wild plant species, (Monocot/dicot). ER25 endpoints. Parameter selection see within table, further explanations see text and figure above.

	Selection	Group		n	Min	Quotient	GeoMean	Quotient
Active	**SGI3**	Mono-cots	Crop	5 (species)	0.3043	*1.1*	0.793	0.96
Endpoint	ER25		Wild	7 (species)	0.2679		0.826	
Measured	BM	Dicots	Crop	8 (species)	0.9236	*3.9*	1.773	*2.5*
Lab/field	L&F&I		Wild	6 (species)	0.2382		0.706	
SE, VV or all	VV	All	Crop	13 (species)	0.3043	*1.3*	1.301	*1.7*
Censored val.	numeric		Wild	13 (species)	0.2382		0.768	

14 Appendix 6 - Additional statistical analysis of Crop and Wild Plants (multiple regression) Z. Gao et al.

Additional statistical analysis of Crop and Wild Plants

Zhenglei Gao, Ulrich Zumkier & Heino Christl

Revised 07 September 2016

Todos

1. Multiple regression for different factors
2. MDD
3. Mixed effect for MoA with AS defined as random factor

Introduction

The working hypothesis is that wild species are more sensitive than crop species, hence **lower** values in ER10, ER25 and ER50 to be expected.

- Find out if there's a significant difference between crop species' and wild species sensitivity (and by which magnitude they differ)
- Compare endpoints of a certain a.s./endpoint combination in terms of crop and wild species (C/W)
- (For some a.s. we have hence two or even three pairs of data each, for others only one (e.g. only ER25 but no ER10 or ER50 values)
- The endpoints are available either as original rates [g/ha] or log-transformed [log(g/ha)] - pick one. Either use the log-transformed values or use the original values and transform within the model (Note ZG transformed within the model, using the natural logarithm, and we maintained this in later revisions).
- (Factorial model, halving the rates reduces the effect as much as doubling the rates increases the effect.)
- Four confounding factors included that are expected to add noise to the data

Results

1. Read in new data

```
require(ProjectTemplate)

## Loading required package: ProjectTemplate

## Warning: package 'ProjectTemplate' was built under R version 3.2.5
```

```
##create.project("P13096D")

#if(Sys.getenv("computername")!="NT3-016") setwd("D:/P13096D/P13096D/") el
se setwd("P13096D/") #UZ correct WD (Personal folder/Heino/?)
rm(list=ls())
load.project()

## Loading project configuration

## Warning in .load.config(override.config): You are missing a configurati
on
## file: config/global.dcf . Defaults will be used.

## Warning in .check.version(config): Your configuration is compatible wit
h version 0.5 of the ProjectTemplate package.
##   Please run ProjectTemplate::migrate.project() to migrate to the insta
lled version 0.7.

## Autoloading cache

## Autoloading data

## Munging data

evalhtml <- FALSE
require(knitr)

## Loading required package: knitr

## Warning: package 'knitr' was built under R version 3.2.5

opts_chunk$set(fig.width=6.5)
opts_chunk$set(comment=NA)
opts_chunk$set(echo=FALSE)
opts_chunk$set(warnings=FALSE)
opts_chunk$set(message=FALSE)

Warning: package 'pander' was built under R version 3.2.5
```

Using LogValue as value column: use value.var to override.

CvW	F_AASI_D	F_AASI_M	F_CMD_D	F_CMD_M	F_GW_D	F_PHI_M	i_AASI_D	i_AASI_M	i_GW_D	L_AASI_D	L_AASI_M	L_ACI_D	L_CMD_D	L_GW_D	L_GW_M	L_OTH_D	L_PHI_D	L_PHI_M	L_SGI_D	L_SGI_M
C	10	0	0	1	15	1	0	0	0	76	35	0	0	10	6	0	10	6	12	6
W	6	2	12	1	27	0	14	2	8	7	18	4	1	10	2	2	0	2	0	2

Using LogValue as value column: use value.var to override.

Cv_W	F_AASI_D	F_AASI_M	F_CMD_D	F_CMD_M	F_GW_D	F_GW_M	F_LSI_D	F_OTH_D	F_OTH_M	F_PHI_D	F_PHI_M	F_SGI_D	F_SGI_M	L_AASI_D	L_AASI_M	L_ACI_D	L_ACI_M	L_CMD_D	L_CMD_M	L_GW_D	L_GW_M	L_ICD_D	L_LSI_D	L_LSI_M	L_PHI_D	L_PHI_M	L_SGI_D	L_SGI_M
C	23	3	0	0	39	16	0	14	7	3	1	12	8	110	51	0	2	1	1	67	23	1	11	11	70	27	21	12
W	22	20	32	1	39	26	1	15	16	12	2	8	10	188	71	9	21	0	0	33	11	0	0	4	58	12	0	3

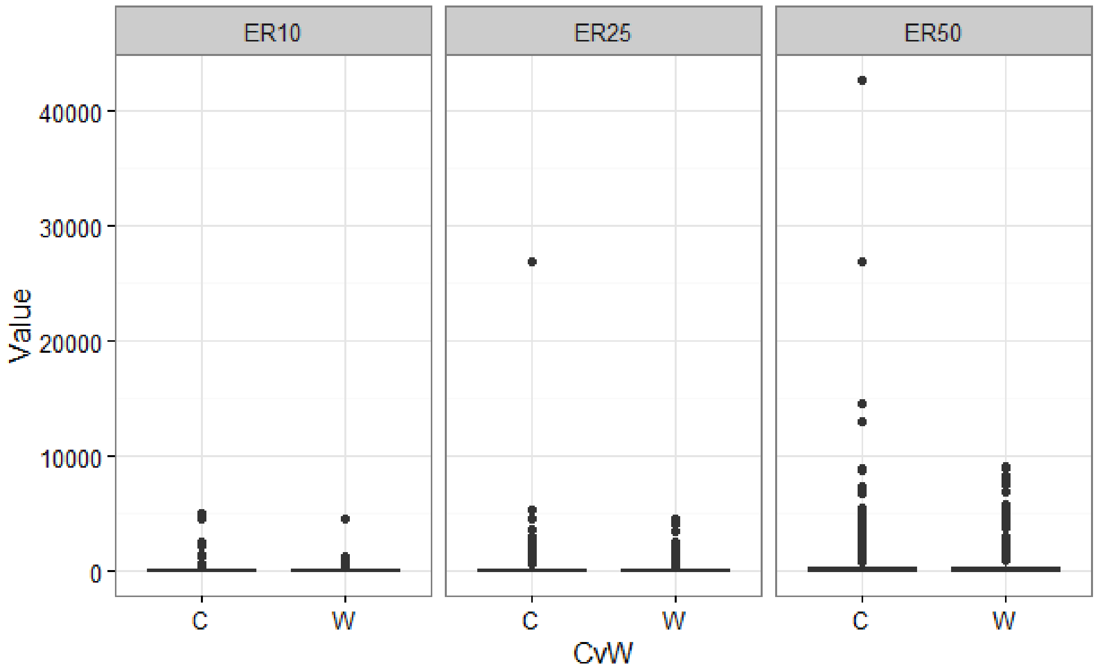
Using LogValue as value column: use value.var to override.

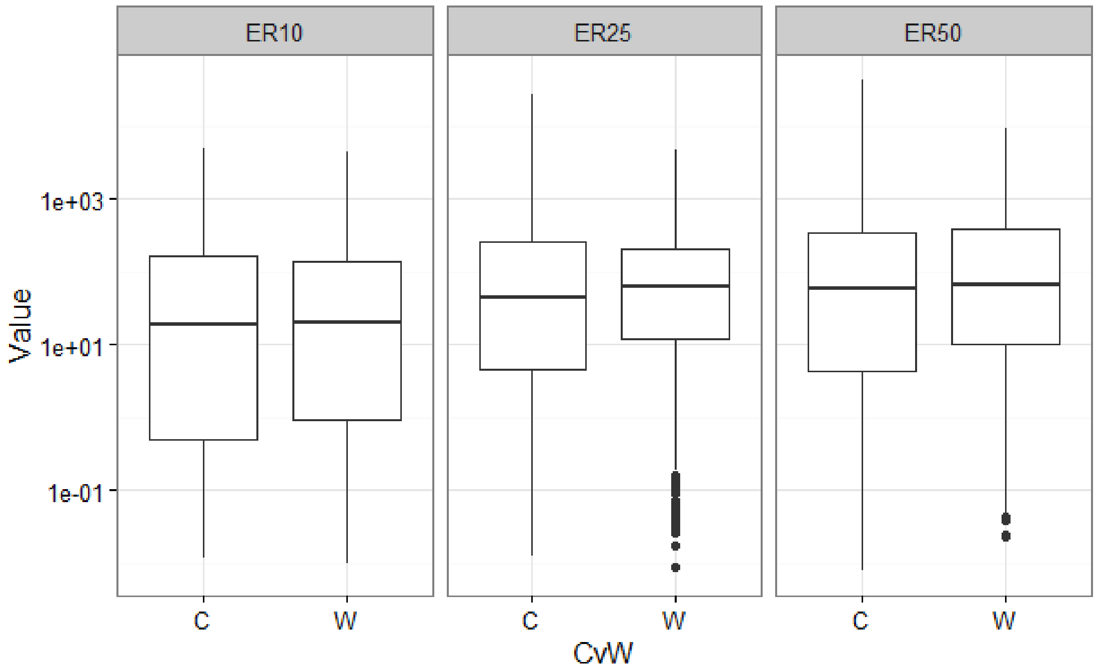
Cv_W	F_AASI_D	F_AASI_M	F_ACI_M	F_CMD_D	F_CMD_M	F_GW_D	F_GW_M	F_LSI_M	F_OTH_D	F_OTH_M	F_PHI_D	F_PHI_Gymn	F_PHI_M	F_SGI_M	i_AASI_D	i_AASI_M	i_GW_D	L_AASI_D	L_AASI_M	L_ACI_D	L_ACI_M	L_CMD_D	L_CMD_M	L_GW_D	L_GW_M	L_ICD_D	L_ICD_M	L_LSI_D	L_LSI_M	L_OTH_D	L_PHI_D	L_PHI_Gymn	L_PHI_M	L_SGI_D	L_SGI_M
C	7	0	1	0	2	15	2	1	1	0	1	0	1	0	0	0	0	224	100	0	6	31	15	85	24	5	6	10	13	0	66	0	22	17	8
W	31	10	16	36	12	24	2	5	4	2	24	3	11	3	25	2	16	107	44	4	25	35	26	57	9	0	0	0	7	2	30	2	8	28	8

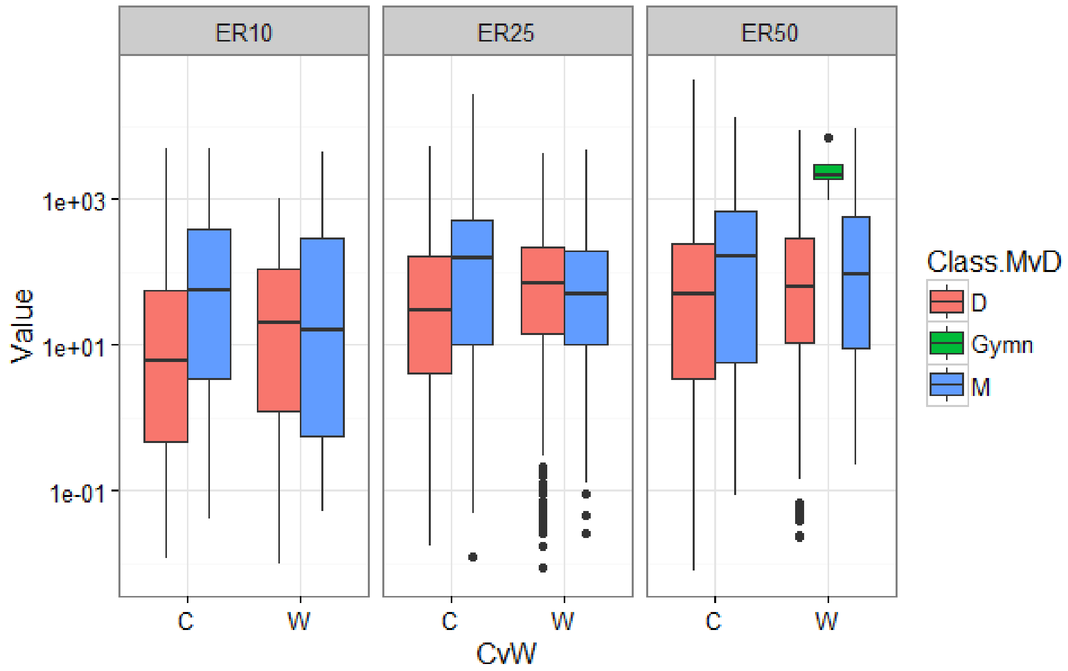
2. EDA

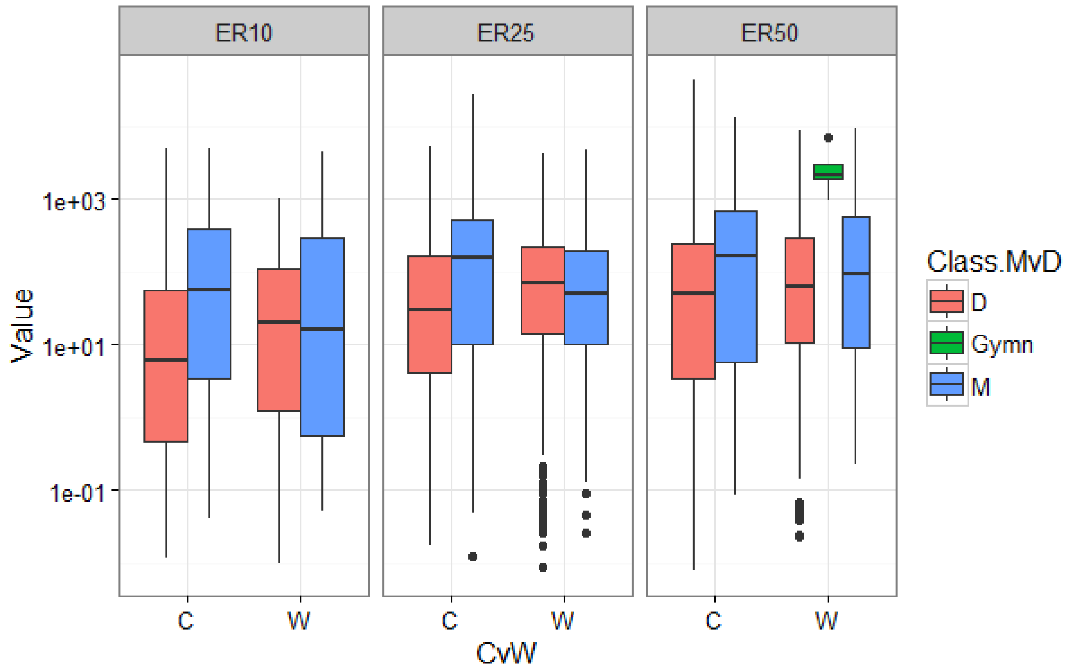
2.1 Compare different Categorization for C and W.

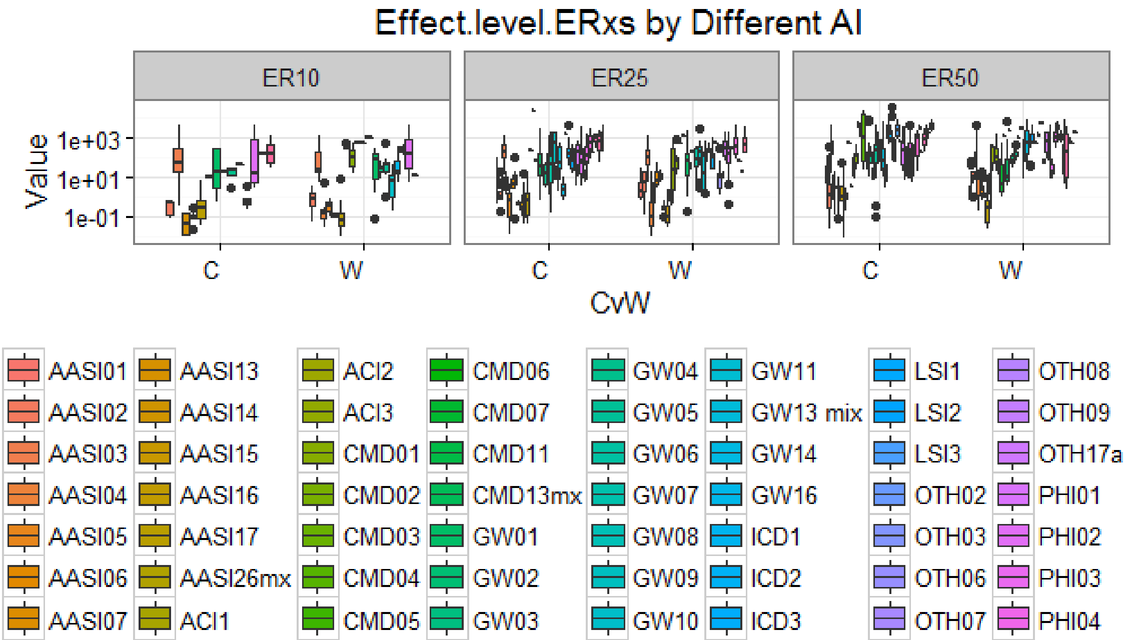
```
Warning: package 'ggplot2' was built under R version 3.2.5
```


ER10
ER25
ER50
Value
40000
30000
20000
10000
0
C
W
CvW

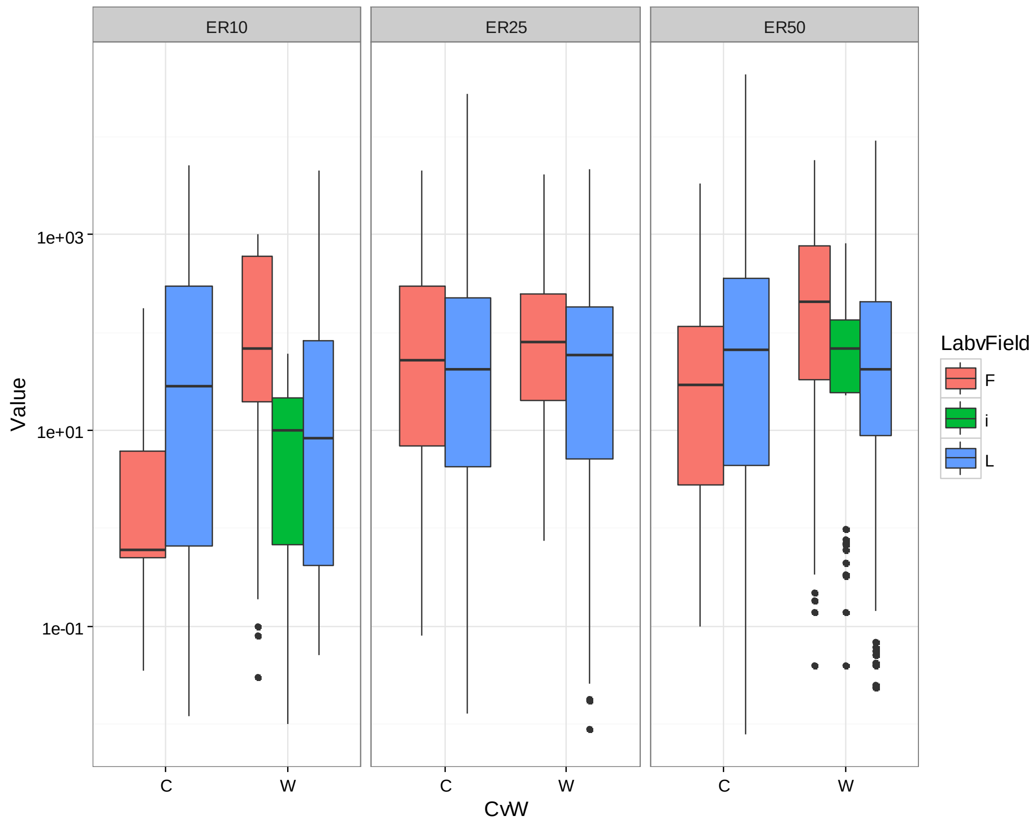

ER10
ER25
ER50
Value
1e+03
1e+01
1e-01
C
W
CvW


ER10
ER25
ER50
1e+03
1e+01
1e-01
Value
C
W
CvW
Class.MvD
D
Gymn
M

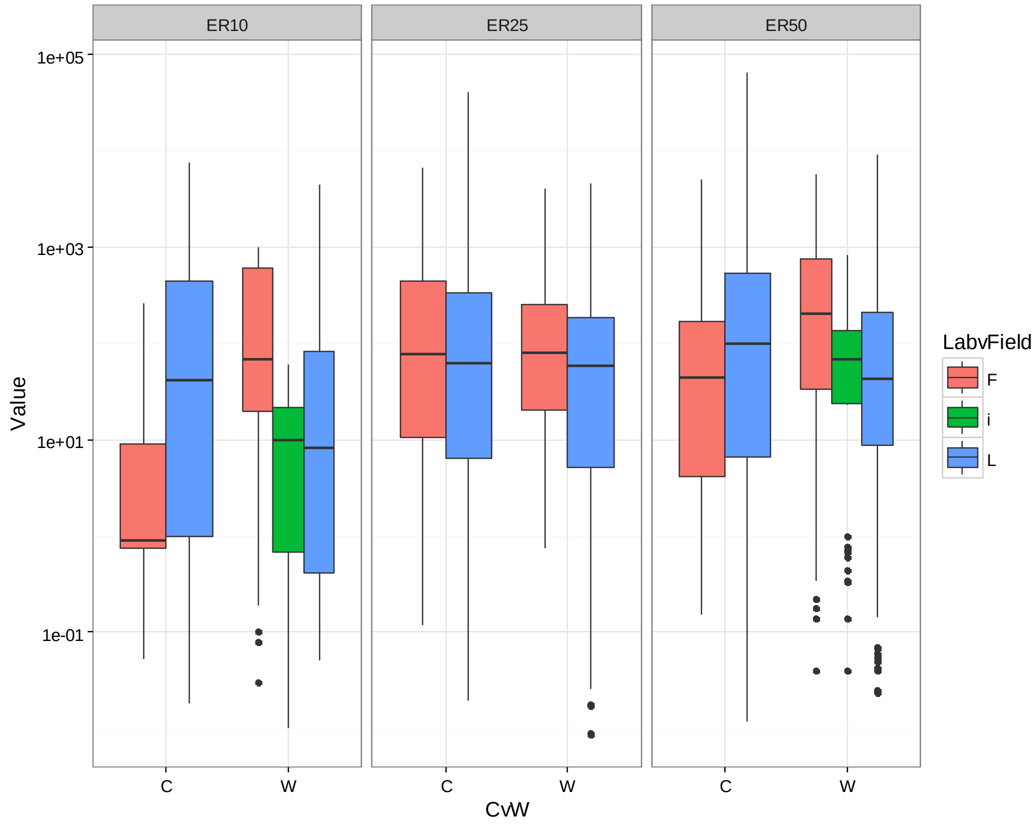

ER10
ER25
ER50
1e+03
1e+01
1e-01
Value
C
W
CvW
Class.MvD
D
Gymn
M



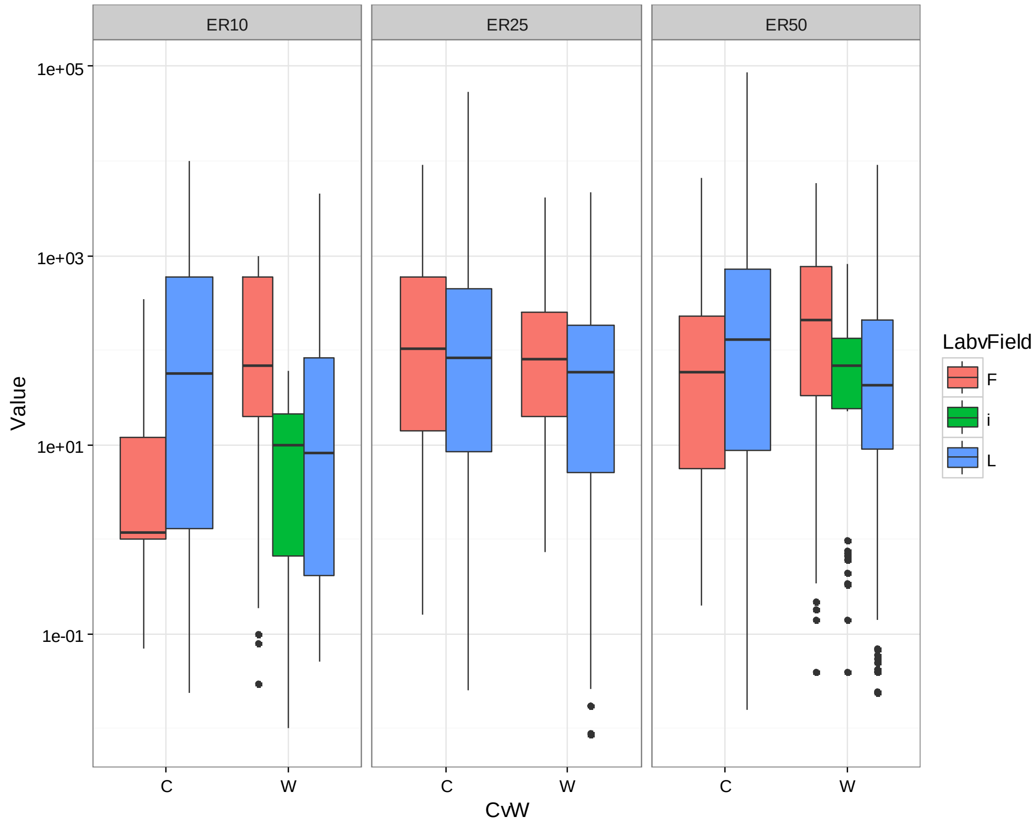
Below: Original data:



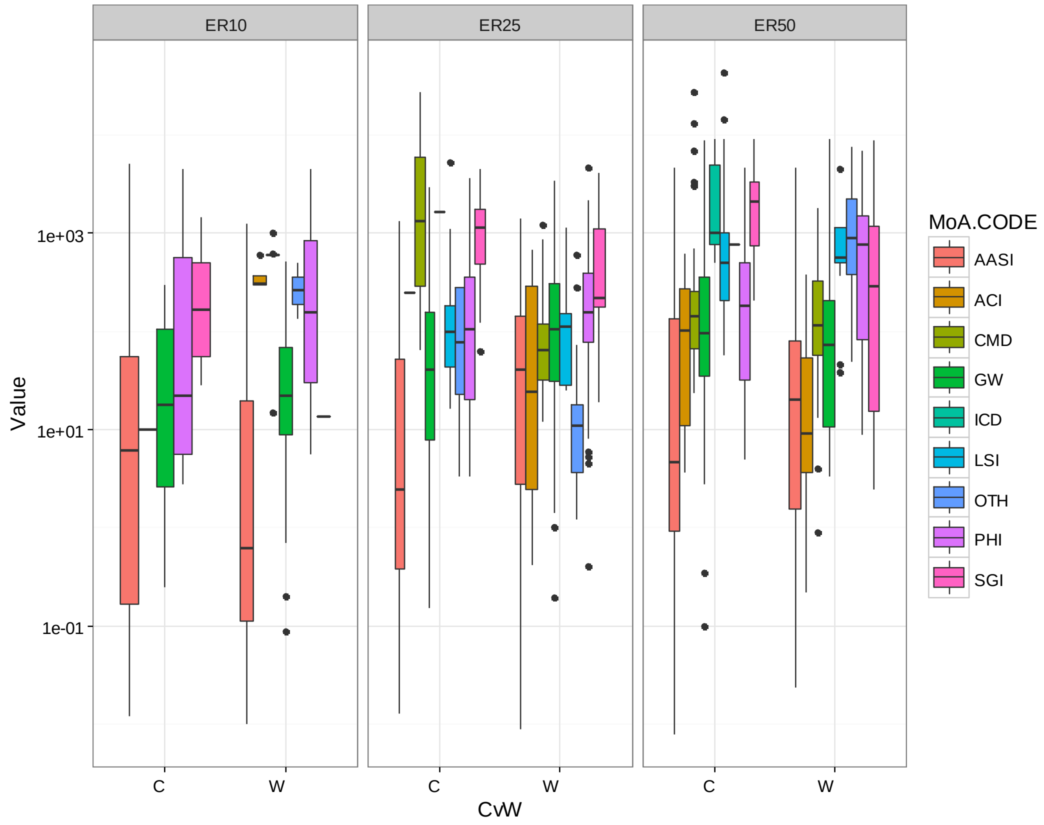
Below: Modified data, crop endpoints increased, (f = 1.5)



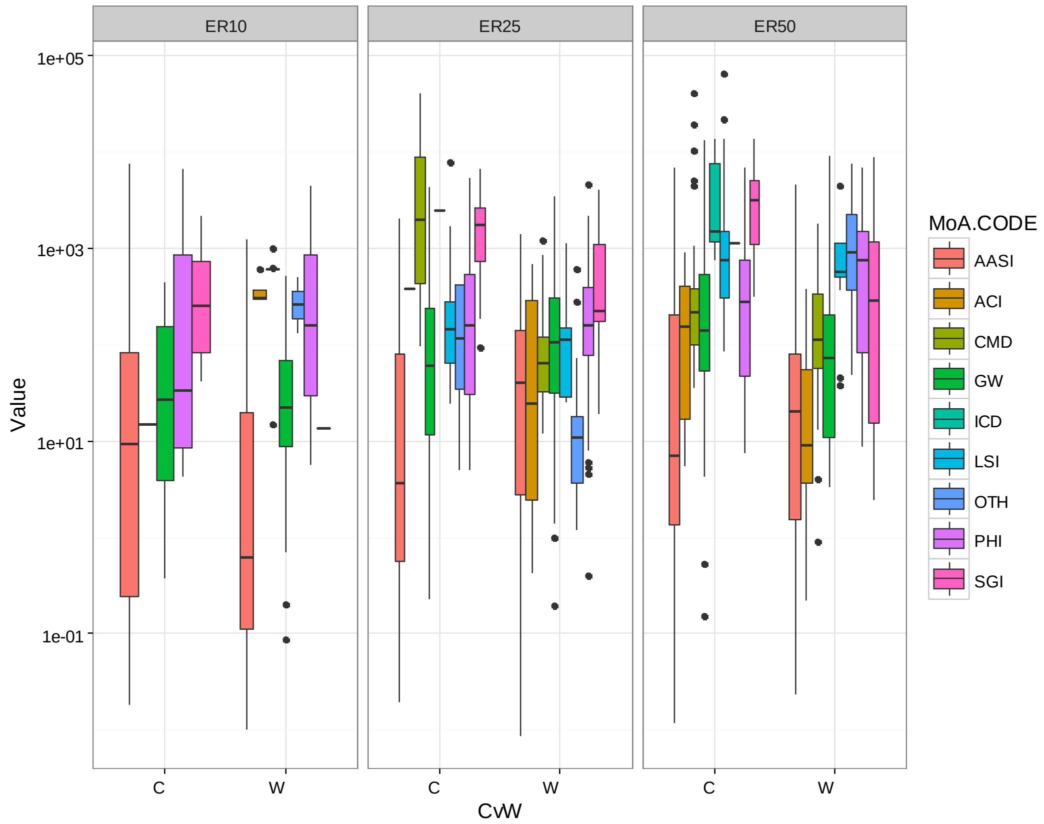
Below: Modified data, crop endpoints increased (f = 2.0)



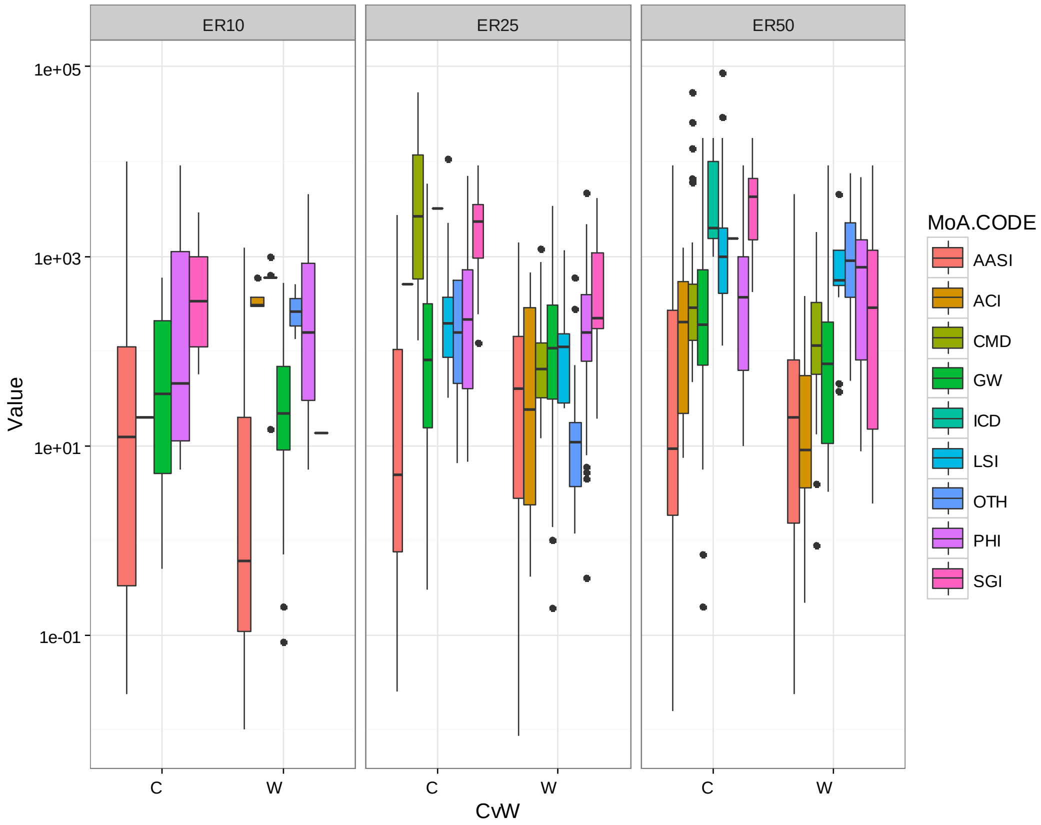
Below: Original data:



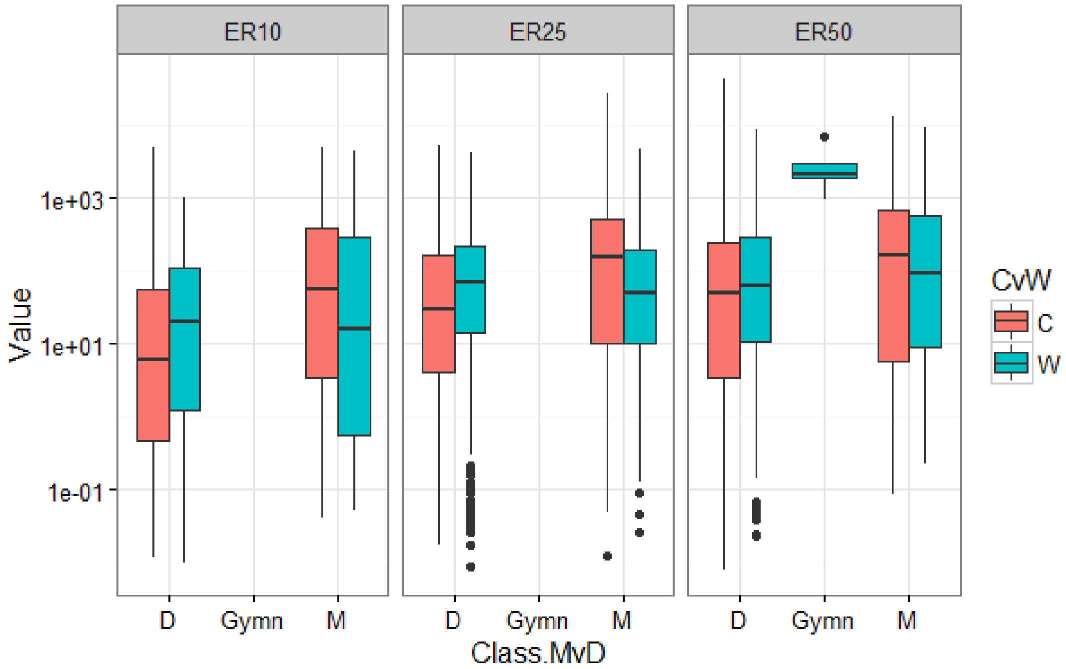
Below: Modified data, crop endpoints increased, (f = 1.5)

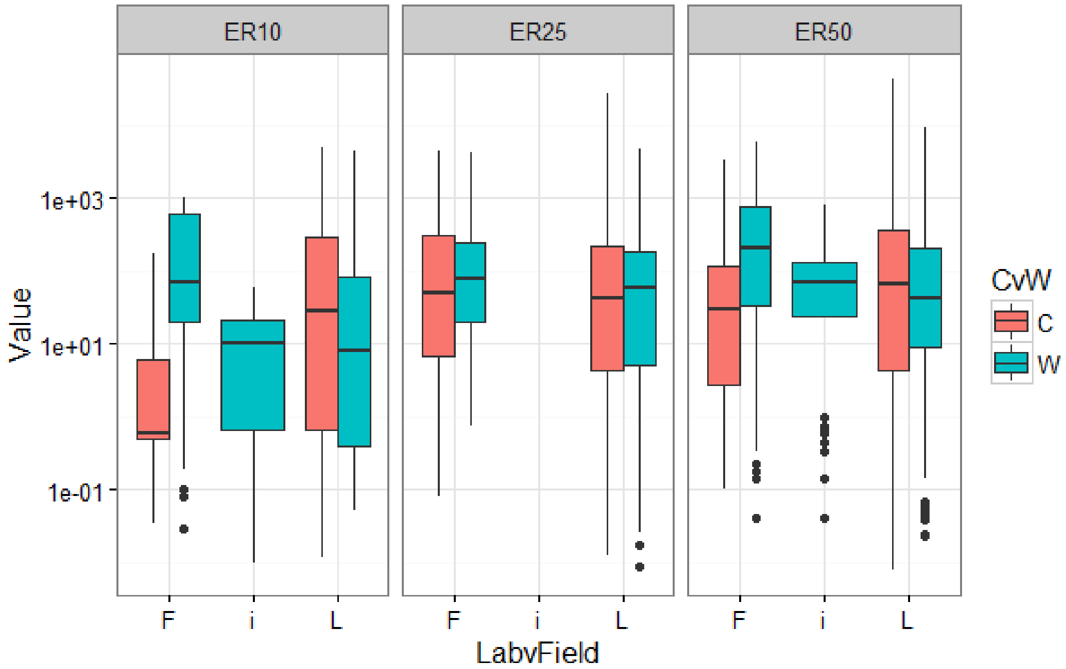


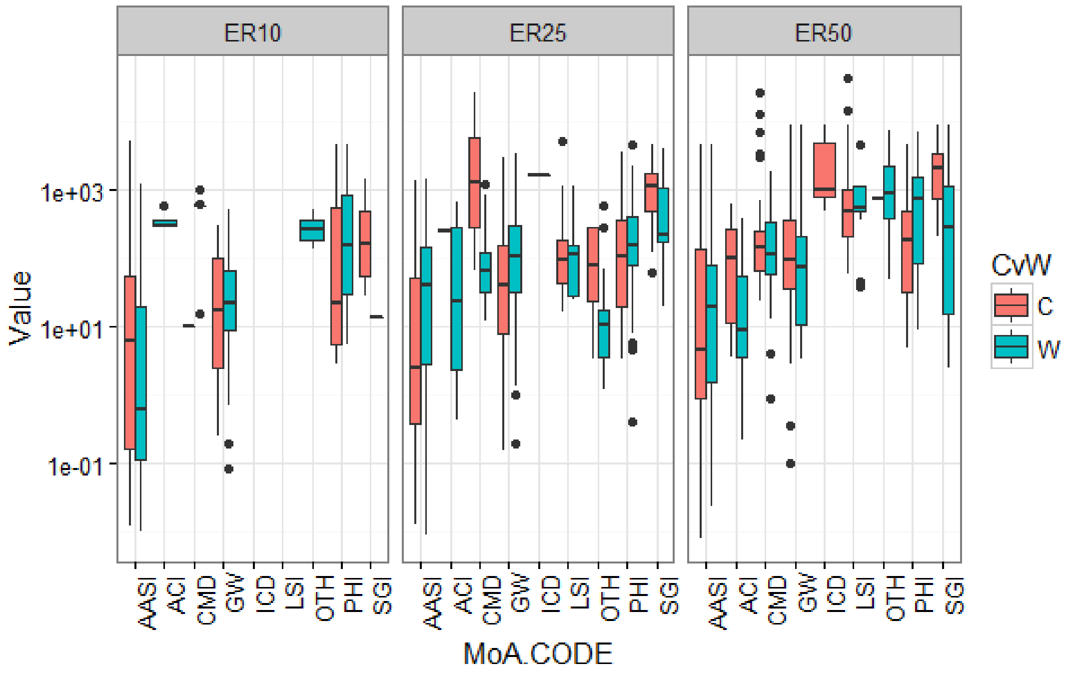
Below: Modified data, crop endpoints increased (f = 2.0)

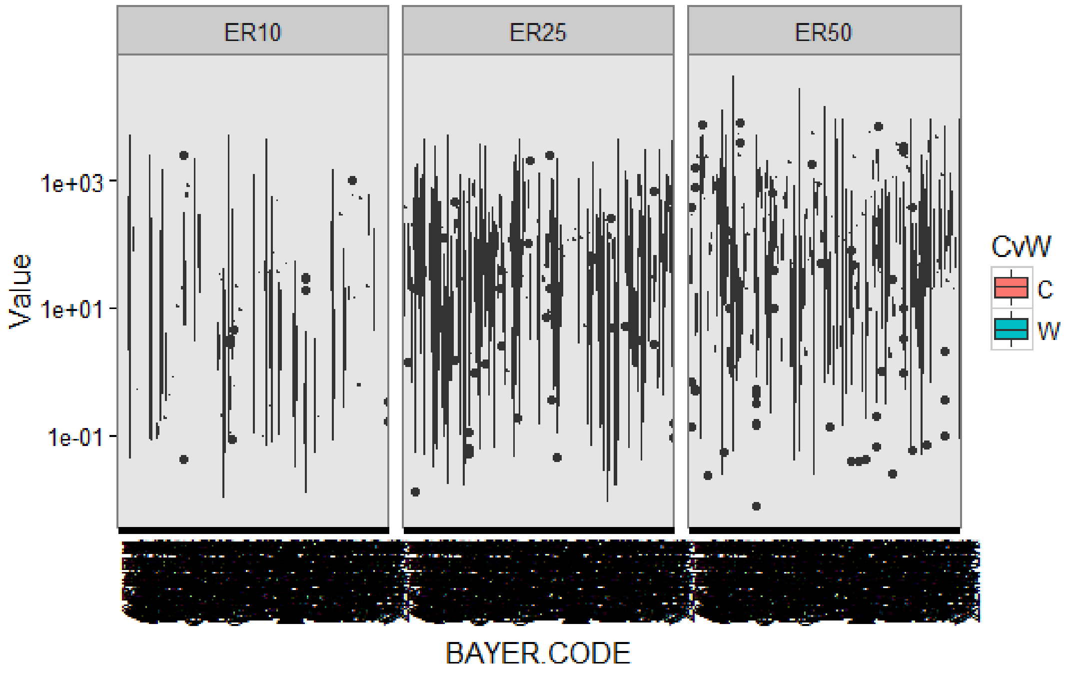


2.1 Compare C and W in different Categorizations.

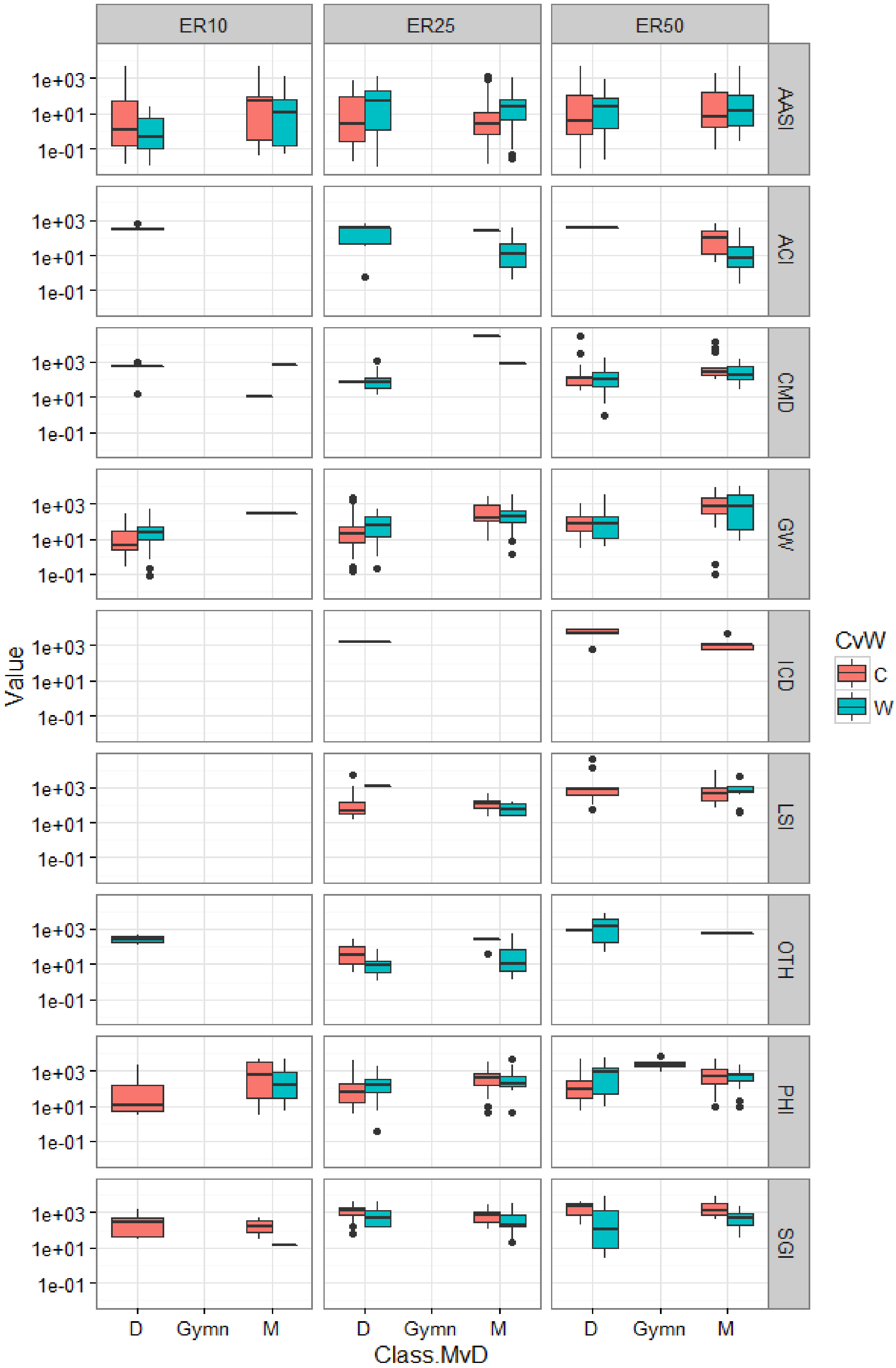


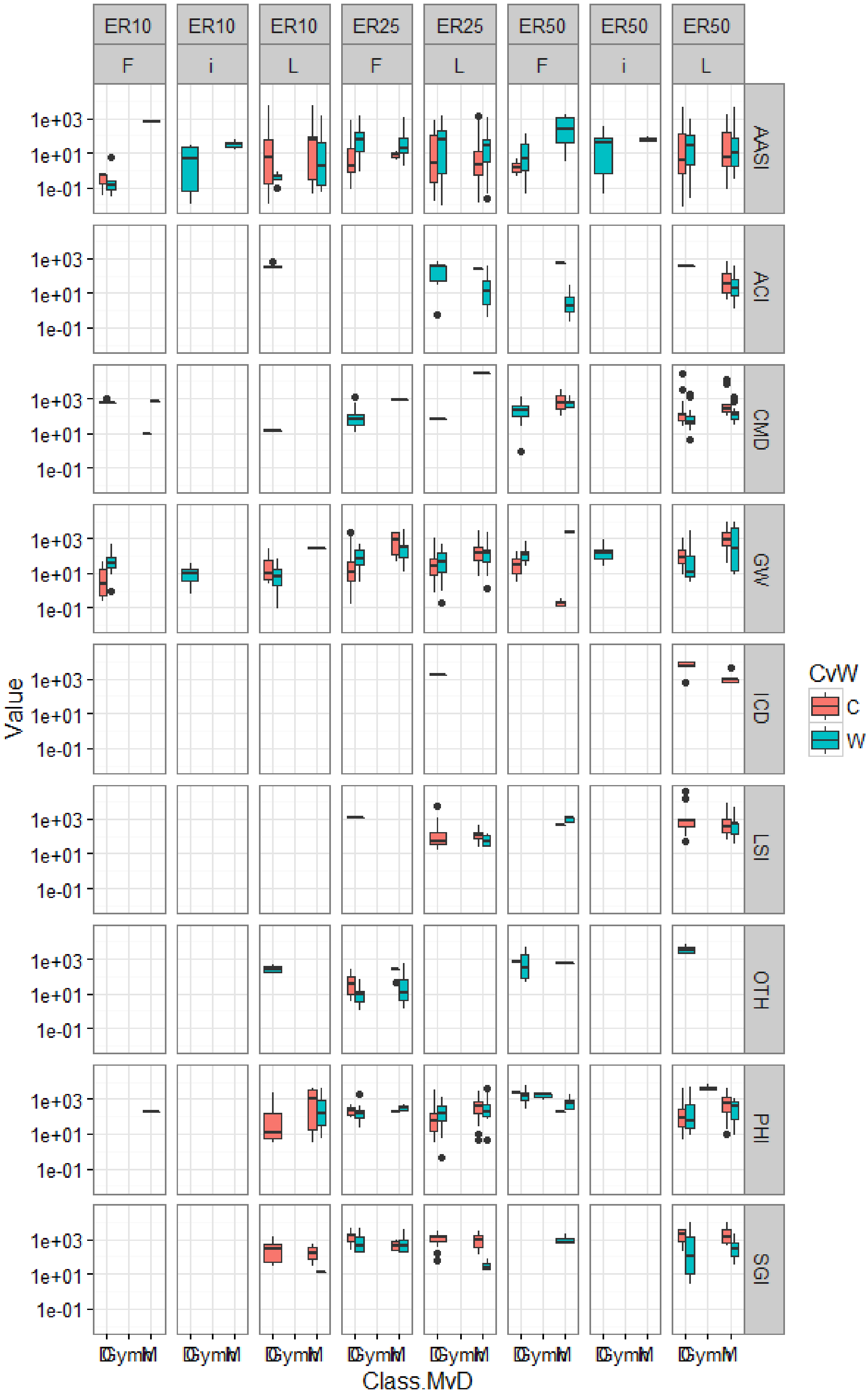



ER10
ER25
ER50
Value
1e+03
1e+01
1e-01
AASI
ACI
CMD
GW
ICD
LSI
OTH
PHI
SGI
MoA.CODE
CvW
C
W


ER10
ER25
ER50
Value
1e+03
1e+01
1e-01
BAYER.CODE
CvW
C
W

2.3 Interaction




ER10
ER10
ER10
ER25
ER25
ER50
ER50
ER50
F
i
L
F
L
F
i
L
AASI
ACI
CMD
GW
ICD
LSI
OTH
PHI
SGI
1e+03
1e+01
1e-01
Value
Class.MvD
CvW
C
W

3. Model Selection

```
Warning: package 'texreg' was built under R version 3.2.5


================================================
                  ER10 Best   ER10 no interaction
------------------------------------------------
(Intercept)        -1.37 *      0.45
                   (0.59)      (0.50)
CvWW                2.32 ***   -0.37
                   (0.69)      (0.43)
MoA.CODEACI         7.23 ***    5.27 ***
                   (1.50)      (1.50)
MoA.CODECMD         5.01 ***    5.59 ***
                   (0.87)      (0.89)
MoA.CODEGW          2.51 ***    2.23 ***
                   (0.45)      (0.45)
MoA.CODEOTH         6.90 ***    4.94 *
                   (2.03)      (2.07)
MoA.CODEPHI         2.66 ***    2.74 ***
                   (0.66)      (0.69)
MoA.CODESGI         3.17 ***    3.38 ***
                   (0.65)      (0.68)
LabvFieldi         -0.91        0.17
                   (0.74)      (0.73)
LabvFieldL          2.72 ***    0.55
                   (0.62)      (0.50)
Class.MvDM          1.20 **     1.38 ***
                   (0.45)      (0.40)
CvWW:LabvFieldL    -5.01 ***
                   (0.91)
CvWW:Class.MvDM     1.62
                   (0.86)
------------------------------------------------
R^2                 0.32        0.26
Adj. R^2            0.30        0.23
Num. obs.         308         308
RMSE                2.72        2.85
================================================
*** p < 0.001, ** p < 0.01, * p < 0.05


========================================================
                         ER25 Best    ER25 no interaction
--------------------------------------------------------
(Intercept)                 1.41 ***     2.00 ***
                           (0.29)       (0.23)
CvWW                        2.24 ***     0.65 ***
                           (0.40)       (0.15)
MoA.CODEACI                 5.95 **      0.62
```

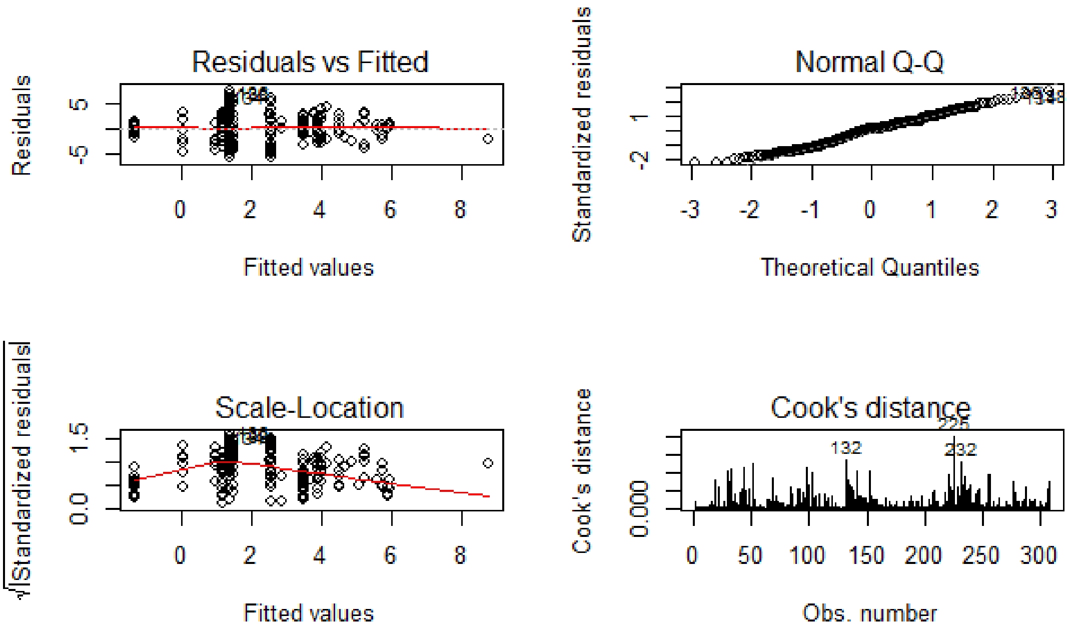
```
                          (1.89)       (0.44)
MoA.CODECMD                3.86 *       1.90 ***
                          (1.89)       (0.45)
MoA.CODEGW                 1.74 ***     1.68 ***
                          (0.28)       (0.20)
MoA.CODEICD                6.09 **      5.75 *
                          (2.30)       (2.38)
MoA.CODELSI                3.41 ***     2.60 ***
                          (0.70)       (0.48)
MoA.CODEOTH                2.17 ***     0.53
                          (0.62)       (0.38)
MoA.CODEPHI                2.81 ***     2.62 ***
                          (0.31)       (0.21)
MoA.CODESGI                5.59 ***     4.28 ***
                          (0.42)       (0.31)
LabvFieldL                -0.11        -0.35
                          (0.27)       (0.19)
Class.MvDM                 0.11         0.48 **
                          (0.30)       (0.16)
CvWW:MoA.CODEACI          -4.10 *
                          (1.73)
CvWW:MoA.CODECMD          -3.26
                          (1.92)
CvWW:MoA.CODEGW           -1.29 ***
                          (0.39)
CvWW:MoA.CODELSI          -1.18
                          (1.20)
CvWW:MoA.CODEOTH          -3.92 ***
                          (0.77)
CvWW:MoA.CODEPHI          -0.99 *
                          (0.40)
CvWW:MoA.CODESGI          -2.65 ***
                          (0.68)
CvWW:LabvFieldL           -0.74 *
                          (0.38)
CvWW:Class.MvDM           -0.53
                          (0.32)
MoA.CODEACI:Class.MvDM    -1.83
                          (0.95)
MoA.CODECMD:Class.MvDM     3.93 *
                          (1.91)
MoA.CODEGW:Class.MvDM      1.99 ***
                          (0.39)
MoA.CODELSI:Class.MvDM    -0.25
                          (0.96)
MoA.CODEOTH:Class.MvDM     1.52 *
                          (0.69)
MoA.CODEPHI:Class.MvDM     1.34 **
                          (0.47)
MoA.CODESGI:Class.MvDM    -0.54
                          (0.61)
--------------------------------------------------------
```

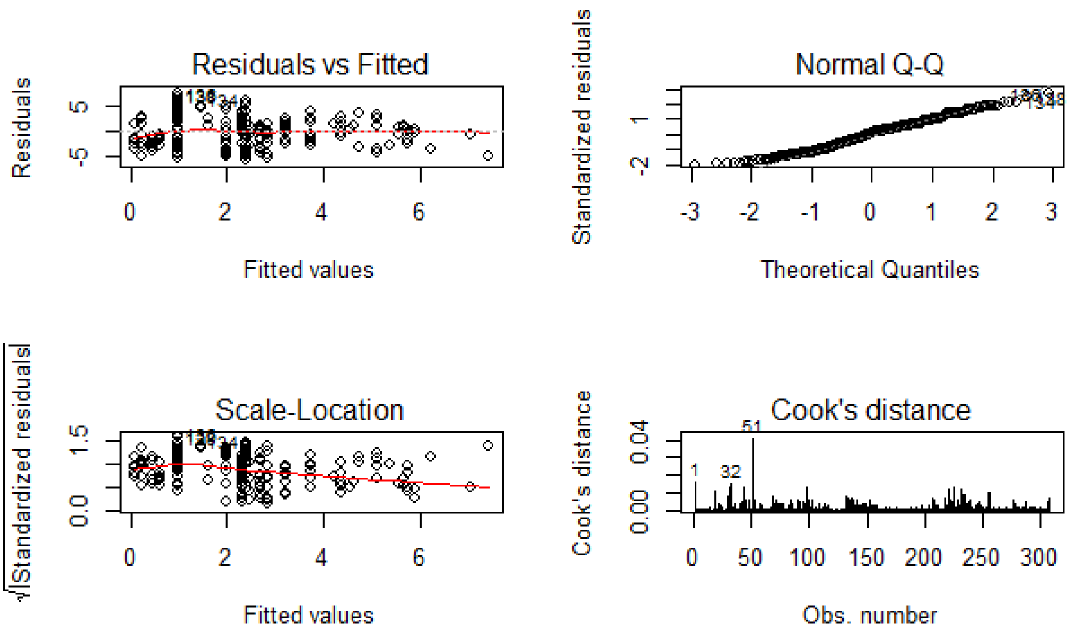
```
R^2                        0.31         0.24
Adj. R^2                   0.29         0.24
Num. obs.               1148         1148
RMSE                       2.29         2.38
========================================================
*** p < 0.001, ** p < 0.01, * p < 0.05


===================================================================
                                   ER50 best    ER50 no interaction
-------------------------------------------------------------------
(Intercept)                           0.37          2.25 ***
                                     (0.82)        (0.22)
CvWW                                  1.25         -0.01
                                     (0.91)        (0.14)
MoA.CODEACI                          11.14 ***      0.03
                                     (3.31)        (0.35)
MoA.CODECMD                          10.69 ***      2.65 ***
                                     (2.65)        (0.21)
MoA.CODEGW                            2.81 **       2.29 ***
                                     (0.99)        (0.18)
MoA.CODEICD                           6.40 ***      5.18 ***
                                     (0.98)        (0.70)
MoA.CODELSI                           9.44 **       3.74 ***
                                     (3.22)        (0.40)
MoA.CODEOTH                           6.26 **       4.31 ***
                                     (2.33)        (0.78)
MoA.CODEPHI                           9.45 ***      3.16 ***
                                     (1.99)        (0.21)
MoA.CODESGI                           4.55 **       3.76 ***
                                     (1.72)        (0.31)
LabvFieldi                            0.58          0.18
                                     (0.59)        (0.39)
LabvFieldL                            1.55         -0.27
                                     (0.83)        (0.19)
Class.MvDGymn                         0.01          2.48 *
                                     (1.33)        (1.04)
Class.MvDM                           -2.45          0.74 ***
                                     (2.02)        (0.15)
CvWW:MoA.CODEACI                    -12.55 ***
                                     (3.05)
CvWW:MoA.CODECMD                     -7.19 **
                                     (2.65)
CvWW:MoA.CODEGW                       0.46
                                     (1.15)
CvWW:MoA.CODELSI                     -6.40 *
                                     (3.16)
CvWW:MoA.CODEOTH                     -1.86
                                     (2.60)
CvWW:MoA.CODEPHI                     -3.75
                                     (2.04)
CvWW:MoA.CODESGI                     -2.92 ***
```

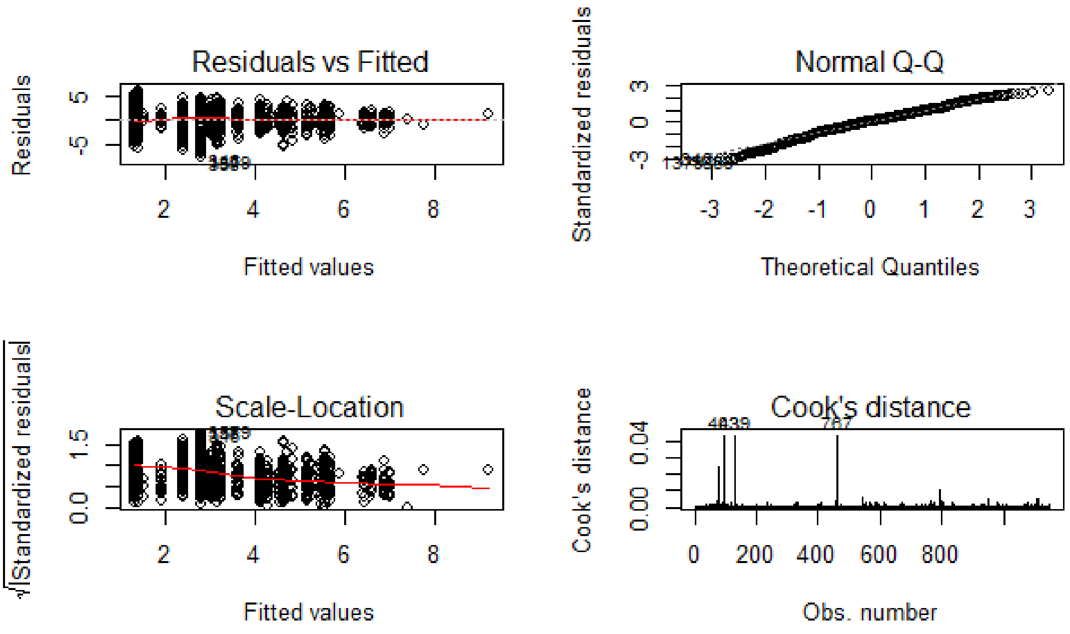
	(0.61)
CvWW:LabvFieldL	-0.63
	(0.94)
MoA.CODEGW:LabvFieldi	-0.54
	(0.91)
MoA.CODEACI:LabvFieldL	-7.03 *
	(3.21)
MoA.CODECMD:LabvFieldL	-7.82 **
	(2.68)
MoA.CODEGW:LabvFieldL	-0.28
	(1.03)
MoA.CODELSI:LabvFieldL	-4.57
	(3.15)
MoA.CODEOTH:LabvFieldL	1.23
	(1.94)
MoA.CODEPHI:LabvFieldL	-6.78 ***
	(2.01)
MoA.CODESGI:LabvFieldL	0.68
	(1.58)
CvWW:Class.MvDM	5.84 **
	(1.86)
MoA.CODEACI:Class.MvDM	-2.83 *
	(1.22)
MoA.CODECMD:Class.MvDM	-2.26 *
	(1.07)
MoA.CODEGW:Class.MvDM	-1.43
	(1.66)
MoA.CODEICD:Class.MvDM	-1.88
	(1.34)
MoA.CODELSI:Class.MvDM	-1.28
	(0.95)
MoA.CODEOTH:Class.MvDM	-3.02
	(2.04)
MoA.CODEPHI:Class.MvDM	-4.10 ***
	(1.11)
MoA.CODESGI:Class.MvDM	0.18
	(0.68)
LabvFieldL:Class.MvDGymn	3.87
	(2.07)
LabvFieldi:Class.MvDM	-1.47
	(1.78)
LabvFieldL:Class.MvDM	2.92
	(2.03)
CvWW:MoA.CODEACI:LabvFieldL	11.82 ***
	(3.22)
CvWW:MoA.CODECMD:LabvFieldL	5.82 *
	(2.70)
CvWW:MoA.CODEGW:LabvFieldL	-2.05
	(1.22)
CvWW:MoA.CODELSI:LabvFieldL	5.97
	(3.33)
CvWW:MoA.CODEPHI:LabvFieldL	3.07

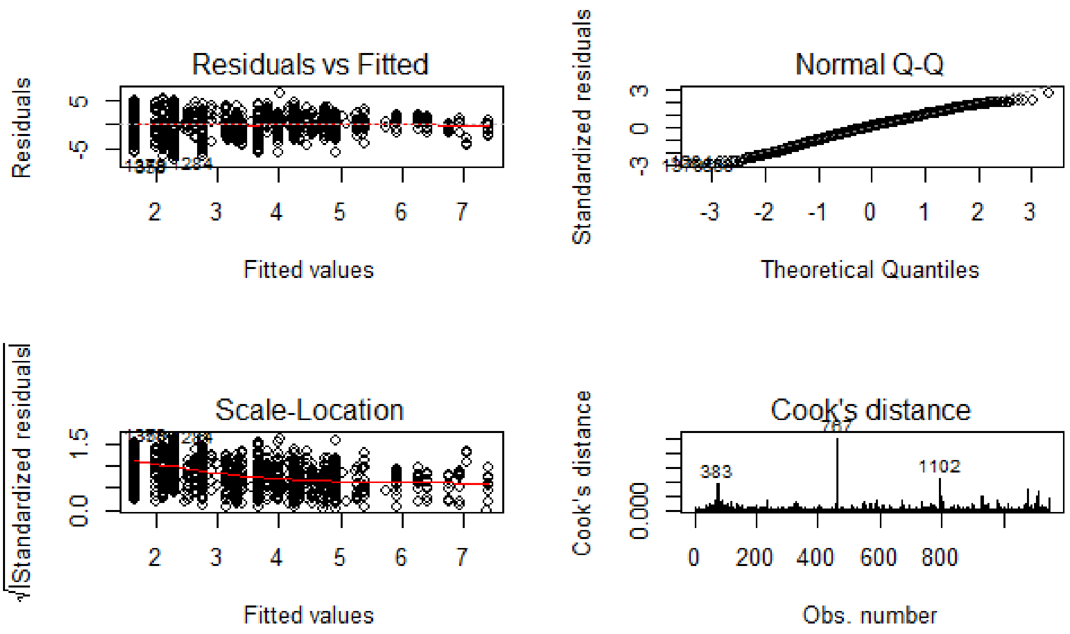
```
                                     (2.10)
CvWW:LabvFieldL:Class.MvDM           -6.21 **
                                     (1.89)
MoA.CODECMD:LabvFieldL:Class.MvDM     2.97 *
                                     (1.18)
MoA.CODEGW:LabvFieldL:Class.MvDM      3.29
                                     (1.73)
MoA.CODEPHI:LabvFieldL:Class.MvDM     5.01 ***
                                     (1.22)
-------------------------------------------------------------------
R^2                                   0.40         0.32
Adj. R^2                              0.38         0.31
Num. obs.                          1281         1281
RMSE                                  2.18         2.29
===================================================================
*** p < 0.001, ** p < 0.01, * p < 0.05
```

4. Model Validation

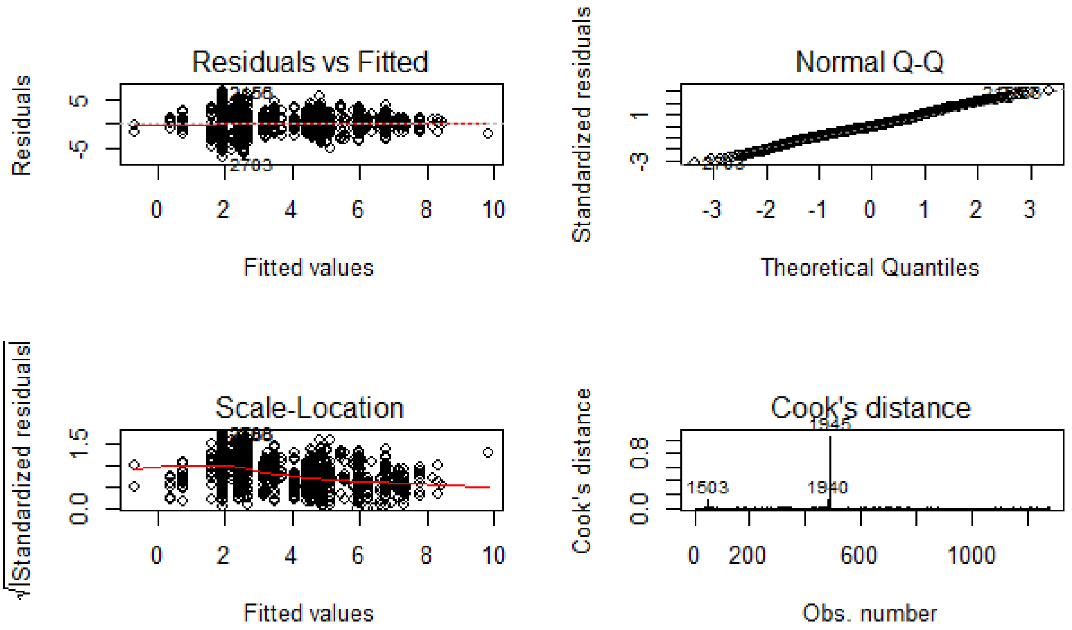

Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Cook's distance
Cook's distance
Obs. number

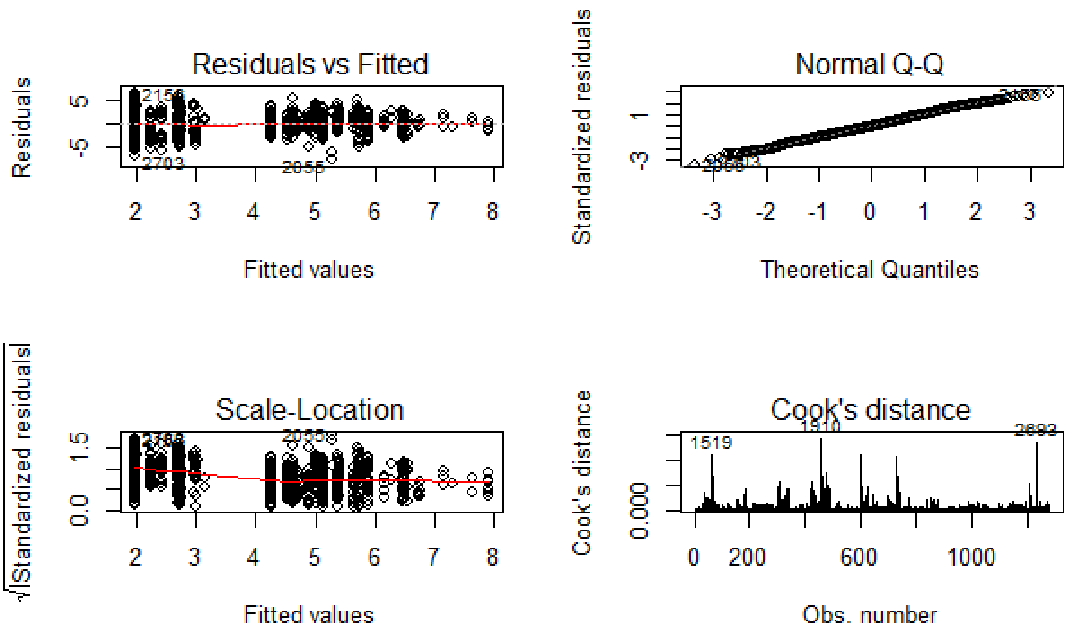

Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Cook's distance
Cook's distance
Obs. number






```
Warning: not plotting observations with leverage one:
  490, 649
```





5. Interpretation/Model for all effect levels (ER10, ER25 and ER50) pooled.

As difference between crop species and wild species were identified significantly positive based on ER25 but negative on ER10 or on ER50, we took another step to pool together the different effect levels (ER10 - ER50) data, assuming a common difference between crop species and wild species.

Some of the ER10, ER25 or ER50 for the same substance derived from the same experimental dose-response curve. Such cases are not independent but may be related by formulas such as

$$ER25 = {\frac{25}{100-25}}^{1/H} \cdot ER50$$

where, H is the hill slope of the dose-response curve. We assume the dose response curve can be described by a four parameter logistic equation here.

NOTE HC: However, it must be taken into account that this affects only a fraction of data. For many experiments just one effect level is reported, e.g. the ER25 was reported in a study submitted to the US-EPA, an ER50 in a study generated for the European notification procedure, and an ER10 generated by an independent scientist. For some experiments endpoints were available for two or even all three effect levels, only these are not independent.

If all different effect levels were dependent, then theoretically for wild species, although the average ER50 could be lower than that of the crop species, the ER25 could be significantly higher also due to the different hill slopes of the dose-response curves.

Pooled analysis is thus not completely appropriate since we do not have independent data throughout. Ideally we should consider which ER10, ER25 or ER50 are pairs/triplets to model accordingly, but this was not implemented for time and budget constraints.

In the table below the first term each is considered as the baseline, i.e. in case of the effect levels ERx the ER10. The deviations between that baseline and the other effect levels are listed in the table below (SD in brackets).

For interpretation note that the ERx values had been logarithmized, so to assess the actual difference between baseline and variable assessed it has to be backtransformed. The coefficient of 0.56 for ER25 (Pooled best) indicates that ER25 were a factor of 1.75 ($e^{0.56}$) higher than the ER10 in this database.

	Pooled best	Pooled no interaction
(Intercept)	0.04	1.15 ***
	(0.38)	(0.19)
Effect.level.ERxER25	0.56 ***	0.68 ***
	(0.16)	(0.16)
Effect.level.ERxER50	1.07 ***	1.01 ***
	(0.15)	(0.16)
CvWW	1.28 **	0.32 **
	(0.45)	(0.10)
MoA.CODEACI	5.15 *	0.38
	(2.52)	(0.28)
MoA.CODECMD	2.67	2.58 ***
	(1.56)	(0.19)
MoA.CODEGW	1.98 ***	1.93 ***
	(0.44)	(0.13)
MoA.CODEICD	6.23 ***	5.23 ***
	(0.96)	(0.71)
MoA.CODELSI	5.24	3.18 ***

	(3.49)	(0.32)
MoA.CODEOTH	3.21 ***	1.39 ***
	(0.65)	(0.33)
MoA.CODEPHI	4.98 ***	2.86 ***
	(1.00)	(0.15)
MoA.CODESGI	6.24 ***	3.93 ***
	(0.70)	(0.21)
LabvFieldi	-0.50	-0.33
	(0.48)	(0.32)
LabvFieldL	0.93 *	-0.15
	(0.38)	(0.13)
Class.MvDGymn	0.50	2.49 *
	(1.40)	(1.10)
Class.MvDM	2.53 ***	0.66 ***
	(0.52)	(0.10)
CvWW:MoA.CODEACI	-6.23 *	
	(2.46)	
CvWW:MoA.CODECMD	0.30	
	(1.57)	
CvWW:MoA.CODEGW	0.30	
	(0.56)	
CvWW:MoA.CODELSI	-0.07	
	(2.61)	
CvWW:MoA.CODEOTH	-2.48 **	
	(0.78)	
CvWW:MoA.CODEPHI	-0.53	
	(1.05)	
CvWW:MoA.CODESGI	-1.33	
	(0.88)	
CvWW:LabvFieldL	-0.38	
	(0.48)	
MoA.CODEGW:LabvFieldi	0.13	
	(0.72)	
MoA.CODEACI:LabvFieldL	-0.74	
	(2.55)	
MoA.CODECMD:LabvFieldL	0.13	
	(1.61)	
MoA.CODEGW:LabvFieldL	0.07	
	(0.49)	
MoA.CODELSI:LabvFieldL	-1.35	
	(3.53)	
MoA.CODEOTH:LabvFieldL	3.74 **	
	(1.31)	
MoA.CODEPHI:LabvFieldL	-2.43 *	
	(1.02)	
MoA.CODESGI:LabvFieldL	-1.25	
	(0.78)	
CvWW:Class.MvDM	-0.52 *	
	(0.22)	
MoA.CODEACI:Class.MvDM	-2.55 ***	
	(0.70)	
MoA.CODECMD:Class.MvDM	-0.94	

```
                                      (0.83)
MoA.CODEGW:Class.MvDM                 -0.35
                                      (0.62)
MoA.CODEICD:Class.MvDM                -1.71
                                      (1.36)
MoA.CODELSI:Class.MvDM                -2.79
                                      (2.61)
MoA.CODEOTH:Class.MvDM                -1.26
                                      (0.79)
MoA.CODEPHI:Class.MvDM                -2.53 **
                                      (0.85)
MoA.CODESGI:Class.MvDM                -2.83 **
                                      (0.89)
LabvFieldL:Class.MvDGymn               2.94
                                      (2.18)
LabvFieldi:Class.MvDM                  0.42
                                      (1.32)
LabvFieldL:Class.MvDM                 -2.15 ***
                                      (0.52)
CvWW:MoA.CODEACI:LabvFieldL            4.64
                                      (2.61)
CvWW:MoA.CODECMD:LabvFieldL           -1.99
                                      (1.64)
CvWW:MoA.CODEGW:LabvFieldL            -1.72 **
                                      (0.64)
CvWW:MoA.CODELSI:LabvFieldL           -0.55
                                      (2.74)
CvWW:MoA.CODEPHI:LabvFieldL            0.03
                                      (1.10)
CvWW:MoA.CODESGI:LabvFieldL           -1.67
                                      (1.00)
MoA.CODECMD:LabvFieldL:Class.MvDM      1.96 *
                                      (0.96)
MoA.CODEGW:LabvFieldL:Class.MvDM       2.16 **
                                      (0.71)
MoA.CODELSI:LabvFieldL:Class.MvDM      2.11
                                      (2.70)
MoA.CODEPHI:LabvFieldL:Class.MvDM      3.67 ***
                                      (0.92)
MoA.CODESGI:LabvFieldL:Class.MvDM      2.58 *
                                      (1.02)
---------------------------------------------------------------------
R^2                                    0.33         0.27
Adj. R^2                               0.32         0.27
Num. obs.                           2737         2737
RMSE                                   2.34         2.42
=====================================================================
*** p < 0.001, ** p < 0.01, * p < 0.05
```

```
==================================================================
                      ER10         ER25          ER50          Pooled
------------------------------------------------------------------
(Intercept)             0.45         2.00 ***      2.25 ***      1.15 ***
                       (0.50)       (0.23)        (0.22)        (0.19)
CvWW                   -0.37         0.65 ***     -0.01          0.32 **
                       (0.43)       (0.15)        (0.14)        (0.10)
MoA.CODEACI             5.27 ***     0.62          0.03          0.38
                       (1.50)       (0.44)        (0.35)        (0.28)
MoA.CODECMD             5.59 ***     1.90 ***      2.65 ***      2.58 ***
                       (0.89)       (0.45)        (0.21)        (0.19)
MoA.CODEGW              2.23 ***     1.68 ***      2.29 ***      1.93 ***
                       (0.45)       (0.20)        (0.18)        (0.13)
MoA.CODEOTH             4.94 *       0.53          4.31 ***      1.39 ***
                       (2.07)       (0.38)        (0.78)        (0.33)
MoA.CODEPHI             2.74 ***     2.62 ***      3.16 ***      2.86 ***
                       (0.69)       (0.21)        (0.21)        (0.15)
MoA.CODESGI             3.38 ***     4.28 ***      3.76 ***      3.93 ***
                       (0.68)       (0.31)        (0.31)        (0.21)
LabvFieldi              0.17                       0.18         -0.33
                       (0.73)                     (0.39)        (0.32)
LabvFieldL              0.55        -0.35         -0.27         -0.15
                       (0.50)       (0.19)        (0.19)        (0.13)
Class.MvDM              1.38 ***     0.48 **       0.74 ***      0.66 ***
                       (0.40)       (0.16)        (0.15)        (0.10)
MoA.CODEICD                          5.75 *        5.18 ***      5.23 ***
                                    (2.38)        (0.70)        (0.71)
MoA.CODELSI                          2.60 ***      3.74 ***      3.18 ***
                                    (0.48)        (0.40)        (0.32)
Class.MvDGymn                                      2.48 *        2.49 *
                                                  (1.04)        (1.10)
Effect.level.ERxER25                                             0.68 ***
                                                                (0.16)
Effect.level.ERxER50                                             1.01 ***
                                                                (0.16)
------------------------------------------------------------------
R^2                     0.26         0.24          0.32          0.27
Adj. R^2                0.23         0.24          0.31          0.27
Num. obs.             308         1148          1281          2737
RMSE                    2.85         2.38          2.29          2.42
==================================================================
*** p < 0.001, ** p < 0.01, * p < 0.05
```

<u>Below: Modified, crops multiplied by Factor 1.5. Only pooled</u>

```
> screenreg(list(mod3,mod3.0),custom.model.names = c("Pooled
best","Pooled no interaction"))

==========================================================================
=
```

```
                                    Pooled best  Pooled no
interaction
--------------------------------------------------------------------
-
(Intercept)                            0.44         1.56 ***
                                      (0.38)       (0.19)
Effect.level.ERxER25                   0.56 ***     0.68 ***
                                      (0.16)       (0.16)
Effect.level.ERxER50                   1.07 ***     1.01 ***
                                      (0.15)       (0.16)
CvWW                                   0.88        -0.09
                                      (0.45)       (0.10)
MoA.CODEACI                            5.15 *       0.38
                                      (2.52)       (0.28)
MoA.CODECMD                            2.67         2.58 ***
                                      (1.56)       (0.19)
MoA.CODEGW                             1.98 ***     1.93 ***
                                      (0.44)       (0.13)
MoA.CODEICD                            6.23 ***     5.23 ***
                                      (0.96)       (0.71)
MoA.CODELSI                            5.24         3.18 ***
                                      (3.49)       (0.32)
MoA.CODEOTH                            3.21 ***     1.39 ***
                                      (0.65)       (0.33)
MoA.CODEPHI                            4.98 ***     2.86 ***
                                      (1.00)       (0.15)
MoA.CODESGI                            6.24 ***     3.93 ***
                                      (0.70)       (0.21)
LabvFieldi                            -0.50        -0.33
                                      (0.48)       (0.32)
LabvFieldL                             0.93 *      -0.15
                                      (0.38)       (0.13)
Class.MvDGymn                          0.50         2.49 *
                                      (1.40)       (1.10)
Class.MvDM                             2.53 ***     0.66 ***
                                      (0.52)       (0.10)
CvWW:MoA.CODEACI                      -6.23 *
                                      (2.46)
CvWW:MoA.CODECMD                       0.30
                                      (1.57)
CvWW:MoA.CODEGW                        0.30
                                      (0.56)
CvWW:MoA.CODELSI                      -0.07
                                      (2.61)
CvWW:MoA.CODEOTH                      -2.48 **
                                      (0.78)
CvWW:MoA.CODEPHI                      -0.53
                                      (1.05)
```

```
CvWW:MoA.CODESGI                     -1.33
                                     (0.88)
CvWW:LabvFieldL                      -0.38
                                     (0.48)
MoA.CODEGW:LabvFieldi                 0.13
                                     (0.72)
MoA.CODEACI:LabvFieldL               -0.74
                                     (2.55)
MoA.CODECMD:LabvFieldL                0.13
                                     (1.61)
MoA.CODEGW:LabvFieldL                 0.07
                                     (0.49)
MoA.CODELSI:LabvFieldL               -1.35
                                     (3.53)
MoA.CODEOTH:LabvFieldL                3.74 **
                                     (1.31)
MoA.CODEPHI:LabvFieldL               -2.43 *
                                     (1.02)
MoA.CODESGI:LabvFieldL               -1.25
                                     (0.78)
CvWW:Class.MvDM                      -0.52 *
                                     (0.22)
MoA.CODEACI:Class.MvDM               -2.55 ***
                                     (0.70)
MoA.CODECMD:Class.MvDM               -0.94
                                     (0.83)
MoA.CODEGW:Class.MvDM                -0.35
                                     (0.62)
MoA.CODEICD:Class.MvDM               -1.71
                                     (1.36)
MoA.CODELSI:Class.MvDM               -2.79
                                     (2.61)
MoA.CODEOTH:Class.MvDM               -1.26
                                     (0.79)
MoA.CODEPHI:Class.MvDM               -2.53 **
                                     (0.85)
MoA.CODESGI:Class.MvDM               -2.83 **
                                     (0.89)
LabvFieldL:Class.MvDGymn              2.94
                                     (2.18)
LabvFieldi:Class.MvDM                 0.42
                                     (1.32)
LabvFieldL:Class.MvDM                -2.15 ***
                                     (0.52)
CvWW:MoA.CODEACI:LabvFieldL           4.64
                                     (2.61)
CvWW:MoA.CODECMD:LabvFieldL          -1.99
                                     (1.64)
```

```
CvWW:MoA.CODEGW:LabvFieldL                  -1.72 **
                                            (0.64)
CvWW:MoA.CODELSI:LabvFieldL                 -0.55
                                            (2.74)
CvWW:MoA.CODEPHI:LabvFieldL                  0.03
                                            (1.10)
CvWW:MoA.CODESGI:LabvFieldL                 -1.67
                                            (1.00)
MoA.CODECMD:LabvFieldL:Class.MvDM            1.96 *
                                            (0.96)
MoA.CODEGW:LabvFieldL:Class.MvDM             2.16 **
                                            (0.71)
MoA.CODELSI:LabvFieldL:Class.MvDM            2.11
                                            (2.70)
MoA.CODEPHI:LabvFieldL:Class.MvDM            3.67 ***
                                            (0.92)
MoA.CODESGI:LabvFieldL:Class.MvDM            2.58 *
                                            (1.02)
--------------------------------------------------------------------
-
R^2                                          0.33          0.27
Adj. R^2                                     0.32          0.27
Num. obs.                                 2737          2737
RMSE                                         2.34          2.42
====================================================================
=
*** p < 0.001, ** p < 0.01, * p < 0.05
```

Below: Modified, crops multiplied by Factor 2. Only pooled

```
> screenreg(list(mod3,mod3.0),custom.model.names = c("Pooled
best","Pooled no interaction"))

====================================================================
=
                                          Pooled best   Pooled no
interaction
--------------------------------------------------------------------
-
(Intercept)                                  0.73          1.85 ***
                                            (0.38)        (0.19)
Effect.level.ERxER25                         0.56 ***      0.68 ***
                                            (0.16)        (0.16)
Effect.level.ERxER50                         1.07 ***      1.01 ***
                                            (0.15)        (0.16)
CvWW                                         0.59         -0.38 ***
                                            (0.45)        (0.10)
MoA.CODEACI                                  5.15 *        0.38
```

	(2.52)	(0.28)
MoA.CODECMD	2.67	2.58 ***
	(1.56)	(0.19)
MoA.CODEGW	1.98 ***	1.93 ***
	(0.44)	(0.13)
MoA.CODEICD	6.23 ***	5.23 ***
	(0.96)	(0.71)
MoA.CODELSI	5.24	3.18 ***
	(3.49)	(0.32)
MoA.CODEOTH	3.21 ***	1.39 ***
	(0.65)	(0.33)
MoA.CODEPHI	4.98 ***	2.86 ***
	(1.00)	(0.15)
MoA.CODESGI	6.24 ***	3.93 ***
	(0.70)	(0.21)
LabvFieldi	-0.50	-0.33
	(0.48)	(0.32)
LabvFieldL	0.93 *	-0.15
	(0.38)	(0.13)
Class.MvDGymn	0.50	2.49 *
	(1.40)	(1.10)
Class.MvDM	2.53 ***	0.66 ***
	(0.52)	(0.10)
CvWW:MoA.CODEACI	-6.23 *	
	(2.46)	
CvWW:MoA.CODECMD	0.30	
	(1.57)	
CvWW:MoA.CODEGW	0.30	
	(0.56)	
CvWW:MoA.CODELSI	-0.07	
	(2.61)	
CvWW:MoA.CODEOTH	-2.48 **	
	(0.78)	
CvWW:MoA.CODEPHI	-0.53	
	(1.05)	
CvWW:MoA.CODESGI	-1.33	
	(0.88)	
CvWW:LabvFieldL	-0.38	
	(0.48)	
MoA.CODEGW:LabvFieldi	0.13	
	(0.72)	
MoA.CODEACI:LabvFieldL	-0.74	
	(2.55)	
MoA.CODECMD:LabvFieldL	0.13	
	(1.61)	
MoA.CODEGW:LabvFieldL	0.07	
	(0.49)	
MoA.CODELSI:LabvFieldL	-1.35	

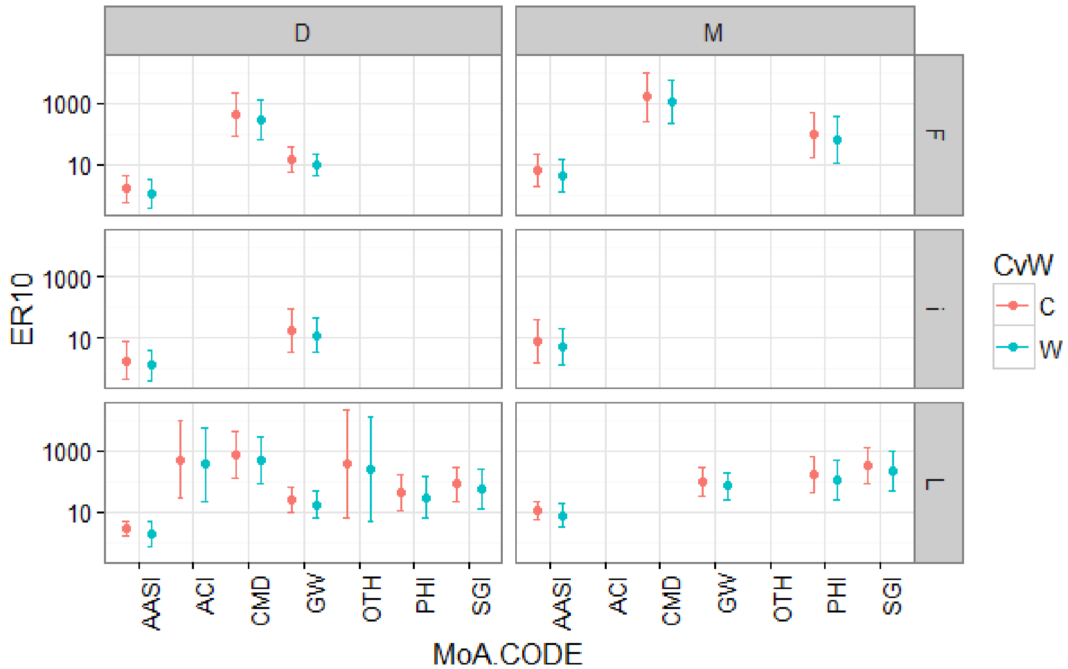
```
                                        (3.53)
MoA.CODEOTH:LabvFieldL                   3.74 **
                                        (1.31)
MoA.CODEPHI:LabvFieldL                  -2.43 *
                                        (1.02)
MoA.CODESGI:LabvFieldL                  -1.25
                                        (0.78)
CvWW:Class.MvDM                         -0.52 *
                                        (0.22)
MoA.CODEACI:Class.MvDM                  -2.55 ***
                                        (0.70)
MoA.CODECMD:Class.MvDM                  -0.94
                                        (0.83)
MoA.CODEGW:Class.MvDM                   -0.35
                                        (0.62)
MoA.CODEICD:Class.MvDM                  -1.71
                                        (1.36)
MoA.CODELSI:Class.MvDM                  -2.79
                                        (2.61)
MoA.CODEOTH:Class.MvDM                  -1.26
                                        (0.79)
MoA.CODEPHI:Class.MvDM                  -2.53 **
                                        (0.85)
MoA.CODESGI:Class.MvDM                  -2.83 **
                                        (0.89)
LabvFieldL:Class.MvDGymn                 2.94
                                        (2.18)
LabvFieldi:Class.MvDM                    0.42
                                        (1.32)
LabvFieldL:Class.MvDM                   -2.15 ***
                                        (0.52)
CvWW:MoA.CODEACI:LabvFieldL              4.64
                                        (2.61)
CvWW:MoA.CODECMD:LabvFieldL             -1.99
                                        (1.64)
CvWW:MoA.CODEGW:LabvFieldL              -1.72 **
                                        (0.64)
CvWW:MoA.CODELSI:LabvFieldL             -0.55
                                        (2.74)
CvWW:MoA.CODEPHI:LabvFieldL              0.03
                                        (1.10)
CvWW:MoA.CODESGI:LabvFieldL             -1.67
                                        (1.00)
MoA.CODECMD:LabvFieldL:Class.MvDM        1.96 *
                                        (0.96)
MoA.CODEGW:LabvFieldL:Class.MvDM         2.16 **
                                        (0.71)
MoA.CODELSI:LabvFieldL:Class.MvDM        2.11
```

```
                                          (2.70)
MoA.CODEPHI:LabvFieldL:Class.MvDM          3.67 ***
                                          (0.92)
MoA.CODESGI:LabvFieldL:Class.MvDM          2.58 *
                                          (1.02)
--------------------------------------------------------------------
-
R^2                                        0.34         0.27
Adj. R^2                                   0.32         0.27
Num. obs.                               2737         2737
RMSE                                       2.34         2.42
====================================================================
=
*** p < 0.001, ** p < 0.01, * p < 0.05
```

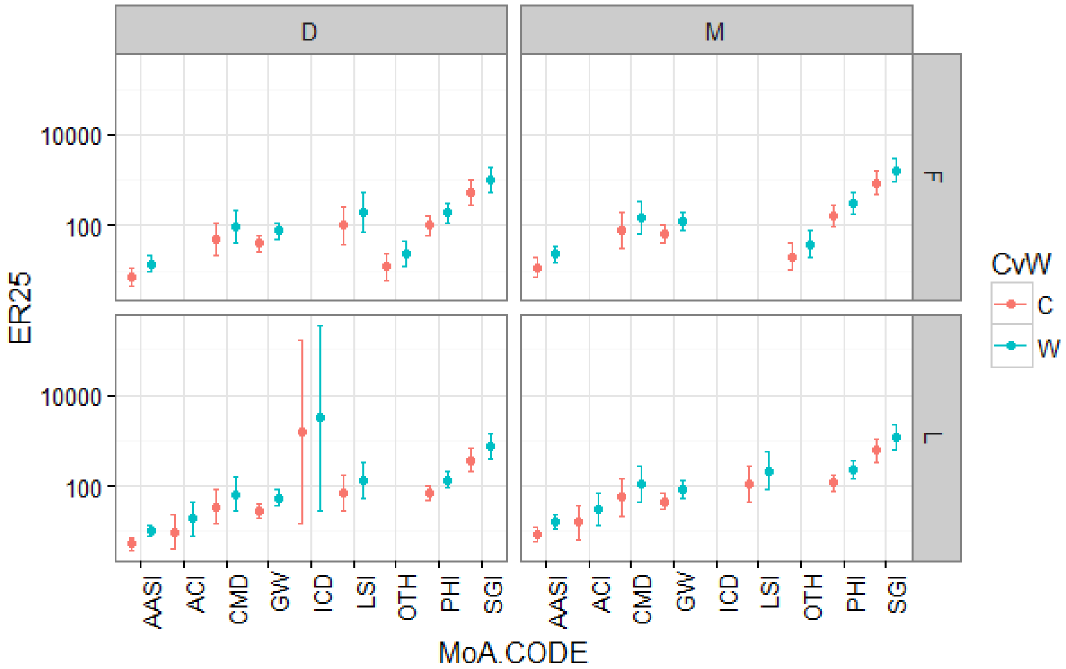
(End Section with modified data)

MDD Calculation

ER10



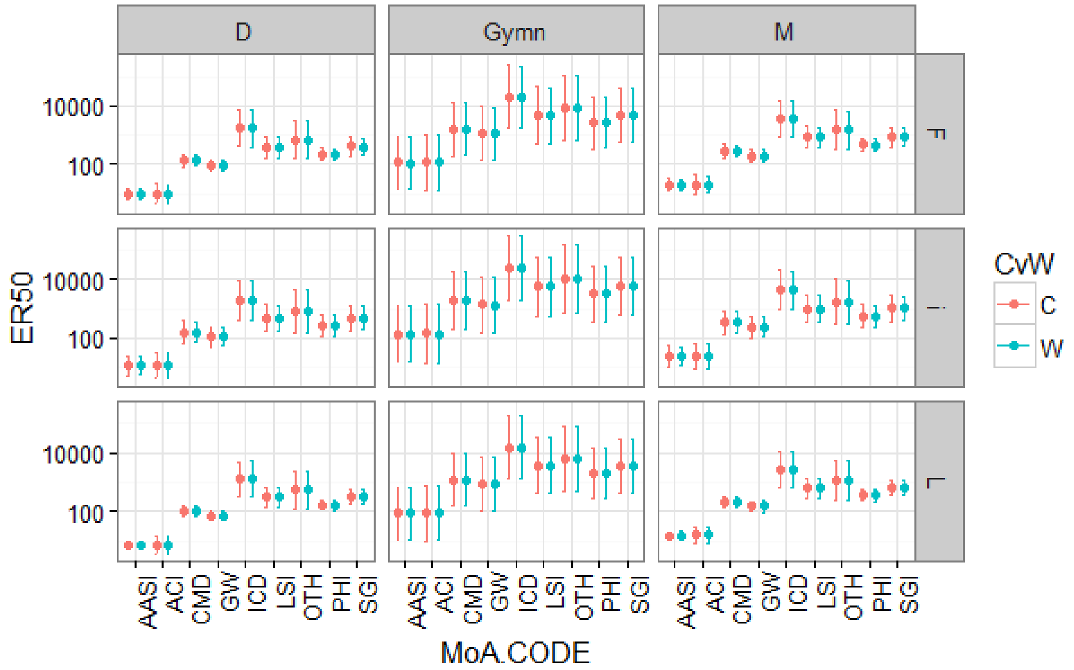
ER25

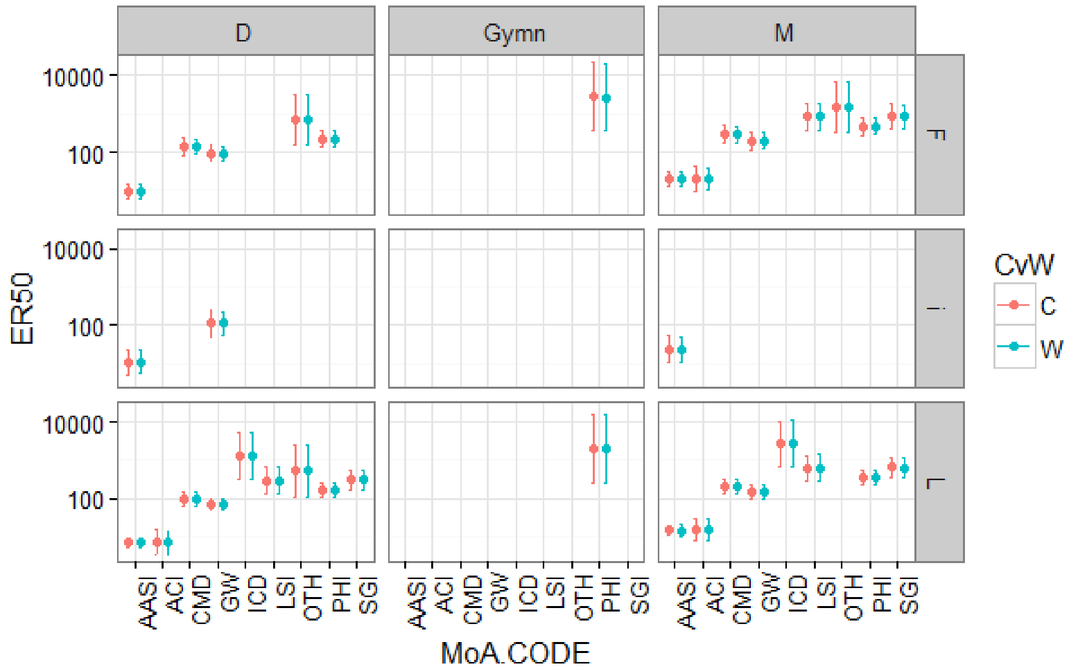


MoA.CODE	LabvField	Class.MvD	MDD1	MDD%1	MDD2	MDD%2	MDD3	MDD%3	MDD4	MDD%4	Crop	Wild
AASI	L	M	3.207	38.29	5.197	62.06	6.142	38.29	9.954	62.06	8.374	16.04
AASI	L	D	1.682	32.49	2.492	48.12	3.222	32.49	4.772	48.12	5.178	9.917
ACI	L	M	10.92	70.02	36.42	233.5	20.91	70.02	69.75	233.5	15.6	29.87
ACI	L	D	6.842	70.96	23.56	244.4	13.11	70.96	45.13	244.4	9.642	18.47
GW	F	D	16.97	42.68	29.61	74.46	32.51	42.68	56.71	74.46	39.76	76.16
CMD	F	D	33.74	68.33	106.6	215.8	64.63	68.33	204.1	215.8	49.38	94.58
GW	L	D	11.67	41.85	20.08	71.98	22.36	41.85	38.46	71.98	27.89	53.42
CMD	F	M	56.3	70.5	190.8	238.9	107.8	70.5	365.5	238.9	79.86	153
AASI	F	M	5.744	48.11	11.07	92.73	11	48.11	21.2	92.73	11.94	22.87
PHI	F	M	88.47	54	192.4	117.4	169.5	54	368.4	117.4	163.8	313.8
CMD	L	D	24.63	71.1	85.2	246	47.17	71.1	163.2	246	34.64	66.34
AASI	F	D	3.356	45.47	6.154	83.37	6.428	45.47	11.79	83.37	7.382	14.14
PHI	L	D	29.61	41.67	50.77	71.45	56.71	41.67	97.24	71.45	71.05	136.1
PHI	L	M	53.55	46.6	100.3	87.28	102.6	46.6	192.1	87.28	114.9	220.1
GW	L	M	20.68	45.83	38.17	84.61	39.6	45.83	73.11	84.61	45.11	86.41
PHI	F	D	51.91	51.25	106.5	105.1	99.42	51.25	203.9	105.1	101.3	194
CMD	L	M	40.95	73.1	152.2	271.8	78.44	73.1	291.6	271.8	56.02	107.3
ICD	L	D	1617	99.86	1190929	73537	3098	99.86	2281068	73537	1619	3102
GW	F	M	29.47	45.82	54.39	84.57	56.44	45.82	104.2	84.57	64.31	123.2
LSI	F	D	75.02	75.5	306.2	308.1	143.7	75.5	586.4	308.1	99.37	190.3
OTH	F	M	12.6	62.06	33.2	163.6	24.12	62.06	63.59	163.6	20.29	38.87
OTH	F	D	7.732	61.62	20.15	160.6	14.81	61.62	38.59	160.6	12.55	24.03
SGI	F	D	308.7	57.68	729.6	136.3	591.4	57.68	1397	136.3	535.3	1025
SGI	F	M	501.6	57.94	1193	137.8	960.8	57.94	2285	137.8	865.7	1658
LSI	L	D	51	73.17	190.1	272.7	97.68	73.17	364	272.7	69.7	133.5

LSI	L	M	82.22	72.94	303.8	269.5	157.5	72.94	581.9	269.5	112.7	215.9
SGI	L	D	217.9	58.02	519	138.2	417.3	58.02	994	138.2	375.5	719.2
SGI	L	M	356.2	58.66	861.6	141.9	682.3	58.66	1650	141.9	607.3	1163

ER50

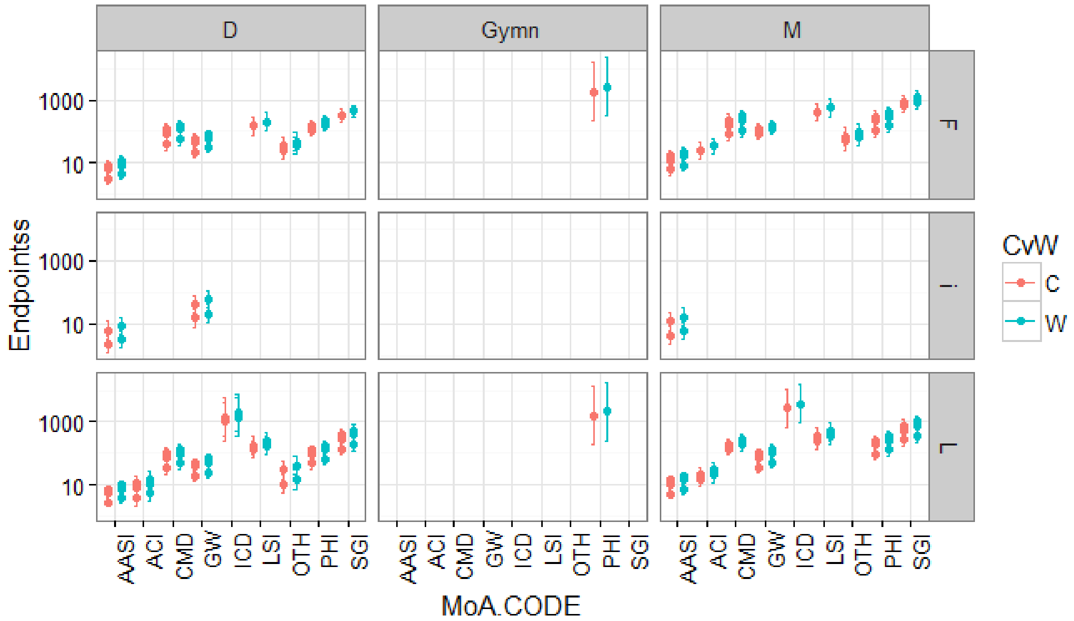


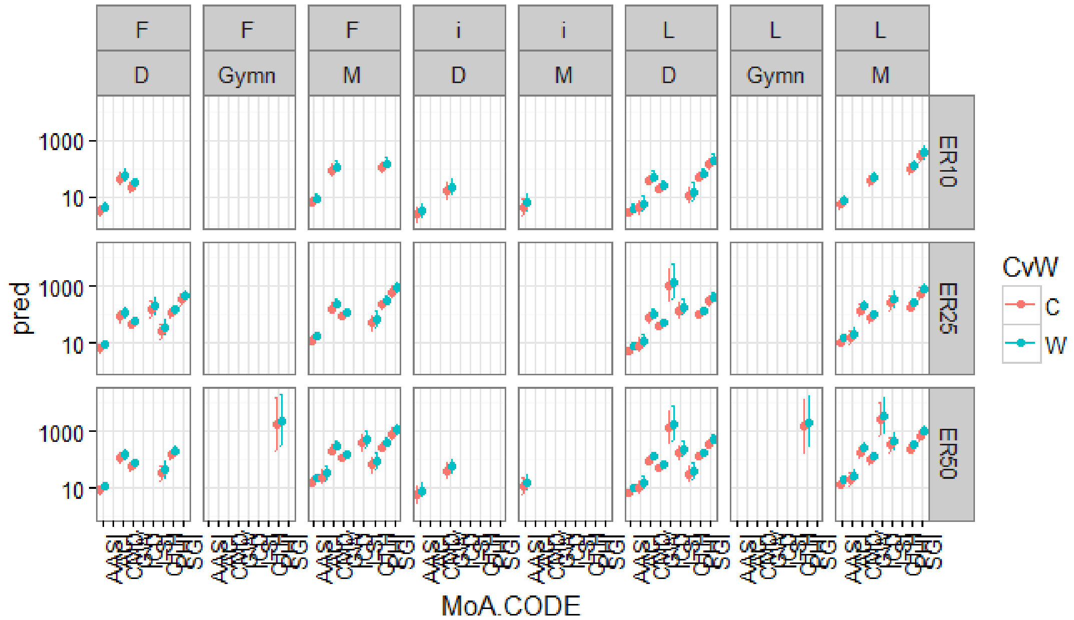


MoA.CODE	LabvField	Class.MvD	MDD1	MDD% 1	MDD2	MDD% 2	MDD3	MDD% 3	MDD4	MDD% 4	Crop	Wild
CMD	L	M	100.4	46.58	188	87.21	99.6	46.58	186.5	87.21	215.6	213.8
CMD	L	D	45.63	44.4	82.06	79.85	45.25	44.4	81.38	79.85	102.8	101.9
GW	F	D	43.93	46.6	82.27	87.27	43.57	46.6	81.59	87.27	94.28	93.49
CMD	F	D	66.78	49.59	132.5	98.39	66.23	49.59	131.4	98.39	134.7	133.5

GW	L	D	27.67	38.46	44.97	62.5	27.44	38.46	44.6	62.5	71.95	71.36
CMD	F	M	145.4	51.46	299.4	106	144.2	51.46	297	106	282.5	280.1
AASI	F	M	9.495	47.59	18.12	90.8	9.416	47.59	17.96	90.8	19.95	19.79
PHI	F	M	244.8	52.17	511.7	109.1	242.7	52.17	507.5	109.1	469.2	465.3
AASI	L	D	2.249	30.98	3.259	44.9	2.23	30.98	3.232	44.9	7.259	7.198
AASI	i	D	7.284	64.17	20.33	179.1	7.223	64.17	20.16	179.1	11.35	11.26
GW	i	D	73.42	65.25	211.3	187.8	72.81	65.25	209.5	187.8	112.5	111.6
AASI	L	M	5.578	36.63	8.802	57.8	5.531	36.63	8.729	57.8	15.23	15.1
AASI	F	D	4.213	44.3	7.564	79.54	4.178	44.3	7.502	79.54	9.51	9.431
GW	L	M	67.2	44.52	121.1	80.24	66.64	44.52	120.1	80.24	151	149.7
AASI	i	M	15.82	66.42	47.09	197.8	15.68	66.42	46.7	197.8	23.81	23.61
PHI	L	D	73.34	42.97	128.6	75.35	72.73	42.97	127.5	75.35	170.7	169.3
PHI	L	M	167.4	46.75	314.4	87.81	166	46.75	311.8	87.81	358.1	355.1
SGI	L	D	175	56.35	400.9	129.1	173.5	56.35	397.6	129.1	310.5	308
SGI	L	M	378.4	58.08	902.5	138.5	375.2	58.08	895	138.5	651.5	646.1
LSI	L	M	423.9	66.39	1261	197.6	420.4	66.39	1251	197.6	638.5	633.1
LSI	L	D	205.5	67.52	632.7	207.9	203.8	67.52	627.5	207.9	304.3	301.8
ICD	L	M	2314	85.7	16175	599.1	2294	85.7	16040	599.1	2700	2677
ICD	L	D	1103	85.74	7740	601.5	1094	85.74	7676	601.5	1287	1276
OTH	L	D	476.7	88.5	4146	769.9	472.7	88.5	4112	769.9	538.6	534.1
PHI	L	Gymn	1927	94.3	33786	1653	1911	94.3	33505	1653	2044	2027
ACI	L	M	9.444	60.25	23.76	151.6	9.366	60.25	23.56	151.6	15.67	15.54
PHI	F	D	109.9	49.16	216.3	96.71	109	49.16	214.5	96.71	223.6	221.8
OTH	F	D	622.6	88.23	5289	749.6	617.4	88.23	5245	749.6	705.7	699.8
LSI	F	M	575	68.73	1839	219.8	570.2	68.73	1824	219.8	836.5	829.6
SGI	F	M	543.2	63.64	1494	175	538.7	63.64	1481	175	853.6	846.5
PHI	F	Gymn	2526	94.31	44369	1657	2505	94.31	44000	1657	2678	2656
GW	F	M	100.9	51.01	206	104.1	100.1	51.01	204.2	104.1	197.8	196.1
ACI	F	M	12.97	63.14	35.18	171.3	12.86	63.14	34.89	171.3	20.54	20.37
OTH	F	M	1310	88.48	11374	768.3	1299	88.48	11279	768.3	1480	1468
ACI	L	D	4.721	63.18	12.82	171.6	4.682	63.18	12.72	171.6	7.472	7.409

Pooled





Effect. level. ERx	MoA.CODE	LabvField	Class. MvD	MDD1	MDD %1	MDD2	MDD %2	MDD3	MDD %3	MDD4	MDD %4	Crop	Wild
ER10	CMD	L	D	17.35	48.1	33.43	92.7	23.8	48.1	45.85	92.7	36.07	49.47
ER10	GW	F	D	8.876	40.79	14.99	68.88	12.17	40.79	20.56	68.88	21.76	29.84
ER10	CMD	F	D	20	47.93	38.41	92.06	27.43	47.93	52.68	92.06	41.73	57.22
ER10	GW	L	D	7.517	39.96	12.52	66.56	10.31	39.96	17.17	66.56	18.81	25.8
ER10	CMD	F	M	39.83	49.34	78.62	97.41	54.62	49.34	107.8	97.41	80.71	110.7
ER10	AASI	F	M	2.61	42.63	4.549	74.32	3.579	42.63	6.239	74.32	6.121	8.395
ER10	PHI	F	M	50.89	47.64	97.19	91	69.79	47.64	133.3	91	106.8	146.5
ER10	AASI	L	D	0.9864	36.06	1.543	56.4	1.353	36.06	2.116	56.4	2.735	3.752
ER10	AASI	i	D	1.35	59.18	3.308	145	1.852	59.18	4.536	145	2.281	3.129

ER10	GW	i	D	9.396	59.9	23.43	149.3	12.89	59.9	32.13	149.3	15.69	21.51
ER10	AASI	F	D	1.291	40.78	2.179	68.87	1.77	40.78	2.989	68.87	3.165	4.34
ER10	AASI	L	M	2.026	38.29	3.284	62.06	2.779	38.29	4.503	62.06	5.291	7.256
ER10	AASI	i	M	2.678	60.69	6.812	154.4	3.673	60.69	9.342	154.4	4.413	6.052
ER10	GW	L	M	15.47	42.52	26.92	73.98	21.22	42.52	36.92	73.98	36.39	49.9
ER10	SGI	L	D	69.13	49.45	136.8	97.83	94.8	49.45	187.5	97.83	139.8	191.7
ER10	SGI	L	M	135.7	50.19	272.4	100.8	186.1	50.19	373.6	100.8	270.4	370.8
ER10	PHI	L	D	20.4	42.73	35.62	74.62	27.97	42.73	48.85	74.62	47.73	65.46
ER10	PHI	L	M	41.33	44.76	74.83	81.04	56.68	44.76	102.6	81.04	92.33	126.6
ER10	ACI	L	D	2.324	58.02	5.536	138.2	3.187	58.02	7.593	138.2	4.006	5.493
ER10	OTH	L	D	7.004	63.74	19.32	175.8	9.606	63.74	26.49	175.8	10.99	15.07
ER25	AASI	L	M	3.128	29.82	4.458	42.49	4.29	29.82	6.113	42.49	10.49	14.39
ER25	AASI	L	D	1.429	26.34	1.939	35.75	1.959	26.34	2.66	35.75	5.424	7.439
ER25	ACI	L	M	8.218	53.49	17.67	115	11.27	53.49	24.23	115	15.36	21.07
ER25	ACI	L	D	4.371	55.03	9.719	122.4	5.994	55.03	13.33	122.4	7.942	10.89
ER25	GW	F	D	14.46	33.5	21.74	50.38	19.83	33.5	29.82	50.38	43.15	59.18
ER25	CMD	F	D	36.4	43.99	64.98	78.54	49.91	43.99	89.11	78.54	82.74	113.5
ER25	GW	L	D	11.85	31.78	17.38	46.58	16.26	31.78	23.83	46.58	37.3	51.15
ER25	CMD	F	M	73.03	45.64	134.3	83.95	100.2	45.64	184.2	83.95	160	219.5
ER25	AASI	F	M	4.437	36.55	6.993	57.62	6.085	36.55	9.59	57.62	12.14	16.65
ER25	PHI	F	M	87.5	41.31	149.1	70.39	120	41.31	204.5	70.39	211.8	290.5
ER25	CMD	L	D	31.41	43.91	55.99	78.29	43.07	43.91	76.79	78.29	71.52	98.08
ER25	AASI	F	D	2.139	34.09	3.246	51.73	2.934	34.09	4.452	51.73	6.275	8.606
ER25	PHI	L	D	32.07	33.89	48.51	51.26	43.99	33.89	66.54	51.26	94.65	129.8
ER25	PHI	L	M	67.52	36.88	107	58.43	92.6	36.88	146.7	58.43	183.1	251.1
ER25	GW	L	M	25.57	35.45	39.62	54.91	35.07	35.45	54.33	54.91	72.15	98.95
ER25	PHI	F	D	42.57	38.88	69.64	63.6	58.38	38.88	95.51	63.6	109.5	150.2
ER25	CMD	L	M	62.96	45.51	115.5	83.52	86.34	45.51	158.5	83.52	138.3	189.7
ER25	ICD	L	D	870	86.08	6250	618.4	1193	86.08	8571	618.4	1011	1386
ER25	GW	F	M	30.85	36.97	48.95	58.65	42.32	36.97	67.13	58.65	83.46	114.5
ER25	LSI	F	D	91.58	60.78	233.5	155	125.6	60.78	320.3	155	150.7	206.6
ER25	OTH	F	M	28.51	58.46	68.63	140.8	39.09	58.46	94.12	140.8	48.76	66.87
ER25	OTH	F	D	14.63	58.06	34.89	138.4	20.07	58.06	47.85	138.4	25.21	34.57
ER25	SGI	F	D	149.6	46.66	280.5	87.47	205.2	46.66	384.6	87.47	320.6	439.8
ER25	SGI	F	M	294.6	47.5	561.3	90.49	404.1	47.5	769.7	90.49	620.2	850.6
ER25	LSI	L	D	77.16	59.24	189.3	145.4	105.8	59.24	259.7	145.4	130.2	178.6
ER25	LSI	L	M	147.7	58.64	357.2	141.8	202.6	58.64	489.9	141.8	251.9	345.5
ER25	SGI	L	D	123.8	44.67	223.7	80.72	169.8	44.67	306.8	80.72	277.2	380.1
ER25	SGI	L	M	244.2	45.54	448.3	83.63	334.9	45.54	614.9	83.63	536.1	735.3
ER50	CMD	L	M	81.86	42.59	142.6	74.19	112.3	42.59	195.5	74.19	192.2	263.6
ER50	CMD	L	D	40.4	40.66	68.07	68.51	55.4	40.66	93.36	68.51	99.36	136.3
ER50	GW	F	D	20.95	34.95	32.21	53.72	28.73	34.95	44.17	53.72	59.95	82.21
ER50	CMD	F	D	48.23	41.96	83.1	72.3	66.14	41.96	114	72.3	114.9	157.6
ER50	GW	L	D	16.3	31.45	23.78	45.88	22.35	31.45	32.61	45.88	51.82	71.07
ER50	CMD	F	M	97.47	43.84	173.6	78.06	133.7	43.84	238	78.06	222.3	304.9
ER50	AASI	F	M	6.366	37.75	10.23	60.66	8.731	37.75	14.03	60.66	16.86	23.12
ER50	PHI	F	M	125.5	42.64	218.8	74.35	172.1	42.64	300	74.35	294.2	403.5
ER50	AASI	L	D	1.932	25.63	2.597	34.47	2.649	25.63	3.562	34.47	7.535	10.33
ER50	AASI	i	D	3.656	58.19	8.744	139.1	5.015	58.19	11.99	139.1	6.284	8.618
ER50	GW	i	D	25.42	58.81	61.71	142.8	34.86	58.81	84.63	142.8	43.21	59.27
ER50	AASI	L	M	4.283	29.39	6.065	41.61	5.874	29.39	8.318	41.61	14.58	19.99
ER50	AASI	F	D	3.081	35.35	4.766	54.67	4.226	35.35	6.536	54.67	8.717	11.96

ER50	GW	L	M	35.36	35.28	54.64	54.51	48.5	35.28	74.93	54.51	100.2	137.5
ER50	AASI	i	M	7.266	59.77	18.06	148.6	9.965	59.77	24.77	148.6	12.16	16.67
ER50	PHI	L	D	44.72	34.01	67.77	51.54	61.33	34.01	92.94	51.54	131.5	180.3
ER50	PHI	L	M	94.31	37.08	149.9	58.94	129.3	37.08	205.6	58.94	254.3	348.8
ER50	SGI	L	D	171.8	44.63	310.4	80.6	235.7	44.63	425.6	80.6	385.1	528.1
ER50	SGI	L	M	339.4	45.57	623.6	83.72	465.5	45.57	855.2	83.72	744.8	1021
ER50	LSI	L	M	204.3	58.38	491	140.3	280.2	58.38	673.4	140.3	350	480
ER50	LSI	L	D	106.7	58.96	260	143.7	146.3	58.96	356.5	143.7	180.9	248.2
ER50	ICD	L	M	2332	85.86	16486	607	3198	85.86	22610	607	2716	3725
ER50	ICD	L	D	1206	85.86	8523	607	1653	85.86	11689	607	1404	1926
ER50	OTH	L	D	18.48	61.06	47.46	156.8	25.35	61.06	65.09	156.8	30.27	41.52
ER50	PHI	L	Gymn	1515	95.11	30959	1944	2077	95.11	42458	1944	1593	2184
ER50	ACI	L	M	11.24	52.64	23.72	111.2	15.41	52.64	32.54	111.2	21.34	29.27
ER50	PHI	F	D	61.29	40.29	102.6	67.48	84.05	40.29	140.8	67.48	152.1	208.6
ER50	OTH	F	D	20.78	59.33	51.08	145.9	28.49	59.33	70.05	145.9	35.02	48.03
ER50	LSI	F	M	244.9	60.49	619.9	153.1	335.9	60.49	850.2	153.1	404.9	555.3
ER50	SGI	F	M	417.1	48.41	808.5	93.83	572	48.41	1109	93.83	861.6	1182
ER50	PHI	F	Gymn	1752	95.1	35796	1943	2403	95.1	49092	1943	1842	2527
ER50	GW	F	M	44.38	38.28	71.91	62.01	60.87	38.28	98.61	62.01	116	159
ER50	ACI	F	M	13.73	55.59	30.91	125.2	18.82	55.59	42.39	125.2	24.69	33.86
ER50	OTH	F	M	40.47	59.74	100.5	148.4	55.5	59.74	137.8	148.4	67.74	92.9
ER50	ACI	L	D	5.981	54.2	13.06	118.4	8.202	54.2	17.91	118.4	11.03	15.13

Models	Min1	Max1	Min2	Max2	Min3	Max3	Min4	Max4
ER10	n.d	n.d	n.d	n.d	n.d	n.d	n.d	n.d
ER25	32.49	99.86	48.12	73537	32.49	99.86	48.12	73537
ER50	30.98	94.31	44.9	1657	30.98	94.31	44.9	1657
Pooled	25.63	95.11	34.47	1944	25.63	95.11	34.47	1944

Mixed Effect Models

```
Warning: package 'lme4' was built under R version 3.2.5


=============================================================================
                         Model 1       Model 2       Model 3       Model 4
-----------------------------------------------------------------------------
(Intercept)                -0.88          0.55          1.39 **       0.02
                           (0.78)        (0.50)        (0.48)        (0.42)
CvWW                       -0.03         -0.24 *       -0.43 ***     -0.34 ***
                           (0.33)        (0.11)        (0.11)        (0.07)
MoA.CODEACI                 7.04 ***      2.65          2.39 *        2.41 *
                           (2.12)        (1.61)        (0.98)        (0.97)
MoA.CODECMD                 6.64 ***      5.65 ***      4.59 ***      5.45 ***
                           (1.34)        (0.97)        (0.75)        (0.68)
MoA.CODEGW                  3.55 ***      3.19 ***      3.73 ***      3.78 ***
                           (1.01)        (0.69)        (0.67)        (0.57)
MoA.CODEOTH                 6.71 **       3.18 **       5.54 ***      4.87 ***
                           (2.29)        (1.11)        (0.93)        (0.76)
MoA.CODEPHI                 4.90 **       4.29 ***      4.32 ***      4.67 ***
                           (1.64)        (0.81)        (0.68)        (0.64)
MoA.CODESGI                 5.56 **       5.77 ***      5.32 ***      5.59 ***
                           (1.95)        (1.02)        (0.83)        (0.79)
LabvFieldi                 -1.22 *                     -0.33         -0.74 **
                           (0.61)                      (0.28)        (0.23)
LabvFieldL                 -0.25         -0.23         -0.31 *       -0.42 ***
```

```
                            (0.40)        (0.17)        (0.14)        (0.10)
Class.MvDM                   1.14 ***      0.65 ***      0.73 ***      0.74 ***
                            (0.26)        (0.11)        (0.10)        (0.07)
MoA.CODEICD                                7.06 **       6.06 ***      6.47 ***
                                          (2.21)        (1.21)        (1.09)
MoA.CODELSI                                4.01 ***      5.10 ***      4.97 ***
                                          (1.21)        (0.99)        (0.97)
Class.MvDGymn                                            2.37 **       2.21 **
                                                        (0.81)        (0.84)
Effect.level.ERxER25                                                   0.56 ***
                                                                      (0.11)
Effect.level.ERxER50                                                   1.27 ***
                                                                      (0.11)
CvWI                                                                  -0.61
                                                                      (1.61)
-------------------------------------------------------------------------------
AIC                       1275.38       4370.28       4859.41      10543.39
BIC                       1323.87       4440.93       4941.90      10655.78
Log Likelihood            -624.69      -2171.14      -2413.71      -5252.70
Num. obs.                  308          1148          1281          2738
Num. groups: AS.Code        23            42            55            67
Var: AS.Code (Intercept)     3.13          2.30          1.89          1.96
Var: Residual                3.14          2.37          2.34          2.57
===============================================================================
*** p < 0.001, ** p < 0.01, * p < 0.05
```

Conclusion

The working hypothesis could not be verified, overall there were no pronounced differences in sensitivity between wild plant species and crop species.

Differences varied depending on the effect level analyzed, i.e. the outcome based on ER25 was somewhat different than the one based on comparison of ER10 and ER50. However, supposing a consistent difference in sensitivity between crops and wild species at least the direction of the slope should have been similar irrespective of the effect level assessed.

The revised pooled analysis (all three effect levels) which is considered most relevant, (considering that the fundamental relationship should be visible with any of the three effect levels) indicates statistical significant differences between crop and wild plant species. The coefficient between logs of crop and wild endpoints are 1.28 for the model with interactions, and 0.32 for the standard model (no interactions). Taking into account that the model with interactions ("best model") tries to optimize purely on mathematics, disregarding the actual relations between different explanatory variables, the results of the standard model (no interactions) are considered more reliable.

For interpretation note that the ERx values had been logarithmized, so to assess the actual difference between baseline and variable assessed it has to be back-transformed.

Wild endpoints were thus on average by a factor of 3.6 ($e^{1.28}$) higher compared to the baseline i.e. the crops (model with interactions "best"), but only by a factor of 1.37 ($e^{0.32}$) when based on the standard model (no interaction)

The coefficient of 0.56 for ER25 (Pooled best) indicates that ER25 were a factor of 1.75 higher than the ER10, and ER50 were a factor of 2.9 ($e^{1.07}$) higher than the ER10 in this database. All these factor levels were statistically significantly different from the baseline.

When comparing different modes of action, it should be noted that the base line (AASI) was also the one with by far most data (about half of all endpoints). Also they included the most potent active substances, where only a few grams per hectare achieve a similar level of weed control as other modes of actions where rather kg per hectare are needed to achieve a similar level of weed control. Consequently all coefficients were positive, with back-transformed values ranging from 4 ($e^{1.39}$) to 512 ($e^{6.24}$). Compared to these large differences those between lab and field tests were only minor, and generally not significantly different. Monocots and dicots different significantly, with monocots on average less sensitive than dicots, details see first coefficient table.

In the second coefficient table, results of the standard model (no interactions) are presented for individual effect levels, and for pooled data. Based solely on ER10, wild endpoints were by a factor of 1.45 ($e^{-0.37}$) **lower** than crop endpoints, based on ER25 by a factor of 1.9 ($e^{0.65}$) **higher** and based on ER50 basically identical, factor 1.01 ($e^{-0.01}$). All effect levels pooled indicated that wild plants endpoints were by a factor of 1.37 ($e^{0.32}$) higher than crops, so crops were slightly more sensitive than wild plants.

In two additional runs the pooled models were repeated with modified data. All crop species endpoints were increased by factors of 1.5 or 2.0, while the wild species' endpoints were left unchanged. Considering that the analysis of the original data showed crop endpoints to be on average slightly lower than wild endpoints, the manipulation by a factor of 1.5 should reduce these differences, and the canonical coefficients for the variable 'CvW' should be reduced by a factor of 0.405 (ln of 1.5) and manipulation by a factor of 2 should reduce the coefficients by 0.69 (ln of 2.0), respectively. The actual results were close to these theoretical expectations, see Table . Based on the pooled standard model (no interactions), original, factor 1.5- and factor 2-modified data resulted in coefficients of 0.32, -0.09 and -0.38s, which are very close to the predicted changes. this indicates that the method was able to detect differences in sensitivity between crops and wild plants around a factor of 1.5.

Mixed effect models

Based on the mixed effect model, coefficients ranged between -0.03 and -0.43, with some models indicating significant differences, albeit on average only marginal. Again differences between different modes of action were conspicuous, those between lab and field tests moderate, with lab endpoints slightly lower than field endpoints, and again clear differences between monocots and dicots, the latter being more sensitive (monocots had higher endpoints). Please note that in case of mixed effect models the coefficients do no longer allow to deduce an average difference between groups (by backtransforming them) for mathematical reasons.

Based on all data (revised, now including vegetative ER10, ER25 and ER50 data), wild plant endpoints thus appear – depending on the model – either to be similar or to be significantly higher than crop endpoints. The hypothesis that wild species are per se more sensitive to agrochemicals than wild crops was thus not supported by this analysis.

Discussion

In contrast to the other project (vegetative/reproductive) the database here does not allow any paired assessment, as there are only a few exceptional instances where both a wild and a crop variant exists from the same species (e.g. Daucus carota was sometimes attributed as "garden carrot" or as "wild carrot"). Generally we had however to match wild data of certain subgroups with crop data, which ever were present. The data are fundamentally heterogeneous, and not all parameters potentially affecting the sensitivity were regularly reported, resulting in a quite patchy database in terms of availability of explanatory variables.

Still these are the data available, and they are considered to allow a useful analysis.

Probably the ultimate protection against non-normality and heterogeneity is to use a non-parametric test. A non-parametric approach for two-way ANOVA is to use ordinal logistic regression.

Right censored data could possibly be compared by Harrington and Fleming's G(rho) tests but we are not sure how to incorporate more than one factors. See however the other statistical Appendix (John W Green) where censored values were considered in an analysis of distributions.

In two additional runs the pooled models were repeated with modified data. All crop species endpoints were increased by factors of 1.5 or 2.0, while the wild species' endpoints were left unchanged. Considering that the analysis of the original data showed crop endpoints to be on average slightly lower than wild endpoints, the manipulation by a factor of 1.5 should reduce these differences, and the canonical coefficients for the variable 'CvW' should be reduced by a factor of 0.405 (ln of 1.5) and manipulation by a factor of 2 should reduce the coefficients by 0.69 (ln of 2.0). The actual results were close to these theoretical expectations, see Table . Based on the pooled standard model (no interactions), original, factor 1.5- and factor 2-modified data resulted in coefficients of 0.32, -0.09 and -0.38s, which are very close to the predicted changes. this indicates that the method was able to detect differences in sensitivity between crops and wild plants around a factor of 1.5.

References

- Herberich E, Sikorski J, Hothorn T (2010) A Robust Procedure for Comparing Multiple Means under Heteroscedasticity in Unbalanced Designs. PLoS ONE 5(3): e9788. doi: 10.1371/journal.pone.0009788
- http://en.wikipedia.org/wiki/Heteroscedasticity-consistent_standard_errors

Useful Links

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

R version 3.2.5 (2016-04-14) -- "Very, Very Secure Dishes"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

15 Appendix 7 - Additional statistical analysis of Crop and Wild Plants (comparisons of distributions – John W. Green)

April 24, 2016

TO: Stefania Loutseti
Heino Christl
FROM: John W. Green
RE: Revised Comparison of Crop and Wild Non-Target Terrestrial Plant Species Sensitivities

Two workshops were held in Wageningen, Netherlands in 2014 and 2015 devoted to issues related to evaluation of possible adverse effects of crop protection products on non-target terrestrial plants. One of several important issues considered at these workshops was whether there is a need to require more wild species to be evaluated beyond the current one wild species, typically wild ryegrass. To answer this question, a large database of results from non-target terrestrial plant studies was collected from across industry and the regulatory community. An analysis was done of this database and a draft report was prepared, entitled Literature *review and analysis Sensitivity of Wild Plant and Crop Species In Context of 1107/2009*. Chen Teel, who participated in the second workshop, and I, who participated in the first workshop, expressed some concerns about the methodology used in that report. As a result, a database was made available to us that contained 1753 species ER25 values, 2056 EC50 values, and 526 ER10 values. This database did not contain all the results used in the cited report because of proprietary concerns from some companies providing the data and it also contained some new results that were not included in the cited report. It is expected that the database available to us is sufficiently similar to the one used in the cited report that conclusions drawn from it will be relevant. A major conclusion of the cited report was that there were no significant differences in sensitivity between wild and crop species. The present report is in general agreement, with some modifications. In particular, for ER50 values, wild species tend to be lower than crop species for some modes of action. Partly this is due to the inability to record 50% non-lethal effects or considerable uncertainty in these estimates. However, for ER25 estimates, there is very little evidence to support a need for additional wild species. One family of wild species, Poaceae, should be included to capture the more sensitive wild species. Current NTTP studies typically already include such a wild species, for example, wild ryegrass, and that or some member of the fescue family should be sufficiently protective. There is likewise little evidence among ER10 values to support a need to add wild species. In this case, there is more limited data for ER10 than for ER25 or ER50. Furthermore, there is much evidence in other databases to suggest that ER10 is often not a viable endpoint in the sense that it is often very difficult to distinguish a 10% effect from natural biological "noise."

Of the 1753 species ER25 values available to us, 611(35%) were either left- or right-censored. In addition, 39% (203 of 526) of EC10 values were censored, and 38% (779 of 2056) EC50 values were censored. Clearly, censored values are very common and ignoring them could bias the conclusions by truncating the distribution at both ends of the range of values in uneven fashion. In particular, the truncated data is likely to underestimate the true variability. Fortunately, there are mathematically sound ways to incorporate censored values in the analysis. The procedure for doing so is described below. Underestimating variability is not entirely eliminated by the adaptation of the method used, but it is reduced.

Censored values should be treated differently from the other values, which are true estimates. The standard way to incorporate such censored values in fitting a distribution is to use maximum likelihood estimates (MLEs) of the mean and standard deviation of the presumed distribution taking the censored values directly into account, so as to reflect the uncertainty in these censored values. A lognormal distribution is generally adequate to describe ERt estimates, t=10, 25, 50. To fit a log-normal distribution, it is convenient to fit a normal distribution to the logarithm of the ERt values. To avoid biases that can arise from log-transforming very small numbers, the actual transform used was LogECt=Log(1+ERt), which ensures positive values. The likelihood function of the log-transformed data has the form

$$L(x;\mu,\sigma)=\prod_{i=1}^{L}\Phi\left(\frac{x_i-\mu}{\sigma}\right)\prod_{i=L+1}^{L+k}\phi\left(\frac{x_i-\mu}{\sigma}\right)\prod_{i=L+k+1}^{nL+k+R}\left[1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]$$

where φ (used in the middle product in the above formula) and Φ are the unit normal probability density function and cumulative distribution function, respectively, where there are L left-censored values, *k* uncensored values and R right-censored values. References on this approach include Helsel (2005), Kom Kan King (2013), Lee and Wang (2013), and their references. The ML estimates of the mean and standard deviations were obtained from SAS version 9.4 Proc Lifereg.

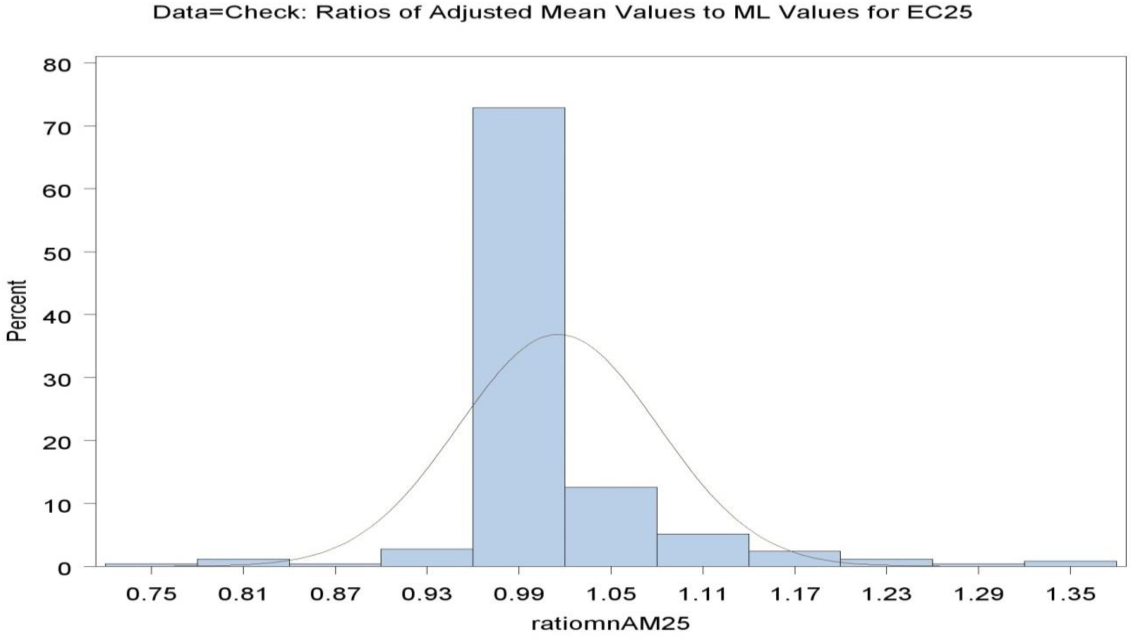
Some grouping of data was needed. For each combination of test venue (greenhouse, hereafter referred to as lab, or field), test type (vegetative vigor or seedling emergence, referred to as VV and SE, respectively), mode of action, monocot/dicot, family, and species type (crop or wild), the above maximum likelihood methodology was used to estimate the mean and standard deviation of the distribution of logECt values for that combination. While different species within the same family can have different sensitivities to the same test substance, there are too many species and too few observations per species to do separate analyses for every species. Restricting species in the same family to the same mode of action (and test venue, test type, and response type) seems a reasonable way to group the data for the purpose at hand.

Combinations that had at most one ERt value were eliminated from further analysis, as were combinations where all values were censored. This left (a) 1279 species EC25 values, of which 507 were censored, in 220 combinations, with 2-56 species EC25 values each; (b) 1608 ER50 values, of which 634 were censored, in 254 combinations, with 2-56 species ER50 values each; (c) 379 ER10 values, of which 177 were censored, in 54 combinations, with 2-56 species EC10 values each.

Once the maximum likelihood estimates of the mean and standard deviation are obtained taking censoring into account, left censored logECt values were replaced by the10th percentile of the fitted distribution, and right-censored values were replaced by the 90th percentile. If the 10th percentile was zero or negative, then instead, the minimum logECt value divided by 100 was used instead, since zero and negative values are biologically impossible. Also, if the 90th percentile was less than the largest logECt value, it was replaced by that largest value. The mean and standard deviation of the resulting logECt values was compared to the ML estimated parameters to make sure this substitution did not drastically alter them.

Figure 1 is a histogram of the ratio of adjusted mean logEC25 to ML estimate, with three extremely large and two negative ML estimates excluded. The agreement is generally good, with a slightly higher likelihood of overestimating logEC25 rather than underestimating it.

Figure 1. Ratio of Adjusted Mean EC25 Means to ML Value



Two large values (5.0 and 5.9) have been omitted.
Figure 2. Ratio of Adjusted EC25 Standard Deviations to ML Value

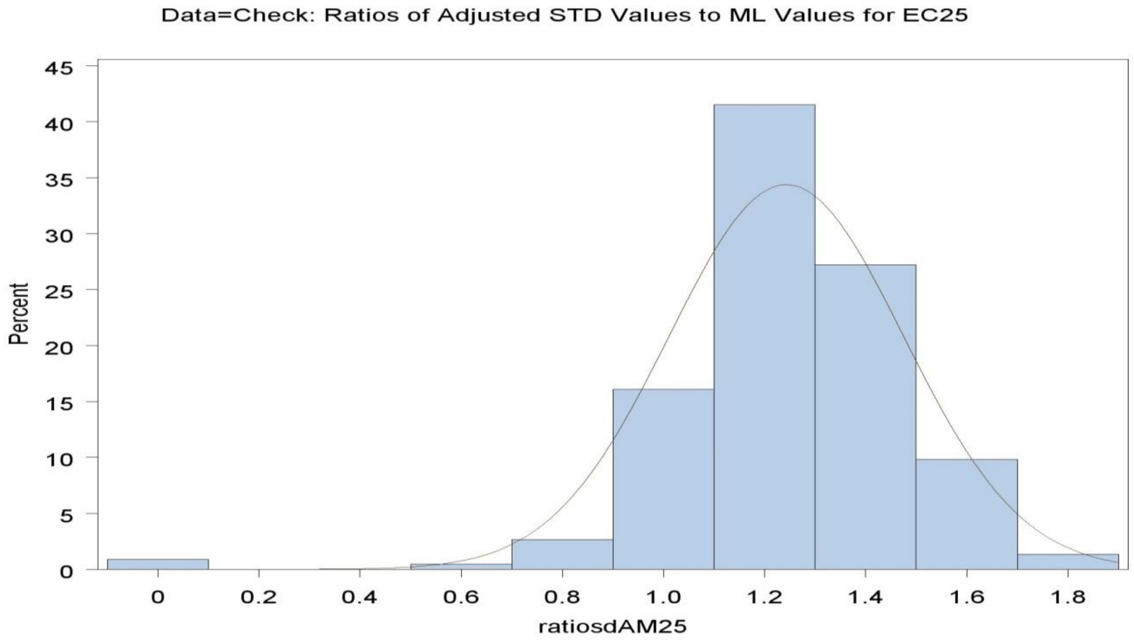
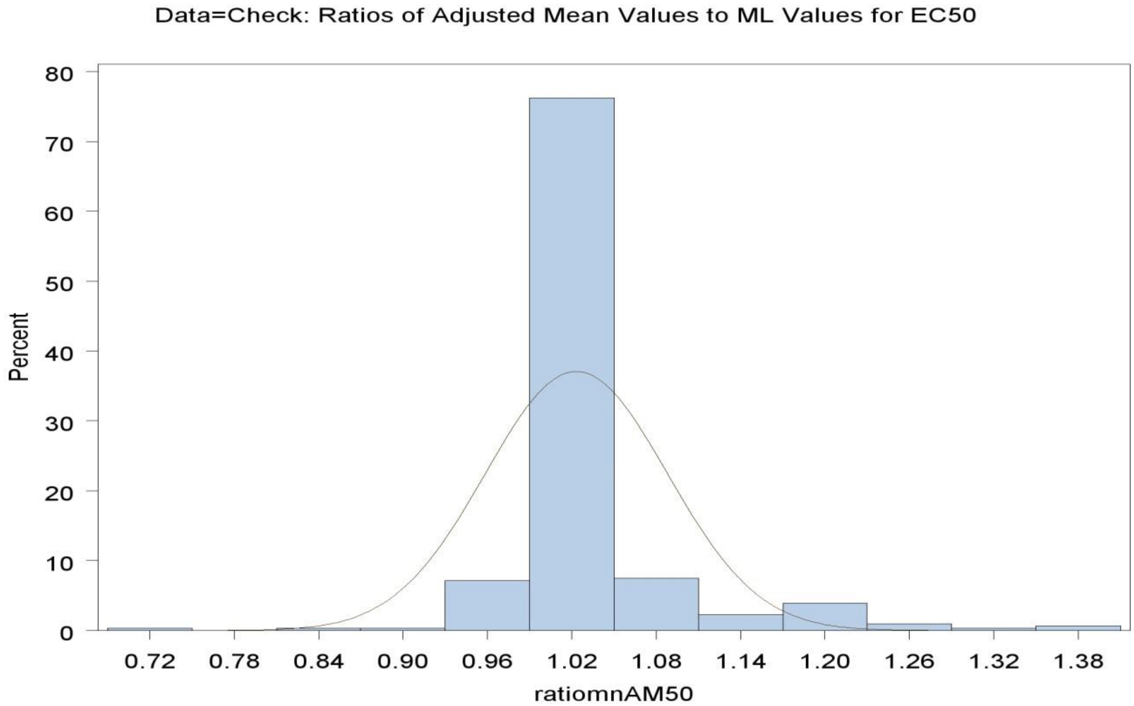


Figure 3. Ratio of Adjusted Mean EC50 Means to ML Value



Two large values (5.0 and 5.9) have been omitted.

Figure 4. Ratio of Adjusted EC50 Standard Deviations to ML Value

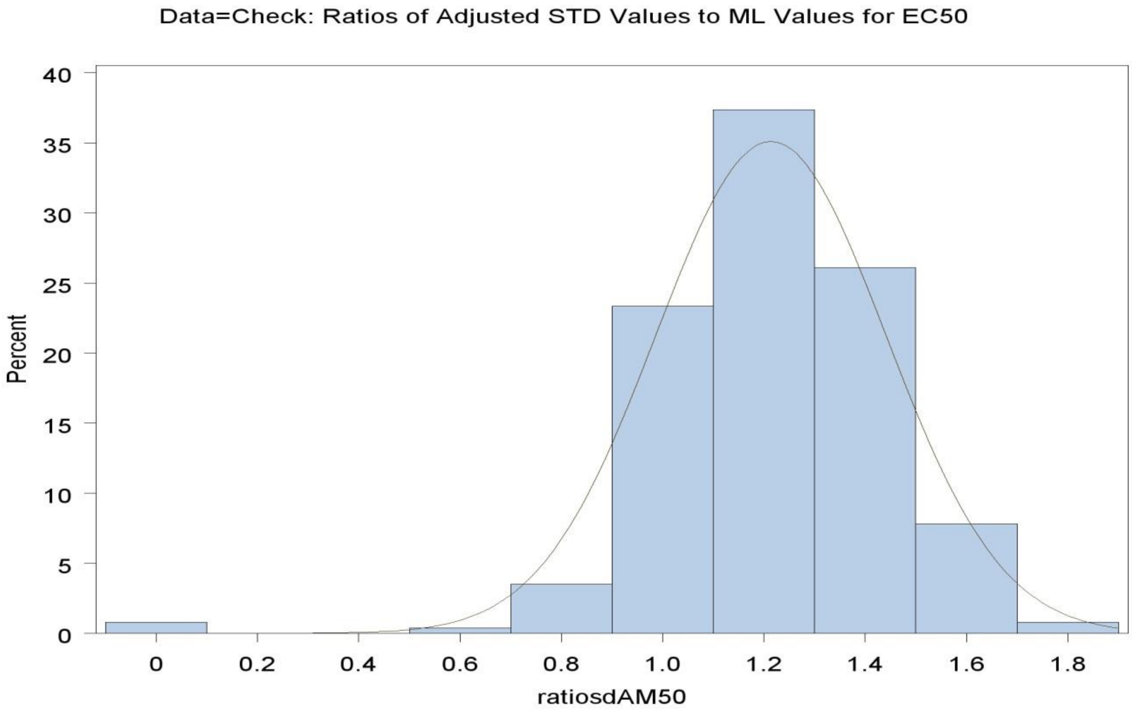
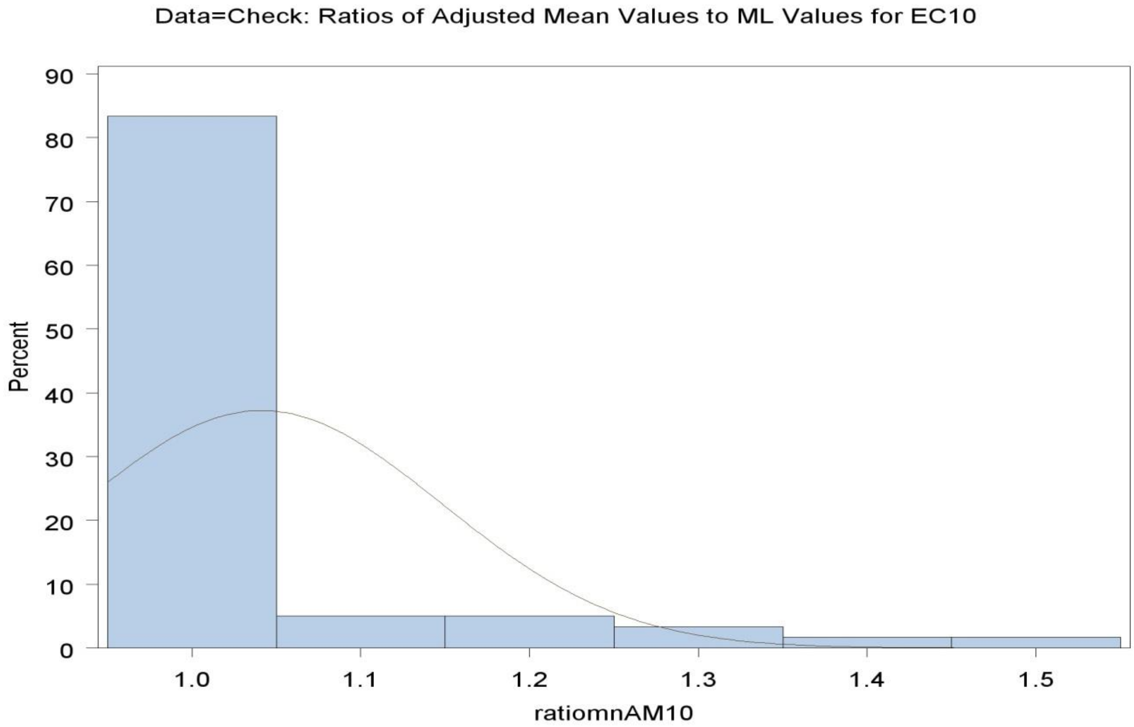
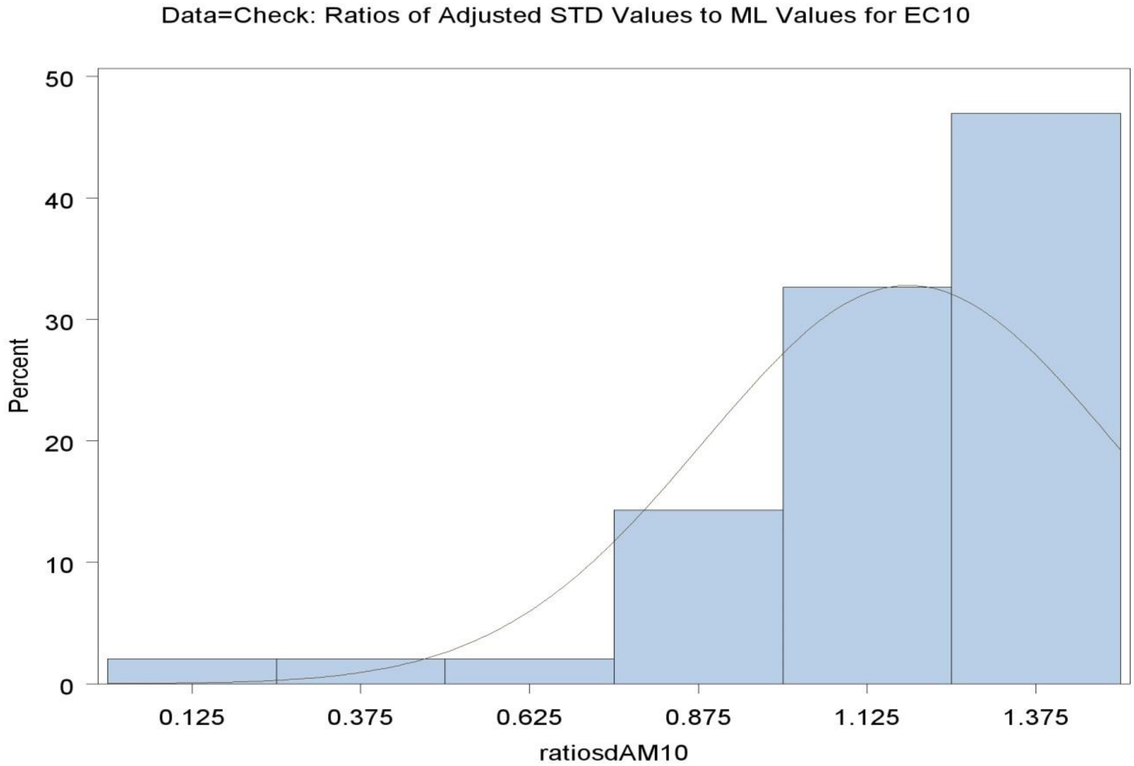

Data=Check: Ratios of Adjusted STD Values to ML Values for EC50
Percent
40
35
30
25
20
15
10
5
0
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
ratiosdAM50

Figure 5. Ratio of Adjusted Mean EC10 Means to ML Value



Two large values (3.8 and 5.9) have been omitted.

Figure 6. Ratio of Adjusted EC10 Standard Deviations to ML Value



In Figures 2, 4, and 6, all of the standard deviation ratios greater than 1.3 come from combinations with at most 2 uncensored values and, with small samples sizes (≤ 2 except for a very few of size 3 and one of size 4). With that understanding, these figures indicate good agreement of adjusted parameter values (mean and standard deviation) with maximum likelihood estimates.

An overall comparison of ERt estimates for crop and wild species is provided by Figures 7-9, where side-by-side histograms of ERt estimates restricted to different orders of magnitude. These figures suggest similar sensitivities for the two types of plant species.

Figure 7. Comparison of ER50 Values for Crop and Wild Species

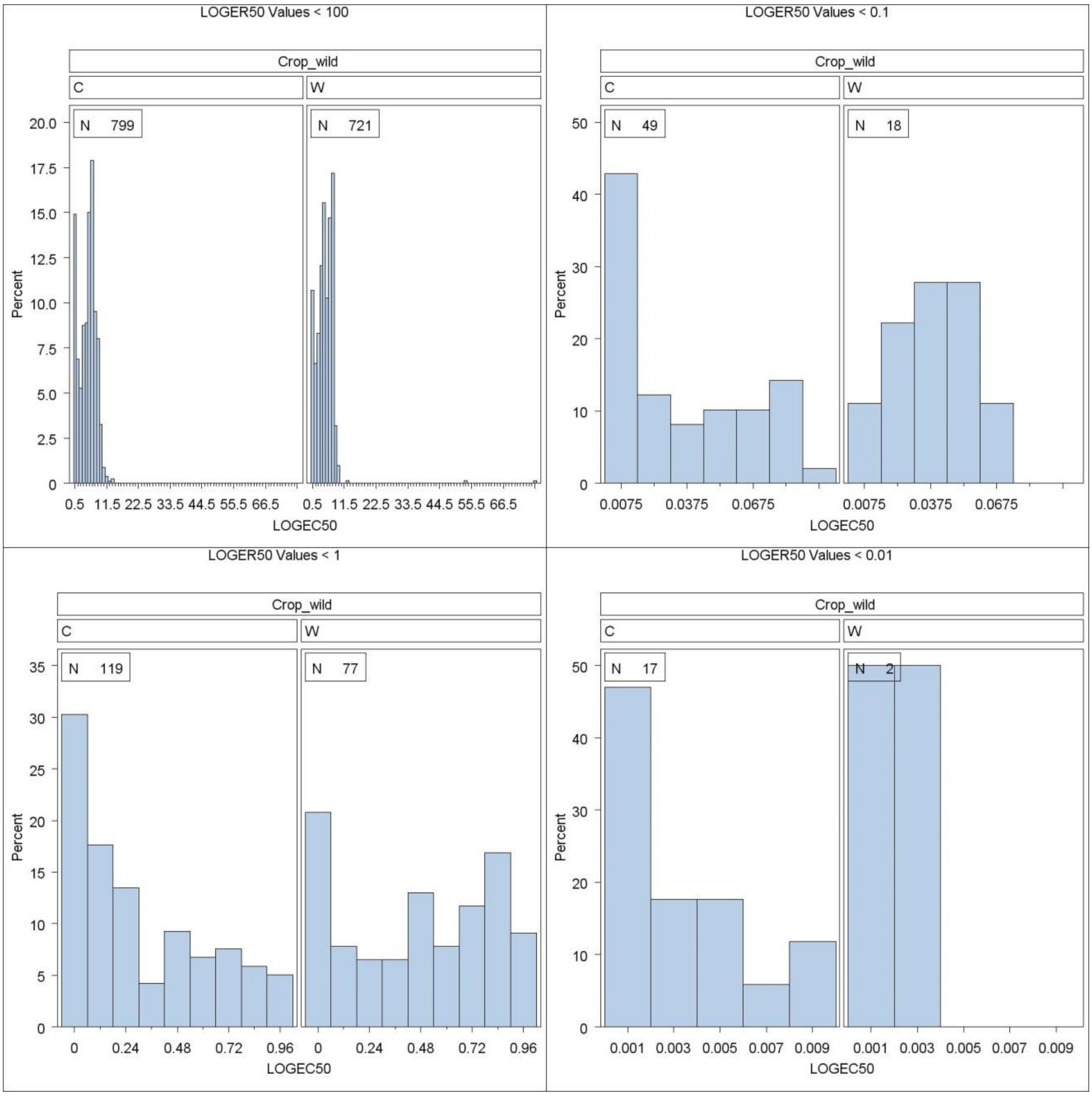


Figure 8. Comparison of ER25 Values for Crop and Wild Species

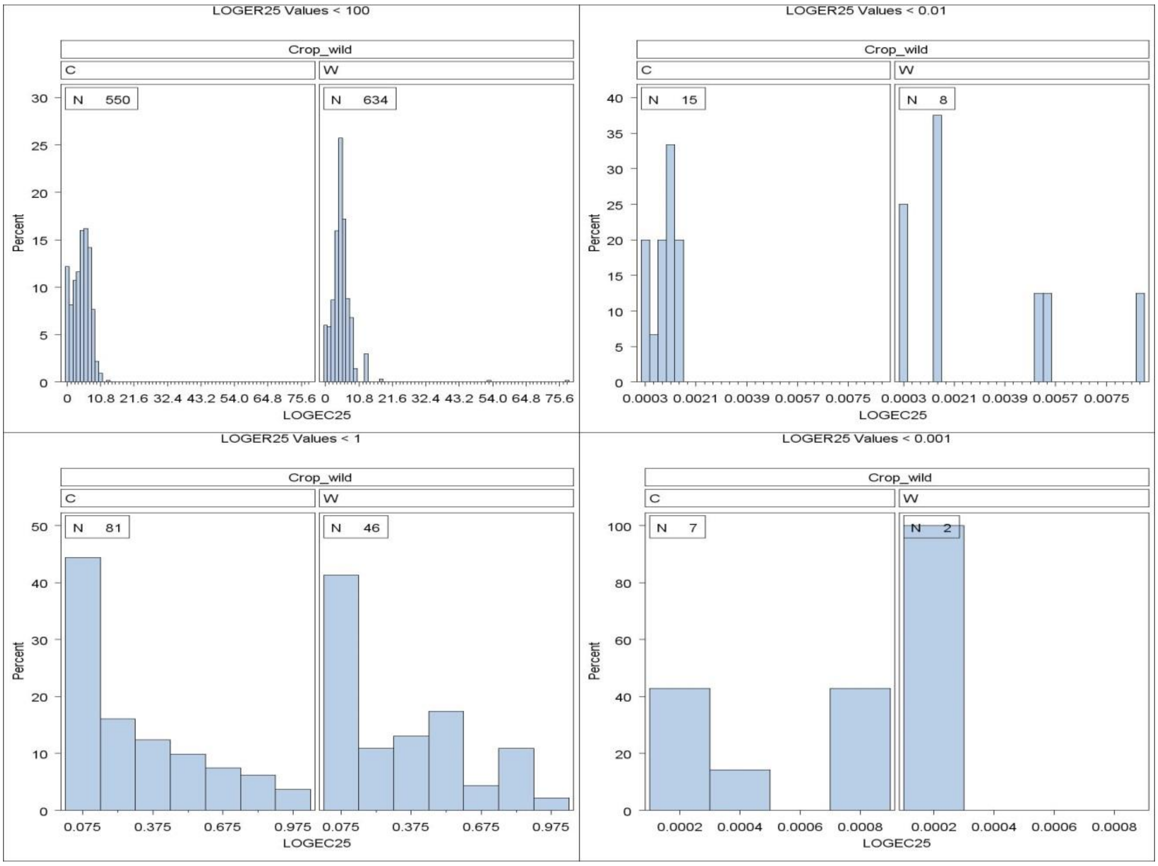
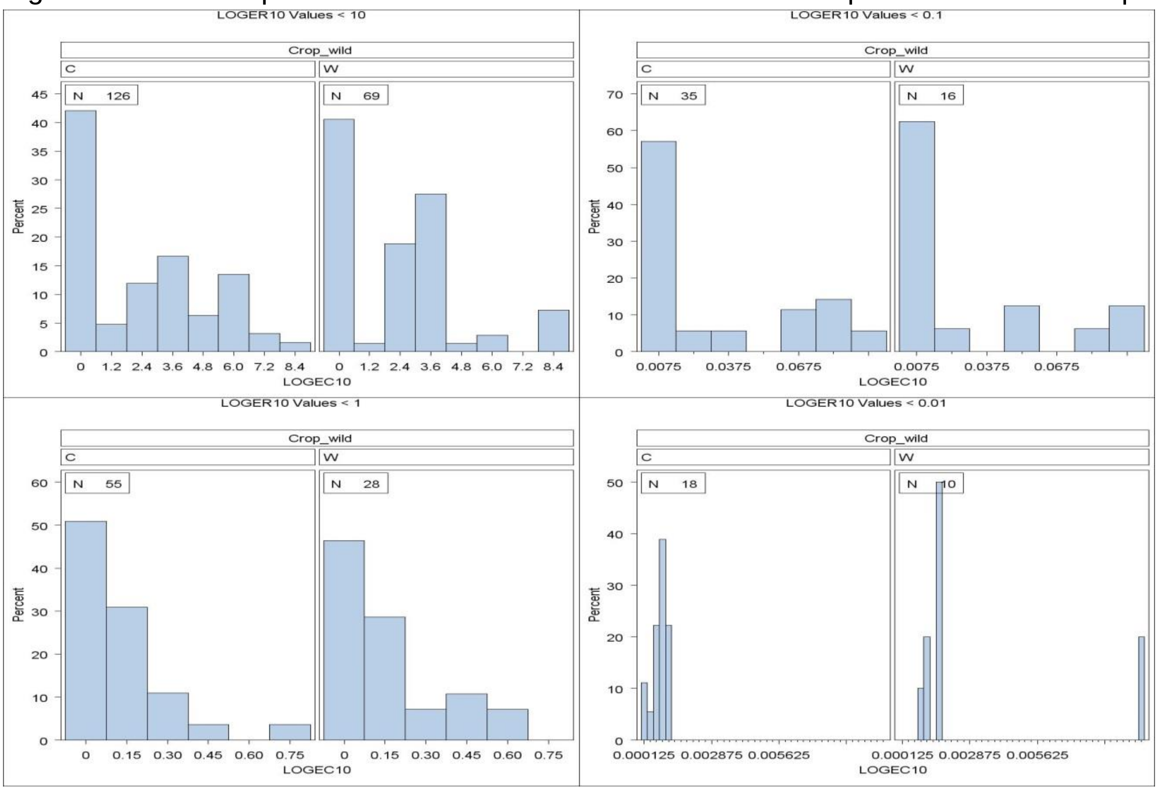


Figure 9. Comparison of ER10 Values for Crop and Wild Species



There may be concern that Figures 7-9 are too crude and that a more informative assessment would require comparisons by endpoint (BM or SH), mode of action, or other characteristics. To that end, the next step was to do a simple ANOVA of logECt with explanatory factors of test venue, test type, family, mode of action, monocot/dicot, and species type. No significant difference was found between monocots and dicots once adjustments for the other factors were taken into account and this category was dropped from further analysis. The next step was to do separate ANOVAs of logECt with sole factor the interaction of mode of action and species type (crop/wild) for each combination of test type (VV or SE), species type, family, and response (biomass, shoot height, survival). Within each such ANOVA, there was little concern about variance heterogeneity and it was possible to compare wild and crop species for each mode of action.

Each ANOVA produced one test indicating whether there was a significance difference between wild and crop species for the indicated family, mode of action, and venue. The mean responses for crop (cmean) and wild (wmean) and their ratio (cmean/wmean) were calculated. Only those tests that were significant at the 0.1 level are shown.

ER50 VV BM

Table 1. Compare Crop and Wild Adjusted Mean EC50 Values

Family	Lab_field	MoA_CODE	ProbF	Cmean	Wmean	Ratiocw
Brassicaceae	L	AASI	0.034	3.056	1.330	2.30
Brassicaceae	L	CMD	0.023	6.815	3.693	1.85
Brassicaceae	L	GW	0.032	4.990	2.874	1.74
Chenopodiaceae	L	AASI	0.025	0.851	3.044	0.28
Fabaceae	L	AASI	0.050	2.440	4.195	0.58
Fabaceae	L	CMD	0.004	9.105	3.666	2.48
Poaceae	L	CMD	0.032	6.672	4.916	1.36
Poaceae	L	SGI	0.007	8.772	5.753	1.52

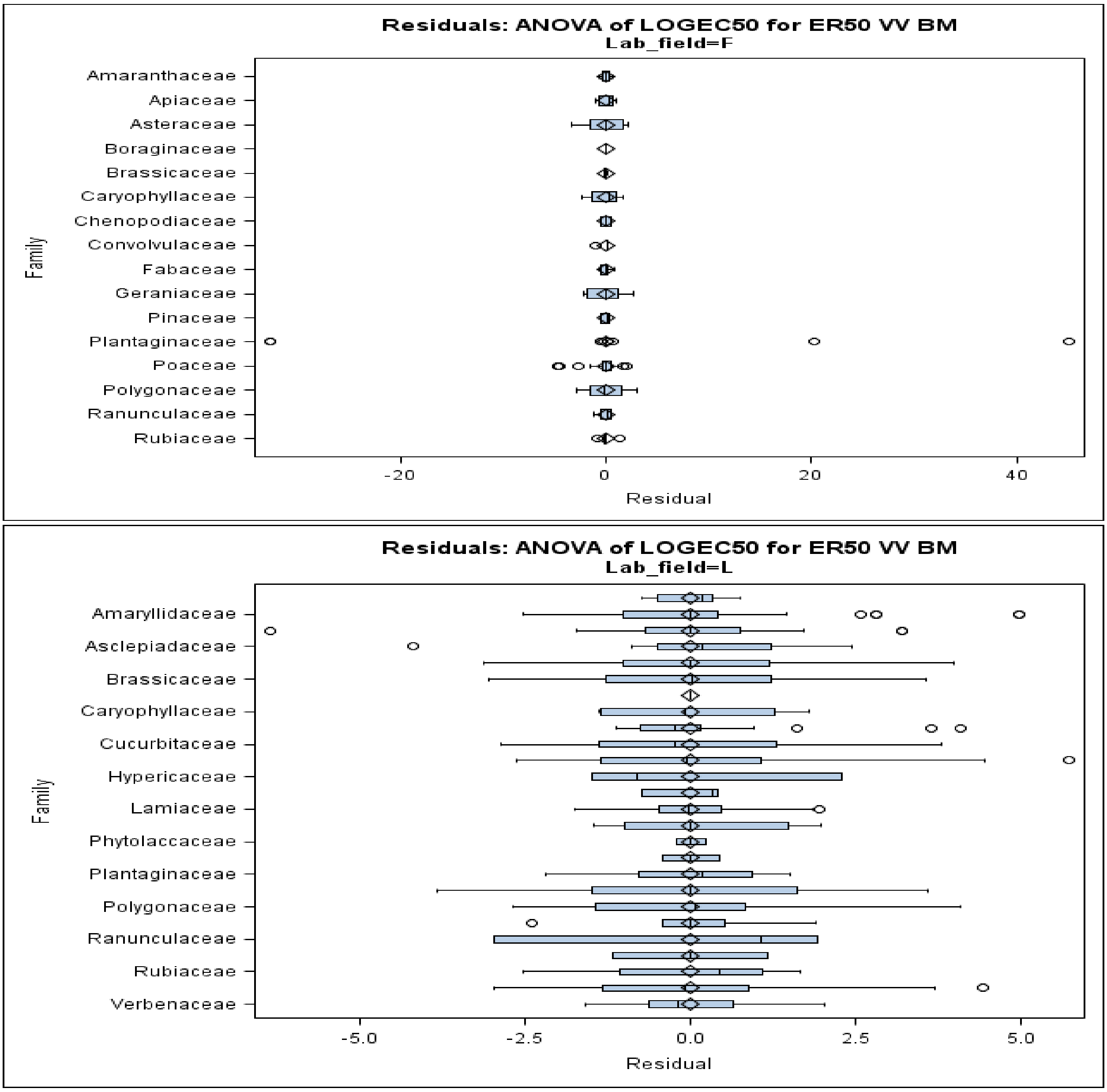
Visual assessment of variances across families for each venue (lab or field) is shown in Figure 10. While differences are apparent, they are not great and, more importantly, no two families were included in the same model. Clearly there are differences in variability between lab and field studies. This is due to the much smaller number of VV BM field studies for which ER50 was reported.

For family Brassicaceae with MoA_Code=CMD, there was only one crop species, cabbage (*Brassica oleracea*) and one wild species, *Capsella bursa-pastoris* and small numbers of each. For MoA_Code=AASI, there were three species of each type. Crop ER50 values lower, by an order of magnitude, than any observed ER50 values among wild species were present, but there was much greater variation among ER50 values even for the same species (mostly rapeseed). No explanation for this variation was evident in the database, other than the wild results were based on much smaller data sets. For MoA_Code=GW, there were only four wild species ER50 values and the lowest ER50 is high enough that it is unlikely to drive the risk assessment.

For Fabaceae with MoA_Code=CMD, there were only three wild species ER50 values and two crop values, in each case all of the same species. While the crop values are much higher, lowest wild ER50 is unlikely to drive the risk assessment.

For Poaceae with MoA_Code=CMD, the crop values tend to be higher, but even lowest wild ER50 is unlikely to drive the risk assessment. With MoA_Code=SGI, again the crop values tend to be higher, but even lowest wild ER50 is unlikely to drive the risk assessment.

Figure 10. Residuals from ANOVA of VV Biomass ANOVA of Log(1+EC50)



ER25 VV BM

Table 2. Compare Crop and Wild Adjusted Mean EC25 Values

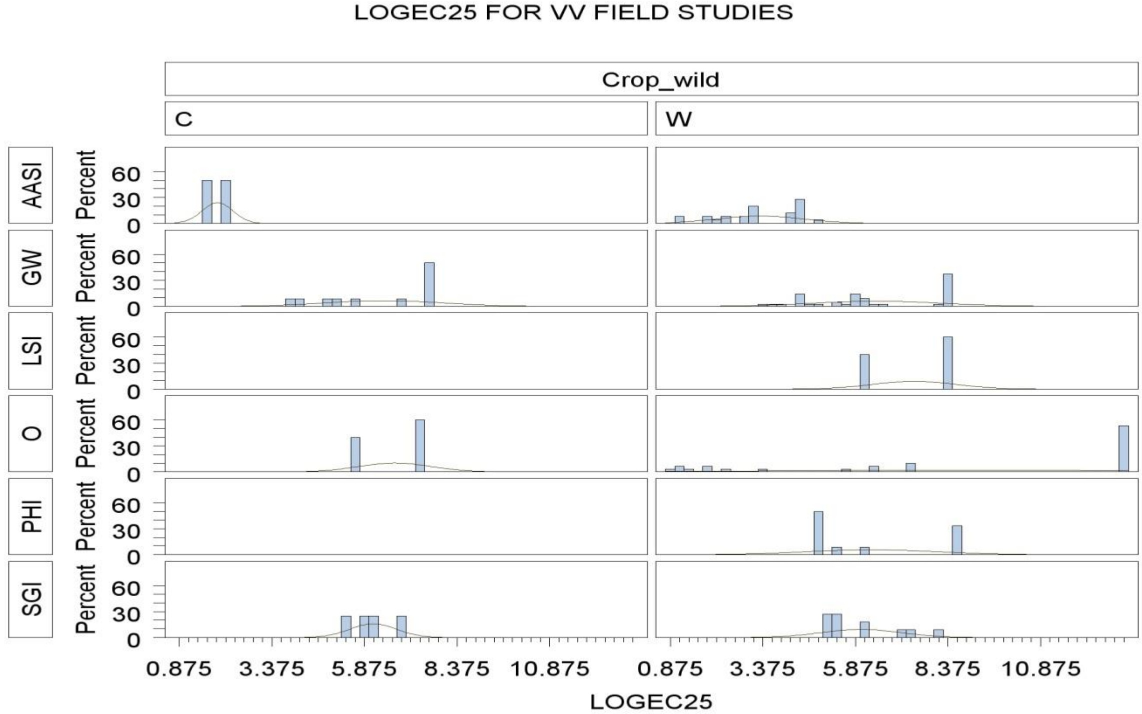
Family	Lab_field	MoA_CODE_	ProbF	cmean	wmean	ratiocw
Fabaceae	F	GW	0.045	2.361	4.019	0.59
Poaceae	F	O	0.086	5.631	8.986	0.63
Asteraceae	L	AASI	0.040	3.517	4.600	0.76
Asteraceae	L	PHI	0.078	4.001	5.150	0.78
Chenopodiaceae	L	AASI	0.004	0.242	4.047	0.06
Poaceae	L	AASI	0.041	3.197	4.374	0.73
Poaceae	L	SGI	0.019	7.764	3.153	2.46

The only instance in which the crop ER25 value is significantly larger than the wild value is for the family Poaceae with MoA_Code=SGI.

The family Poaceae is the only family of non-target terrestrial plants found in vegetative vigor (VV) studies where some summary measures indicate that wild species may have lower EC25 values than crop species for biomass. The crop species in this family are millet, maize, and wheat. The most sensitive wild species, relative to biomass EC25, are *Elymus riparius* (riverside wild rye), *Lolium perenne* (English ryegrass), and *Setaria faberi* (Japanese bristlegrass).

Whether crop or wild species are more sensitive could depend on mode of action. Figure 11 compares crop and wild species within the Poaceae family by mode of action. There were only two wild ER25 values for MoA_Code=SGI, both from *Lolium* sp., but they are notably lower than the crop values. It is not clear that any inclusion of wild species to capture this one difference is needed, given that the crop values for other modes of action yield lower ER25 values than those observed for *Lolium* sp., but it is already common to include a wild species from this family.

Figure 11. Comparison of crop and wild species in family Poaceae within VV field studies



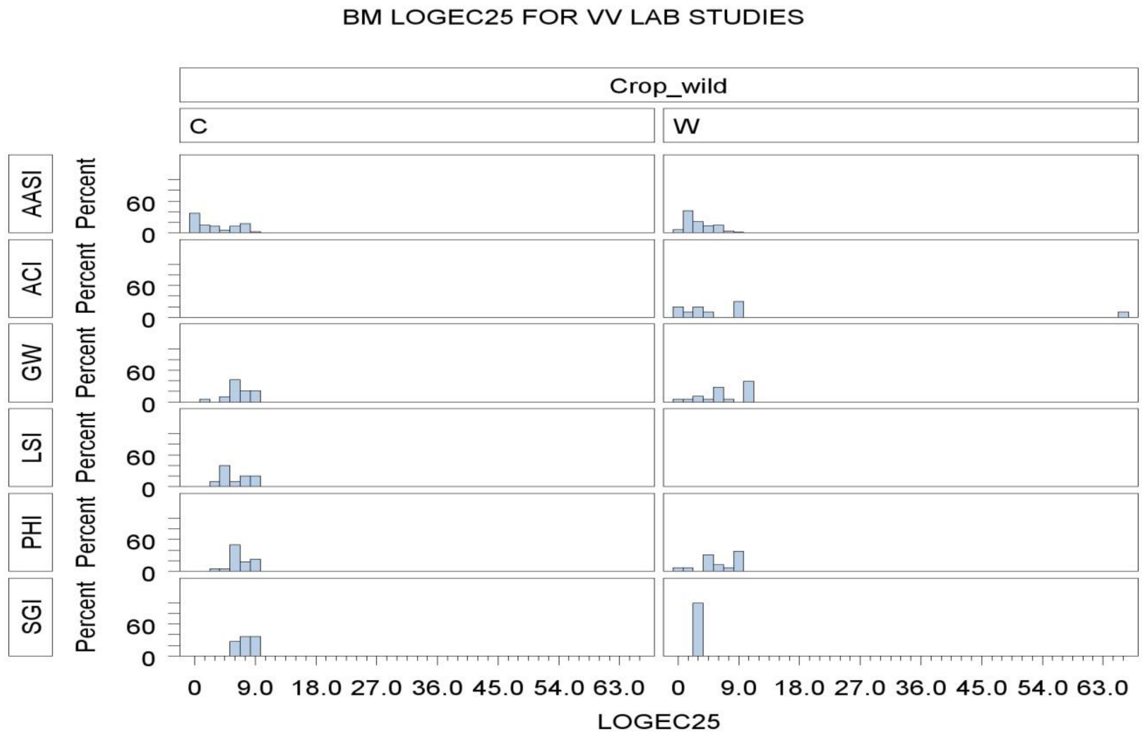
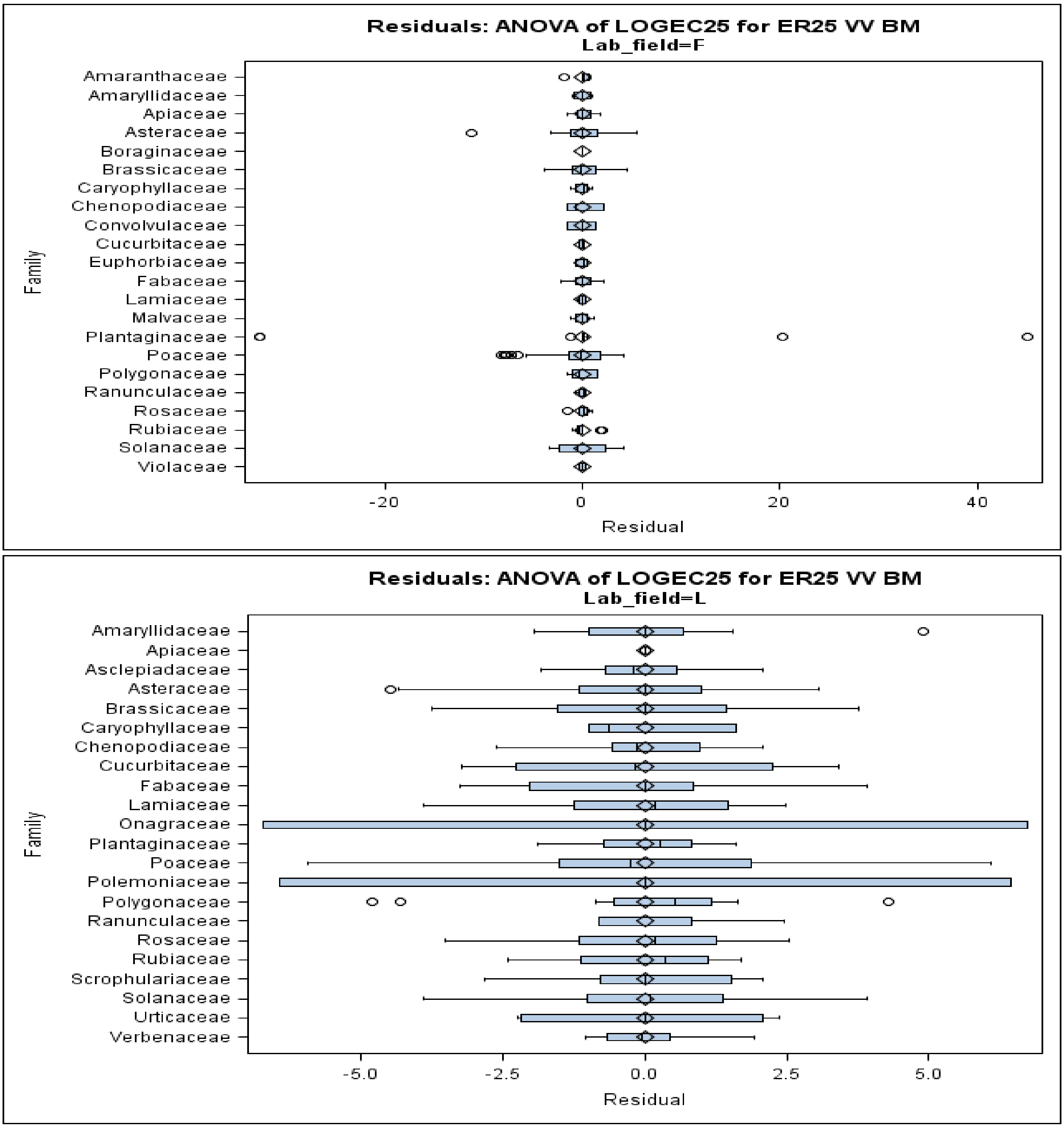


Figure 12. Residuals from ANOVA of VV Biomass ANOVA of Log(1+EC25)



ER10 VV BM

There was limited data and no combinations of family and mode of action for which there was a significant difference between wild and crop ER10 values.

ER50 SE BM

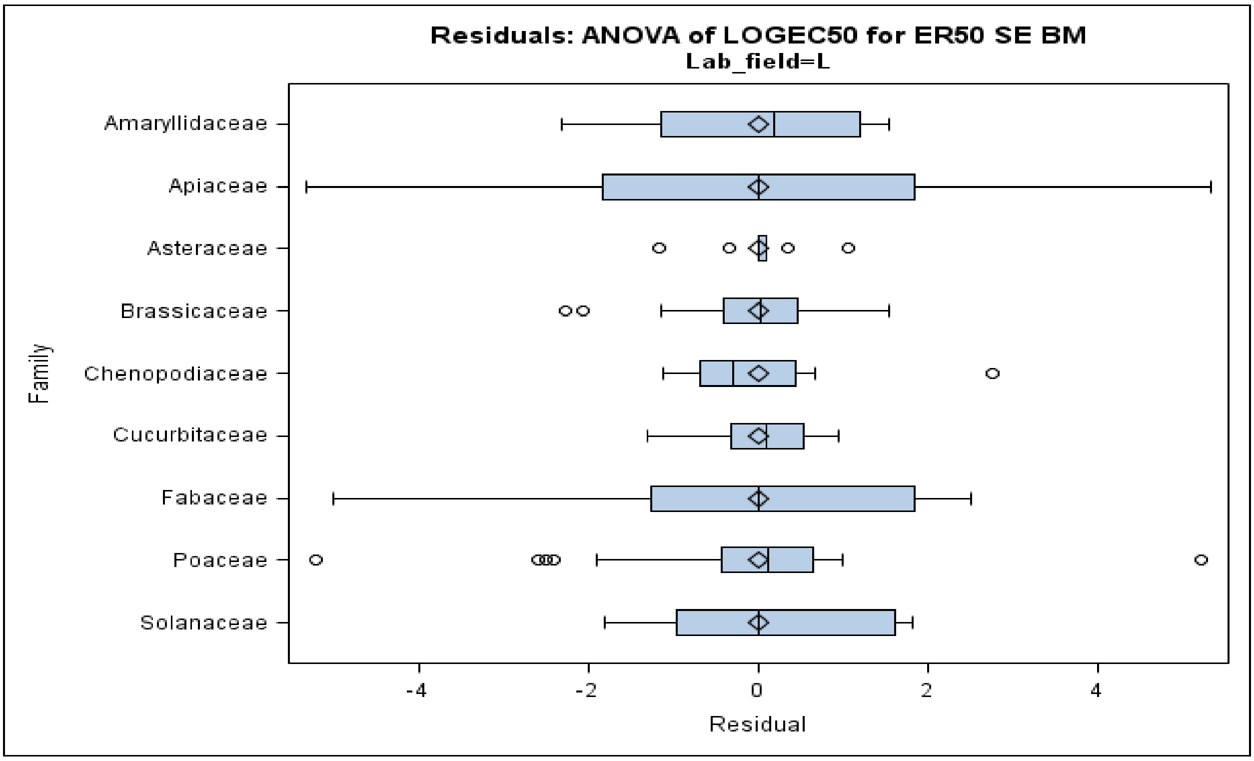
Table 3. Compare Crop and Wild Adjusted Mean SE BM EC50 Values

Family	Lab_field	MoA_CODE_	ProbF	cmean	wmean	ratiocw
Poaceae	L	AASI	0.011	4.561	1.995	2.29
Poaceae	L	PHI	0.041	4.322	7.516	0.58
Poaceae	L	SGI	0.005	8.508	4.359	1.95

Only comparisons where the difference between crop and wild ER50 values is significant are shown.

For family Poaceae with Mode of action=AASI, all four wild species ER50 values are from *Lolium* sp. Those values lie below the crop values, most of which are right-censored. All of the values are relatively high and may not drive risk assessment. For mode of action =PHI, there were only two wild values. One of those was right-censored and its estimated value is extremely large. The only non-censored value was from *Lolium* sp. and its value was much lower than the crop values. Though as in other instances, this ER50 is high enough that it may not drive the risk assessment. For mode of action =SGI, there were only two wild values, both *Lolium* sp. and both considerably lower than the crop values. Three of the five crop values were right-censored but all were higher than both wild values.

Figure 13. Residuals from ANOVA of SE Biomass ANOVA of Log(1+EC50)



Variability in the Asteraceae family was much less than in the other families due to a very small dataset. However, no other family was analyzed together with this one, so there is no problem with the analysis. Otherwise, variability in SE BM ER50 values is similar across families.

ER25 SE BM

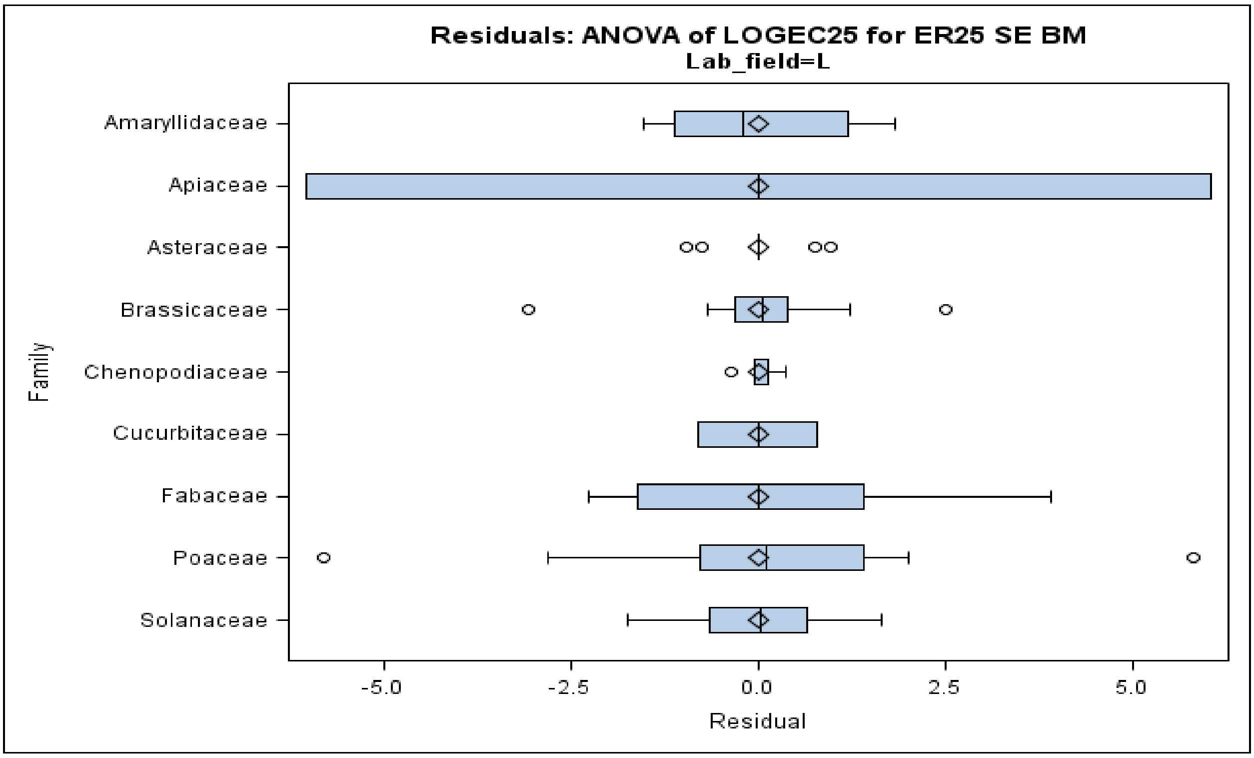
Table 4. Compare Crop and Wild Adjusted Mean SE BM EC25 Values

Family	Lab_field	MoA_CODE_	ProbF	cmean	wmean	ratiocw
Poaceae	F	O	0.002	5.631	3.103	1.81
Poaceae	L	SGI	0.093	7.625	3.631	2.10

Only comparisons where the difference between crop and wild ER25 values is significant are shown.

For Poaceae with MoA=O, there were six wild species SE BM ER25 values, all of which were 4 to 20 times smaller than all three of the crop values. None of these were estimates from censored data, and no two values of either type were from the same species. For MoA=SGI, there were two *Lolium* sp. Wild SE BM ER25 values, both of which were notably smaller than the four crop values.

Figure 14. Residuals from ANOVA of SE Biomass ANOVA of Log(1+EC25)



ER10 SE BM

There were no wild species SE BM ER10 values available, so no comparisons could be made.

Summary for biomass

Based on the preceding analyses that show greater variability in ER50 values, it is concluded that ER25 is a more reliable endpoint than ER50 and crop species values are as low as wild species values, with the possible exception of the Poaceae family. There appears to be sufficient justification for inclusion of a wild species from this family and one or more such species are currently typically included in both vegetative vigor and seedling emergence studies. There is too little data to determine conclusively whether ER10 would be a reliable endpoint, but such data as is available offers no reason to favor its use.

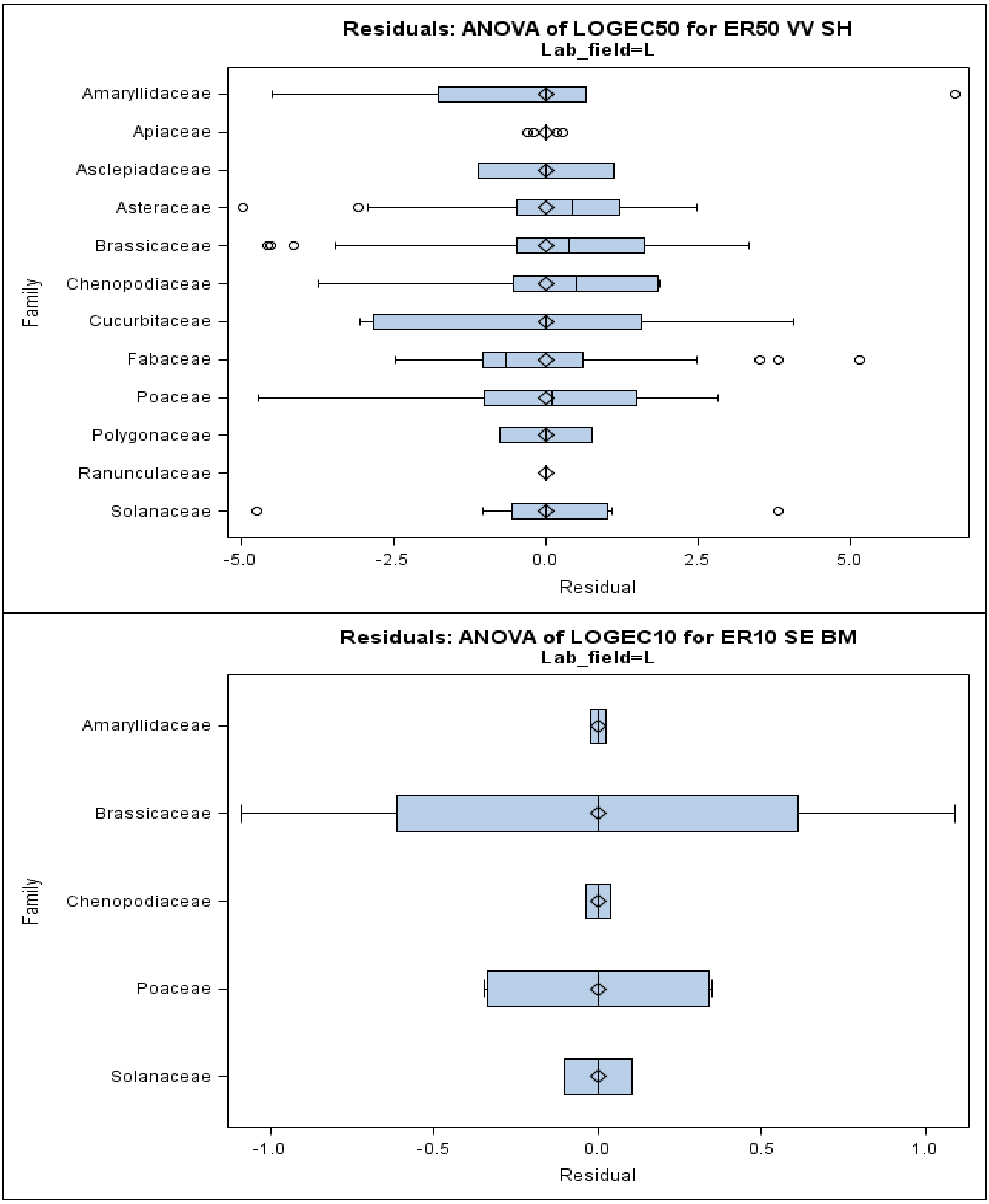
ER50 VV SH

Table 5. Compare Crop and Wild Adjusted Mean EC50 Values

Family	Lab_field	MoA_CODE_	ProbF	cmean	wmean	ratiocw
Apiaceae	L	AASI	0.082	5.860	5.384	1.09
Asteraceae	L	AASI	0.106	5.682	4.302	1.32
Poaceae	L	AASI	0.104	4.378	5.521	0.79
Poaceae	L	PHI	0.963	7.687	7.619	1.01

The difference in VV SH ER50 values between wild and crop species was small in every family. For only one family, Apiaceae, and one MoA, AASI, were these differences significant. As Figure 15 indicates, variability of lab values among Apiaceae species ER50 values was low, perhaps in part because the two crop species and the two wild species were variations on a single species, *Daucus carota* (either labelled as ‘garden carrot’ or as ‘wild carrot’). The small variance accounts for the small difference between crop and wild species being statistically significant.

Figure 15. Residuals from ANOVA of VV Shoot Height ANOVA of Log(1+EC50)



There were significant differences in ER50 values between crop and wild plants for the family Asteraceae, with the crop values larger on average than the wild values. With one exception, the high crop values were right-censored and all are from lettuce (*Lactuca sativa*). Other crop values were similar to those for wild species.

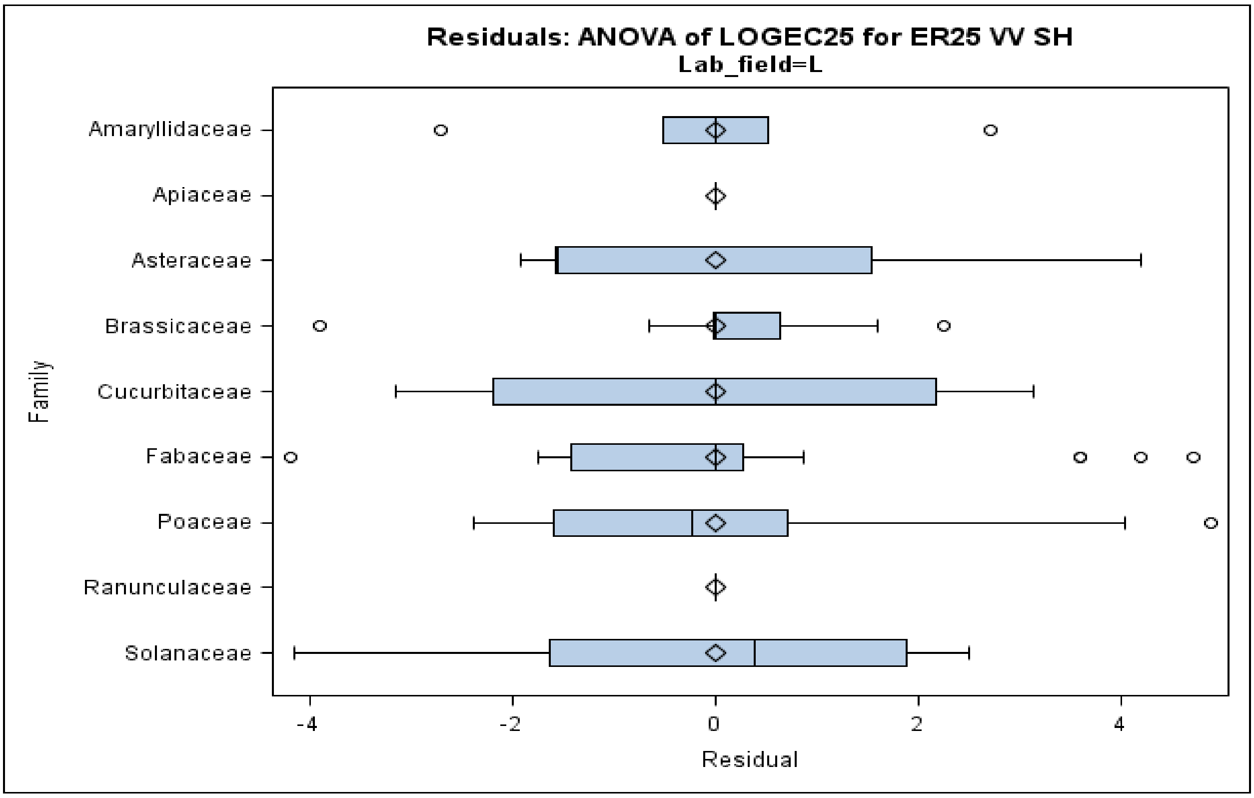
ER25 VV SH

Table 5. Compare Crop and Wild Adjusted Mean VV SH EC25 Values

Family	Lab_field	MoA_CODE_	ProbF	cmean	wmean	ratiocw
Poaceae	L	AASI	0.805	2.381	2.086	1.14
Poaceae	L	PHI	0.719	6.817	7.498	0.91

No significant difference was found between crop and wild species VV SH ER25 values.

Figure 16. Residuals from ANOVA of VV Shoot Height ANOVA of Log(1+EC25)



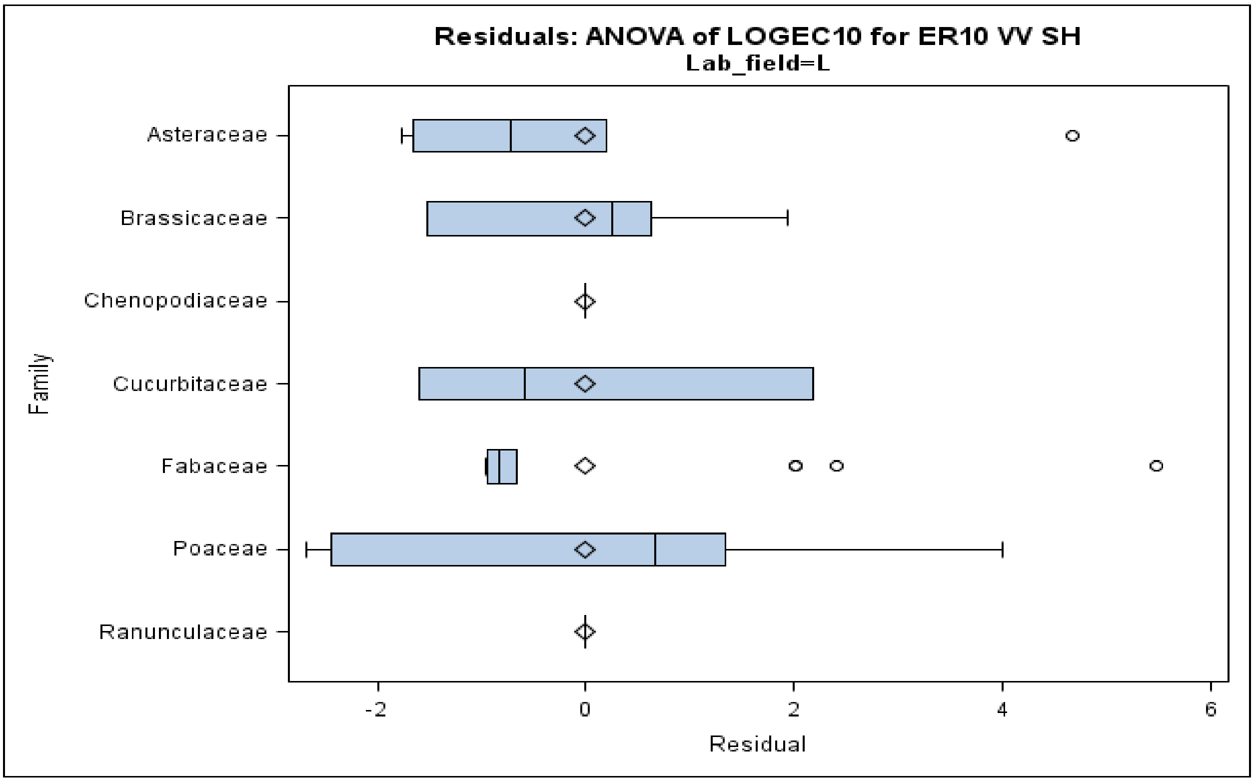
ER10 VV SH

Table 6. Compare Crop and Wild Adjusted Mean VV SH EC10 Values

Family	Lab_field	MoA_CODE_	ProbF	cmean	wmean	ratiocw
Poaceae	L	AASI	0.839	2.690	2.440	1.10

No significant difference was found between the crop and wild species VV SH ER10 values. However, only one combination of family and mode of action was included in the database.

Figure 17. Residuals from ANOVA of VV Shoot Height ANOVA of Log(1+EC10)



ER50 SE SH
There were only crop species tested and only in the lab, so no comparisons of wild and crop were possible.

ER25 SE SH
There were only crop species tested and only in the lab, so no comparisons of wild and crop were possible.

ER10 SE SH
No shoot height ER10 values from seedling emergence studies were in the database.

References

Helsel DR. 2005. Nondetects and Data Analysis. Statistics for Censored Environmental Data. Wiley-Interscience.

Kon Kam King, G., P. Veber, S. Charles, and M.L. Delignette-Muller (2013). MOSAIC_SSD: a new web-tool for the Species Sensitivity Distribution, allowing to include censored data by maximum Likelihood. http://arxiv.org/ftp/arxiv/papers/1311/1311.5772.pdf

Lee, ET., JW. Wang 2013. Statistical Methods for Survival Data Analysis 4th Edition. Wiley. ISBN-13: 978-1118095027; ISBN-10: 1118095022

SAS Institute Inc. 2015. SAS® 9.4 SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513-2414. Cary, NC.