

# **Modeling Microbial Regulation of Pesticide Turnover in Soils**

**Dissertation to obtain the doctoral degree of Agricultural Sciences**

**(Dr. sc. agr.)**

**Faculty of Agricultural Sciences**

**University of Hohenheim**

Institute of Soil Science and Land Evaluation

submitted by

*Luciana Chavez Rodriguez*

from Caraz, Ancash, Peru

2021

This thesis was accepted as a doctoral dissertation in fulfillment of the requirements for the degree “Doktor der Agrarwissenschaften” by the Faculty of Agricultural Sciences at the University of Hohenheim on 10 of July 2021.

Date of examination: 15 of September 2021

Dean of Faculty: Prof. Dr. Ralf Vögele

**Examination committee**

Head of the committee: Prof. Dr. Andrea Knierim

1<sup>st</sup> examiner and reviewer: Prof. Dr. Thilo Streck

2<sup>nd</sup> examiner and reviewer: Prof. Brian P. Ingalls

3<sup>rd</sup> examiner: Prof. Dr. Ellen Kandeler

## Contents

<b>1</b>	<b>Summary</b>	<b>1</b>
<b>2</b>	<b>Zusammenfassung</b>	<b>3</b>
<b>3</b>	<b>General Introduction</b>	<b>7</b>
3.1	Pesticides: application and environmental concerns . . . . .	7
3.2	Fate of pesticides in the environment . . . . .	8
3.3	How to assess pesticide fate in the environment . . . . .	10
3.4	Model data integration using omics and isotopic data . . . . .	12
3.5	Model uncertainty and sensitivity analysis . . . . .	13
3.6	Model-based optimal design of experiments . . . . .	14
<b>4</b>	<b>Research questions</b>	<b>16</b>
<b>5</b>	<b>Research design</b>	<b>18</b>
5.1	Model assumptions . . . . .	18
5.2	Model pesticides . . . . .	18
5.3	Methods for model uncertainty quantification and sensitivity analysis . . . . .	19
5.4	Thesis outline . . . . .	20
<b>6</b>	<b>Gene-centric model approaches for accurate prediction of pesticide biodegradation in soils (Paper 1)</b>	<b>22</b>
6.1	Abstract . . . . .	23
6.2	Introduction . . . . .	23
6.3	Theory . . . . .	25
6.3.1	Model Structure . . . . .	25
6.3.2	Process formulations . . . . .	26

---

6.4	Materials and Methods . . . . .	32
6.4.1	Model Reduction . . . . .	32
6.4.2	Model calibration . . . . .	33
6.4.3	Model comparison . . . . .	36
6.4.4	Model validation . . . . .	37
6.4.5	Local sensitivity analysis . . . . .	37
6.4.6	Global sensitivity analysis – Morris method . . . . .	37
6.5	Results and discussion . . . . .	38
6.5.1	Calibration, parameterization and model dynamics of full model (V0) . . . . .	38
6.5.2	Model comparison . . . . .	43
6.5.3	Model validation . . . . .	44
6.5.4	Implications for biogeochemical modeling informed by genetic data . . . . .	48
<b>7</b>	<b>Modeling bioavailability limitations of atrazine degradation in soils (Paper 2)</b>	<b>50</b>
7.1	Abstract . . . . .	51
7.2	Introduction . . . . .	51
7.3	Material and Methods . . . . .	53
7.3.1	Model Description . . . . .	53
7.3.2	Process Formulations . . . . .	54
7.3.3	Model calibration . . . . .	59
7.3.4	Soil measurements . . . . .	63
7.3.5	Soil predictions . . . . .	65
7.3.6	Global sensitivity analysis . . . . .	66
7.3.7	Local sensitivity analysis . . . . .	67
7.4	Results . . . . .	67
7.4.1	Calibration to chemostat and retentostat data . . . . .	67
7.4.2	Predictions of atrazine and hydroxyatrazine fate in soils and comparison against field data . . . . .	70



7.5	Discussion . . . . .	73
7.5.1	Bacterial adaption to low nutrient availability affects model parameterization	73
7.5.2	Pesticide persistence in soil . . . . .	75
<b>8</b>	<b>Temperature and soil moisture change microbial allocation of pesticide-derived carbon (Paper 3)</b>	<b>77</b>
8.1	Abstract . . . . .	78
8.2	Introduction . . . . .	78
8.3	Materials and Methods . . . . .	81
8.3.1	Soil origin and sampling . . . . .	81
8.3.2	Experimental Design . . . . .	82
8.3.3	MCPA analysis in soil . . . . .	82
8.3.4	MCPA mineralization ( $^{14}\text{CO}_2$ ) . . . . .	83
8.3.5	Microbial biomass ( $C_{\text{mic}}$ ) . . . . .	84
8.3.6	MCPA degrader abundance and activity . . . . .	85
8.3.7	Gene-centric modeling of MCPA biodegradation . . . . .	86
8.3.8	Carbon use efficiency (CUE) . . . . .	90
8.3.9	Statistical analyses . . . . .	91
8.4	Results and discussion . . . . .	92
8.4.1	Enhanced MCPA mineralization by elevated temperature and moisture	92
8.4.2	Invariable microbial dynamics under limiting temperature and moisture	94
8.4.3	CUE dependency on temperature, moisture, and MCPA concentration	95
8.4.4	Effect of temperature and soil moisture on pesticide $\text{DT}_{50}$ . . . . .	99
<b>9</b>	<b>Optimal design of experiments for effective modeling of atrazine degradation in soils (Paper 4)</b>	<b>102</b>
9.1	Abstract . . . . .	103
9.2	Introduction . . . . .	104

---

9.3	Material and Methods . . . . .	106
9.3.1	Atrazine degradation models . . . . .	106
9.3.2	Prospective optimal design of experiments (OD) . . . . .	113
9.4	Results and Discussion . . . . .	119
9.4.1	Can we distinguish active AT degradation pathways based on observations of metabolite concentrations? . . . . .	119
9.4.2	Which experimental designs provide the most informative data for model discrimination? . . . . .	122
9.4.3	Can measuring pools not commonly measured improve model discrimination? . . . . .	125
9.4.4	Role of NI-NE measurements: . . . . .	127
9.4.5	Implications for biogeochemical modeling and data integration . . . . .	129
<b>10</b>	<b>General Discussion</b>	<b>131</b>
10.1	Improving process understanding and prediction of pesticide degradation in soils . . . . .	131
10.2	Challenges in mechanistic model development and how to overcome them . . . . .	134
10.3	Towards a predictive model: applicability and further extensions of the work . . . . .	135
10.4	Research perspectives: Deep neural networks . . . . .	138
<b>11</b>	<b>Appendices</b>	<b>140</b>
11.1	Supplementary Information for Chapter 6 (Paper 1) . . . . .	140
11.1.1	Model Formulations for reduced Models . . . . .	140
11.1.2	Methods: Sampling data points . . . . .	144
11.1.3	Methods: Local and Global Sensitivity and uncertainty analysis . . . . .	144
11.1.4	Additional Results . . . . .	148
11.1.5	Residual pesticide (2,4-D and MCPA) concentration in soil . . . . .	151
11.1.6	Local Sensitivity Analysis of reduced model variants . . . . .	153
11.1.7	Global Sensitivity Analysis Results - Morris Method . . . . .	156

11.1.8	Prior and posterior distribution of the model parameters on 2,4-D and MCPA data . . . . .	157
11.1.9	Correlation matrix of model variants . . . . .	173
11.2	Supplementary Information for Chapter 7 (Paper 2) . . . . .	176
11.2.1	Illustration of the degradation rate vs. Substrate concentration for both model variants M and T: . . . . .	176
11.2.2	Determination of correction factor: . . . . .	176
11.2.3	Determination of standard deviation of data $\sigma_i^2$ (eq. 74 of the main text) . . . . .	178
11.2.4	Soil Observations in Poltringen and Tailfingen, Germany . . . . .	180
11.2.5	Seepage pesticide (atrazine and hydroxyatrazine) concentration results . . . . .	182
11.2.6	Calibration to chemostat and retentostat data of alternative Monod-Model . . . . .	184
11.2.7	Correlation tables . . . . .	185
11.2.8	Global sensitivity results . . . . .	187
11.2.9	Local sensitivity analysis . . . . .	191
11.2.10	Alternative model structure for engineered systems . . . . .	191
11.3	Supplementary Information for Chapter 8 (Paper 3) . . . . .	196
11.3.1	Main characteristics of the sampled soil . . . . .	196
11.3.2	MCPA degrader abundance and activity . . . . .	196
11.3.3	Model parameters . . . . .	197
11.3.4	$^{14}\text{C}$ incorporation to $^{14}\text{CO}_2$ respiration . . . . .	200
11.3.5	$CUE_C$ . . . . .	205
11.4	Supplementary Information for Chapter 9 (Paper 4) . . . . .	206
11.4.1	Model scenario descriptions . . . . .	206
11.4.2	Modified constraint-based parameter search algorithm - Sampling results . . . . .	214
11.4.3	Energy distance (ED) robustness results . . . . .	230
	<b>References</b>	<b>231</b>

<b>12 Acknowledgements</b>	<b>282</b>
<b>13 Curriculum Vitae</b>	<b>285</b>
<b>14 Declaration</b>	<b>286</b>

## List of Figures

1	Transfer/transport fate of pesticide in the environment. 1: volatilization into the air, 2: Surface and subsurface runoff, 3: Leaching into groundwater, 4: plant uptake, 5: Sorption onto soil particles, 6: wet and dry deposition, 7: biological transformations . . . . .	9
2	“Optimal pesticide degradation” scenario. In this scenario, bacteria and pesticide co-occur in the same spots, facilitating degradation. Water transport in soil, soil spatial heterogeneity and additional carbon sources are not included . . . . .	18
3	Model schematic. Green boxes: independent state variables calculated directly; orange boxes: variables related to processes in quasi-steady state; purple box: additional pool that is modeled indirectly. Solid arrows indicate mass transfer; dashed arrows indicate regulation. . . . .	26
4	Model calibration on 2,4-D data. Time series of pesticide mineralization (A, B), <i>tfdA</i> mRNA copies (transcripts) per g of soil (C, D), <i>tfdA</i> gene copies per g of soil (E, F). Error bars show the standard deviation of the data and of the simulations (based on MCMC ensembles, see Materials and Methods 6.4.2). The inset in panel E shows calibrated model simulations of <i>tfdA</i> DNA, active, inactive and dead bacteria expressed in $\text{mmol g}^{-1}$ (based on 95 % confidence interval of MCMC ensembles, see Materials and Methods 6.4.2). Bold lines represent the mean value of each bacterial pool. . . . .	40

5	Model validation on MCPA data. Times series of pesticide mineralization (A, B), <i>tfdA</i> mRNA copies (transcripts) per g of soil (C, D), <i>tfdA</i> gene copies per g of soil (E, F). Error bars show the standard deviation of the data and of the simulations (based on MCMC ensembles, see Materials and Methods 6.4.2). . . . .	45
6	<i>tfdA</i> mRNA vs. rate of mineralization. Panel A shows reconstructed mineralization rates determined as the first derivatives of the “3/2-exp” and “3/2-lin” models in Bælum et al. 2008 [128] (with parameters as in Table 1 of that reference), and measured <i>tfdA</i> mRNA after 2,4-D application. Two arbitrarily selected relationships from the MCMC ensembles are plotted for full model version V1 (B, C) and parsimonious model version V3 (D, E), showing hysteresis and linear relationships between these two variables. Blue lines indicate the first application of 2,4-D, and the red lines the second application. Black arrows in panels A and B show the direction of the hysteretic curve. . . . .	49
7	Model structure for engineered (chemostat/retentostat) systems (green) and extension for soil (blue). The model explicitly accounts for light “l” and heavy “h” isotopologues ( $^{12}\text{C}/^{13}\text{C}$ ) of AT due to enzymatic transformation in the intracellular “i” and extracellular “e” compartments, as well as in equilibrium sorption in soil “e, S” . . . . .	54
8	Simulations (boxplots) of model variant T (thermodynamic growth constraints) and measured data (blank diamonds + estimated standard deviation). <b>A-C.</b> Steady-state concentrations for the chemostat (five dilution rates: C1-C5: 0.023, 0.032, 0.048, 0.056, 0.068 h <sup>-1</sup> , respectively) and the retentostat (dilution rate: R: 0.020 h <sup>-1</sup> ). The middle line in the boxplot is the median of the ensemble outputs from the MCMC simulation ensemble (see M&M 7.4); boxes represent 25% and 75% percentiles; whiskers corresponds to +/- 1.5 x IQR (interquartile range). <b>D.</b> Enrichment factors ( $\epsilon$ ) were reported only for the lowest dilution rate of the chemostat (C1) and the retentostat (R), but simulated for C2-C5. . . . .	68

9	Posterior distributions for calibrations with chemostat (8 parameters) and retentostat (4 parameters) data. All parameters are expressed in log scale with the exception of the growth yield $Y$ . . . . .	69
10	Simulated residual concentrations of atrazine AT (A,B) and hydroxyatrazine HY (C,D) in topsoils (0-30 cm), and observations (grey background) in topsoils of two field sites Poltringen (P) and Taifingen (T) after 30 years. (C) represent simulations using chemostat fitted parameters, and (R) simulations using retentostat fitted parameters. The middle line of boxplots is the median of the ensemble outputs from the MCMC simulation ensemble (see M&M 7.4); boxes represent 25% and 75% percentiles; whiskers corresponds to $\pm 1.5 \times$ IQR (interquartile range). . . . .	73
11	Measured (dots) and simulated (lines) of cumulative $^{14}\text{CO}_2$ mineralization at two MCPA concentrations as a function of temperature and soil moisture over time (A, B, C, D), Residual MCPA expressed as $\text{mg kg}^{-1}$ over time (E, F, G, H), <i>tfdA</i> genes during the MCPA biodegradation experiment expressed as gene copies $\text{g}^{-1}$ dry weight (I, J, K, L), <i>tfdA</i> transcripts quantities during MCPA degradation expressed as transcripts copies $\text{g}^{-1}$ dry weight (M, N, O, P). <i>tfdA</i> genes and transcripts are expressed at log-scale. Error bars represent standard errors of the mean values for soil triplicates (see M&M). . . . .	95
12	$CUE$ vs. time (d) shown in panels A to D, and $CUE$ vs normalized residual MCPA concentration in soil shown in panels E to H. $CUE_M$ (Eq. 105) is presented as points and $CUE_E$ (Eq. 106) as lines . . . . .	99

- 13 Atrazine (AT) degradation in soils: model framework, describing AT, its intermediate metabolites (in blue), sink pools (in grey), and the bacterial guilds involved in the degradation process. Arrow colors indicate activity of the bacterial guilds. Black arrows represent abiotic processes: degradation of AT and nitrogen-dependent degradation of CA. Dashed line represents the unintended HY leaked out of the degradation carried out by guild A [62, 63]. Dashed-dot line shows a degradation step uncoupled from growth (carried out by guild B);  $f_D$  represented the proportion of DIA formed during AT degradation carried out by guild D . 107
- 14 Pairwise energy distances (expressed in standard deviation units) over the candidate designs. A. Model M1 against other models; B. Model M2 against other models; C. Model M3 against other models; D. Model M4 against other models; E. Model M5 against other models; F. Model M6 against other models. Vertical lines represent transition from one sampling frequency/duration to another: left is daily over 25 days; middle is every second day over 50; right is every four days over 100. The horizontal line represents the selected minimum energy distance threshold for model discrimination (distance of two standard deviations) 121
- 15 Optimal designs from the Pareto front based only on AT, metabolite (HY, DEA, DIA) and sinks (CA, CO<sub>2</sub>) measurements (Y = measured, N = not measured). Time column 's values are: (1) short-term, every day sampling until day 25, (2) middle-term, every two days sampling until day 50, (3) long-term, every four days sampling until day 100 . . . . . 123

16	Pairwise energy distances (expressed in standard deviation units) over the candidate designs including biomass measurements. A. Model M1 against other models; B. Model M3 against other models; C. Model M2 against other models; D. Model M4 against other models. Vertical lines represent transition from one sampling frequency/duration to another: left is daily over 25 days; middle is every second day over 50; right is every four days over 100. The horizontal line represents the selected minimum energy distance threshold for model discrimination (distance of two standard deviations) . . . . .	126
17	Optimal designs including biomass measurements (Biomass of guild D and total Biomass). Y = measured, N = not measured. Time column's values are: (1) short-term, every day sampling until day 25, (2) middle-term, every two days sampling until day 50, (3) long-term, every four days sampling until day 100 .	127
18	Pairwise energy distances (expressed in standard deviation units) over the candidate designs including i) NI: A. Model M1 against other models; B. Model M2 against other models; C. Model M3 against other models; and ii) NE: D. Model M1 against other models; E. Model M2 against other models; F. M4 against other models. Vertical lines represent transition from one sampling frequency/duration to another: left is daily over 25 days; middle is every second day over 50; right is every four days over 100. The horizontal line represents the selected minimum energy distance threshold for model discrimination (distance of two standard deviations) . . . . .	128
19	Modeling microbial regulation of pesticide turnover. Pipeline to develop a predictive model. Steps included in this thesis are marked with red circles . .	137
20	Theoretical pipeline to derive large scale predictions combining machine learning techniques and mechanistic modeling approaches. Adapted from Baker et al. [381]. Further information on Kbase tools can be found in the Kbase website	139



S1	Model calibration on MCPA data. Time series of pesticide mineralization (A, B), <i>tfdA</i> mRNA copies (transcripts) per g of soil (C, D), <i>tfdA</i> DNA gene copies per g of soil (E, F). Error bars show the standard deviation of the data and of the simulations (based on MCMC ensembles, see Material and Methods 6.4.2) . . .	148
S2	<i>tfdA</i> gene dynamics for V0 (A, A1, A2), V1 (B), V2 (C). V1 and V2 only include dynamics of 2,4-D calibration dataset. Panels showed the mean value of the corresponding dynamics (bold line) and the 95% confidence interval (based on MCMC ensembles, see Material and Methods 6.4.2) . . . . .	149
S3	<i>tfdA</i> gene dynamics for V3 (D, D1, D2), V4 (E, E1, E2), V4' (F, F1, F2). Panels showed the mean value of the corresponding dynamics (bold line) and the 95% confidence interval (based on MCMC ensembles, see Material and Methods 6.4.2).	150
S4	Total residual pesticides in soils for 2,4-D and MCPA for models: V0 (A, A1, A2), V1 (B), V2 (C): Panels showed the mean value of the corresponding dynamics (bold line) and the 95% confidence interval (based on MCMC ensembles, see Material and Methods 6.4.2) . . . . .	151
S5	Total residual pesticides in soils for 2,4-D and MCPA for models: V3 (D, D1, D2), V4 (E, E1, E2), V4' (F, F1, F2). Panels showed the mean value of the corresponding dynamics (bold line) and the 95% confidence interval (based on MCMC ensembles, see Material and Methods 6.4.2) . . . . .	152
S6	Morris Method results for V0. X axis shows the model parameters and Y axis shows the outputs chosen for the Morris analysis . . . . .	156
S7	Prior and posterior distribution of parameters (1-9) for V0 with 2,4-D data . .	157
S8	Prior and posterior distribution of parameters (10-17) for V0 with 2,4-D data .	158
S9	Prior and posterior distribution of parameters (1-9) for V0 with MCPA data . .	159
S10	Prior and posterior distribution of parameters (10-17) for V0 with MCPA data	160
S11	Prior and posterior distribution of parameters (1-6) for V3 with 2,4-D data . .	161
S12	Prior and posterior distribution of parameters (7-10) for V3 with 2,4-D data . .	162

S13	Prior and posterior distribution of parameters (1-6) for V3 with with MCPA data	163
S14	Prior and posterior distribution of parameters (7-10) for V3 with with MCPA data	164
S15	Prior and posterior distribution of parameters for V4 with 2,4-D data . . . . .	165
S16	Prior and posterior distribution of parameters for V4 with MCPA data . . . . .	166
S17	Prior and posterior distribution of parameters for V4' with 2,4-D data . . . . .	167
S18	Prior and posterior distribution of parameters for V4' with MCPA data . . . . .	168
S19	Prior and posterior distribution of parameters (1-9) for V1 with 2,4-D data . .	169
S20	Prior and posterior distribution of parameters (10-13) for V1 with 2,4-D data .	170
S21	Prior and posterior distribution of parameters (1-9) for V2 with 2,4-D data . .	171
S22	Prior and posterior distribution of parameters (10-15) for V2 with 2,4-D data .	172
S23	Correlation matrix for full model V0. . . . .	173
S24	Correlation matrix for the model variant V1. . . . .	173
S25	Correlation matrix for the model variant V2. . . . .	174
S26	Correlation matrix for the model variant V3. . . . .	174
S27	Correlation matrix for the model variant V4. . . . .	175
S28	Correlation matrix for the model variant V4'. . . . .	175
S29	Illustration of the degradation rate vs. Substrate concentration for both model variants M (1) and T (2) at high concentrations and low concentration (inset), for $k_{HY} = 71 \text{ d}^{-1}$ and $K_M^{HY} = 20 \text{ } \mu\text{g L}^{-1}$ . . . . .	176
S30	Sorption fitting for both study sites . . . . .	182
S31	Seepage water flux at A. Poltringen and B. Tailfingen at 50cm depth measured at 2019.	183

S32	Simulations (boxplots) using model variant M (including Monod-kinetics) and measurements (blank diamonds + estimated standard deviation). <b>A-C.</b> Steady-state concentrations for the chemostat (for five dilution rates: C1-C5: 0.023, 0.032, 0.048, 0.056, 0.068 h <sup>-1</sup> ) and the retentostat (dilution rate: R: 0.020 h <sup>-1</sup> ). The middle line in the boxplot is the median of the ensemble outputs from the MCMC (see M&M 7.4); boxes represent 25% and 75% percentiles, and whiskers corresponds to +/- 1.5 x IQR (interquartile range). <b>D.</b> Enrichment factors ( $\epsilon$ ), reported only for the lowest dilution rate of the chemostat (C1) and retentostat (R), but simulated for C2-C5. . . . .	184
S33	Posterior distributions for calibrations with chemostat (8 parameters) and retentostat (4 parameters) data. All parameters are expressed n log scale with the exception of the parameter Y. . . . .	185
S34	$l_2$ - norm values from Morris Method [73, 105] for thermodynamic and Monod Model . . . . .	187
S35	Sobol Indices [202]: 1: Final Biomass, 2 Final AT in the system:, 3 Final AT inside the cell:, 4: Final HY in the system, 5: Final HY inside the cell, 6: Enrichment Factor $\epsilon$ fit. . . . .	188
S36	Alternative model structure for engineered systems, including a pool of growing and non-growing bacteria. . . . .	191
S37	Chemostat and Retentostat fit: Atrazine, hydroxyatrazine and biomass . . . . .	195
S38	Chemostat and Retentostat fit: Enrichment factor. Dashed line indicate the enrichment value for chemostat -5.4‰ and retentostat -0.45‰ . . . . .	195
S39	Cumulative <sup>14</sup> CO <sub>2</sub> mineralization of two MCPA concentrations as a function of soil temperature and soil moisture over time. Mineralization of MCPA is represented by the percentage of initial <sup>14</sup> C-MCPA. Curves were fitted to the data points via a logistic model. . . . .	200

---

S40	$CUE_C$ (eq. 107) vs. time (d) showed in panels A to D, and $CUE_C$ vs normalized residual MCPA concentration in soils showed in panels E to H . . . . .	205
S41	Sampled parameters for model scenario M1 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results . . . . .	214
S42	Sampled parameters (continuation) for model scenario M1 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results	215
S43	Sampled parameters for model scenario M2 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results . . . . .	216
S44	Sampled parameters (continuation) for model scenario M2 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results	217
S45	Sampled parameters for model scenario M3 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results . . . . .	218
S46	Sampled parameters (continuation) for model scenario M3 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results	219

---

S47	Sampled parameters for model scenario M4 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of three independent runs of the sampling algorithm (pink, yellow and purple). Both independent runs show similar sampling results . . .	220
S48	Sampled parameters (continuation) for model scenario M4 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of three independent runs of the sampling algorithm (pink, yellow and purple). Both independent runs show similar sampling results . . . . .	221
S49	Sampled parameters for model scenario M5 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results . . . . .	222
S50	Sampled parameters (continuation) for model scenario M5 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results	223
S51	Sampled parameters for model scenario M6 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results . . . . .	223
S52	Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO <sub>2</sub> )) presented as spaghetti plots for model scenario M1. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm. Pools of DIA and DEA lie under the detection limit of $1 \cdot 10^{-7}$ mg cm <sup>-3</sup> . . . . .	224

S53	Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO <sub>2</sub> )) presented as spaghetti plots for model scenario M2. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm. Pools of DIA and DEA lie under the detection limit of $1 \cdot 10^{-7}$ mg cm <sup>-3</sup> . . . . .	225
S54	Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO <sub>2</sub> )) presented as spaghetti plots for model scenario M3. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm. Pools of DIA and DEA lie under the detection limit of $1 \cdot 10^{-7}$ mg cm <sup>-3</sup> . . . . .	226
S55	Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO <sub>2</sub> )) presented as spaghetti plots for model scenario M4. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm. Pools of DIA and DEA lie under the detection limit of $1 \cdot 10^{-7}$ mg cm <sup>-3</sup> . . . . .	227
S56	Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO <sub>2</sub> )) presented as spaghetti plots for model scenario M5. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm . . . . .	228
S57	Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO <sub>2</sub> )) presented as spaghetti plots for model scenario M6. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm . . . . .	229
S58	Robustness test. Energy distance values per design candidates, including AT, metabolites (HY, DEA, DIA), and sinks (CA, CO <sub>2</sub> ) for the different ensemble sizes. Ensemble size of 100 differs from the bigger ensemble sizes, but an ensemble size of 1,000, 5,000, 10,000 seems to produce similar results . . . . .	230

## List of Tables

1	Initial parameter ranges for model calibration. . . . .	34
2	Uncertainty analysis for the full model variant (V0) on 2,4-D data. . . . .	41
3	Model comparison based on the 2,4-D data. . . . .	44
4	Simulated 2,4-D/MCPA mean residual concentration in soil in $\text{mg}_C \text{kg}^{-1}$ (total concentration at the end of the experiment). . . . .	47
5	Model Parameters . . . . .	61
6	Details of accelerated solvent extraction method . . . . .	64
7	Soil Parameters . . . . .	66
8	Calibrated parameter values of the model variant T based on an MCMC ensemble for Chemostat and Retentostat fits . . . . .	71
9	Half-life $\text{DT}_{50\text{RES}}$ derived from the residual MCPA concentration in soils and $\text{DT}_{50\text{MIN}}$ derived from mineralization kinetics as a function of soil water content, concentration, and temperature . . . . .	100
10	Scenario models of AT degradation in soils . . . . .	113
11	Model Parameters for Monod and first-order kinetic models . . . . .	117
12	Single elements of the energy distance (ED) between model M1 (X) and M6 (Y) . . . . .	124
13	Optimal designs after multiobjective pareto analysis of candidate designs including NI and NE in addition to AT, metabolites (HY, DIA, DEA) and sinks (CA, $\text{CO}_2$ ) . . . . .	129
S1	List of sampling points batch degradation experiments . . . . .	144
S2	Uncertainty analysis for V1 for 2,4 D data. . . . .	153
S3	Uncertainty analysis for V2 for 2,4 D data. . . . .	154
S4	Uncertainty analysis for V3 for 2,4 D data. . . . .	155
S5	Uncertainty analysis for V4 for 2,4 D data. . . . .	155
S6	Uncertainty analysis for V4' for 2,4 D data. . . . .	156
S7	Recalculated standard deviations for chemostat and retentostat calibration based on [63, 238] . . . . .	179

---

S8	Pesticide Inventory - Atrazine and Hydroxyatrazine in soils . . . . .	180
S9	Sorption test results . . . . .	181
S11	Parameter correlation for model variant T data using MCMC ensemble . . . . .	185
S10	Calibrated parameter values of the model variant M based on a MCMC ensemble for Chemostat and Retentostat fits . . . . .	186
S12	Parameter correlation for model variant M data using MCMC ensemble . . . . .	187
S13	Local sensitivity scores for soil model variants of residual AT in soils . . . . .	189
S14	Local sensitivity scores for soil model variants of residual HY in soils . . . . .	190
S15	Chemical and physical soil properties . . . . .	196
S16	Reverse transcription . . . . .	196
S17	Digestion . . . . .	197
S19	Model Parameters . . . . .	197
S18	Primer sequence . . . . .	198
S20	Calibrated model parameters . . . . .	199
S21	Contrast of the estimated marginal means of mineralization on day 28 as a function of temperature (the contrast function setting interaction = "tukey") . . . . .	200
S22	Interaction contrast of the estimated marginal means of mineralization on day 28 as a function of temperature and MCPA concentration (the contrast function setting interaction = "tukey") . . . . .	201
S23	Contrast of the estimated marginal means of mineralization on day 28 as a function of soil moisture (the contrast function setting interaction = "tukey") . . . . .	201
S24	Interaction contrast of the estimated marginal means of mineralization on day 28 as a function of soil moisture and MCPA concentration (the contrast function setting interaction = "tukey") . . . . .	201
S25	Interaction contrast of the estimated marginal means of <i>tfdA</i> copies $g^{-1}$ as a function of temperature, MCPA concentration, time and soil moisture (the contrast function setting interaction = "tukey") . . . . .	202



---

S26	Interaction contrast of the estimated marginal means of <i>tfdA</i> gene transcript as a function of temperature, MCPA concentration, time and soil moisture (the contrast function setting interaction = "tukey") . . . . .	203
S27	Contrast of the estimated marginal means of <sup>14</sup> C incorporation as a function of soil moisture (the contrast function setting interaction ="tukey") . . . . .	204
S28	Contrast of the estimated marginal means of <sup>14</sup> C incorporation as a function of temperature (the contrast function setting interaction ="tukey") . . . . .	204
S29	Contrast of the estimated marginal means of $CUE_M$ as a function of MCPA concentration (the contrast function setting interaction = "tukey") . . . . .	204
S30	Contrast of the estimated marginal means of $CUE_M$ as a function of temperature (the contrast function setting interaction = "tukey") . . . . .	204
S31	Contrast of the estimated marginal means of $CUE_M$ as a function of soil moisture (the contrast function setting interaction = "tukey") . . . . .	205

## 1 Summary

Pesticides are widely used for pest control in agriculture. Besides their intended use, their long-term fate in real systems is not well understood. They may persist in soils, thereby altering ecosystem functioning and ultimately affecting human health. Pesticide fate is assessed through dissipation experiments in the laboratory or the field. While field experiments provide a close representation of real systems, they are often costly and can be influenced by many unknown or uncontrollable variables. Laboratory experiments, on the other hand, are cheaper and have good control over the governing variables, but due to simplification, extrapolation of the results to real systems can be limited. Mechanistic models are a powerful tool to connect lab and field data and help us to improve our process understanding. Therefore, I used mechanistic, process-based models to assess key microbial regulations of pesticide degradation. I tested my model hypotheses with two pesticide classes: i) chlorophenoxy herbicides (MCPA (2-methyl-4-chlorophenoxyacetic acid) and 2,4-D (2,4-Dichlorophenoxyacetic acid)), and ii) triazines (atrazine (AT)), in an ideal scenario, where bacterial degraders and pesticides are co-localized. This thesis explores some potential controls of pesticide degradation in soils: i) regulated gene expression, ii) mass-transfer process across the bacterial cell membranes, iii) bioenergetic constraints, and iv) environmental factors (soil temperature and moisture).

The first part of this thesis describes a set of gene-centric models that explicitly incorporate gene and enzyme expression. I calibrated and validated the model variants with data from two batch experiments of 2,4-D and MCPA pesticide degradation. I compared the performance of the model variants against each other and a standard Monod model. Results highlight that regulated gene expression controls 2,4-D and MCPA degradation in soils. The novel gene-centric models predict pesticide mineralization as good as the standard Monod model, but additionally account for pesticide-triggered gene regulation, allowing us to better capture microbial dynamics during pesticide mineralization. This way, the gene-centric models could be used to explore the relationship between transcription of functional genes and process rates, thereby offering an advantage over the standard Monod model.

The second part of this thesis extends and improves existing chemostat/retentostat models to evaluate the role of biophysical limitations (mass transfer across the cell membrane) and bioenergetic growth constraints of pesticide degradation, both hypothesized to be responsible for pesticide persistence at low concentrations in real systems. The target pesticide was atrazine, which is highly persistent. Results point out that sorption-limited bioavailability could explain the long-term fate and persistence of the main degradation metabolite hydroxyatrazine. However, my model overestimated the long-term biodegradation of atrazine in soils, indicating that more processes than bioavailability are regulating atrazine degradation.

The third part of this thesis explores the role of environmental factors (soil temperature, soil moisture, and substrate concentration) for the fate of pesticides. Through a combination of lab experiments and modeling, MCPA degradation was investigated under different soil temperature (10°C and 20°C) and moisture (pF 1.8 and 3.5) regimes, and substrate concentrations (1 and 20 mg kg<sup>-1</sup>). Results show that microbial degrader populations degrade the pesticide even in colder and drier soils and at low substrate concentrations. By measuring and simulating a higher carbon use efficiency (CUE), I could confirm that microbial degraders are able to cope with such limiting conditions by allocating more carbon to their biomass as a result of potential physiological adaptation. Therefore, extreme environmental conditions do not explain pesticide persistence in soils.

The models presented in this thesis show that including microbial regulations improves predictions of pesticide degradation, compared to conventional models based on Monod kinetics. The gene-centric models achieved a better representation of microbial dynamics and enable us to explore the relationship between functional genes and process rates, and the models that used transition state theory to account for bioenergetic constraints improved the description of degradation at low concentrations. However, the lack of informative data for the validation of model processes hampered model development. Therefore, in the fourth part of this thesis, I used atrazine with its rather complex degradation pathway to apply a prospective optimal design method to find the optimal experimental designs to enable us identifying the degradation

pathway present in a given environment. The optimal designs found suggest to prioritize determining metabolites and biomass of specific degraders, which are not typically measured in environmental fate studies. These data will lead to more robust model formulations for risk assessment and decision-making.

With this thesis, I revealed important regulations of pesticide degradation in soils that help to improve process understanding and model predictions. I provided simple model formulations, for example the Hill function for gene expression and transition state theory for bioenergetic growth constraints, which can easily be integrated into biogeochemical models. My thesis covers initial but essential steps towards a predictive pesticide degradation model usable for risk assessment and decision-making. I also discuss implication for further research, in particular how mechanistic process-based modeling could be combined with new technologies like omics and machine learning.

## **2 Zusammenfassung**

Pestizide sind weit verbreitet in der landwirtschaftlichen Schädlingsbekämpfung. Anders als ihre Wirkungsweise, ist ihr Langzeitverbleib in der Umwelt nicht gut verstanden. Sie gelangen in den Boden und können sich dort anreichern und die Bodenfunktionen beeinträchtigen und letztendlich auch die menschliche Gesundheit gefährden. Die Ausbreitung von Pestiziden wird anhand von Abbauprobungen in Labor- und Feldexperimenten ermittelt. Feldexperimente bieten ein relativ genaues Abbild natürlicher Systeme, sind jedoch meist teuer und können durch unbekannte oder nicht kontrollierbare Faktoren stark beeinflusst werden. Laborexperimente sind in dieser Hinsicht kostengünstiger und bieten eine gute Kontrolle der einwirkenden Faktoren. Allerdings lassen sich die Ergebnisse nur begrenzt auf natürliche Systeme übertragen. Mechanistische Modelle sind ein mächtiges Werkzeug, um Labor- und Felddaten zusammenzuführen und helfen uns dabei, die mikrobiellen Regulationsmechanismen des Pestizidabbaus im Boden besser zu verstehen. Aus diesem Grund habe ich mechanistische, prozess basierte Modelle eingesetzt. Ich habe meine Modellhypothesen bei zwei Pestizidgruppen

getestet: i) Chlorphenoxyherbiziden (MCPA (2-Methyl-4-chlorphenoxyessigsäure) und 2,4-D (2,4-Dichlorphenoxyessigsäure)) und ii) Triazinen (Atrazin (AT)), in einem Idealszenario, wo bakterielle Abbauer und Pestizid kolokalisiert auftreten. Meine Doktorarbeit konzentriert sich auf einige der potenziellen Kontrollmechanismen des Pestizidabbaus im Boden: i) regulierte Genexpression, ii) Massentransferprozesse durch die Zellmembran, iii) bioenergetische Limitierungen und iv) Umweltfaktoren (Bodentemperatur und Bodenfeuchte).

Der erste Teil dieser Doktorarbeit beschreibt eine Reihe Modelle, die explizit Gen- und Enzymexpression beinhalten. Kalibriert und validiert habe ich die Modellvarianten mit Daten aus zwei Batch-Experimenten über Pestizidabbau von 2,4-D und MCPA. Ich verglich die Leistungsfähigkeit der Modellvarianten gegeneinander und gegenüber einem herkömmlichen Monod-Modell. Die Ergebnisse zeigen, dass die Genexpression den Abbau von 2,4-D und MCPA reguliert. Die neuartigen gen-basierten Modelle sagen die Pestizidmineralisierung ebenso gut voraus wie ein herkömmliches Monod-Modell. Darüber hinaus sorgt die Berücksichtigung einer pestizidabhängigen Genregulierung dafür, die mikrobielle Dynamik während der Pestizidmineralisierung besser widerzuspiegeln.

Der zweite Teil dieser Doktorarbeit erweitert und verbessert bestehende Chemostat-/Retentostat-Modelle, um zu evaluieren, welche Rolle biophysikalische Limitierungen (Massentransfer durch die Zellmembran) und bioenergetische Wachstumslimitierungen beim Pestizidabbau spielen, da beide vermutet wird, dass sie für Pestizidpersistenz verantwortlich zu sein. Das untersuchte Pestizid war Atrazin, das recht persistent ist. Die Ergebnisse zeigen, dass die sorptionslimitierte Bioverfügbarkeit das Langzeitverhalten und die Persistenz des Hauptmetaboliten Hydroxyatrazin erklären konnten. Jedoch überschätzte das Modell den biologischen Langzeitabbau von Atrazin, was darauf hinweist, dass noch weitere Prozesse den Atrazinabbau regulieren.

Der dritte Teil dieser Doktorarbeit untersucht die Rolle von Umweltfaktoren (Bodentemperatur, Bodenfeuchte und Substratkonzentration) auf den Abbau von MCPA bei verschiedenen Bodentemperatur- (10°C und 20°C) und Bodenfeuchtereimen (pF 1,8 und pF 3,5)

und Substratkonzentrationen (1 und 20 mg kg<sup>-1</sup>). In Laborexperimenten und Simulationen zeigte sich dass Populationen mikrobieller Abbauer auch in kälteren und trockeneren Böden und unter geringen Substratkonzentrationen MCPA abbauen. Durch die Messung und Simulation einer höheren Kohlenstoffnutzungseffizienz (CUE, engl.: carbon use efficiency) konnten wir bestätigen, dass mikrobielle Abbauer unter limitierenden Bedingungen auf Stress reagieren, indem sie mehr Kohlenstoff in ihre Biomasse verlagern. Aus diesem Grund erklären extreme Umweltfaktoren nicht die Persistenz von Pestiziden im Boden.

Die in dieser Doktorarbeit vorgestellten Modelle zeigen, dass die Berücksichtigung mikrobieller Regulationen Vorhersagen des Pestizidabbaus verbessert, gegenüber herkömmlichen, auf Monod-Kinetik-basierenden Modellen. Die gen-basierten Modelle erreichten eine bessere Repräsentation der mikrobiellen Dynamik und geben uns die Möglichkeit, den Zusammenhang zwischen funktionellen Genen und Prozessraten herzustellen, wohingegen Modelle, die die Abbaugeschwindigkeit auf Grundlage der Theorie des Übergangszustandes limitieren, eine genauere Konzentrationen liefern. Der Mangel an Messdaten zur Validierung behinderte allerdings die Modellentwicklung. Daher benutzte ich im vierten Teil dieser Arbeit, am Beispiel von Atrazin, mit seinem eher komplexen Abbauweg, eine Methode des prospective optimal design, um das bestmögliche Experimentaldesign zu finden, mit dem wir den in einer bestimmten Umgebung vorherrschenden Abbauweg identifizieren können. Die gefundenen optimalen Designs weisen auf die Erfordernis hin, die Messung von Hauptmetaboliten und Biomasse von spezifischen Abbauern zu priorisieren, welche in Abbauversuchen typischerweise nicht gemessen werden. Die Informationen aus diesen Daten werden zu besseren Modellformulierungen führen, die sich für Risikoabschätzung und Entscheidungsfindung nutzen lassen.

Mit dieser Doktorarbeit konnte ich für den Pestizidabbau im Boden wichtige Regulationsmechanismen aufdecken, und so, unser Verständnis und Vorhersagen solcher Prozesse verbessern. Ich stelle einfache Modellformulierungen bereit, beispielsweise die Hill-Funktion für Genexpression und eine Implementierung der Theorie des Übergangszustands,

welche sich einfach in biogeochemische Modelle integrieren lassen. Meine Arbeit liefert grundlegende und entscheidende Schritte zur Entwicklung eines Vorhersagemodells für den Pestizidabbau und dessen Einsatz in Risikoabschätzung und Entscheidungsfindung. Darüber hinaus gebe ich einen Ausblick auf weiterführende Forschungsansätze, insbesondere wie sich mechanistische, prozess-basierte Modellansätze mit neuen Technologien wie omics und Machine Learning verbinden lassen könnten.

## 3 General Introduction

### 3.1 Pesticides: application and environmental concerns

The increasing pressure for food supply worldwide has led to the intensification of agriculture and the search for ways to increase yields on limited arable land [1]. In this context, pesticides have become a relevant component of modern agriculture aiming to enhance productivity by reducing potential plant stressors such as pests, weeds or diseases [2–5]. Besides agriculture, their use also extends to public health programs as a way to deal with vector-borne diseases [1, 6].

Typically, pesticides can be classified according to the target pests (fungicides, bactericides, herbicides, acaricides), the chemical composition (organochlorines, organophosphates, carbamates, pyrethroids, etc.), and mode of entry into the target pest (systemic, contact, stomach poisoning, fumigants, etc.) [7]. Recently, the World Health Organization has recommended [8] that pesticides be classified according to hazard: from unlikely hazard, to present acute hazard, to extremely hazardous.

Over the last thirty years, pesticide application has increased [9]. Estimations show an annual use of pesticides of about 2 million tons per year, with projections to surpass 3.5 million tons in the coming years [10]. Moreover, high-income countries apply more pesticides, and the predominance of China and the United States as the major pesticide users seems to validate this statement [11]. However, the increment in application does not necessarily translate to an increase in productivity [12], as high pesticide applications can negatively impact non-targeted organisms [13–15], and ecosystems [16].

First signs of pesticide impact on non-target organisms were reported in the book *Silent Spring* [17] published in 1962, alerting the world about potential effects of the pesticide DDT on wildlife, especially birds and bees. Studies have also suggested negative effects on humans ranging from endocrine [18, 19] and reproductive effects [20, 21] to even cancer [22, 23]. Therefore, pesticide use in the EU is increasingly regulated [24], and, in some cases, led to the



removal of some pesticides from the market [25]. New movements towards a more holistic and environmental friendly management, as well as organic farming are arising [26] as a response to the overuse of pesticides in modern agriculture. This involves biological pest control as well as the use of alternative natural and biodegradable pesticides [27–29]. Nevertheless, as pesticides are still used, research focused on the effects of pesticides continues to be relevant as well.

### **3.2 Fate of pesticides in the environment**

The environmental fate of pesticides is driven by a complex interplay of microbial and physicochemical processes as well as physicochemical properties of the pesticides in soil and water, influenced by environmental conditions [30]. Pesticides undergo different degradation pathways in the environment. During their lifetime, pesticides are subject to transfer/transport and transformation processes [31].

Transfer/transport processes (Figure 1) control the movement of pesticides through the three main environmental compartments (air, water, and soils), and how they eventually end up in living organisms (plants and animals) [31]. Pesticides enter the atmosphere through volatilization [32], and go back to the surface through dry or wet deposition [33, 34]. From soil and plant surfaces, pesticides can also enter the water systems (rivers, lakes) via surface runoff [35], or get into groundwater through leaching processes [36]. In the pedosphere, pesticides can be sorbed onto soil particles. Finally, pesticides can enter and further bioaccumulate in the biosphere through uptake by plant roots [37], or ingestion by animals [38].

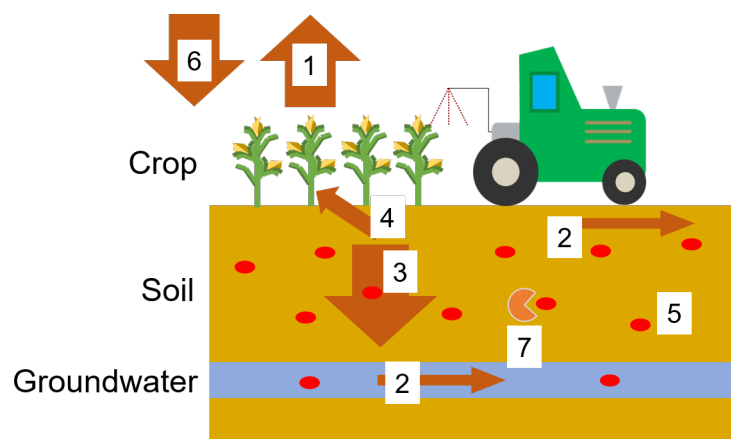


Figure 1: Transfer/transport fate of pesticide in the environment. 1: volatilization into the air, 2: Surface and subsurface runoff, 3: Leaching into groundwater, 4: plant uptake, 5: Sorption onto soil particles, 6: wet and dry deposition, 7: biological transformations

Transformation processes (abiotic and biotic) are responsible for the degradation of pesticides and the formation of intermediate metabolites. Abiotic pesticide transformation includes photolysis [39], hydrolysis [40], and redox reactions [41]. Biological transformations include the use of the pesticide as a carbon and/or nitrogen source for growth and maintenance [42–45], energy source [46, 47], and electron donor for redox reactions [48]. Many bacterial strains have evolved to degrade pesticides due to long-term exposure to such substances [49]. Cometabolic degradation [42, 50] that uses pesticide as “non-growth or fortuitous substrate” [51], as well as plant-mediated transformations [52–54] have also been observed.

Despite the existence of different degradation pathways and the observable biodegradability of pesticides, some pesticides can remain in soils in significant concentrations to affect human health and soil biota [55]. According to Silva et al. [55], around 80% of the studied topsoils across Europe contained pesticide residues and mixtures of them at very low concentrations. The herbicide atrazine is one particular example of long-term persistence in real systems. This herbicide can still be found in soils and groundwater even after 30 years of absence of application due to bans [56, 57]. Many physical and biological processes have been hypothesized to limit pesticide degradation in soils. Physical processes such as sorption onto soil particles [58], or onto

humic substances [59] retard degradation. Spatial heterogeneity (of pesticide [60] and degrader populations [61]) also reduces degradation rates leading to persistence. Biologically mediated processes including active [35] and passive [62, 63] transport across the cell membrane could represent the first step in a predominantly intracellular degradation of pesticide. Additionally, metabolic demands surpassing the catabolic energy obtained from pesticide degradation might stop degradation below certain threshold concentrations [64]. However, the unexpected findings of pesticide residues in soils demonstrate that our understanding of pesticide degradation in the environment is still incomplete.

### **3.3 How to assess pesticide fate in the environment**

Degradation experiments are designed to investigate degradation pathways, the formation of main metabolites, and the half-lives of pesticides [65]. They are generally performed with laboratory experiments (controlled conditions) and field studies (closer to natural conditions).

Laboratory experiments in soil microcosms or chemostat/retentostat reactors, provide a simplified representation of processes in nature, and conclusions from these experiments, therefore, cannot be simply extrapolated to real systems [66, 67]. One proof is the case of atrazine, which is readily degradable under controlled conditions but persists in soils and groundwater [56]. However, due to the simplicity of laboratory systems, specific processes can be studied without the interference of uncontrolled factors [68, 69]. Field studies are closer representations of real systems but are usually linked with high costs and the encounter of many uncontrolled factors that can affect the target measurements [70]. Because both methods have limitations, modeling tools can bridge the gap and connect them to real systems.

Kinetic models, representing biochemical reactions through equations [71] can be used to describe pesticide degradation in different real systems (soil, groundwater, water bodies) for prediction and for process understanding.

Prediction or risk assessment models for pesticide degradation are simple models generally varying from first-order-like models to lag-phase models. The forum for the coordination of

pesticide fate models and their use (FOCUS) [65] compiles different kinetic models that are used for the assessment of new pesticides. FOCUS models evaluate pesticide persistence using degradation endpoints (standard endpoints evaluated are  $DT_{50}$  and  $DT_{90}$ , which refers to the time that it takes to dissipate 50% to 90% of the pesticide) of the parental compound and metabolites. Some models additionally describe the influence of soil heterogeneity in pesticide degradation (Gustafson and Holden model) and the role of bacterial degraders (Logistic model) [65]. Despite their effectiveness, in some cases, endpoints are underestimated, and residual pesticides can still be found in soils and groundwater [55, 72]. This suggests that key drivers of pesticide degradation might be misrepresented in such models [73].

A deeper investigation of the mechanisms controlling pesticide degradation can be achieved with biogeochemical mechanistic modeling [74, 75]. This approach can not only improve our mechanistic understanding of reactive processes, but also produces benchmark models that can then be simplified to be applicable to large scales. Some examples of mechanistic approaches with the potential for modeling pesticides in soils are described as follows:

1. **Metabolic flux modeling** is an approach that uses genome sequences to derive the potential biogeochemical reactions related to them. It is mainly focused on single species but can be extended to multiple species assuming a supra-organism [76]. An example of this approach has been applied to atrazine in contaminated soils coupled with biostimulation strategies [77].
2. **Gene-centric modeling** is an approach that uses genetic information on specific functional groups to make quantitative predictions of genes and mRNA dynamics, and relating them to substrate dynamics [76]. One example of this approach for pesticide degradation is the PECCAD model [43, 78].
3. **Agent-based modeling** is an approach that describes individuals and their interactions with detailed process descriptions, but it can be complex and difficult to apply [76]. With pesticides, they have mainly been used for risk assessment [79].

### 3.4 Model data integration using omics and isotopic data

The lack of data to validate mechanistic models is currently the main limitation in the applicability of biogeochemical modeling [73, 75]. However, modern techniques in biology continuously extend data availability and exploit new sources of information that can be useful for modeling different compounds and systems.

New techniques such as omics technologies (i.e., genomics, transcriptomics, proteomics, metagenomics, metabolomics, etc.) [80] provide promising information to analyze biological systems to far greater detail [80, 81]. For example, bacterial degraders can be represented in models by their functional groups or genes [43, 82, 83], transcripts or mRNA [83], and enzymes [83–85]. The explicit integration of gene abundances into mechanistic models helps to account for ecological functions of the species that might be misrepresented due to the low number of cultivatable microorganisms [82]. The explicit integration of enzymatic regulations is also key to improve predictions of organic matter dynamics [85], and to describe complex microbial communities [84]. This approach describes the microbial community as a “collective assembly of metabolic capabilities” represented by functional enzymes that do not depend on particular bacterial guilds. Finally, large datasets from metabolomics or transcriptomics have been used to derive degradation rates [77] and microbiome responses to environmental perturbations [86] through a combination of biogeochemical models and machine learning tools.

Experiments, involving stable isotopes (defined as elements with the same properties but different atomic masses, attributed to differences in their amount of neutrons) [87], are used in modeling the fate of compounds in real systems and have the potential to reduce model uncertainty and equifinality of model parameters [88, 89]. The most commonly used stable isotopes are nitrogen, carbon, phosphorus [87], and oxygen [90]. For example, isotopes of  $^{13}\text{C}$  have been used for carbon turnover models in soils [43, 88],  $^2\text{H}$  and  $^{18}\text{O}$  for soil water flow models [91], and  $^{15}\text{N}$  for nitrogen cycle models [92]. Recently, “compound-specific isotope analysis (CSIA)” has been fundamental to determine the degradation pathway of atrazine

in engineered systems, and thus the rate-limiting step of the degradation of this compound, especially at low substrate concentrations [62, 63, 93].

### **3.5 Model uncertainty and sensitivity analysis**

Mathematical models and models in general are always simplified representations, limited by our understanding of how complex systems work [94, 95]. Unfortunately, the simplification [95] introduces an error into our model formulations, called structural uncertainty [96]. On top of that, lab and field measurements used to calibrate and validate our models usually carry noise as a product of human error or the inability to control sources of variations that influence the processes to be measured [96]. Finally, complex models for pesticide degradation are typically of the mechanistic type and include many model parameters whose values cannot be accurately identified with the available data (equifinality problem) [73], which is another source of uncertainty.

Model uncertainty impairs the reliability of model predictions/simulations and might distort the interpretation of model results. The quantification of model uncertainty is therefore an important step to assess model reliability, especially when policymakers shall use model predictions to establish management plans for pesticide use [97] or mitigation strategies against pollution [98]. Various methods of uncertainty quantification have been established in different research fields. Often, methods are based on Bayesian theory [99], such as the Bayesian multi-model ensemble analysis [95]. Frequently, mechanistic models for pesticide degradation are sloppy [73], meaning that parameters may not be identifiable. An identifiability analysis helps to determine those parameters and thus the processes that cannot be identified with the given data, leading to a further simplification of the model formulations [100, 101].

Sensitivity analyses provide the information on how the uncertainty of model inputs (model parameters) impacts the uncertainty of model outputs [102, 103]. Sensitivity analyses can be local, providing a limited picture of the parameters that the larger impact on model output [104]. Global sensitivity analysis, on the other hand, evaluates changes within the entire parameter

space, affected by all parameters at once [105]. Techniques for global sensitivity analyses are the Morris Method [103, 106, 107] and the variance-based “Sobol” method [103].

### 3.6 Model-based optimal design of experiments

In the intent to describe pesticide degradation, many distinct and competing model formulations have been developed to account for the complexity of pesticide degradation in the environment, our incomplete understanding of the processes that control degradation, and the lack of sufficient data to validate model assumptions. Finding the best representation of pesticide degradation is hence a non-trivial and challenging problem, especially if the models are later used for predictions [108]. Model selection techniques aim to guide this selection process by finding a trade-off between model complexity and goodness of fit against the available data [109]. If the data used for model selection is insufficient, the original experimental setup can be refined through a retrospective optimal design of experiments.

On the other hand, a prospective optimal design of experiments can allow for model selection prior to the execution of the experiment [110]. Its objective is to find a single design  $d_{opt}$  from a collection of designs  $D$  that maximizes the information gain towards a specific goal ( $\phi$ ) [111], which could be model discrimination. Additionally, the prospective method can be targeted to improve model calibration and reduce uncertainty of predictions.

When working with competing models, it is important to notice that including a completely inappropriate model can easily skew the results by indicating an easy discrimination of that model. Therefore, it is essential to carefully select for models with the best possible representation of the system to work with [112].

There are different methods for prospective optimal design that have been proposed in various fields. Leube et al. [113] introduced the preDIA method (posterior data impact assessor) that combines Monte Carlo simulations, Bayes’ theorem and Bayesian model averaging to “average the utility of designs over all possible measurements that a given sample can produce” [113]. This method was used on steady-state simulations to evaluate the optimal design that could

predict the long-term reduction of a pollutants' concentrations in groundwater [113], to evaluate the location of wells for a better determination of the groundwater divide location [111].

The main drawback of methods like preDIA is the computational effort to: i) compute the marginal utility of the designs, and ii) produce non-biased plausible model simulations that allow us to generalize the optimal design results. To solve the problem of computational effort, we can use other metrics to evaluate the benefit of the proposed designs. Because designs can be interpreted as vectors containing measurements [111], metrics utilizing the Euclidean distance can be adapted with relatively low computational effort, for example the concepts of energy distance [114],  $L_2$ -distance, and  $L_2$ -norm [115].

Producing plausible model outputs prior to the experiment, while also expending reasonable computational effort can be challenging, especially when the behavior of the model simulations cannot be constrained by data. A combination of rule-based and Markov chain Monte Carlo (MCMC) methods can provide a solution. Rules can be derived from expert knowledge of the system (i.e., half-life of pesticide) and be used to only keep relevant system behaviors [116]. Complementary, MCMC-based methods such as the DREAM algorithm [117], or the constraint-based search algorithm [118] are valuable to efficiently sample parameters leading to the desired behavior.



## 4 Research questions

Even though we know that bacteria and other soil organisms as well as abiotic processes have the potential to degrade pesticides, some persist in soils for long periods [55–57]. This shows that our understanding of the processes controlling pesticide degradation in soils is still incomplete. Moreover, potential microbial and biophysical limitations and environmental factors such as soil temperature and soil moisture are not fully considered into current pesticide degradation models, or their relevance for pesticide persistence has not been fully evaluated.

To improve our process understanding of pesticide fate in soils, I explored three potential mechanisms controlling pesticide degradation, microbial regulation, biophysical limitations, and the influence of environmental factors (temperature and soil moisture), to answer the following research questions.

**(R1):** What is the role of regulated gene expression as a microbial control of pesticide degradation in soils, and what is the benefit of explicitly including this process into biogeochemical models for process understanding and model predictions?

**(R2):** Do biophysical limitations (mass transfer across the cell membrane) and bioenergetic growth constraints control the degradation of pesticides in soil?

**(R3):** How do different soil temperature and soil moisture levels affect the overall pesticide fate in soils?

Pesticide degradation in soils can be carried out by different bacterial strains and is affected by physicochemical processes such as sorption and abiotic degradation. These processes require model formulations that differ in complexity and accuracy of the predictions. To identify the predominant pathway of pesticide degradation in soils, and to select the most suitable model among competing models, I investigated which data is needed to distinguish the competing degradation pathways. To this end, I aimed to answer the following additional research questions:

**(R4):** What type of measurements should be prioritized to distinguish competing degradation pathways of pesticides in soils?

**(R5):** What level of complexity is needed to represent pesticide degradation in soils?

I answered these questions with a data-model integration approach using mechanistic, process-based models. With this thesis, I expect to improve process understanding related to pesticide degradation in soils. Further, I provide equations/new approaches of how to model limitations of pesticide degradation that can easily be implemented into biogeochemical models.

## 5 Research design

### 5.1 Model assumptions

With my thesis, I aimed to explore the role of biokinetic processes that might explain pesticide persistence. I assumed a so-called “optimal pesticide degradation scenario” (Figure 2), focusing on microbial regulations to be the main drivers of pesticide degradation. Therefore, I excluded processes that may retard or even enhance pesticide degradation, such as i) spatial soil heterogeneity, ii) preferential water transport (advective–dispersive transport), iii) competing carbon sources (I assume pesticides as the sole carbon and energy source).

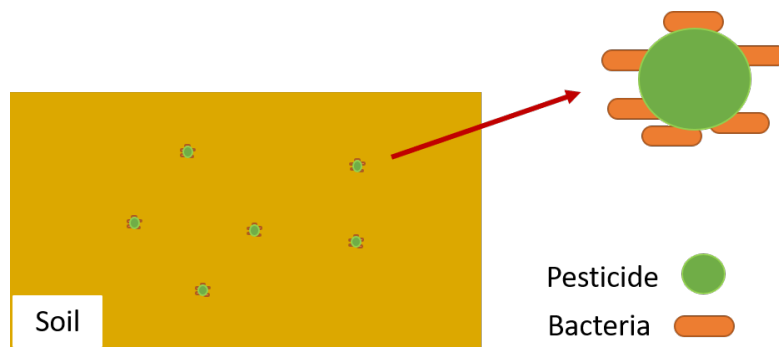


Figure 2: “Optimal pesticide degradation” scenario. In this scenario, bacteria and pesticide co-occur in the same spots, facilitating degradation. Water transport in soil, soil spatial heterogeneity and additional carbon sources are not included

### 5.2 Model pesticides

I used two model pesticides from two classes: i) Chlorophenoxy herbicides, represented by 2,4-D and MCPA, and ii) Triazines, represented by atrazine. Based on the characteristics of these two pesticide classes and their fate in the environment, I evaluated different hypotheses and degradation mechanisms specific for each pesticide.

**2,4-Dichlorophenoxyacetic acid (2,4-D) and 2-methyl-4-chlorophenoxyacetic acid (MCPA)** are auxin active molecules that disturb tissue growth of higher plants [119, 120], and which therefore are used as herbicides. 2,4-D and MCPA are highly soluble substances and

prone to leaching due to their low sorption to soil particles [121]. The complete degradation pathway of both components has been described in the literature [122–125], as well as the genes controlling degradation. The *tfdA* [126–128], *cadA* and *r/sdpA* genes [129, 130] are identified as the genes responsible for 2,4-D and MCPA degradation, with *tfdA* genes being the most abundant in soils [131]. The well-studied degradation pathway at the molecular level made these two herbicides suitable model pesticides to explore microbial regulations of pesticide degradation.

**Atrazine** herbicide that despite having been banned for over thirty years in the EU, it is still detected in relevant concentrations in soils and groundwater [56, 57]. Thus, atrazine is a very good model pesticide to explore pesticide persistence in soils. In the environment, atrazine undergoes different abiotic [132] and biotic [49] degradation pathways. Many bacterial strains are involved in atrazine degradation, using it either as a carbon source [133], nitrogen source [134], or as both carbon and nitrogen source [135]. Bacterial strains carrying the genes *atzABC*, *trzN–atzBC*, or *trzN–atzC* [49, 63, 136] can grow on the side chains of atrazine and degrade it to cyanuric acid. Strains with the genes *atzDEF* are able to further reduce the intermediate metabolite cyanuric acid to CO<sub>2</sub> under absence of alternative nitrogen sources [137–139]. The absence of *atzB* or *atzC* genes leads to the accumulation of hydroxyatrazine [140–142], the main metabolite of atrazine by dechlorination. Hydroxyatrazine is also persistent in real systems [143, 144]. Additionally, cometabolic degradation pathways of atrazine produce the metabolites deethylatrazine (DEA) and deisopropylatrazine (DIA), which also persist in soils [145, 146]. This diversity of degradation pathways makes atrazine an interesting model pesticide for the identification of degradation mechanisms.

### 5.3 Methods for model uncertainty quantification and sensitivity analysis

In this thesis, parameter uncertainty was determined using the DREAM toolbox in Matlab [117]. The DiffeREntial Evolution Adaptive Metropolis (DREAM) algorithm for model calibration calculates parameter uncertainty in the form of a posterior parameter distribution based

on a Bayesian approach (prior knowledge of the model parameters or uninformative prior distributions and likelihood of the data). Identifiability analysis [100, 101] was performed to find parsimonious model formulations. Additionally, I used both local and global sensitivity approaches. In local sensitivity analysis, I determined the local sensitivity score [147]. In global sensitivity analysis, I used the Morris method as inexpensive screening of important parameters in high-dimensional problems [103, 106, 107] and the variance-based Sobol method to quantify the contribution to the variance of model outputs coming from single parameters and the interaction of parameters [148].

#### 5.4 Thesis outline

My thesis investigates microbial regulations of pesticide turnover in soil and includes four sections corresponding to four scientific papers (sections 6 to 9). A small summary of the main findings of each paper and their connections to the research questions of my thesis are presented in this section:

**Paper 1 (section 6)** addresses research question R1 and describes a set of gene-centric models that explicitly incorporate microbial regulation (gene expression). I calibrated and validated these model variants with published data from two degradation experiments involving two model pesticides, 2,4-Dichlorophenoxyacetic acid (2,4-D) and 2-methyl-4-chlorophenoxyacetic acid (MCPA) [128]. I compared the performance of the model variants with each other and a standard Monod model to determine the role of gene expression in predicting pesticide degradation in soil and a potential parsimonious model.

**Paper 2 (section 7)** addresses research question R2. It extends and improves existing chemostat/retentostat models by including refined representations of mass-transfer processes across the cell membrane as well as energetic growth constraints through transition state theory. The target pesticide was atrazine. After calibrating the model, I used it to produce site-specific predictions for soils and compared them to field observations of residual atrazine concentrations from two arable topsoils in southern Germany.

**Paper 3 (section 8)** addresses research question R3 through a combined lab and modeling study of MCPA degradation under different soil temperature (10°C and 20°C) and soil moisture regimes (pF = 1.8 and 3.5), involving two substrate concentrations (1 and 20 mg kg<sup>-1</sup>). Genes and transcripts as well as MCPA mineralization and residual concentration in soils were measured in a microcosm experiment of thirty days. The modeling work was based on gene-centric approaches and included additional features hypothesized to have a substantial impact under low substrate concentrations, such as maintenance fluxes, production of non-extractable residues (NER), and a constitutive gene expression.

**Paper 4 (section 9)** addresses research questions R4 and R5. As model pesticide, I used atrazine, as it has a rather complex degradation pathway driven by different bacterial guilds, physicochemical and abiotic processes. Different degradation pathways of atrazine were represented by different competing models, and, in order to identify the predominant degradation pathway, I used prospective optimal design.

## **6 Gene-centric model approaches for accurate prediction of pesticide biodegradation in soils (Paper 1)**

This chapter includes the following publications:

1. Adapted with permission from Chavez Rodriguez, L., Ingalls, B., Schwarz, E., Streck, T., Uksa, M., Pagel, H. (2020). Gene-Centric Model Approaches for Accurate Prediction of Pesticide Biodegradation in Soils. *Environmental Science & Technology*, 54(21), 13638–13650. <https://doi.org/10.1021/acs.est.0c03315>. Copyright 2020 American Chemical Society.
2. Adapted with permission from Correction to the original Paper as published in Chavez Rodriguez, L., Ingalls, B., Schwarz, E., Streck, T., Uksa, M., & Pagel, H. (2021). Correction to “Gene-Centric Model Approaches for Accurate Prediction of Pesticide Biodegradation in Soils.” *Environmental Science & Technology*, 55(9), 6524. <https://doi.org/10.1021/acs.est.1c01972>. Copyright 2020 American Chemical Society.

with the following modifications:

1. Correction to the original Paper as published in Chavez Rodriguez, L., Ingalls, B., Schwarz, E., Streck, T., Uksa, M., & Pagel, H. (2021). Correction to “Gene-Centric Model Approaches for Accurate Prediction of Pesticide Biodegradation in Soils.” *Environmental Science & Technology*, 55(9), 6524. <https://doi.org/10.1021/acs.est.1c01972>
2. Numbers of figures, tables, and equations are relative to this thesis and not to the original publication.

## 6.1 Abstract

Pesticides are widely used in agriculture despite their negative impact on ecosystems and human health. Biogeochemical modeling facilitates the mechanistic understanding of microbial controls on pesticide turnover in soils. We propose to inform models of coupled microbial dynamics and pesticide turnover with measurements of the abundance and expression of functional genes. To assess the advantages of informing models with genetic data, we developed a novel “gene-centric” model and compared model variants of differing structural complexity against a standard biomass-based model. The models were calibrated and validated using data from two batch experiments in which the degradation of the pesticides dichlorophenoxyacetic acid (2,4-D) and 2-methyl-4-chlorophenoxyacetic acid (MCPA) were observed in soil. When calibrating against data on pesticide mineralization, the gene-centric and biomass-based models performed equally well. However, accounting for pesticide-triggered gene regulation allows improved performance in capturing microbial dynamics and in predicting pesticide mineralization. This novel modeling approach also reveals a hysteretic relationship between pesticide degradation rates and gene expression, implying that the biodegradation performance in soils cannot be directly assessed by measuring the expression of functional genes. Our gene-centric model provides an effective approach for exploiting molecular biology data to simulate pesticide degradation in soils.

## 6.2 Introduction

Pesticides are important agrochemicals used for plant protection and yield optimization [2–5]. Despite their intended beneficial use, many of the applied active components end up in soils, groundwater or surface water [149–151], where they are potentially harmful for living organisms and the environment [15, 152]. Soil microorganisms (fungi and bacteria) are known to be the main drivers of pesticide degradation in soils [43, 149]; they have evolved to use pesticides as both carbon (C) and energy sources [49]. In this context, the most important microbial detoxification process in soils is the enzyme-catalyzed biotic transformation of pesticides [153].



To predict the fate of pesticides in the environment, we need to improve our understanding of the microbial control of pesticide degradation, particularly at low concentrations [154]. Biogeochemical modeling is an established approach for testing our understanding of bioreactive processes, as well as for quantifying and predicting the biodegradation of pesticides in soils [74, 155]. Current biogeochemical models incorporate important rate limiting factors such as microbial dynamics (growth, metabolism, and physiology) as well as sorption-controlled substrate availability. Recent modeling approaches seek to improve the representation of microbial pesticide degradation by exploiting experimental assays of marker genes that encode enzymes that catalyze specific reactions [156]. Measurements of DNA and transcript abundance of functional genes facilitate an improved understanding of biochemical processes by providing a direct link between specific microorganisms and biochemical functions. This quantitative gene data should thus facilitate a more robust estimation of biokinetic parameters of biogeochemical models [75, 82] in comparison with more traditional approaches [73].

Some biogeochemical modeling approaches in marine and groundwater systems incorporate and simulate either functional genes and transcripts (gene-centric model [82]), or enzyme concentration and transcripts (cybernetic or enzyme-based approach [84, 157]). Their results highlight the potential of incorporating molecular data into modeling to improve process understanding and model predictions [82]. Existing gene-centric models of pesticide turnover in soil improved the representation of microbial dynamics in soil [43, 158], but misrepresent important limiting factors of pesticide degradation [73] such as pesticide-dependent gene expression.

In order to address this problem, we present a novel modeling approach that exploits data on the abundance and the expression of functional genes involved in pesticide degradation in soils. We expect that a complete description of transcription of specific genes and translation of targeted enzymes in our modeling approach will improve the representation of the controllers of pesticide degradation in soils. We used previously published data from laboratory experiments [128] to calibrate and validate a suite of model formulations. These experiments

involved observation of the degradation of the pesticides 2,4-dichlorophenoxyacetic acid (2,4-D) and 2-methyl-4-chlorophenoxyacetic acid (MCPA). These pesticides have a similar chemical structure and a simple degradation pathway mediated by the same functional genes, which have been extensively studied [122–125, 127, 130, 159–162]. Therefore, they provide a straightforward test case for our novel modeling approach.

We compare the model performance of a collection of gene-based models against a traditional biomass-based model to test whether our extended modeling approach provides better predictions. Finally, we used identifiability and uncertainty analysis to compare our gene-based model variants. We identified the model variant that is best supported by the available data, and that we can recommend as the most parsimonious tool to be used for description and prediction of this degradation process.

## 6.3 Theory

### 6.3.1 Model Structure

The model structure is shown in Figure 3. The processes are assumed to occur in a spatially homogeneous environment [128]. Pesticide is assumed to equilibrate rapidly between the sorbed and solution phase concentration. The model accounts for a single microbial population, which relies on the the pesticide as its sole carbon (C) and energy source. The microbial pool is partitioned into three subpopulations: active cells, inactive cells (dormant bacteria), and dead cells (relic cell pool), based on observation of typical bacterial states found in environmental systems [163–167]. Activation/inactivation is regulated by the pesticide concentration. Active cells respond to the presence of pesticide by expressing specific functional genes. We assume this occurs via upregulation of transcription (formation of messenger RNA (mRNA)). Pesticide uptake and metabolism are described as a single process, which leads to growth of active cells and mineralization of pesticide to CO<sub>2</sub>. Active and inactive cells die at constitutive rates. Finally, dead cells decay, releasing CO<sub>2</sub>. If not otherwise stated, all concentrations are given in mmol g<sup>-1</sup> or mmol cm<sup>-3</sup> and refer to C in mass of dry soil or in volume of soil solution, respectively.

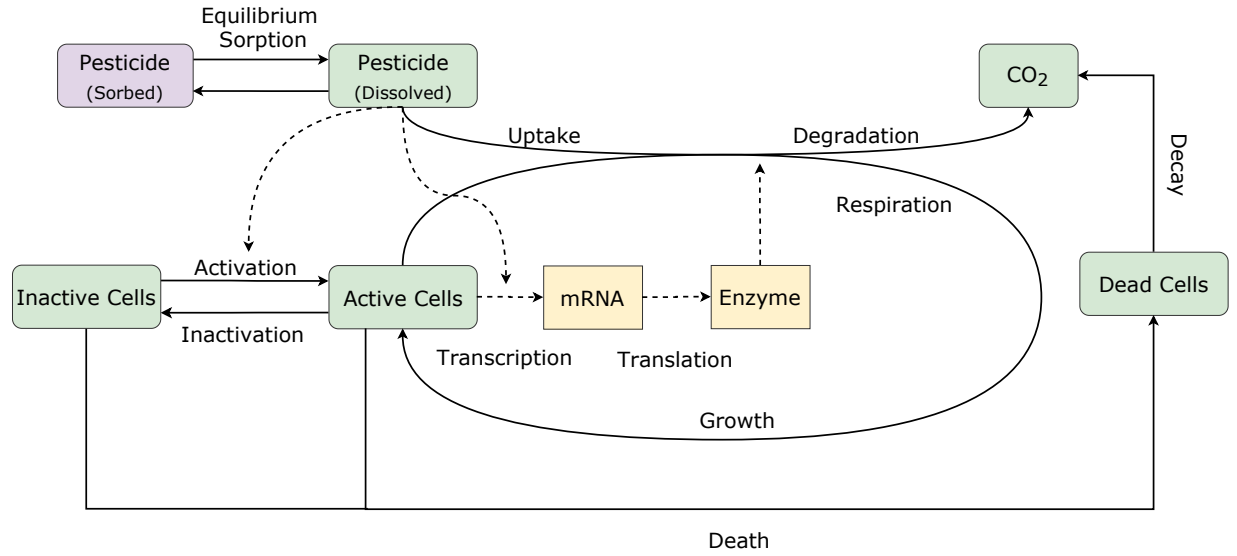


Figure 3: Model schematic. Green boxes: independent state variables calculated directly; orange boxes: variables related to processes in quasi-steady state; purple box: additional pool that is modeled indirectly. Solid arrows indicate mass transfer; dashed arrows indicate regulation.

### 6.3.2 Process formulations

#### Sorption

The total concentration of pesticide  $C_p^T$  [mmol cm<sup>-3</sup>] in soil is partitioned into two pesticide pools: solution phase concentration  $C_p^L$  [mmol cm<sup>-3</sup>] and sorbed phase concentration  $C_p^S$  [mmol g<sup>-1</sup>]:

$$C_p^T = \theta \cdot C_p^L + \rho \cdot C_p^S \quad (1)$$

where:  $\theta$  [1] is the water content in soil and  $\rho$  [g cm<sup>-3</sup>] is soil bulk density.

We assume that  $C_p^L$  and  $C_p^S$  are related by the Freundlich isotherm  $C_p^S = K_{FP} \cdot (C_p^L)^{n_{FP}}$  (bioavailability limitation of pesticide degradation). Freundlich sorption enters the model by the retardation factor [149, 168]:

$$RF := \frac{\frac{dC_P^T}{dt}}{\theta \cdot \frac{dC_P^L}{dt}} = 1 + \frac{\rho}{\theta} \cdot K_{FP} \cdot n_{FP} \cdot (C_P^L)^{(n_{FP}-1)} \quad (2)$$

where  $K_{FP}$  [ $\text{mmol}^{(1-n_{FP})} \text{g}^{-1} \text{cm}^{3n_{FP}}$ ] and  $n_{FP}$  [1] are the Freundlich coefficient and exponent respectively.

### Bacterial subpopulations

Our model incorporates three bacterial subpopulations: active degrader bacteria  $C_B^a$  [ $\text{mmol g}^{-1}$ ], inactive bacteria  $C_B^i$  [ $\text{mmol g}^{-1}$ ], and dead bacteria  $C_B^d$  [ $\text{mmol g}^{-1}$ ]. Inactivation (dormancy) is used by microbes as a bet-hedging strategy to cope with unfavorable conditions, including substrate limitation [163]. We included the dead bacteria (relic bacterial population) to avoid overestimation of active degraders [165–167] when comparing simulation results with experimental observations. These pools are depicted in Figure 3 as active, inactive and dead cells. The growth rate [ $\text{mmol g}^{-1} \text{d}^{-1}$ ] of active bacteria ( $r_{growth}$ ) is:

$$r_{growth} = \mu_P \cdot C_B^a \quad (3)$$

where  $\mu_P$  [ $\text{d}^{-1}$ ] is the specific growth rate coefficient (Eq. 18).

Activation and inactivation rates [169, 170] depend on the concentration of pesticide through a thresholding function. The corresponding rates [ $\text{mmol g}^{-1} \text{d}^{-1}$ ] are defined as follows:

$$r_{activation} = \tau \cdot k_r \cdot C_B^i \quad (4)$$

$$r_{inactivation} = (1 - \tau) \cdot k_d \cdot C_B^a \quad (5)$$

where  $k_r$  [ $\text{d}^{-1}$ ] and  $k_d$  [ $\text{d}^{-1}$ ] are the coefficients of activation/deactivation for inactive/active cells, and  $\tau$  [1] is a switch function: [169, 170]:

$$\tau = \left[ \exp \left( \frac{C_T - C_P^L}{n \cdot C_T} \right) + 1 \right]^{-1} \quad (6)$$

where  $C_T$  [mmol cm<sup>-3</sup>] is a pesticide concentration threshold, and  $n$  [1] modulates the steepness of the curve [169, 170].

The rate of decay of all bacterial pools (mmol g<sup>-1</sup> d<sup>-1</sup>) is described with a first-order function:

$$r_{decay}^j = C_B^j \cdot a_j \quad (7)$$

where  $j$  could denote active, inactive and dead bacteria, and  $a_j$  is the decay rate coefficient of the corresponding population.

### Gene expression

Active bacteria respond to the presence of pesticide by transcribing and translating specific functional genes for pesticide degradation. In the case of 2,4-D and MCPA, we described the expression of the functional gene *tfdA*, which encodes for an  $\alpha$ -ketoglutarate-dependent dioxygenase [128, 171, 172]. This enzyme catalyzes the cleavage of the ether bond between the phenol ring and the acetic acid side chain of 2,4-D and MCPA [130]. We assumed the first degradation step to be the rate determining step despite the fact that more genes *tfdABCDEF* are involved in the full degradation of both pesticides. This is a reasonable assumption given the general understanding of the *tfdABCDEF* mediated degradation pathway [171, 173]. The degradation of 2,4-D and MCPA involving a constitutive gene expression at low concentration followed by a pesticide-dependent gene expression was not included as an alternative microbial control mechanism, because this process has only been reported for one bacteria strain [126]. Alternative genes (i.e., *cadA* [174] and *RdpA*, *SdpA* [175]) capable of degrading 2,4-D and MCPA following the same degradation pathway as *tfdA* gene were not included because of their comparatively low abundance in soil samples [131, 173, 176].

The rate of transcription of specific genes is assumed to be pesticide-dependent, described by a Hill function [177]:

$$r_{transcription} = \beta_{max} \cdot \left( \frac{(C_P^L)^{n_H}}{K_G^{n_H} + (C_P^L)^{n_H}} \right) \quad (8)$$

where  $\beta_{max}$  [transcripts gene<sup>-1</sup> d<sup>-1</sup>] is the rate coefficient of transcription,  $K_G$  [mmol cm<sup>-3</sup>] is the half-maximal triggering concentration, and  $n_H$  [1] is the Hill coefficient.

Translation follows a first order function based on the concentration of transcripts [transcripts gene<sup>-1</sup>]:

$$r_{translation} = k_E \cdot mRNA \quad (9)$$

where  $k_E$  [mmol<sub>E</sub> transcripts<sup>-1</sup> d<sup>-1</sup>] is the rate coefficient of translation of transcripts (mRNA) into enzymes (E) [mmol<sub>E</sub> gene<sup>-1</sup>].

Decay of transcripts (mRNA) and enzymes (E) is assumed to be first order:

$$r_{mRNA-decay} = d_G \cdot mRNA \quad (10)$$

$$r_{enzyme-decay} = d_E \cdot E \quad (11)$$

where  $d_G$  [d<sup>-1</sup>] and  $d_E$  [d<sup>-1</sup>] are first order decay coefficients.

Experimental data on 2,4-D and MCPA degradation [128] have shown that the timescale of pesticide mineralization and bacteria growth is days, whereas transcription and translation are processes on the timescale of hours to seconds [178]. Therefore, we assumed quasi-steady-state (QSS) for gene expression [177]. The QSS formulation for transcripts reads:

$$\frac{dmRNA}{dt} = r_{transcription} - r_{mRNA-decay} = 0 \quad (12)$$

resulting in the following formulation [transcripts gene<sup>-1</sup>]:

$$\widehat{mRNA} = \frac{\beta_{max}}{d_G} \cdot \left( \frac{(C_P^L)^{n_H}}{K_G^{n_H} + (C_P^L)^{n_H}} \right) \quad (13)$$

We defined  $f_T = \frac{\beta_{max}}{d_G}$  as the number of transcripts per gene. The parameter  $f_T$  can take values higher than 1 to express at least one gene or transcript per cell, or lower than 1 to compensate for extraction bias of mRNA and DNA. The extraction bias was assumed constant for the soil

used in the experiment.

The QSS formulation for enzymes reads:

$$\frac{dE}{dt} = r_{translation} - r_{enzyme-decay} = 0 \quad (14)$$

resulting in the following formulation [ $\text{mmol}_E \text{ gene}^{-1}$ ]:

$$\hat{E} = \frac{k_E \cdot f_T \cdot \left( \frac{(C_P^L)^{n_H}}{K_G^{n_H} + (C_P^L)^{n_H}} \right)}{d_E} \quad (15)$$

### Pesticide uptake

We treat pesticide uptake and biotic degradation as a single process, with a degradation rate coefficient ( $\mu_P$  [ $\text{d}^{-1}$ ]) given by a Michaelis-Menten term:

$$\mu_P = \frac{k_{vmax} \cdot C_P^L \cdot \hat{E}}{(K_M + C_P^L) \cdot f_1} \quad (16)$$

where  $k_{vmax}$  [ $\text{mmol mmol}_E^{-1} \text{ d}^{-1}$ ] describes the rate of degradation of pesticide,  $K_M$  [ $\text{mmol cm}^{-3}$ ] is the half-maximal pesticide concentration, and  $f_1$  is a conversion factor from gene to C [ $\text{mmol gene}^{-1}$ ]. We do not consider chemical degradation. Substituting  $\hat{E}$  from eq. 15 gives:

$$\mu_P = \left( \frac{k_{vmax} \cdot k_E \cdot f_T}{d_E \cdot f_1} \right) \cdot \left( \frac{\left( \frac{(C_P^L)^{(n_H+1)}}{K_G^{n_H} + (C_P^L)^{n_H}} \right)}{K_M + C_P^L} \right) \quad (17)$$

Finally, defining  $\mu_{max} = \frac{k_{vmax} \cdot k_E \cdot f_T}{d_E \cdot f_1}$  (the maximum growth rate coefficient [ $\text{d}^{-1}$ ]), we have a concise description of growth rate as:

$$\mu_P = \mu_{max} \cdot \left( \frac{(C_P^L)^{(n_H+1)}}{K_G^{n_H} + (C_P^L)^{n_H}} \right) \cdot (K_M + C_P^L)^{-1} \quad (18)$$

The uptake rate is dependent only on the bioavailable pesticide (RF from eq. 2) and is scaled by the parameter  $Y_P$  (uptake efficiency specific for the degraders and the substrate) as follows:

$$r_{uptake} = \frac{-\mu_P \cdot C_B^a \cdot \left(\frac{1}{Y_P}\right) \cdot \left(\frac{\rho}{\theta}\right)}{RF} \quad (19)$$

The pesticide is used as both C and energy source. The growth yield,  $Y_P$ , determines the relative proportion of respiration and growth based on the total pesticide uptake [88].

### CO<sub>2</sub> accumulation

The final product of pesticide degradation is CO<sub>2</sub> (in mmol g<sup>-1</sup> soil) produced by bacterial respiration:

$$r_{respiration} = \mu_P \cdot C_B^a \cdot \left(\frac{1 - Y_P}{Y_P}\right) \quad (20)$$

A fraction of the decaying dead cells also contributes to the CO<sub>2</sub> pool through the parameter  $a_{CO_2}$  [1]:

$$r_{cell-decomposition} = r_{decay}^d \cdot a_{CO_2} \quad (21)$$

This flux consists of the carbon released from decaying dead cells that have incorporated the pesticide into their biomass. Autotrophic and heterotrophic fixation of mineralized CO<sub>2</sub> from 2,4-D/MCPA have been shown to play a minor role [179]. Therefore, we did not explicitly consider these processes in the model.



The full model is described by the following ordinary differential equation (ODE) system:

$$\frac{dC_B^a}{dt} = r_{\text{growth}} - r_{\text{decay}}^a + r_{\text{activation}} - r_{\text{deactivation}} \quad (22)$$

$$\frac{dC_B^i}{dt} = r_{\text{deactivation}} - r_{\text{activation}} - r_{\text{decay}}^i \quad (23)$$

$$\frac{dC_B^d}{dt} = r_{\text{decay}}^a + r_{\text{decay}}^i - r_{\text{decay}}^d \quad (24)$$

$$\frac{dC_P^L}{dt} = r_{\text{uptake}} \quad (25)$$

$$\frac{dCO_2}{dt} = r_{\text{respiration}} + r_{\text{cell-decomposition}} \quad (26)$$

## 6.4 Materials and Methods

### 6.4.1 Model Reduction

Starting out from the full model formulation (V0), we define three model variants considering pesticide-dependent gene expression: V1, V2 and V3, and also one variant considering unregulated gene expression V4 and a biomass-based model V4':

- V1: The inactive bacteria pool is set to zero ( $C_B^i = 0$ ).
- V2: The dead bacteria pool is set to zero ( $C_B^d = 0$ ). To account for cellular decay, a fraction of the decaying active and inactive bacteria is set to directly contribute to the  $CO_2$  pool as follows:

$$r_{\text{cell-decomposition}} = (C_B^a \cdot a_a + C_B^i \cdot a_i) \cdot a_{CO_2} \quad (27)$$

- V3: The inactive bacteria pool and the dead bacteria pool are set to zero ( $C_B^i = 0$  and  $C_B^d = 0$ ). The parameter  $K_M$  is also set to zero ( $K_M = 0$ ) to neglect pesticide-dependent growth and keep pesticide-dependent gene expression. The growth rate coefficient ( $\mu_P$ ) is calculated as follows:

$$\mu_P = \mu_{\text{max}} \cdot \left( \frac{(C_P^L)^{n_H}}{(K_G)^{n_H} + (C_P^L)^{n_H}} \right) \quad (28)$$

To account for cellular decay, a fraction of the decaying active cells is set to directly

contribute to the CO<sub>2</sub> pool as follows:

$$r_{cell-decomposition} = C_B^a \cdot a_a \cdot a_{CO_2} \quad (29)$$

- V4: The inactive bacteria and the dead bacteria are set to zero ( $C_B^i = 0$  and  $C_B^d = 0$ ). The parameter  $K_G$  is also set to zero ( $K_G = 0$ ) to account for pesticide-dependent growth kinetics. The growth rate coefficient ( $\mu_P$ ) is calculated as follows:

$$\mu_P = \mu_{max} \cdot \left( \frac{C_P^L}{K_M + C_P^L} \right) \quad (30)$$

This model variant considers an unregulated mRNA expression in QSS (eq. 13):

$$\widehat{mRNA} = f_T \cdot C_B^a \quad (31)$$

- V4': Same structure as V4 without unregulated mRNA gene expression. This model variant is close to the standard Monod-based model that takes gene abundances as a proxy of bacterial biomass.

For convenience, the full description of each model can be found in Supporting Information.

## 6.4.2 Model calibration

### Description of the experiment

We used published data from a batch degradation experiment in microcosms applying the <sup>14</sup>C-labelled 2,4-D and MCPA to a Typic Argiudoll with a pH of 7.2, 19% clay, 18% silt, 62% sand and 1.2% carbon [128]. The experiment consisted of one application of 20 mg kg<sup>-1</sup> (8.8 mg kg<sup>-1</sup> soil of 2,4-D, and 10.8 mg kg<sup>-1</sup> soil of MCPA; pesticide in C equivalent) of the corresponding pesticide at day 0, and a second application of 20 mg kg<sup>-1</sup> when mineralization of the first application stopped (at day 24 for 2,4-D and day 33 for MCPA.) The experiments ran until

day 34 for 2,4-D and day 67 for MCPA (the sampling protocol can be found in the Supporting Information Table S1) The dataset consisted of time series with three replicates of pesticide mineralization (%), abundance of *tfdA* genes and expressed *tfdA* genes in soil.

### Model calibration

We calibrated each model variant against the data from the 2,4-D experiment. We started calibration with a manual exploration of the parameter space within the ranges shown in Table 1 to achieve visually acceptable fits. Nominal parameter ranges were based on literature [158, 165, 180–190]. We extended the ranges of all parameters by four orders of magnitude, which allowed us to capture sufficient parameter variation.

Table 1: Initial parameter ranges for model calibration.

Parameters	Definition	Units	Minimum	Maximum
$f_T$	Number of transcripts per gene [187]	transcripts gene <sup>-1</sup>	10 <sup>-4</sup>	10 <sup>3</sup>
$n_H$	Hill coefficient [185]	1	1	10
$K_G$	Half-maximal triggering concentration [188]	mmol cm <sup>-3</sup>	10 <sup>-10</sup>	10 <sup>6</sup>
$\mu_{max}$	Maximum growth rate [158, 181]	d <sup>-1</sup>	10 <sup>-4</sup>	10 <sup>5</sup>
$f_i$	Conversion factor (gene to C) [183]	mmol gene <sup>-1</sup>	10 <sup>-14</sup>	10 <sup>-8</sup>
$K_M$	Half-maximal pesticide concentration [184]	mmol cm <sup>-3</sup>	10 <sup>-8</sup>	10 <sup>4</sup>
$C_T$	Pesticide concentration threshold [180, 186]	mmol cm <sup>-3</sup>	10 <sup>-10</sup>	10 <sup>-2</sup>
$a_a$	Decay rate coeff. for active bacteria [158, 181]	d <sup>-1</sup>	10 <sup>-5</sup>	10 <sup>2</sup>
$a_i$	Decay rate coeff. for inactive bacteria [158, 181]	d <sup>-1</sup>	10 <sup>-7</sup>	10 <sup>-2</sup>
$k_r$	Coefficient of activation [170]	d <sup>-1</sup>	10 <sup>-5</sup>	10 <sup>2</sup>
$k_d$	Coefficient of deactivation [170]	d <sup>-1</sup>	10 <sup>-5</sup>	10 <sup>2</sup>
$a_d$	Decay rate coeff. for dead bacteria [189]	d <sup>-1</sup>	10 <sup>-4</sup>	10 <sup>4</sup>
$Y_P$	Uptake efficiency [158]	1	0.1	0.9
$a_{CO_2}$	Fraction of bacteria contributing to CO <sub>2</sub>	1	0.1	0.9
$K_{FP}$	Freundlich coefficient [158, 182]	mmol <sup>(1-n<sub>FP</sub>)</sup> g <sup>-1</sup> cm <sup>3n<sub>FP</sub></sup>	10 <sup>-2</sup>	1
$n_{FP}$	Freundlich exponent [158, 182]	1	0.8	1
$a_r$	Initial fraction of dead bacteria [165]	1	0.1	0.9

Using the manually determined preliminary fit as the initial parametrization, we calibrated

each model variant by minimizing the weighted sum of squared errors (SSE) following a hybrid optimization method [191]. A hybrid method consists of an initial search using the global optimization algorithm `Simulated Annealing`, followed by the local optimization `fmincon` of Matlab. We defined the SSE as:

$$SSE = \sum_i \frac{(y_{obs}^i - y_{sim}^i)^2}{\sigma_i^2} \quad (32)$$

where  $y_{obs}^i$  are the observations,  $y_{sim}^i$  the corresponding model outputs, and  $\sigma_i$  is the standard deviation of the corresponding observations. Initial calibration attempts were unsuccessful due to the wide ranges of variation within replicates, especially for measured expressed *tfdA* genes in soil. Consequently, we set a minimum threshold of 15% as coefficient of variation for the replicates (without which the genes and transcripts observations would have had a negligible contribution to SSE).

The model outputs corresponding to the measured data were calculated from the state variables as follows:

$$\text{Mineralization [\%]} = \frac{\text{CO}_2 \cdot 100}{\text{Initial Pesticide Concentration}} \quad (33)$$

$$\text{Genes [copies g}^{-1}\text{]} = \frac{C_B^a + C_B^i + C_B^d}{f_1} \quad (34)$$

$$\text{Transcripts [copies g}^{-1}\text{]} = \frac{mRNA \cdot C_B^a}{f_1} \quad (35)$$

We applied a Markov Chain Monte Carlo simulation using the Bayesian DREAM<sub>(ZS)</sub> algorithm within the DREAM Matlab toolbox [117] to estimate parameter and simulation uncertainty. Uniform parameter distributions (see Table 1 for ranges) were chosen as flat/uninformative priors for Bayesian inference using MCMC sampling. The starting values of the MCMC chains were drawn from a multinormal distribution of the parameters in log-space with mean values equal to the best fit from the hybrid method, arbitrary variances of 2.5, and zero covariances.

The selected variance was set to capture sufficient variation of the parameters. Convergence of chains was assumed for a  $\hat{R}$ -diagnostic [192] lower than 1.2 [117]. Minimum and maximum parameter values were taken from Table 1, and we chose the option *Reflect* as a boundary handling method in DREAM<sub>(ZS)</sub>. We used a Gaussian likelihood considering heteroscedastic measurement errors as implemented in DREAM<sub>(ZS)</sub>:

$$L(x|\tilde{Y}) = -\frac{n}{2} \cdot \log(2\pi) - \sum_{i=1}^n \{\log(\sigma_i)\} - \frac{1}{2} \cdot \sum_{i=1}^n \left( \frac{y_{obs}^i - y_{sim}^i}{\sigma_i} \right)^2 \quad (36)$$

where  $\sigma_i$  are the standard deviations of the observations, and  $y_{obs}^i - y_{sim}^i$  are the residuals.

We ran DREAM<sub>(ZS)</sub> in parallel, using three Markov chain trajectories with 100,000 and 300,000 simulations per chain, achieving convergence for all models.

The same process was followed for all model versions with the exception of V4', for which we only used Pesticide Mineralization (%) and *tfdA* genes for calibration.

### 6.4.3 Model comparison

We chose the Akaike Information Criterion (AICc) [101, 193] and the Bayesian Information Criterion (BIC) for the numerical evaluation of the Bayesian Model Evidence [194, 195]. The AICc is a measure of the predictive capability of a model, and the BIC indicates the identifiability of the parameters of a model for the given data [109]. AICc and the BIC were calculated as:

$$\text{AICc} = 2 \cdot m + n \cdot \ln \left( \frac{\text{SSE}}{n} \right) + \frac{2 \cdot m \cdot (m + 1)}{n - m - 1} \quad (37)$$

$$\text{BIC} = n \cdot \ln \left( \frac{\text{SSE}}{n} \right) + m \cdot \ln(n) \quad (38)$$

where  $m$  is the number of free parameters (including the initial conditions used as free parameters for model calibration),  $n$  is the number of observations (this evaluation was done on 2,4-D data with 54 observation points), and SSE is the sum of squared errors previously defined (eq. 32).

#### 6.4.4 Model validation

We validated the selected models against the MCPA data. We visually evaluated how well the models captured the main trends of the measured data: mineralized MCPA (%), *tfdA* genes and transcripts in soils. We compared simulated residual concentrations of 2,4-D and MCPA (total concentration at the end of the experiment) against reported measurements in short-term degradation experiments using these pesticides [196–198] as an additional validation procedure.

#### 6.4.5 Local sensitivity analysis

We performed a local sensitivity analysis based on the best fit obtained for the models. The analysis included: local parametric sensitivity coefficient [199, 200], identifiability score [100, 101], percentage error of the estimation [101], and parameter correlation matrix [101] (see Supporting Information 11.1.3, Methods: Local and Global Sensitivity and uncertainty analysis). We did not include the initial conditions in our analysis [101].

#### 6.4.6 Global sensitivity analysis – Morris method

We performed a global sensitivity analysis using the Morris Method [103, 105–107, 201], implemented in the SAFE toolbox of Matlab [202] (Supporting Information 11.1.3). We calculated two sensitivity metrics [201]: the mean of the elementary effects ( $\mu^*$ ) and the standard deviation of the elementary effects ( $\sigma$ ) at 20,000 points in the parameter space, corresponding to 380,000 model runs.

We ran the sensitivity analysis up to the day 24 to cover the first pesticide application of 2,4-D. The output variables considered were: maximum pesticide mineralization, maximum abundance of genes and transcripts, maximum active genes, minimum pesticide in solution, time of inflection point of mineralization, time to achieve the maximum mineralization rate and maximum gene expression, and SSE (eq. 32).

Parameters for sensitivity analysis were sampled from a normal distribution with the mean taken from the best model fit on 2,4-D data. The standard deviation was approximated so that

the resulting normal distribution of each parameter fitted into the upper and lower boundaries previously set in Table 1. As sampling strategy, we used the Latin hypercube sampling with radial trajectory [73].

Because the Morris method only allows a ranked classification of the parameters according to the values of  $\mu^*$  and  $\sigma$  obtained per parameter per output [73, 202], we normalized by dividing by the maximum  $\mu^*$  and  $\sigma$  observed for each parameter. We used the normalized  $\mu^*$  and  $\sigma$  to calculate the  $l_2$  - norm ( $l_2 = \sqrt{\mu^{*2} + \sigma^2}$ ) of each parameter [73, 105, 107]. The high leverage parameters [73, 105] are those with an  $l_2$  - norm higher than 0.5.

## 6.5 Results and discussion

### 6.5.1 Calibration, parameterization and model dynamics of full model (V0)

#### Model Calibration

We were able to visually calibrate the full model variant (V0) with respect to 2,4-D mineralization over the entire experiment (Figure 4A, black curve) with reasonable uncertainty comparable to the standard deviation of the data (Figure 4B, black dots with black error bars). The model, however, failed to reproduce the peaks of gene expression (mRNA), especially after the second pesticide application (Figure 4C, black curve).

Simulated *tfdA* gene abundances matched well with the observed data during the first degradation phase (Figure 4E, black curve). However, at the end of the experiment (day 34), a clear decay of genes was shown but not captured by the simulation. In contrast, the simulated *tfdA* gene abundances stabilized, indicating underestimation of bacterial decay. The uncertainty of the simulations was low compared to the variability of the data (Figure 4F, black dots with black error bars). The simulated behavior of the *tfdA* gene abundances from Figure 4E (black curve) can be understood in terms of the dynamics of the active, inactive and dead bacterial pool (Figure 4, inset panel E). We emphasized that active bacteria did decay at the end of the experiment, but the total DNA pool remained constant because of the slow decay rate of inactive bacteria, which was the predominant physiological bacterial state at the end of the

simulation.



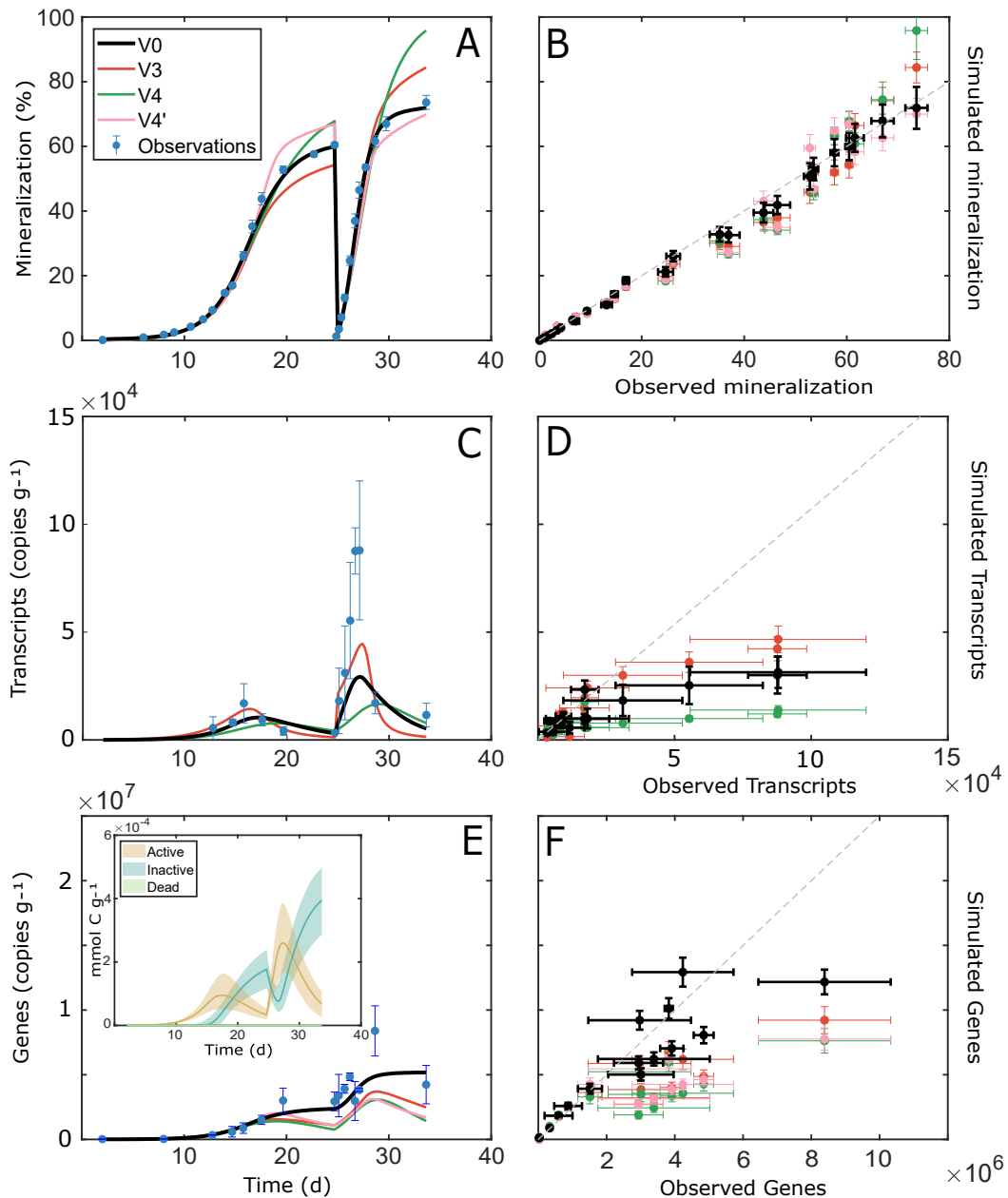


Figure 4: Model calibration on 2,4-D data. Time series of pesticide mineralization (A, B), *tfdA* mRNA copies (transcripts) per g of soil (C, D), *tfdA* gene copies per g of soil (E, F). Error bars show the standard deviation of the data and of the simulations (based on MCMC ensembles, see Materials and Methods 6.4.2). The inset in panel E shows calibrated model simulations of *tfdA* DNA, active, inactive and dead bacteria expressed in  $\text{mmol C g}^{-1}$  (based on 95 % confidence interval of MCMC ensembles, see Materials and Methods 6.4.2). Bold lines represent the mean value of each bacterial pool.

## Sensitivity analysis for model reduction

Table 2: Uncertainty analysis for the full model variant (V0) on 2,4-D data.

Parameter	Best Fit	SC	IS	PE	MV	SD
$\mu_{max}$	0.5	554.6	423.9	59.7	0.9	2.0
$a_r$	0.7	94.7	10.0	$1.5 \cdot 10^3$	0.9	0.04
$Y_P$	0.4	58.5	28.9	318.6	0.3	0.05
$n_{FP}$	0.9	55.6	46.8	$2.5 \cdot 10^3$	0.9	0.06
$f_1$	$8.2 \cdot 10^{-11}$	44.7	1.3	523.4	$8.9 \cdot 10^{-11}$	1.2
$a_a$	0.1	29.6	4.0	94.1	$5.0 \cdot 10^{-4}$	10.8
$K_G$	$7.0 \cdot 10^{-4}$	13.9	0.0	$1.1 \cdot 10^4$	$5.0 \cdot 10^{-7}$	234.4
$f_T$	0.02	13.0	7.0	96.4	0.01	2.1
$n_H$	5.6	6.6	0.5	258.4	4.7	2.9
$K_{FP}$	0.1	1.4	$2.8 \cdot 10^{-6}$	$5.3 \cdot 10^4$	0.1	4.2
$a_{CO_2}$	0.8	0.6	0.1	455.6	0.8	0.1
$k_r$	2.2	0.4	$1.1 \cdot 10^{-3}$	$4.4 \cdot 10^3$	0.5	1.5
$K_M$	$2.9 \cdot 10^{-5}$	0.2	$2.0 \cdot 10^{-4}$	$1.3 \cdot 10^4$	$2.5 \cdot 10^{-4}$	16.9
$a_s$	14.7	$7.4 \cdot 10^{-4}$	$6.7 \cdot 10^{-6}$	$1.3 \cdot 10^4$	89.3	17.3
$a_i$	$9.3 \cdot 10^{-5}$	$2.5 \cdot 10^{-8}$	$3.5 \cdot 10^{-13}$	$7.7 \cdot 10^7$	$2.4 \cdot 10^{-5}$	25.9
$C_T$	$1.0 \cdot 10^{-6}$	0.0	0.0	0.0	$6.0 \cdot 10^{-4}$	2.5
$k_d$	0.03	0.0	0.0	0.0	0.3	1.2

SC = Sensitivity coefficient, IS = Identifiability score, PE = Percentage Error, MV and SD = mean and standard deviation of the estimation from DREAM<sub>(ZS)</sub> (See Materials and Methods 6.4.2). Parameters in yellow were candidates to be reduced.

Local and global sensitivity analysis showed many parameters to be low-leverage and poorly identifiable, and the percentage errors of the parameter fits were overall high (Table 2 and Supporting Information, Figure S6). Parameters controlling the dynamics of the inactive and dead bacteria pools ( $k_r$ ,  $k_d$ ,  $a_i$ ,  $a_r$  and  $a_s$ ) were poorly identifiable with low impact (Table 2 highlighted in yellow), suggesting elimination of these parameters and simplification of the model structure. The calibrated values of  $C_T$  (the pesticide concentration threshold) ranged

from  $\mu\text{g L}^{-1}$  to  $\text{mg L}^{-1}$  (Supporting Information, Figure S8).  $C_T$  was found to impact the inactive bacteria pool at some point of the parameter space (Supporting Information, Figure S6), but no impact was observed when values are in the order of  $\mu\text{g L}^{-1}$  (Table 2). Threshold concentrations for activation of pesticide degrading bacteria have not yet been reported in the literature, but estimated threshold concentration values for activation of *Escherichia coli* [180, 186] are in the order of  $\mu\text{g L}^{-1}$ , similar to typical residual pesticide concentrations in soils [57, 203]. This suggests that reasonable values of  $C_T$  may be in the order of  $\mu\text{g L}^{-1}$ , and would thus have minimal impact on the model outputs (active, inactive and dead bacteria, pesticide concentration and  $\text{CO}_2$ ). The parameters  $K_M$  and  $K_G$  (which characterize the threshold for pesticide-dependent growth (eq. 16) and pesticide-dependent gene expression (eq. 8), respectively) had minimal impact on model dynamics (low leverage and low identifiability with a high percentage error). Despite being relatively low-leverage and low identifiable, sorption parameters ( $K_{FP}$  and  $n_{FP}$ ) were not considered for reduction due to the importance of sorption of pesticide in soil [204]. Moreover, these two parameters could be directly measured through sorption kinetic experiments [205–208] (not performed in the current work). Additionally, the analysis suggests that parameter  $a_{\text{CO}_2}$  can be eliminated due to its low impact on model outputs. However, we determined that this reduction results in overestimation of mineralization from dead bacteria (preliminary analysis; data not shown). Therefore, these three parameters were not included for model reduction.

Based on these sensitivity results, we select four reduced gene-centric model variants (V1, V2, V3 and V4) and a biomass-based model variant (V4'). In model variant V1, we eliminated dormancy by setting the inactive bacteria pool to zero. This reduction appears to contradict published reports that up to 60 % of bacteria are dormant in low nutrient systems [209]. This discrepancy could be resolved by further subdividing the active population, allowing for a 'potentially active' subpopulation [164], although we did not explore this option. In the second model variant, we removed the dead bacteria pool (relic bacteria) by setting it to zero. Neglecting relic bacteria in our model formulation could inflate the *tfdA* gene abundance simulated in soils [165, 166], and therefore, the pesticide degradation capacity of the soil (measured in terms

of mineralized pesticide). However, this effect was not observed in our calibration results (Figure 4A, C and E, red line for model variant V3 that exhibits the same features as model V2). In our initial development of the third model variant V3, we removed the inactive and dead bacteria. This, however, did not improve the sensitivity analysis results of the remaining parameters, and the parameters  $K_G$  and  $K_M$  were still low-leverage and poorly identifiable (data not shown). We chose to set only the parameter  $K_M$  to zero to keep a gene-centric model formulation. Because of this, this model variant has gene-expression as the only pesticide-dependent process.

In the two additional model variants (V4 and V4'), we removed inactive and dead bacteria. We also set the parameter  $K_G$  to zero and kept the parameter  $K_M$  to consider substrate-limited growth as the only pesticide-dependent process. Variant V4 describes a non-regulated (constitutive) gene expression. We developed this variant to address whether this approximation could also be a valid representation of the *tfdA* transcript dynamics. Variant V4' has the same model structure as V4, but without the non-regulated gene expression, keeping only the Monod-kinetics. Further sensitivity analysis of these two models (Supporting Information, Table S5 and S6), revealed a low-leverage Monod parameter  $K_M$ , suggesting a further reduction to a first-order-like model variant. This simple model, however, could not be successfully calibrated with the given data (not shown). Therefore, further model reduction steps were not considered.

### 6.5.2 Model comparison

We calibrated all model variants following the same methodology as with the full model with the exception of V4' that was calibrated using only pesticide mineralization and *tfdA* gene abundance. We evaluated the performance of the model variants based on visual fits, SSE (eq. 79) and the information criteria AICc (eq. 37) and BIC (eq. 38) (See Materials and Methods 6.4.2).

Table 3: Model comparison based on the 2,4-D data.

Model	FP	SSE	AICc	BIC
V0	18	192.9	124.3	140.6
V1	14	171.8	101.3	118.3
V2	16	189.8	114.6	131.7
V3	11	189.6	96.1	111.7
V4	10	229.5	103.3	118.0

FP = free parameters, AICc = corrected Akaike information criterion, and BIC = Bayesian information criterion.

Despite the expected high uncertainty of parameter estimates (see marginal posterior parameter distributions in Figures S8-S22 in the supplementary material), our predictions exhibited only moderate uncertainty (error bars in Figure 4B, Figure 4D and Figure 4F). This feature is typical of “sloppy” biogeochemical models [73], and in consequence allowed us to distinguish model performance of the model variants using visual fits, SSE (eq. 79), and the information criteria AICc (eq. 37) and BIC (eq. 38) (See Materials and Methods). By visual inspection (Figure 4; visual fits for model variants V1 and V2 not shown), as well as SSE and AICc (Figure 3), all model variants showed similar performance compared to the full model variant V0. A reasonable compromise among the three information criteria used (SSE, AIC, BIC) suggests that model variant V3 is the best model.

### 6.5.3 Model validation

We validated the models against the MCPA data, using the parameter values calibrated on 2,4-D (see Materials and Methods 6.4.2). We applied the validation procedure to the full model version V0, the parsimonious model version V3 and the two biomass-based models V4 and V4'. In addition, we compared simulated residual concentrations of pesticides in soils with all models against typical measured values.

## Validation on MCPA data

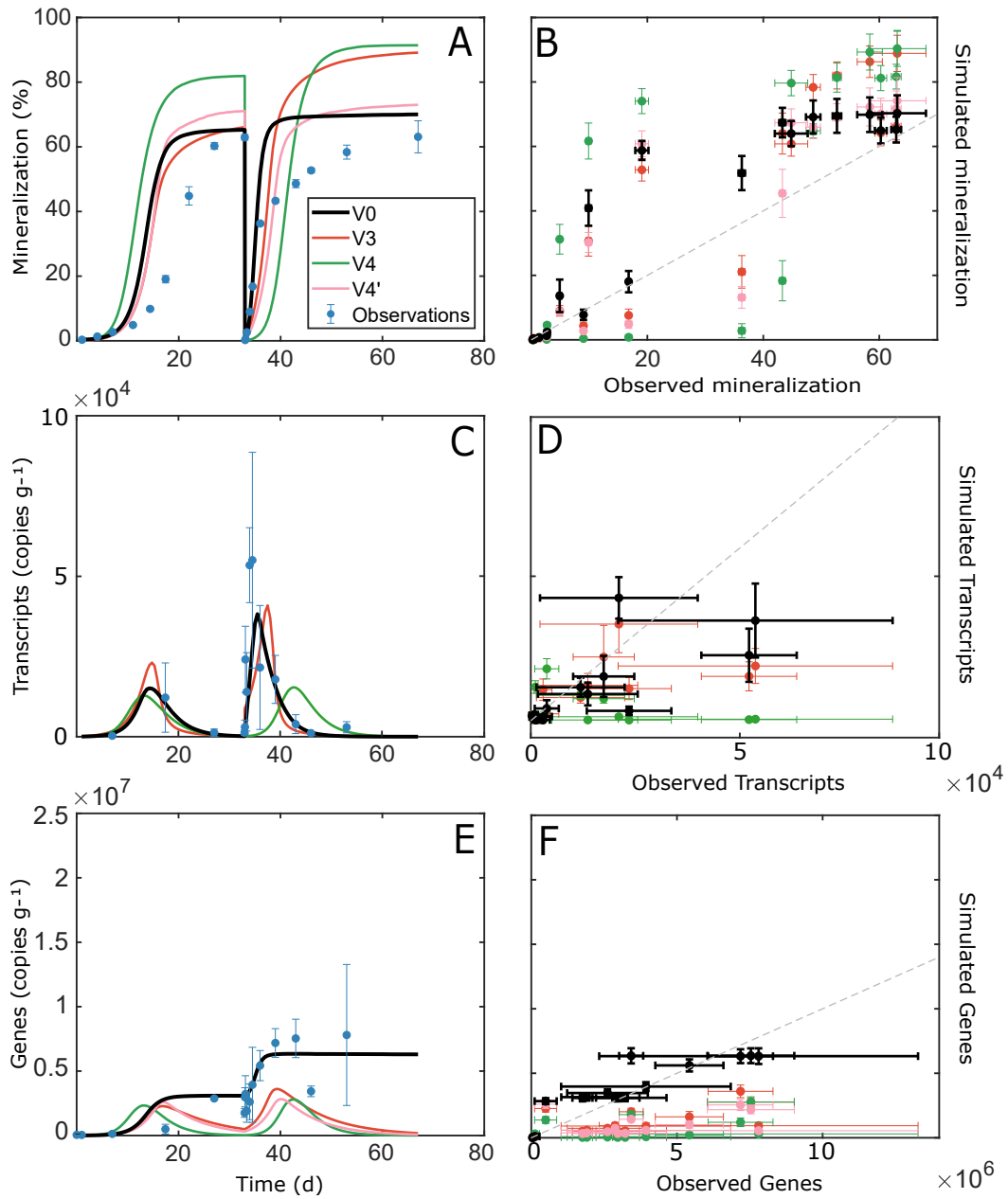


Figure 5: Model validation on MCPA data. Times series of pesticide mineralization (A, B), *tfdA* mRNA copies (transcripts) per g of soil (C, D), *tfdA* gene copies per g of soil (E, F). Error bars show the standard deviation of the data and of the simulations (based on MCMC ensembles, see Materials and Methods 6.4.2).

2,4-D and MCPA are chemically very similar with slight differences in substrate affinities [124, 210]. However, experimental measurements showed faster degradation dynamics of 2,4-D compared to MCPA [128]. Nevertheless, we expected that the models calibrated against the 2,4-D data could be usefully validated against the MCPA degradation data. As expected, all model variants predicted faster MCPA mineralization compared to the measurements (Figure 5A). (These time mismatches were not an artifact of the model formulation: an independent calibration against MCPA data was successful for all measured variants, see Supporting Information, Figure S1). Visually, the full model variant V0 and variant V4' were the best performing models regarding mineralization along the whole experiment and closely matched the data at the end of the first and second degradation phase (Figure 5A, black and pink line, respectively). Model variants V3 and V4 failed to predict MCPA mineralization, showing a strong overestimation of mineralization at the end of the experiment (day 67) (Figure 5A, red and green line, respectively), despite the success of variant V3 on predicting the first degradation phase. Our interpretation is that the overestimation of mineralization is linked to the underestimation of biomass growth (Figure 5E, red and green line, respectively). This underestimation leads to an increase of residual MCPA concentration in soil; therefore increasing total mineralization.

The model variants V0 and V3 captured the main trend of the expressed *tfdA* genes, including the high peaks of gene expression of both degradation phases (Figure 5C, black and red line respectively), with low uncertainty (Figure 5D). Model variant V4 failed in reproducing the expressed gene data (Figure 5C, green line). This confirms that the expression of *tfdA* genes is pesticide-regulated [78, 128], and that constitutive gene expression of *tfdA* is not a valid assumption for this process.

On the other hand, model variant V0 was the only model able to accurately reproduce data on *tfdA* genes (Figure 5E, black line) with low uncertainty of the simulations (Figure 5F). The fact that V0 was able to predict *tfdA* genes in soil can be explained by the dynamics of the active and inactive bacteria pools, similarly as with 2,4-D (Inset of Figure 4F). Model variants V3, V4 and V4' (all of which described a single bacterial subpopulation) predicted a pronounced

decay of bacteria at the end of the first phase. This explains these models' predictions of a slow bacterial recovery at the beginning of the second degradation phase, which made it impossible for their outputs to match the data (Supporting Information, Figure S3).

### Validation of simulated residual pesticide concentration

Residual pesticide in soil (total concentration at the end of the experiment) was not a measured variable in the experimental dataset we used [128]. Therefore, we used experimental data from previous studies to evaluate models' performance. Short-term experiments showed that both 2,4-D and MCPA are readily degradable compounds which do not persist in soil [196–198]. Some experiments have reported no 2,4-D and MCPA residues or residues below detection limit after 1 to 5 weeks [124, 211–214].

Table 4: Simulated 2,4-D/MCPA mean residual concentration in soil in  $\text{mgC kg}^{-1}$  (total concentration at the end of the experiment).

Model versions	Pesticides	
	2,4-D	MCPA
V0	0.5 – 2.3	$\leq 0.02$
V1	4.2 – 6.0	-
V2	0.1 – 1.4	-
V3	8.6 – 9.8	1.2 – 3.1
V4	3.0 – 4.8	$\leq 0.03$
V4'	$\leq 0.01$	$\leq 0.03$
Short-Term experiments [124, 211, 213, 214]	0.02 – 0.05	
Field studies [212]	0.01	

In comparison to literature values, all gene-centric model variants (V0, V1, V2 and V3), including biomass-based model variant V4 overestimated the residual pesticide concentration, especially for 2,4-D (Table 4). The models that include an inactive bacteria pool (V0 and V2) performed better, especially when simulating residual MCPA concentrations in soil. The average total pesticide dissipation simulated by these models was up to 98%, including both



pesticide mineralization and pesticide used for biomass formation and for production of non-extractable residues NER [179, 215, 216] (not explicitly accounted for in our model formulations). Model variants V1 and V3, which exclude inactive bacteria pool, overestimated residual pesticide concentrations by more of a hundred times (for both pesticides), resulting in predicted residual concentration that are even higher than residual concentrations reported for persistent pesticides [203]. These results suggest that dormancy is an important feature which should be accounted for.

The simple Monod-kinetics-based model variant V4' simulated negligible residual concentrations of both pesticides (Table 4), and thus outperformed the gene-centric model variants.

#### **6.5.4 Implications for biogeochemical modeling informed by genetic data**

Our gene-centric models can be used to explore the relationship between transcription of functional genes and process rates, offering an advantage over traditional models. We observed a non-linear hysteretic relationship between gene transcripts and mineralization rate (Figure 6). Although certain valid model parametrizations lead to a narrow hysteretic behavior close to a linear relationship (Figure 6), the non-linear hysteretic behavior in our findings challenges the common assumption of a simple linear relationship between functional gene transcripts and process rates [217–219], which could also be observed in the data (see Figure 6A). In addition to genetic data, proteomics data could be readily used. Thus, our approach provides a quantitative framework to couple gene and enzyme dynamics with pesticide dynamics, allowing for an estimation of reaction rates, which are difficult to measure directly.

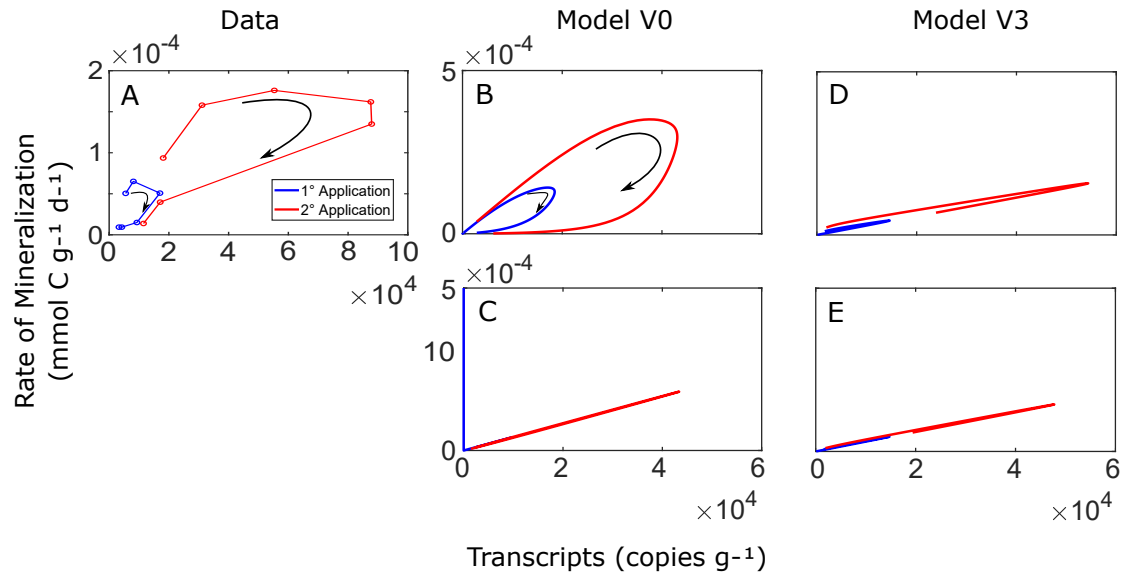


Figure 6: *tfdA* mRNA vs. rate of mineralization. Panel A shows reconstructed mineralization rates determined as the first derivatives of the “3/2-exp” and “3/2-lin” models in Bælum et al. 2008 [128] (with parameters as in Table 1 of that reference), and measured *tfdA* mRNA after 2,4-D application. Two arbitrarily selected relationships from the MCMC ensembles are plotted for full model version V1 (B, C) and parsimonious model version V3 (D, E), showing hysteresis and linear relationships between these two variables. Blue lines indicate the first application of 2,4-D, and the red lines the second application. Black arrows in panels A and B show the direction of the hysteric curve.

Gene-centered models thus provide mechanistic insights, despite being more complex than traditional approaches. This complexity poses challenges for inverse parameter identification from experimental data [73, 220, 221]. Achieving a robust model parametrization is hampered by the significant uncertainty in current measurements of functional genes and transcripts. We expect that better estimates of biokinetic parameters will be achieved with highly resolved time-series of genetic data and further advancements of molecular methods. Gene-centric biogeochemical modeling then provides a promising toolset to improve mechanistic simulations of biodegradation processes in soils, especially when coupled with reactive transport models in soil, and used for scenario simulations with other competing carbon sources and cometabolic pathways. Moreover, our approach can be transferred to other pesticides if degradation pathways and involved functional genes are known.

## **7 Modeling bioavailability limitations of atrazine degradation in soils (Paper 2)**

This chapter includes the following publication:

Luciana Chavez Rodriguez, Brian Paul Ingalls, Jana Meierdierks, Kankana Kundu, Thilo Streck and Holger Pagel (2021). Modeling bioavailability limitations of atrazine degradation in soils. *Front. Environ. Sci. - Biogeochemical Dynamics*.

with the following modification:

1. Numbers of figures, tables, and equations are relative to this thesis and not to the original publication.

## 7.1 Abstract

Pesticide persistence in soils is a widespread environmental concern in agro-ecosystems. One particularly persistent pesticide is atrazine, which continues to be found in soils and groundwater in the EU despite having been banned since 2004. A range of physical and biological barriers, such as sorption and mass-transfer into bacterial cells, might limit atrazine degradation in soils. These effects have been observed in experiments and models working with simplified systems. We build on that work by developing a biogeochemical model of the degradation process. We extended existing engineered system models by including refined representations of mass-transfer processes across the cell membrane as well as thermodynamic growth constraints. We estimated model parameters by calibration with data on atrazine degradation, metabolite (hydroxyatrazine) formation, biomass, and isotope fractionation from a set of controlled retentostat/chemostat experiments. We then produced site-specific model predictions for arable topsoil and compared them with field observations of residual atrazine concentrations. We found that the model overestimated long-term atrazine biodegradation in soils, indicating that this process is likely not limited by bioavailability or energetic constraints of microbial growth. However, sorption-limited bioavailability could explain the long-term fate and persistence of the main degradation metabolite hydroxyatrazine. Future studies should seek alternative controls that drive the observed atrazine persistence in soil. This work helps to bridge the gap between engineered and real systems, allowing us to use laboratory setups to gain insight into real environmental systems.

## 7.2 Introduction

The worldwide intensification of agriculture is closely linked to increased use of pesticides [222]. Persistent pesticides are defined as those that remain in soils “in significant concentrations until the next growing season” [223]. Field monitoring campaigns have demonstrated the presence of residual pesticides across Europe [55].

Atrazine (AT) is a herbicide in common use worldwide. AT was banned in Germany in 1991

and in the EU in 2004 [25, 224] because of its potential toxic effects on non-target organisms [13, 14, 16, 225, 226], and on human health [21, 25]. Despite the ban, AT persists in soils and groundwater [25, 57]: AT and its degradation metabolites (hydroxyatrazine, deethylatrazine, deisopropylatrazine) are still found in Europe at low concentrations (about 1-40  $\mu\text{g kg}^{-1}$ ) in soils [203], and ( $< 0.1 \mu\text{g L}^{-1}$ ) in groundwater [57]). These concentrations might still be relevant for human and ecosystem health [25, 57, 227]. This persistence is surprising, given that studies have confirmed (i) the frequent presence of bacterial strains able to completely degrade atrazine (under controlled conditions) [49, 63, 135, 228]; and (ii) alternative photolytic degradation of atrazine in soil [132, 229, 230]. The persistence of atrazine and other pesticides in the environment demands a better understanding of degradation processes to improve long-term monitoring and pollution mitigation strategies [98].

Pesticide degradation in the environment may be impeded by a range of physical and biological constraints. For instance, sorption of pesticides onto soil particles limits microbial access to pesticides, retarding degradation [58, 231, 232]. Moreover, spatial heterogeneity and separation of microorganisms and pesticides in soil reduces biodegradation rates [233, 234]. Diffusion-limited transport across the cell membrane has been identified as a potential limiting step of pesticide degradation under low concentrations, based on observations made in engineered (chemostat and retentostat) systems [63, 228, 235–238]. Likewise, under specific conditions, the energy produced from catabolism of some pesticides may be insufficient to support cellular energy needs, leading to pesticide persistence despite microbial accessibility [64]. To date, the effect of these barriers has only been explored in the lab under controlled conditions [62, 63, 93, 228] or in simulation studies based on simplified systems [232, 238–240].

In this work, we apply biogeochemical modeling to investigate potential factors of long-term pesticide persistence in soils. We extended existing chemostat/retentostat models [239] by the (i) introduction of thermodynamic growth constraints [241, 242] (the alternative model formulation uses a simple Monod kinetics growth [239]), (ii) a refined formulation of mass-transfer processes across cell membranes, and (iii) calibration against isotope fractionation

data. We then extended the model by including equilibrium sorption and leaching in soils, and ran site-specific predictions of pesticide degradation in soil over 30 years. We compare our model predictions with residual atrazine concentration of topsoils at two study sites (arable soil) in Germany at which no atrazine has been applied for over 30 years. Albeit the long-term predictions show considerable discrepancies with the field data, our analysis provides insight into the relative contributions of model features toward long-term atrazine persistence in soils.

### 7.3 Material and Methods

#### 7.3.1 Model Description

Our model (Figure 7) describes a single bacterial population ( $C_B$ ) that uses atrazine ( $AT$ ) as its sole carbon ( $C$ ) and energy source. The core model (green background), describes behavior in engineered systems (chemostat/retentostat); it incorporates intracellular and extracellular compartments, each of which contain concentrations of both  $AT$  and hydroxyatrazine ( $HY$ ). (Hydroxyatrazine is produced by dechlorination of the side chain of  $AT$ . This is the first metabolic step of  $AT$  degradation.) We extended the model to soil (blue background) by incorporating equilibrium sorption and leaching for each component in the extracellular compartment.

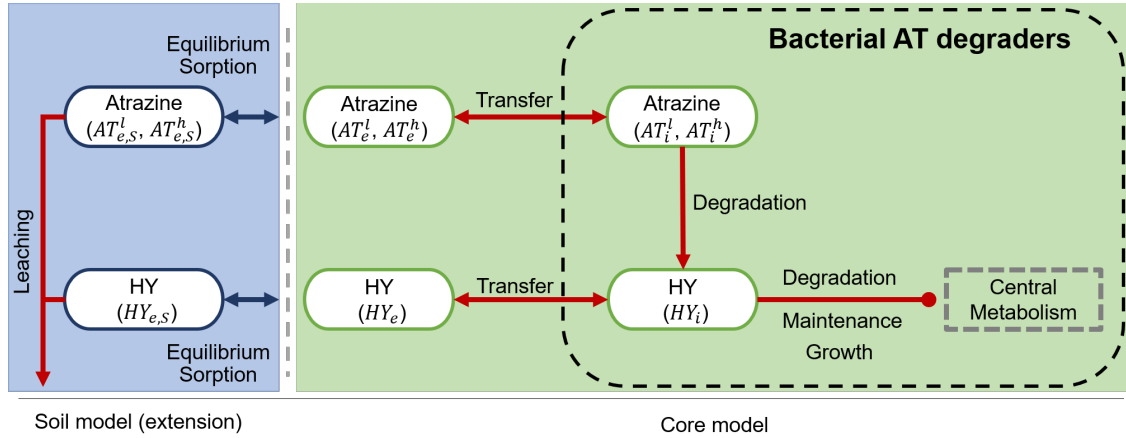


Figure 7: Model structure for engineered (chemostat/retentostat) systems (green) and extension for soil (blue). The model explicitly accounts for light “l” and heavy “h” isotopologues ( $^{12}\text{C}/^{13}\text{C}$ ) of AT due to enzymatic transformation in the intracellular “i” and extracellular “e” compartments, as well as in equilibrium sorption in soil “e, S”

### 7.3.2 Process Formulations

#### Atrazine and hydroxyatrazine degradation

The model describes pools of atrazine (AT) [ $\mu\text{g L}^{-1}$ ] and hydroxyatrazine (HY) [ $\mu\text{g L}^{-1}$ ] in the intracellular and extracellular compartments:  $AT_i/HY_i$  and  $AT_e/HY_e$ , respectively. To take advantage of available data on isotope fractionation of AT, we split the AT pools into light ( $AT^l$ ) and heavy ( $AT^h$ ) isotopologues ( $^{12}\text{C}/^{13}\text{C}$ ) in each compartment.

We modelled degradation of both isotopologues of AT with Michaelis-Menten kinetics, allowing for competition for binding sites. For the light isotopologue:

$$r_{degradation}^{AT^l} = \frac{k_{AT} \cdot AT_i^l}{K_M^{AT} + AT_i^l + AT_i^h} \quad (39)$$

where  $k_{AT}$  [ $\text{d}^{-1}$ ] is the maximum degradation rate of AT and  $K_M^{AT}$  [ $\mu\text{g L}^{-1}$ ] is the half-saturation concentration.

The slightly slower degradation of the heavy isotopologue is captured by scaling the maximal

degradation rate by  $\beta \lesssim 1$  as follows:

$$r_{degradation}^{AT^h} = \frac{\beta \cdot k_{AT} \cdot AT_i^h}{K_M^{AT} + AT_i^l + AT_i^h} \quad (40)$$

We considered two separate formulations of HY degradation. Model variant M employs standard Monod kinetics:

$$\text{Variant M : } r_{degradation}^{HY} = \frac{k_{HY} \cdot HY_i}{K_M^{HY} + HY_i} \quad (41)$$

where  $k_{HY}$  [ $d^{-1}$ ] is the maximum degradation rate and  $K_M^{HY}$  [ $\mu g L^{-1}$ ] is the half-saturation concentration for HY.

Because metabolism of pesticide at low concentrations might not be energetically favorable for bacterial growth [64], we considered a second model variant in which degradation of HY is described by transition state theory [241, 242], using HY as the carbon and energy source.

$$\text{Variant T : } r_{degradation}^{HY} = k_{HY} \cdot e^{\left(-\frac{K_M^{HY}}{HY_i}\right)} \quad (42)$$

where again  $k_{HY}$  [ $d^{-1}$ ] is the maximum degradation rate, but now  $K_M^{HY}$  [ $\mu g L^{-1}$ ] is a reference concentration for growth.

These two variants (Monod (M), Thermodynamic (T)) show similar behaviour at high HY concentrations (such as in chemostat/retentostat systems), but differ considerably at low HY concentrations (such as in soil).

### Mass-transfer

We account for diffusive transport of AT and HY across the cell membrane [62, 239] by writing:

$$r_{mass-transfer}^{AT^l} = r_e \cdot (AT_e^l - AT_i^l) \quad (43)$$



$$r_{mass-transfer}^{AT^h} = r_e \cdot (AT_e^h - AT_i^h) \quad (44)$$

$$r_{mass-transfer}^{HY} = r_e \cdot (HY_e - HY_i) \quad (45)$$

where  $l$  indicates the light isotopologue, and  $h$  the heavy isotopologue, and  $r_e$  [ $L d^{-1} \mu g^{-1}$ ] is the mass-transfer rate coefficient assumed to be the same for both compounds.

### Maintenance

We incorporate metabolic maintenance requirements following the Pirt model [239, 243]:

$$r_{maintenance} = m \cdot Y \cdot C_B \quad (46)$$

where  $m$  [ $d^{-1}$ ] is the maintenance coefficient.

### Input and washout of AT, HY, biomass

For engineered systems (chemostat/rententostat), we include a constant input of AT as:

$$r_{input}^{AT^l} = r_D \cdot AT_I^l \quad (47)$$

$$r_{input}^{AT^h} = r_D \cdot AT_I^h \quad (48)$$

where  $r_D$  [ $d^{-1}$ ] is the dilution rate. Additionally, we define washout terms for biomass, and AT and HY:

$$r_{cell-washout} = r_D \cdot \alpha \cdot C_B \quad (49)$$

$$r_{washout}^{AT_e^l} = r_D \cdot AT_e^l \quad (50)$$

$$r_{washout}^{AT_e^h} = r_D \cdot AT_e^h \quad (51)$$

$$r_{washout}^{HY_e} = r_D \cdot HY_e \quad (52)$$

where  $\alpha$  [-] is 1 for a chemostat (from which biomass is washed out) and 0 for a retentostat system (where biomass is retained).

The core model is described by the following system of ordinary differential equations (ODE):

$$\frac{dC_B}{dt} = r_{degradation}^{HY} \cdot Y - r_{maintenance} - r_{cell-washout} \quad (53)$$

$$\frac{dAT_i^l}{dt} = r_{mass-transfer}^{AT^l} \cdot \frac{f_{cell}}{V_u} - r_{degradation}^{AT^l} \cdot \frac{f_{cell}}{V_u} - \frac{AT_i^l}{C_B} \cdot \frac{dC_B}{dt} \quad (54)$$

$$\frac{dAT_i^h}{dt} = r_{mass-transfer}^{AT^h} \cdot \frac{f_{cell}}{V_u} - r_{degradation}^{AT^h} \cdot \frac{f_{cell}}{V_u} - \frac{AT_i^h}{C_B} \cdot \frac{dC_B}{dt} \quad (55)$$

$$\frac{dAT_e^l}{dt} = r_{input}^{AT^l} - r_{washout}^{AT_e^l} - r_{mass-transfer}^{AT^l} \cdot C_B \quad (56)$$

$$\frac{dAT_e^h}{dt} = r_{input}^{AT^h} - r_{washout}^{AT_e^h} - r_{mass-transfer}^{AT^h} \cdot C_B \quad (57)$$

$$\frac{dHY_i}{dt} = (r_{mass-transfer}^{HY} + r_{degradation}^{AT^l} + r_{degradation}^{AT^h} - r_{degradation}^{HY}) \cdot \frac{f_{cell}}{V_u} - \frac{HY_i}{C_B} \cdot \frac{dC_B}{dt} \quad (58)$$

$$\frac{dHY_e}{dt} = -r_{mass-transfer}^{HY} \cdot C_B - r_{washout}^{HY_e} \quad (59)$$

where  $f_{cell}$  [ $\mu\text{g cell}^{-1}$ ] is a conversion factor from cells to carbon, and  $V_u$  [L] is the volume of a single bacterium, set to  $1 \cdot 10^{-15}$  [63] (full details in the Supplementary Section 11.2.2). The last terms in eqs. 54, 55 and 58 account for changes in inner cell concentrations as the total bacterial volume changes due to growth and decay.

### Extension for soil

As shown in Figure 7, we extend the core model by including equilibrium sorption and transport.

We partition the extracellular concentrations of both AT isotopologues, as well as HY, into

solution phase and sorbed phase concentrations:

$$C^T = \theta \cdot C^L + \rho \cdot C^S \quad (60)$$

where  $C^T$  [ $\mu\text{g L}^{-1}$ ] is the total extracellular concentration (AT and HY),  $C^L$  [ $\mu\text{g L}^{-1}$ ] is the solution phase concentration ( $AT_e^l$ ,  $AT_e^h$ ,  $HY_e$ ),  $C^S$  [ $\mu\text{g kg}^{-1}$ ] is the sorbed phase concentration ( $AT_{e,S}^l$ ,  $AT_{e,S}^h$ ,  $HY_{e,S}$ ),  $\theta$  [-] is the water content in soils, and  $\rho$  [ $\text{kg L}^{-1}$ ] is the soil bulk density.

We relate  $C^L$  and  $C^S$  by the Freundlich isotherm:

$$C^S = K_F \cdot (C^L)^{n_F}, \quad (61)$$

implemented in the model via the retardation factor:

$$RF := 1 + \frac{\rho}{\theta} \cdot \frac{dC^S}{dC^L} = 1 + \frac{\rho}{\theta} \cdot K_F \cdot n_F \cdot (C^L)^{(n_F-1)} \quad (62)$$

where  $K_F$  ( $K_{AT}$  and  $K_{HY}$  for AT and HY respectively) [ $\mu\text{g}^{(1-n_F)}\text{kg}^{-1}\text{L}^{n_F}$ ] is the Freundlich coefficient and  $n_F$  ( $n_{AT}$  and  $n_{HY}$  for AT and HY respectively) [-] is the Freundlich exponent.

Additionally, *Arthrobacter aurescens* TC1 and other AT degraders utilize other organic substances as C and energy source. We, therefore, assume that a minimum AT degrader biomass is maintained in soil [244]):

$$r'_{\text{maintenance}} = m \cdot Y \cdot (C_B - M) \quad (63)$$

where  $M$  [ $\mu\text{g L}^{-1}$ ] is the minimum bacterial biomass in soil.

Transport is restricted to convective flow:

$$r_{\text{leaching}}^{AT_e^l} = \frac{v_v}{\theta} \cdot AT_e^l \quad (64)$$

$$r_{leaching}^{AT_i^h} = \frac{v_v}{\theta} \cdot AT_e^h \quad (65)$$

$$r_{leaching}^{HY_i} = \frac{v_v}{\theta} \cdot HY_e \quad (66)$$

where  $v_v$  [ $d^{-1}$ ] is the water flow per soil volume in the plough layer.

We did not include abiotic degradation of AT [132, 229, 230], which has been observed to have a relatively small contribution compared to biotic degradation [237].

The full model for soil is described by the following system of ODEs

$$\frac{dC_B}{dt} = r_{degradation}^{HY_i} \cdot Y - r'_{maintenance} \quad (67)$$

$$\frac{dAT_i^l}{dt} = r_{mass-transfer}^{AT^l} \cdot \frac{f_{cell}}{V_u} - r_{degradation}^{AT_i^l} \cdot \frac{f_{cell}}{V_u} - \frac{AT_i^l}{C_B} \cdot \frac{dC_B}{dt} \quad (68)$$

$$\frac{dAT_i^h}{dt} = r_{mass-transfer}^{AT^h} \cdot \frac{f_{cell}}{V_u} - r_{degradation}^{AT_i^h} \cdot \frac{f_{cell}}{V_u} - \frac{AT_i^h}{C_B} \cdot \frac{dC_B}{dt} \quad (69)$$

$$\frac{dAT_e^l}{dt} = - \frac{r_{mass-transfer}^{AT^l} \cdot C_B + r_{leaching}^{AT_e^l}}{RF} \quad (70)$$

$$\frac{dAT_e^h}{dt} = - \frac{r_{mass-transfer}^{AT^h} \cdot C_B + r_{leaching}^{AT_e^h}}{RF} \quad (71)$$

$$\frac{dHY_i}{dt} = (r_{mass-transfer}^{HY} + r_{degradation}^{AT_i^l} + r_{degradation}^{AT_i^h} - r_{degradation}^{HY_i}) \cdot \frac{f_{cell}}{V_u} - \frac{HY_i}{C_B} \cdot \frac{dC_B}{dt} \quad (72)$$

$$\frac{dHY_e}{dt} = - \frac{r_{mass-transfer}^{HY} \cdot C_B + r_{leaching}^{HY_e}}{RF} \quad (73)$$

### 7.3.3 Model calibration

#### Engineered systems: experimental details

We calibrated two model variants (M: employing Monod-kinetics for HY degradation; T: employing thermodynamic HY biodegradation constraints) against published data from chemostat and retentostat experiments (with two replicates per experiment). Atrazine was provided as the sole C and energy source for the bacterial strain *Arthrobacter aurescens* TC1 [63,

228]. Both engineered systems were fed with an AT solution ( $30 \text{ mg L}^{-1}$ ), with dilution rates, for the chemostat, of 0.023, 0.032, 0.048, 0.056,  $0.068 \text{ d}^{-1}$ , and, for the retentostat, of  $0.02 \text{ d}^{-1}$ . For each system at each dilution rate, concentrations of AT [ $\mu\text{g L}^{-1}$ ], HY [ $\mu\text{g L}^{-1}$ ], and living biomass [ $\text{cell L}^{-1}$ ] were reported at steady-state (details in the Supplementary Section 11.2.3). Additionally, the isotope fractionation coefficient ( $\epsilon$ ) was measured at the outlet of the first dilution rate of the chemostat ( $-5.4\text{‰}$ , only at the lowest dilution rate), and retentostat ( $-0.45\text{‰}$ )

### **Calibration strategy**

Our initial intent was to estimate a single set of model parameters for both engineered systems. This was not possible, however, most likely due to differences in bacterial physiology [238, 245]. In our next attempt, we introduced a switch function [169, 170], allowing for environmental-specific transition between the two conditions (chemostat and retentostat) (Supplementary Section 11.2.10). This model, despite its high complexity and many degrees of freedom, was still unable to simulate both engineered systems together (Supplementary Section 11.2.10). Therefore, we exhaustively investigated (using fits for both systems and sensitivity analysis) subsets of parameters that could be kept fixed at the chemostat fit while still capturing bacterial behaviour in the retentostat in a two-step calibration process, as follows.

STEP 1 - pre-calibration step: We started by using the five steady-states (one with each dilution rate) measured in the chemostat, and the isotope fractionation of the lowest dilution rate (16 data points). We considered the parameter ranges shown in Table 5. The nominal values were taken from literature (Table 5). Ranges were selected as to capture parameter variation.

Table 5: Model Parameters

Param	Description	Units	Nominal	Minimum	Maximum
$k_{AT}$	Maximum degradation rate of AT	$\text{d}^{-1}$	0.10(a)	$1 \cdot 10^{-4}$	$1 \cdot 10^4$
$K_M^{AT}$	Half saturation concentration for AT degradation	$\mu\text{g L}^{-1}$	237(b)	$1 \cdot 10^{-5}$	$1 \cdot 10^4$
$k_{HY}$	Maximum degradation rate of HY	$\text{d}^{-1}$	0.10(a)	$1 \cdot 10^{-5}$	300
$K_M^{HY}$	Reference/half-saturation concentration	$\mu\text{g L}^{-1}$	0.05(c)	$1 \cdot 10^{-4}$	$1 \cdot 10^4$
$Y$	Growth yield	–	0.04(b)	0.01	0.15
$m$	Maintenance coefficient	$\text{d}^{-1}$	0.10(b)	$1 \cdot 10^{-4}$	$1 \cdot 10^4$
$f_{cell}$	Conversion factor cell to C	$\mu\text{g cell}^{-1}$	$2.6 \cdot 10^{-8}$ (d)	$4 \cdot 10^{-9}$	$5 \cdot 10^{-7}$
$r_e$	Mass-transfer rate coefficient	$\text{L d}^{-1} \mu\text{g}^{-1}$	0.003(b)	$1 \cdot 10^{-4}$	$1 \cdot 10^8$

Highlighted parameters were estimated for the retentostat system. References: (a) [246], (b) [239], (c) [241], (d) [183].

We used the global optimization algorithm Simulated Annealing [simannealbnd] of MATLAB to minimize the weighted sum of squared errors (SSE) :

$$SSE = \sum_{i=1}^n \frac{(y_{obs}^i - y_{sim}^i)^2}{\sigma_i^2} \quad (74)$$

where  $y_{obs}^i$  and  $y_{sim}^i$  are the mean values per observation type and dilution rate, and the corresponding model output for the  $i^{th}$  data point from  $n$  total data points.  $\sigma_i^2$  is the recalculated standard deviation per observation type and dilution rate (details are given in Supplementary Section 11.2.3, Table S7).

We then calibrated the retentostat system at the steady-state (4 data points) using Simulated Annealing again. An acceptable description could be reached by fixing four parameters and allowing the other four to vary:  $k_{AT}$ ,  $K_M^{HY}$ ,  $f_{cell}$ , and  $r_e$  (highlighted yellow in Table 5). Details are given in the Discussion.

The model outputs corresponding to the measurements were:

$$\text{AT Concentration } [\mu\text{g L}^{-1}] = AT_e^l + AT_e^h \quad (75)$$

$$\text{HY Concentration } [\mu\text{g L}^{-1}] = HY_e \quad (76)$$

$$\text{Biomass } [\text{cell L}^{-1}] = \frac{C_B}{f_{cell}} \quad (77)$$

Isotope fractionation was determined as:

$$\varepsilon = \delta_{inlet} - \delta_{outlet} \quad (78)$$

where  $\delta_{inlet}$  is the isotope ratio of the heavy and the light isotopologues of AT at the inlet, given as  $-29\%$  [62, 228], and  $\delta_{outlet}$  was determined as

$$\delta_{outlet} = \left( \frac{\frac{AT_e^h}{AT_e^l}}{R} - 1 \right) \cdot 1000 \quad (79)$$

where  $R$  is the reference  $^{13}\text{C}/^{12}\text{C}$  isotopoe ratio of Vienna Pee Dee Belemnite [147]. The parameter  $\beta$  (eq. 40) can be directly derived from the enzymatic fractionation coefficient of AT ( $\varepsilon = -5.4\%$ ) measured for a particular bacterial strain [228, 239]:

$$\varepsilon = \beta - 1 \quad (80)$$

**STEP 2 - Full calibration:** For both systems, a full calibration step, including parameter and output uncertainty were determined with the Markov Chain Monte Carlo (MCMC) algorithm of the DREAM MATLAB toolbox [117]. We fitted the 8 chemostat system parameters and the 4 differing retentostat system parameters simultaneously (marked in yellow in Table 5) in one optimization run. We chose a flat and uninformative prior distribution for the MCMC. The starting values of the MCMC chains were drawn from a normal distribution of the parameters

in log-space (mean value equal to the best fit of the Simulated Annealing (step 1), an arbitrary variance of 1, and zero covariance between the parameters). Minimum and maximum parameter values were taken from Table 5, and the option “reflect” was selected as a method for handling parameter boundaries. The  $\hat{R}$ -diagnostic [192] lower than 1.2 [117] was used as convergence diagnostics. We used a Gaussian likelihood considering heteroscedastic measurement errors as implemented in DREAM:

$$L(x|\tilde{Y}) = -\frac{n}{2} \cdot \log(2\pi) - \sum_{i=1}^n \{\log(\sigma_i)\} - \frac{1}{2} \cdot \sum_{i=1}^n \left( \frac{y_{obs}^i - y_{sim}^i}{\sigma_i} \right)^2 \quad (81)$$

### 7.3.4 Soil measurements

#### Soil sampling

Topsoil was sampled from the plough layer (0-30 cm) of two agricultural fields (Poltringen and Tailfingen) in the vicinity of Tübingen, Germany. The soils were classified as Vertic Cambisol on gypsum keuper (Poltringen), and eroded Luvisol (siltic) on loess (Tailfingen). To obtain representative samples, 20 individual samples were drilled by hand down to 30 cm depth and combined in the field to one composite sample. In the lab, samples were thoroughly mixed (using a sample-splitter; Retsch GmbH, Germany), freeze-dried, and ground before further processing (exhaustive extraction and sorption test).

#### Exhaustive soil extraction

Pesticides (atrazine and hydroxyatrazine) were extracted from soil with an accelerated Solvent Extractor (ASE 300 Dionex, Thermo Scientific) at 80°C and 150 bar, using acetone as the main solvent (parameters in Table 6). To ensure a homogeneous flow through the extraction cells, soil samples were mixed with 80% (mass) clean quartz sand before extraction. To control for potential losses of pesticide during the processing (enrichment and clean-up) of the extracts, 10 ng of Isoproturon-D6 were added to each extract. Subsequently, the extracts were reduced with a rotational evaporator until acetone was evaporated completely. The residual aqueous sample



was filtered through 0.25 µm PTFE syringe filters (Agilent, Waldbronn Germany) and 10% (Vol.) of MeOH was added before the measurement at the liquid chromatography–mass spectrometry (HPLC-MS/MS). The target compounds were separated with an Agilent 1290 Infinity HPLC (Agilent, Waldbronn, Germany) using a reversed phase column (Agilent Poroshell 12 EC-C18, 2.7 µm, 2.1 x 100 mm). The quantification of the target compounds was done based on an external calibration using 10 standards with concentrations between 0.02 and 10 µg L<sup>-1</sup>. As control for a potential shift during the measurement, every 15 samples, one external standard was monitored, with a concentration of 2.5 µg L<sup>-1</sup> (Measurements are shown in Supplementary Section 11.2.4, Table S7)

Table 6: Details of accelerated solvent extraction method

Parameter	Settings
Solvent	Acetone:MilliQ(9:1)
Temperature [°C]	80
Pressure [psi]	1500
Heat [min]	5
Static time [min]	10
Flush vol. [%]	70
Purge [s]	100
Static cycles	2

### Sorption test

Six initial concentrations of atrazine (0.06, 0.4, 4, 36, 420 and 2060 µg L<sup>-1</sup>) were prepared from a stock solution of atrazine in MilliQ water (using a pure, analytical standard from Sigma Aldrich). The solutions were spiked with CaCl<sub>2</sub> (0.5 g L<sup>-1</sup>) and NaN<sub>3</sub> (0.25 g L<sup>-1</sup>) to provide a stable ionic strength and minimize bacterial activity. The sorption test was conducted in triplicates in 50 mL glass vials (with teflon-lined caps), containing 15 g of soil and 30 mL of spiking solution. The vials were kept on a horizontal shaker (150 rpm) for 10 days in the dark and at 20°C. To separate soil solids from water, the vials were kept standing for three days until all fine particles

were settled. A small test with filtering the aqueous phase had confirmed this approach as valid. Subsequently, the aqueous phase was transferred into clean vials using glass pipettes.

After separating soil solids from water, 20 ng of atrazine-D6 was added as an internal standard to the aqueous phase. Processing of the aqueous samples varied for the different concentrations: Samples with lowest concentrations were enriched via solid phase extraction (Waters OASIS HLB). Samples with expected concentrations between 0.2 and 10  $\mu\text{g L}^{-1}$  were filtered through 0.25  $\mu\text{m}$  PTFE syringe filters and 2% (Vol.) of acetonitrile was added. For concentrations above 10  $\mu\text{g L}^{-1}$ , the samples were filtered and then diluted with MilliQ:acetonitrile (98:2) before LC-MS/MS measurements. As quality control, blanks with ultra-pure water, leaching blanks with ultrapure-water and soil, and controls with spiking solution without soil were analysed in triplicates confirming no relevant loss of atrazine or contamination (Supplementary Section 11.2.4, Table S8).

We determined the Freundlich sorption parameters ( $K_{AT}$  and  $n_{AT}$ ) for atrazine at both sites by regressing the sorbed concentration on the solution concentration (eq. 61, and 62). We used the Nonlinear regression function [nlinfit] of Matlab (Supplementary Section 11.2.4, Figure S30 and Table S9). The sorption coefficient of hydroxyatrazine ( $K_{HY}$ ) was calculated by dividing the normalized sorption coefficient of atrazine  $K_{AT}^*$  (sorption coefficient  $K_{AT}$  divided by the water solubility of atrazine  $S_{AT}$ ) by the water solubility of hydroxyatrazine ( $S_{HY}$ ) at the power of  $n_{AT}$  (eq. 82) [247, 248]. The sorption exponent for hydroxyatrazine was assumed to be equal to atrazine because the Freundlich exponent is rather soil- than compound- specific:

$$K_{HY} = \frac{K_{AT}^*}{(S_{HY})^{(n_{AT})}} \quad (82)$$

### 7.3.5 Soil predictions

We ran simulations in soils using both sets of calibrated parameters (chemostat and retentostat) for four different model configurations: i) with Monod-kinetics and without leaching (Variant M-NL), ii) with thermodynamic growth constraint and without leaching (Variant T-NL), iii) with

Monod-kinetics and leaching (Variant M-L), iv) with thermodynamic growth constraint and leaching (Variant T-L). We fixed the equilibrium sorption parameters ( $K_F$  and  $n_F$ ) to the means of sorption parameters at the sites Poltringen and Tailfingen (Supplementary Section 11.2.4, Table S9). We fixed the minimum bacterial biomass in soil  $M$  according to Klier et al. [244]. We derived the water flow ( $v_v$ ) from the mean daily water flux of both sites ( $0.56 \text{ mm d}^{-1}$ ) divided by the ploughing depth of 30 cm. The values of the soil parameters are shown in Table 7:

Table 7: Soil Parameters

Param.	Description	Units	Value
$K_{AT}$	Sorption coefficient for AT	$\mu\text{g}^{(1-n_{AT})}\text{kg}^{-1}\text{L}^{n_{AT}}$	3.2
$n_{AT}$	Sorption exponent for AT	–	0.85
$K_{HY}$	Sorption coefficient for HY	$\mu\text{g}^{(1-n_{HY})}\text{kg}^{-1}\text{L}^{n_{HY}}$	17.4
$n_{HY}$	Sorption exponent for HY	–	0.85
$v_v$	Water flow per soil volume in the plough layer	$\text{d}^{-1}$	0.00188
$M$	Minimum bacterial biomass per volume of soil solution	$\mu\text{g L}^{-1}$	0.03

To compare with the field monitoring data from the sites Poltringen and Tailfingen, we ran simulations with all four variants of the soil model, assuming an initial application of  $1,000 \mu\text{g kg}^{-1}$  [249, 250] and predicting residual concentrations after 30 years.

### 7.3.6 Global sensitivity analysis

We determined the Morris and Sobol indices [103, 106, 107, 202] for the two core model variants (M and T), using the SAFE toolbox of MATLAB [202, 251]. We calculated the mean of the elementary effects ( $\mu^*$ ) and the standard deviation of the elementary effects ( $\sigma$ ) for the Morris Method, as well as main and total effects for Sobol indices with a total of 15,000 sample inputs in both cases.

We sampled parameters from a uniform distribution taken from the posterior distribution of the fitted parameters against the chemostat and retentostat data combined (Table 8 from the Results section 7.4). We used a latin hypercube sampling strategy [73]. Additionally for the

Morris Method, we calculated the  $l_2$  norm ( $l_2 = \sqrt{\mu^{*2} + \sigma^2}$ ) of each parameter [73, 105, 107] and considered parameters with  $l_2$ -norm higher than 0.5 as high leverage.

We selected the following outputs: steady state biomass, AT, and HY (extracellular and intracellular), and isotope fractionation  $\varepsilon$  (eq. 78). We ran the model for 200 d to guarantee steady-state in the simulations.

### 7.3.7 Local sensitivity analysis

We performed a local parametric sensitivity analysis [199, 200] for the four soil model variants as described above, based on the best fit against the chemostat and retentostat observations. The target outputs were the residual concentration of AT and HY after 30 years. We addressed all kinetics (Table 5) and soil parameters (Table 7), as well as the initial AT application.

## 7.4 Results

### 7.4.1 Calibration to chemostat and retentostat data

The two core model variants behave equally in engineered environments, and so we present the results only for Variant T. (Results corresponding to Variant M are presented in the Supplementary Section 11.2.6, Figures S32 and S33, and Table S10). Following a two-step approach, we calibrated the 8 chemostat system parameters and the 4 differing retentostat system parameters simultaneously.

#### **Concentrations:**

Our simulations were in good agreement with observed data for the chemostat (Figure 8 A-C). After the partial re-calibration, we found acceptable agreement for the retentostat system, but with a slightly higher model output uncertainty for the biomass (this was not unexpected, given the relative lack of data for calibration).

**Fractionation:**

Simulations showed agreement with the observed isotope fractionation for both systems, with slightly higher uncertainty for the retentostat (Figure 8D). Isotope fractionation of AT occurs when enzymatic transformation is the rate-limiting step. In this case, the enzymatic fractionation coefficient of AT ( $\epsilon$ ) lies close to -5.4‰ (chemostat). At low AT concentrations, the mass transfer across the cell membrane becomes rate-limiting, and no isotope fractionation is observed ( $\epsilon$  of just -0.45‰; retentostat) [63, 228, 239].

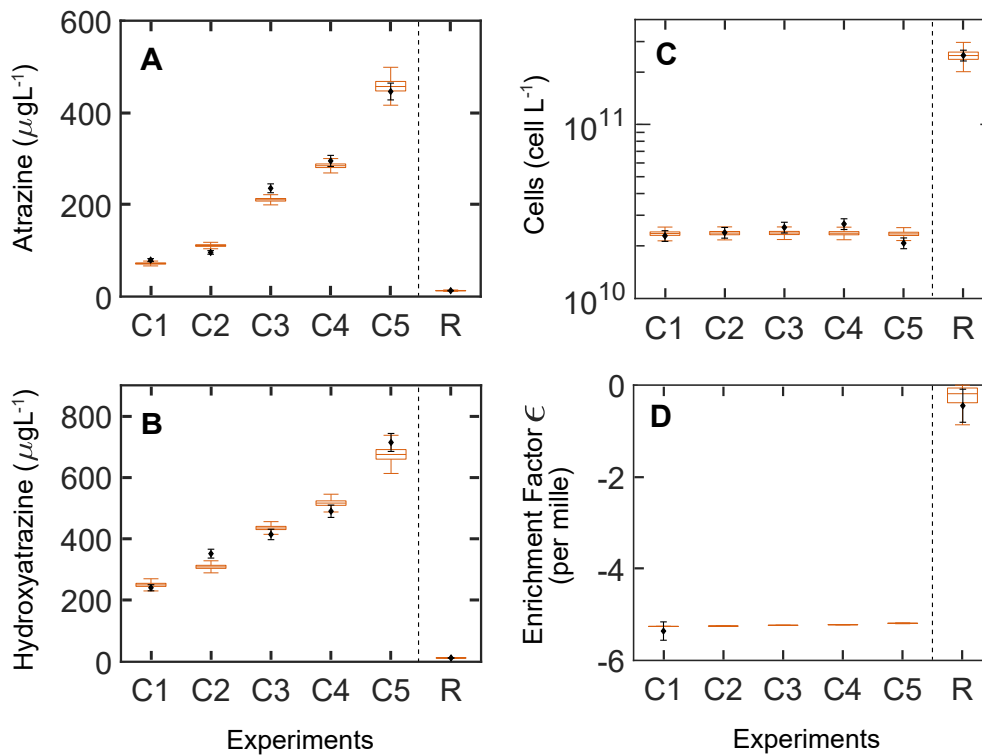


Figure 8: Simulations (boxplots) of model variant T (thermodynamic growth constraints) and measured data (blank diamonds + estimated standard deviation). **A-C.** Steady-state concentrations for the chemostat (five dilution rates: C1-C5: 0.023, 0.032, 0.048, 0.056, 0.068  $\text{h}^{-1}$ , respectively) and the retentostat (dilution rate: R: 0.020  $\text{h}^{-1}$ ). The middle line in the boxplot is the median of the ensemble outputs from the MCMC simulation ensemble (see M&M 7.4); boxes represent 25% and 75% percentiles; whiskers corresponds to  $\pm 1.5 \times \text{IQR}$  (interquartile range). **D.** Enrichment factors ( $\epsilon$ ) were reported only for the lowest dilution rate of the chemostat (C1) and the retentostat (R), but simulated for C2-C5.

**Parameter estimates and uncertainty:**

Kinetic parameters related to AT and HY degradation (chemostat:  $k_{AT}$ ,  $K_M^{AT}$ ,  $k_{HY}$ ,  $K_M^{HY}$ , retentostat:  $k_{AT}$ ,  $K_M^{HY}$ ) appear to be well-informed by the data, showing relatively well-constrained posterior distributions (Figure 9 A, B, C, D, I, J), low standard deviations (Table 8), and considerable impact on model outputs according to the Sobol analyses (especially  $k_{AT}$  and  $K_M^{AT}$ , Supplementary Figure S35). The maintenance parameter  $m$  was interestingly well constrained by the chemostat data (Figure 9 E, Table 8); the global sensitivity analysis confirmed this parameter to be low leverage (Supplementary Figure S33 and S34). The mass-transfer rate parameter  $r_e$  was not well-constrained for the chemostat data (Figure 9 H), but fitted relatively well to the retentostat data (Figure 9 L), especially with the model variant M (Section 11.2.6, Figures S33). This parameter showed a considerable impact on model outputs (Supplementary Figure S34 and S35). The yield parameter  $Y$  and conversion parameter  $f_{cell}$  were highly uncertain and not well-constrained for either system, probably due to the high correlation with other parameters like kinetic parameters  $k_{AT}$  and  $k_{HY}$  (Supplementary Tables S11 and S12 for model variant M).

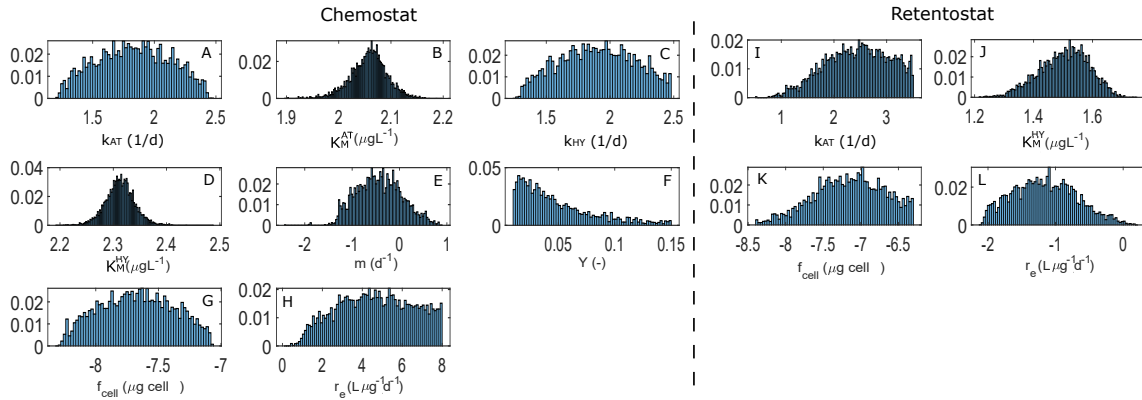


Figure 9: Posterior distributions for calibrations with chemostat (8 parameters) and retentostat (4 parameters) data. All parameters are expressed in log scale with the exception of the growth yield  $Y$ .

### Comparison of parameter estimates between the chemostat and retentostat:

Comparing the mean and MAP calibrated parameter values in Table 8, we see that the per-cell AT degradation rate ( $k_{AT}$ ) is estimated to be higher for cells living in the low nutrient retentostat system. Conjectured physiological adaptations [238] in the retentostat environment may be responsible for the difference in the estimated value of  $f_{cell}$  compared to the chemostat, reflecting changes in cell volume. However, this estimate is highly uncertain and highly correlated to other parameters' values. Physiological adaptations might also be responsible for a reduced value of parameter  $K_M^{HY}$  in the retentostat system, possibly reflecting a change in nutrient demand. The estimate of  $r_e$  was higher in the chemostat than in the retentostat, indicating a change in membrane properties leading to strong mass transfer limitations across the cell membrane. The estimates of the parameters  $k_{AT}$ ,  $K_M^{HY}$  and  $r_e$  using model variant M show the same tendencies, but exhibit stronger changes (increase/decrease) from chemostat to retentostat (Supplementary Section 11.2.6, Table S10). The main difference is in parameter  $f_{cell}$ , which shows a clear reduction in the retentostat, strongly supporting the findings of Kundu et al. [238].

### 7.4.2 Predictions of atrazine and hydroxyatrazine fate in soils and comparison against field data

We simulated the fate of AT and HY in soil for 30 years, assuming a single initial AT input of  $1,000 \mu\text{g kg}^{-1}$  [249, 250]. For this, we used the full posterior parameter estimates from the chemostat and retentostat systems for four model variants (Figure 10). All model variants predicted very low residual AT concentrations, considerably underestimating the observed concentrations of  $0.3$  and  $0.6 \mu\text{g kg}^{-1}$  in the top soil of both field sites (Poltringen and Taiflingen respectively) (Figure 10 A, B).

In contrast, predictions of residual HY mainly overestimated the observed HY concentration at both study sites (around  $2 \mu\text{g kg}^{-1}$  in both sites) (Figure 10 C, D). Predictions using retentostat fitted parameters in combination with thermodynamically constrained growth and leaching (Figure 10 D) predicted long term persistence of HY, with mean values around  $36 \mu\text{g kg}^{-1}$ .

Table 8: Calibrated parameter values of the model variant T based on an MCMC ensemble for Chemostat and Retentostat fits

Parameters	Chemostat				Retentostat			
	MAP	Mean	Std	Range	MAP	Mean	Std	Range
$k_{AT}$ [ $\text{d}^{-1}$ ]	71.1	67.9	2.0	[16.4 – 289.3]	239.4	276.2	4.0	[3.3 – $3.2 \cdot 10^3$ ]
$K_M^{HY}$ [ $\mu\text{g L}^{-1}$ ]	199.9	205.4	1.1	[155.9 – 306.7]	26.9	31.9	1.2	[16.4 – 55.9]
$f_{cell}$ [ $\mu\text{g cell}^{-1}$ ]	$1.9 \cdot 10^{-8}$	$2.1 \cdot 10^{-8}$	2.0	[ $4.9 \cdot 10^{-9}$ – $8.6 \cdot 10^{-8}$ ]	$2.8 \cdot 10^{-8}$	$6.9 \cdot 10^{-8}$	2.9	[ $4.0 \cdot 10^{-9}$ – $5.0 \cdot 10^{-7}$ ]
$r_e$ [ $\text{L d}^{-1} \mu\text{g}^{-1}$ ]	544.7	$4.8 \cdot 10^4$	71.2	[ $1.5 \cdot 1.0 \cdot 10^8$ ]	0.2	0.1	2.9	[ $7.8 \cdot 10^{-3}$ – 1.5]
$K_M^{AT}$ [ $\mu\text{g L}^{-1}$ ]	108.6	114.2	1.1	[78.9 – 157.2]				
$k_{HY}$ [ $\text{d}^{-1}$ ]	83.3	79.2	2.0	[17.7 – 299.9]				
$m$ [ $\text{d}^{-1}$ ]	1.0	0.4	3.0	[0.0 – 7.5]				
$Y$ [-]	0.04	0.04	2.0	[0.01 – 0.1]				

Maximum a-posteriori (MAP), mean, standard deviation (Std) and range of the MCMC ensemble for both engineered systems. Calibrated parameters of model variant M are shown in the Supplementary Section 11.2.6, Table S10



However, model variants with Monod kinetics (M-NL(R) and M-L(R)) performed better and predicted residual HY concentrations much closer to the measurements (9 and 20  $\mu\text{g kg}^{-1}$ , Figure 10 D).

As expected, simulations of this simple model over 30 years are highly uncertain. Based on our local sensitivity analysis (Supplementary Section 11.2.9, Table S13 and S14), the sorption exponent of both chemicals ( $n_{AT}$  and  $n_{HY}$ ) showed the highest impact on the residual concentrations of AT and HY after 30 years, revealing a strong dependency on sorption characteristics of the soils. Surprisingly, the initial application of AT only impacted the residual concentration of AT in model variants incorporating thermodynamic growth constraints. Water flow ( $v_v$ ) and minimum bacterial biomass ( $M$ ) had low impact on the residual concentrations, despite their role to improve model predictions (our best model predictions include leaching; recall that parameter  $M$  accounts for alternative carbon sources for soil bacterial biomass). As to be expected, the kinetic parameters, in contrast to the sorption parameters, had a negligible impact on the target outputs.

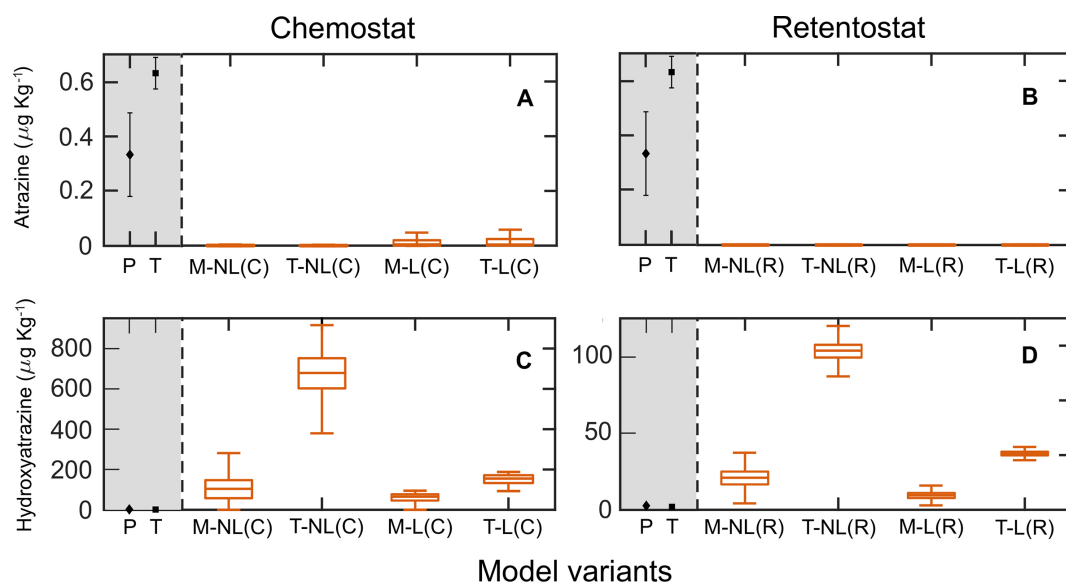


Figure 10: Simulated residual concentrations of atrazine AT (A,B) and hydroxyatrazine HY (C,D) in topsoils (0-30 cm), and observations (grey background) in topsoils of two field sites Poltringen (P) and Taifingen (T) after 30 years. (C) represent simulations using chemostat fitted parameters, and (R) simulations using retentostat fitted parameters. The middle line of boxplots is the median of the ensemble outputs from the MCMC simulation ensemble (see M&M 7.4); boxes represent 25% and 75% percentiles; whiskers corresponds to  $\pm 1.5 \times \text{IQR}$  (interquartile range).

## 7.5 Discussion

### 7.5.1 Bacterial adaption to low nutrient availability affects model parameterization

Due to the apparent similarities between the chemostat and retentostat systems, our initial intent was to achieve a joint fit for both systems. In particular, by including a flexible formulation of the mass-transfer rate, as well as a thermodynamically constrained growth rate instead of a Monod formulation, we aimed to represent systems with or without mass-transfer limitations across the cell membrane by one model. However, we found that goal unattainable. Recent publications [238, 245] show evidence of a phenotypic differentiation of a single population into separate growing and non-growing (i.e. energy used only for maintenance) bacterial subpopulations [238]. Thus, we focused on the key parameters that have to be re-calibrated

between the two systems using two model variants that exhibit equivalent behavior over the range of inputs in the engineered systems (Table 8).

After fitting model parameters to the chemostat data, we systematically tested which parameters had to be re-calibrated to capture the retentostat behavior. We were guided by sensitivity analyses, as well as our understanding of the role of the parameters in our model. We fixed the maximum degradation rate of HY ( $k_{HY}$ ), the growth yield ( $Y$ ), the half-saturation concentration for AT degradation ( $K_M^{AT}$ ) and the maintenance parameter ( $m$ ) because of their low impact on model output (Supplementary Information, Figure S33 and S34). Similar sensitivities were previously reported in the literature [239]. Summing up, the parameters that had to be re-calibrated to capture the retentostat behavior are:  $k_{AT}$ ,  $K_M^{HY}$ ,  $f_{cell}$ , and  $r_e$ . We justify the requirement of these needed adjustments in the following.

The parameters  $k_{AT}$  (maximum degradation rate of AT), and  $K_M^{HY}$  (reference/half-saturation concentration) represent physiological features that can be expected to change under starvation conditions [238, 252]. Relative to the chemostat conditions, in the low-HY retentostat environment, we estimate a higher values of  $k_{AT}$  and lower values of  $K_M^{HY}$  (Table 8), indicating faster AT transformation to HY, and physiological adaptation of microorganisms to use of HY, respectively. While the fitted value of  $k_{AT}$  was about twice the value of  $k_{HY}$  (maximum degradation rate coefficient of HY) in the retentostat, both parameters ( $k_{AT}$  and  $k_{HY}$ ) were similar in the chemostats (Table 8). This difference in the parameterization of both systems shows that the physiological adaption of microorganisms to low concentrations affects the regulation of the AT degradation reaction network such that HY transformation becomes rate-limiting for microbial growth.

We found that re-calibration of the parameter  $f_{cell}$  is an efficient way to capture specific bacterial differentiation for low nutrient systems [63]. The parameter  $f_{cell}$  is a scaling factor used to convert cells to C [183] and might suggest morphological changes (shape and volume) observed in *Arthrobacter aurescens* to cope with stressful starvation conditions [136, 252]. Due to the high uncertainty in parameter estimation, more experiments are needed to identify the

underlying mechanism.

Changes in the value of  $r_e$  (mass-transfer rate coefficient) between chemostat and retentostat system could reflect morphological/physiological changes in the cell membrane (Table 8). The relatively lower value of  $r_e$  in the retentostat suggests a strong mass-transfer limitation across the cell membrane in that case.

### 7.5.2 Pesticide persistence in soil

The main objective of our work was to accurately represent Atrazine (AT) degradation in soils, and especially to capture the long-term persistence of AT and its main metabolite Hydroxyatrazine (HY).

Despite the related uncertainty for long-term predictions, persistence of HY even after 30 years was consistently predicted by model variant M-L calibrated with retentostat data (Figure 10 D). In general, retentostat concentrations are closer to the soil environment, so that more accurate predictions are to be expected (biomass retention, low nutrient levels). Additionally, incorporation of leaching gave a better representation of the pesticide losses over time. Simulation with a simple model incorporating only leaching over the 30 years leads to a residual concentration of AT of about  $2 \mu\text{g kg}^{-1}$ . This value is close to the measured residual concentrations indicating that only low AT degradation might have occurred at the field sites (Figure 10). Standard Monod model variants predicted HY concentrations after 30 years better than thermodynamic models. Therefore, energetic constraints of microbial growth likely do not limit HY degradation in soils. In contrast, all model configurations predicted a nearly complete consumption of AT after 30 years, a behavior not observed in field surveys [25, 57], including the field measurements of this study (Figure 10 A).

A range of biological and physical processes in soil have been hypothesized as potential mechanisms of pesticide persistence in real systems. These include physicochemical control mechanisms limiting bioavailability, such as chemisorption onto humic substances [59], physical stabilization in soil microaggregates [253], or the spatial encounter of substrates

and degraders [60]. Including these additional mechanisms by applying better sorption and stabilization model formulations [153, 254–258] and spatially resolved modeling approaches [259–262] might further improve predicting the persistence of AT and other pesticides in soil. Our study investigated to what extent mass-transfer limitations and bioenergetic constraints can explain the long-term fate of atrazine and its major metabolite hydroxyatrazine in soils. We found evidence against the hypothesis that passive diffusion across the cell membrane of bacterial degraders limits atrazine degradation in the long term. Atrazine is not degraded to HY for the energy gain by microorganisms and our results suggest that sorption-limited bioavailability and not energetic growth constraints control the persistence of hydroxyatrazine. Hence, standard Monod kinetics for bacterial growth can predict the long-term fate of organic chemicals if soil microorganisms directly utilize them as an energy source. Further research should prioritize the analysis of energetic costs of biogeochemical transformations without a direct microbial energy gain (atrazine dechlorination).

## **8 Temperature and soil moisture change microbial allocation of pesticide-derived carbon (Paper 3)**

This chapter includes the following publication submitted as:

Johannes Wirsching, Luciana Chavez Rodriguez, Franziska Ditterich, Holger Pagel, Rushan He, Christian Zwiener, Marie Uksa, Ellen Kandeler, Christian Poll (2021). Temperature and soil moisture change microbial allocation of pesticide-derived carbon. *Environmental Science & Technology*.

with the following modification:

1. Numbers of figures, tables, and equations are relative to this thesis and not to the original publication.

## 8.1 Abstract

The influence of temperature and soil moisture on the mineralization of pesticides has been studied extensively, indicating, in most cases, longer half-live times ( $DT_{50}$ ) in soils with lower temperature and moisture. However, how the underlying metabolic processes of specific degrader microorganisms change under altered environmental conditions (temperature and soil moisture) is yet unknown. This study aimed to link changes in carbon (C) use efficiency ( $CUE$ ) under optimal (20°C, pF 1.8) and limiting conditions (10°C, pF 3.5) to the activities ( $tfdA$  mRNA) and abundances ( $tfdA$  DNA) of pesticide degraders during  $^{14}C$ -labeled 2-methyl-4-chlorophenoxyacetic acid (MCPA) degradation. We performed a laboratory incubation experiment at two MCPA concentrations (1 and 20 mg kg<sup>-1</sup>) and used a mechanistic gene-based biodegradation model to support data interpretation. After four weeks, mineralization reached almost 70% under optimal conditions but less than 25% under limiting conditions. Estimated  $CUE$  and measured  $tfdA$  genes suggest a metabolism that favors anabolic processes under limiting conditions and reallocation of MCPA-C from growth to tolerance mechanisms. Our work suggests that, at low initial concentrations, the derivation of  $DT_{50}$  values should not be based on mineralization kinetics alone, since they fail to account for the contribution of more efficient carbon utilization, leading to overestimation of the residence time.

## 8.2 Introduction

Over 80% of 317 topsoils tested in the European Union contained pesticide residues that, in some cases, exceeded predicted environmental concentrations [55]. The persistence of pesticides in soil depends, amongst other processes, on its biodegradation under varying environmental conditions. Degradation rates are often determined from mineralization of  $^{14}C$ -labeled pesticides [211, 263–265]. However, the problem with assessing pesticide persistence based on  $^{14}C$  mineralization is that it neglects shifts in microbial C-allocation depending on temperature and soil water content, which may result in an overestimation of the half-life ( $DT_{50}$ ).

Carbon use efficiency (CUE) describes microbial C-allocation into either CO<sub>2</sub> respiration or microbial biomass. Generally, CUE increases with declining temperature, indicating an increase in relative allocation of assimilated C to growth [266]. Given this understanding, and despite information in the literature that a temperature increase of 10°C can accelerate pesticide mineralization by a factor of almost 2 [211, 267–269], lower temperatures might not always lead to significantly decreased degradation; they may instead indicate a C redistribution within the microbial cell where less CO<sub>2</sub> is emitted, as more of it is used to build additional living biomass [266].

Muskus et al. [270] found that a temperature drop from 20°C to 10°C resulted in less mineralization of labeled <sup>13</sup>C<sup>15</sup>N-glyphosate, but promoted the formation of <sup>13</sup>C non-extractable residues (NER; proteins + other remaining biomass residues (bioNER) + sorbed and sequestered starting compounds (xenoNER)). However, if BioNERs are determined at the end of an experiment, after the death of the microorganisms, information on C uptake dynamics during pesticide turnover is lost. In such a case, uptake of pesticide-derived C is only considered as an additional C reservoir and not as a driver of decomposition. The dynamics of CUE, i.e., the proportion of the substrate that, over time, immediately goes into the microbial biomass in relation to the C lost as CO<sub>2</sub> [271], could be an important addition to the conventional mineralization-based approach and provide a more accurate assessment of pesticide degradation at different temperatures.

Soil moisture content is one of the most important factors regulating biological activities in soils [264] and serves as a solubilizer for the movement and distribution of pesticides [272]. According to Pinheiro et al. [273], below the centimeter scale, the fate of pesticides in soils depends on the spatial distribution of pesticide and degrader microorganisms. In unsaturated soils, where the contact between pesticide and microorganisms is established only by diffusion or mass flow [274] due to the heterogeneous soil matrix [275], molecular diffusion represents the dominant mode of transport at the smallest spatial scale. Furthermore, the diffusion of dissolved substances, e.g., pesticides, is limited by the proportion of water-filled pores [276]



and tortuosity; determination of water content makes it possible to compare microbial reaction rates between diffusion- and non-diffusion-limited systems [264].

Most of the relevant literature reports that increasing water content, within a soil water potential range of -1.5 to -0.015 MPa [264, 277], intensifies the degradation of pesticides [211, 278]. For example, MCPA persists ten times longer in dry soils than in moist soils [279] due to the moisture-sensitive exponential growth of microbes [211]. As aridity increases, microorganisms must invest more energy [280] to overcome the suction holding the water in the soil [281]. As energy requirements in drier soils may therefore increase, it is conceivable that the way microbes allocate C will have a profound impact on pesticide mineralization rates. This may mean that, due to physiological trade-offs between C-assimilation and dissimilation under drought conditions [282],  $^{14}\text{CO}_2$  mineralization may not be linearly related to total pesticide turnover [283].

Specifically, drought-tolerant microbes invest heavily in the formation of extracellular enzymes to maintain carbon uptake for the synthesis of stress response compounds such as osmolytes, cryoprotectants, and chaperones [282] to stabilize cell pressure [284]. This would imply that pesticide mineralization is not synonymous with microbial pesticide degradation since microbial C utilization plays a decisive role. This mechanism has already been demonstrated for soil turnover by Zeglin et al. [283], in which soil C sequestration was higher under dry conditions.

Degradation rate and C allocation are affected not only by environmental conditions, but also by initial pesticide concentration [285, 286]. Pesticide degradation at low concentrations usually follows a first order kinetics and is often astonishingly fast [287, 288]. In contrast, pesticide degradation at high concentrations is slower and accompanied by a simultaneous increase in degradation activity and genetic degradation potential [131]. With respect to C-allocation, in a previous study we [289] demonstrated that the predominantly catabolic use of MCPA at concentrations  $\leq 1 \text{ mg kg}^{-1}$  shifted towards an gradually increasing anabolic metabolism at concentration  $\geq 5 \text{ mg kg}^{-1}$ . If initial pesticide concentration determines C use by microbial

pesticide degraders, the impacts of temperature and moisture on C allocation may also depend on pesticide concentration.

The objective of this study was to examine the impact of environmental factors (temperature and soil moisture) on pesticide degradation and derived carbon use efficiency, and to link them to the associated abundances and activities of degraders. Specifically, we were interested in the deviations in half-lives derived from actual concentration decrease and mineralization kinetics. To address these questions, we used  $^{14}\text{C}$ -labeled MCPA, a weakly adsorbed pesticide [290] that is readily soluble in water [291] and highly biodegradable [292]. An additional advantage of using MCPA is that the entire degradation pathway and the functional genes involved have been characterized [128]. We hypothesized that increased *CUE* at lower temperatures and soil moisture i) leads to increased  $^{14}\text{C}$  content in the microbial biomass, and ii) could demonstrate a significant overestimation of pesticide half-life time when derived only from  $^{14}\text{C}$  mineralization curves. We expected iii) that the effect size will be more pronounced at higher initial concentration. To address these hypotheses, we determined the temporal relationship between mineralization ( $^{14}\text{CO}_2$ ) and biomass ( $^{14}\text{C}_{\text{mic}}$ ) formation in distinct phases, *CUE*, alteration of MCPA degradation activity (expressed genes), MCPA-degrading genetic potential (functional genes), and the half-life of MCPA under optimal and limiting environmental conditions. In addition to the experimentally-based *CUE*, we applied a gene-centric model with the experimental data and calculated two model-based carbon use efficiencies for interpretation of pesticide-derived C utilization of the specific pesticide degraders.

## **8.3 Materials and Methods**

### **8.3.1 Soil origin and sampling**

The study site was in the central region of the Ammer catchment in southwest Germany (48°33'24.664", 8°52'31.259"). Soil samples were taken in March 2019 from an Ap-horizon (0-5 cm) of a silty Luvisol (World Reference Basis for Soil Resources). According to the farmer's records of their cultivation and spraying programs, MCPA was never applied to the agricultural

field. The main pesticides applied were chloridazone and metamitron. After sampling, the soil was sieved (<2 mm), homogenized, and stored at -20 °C ( -80 °C for *mRNA* samples) to prevent further biological reactions. The main characteristics of the soil are shown in Table S15.

### 8.3.2 Experimental Design

The experimental set-up consisted of two temperatures (10 and 20°C), two water treatments (pF 1.8 and 3.5), and three concentrations of ring <sup>14</sup>C-labeled MCPA (0, 1, and 20 mg kg<sup>-1</sup> soil). In this study, we defined 10°C, pF 3.5 and the two lowest MCPA concentrations as limiting conditions, and 20°C, pF 1.8 and the highest MCPA concentrations as optimal conditions. The experimental set-up consisted of 36 total samples, with three replicates for each treatment. At three time points (5, 15 and 28 days) representing specific phases of MCPA mineralization, i.e., initial lag phase, phase of exponential growth, and final saturation phase, we sampled independent sets of microcosms. MCPA solution with a <sup>14</sup>C activity of 15 kBq (99% purity, specific activity 50-60 mCi mmol<sup>-1</sup>; BIOTREND Chemikalien GmbH, Germany) was uniformly applied to adjust gravimetric soil water content to 39.6% (pF 1.8) and 29.1% (pF 3.5). Subsequently, after thoroughly mixing the soil with the MCPA solution, cylinders (diameter = 5.6 cm, height = 4 cm) were filled with 100 g of soil and compacted to a bulk density of 1.2 g cm<sup>-3</sup> (height of the soil core was 3 cm). In addition, there was one set of 36 microcosms that contained unlabeled MCPA (analytical MCPA purity 99.2%, Sigma-Aldrich, Germany), from which a series of subsamples (on every second day) for <sup>14</sup>C-free RNA/DNA co-extraction and MCPA quantification was taken and stored at -20°C until the analyses.

### 8.3.3 MCPA analysis in soil

A soil suspension of two g soil mixed with 10mL methanol/water (1:1) was homogenized on a horizontal shaker at 200 rev min<sup>-1</sup> for 10 min, then heated in a water bath for 60 min at 50 °C. After shaking again at 200 rev min<sup>-1</sup> for 10 min, the mixture was centrifuged at 2500 g for 10 min, and 2 mL of the supernatant was removed and filtered (0.45 µm pore size). The extraction

recovery of MCPA was >98%. Before the HPLC-MS/MS analysis, the extracts were sonicated and homogenized for 5 min by vortexing. On a 1260-Infinity system from Agilent Technologies, one  $\mu\text{L}$  of the sample was injected onto a reversed-phase column (Agilent Poroshell 120 C18, 2.1 mm internal diameter, 100 mm length, 2.7  $\mu\text{L}$  particle size) at a temperature of 40°C. MCPA was eluted isocratically within 5 min using 50% water and acetonitrile (both acidified with 0.1% formic acid) at a flow rate of 0.4  $\text{mL min}^{-1}$ . After chromatographic separation, MCPA was detected by tandem mass spectrometry using an Agilent 6490 iFunnel Triple Quadrupole (QQQ) instrument. The analyte was ionized using negative electrospray ionization (ESI-) by applying 12  $\text{L min}^{-1}$  sheath gas ( $\text{N}_2$ ) at 400°C, 16  $\text{L min}^{-1}$  drying gas ( $\text{N}_2$ ) at 150°C, 30 psi nebulizer pressure, 4.2 kV capillary voltage, and 1.2 kV nozzle voltage. MS/MS experiments were conducted by MRM (Multi Reaction Monitoring), using  $\text{N}_2$  as collision gas and collision energy (CE) dependent mass transitions (MCPA: quantifier 198.9/140.9 at 10 eV, qualifier: 198.9/34.9 at 45 eV). The limit of quantification (LOQ) was defined as 13  $\mu\text{g kg}^{-1}$  MCPA in soil.

#### **8.3.4 MCPA mineralization ( $^{14}\text{CO}_2$ )**

The  $^{14}\text{CO}_2$  evolution from the microcosms was determined via titration (DIN EN ISO 16072:2011-09). First, a 0.5 mL aliquot was taken from a  $\text{CO}_2$  trap containing 2 mL 1M NaOH which was set up in the microcosm. The actual respiration measurement was carried out by adding 0.5 mL of 1 M  $\text{BaCl}_2$  and two drops of phenolphthalein. In the following titration with 0.1 M HCl, the end point of the neutralization reaction was indicated by a color change to transparency. To determine the  $^{14}\text{CO}_2$  content, an aliquot of 1 mL was taken from the same  $^{14}\text{CO}_2$  trap and mixed with 4 ml scintillation liquid (Rotiszint Eco Plus, Carl Roth GmbH + Co. KG) in a 5 mL scintillation vial (LDPE). The decay rate in Bequerel (Bq) was measured using a scintillation counter (Wallac 1411, liquid scintillation counter, USA). To account for interfering substances, a quenching adjustment with  $^{14}\text{C}$  aqueous standards was used to improve the accuracy of the actual counts per second (cps) for the entire energy band.

The half-lives ( $\text{DT}_{50\text{MIN}}$ ) derived from the cumulative mineralization curves were calculated

from the estimated parameters of the fitted eqs.83 and 84 (see also Duo-Sen and Shui-Ming [293] and Wirsching et al. [289]):

$$C = C_0 \cdot \left( 1 - \frac{1}{(1 - f_k) \cdot e^{k_1 \cdot t} + f_k} \right) \quad (83)$$

$$DT_{50MIN} = \frac{1}{k_1} \cdot \ln \left[ \frac{1}{1 - f_k} + 1 \right] \quad (84)$$

Where C is the MCPA-derived  $^{14}\text{CO}_2$  (% of MCPA initially applied),  $C_0$  is the total mineralizable MCPA that was not immediately incorporated into the microbial biomass or bound to the soil organic matter after application (% of MCPA initially applied),  $k_1$  ( $\text{d}^{-1}$ ) is the rate constant of MCPA degradation per day, and  $f_k$  is a dimensionless parameter constrained between 0 and 1 [289].

### 8.3.5 Microbial biomass ( $C_{\text{mic}}$ )

Microbial biomass was estimated using the chloroform fumigation extraction method (CFE) developed by Vance et al. [294], adapted by Poll et al. [295] for additional  $^{14}\text{C}$  determination. Prior to extraction, 10 g soil was first weighed to ensure the release of the microbially bound C after a 24-hour fumigation with ethanol-free chloroform. After removal of the chloroform, 40 mL of 0.5 M  $\text{K}_2\text{SO}_4$  solution was added, shaken on a horizontal shaker at  $250 \text{ rev min}^{-1}$  for 30 min and centrifuged at 4420 g for 30 min. The clear supernatant was then passed through a 20  $\mu\text{m}$  filter, diluted 1:4 with deionized water to avoid high salinity during detection, and measured with a Multi-N/C 2100S TOC-TNb analyzer (AnalytikJena, Jena, Germany). A second subsample of 10 g soil was not fumigated with chloroform to determine only the amount of extractable organic carbon. C content of the contro (non-fumigated) samples was subtracted from the C content of the fumigated samples to determine  $C_{\text{mic}}$  content. The  $k_{\text{EC}}$  factor of 0.45 was used to estimate the extractable portion of microbial biomass C [296]. To obtain the  $^{14}\text{C}$  content in  $C_{\text{mic}}$ , 1 mL of the CFE supernatant was mixed with 4 mL scintillation liquid

(Rotiszint Eco Plus, Carl Roth GmbH+Co. KG) in a 5 mL scintillation vial (LDPE). Calculation of the incorporated  $^{14}\text{C}$  was performed as described for the  $C_{\text{mic}}$  content, here using the activity difference between the fumigated and non-fumigated samples. For the non-fumigated samples, the undiluted supernatant was used. Total  $^{14}\text{C}$  utilization was estimated by adding the  $^{14}\text{CO}_2$  mineralization and  $^{14}\text{C}$  incorporation on days 5, 15 and 28.

### 8.3.6 MCPA degrader abundance and activity

#### DNA/RNA co-extraction

For RNA and DNA extraction, 2 g frozen soil was weighed into 15 mL bead-beating tubes and extracted using the RNeasy PowerSoil Total RNA Kit for soil (Qiagen, Germany) and the RNeasy PowerSoil DNA Elution Kit (Qiagen, Germany) in a co-extraction method following the user manual. After extraction, the RNA and DNA samples were aliquoted and stored at  $-80^\circ\text{C}$  (RNA samples) or at  $-20^\circ\text{C}$  (DNA samples) for further use. The DNA and RNA concentrations were measured using a fluorescent dye and microplate reader (Synergy HTX Multi-Mode Reader, Bio-Tek Instruments Inc., Germany). For DNA and RNA quantification, the Quant-iT<sup>TM</sup> PicoGreen<sup>TM</sup> dsDNA Assay Kit and the Quant-iT<sup>TM</sup> RNA Assay Kit (Thermo Fisher Scientific, Germany), respectively were used following the user manuals. Before using the RNA-samples for Real-Time quantitative PCR (qPCR), possible remaining DNA in the RNA samples was digested using the TURBO DNA-free<sup>TM</sup> Kit (Invitrogen, Thermo Fisher Scientific, Germany) following the manufacturer's protocol (Table S1). After digestion, RNA samples were divided in two subsamples of 11  $\mu\text{L}$  each and labeled as + and - subsamples. For the following reverse transcription, the SuperScript<sup>TM</sup> III Reverse Transcriptase Kit with Random Primers and RNase (Invitrogen, Thermo Fisher Scientific, Germany), and dNTPs (10mM; Genaxxon, Germany) were used. The reaction was carried out according to the user manual of SuperScript<sup>TM</sup> III Reverse Transcriptase from Invitrogen. Reaction conditions and temperature program are described in Table S16. The + subsamples served as cDNA, whereas the - subsample served as a negative control for the remaining DNA after digestion (Table S17).

### Real-Time quantitative PCR (qPCR)

For gene quantification (bacterial 16S rRNA and functional genes), qPCR assays were applied using an ABI Prism 7500 Fast system (Applied Biosystems, Germany) with SYBR Green detection. The primer and qPCR conditions used are listed in Table S18. Each SYBR Green reaction contained 7.5  $\mu\text{L}$  of Power SYBR® Green PCR master mix (Applied Biosystems, Germany), 0.75  $\mu\text{L}$  of each primer (5  $\mu\text{M}$ ), 0.375  $\mu\text{L}$  of T4gp32 (MP Biomedicals, Germany), 3.625  $\mu\text{L}$  water and 2  $\mu\text{L}$  diluted template DNA or cDNA (3  $\text{ng } \mu\text{L}^{-1}$ ) for functional genes (*tfdA* and *cadA*). For 16S rRNA, 1  $\mu\text{L}$  diluted template DNA or cDNA (3  $\text{ng } \mu\text{L}^{-1}$ ) and 4.625  $\mu\text{L}$  of water was used. For quantification, standard plasmid DNA was used with a dilution series from  $10^8$  to  $10^1$  copies  $\mu\text{L}^{-1}$  according to Ditterich et al. [131]. *CadA* showed no response to MCPA addition and was therefore not discussed further in the course of the study.

### 8.3.7 Gene-centric modeling of MCPA biodegradation

We used a recently developed modeling approach (ref. to Chavez Rodriguez et al. [297]) to simulate MCPA mineralization ( $C_P$  [ $\text{mmol g}^{-1}$ ]), *tfdA* genes (proxy for active bacterial biomass  $C_B$  [ $\text{mmol g}^{-1}$ ]) and transcripts, and CUE. The original modeling approach was extended to account for constitutive gene expression and to include a temperature response function.

#### Model description

We assumed gene expression to be in quasi-steady state described by the Hill function, including constitutive gene expression that is potentially important at low concentrations:

$$\widehat{mRNA} = f_T \cdot \left( \frac{(C_P^L)^{n_H}}{(C_P^L)^{n_H} + (K_G)^{n_H}} + \frac{\alpha}{f_T} \right) \quad (85)$$

where  $f_T$  represents the number of transcripts per gene,  $n_H$  [-] and  $K_G$  [ $\text{mmol cm}^{-3}$ ] are the Hill exponent and constant respectively,  $\alpha$  is the constitutive gene expression coefficient set to  $1.2 \cdot 10^{-5}$  transcripts per gene [126], and  $C_P^L$  [ $\text{mmol cm}^{-3}$ ] is the solution phase concentration of

MCPA.

Microbial growth is regulated in three ways (Eq. 98) by: i) MCPA-dependent *tfdA* gene transcription (mRNA, Eq. 85), ii) MCPA-dependent reaction kinetics (Monod term in Eq. 98), and iii) a  $Q_{10}$  temperature response function (Eq. 87).

$$r_{growth} = \mu_{max} \cdot C_B \cdot \left( \frac{(C_P^L)^{n_H}}{(C_P^L)^{n_H} + (K_G)^{n_H}} + \frac{\alpha}{f_T} \right) \cdot \left( \frac{C_P^L}{C_P^L + K_M} \right) \cdot f_R(T) \quad (86)$$

where  $\mu_{max}$  [ $d^{-1}$ ] is the maximum growth rate coefficient,  $K_M$  [ $mmol\ cm^{-3}$ ] is the Monod constant, and  $f_R(T)$  [-] is the temperature response function.

The temperature response function  $f_R(T)$  from Sierra et. al. [298] influences not only microbial growth, but also the decay rate (Eq. 88), maintenance rate (exogenous (Eq. 89) and endogenous (Eq. 90)), and decay rate of non-extractable residues (Eq. 91), and is defined as:

$$f_R(T) = (Q_{10})^{\left( \frac{T - 10^\circ C}{10^\circ C} \right)} \quad (87)$$

where  $Q_{10}$  [-] is the temperature function constant, and T is the temperature in  $^\circ C$ .

The decay rate is defined as:

$$r_{death} = C_B \cdot a_a \cdot f_R(T) \quad (88)$$

where  $a_a$  [ $d^{-1}$ ] is the decay rate coefficient.

The total maintenance rate is partitioned into two different maintenance fluxes: exogenous and endogenous [299]. The exogenous flux describes the fraction of the total maintenance demand that can be met with the available MCPA in the system. We modeled this flux using a simple Michaelis Menten expression [261]:

$$r_{m-exogenous} = \left( \frac{m \cdot C_P^L}{C_P^L + K_M} \right) \cdot C_B \cdot \left( \frac{\rho}{\theta} \right) \cdot f_R(T) \quad (89)$$



where  $m$  [ $\text{d}^{-1}$ ] is the maintenance rate coefficient.

The endogenous maintenance flux describes the fraction of the demand that is met by the biomass under insufficient MCPA levels in the system [299], and it is modeled by subtracting the exogenous maintenance flux from the total maintenance demand.

$$r_{m\text{-endogenous}} = C_B \cdot m \cdot \left( 1 - \frac{C_P^L}{C_P^L + K_M} \right) \cdot f_R(T) \quad (90)$$

We introduced a non-extractable residues pool to account for the delayed release of  $\text{CO}_2$  coming from the decaying biomass:

$$r_{decay}^{NER} = C_B^{NER} \cdot a_{\text{CO}_2} \cdot f_R(T) \quad (91)$$

where:  $C_B^{NER}$  [ $\text{mmol cm}^{-3}$ ] is the non-extractable residues pool, and  $a_{\text{CO}_2}$  [ $\text{d}^{-1}$ ] is the decay rate coefficient of the  $C_B^{NER}$ .

To describe the  $^{14}\text{C}$  dynamics, we incorporated a  $^{14}\text{C}$  pool, which accounts for only the  $^{14}\text{C}$  portion of labeled MCPA. Processes described for the  $^{14}\text{C}$  pool are: growth, maintenance, and respiration. Additionally, the  $C_B^{NER}$  pool traces only the  $^{14}\text{C}$ - $C_B^{NER}$  formed. We calculated each  $^{14}\text{C}$  flux by multiplying the corresponding total flux by the current mass fraction ( $\alpha_{14}$ ) of the source pool [43]. The  $\alpha_{14}$  was in turn derived from the total activity of  $^{14}\text{C}$  per g of soil ( $A_S = 15$  kBq per 100 g of soils) and the mean specific activity of MCPA ( $\alpha_{\text{MCPA}} = 55$  mCi  $\text{mmol}^{-1}$ ) in relation to the initial MCPA ( $C_{\text{MCPA}}^0$ ) applied (either 1 or 20  $\text{mg kg}^{-1}$ ).

The full ODE equations used are:

$$\frac{dC_B}{dt} = r_{growth} - r_{m-endogenous} - r_{decay} \quad (92)$$

$$\frac{d^{14}C_B}{dt} = r_{growth}^{14} - r_{m-endogenous}^{14} - r_{decay}^{14} \quad (93)$$

$$\frac{dC_B^{NER}}{dt} = r_{decay}^{14} - r_{decay}^{NER} \quad (94)$$

$$\frac{dC_P^L}{dt} = -\frac{r_{uptake} + r_{m-exogenous}}{R_F} \quad (95)$$

$$\frac{d^{14}CO_2}{dt} = r_{respiration}^{14} + r_{m-endogenous}^{14} + r_{m-exogenous}^{14} + r_{decay}^{NER} \quad (96)$$

$R_F$  is the retardation factor (ref. to Chavez Rodriguez et al. [297]) introduced to account for nonlinear equilibrium sorption using the Freundlich isotherm:

$$RF := 1 + \frac{\rho}{\theta} \cdot K_P \cdot n_P \cdot (C_P^L)^{(n_P-1)} \quad (97)$$

where  $K_P$  [ $\text{mmol}^{(1-n_P)}\text{g}^{-1}\text{cm}^{(3n_P)}$ ] and  $n_P$  [-] are the Freundlich coefficient and exponent fixed to 0.09 and 0.8576 (adapted from Gawlik et al. [300]),  $\theta$  [ $\text{cm}^3 \text{cm}^{-3}$ ] is the soil water content, and  $\rho$  [ $\text{g cm}^{-3}$ ] is the soil bulk density.

The uptake rate  $r_{uptake}$  (ref. to Chavez Rodriguez et al. [297]) depends on the bioavailable fraction of pesticide as follows:

$$r_{uptake} = \mu_{max} \cdot C_B \cdot \left(\frac{\rho}{\theta}\right) \cdot \left(\frac{(C_P^L)^{n_H}}{(C_P^L)^{n_H} + (K_G)^{n_H}} + \frac{\alpha}{f_T}\right) \cdot \left(\frac{C_P^L}{C_P^L + K_M}\right) \cdot \left(\frac{1}{Y_P}\right) \cdot f_R(T) \quad (98)$$

### Model calibration

We performed a hierarchical model calibration using the parameter ranges from Table S5 and minimized the sum of squared error (SSE) with the optimization algorithm Simulated Annealing from MATLAB:

$$SSE = \sum_{i=1}^n \frac{(y_{measured}^i - y_{simulated}^i)^2}{\sigma_i^2} \quad (99)$$

where  $y_{measured}$  is the mean value of the  $i^{th}$  observation,  $y_{simulated}$  is the corresponding  $i^{th}$  simulated value, and  $\sigma^2$  is the standard deviation of the corresponding observations.

The hierarchy of parameter groups was formed by assuming: i) different bacterial subpopulations under the two different initial concentrations of MCPA applied (C), ii) possible physiological and morphological bacterial changes under different moisture levels (W), not well captured by literature moisture functions, and iii) biological and physicochemical properties of soil (S). Thus, parameters for calibration were grouped according to the proposed hierarchy (Table S19):

Model outputs corresponding to the measured data are:

$$\text{Mineralization [\%]} = \frac{{}^{14}\text{CO}_2 \cdot 100\%}{{}^{14}\text{C}_{MCPA}^0} \quad (100)$$

$$\text{Genes [copies g}^{-1}\text{]} = \frac{C_B}{f_1} \quad (101)$$

$$\text{Transcripts [copies g}^{-1}\text{]} = \frac{\widehat{mRNA} \cdot C_B}{f_1} \quad (102)$$

$$\text{Residual MCPA [mg kg}^{-1}\text{]} = C_P^L \cdot \frac{\theta}{\rho} + K_P \cdot (C_P^L)^{np} \quad (103)$$

$$\text{DT}_{50\text{RES}} [\text{d}^{-1}] = \text{Time}[(\text{Residual MCPA} = 0.5 \cdot C_{MCPA}^0)] \quad (104)$$

where  $f_1$  [mmol gene<sup>-1</sup>] is the conversion factor cell to carbon. Incorporation of <sup>14</sup>C into the microbial biomass ( $C_{mic}$ ) as well as CUE were not used for model calibration

### 8.3.8 Carbon use efficiency (CUE)

We derived  $CUE$  both experimentally and model-based  $CUE_S$ :

1.  $CUE_M$ : experiment-based  $CUE$  used for labeled substances [301].

$$CUE_M = \frac{{}^{14}\text{C}_{mic}}{{}^{14}\text{C}_{mic} + R_{cum}} \quad (105)$$

where:  $^{14}C_{mic}$  is the C uptake in microbial biomass, and  $R_{cum}$  is the cumulative respiration rate.

2.  $CUE_E$ : environmental model-based  $CUE$  adapted from Geyer et al., [302].

$$CUE_E = \frac{^{14}C_B}{^{14}C_B + ^{14}CO_2} \quad (106)$$

3.  $CUE_C$ : community model-based  $CUE$  adapted from Geyer et al. [302] and Manzoni et al. [303].

$$CUE_C = 1 - \frac{r_{respiration}^{14} + r_{m-endogenous}^{14} + r_{m-exogenous}^{14} + r_{decay}^{NER}}{r_{respiration}^{14} + r_{growth}^{14} + r_{m-exogenous}^{14}} \quad (107)$$

### 8.3.9 Statistical analyses

A linear model with mixed effects as part of the "nlme" package using the lme function [304] implemented in R version 3.5.2 was applied, specifying concentration, soil moisture content and interand temperature and their interactions as fixed effects, and microcosms as random effects. To investigate the influences of temperature, soil moisture content and concentration on the  $^{14}CO_2$  mineralization rate, we compared the attained  $^{14}CO_2$  level on day 28 among all treatments. To illustrate the differences in  $CUE$ , we compared all treatment levels. To test the assumption that a temperature reduction to 10°C leads to a significant increase in  $^{14}C$  uptake, we compared  $^{14}C$  uptake at increased and reduced temperatures, at each concentration, and each soil moisture level. The probability that the measurements for a given experimental unit would be temporally correlated also had to be considered. In this case, corAR1 (time) was used to indicate a temporal autocorrelation structure of order one. Since the ANOVA requires a normal distribution and variance homogeneity within the data, an assumption check for the mineralization, incorporation and  $CUE$  was run to ensure that the prerequisites for reliable calculations were met. For specification of contrasts between the influencing factors relevant for the verification of our hypothesis, a post-hoc comparison was conducted using the

package “emmeans” [305]. With this package, the estimated marginal means were calculated, and interaction plots were made by using the “emmip” function to display the interactions between the variables soil moisture, temperature, and concentration. The influences of the variables were compared pairwise with the Tukey method, and the standard error (SE) and p-value for each result was simultaneously computed.

## 8.4 Results and discussion

Pesticide degradation studies often neglect the possibility that the efficiency of microbial C utilization can shift in response to environmental factors [266, 306]. Our study, therefore, analyzed microbial utilization of the pesticide MCPA in response to changes in soil moisture and temperature. In addition to estimating the mineralization of  $^{14}\text{C}$ -labeled MCPA and the dynamics of specific degraders, we calculated the *CUE* of MCPA turnover to evaluate microbial C allocation to catabolic and anabolic processes.

### 8.4.1 Enhanced MCPA mineralization by elevated temperature and moisture

Mineralization of MCPA in soil was quantified as that of the accumulated  $^{14}\text{CO}_2$  at the end of the incubation (Figure 11 A, B, C, D). Under optimal soil conditions (20°C, pH 1.8) and 20  $\text{mg kg}^{-1}$ , nearly 70% of the initially applied  $^{14}\text{C}$ -labeled MCPA was mineralized to  $^{14}\text{CO}_2$ . Under limiting conditions (10°C, pH 3.5) and 1  $\text{mg kg}^{-1}$ , mineralization was significantly reduced and peaked at only 23%. These results were confirmed by our model simulations, which accurately depicted the measured mineralization (Figure 11 A, B, C, D).

As a single factor, a temperature increase from 10 to 20°C resulted in an increase in  $^{14}\text{CO}_2$  mineralization of 10.5% ( $F_{1,16} = 73.9$ ,  $p < 0.01$ ). However, the effect was significantly more pronounced at high MCPA concentrations (+17.7%) than at low MCPA concentrations (+3.4%,  $F_{1,16} = 35.2$ ,  $p < 0.001$ ; Figure 1 A, B, C, D). Comparable temperature-dependent increases in mineralization were demonstrated in studies by Nowak et al. [307] and Muskus et al. [269] for glyphosate, Helweg [263] for mecoprop (MCP), and Bouseba et al. [265] for 2,4-D. These

increases in  $^{14}\text{C}$  mineralization appeared to be independent of the chemical properties and associated behaviours of those pesticides in soils. In our experiment, an explanation could be found in the temperature sensitivity of the enzyme-catalyzed reactions of MCPA degradation, which are associated with inherent kinetic properties (intrinsic temperature sensitivity) and a concentration-dependent response of mineralization rates to temperature (apparent temperature sensitivity) [308].

Additionally, we evaluated the effect of water content as a sole factor, in which a reduction from pF 1.8 to 3.5 resulted in the strongest decrease in  $^{14}\text{CO}_2$  mineralization ( $-16.2\%$   $F_{1,16} = 136$ ,  $p < 0.01$ ). This effect was most pronounced at the high MCPA concentration ( $F_{1,16} = 17.9$ ,  $p < 0.001$ ), where total mineralization was 21.3% higher at pF 1.8 as compared to pF 3.5 (Figure 11 A,B). At the low MCPA concentrations, this increase was only 11.0% (Figure 11 C, D). Generally, microbial activity decreases with increasing osmotic potential, as demonstrated by Sparling et al. [309]. According to Ilstedt et al. [310], the reason why a reduction in water content also reduces the maximum mineralizable  $^{14}\text{C}$  fraction of MCPA is related to the limitation on substrate diffusion due to reduced thickness of the water film on the soil particles as the water content declines [311]. Schroll et al. [264] determined an optimal osmotic potential for aerobically degradable chemicals of  $-0.015$  MPa, which corresponds to a pF value of about 2.2.

Evaluation of MCPA residues (Figure 11 E-H) indicated almost complete degradation, especially in the  $1\text{ mg kg}^{-1}$  treatment at  $20^\circ\text{C}$ , where MCPA was no longer detectable after 10 to 15 days. Only under the treatment combination of  $20\text{ mg kg}^{-1}$ ,  $10^\circ\text{C}$  and pF 1.8, 10% of the initial applied MCPA remained in the soil. In the treatment combination pF 3.5 and  $20^\circ\text{C}$ , no MCPA could be extracted after 20 days. If the incubation took place at  $10^\circ\text{C}$ , MCPA was no longer detectable after 25 days. Similar detection times for MCPA were reported by Bælum et al. [312], Hiller et al. [313], and Peña et al. [314].

#### 8.4.2 Invariable microbial dynamics under limiting temperature and moisture

The *tfdA* gene abundance responded only to the 20 mg kg<sup>-1</sup> and 20°C treatment (Figure 11 I, J). The abundance of *tfdA* genes reached a maximum 10 days after MCPA application; with higher copy numbers in soils at pF 3.5 than in soils at pF 1.8 (concentration:day:temp:pF;  $F_{8,96} = 2.05$ ,  $p = 0.048$ ). After the peak, a slow decline followed until day 28, after which the initial level of 10<sup>4</sup> copies g<sup>-1</sup> was reached again. Similar results were obtained by Vieublé Gonod [315] and Baelum et al. [128], where an initial "lag" period (0 - 8d) with minor mineralization indicated limited microbial pesticide turnover. In a second phase characterized by a sharp increase in mineralization (after day eight), Baelum et al. [128] were able to detect a proliferation of degraders based on *tfdA* copy number to 3.0 · 10<sup>6</sup> per gram of soil, resulting in an approximate sigmoid shape of the mineralization curve after saturation of <sup>14</sup>CO<sub>2</sub> release was attained, consistent with our results. We could associate a maximum increase in *tfdA* copy number to 4.3 · 10<sup>6</sup> per gram soil with comparable mineralization kinetics. The response of *tfdA* transcripts to temperature mirrored the patterns of *tfdA* gene abundance (concentration:day:temp;  $F_{8,96} = 30.01$ ,  $p < 0.001$ ), but transcripts returned after day 15 to the initial level of 10<sup>3</sup> copies g<sup>-1</sup> dw (Figure 11 M,N,O,P). Soil moisture did not affect gene transcription.

The observed patterns of *tfdA* gene and transcript dynamics were well captured by simulations using gene-based mechanistic model. According to Gözdellier et al. [316], different degrader subpopulations are adapted to different MCPA concentrations. Therefore, we allowed the parameters  $f_1$  (conversion factor cell to carbon),  $m$  (maintenance coefficient), and  $Y_p$  (yield coefficient) to take different values at 1 and 20 mg kg<sup>-1</sup>. Calibrated parameters (Table S20) suggested populations with bigger cells, higher maintenance demands and lower yield efficiencies at 20 mg kg<sup>-1</sup> than at 1 mg kg<sup>-1</sup>, which is in accordance with Gözdellier et al. [316]. Additionally, within each concentration level, slightly smaller cells with low maintenance demands and high yield efficiencies might be expected at pF 3.5.

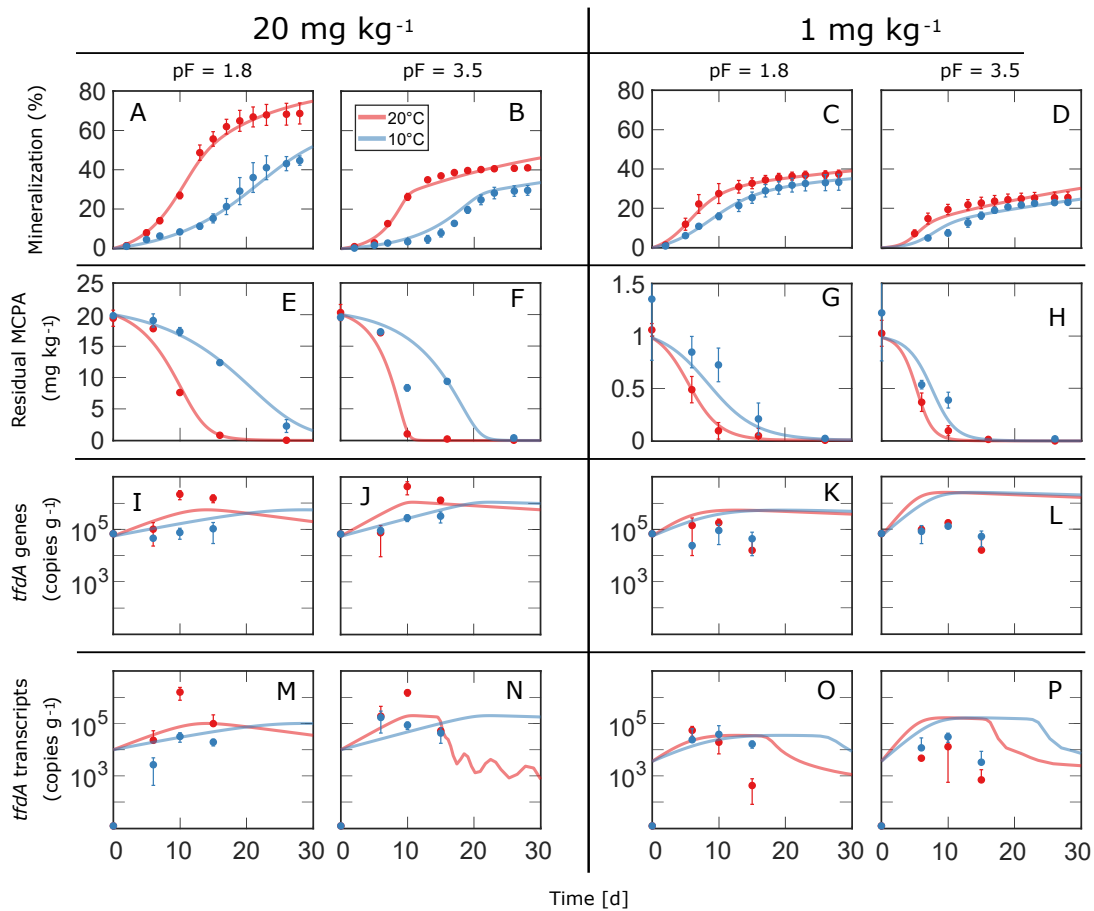


Figure 11: Measured (dots) and simulated (lines) of cumulative  $^{14}\text{CO}_2$  mineralization at two MCPA concentrations as a function of temperature and soil moisture over time (A, B, C, D), Residual MCPA expressed as  $\text{mg kg}^{-1}$  over time (E, F, G, H), *tfdA* genes during the MCPA biodegradation experiment expressed as gene copies  $\text{g}^{-1}$  dry weight (I, J, K, L), *tfdA* transcripts quantities during MCPA degradation expressed as transcripts copies  $\text{g}^{-1}$  dry weight (M, N, O, P). *tfdA* genes and transcripts are expressed at log-scale. Error bars represent standard errors of the mean values for soil triplicates (see M&M).

### 8.4.3 CUE dependency on temperature, moisture, and MCPA concentration

We determined  $CUE_M$  based on measured  $^{14}\text{C}$  incorporation into microbial biomass [301]. Additionally, and taking advantage of our mechanistic gene-centric model, we derived two model-based carbon use efficiencies -  $CUE_E$  and  $CUE_C$ . While  $CUE_M$  accounts for pesticide-



derived C incorporation into the whole microbial community, the two model-based carbon use efficiencies exclusively consider C utilization by specific pesticide degraders.  $CUE_M$  and  $CUE_E$  measure effects of pesticide-C stabilization on carbon utilization over a longer period, taking into account the effects of biomass turnover, substrate recycling, and potential cross-feeding [302].  $CUE_C$  is calculated from simulated process rates and measures the immediate carbon utilization after MCPA uptake.

The prerequisite for  $CUE_M$  assessment is quantification of  $^{14}\text{C}$  incorporation into the biomass ( $^{14}\text{C}_{mic}$ ). Soil moisture did not affect  $^{14}\text{C}_{mic}$  (Table S27). In contrast, a temperature reduction to  $10^\circ\text{C}$  significantly increased  $^{14}\text{C}_{mic}$  during the first five days after MCPA application by 3 percentage points to 10% ( $F_{1,16} = 4.9, p < 0.05$ ), compared to the  $20^\circ\text{C}$  treatment. Microbial uptake of MCPA can occur very quickly, according to Nowak et al. [179], who found a peak in 2,4-D derived  $^{13}\text{C}$  after only two days. They identified bacteria as the main degraders of 2,4-D in the soil. However, an initial high  $^{14}\text{C}_{mic}$  is followed by  $^{14}\text{C}$  losses, since the  $^{14}\text{C}$  is assimilated to form precursor compounds (PreC) for further biosynthesis or is dissimilated for maintenance respiration [302].

This short-term metabolic reaction of degraders is represented by  $CUE_C$  (Figure S40). On average,  $CUE_C$  increased by 0.2 at  $1 \text{ mg kg}^{-1}$  compared to  $20 \text{ mg kg}^{-1}$ .  $CUE_C$  in relation to the remaining MCPA concentration reached zero at  $1 \text{ mg kg}^{-1}$  after about 99% of the initially applied MCPA was degraded in contrast to the  $20 \text{ mg kg}^{-1}$  treatment, where this point was reached earlier (90%). These findings indicate a more efficient utilization of MCPA-derived C at low initial concentrations and a longer-lasting gross production. Gross production is defined as total pesticide uptake minus pesticide-C that is mineralized and used for further biosynthesis processes [302]. Therefore, in contradiction to the constant metabolic flux analysis of Geyer et al. [317], in which no change in the biochemical processes was detected during the incubation of different glucose concentrations, we can confirm a decrease of  $CUE_C$  for MCPA at  $20 \text{ mg kg}^{-1}$ . The lower  $CUE_C$  under the higher MCPA concentration is most likely explained by the two different fitted values of growth yield parameters for each initial MCPA concentration (Table

S20). This finding supports the inherent model assumption made in accordance with Gözdellier et al. [316] that two subpopulations of pesticide degraders with different physiologies trigger concentration-dependent shifts in pesticide-C metabolism. Metabolic regulations leading to increased nutrient-mobilizing extracellular enzymes or carbon-wasting respiratory mechanisms under nutrient limitations could also be responsible for lower  $CUE_C$  [318]. However, these processes can be ruled out because in fertilized soils, nutrient limitations are not expected.

Short term differences in  $CUE_C$  should affect the long-term fate of the MCPA-C, as measured by  $CUE_M$  and  $CUE_E$ .  $CUE_M$  was significantly higher during the MCPA degradation at 20  $\text{mg kg}^{-1}$  compared to 1  $\text{mg kg}^{-1}$  (+ 0.06;  $F_{1,16} = 5.8$ ,  $p < 0.05$ ) during the first 15 days (Figure 12 A, B, C, D).  $CUE_M$  can only be statistically evaluated by comparison at each time point. But, comparing  $CUE_M$  over time is misleading, as different states of degradation dynamics are being compared. To eliminate this deviation,  $CUE_E$  was considered as a function of the relative decrease in MCPA concentration (Figure 12 E, F, G, H). The simulated  $CUE_E$  is about 0.2 higher at low concentrations than at high concentrations, indicating greater carbon stabilization at the ecosystem level at low concentrations. For  $CUE_M$ , this effect was only evident at the end of the incubation ( $CUE = 0.21$ ;  $F_{1,16} = 4.3$ ,  $p = 0.05$ ).

We observed an increase in  $CUE_M$  with decreasing temperature (Figure 12 A, B, C, D), which also has previously been reported [319–321], and is associated with higher growth efficiencies [320, 322] and lower energy costs to maintain existing biomass [301, 316]. An additional temperature effect is that increased microbial activity at 20°C leads to increased MCPA turnover especially at 20  $\text{mg kg}^{-1}$ , which is in agreement with the Arrhenius equation [323]. The substrate concentration was therefore present longer at 10°C and as a result, maintained a higher  $CUE_M$  for a longer time (Figure 12 E, F, G, H). The simulated carbon use efficiencies did not indicate the temperature effect (Figure 12 E, F, G, H). This is because the model assumptions for  $CUE_E$  and  $CUE_C$  assigns the same temperature sensitivity to microbial growth, maintenance and turnover (see Eq. 98, 88, 89, 90, 91).

In addition to concentration and temperature,  $CUE_M$  was increased by the reduction of soil

water content (+ 0.15;  $F_{1,16} = 40.3$ ,  $p < 0.01$ ), especially at the first time point (fifth day). Similarly,  $CUE_E$  and  $CUE_C$  were 0.25 higher at pF 3.5 (Figure 12). Consistent with this finding, Jones et al. [324] found an upward trend in microbial CUE under the following levels: hyper-dry > dry > semi-dry, with the subsequent finding that even under hyper-dry conditions, very low microbial activity and C turnover occurred with altered C allocation. The reason given was a reduced catalytic activity related to a decline in motility of organisms and enzymes across a water film that loses thickness as drought increases.

Interestingly, increased  $CUE$  with reduced temperature and water content was not accompanied by any response of *tfdA* transcript and gene copy numbers in our study. This imbalance may be explained by the fact that microbial use of the substrate is more complex than simply converting it to biomass [271]. Rather, bacterial degraders synthesize a variety of products, e.g., to maintain basic functions, such as extracellular enzymes, extracellular polysaccharides, cell wall polymers, but also stress response compounds, such as osmolytes, to survive under dry conditions [282]. This formation of stress compounds could explain a slight increase in carbon use efficiencies during MCPA degradation under drier conditions compared to the near optimal water content at pF 1.8.

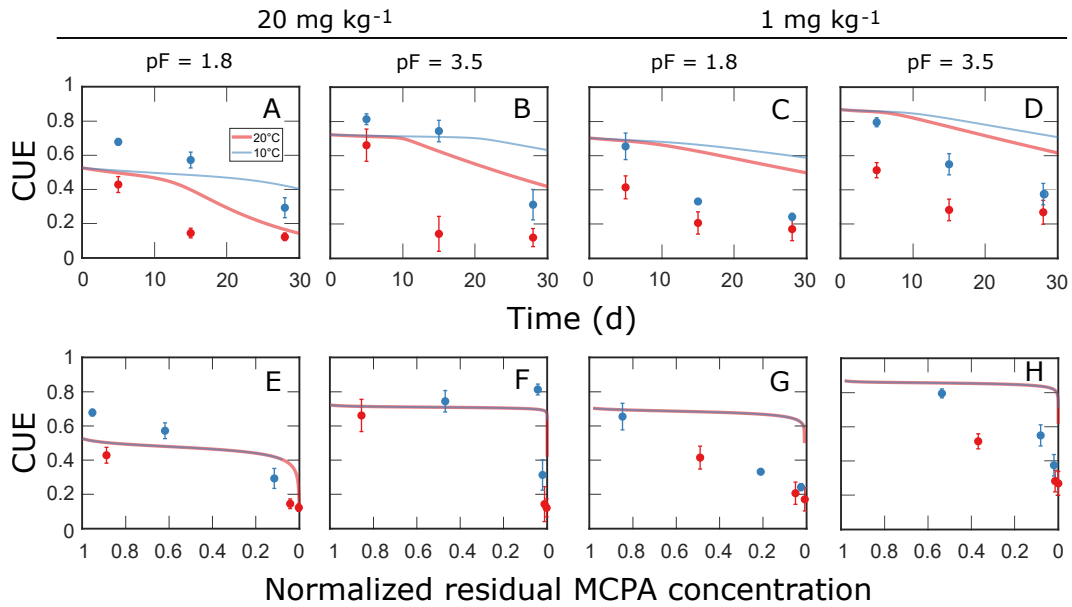


Figure 12:  $CUE$  vs. time (d) shown in panels A to D, and  $CUE$  vs normalized residual MCPA concentration in soil shown in panels E to H.  $CUE_M$  (Eq. 105) is presented as points and  $CUE_E$  (Eq. 106) as lines

#### 8.4.4 Effect of temperature and soil moisture on pesticide $DT_{50}$

Differences in MCPA mineralization were also reflected in  $DT_{50}$ -values, describing the time required to mineralize 50% of the applied MCPA (Table 9). We determined two different  $DT_{50}$ -values; i) a  $DT_{50MIN}$  derived from the mineralization kinetics and typically calculated in dissipation experiments of pesticides, and ii) a  $DT_{50RES}$  derived from the residual MCPA concentration. Under limiting conditions, we observed longer  $DT_{50RES}$  as well as  $DT_{50MIN}$ -values, with temperature exerting a stronger influence than soil moisture.

We observed that lowering the temperature to 10°C at 1 mg kg<sup>-1</sup> and pF 3.5 increased the residence time of MCPA by a factor of 1.9 based on mineralization kinetics (Table 9). In contrast, the  $DT_{50RES}$ -value differed only by a factor of 1.4. This may be explained by an altered temperature-dependent C allocation, namely a disproportionate increase in <sup>14</sup>C incorporation at 10°C versus an increase in mineralization at 20°C. Consequently, this resulted in almost equal MCPA-utilization rates [325]. It is, therefore, important to consider total MCPA turnover, as

$DT_{50RES}$  theoretically includes the dynamics of  $^{14}C$  incorporation and mineralization, whereas  $DT_{50MIN}$  captures only the contribution of mineralization. At pF 1.8, both  $DT_{50}$  values were identical (Table 9). However, at pF 1.8 the  $CUE_C$  and  $CUE_E$  was on average 0.2 lower than at pF 3.5 (Figure 12, Eqs. 107 and 106) reflecting a reduced contribution of  $^{14}C$  incorporation to MCPA-derived C turnover.

Compared to the concentration of  $1 \text{ mg kg}^{-1}$ , the effect of a temperature reduction at  $20 \text{ mg kg}^{-1}$  was independent of soil moisture and the  $DT_{50}$  approach, and increased half lives by a factor of 2 (Table 9). In this case, degradation is initially limited by the number of microorganisms, in contrast to degradation at  $1 \text{ mg kg}^{-1}$ , where the degradation potential is already provided by the autochthonous microbial abundance and rapid first-order degradation can be initiated immediately [171, 289]. According to Babey et al. [259], degradation of 2,4-D is most efficient when the ratio of degraders to instantaneous pesticide concentrations favors of degraders. This was the case for the treatment at  $20^\circ\text{C}$  and high initial pesticide concentration after relatively lower mineralization was observed in the first phase of the experiment (0-5d).

In the absence of growth at  $10^\circ\text{C}$ , as indicated by the lack of increase in *tfdA* copy number, the degradation efficiency was significantly reduced, as reflected in the increased  $DT_{50}$  values (Table 9).

Table 9: Half-life  $DT_{50RES}$  derived from the residual MCPA concentration in soils and  $DT_{50MIN}$  derived from mineralization kinetics as a function of soil water content, concentration, and temperature

	$DT_{50RES}$						$DT_{50MIN}$					
	pF 1.8			pF 3.5			pF 1.8			pF 3.5		
	$10^\circ\text{C}$	$20^\circ\text{C}$	$10^\circ\text{C}/20^\circ\text{C}$	$10^\circ\text{C}$	$20^\circ\text{C}$	$10^\circ\text{C}/20^\circ\text{C}$	$10^\circ\text{C}$	$20^\circ\text{C}$	$10^\circ\text{C}/20^\circ\text{C}$	$10^\circ\text{C}$	$20^\circ\text{C}$	$10^\circ\text{C}/20^\circ\text{C}$
$20 \text{ mg kg}^{-1}$	18.9	9.3	<b>2.0</b>	15.6	7.6	<b>2.1</b>	18.5	9.9	<b>1.9</b>	17.6	8.8	<b>2.0</b>
$1 \text{ mg kg}^{-1}$	9.2	6	<b>1.5</b>	7.5	5.2	<b>1.4</b>	10.3	6.7	<b>1.5</b>	12.2	6.3	<b>1.9</b>

Our results partially refuted previous findings [265, 326–328], stating that the decrease in temperature and soil moisture during biodegradation of MCPA is always accompanied by a significant increase in half-life. The extent to which the residence time of MCPA was affected

by a change in temperature and soil moisture content depended on initial concentration and associated degradation dynamics. In the present study, we demonstrated that  $^{14}\text{C}$  incorporation is not necessarily proportional to mineralization, confirming our hypothesis that under limiting conditions assimilation can be enhanced to support biosynthesis rates. Dissimilation including non-growth maintenance activities [302] increased with temperature, as energy costs became more important to regulate motility or molecular turnover of proteins [329]. As a result, the MCPA-derived carbon will be used more efficiently by microorganisms at low temperatures and reduced soil moisture content. Applying this principle to pesticide degradation, estimating  $\text{DT}_{50}$ -values from cumulative mineralization curves alone could, under certain circumstances, imply a systematic overestimation of persistence time.

## **9 Optimal design of experiments for effective modeling of atrazine degradation in soils (Paper 4)**

This chapter includes the following paper in preparation for publication as:

Luciana Chavez Rodriguez, Ana Gonzalez-Nicolas, Brian Ingalls, Sinan Xiao, Wolfgang Nowak, Thilo Streck, Holger Pagel (2021). Optimal design of experiments for effective modeling of atrazine degradation in soils.

with the following modification:

1. Numbers of figures, tables, and equations are relative to this thesis and not to the original publication.

## 9.1 Abstract

The natural degradation pathways of the herbicide atrazine (AT) are highly complex. These pathways involve the metabolic activity of several bacterial guilds (that use AT as a source of carbon, nitrogen, or both) and abiotic degradation mechanisms. The co-occurrence of multiple degradation pathways, combined with challenges in quantifying bacterial guilds and relevant intermediate metabolites, could be represented by competing model formulations, which all might represent valid descriptions of the fate of AT. A proper understanding of the fate of this complex compound is needed to develop effective management and mitigation strategies. Here, we propose a model discrimination process in combination with a prospective optimal design of experiments. We simulated experimental data using a first-order model that reflects a simple reaction chain of complete AT degradation and a set of Monod-based model variants that consider different bacterial guilds. We used a Bayesian statistical analysis of simulated ensembles to investigate virtual degradation experiments and chemical analysis strategies, thus obtaining predictions on the utility of experiments to deliver conclusive data for model and pathway discrimination. We considered a range of experimental protocols addressing: i) the metabolites or chemicals to measure (AT, metabolites, and  $CO_2$ ), ii) sampling frequency (daily, every two days, or every four days), and iii) features typically not measured (specific bacterial guilds). As a statistical metric for model discrimination we used the energy distance. Our results show that simulated AT degradation pathways following first-order reaction chains can be clearly distinguished from simulations using Monod-based models. Within the Monod-based models, we detected two clusters of models that differ in the number of bacterial guilds involved in AT degradation. Based on our prospective analysis, experimental designs considering the sink cyanuric acid (CA) provided the most informative data to discriminate models. As expected, the inclusion of measurements of specific bacterial guilds improved model discrimination. Our study highlights that environmental fate studies should prioritize measuring metabolites to elucidate active AT degradation pathways in soil and identify robust model formulations supporting risk assessment and mitigation strategies.



**Keywords:** atrazine degradation, monod-kinetics, first-order kinetics, equifinality, model discrimination, optimal design of experiments, energy distance.

## 9.2 Introduction

Pesticides are important chemicals used globally in agriculture to manage plant stressors such as pests, weeds, and diseases [5]. Due to their potential negative effects on ecosystems [226] and human health [21], some pesticides have been banned or otherwise restricted. The pesticide atrazine (AT) was banned in Europe in 2004. However, AT and its metabolites are still found in soils and groundwater in potentially harmful concentrations [57, 203, 330]. In the environment, AT undergoes abiotic [132] and biotic [49] degradation. Several bacterial guilds have been observed to metabolize AT (as carbon source [62, 63], nitrogen source [10] or both [49]), leading to an accumulation of intermediate metabolites, most commonly: hydroxyatrazine (HY) [141, 142, 330], deisopropylatrazine (DIA), and deethylatrazine (DEA) [331–333]. The co-occurrence of multiple AT degradation pathways that can lead to the formation of identical metabolites poses a challenge to determining the fate of AT. This issue confounds our ability to understand why AT persists in real systems, thus hampering future mitigation strategies [98].

Mathematical modeling approaches are valuable tools for the investigation of complex degradation pathways such as AT degradation, allowing for combining the current understanding of AT degradation with mathematical formulations and validating them with real measurement data such as AT and metabolite concentrations and biomass [334]. In the particular case of AT degradation, the limitation of which intermediate metabolites and bacterial guilds involved are measured could lead to distinct mathematical models representing the same system with equivalent accuracy (equifinality problem) [115]. Distinguishing among these competing models can help us to determine which AT degradation pathways are active in a particular environment. When addressing competing models, two cases arise. If all model formulations predict similar behavior for all system elements (AT degradation, metabolite

dissipation, and biomass formation), then the simplest (most parsimonious) model formulation is usually accepted as the best (most valid) representation. Otherwise, it is important to know what observations might provide the most useful information to distinguish the models: to facilitate *model discrimination* [115, 335, 336]. By identifying relative differences between models, we can reduce the number of competing models by clustering together those that are most similar and facilitate model invalidation [115, 336].

Optimal Design (OD) of experiments is a promising tool for addressing the equifinality problem. OD aims to maximize the benefit obtained from experiments [112]. If it is done prior to the execution of the experiment, it is called prospective OD [110, 112]. In our case, we use OD to identify experimental designs that maximize the observed difference between competing models of AT degradation [115]. Among multiple metrics used to distinguish models [115, 335, 336], the concept of energy distance (ED) [114] is a computationally efficient and robust model-distance metric. In this context the design that produces data from which one can maximize total pair-wise (model to model) ED is considered the optimal design for model discrimination.

This work aims to determine the measurements needed to distinguish the active AT degradation pathway in a particular environment, represented by mathematical models. At the same time, we aimed to find the level of model complexity needed to describe AT degradation in soils through model invalidation. For this purpose, we developed a set of mechanistic Monod-based models, representing different degradation pathways of AT in soils and a first-order decay model, typically used to describe degradation at field scale [337]. Later, we applied a prospective OD, using ED for model discrimination, to the set of models, and explored the advantages of including not yet quantified pools in model discrimination.

### 9.3 Material and Methods

#### 9.3.1 Atrazine degradation models

##### Conceptual model

We consider a set of hierarchical, nested models for degradation of atrazine (AT) in soils (Figure 13), including biotic and abiotic degradation, representing common degradation pathways of AT. These models vary in complexity from a complete Monod-model version (M1) to a simple first-order decay model (M6) (which is commonly used to model degradation at field scale) [337].

We assume that degradation processes occur in a well-mixed soil environment that contains a collection of bacterial guilds: labelled A, B, C, and D (see Section Bacterial guilds description). Members of each guild are able to fully or partially metabolize bioavailable AT and its intermediate metabolites as sole carbon and energy sources [139, 338, 339] (Figure 13). Nitrogen use is not considered. The members of each guild are partitioned into two subpools with different physiological states: active and dormant. Activation and deactivation rates are driven by carbon availability in the system. We explicitly account for a dissolved organic carbon pool (DOC) that serves as a collector of dead cells. The last metabolite of the AT transformation is cyanuric acid (CA) [340]. The transformation of CA to carbon dioxide (CO<sub>2</sub>) is regulated by nitrogen availability. At high nitrogen concentrations, CA transformation is strongly inhibited by all guilds [137].

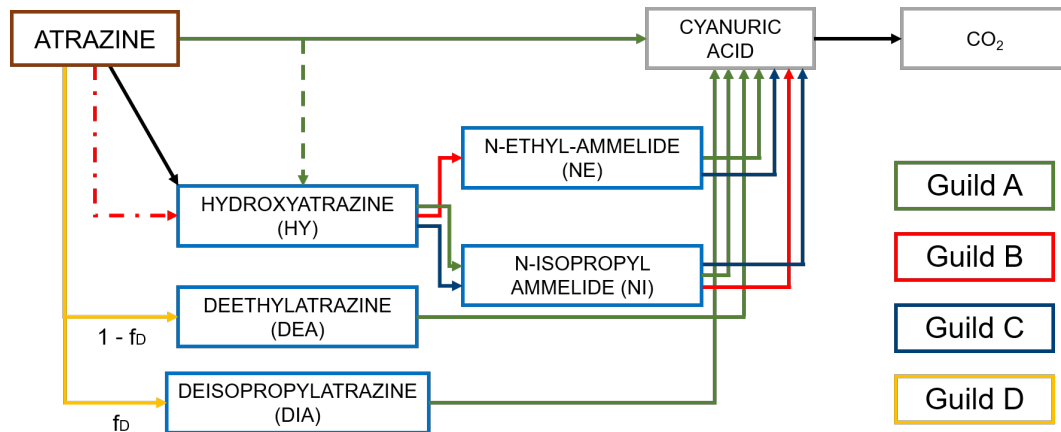


Figure 13: Atrazine (AT) degradation in soils: model framework, describing AT, its intermediate metabolites (in blue), sink pools (in grey), and the bacterial guilds involved in the degradation process. Arrow colors indicate activity of the bacterial guilds. Black arrows represent abiotic processes: degradation of AT and nitrogen-dependent degradation of CA. Dashed line represents the unintended HY leaked out of the degradation carried out by guild A [62, 63]. Dashed-dot line shows a degradation step uncoupled from growth (carried out by guild B);  $f_D$  represented the proportion of DIA formed during AT degradation carried out by guild D

### Bacterial guilds

We defined the four guilds based on genetic information regarding known AT degraders:

1. Guild A is able to use the side chains of AT as carbon source, degrading it to cyanuric acid (CA) [137, 139]. Additionally, this guild can use as carbon sources, the metabolites HY, NE, NI and the products of the dealkylation of AT (DIA and DEA) [146]. Members of this guild constitutively express a range of gene combinations: *atzABC*, *trzN-atzBC* and/or *trzN-atzC* [49, 340]. Examples of members of this guild are: *Arthrobacter aureus* TC1 [62, 63], and *Ensifer* sp [341].
2. Guild B is able to dechlorinate AT to hydroxyatrazine (HY) without gaining either carbon or energy through the activity of genes *atzA* [142] or *trzN* [141]. Additionally, they degrade HY to N-ethylammelide (NE) [141] (via uncharacterize enzymes), or degrade the metabolite N-isopropylammelide (NI) to CA, via the gene *atzC*. Example of members this

guild: *Nocardia* sp [141].

3. Guild C uses HY and NI as main carbon and energy sources by harboring the genes *atzB* and *atzC*, yielding CA [141, 142]. The *atzC* gene also allows for metabolism of NE as carbon source [342]. Examples members of this guild: *Rhizobium* sp [141].
4. Guild D dealkylates AT to the metabolites deethylatrazine (DEA) and deisopropylatrazine (DIA) in a fixed proportion  $f_D$  [343] (Figure 13). Specific genes for this pathway have not been identified; it is believed that this degradation is a cometabolic process [152, 344] mediated by the cytochrome P450 [332]. Examples members of this guild: *Rhodococcus* sp [332].

### Process formulations

1. *AT and metabolite dynamics N*): AT and the intermediate metabolites HY, DEA, DIA, NI and NE (generically referred to by the label  $N$  (for nutrient) below, are each represented by a total concentration  $N^T$  [ $\text{mg cm}^{-3}$ ] segregated into a bioavailable pool (concentration  $N^L$  [ $\text{mg cm}^{-3}$ ]), and a sorbed pool (concentration  $N^S$  [ $\text{mg g}^{-1}$ ]):

$$N^T = \theta \cdot N^L + \rho \cdot N^S \quad (108)$$

where  $\theta$  [ $\text{cm}^3 \text{cm}^{-3}$ ] and  $\rho$  [ $\text{mg cm}^{-3}$ ] are the soil water content and bulk density respectively.

These two pools are related by the Freundlich isotherm with retardation factor ( $RF$ ):

$$RF = 1 + \frac{\rho}{\theta} \cdot K_{FP}^N \cdot n_{FP}^N \cdot (N^L)^{(n_{FP}^N - 1)} \quad (109)$$

where  $K_{FP}^N$  [ $\text{mg}^{(1-n_N)} \text{g}^{-1} \text{L}^{n_N}$ ] is the Freundlich coefficient and  $n_{FP}^N$  [-] is the Freundlich exponent.

Bioavailable carbon sources ( $N^L$ ) are taken up and degraded biotically by active guild

populations  $B_k^a$ . We account for two possible fates for consumed nutrients (metabolite formation and bacterial metabolism). A fraction  $1 - f_N$  of the carbon in nutrient  $N$  is converted converted to the downstream metabolite:

$$r_{metabolite-formation}^{k,N} = \frac{\mu_{k,N} \cdot B_k^a \cdot (1 - f_N) \cdot \frac{\rho}{\theta}}{RF} \quad (110)$$

where  $\mu_{k,N}$  is the growth coefficient (eq. 118).

The remaining fraction  $f_N$  contributes to biomass accumulation and to respiration, with yield factor  $Y_{k,N}$  [-]:

$$r_{nutrient-use}^{k,N} = \frac{\mu_{k,N} \cdot B_k^a \cdot \left( \frac{f_N}{Y_{k,N}} \right) \cdot \frac{\rho}{\theta}}{RF} \quad (111)$$

Together, these give an overall uptake rate:

$$r_{uptake}^{k,N} = \frac{\mu_{k,N} \cdot B_k^a \cdot \left( \frac{f_N}{Y_k} + (1 - f_N) \right) \cdot \frac{\rho}{\theta}}{RF} \quad (112)$$

Specific degradation processes that do not involve biomass accumulation and respiration are described as follows:

- abiotic transformation of AT to HY (photodegradation) has been observed [230].

We model this process (black arrow in Figure 13 by first-order decay):

$$r_{abiotic-degradation}^{AT} = \frac{K_o \cdot AT^L}{RF} \quad (113)$$

where  $K_o$  [ $d^{-1}$ ] is the constant degradation rate.

- It has been observed HY leaks out of members of guild A by passive diffusion [62, 63] to the soil system. We modelled this leak flux as a constant fraction of the AT

that has been uptaken by guild AT:

$$r_{leak}^{AT-HY} = \frac{r_{metabolite-formation}^{A,AT} \cdot f_H}{RF} \quad (114)$$

where:  $f_H$  [-] fraction of the of the uptake AT flux that leaks out.

- Guild B dechlorinates AT to HY without gaining carbon or energy [63]. We modelled this process with Michaelis-Menten kinetics because this step is not coupled to growth (Dashed-point line in Figure 13):

$$r_{dechlorination}^B = \frac{k_{AT-HY} \cdot AT^L \cdot B_B^a}{K_{AT-HY} + AT^L} \cdot RF \quad (115)$$

where  $k_{AT-HY}$  [ $d^{-1}$ ] is the dechlorination rate for Guild B, and  $K_{AT-HY}$  [ $mg\ cm^{-3}$ ] is the half-saturation concentration.

- Guild D metabolizes AT to DIA and to DEA simultaneously [332]. A fraction  $f_D$  of the converted AT is in the form of DIA, while the remaining fraction  $(1 - f_D)$  is in the form of DEA (Figure 13).
2. *CA degradation*: CA is the final metabolite of AT transformation considered in the model because the further breakdown of CA is typically fast, without accumulation of intermediate metabolites [340, 345]. The model reflects CA degradation as inhibitory first-order decay:

$$r_{CA-CO_2} = \frac{CA^L \cdot d_{CA-CO_2} \cdot \frac{K_{in}}{NO_3 + K_{in}}}{RF} \quad (116)$$

where the  $d_{CA-CO_2}$  [ $d^{-1}$ ] is the rate of degradation of CA to  $CO_2$ ,  $K_{in}$  [ $mg\ cm^{-3}$ ] is the inhibition factor, and  $NO_3$  [ $mg\ cm^{-3}$ ] is the nitrogen concentration in the system taken as a model parameter.

3. *Bacterial dynamics and physiology*: We describe two subpopulations of each  $k$  bacterial

guilds ( $k = A, B, C, D$ ) according to their physiological state: active  $B_k^a$  or dormant  $B_k^d$  [ $\text{mg g}^{-1}$ ]. For each guild  $k$ , the active population grows at rate  $r_{growth}^k$  [ $\text{mg g}^{-1} \text{d}^{-1}$ ] on multiple carbon sources modelled with a Monod-kinetics allowing for competition for binding sites [78]:

$$r_{growth}^{k,N} = \left( \sum_{N \in \mathbf{N}_k} \mu_{k,N} \cdot f_N \right) \cdot B_k^a \quad (117)$$

where  $\mathbf{N}_A = \{\text{AT, HY, DEA, DIA, NE, NI}\}$ ,  $\mathbf{N}_B = \{\text{AT, HY, NI}\}$ ,  $\mathbf{N}_C = \{\text{HY, NE, NI}\}$ ,  $\mathbf{N}_D = \{\text{AT}\}$  and  $\mu_{k,N}$  is the growth coefficient defined as:

$$\mu_{k,N} = \frac{\mu_{max}^{k,N} \cdot N^L}{\mu_{max}^{k,N} + \mu_{\mathbf{N}_k}} \quad (118)$$

where:  $\mu_{max}^{k,N}$  [ $\text{d}^{-1}$ ] is the maximum growth rate for the guild  $k$  on the available fraction of the carbon source  $N^L$ . Function  $\mu_{\mathbf{N}_k}$  is defined as:

$$\mu_{\mathbf{N}_k} = \sum_{N=1}^i (N^L + K_{k,N}) \quad (119)$$

where  $i$  is the number of carbon sources that each guild  $k$  can utilize, and  $K_{k,N}$  [ $\text{mg cm}^{-3}$ ] is the half-saturation concentration of each guild  $k$  on each carbon source  $N^L$ .

Dormant populations do not grow. Transitions between dormant and active states are described by a switch-like function proposed by Stolpovsky et al. [169]:

$$r_{activation}^k = \tau_k \cdot k_k^a \cdot B_k^d \quad (120)$$

$$r_{deactivation}^k = (1 - \tau_k) \cdot k_k^d \cdot B_k^a \quad (121)$$

where  $k_k^a$  and  $k_k^d$  [ $\text{d}^{-1}$ ] are the activation and deactivation coefficients for guild  $k$ . Function  $\tau_k$  is defined as:



$$\tau_k = \left[ \exp \left( \frac{N_{th}^k - \sum_{N=1}^i N^L}{n \cdot N_{th}^k} \right) + 1 \right]^{-1} \quad (122)$$

where  $N_{th}^k$  [mg cm<sup>-3</sup>] is the threshold concentration for the guild  $k$ , and  $n$  [-] is the steepness parameter set to 0.1 [170].

Both active and dormant subpopulations are subject to linear decay at rate:

$$r_{decay}^{j,k} = a_{j,k} \cdot B_k^j \quad (123)$$

where  $a_{j,k}$  [d<sup>-1</sup>] is the decay rate coefficient for the guild  $n$  and index  $j$  represents active or dormant bacterial state.

4. *DOC formation and bacterial respiration*: We included two different sink pools:

a) Dissolved organic carbon pool (DOC) which collected dead cells from all guilds

( $r_{decay}^{j,k}$ ). A fraction  $f_R$  of the DOC contributes to the CO<sub>2</sub> pool:

$$r_{DOC-CO_2} = r_{decay}^{j,k} \cdot f_R \quad (124)$$

b) CO<sub>2</sub> [mg g<sup>-1</sup>] accumulates due to bacterial respiration at rate:

$$r_{respiration}^k = \mu_{k,P} \cdot B_k^a \cdot f_N \cdot \left( \frac{1 - Y_k}{Y_k} \right) \quad (125)$$

### Scenario models

AT is commonly found in soils together with three principal intermediate metabolites HY, DIA, and DEA [346–348]. Additional intermediate metabolites NE and NI are also part of some reported degradation pathways [141, 142], but their accumulation has rarely been reported in soils [49]. Therefore, we set six model scenarios based on the presence or absence of the

main bacterial guilds involved in AT degradation so that the metabolites HY, DIA, and DEA, are always present (Figure 13 and Table 10). Additionally, we added a simple first-order decay model M6, which only includes the chemical pools AT, HY, DIA, DEA, and CA and CO<sub>2</sub>. Specific degradation pathways mediated by fungi [349] were not considered. Full ODE equations for each scenario model are presented in the Supplementary information, Section 11.4.1.

Table 10: Scenario models of AT degradation in soils

Model variants	Bacterial guilds	Resulting chemical pools ( <i>N</i> )
M1	A, B, C, D	AT, HY, DIA, DEA, NI, NE
M2	B, C, D	AT, HY, DIA, DEA, NI, NE
M3	A, C, D	AT, HY, DIA, DEA, NI
M4	A, B, D	AT, HY, DIA, DEA, NE
M5	A, D	AT, HY, DIA, DEA
M6	-	AT, HY, DIA, DEA

### 9.3.2 Prospective optimal design of experiments (OD)

#### Model outputs and generation of simulated data

As candidate model outputs, we considered AT and the metabolites HY, DIA, DEA, NI, NE, and sinks CA, and CO<sub>2</sub>, as well as the biomass of the bacterial guild D, (the only guild present in all Monod-based model variants), and the total biomass (of all guilds present in the given system formulation).

Our prospective optimal design analysis is based on simulated data. We chose an initial concentration of AT of 100 mg kg<sup>-1</sup> [350] and an initial total biomass of 0.001 mg kg<sup>-1</sup> for model simulations, equally divided among the guilds present in the system formulation. We set all bacterial guilds to be dormant and all intermediate metabolites to zero at the beginning of the experiment. To restrict to plausible simulations, we defined a set of parameter and process constraints that the model parameters and outputs should satisfy based on expert knowledge and soil observations, as follows.

The parameter constraints ( $Pc$ ) are:

1. For each guild, the maximum growth rate coefficient ( $\mu_{max}^{k,C}$ ) must be higher than death rate coefficient of active bacteria ( $a_{a,k}$ )
2. The Freundlich sorption coefficient of HY ( $K_{FP}^{HY}$ ) must be higher than the Freundlich sorption coefficient of AT, DIA, DEA, CA, NI, and NE [351]
3. The Freundlich sorption coefficient of AT ( $K_{FP}^{AT}$ ) must be higher than Freundlich sorption coefficient of AT, DEA, CA, NE [351]

The process constraints ( $PC$ ) are:

1.  $DT_{50}$  (time that takes to dissipate 50% of the pesticide [65]) of AT between 5 and 25 days [49, 352]
2. AT concentration must be at least  $10^{-8}$  mg mL<sup>-1</sup> [57] at the end of the experiment
3. Mineralization of initially added AT between 20-80% at the end of the experiment
4.  $DT_{50}$  of HY, DIA and DEA between 2-30 days.

The high-dimensional parameter space of the scenario models of AT biodegradation (between 20 and 70 parameters depending on the scenario) implies that the behavioral parameter space satisfying all constraints (viable space [353]) is very small, making simple Monte-Carlo parameter sampling computationally much too expensive. We adopted a constraint-based search (CBS) method [118] and modified it to randomly select parameter sets from the behavioral parameter space. The CBS method is based on an iterative algorithm that successively applies stricter process constraints by increasing the minimum number of process constraints to be satisfied in each iteration. We replaced the original parameter sampling procedure of Gharari et al. [118] with a Metropolis-Hastings algorithm, using a Markov Chain Monte Carlo (MCMC) sampler. As a result, with this new CBS-MCMC method, we achieved reproducible and unbiased sampling of behavioral parameter sets.

The CBS-MCMC method used in our work consists of the following steps:

1. Define a number of parameter constraints ( $Pc$ ) (here,  $Pc = 3$ ).
2. Define the minimum number of process constraints ( $PC$ ) for the initial sampling (here,  $PC = 2$ ).
3. Perform an initial Latin hypercube sampling to draw  $M$  random parameter sets ( $M=500,000$  in this study) using uniform marginal parameter distributions (see Table 11 for ranges).
4. Identify the candidate parameter sets ( $\mathbf{x}_c$ ) that satisfy all given parameter constraints ( $Pc$ ).
5. Run the scenario model with the candidate parameter sets  $\mathbf{x}_c$  and evaluate the number of satisfied process constraints ( $PC'$ ).
6. Accept only the behavioral parameter sets ( $\mathbf{x}'_c$ ) resulting in model runs where  $PC' \geq PC$  and reject all other parameter sets in  $\mathbf{x}_c$ .
7. Increase  $PC$  by one, and use  $M_{MC}$  randomly chosen parameter sets from  $\mathbf{x}'_c$  as starting parameter values ( $\mathbf{x}_s$ ) for MCMC sampling.
8. Apply the Metropolis-Hastings algorithm with  $M_{MC}$  parallel Markov chains ( $M_{MC} = 40$  in this study):
  - a) Generate new candidate parameter sets  $\mathbf{x}_c$  using a Gaussian jumping distribution centered at the parameter values in  $\mathbf{x}_s$  with the standard deviation determined from all corresponding parameter values of  $\mathbf{x}_s$ . Verify if the parameter constraints ( $Pc$ ) are satisfied. We designed the algorithm to repeatedly draw individual parameter sets until all parameter constraints are satisfied. When the generated parameter values are located outside the defined lower and upper bounds (Table 11), they are reflected back into the search space at the respective boundary [117].
  - b) Run the scenario model with  $\mathbf{x}_c$  and evaluate  $PC'$ .

- c) Accept  $\mathbf{x}_c$  as behavioral if  $PC' \geq PC$ , otherwise reject  $\mathbf{x}_c$  and keep  $x_s$ . Update  $\mathbf{x}'_c$  correspondingly.

9. Repeat steps 6 and 7 until all process constraints are satisfied  $PC' = PC$ .

As long as  $PC' \geq PC$ , the length of individual Markov chains ( $L_M$ ) in Step 7 is set to 1,000 draws of candidate parameter sets. If less than  $M_{MC}$  new candidates were accepted, MCMC sampling in Step 7 is repeated, keeping the current value of  $PC$  while successively increasing  $L_M$  by 5,000 until at least  $M_{MC}$  new candidate parameter sets were accepted. As a pre-step to the last iteration of the CBS-MCMC method ( $PC' = PC$ ), repeated MCMC runs were performed to optimally adapt the jumping rate of the Metropolis-Hastings algorithm (Step 7) to achieve an acceptance rate of approximately 0.2 to 0.5 for the final MCMC runs. This way, we generated 30,000 unique model outputs ( $n_{mc}$ ) for further analysis.

Table 11: Model Parameters for Monod and first-order kinetic models

Parameter	Description	Units	Parameter Value	
			Min.	Max.
<b>Monod Parameters</b>				
$\mu_{max}^{n,P}$	Maximal specific growth rate of pesticide degraders	[d <sup>-1</sup> ]	10 <sup>-3</sup>	10 <sup>3</sup>
$K_{n,P}$	Growth substrate affinity coefficient of pesticide degraders	[mg cm <sup>-3</sup> ]	10 <sup>-3</sup>	10 <sup>5</sup>
$k_n^a$	Coefficient rate of activation	[d <sup>-1</sup> ]	10 <sup>-5</sup>	100
$k_n^d$	Coefficient rate of deactivation	[d <sup>-1</sup> ]	10 <sup>-5</sup>	100
$C_T^d$	Threshold concentration	[mg cm <sup>-3</sup> ]	10 <sup>-6</sup>	10 <sup>4</sup>
$a_{a,n}$	Specific death rate of active bacteria	[d <sup>-1</sup> ]	10 <sup>-3</sup>	10 <sup>4</sup>
$a_{i,n}$	Specific death rate of inactive bacteria	[d <sup>-1</sup> ]	10 <sup>-6</sup>	10 <sup>-2</sup>
$Y_n$	Yield parameter	[-]	0.1	1
$k_{AT-HY}$	Dechlorination rate	[d <sup>-1</sup> ]	10 <sup>-4</sup>	10 <sup>3</sup>
$K_{AT-HY}$	Saturation concentration	[mg cm <sup>-3</sup> ]	10 <sup>-5</sup>	10 <sup>4</sup>
<b>Sorption Parameters</b>				
$K_{FP}^P$	Freundlich coefficient	[mg <sup>(1-n<sub>FP</sub><sup>P</sup>)</sup> g <sup>-1</sup> cm <sup>(3n<sub>FP</sub><sup>P</sup>)</sup> ]	0.5	10
$n_{FP}^P$	Freundlich exponent	[-]	0.6	1
<b>First Order Decay Parameters</b>				
$K_o$	Abiotic transformation of Atrazine to HY	[d <sup>-1</sup> ]	10 <sup>-4</sup>	10 <sup>5</sup>
$K_I$	Inhibition factor	[mg cm <sup>-3</sup> ]	10 <sup>-4</sup>	10 <sup>3</sup>
$NO_3$	Nitrogen concentration	[mg cm <sup>-3</sup> ]	10 <sup>-3</sup>	10 <sup>3</sup>
$d_{AT-HY}$	Decay rate of AT to HY	[d <sup>-1</sup> ]	10 <sup>-4</sup>	10 <sup>3</sup>
$d_{AT-DD}$	Decay rate of AT to DD	[d <sup>-1</sup> ]	10 <sup>-4</sup>	10 <sup>3</sup>
$d_{HY-CA}$	Decay rate of HY to CA	[d <sup>-1</sup> ]	10 <sup>-4</sup>	10 <sup>3</sup>
$d_{DIA-CA}$	Decay rate of DIA to CA	[d <sup>-1</sup> ]	10 <sup>-4</sup>	10 <sup>3</sup>
$d_{DEA-CA}$	Decay rate of DEA to CA	[d <sup>-1</sup> ]	10 <sup>-4</sup>	10 <sup>3</sup>
$d_{CA-CO_2}$	Decay rate of CA to CO <sub>2</sub>	[d <sup>-1</sup> ]	10 <sup>-4</sup>	10 <sup>3</sup>
<b>Constant Rate Parameters</b>				
$f_R$	Fraction of dead bacteria which goes to DOC	[-]	0.01	1
$f_H$	Leak flux constant	[-]	0.01	1
$f_D$	Fraction of AT used for DEA formation used by guild D	[-]	0.25	0.75

Next, for each simulation, we generated a relative error  $E_{rel}(t)$  to normalize the output

channels: AT, metabolites (HY, DIA, DEA, NE, NI), and sinks (CA, CO<sub>2</sub>) for all the models at each time,  $y(t)_{m,l}$ . For this purpose, we applied a relative error tied to the mean observed at each time for the whole ensemble of each metabolite and models (30,000 sampled outputs). This way, observations with larger magnitudes will have larger  $E_{rel}(t)$ .

$$E_{rel}(t) = frac \cdot \sqrt{\left(\frac{1}{n_{mod}} \cdot \sum_{m=1}^{n_{mod}} \frac{1}{n_{mc}} \cdot \sum_{l=1}^{n_{mc}} y(t)_{m,l}\right)^2} \quad (126)$$

$$y_{normalized}(t)_{m,l} = \frac{y(t)_{m,l}}{E_{rel}(t)} \quad (127)$$

where *frac* is 10%,  $n_{mod}$  is the number of models ( $n_{mod} = 6$  for analysis including chemicals,  $n_{mod} = 5$  for analysis including biomass, and  $n_{mod} = 3$  for analyses including NI and NE metabolites), and  $n_{mc}$  is the number of realizations (30,000).

### Measuring model separation: Energy distance (ED) for OD

We use energy distance (ED) as a measure of distance between models; more specifically, between normalized model outputs  $y_{normalized}(t)_{m,l}$  (eq. 127). Generically, ED provides a measure of distance between distributions [114]. For our analysis, we generate distributions of normalized model outputs based on particular experimental designs. The energy distance measure accounts for variance by discounting the distance between model outputs by the within-model variance [114]. Because of the normalization of the model outputs (eq. 127), the scale of energy distance will result “in units of error of standard deviations”, making the distances interpretable.

$$ED(X, Y) = \sqrt{2 \cdot E\|X - Y\| - E\|X - X'\| - E\|Y - Y'\|} \quad (128)$$

where X and Y are output distributions from two model instances, respectively, and X' and Y' are separate realizations of the models, respectively. The term  $E\|X - Y\|$  is the expected

Euclidian distance between these distributions, while  $E\|X - X'\|$  and  $E\|Y - Y'\|$  are their expected within-model variances. Discounting by the within-model ensures that noisy outputs that do not differ much between models contribute minimally (or even detract from) the ED measure of model separation.

The goal of our prospective optimal design method was to determine the design that provides the most informative data for model discrimination [112]:

$$d_{opt} = \operatorname{argmax}_{d \in D}(\text{Objective}) \quad (129)$$

where (D) is the set of candidate designs and the objectives are pairwise energy distances (ED) between the set of models under consideration (specified in the Results section).

To make reliable statements about the ED, we chose a subsample size of 10,000 out of the 30,000 simulated data outputs per model variant because this sample size showed a stable ED throughout the candidate designs (Figure S58). We normalized the ED scores by the maximum ED of the candidate designs and applied a multi-objective Pareto optimization (MATLAB's "prtp" function [354]) to determine the non-dominated designs [355] ( $d_{opt}$ ). These are presented as spider plots [356] and were produced in Matlab.

## 9.4 Results and Discussion

### 9.4.1 Can we distinguish active AT degradation pathways based on observations of metabolite concentrations?

To determine the active AT pathway in a particular environment, we explored whether the six proposed models (M1-M6), representing conceptual AT degradation pathways, can be differentiated based on the observation of metabolites and CA and CO<sub>2</sub> (sinks) concentrations. Because AT, the main metabolites (HY, DIA, DEA), and the sink pools (CA and CO<sub>2</sub>) are common to all six model variants (see Figure 13), for candidate experimental designs, we consider measurement of subsets of these, with different sampling strategies, giving 63 possible



combinations of chemical output channels. We do not impose any resource costs for each measurement channel, and so one might expect that the optimal design is to measure all candidate channels to maximize the information gathered. However, the energy distance metric accounts for a trade-off between noise and comparison: the inclusion of noisy, information-poor channels will result in a *drop* in the energy distance between two models (as demonstrated below), incorporating these output channels into the design ‘muddies the waters’; that is the information contained in these measurements makes the model discrimination task more difficult.

With the 63 possible combinations of chemical output channels, we consider three sampling frequencies: 1) every day until day 25, 2) every two days until day 50, 3) every three days until day 100, giving  $63 \times 3 = 189$  designs in total. Figure 14 shows pair-wise model energy distance scores for every candidate design.

Based on a minimum threshold of energy distance of two for model discrimination (horizontal dashed line in panels A to F in Figure 14), the models fell into three groups: i) M1, M2, M3 and M4 (henceforth M|1-4); ii) M5; and iii) the first-order decay model M6. We selected a minimum ED of two because and similarly to standard deviations, a two standard deviation distance would correspond to being outside of the 95% confidence interval.

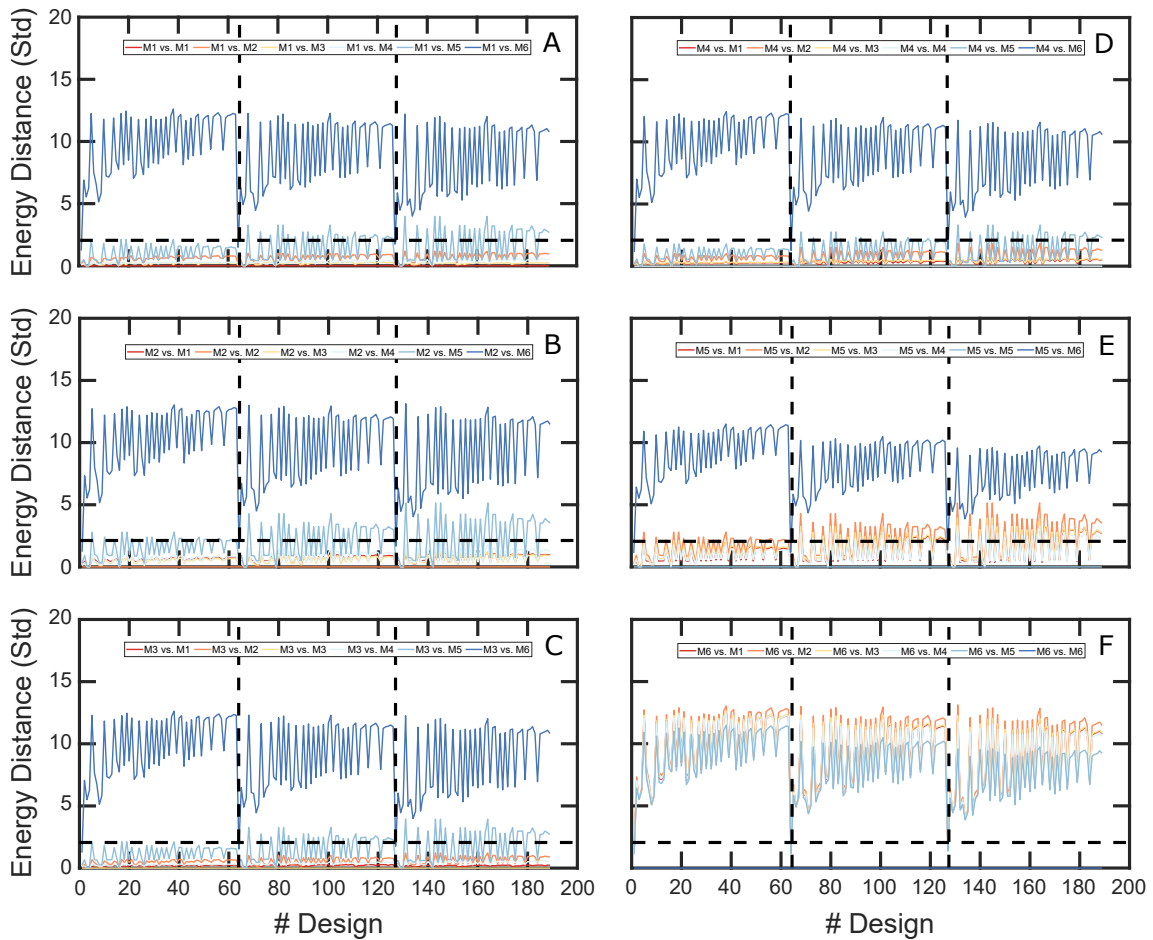


Figure 14: Pairwise energy distances (expressed in standard deviation units) over the candidate designs. A. Model M1 against other models; B. Model M2 against other models; C. Model M3 against other models; D. Model M4 against other models; E. Model M5 against other models; F. Model M6 against other models. Vertical lines represent transition from one sampling frequency/duration to another: left is daily over 25 days; middle is every second day over 50; right is every four days over 100. The horizontal line represents the selected minimum energy distance threshold for model discrimination (distance of two standard deviations)

As expected the simplest, first-order decay model (M6) can be clearly distinguished from the Monod-based models (M1-M5) with all experimental designs, except when only measuring AT (first design in each time-related group) regardless of the sampling frequency.

Likewise, M5 clustered separately. In M5, the complete AT degradation and dealkylation

of AT is mediated by bacterial guilds A and D. There is no formation of NI and NE because the microbial guilds B and C are missing. The absence of these two guilds leads to reduced biodegradation of HY. Additionally, the main source of HY in the soil solution is the leak out of guild A [62, 63] or abiotically produced HY [230, 357].

Models M1, M3 and M4 clustered within a joint group (M|1-4). These three models account for AT dealkylation and dechlorination by guilds A and D, and either guild B or C is present. Interestingly, M2 clusters within the M1-4 group, too, despite the absence of the full degradation pathway of AT carried out by guild A. The absence of guild A eliminates any degradation of the chlorinated AT metabolites DIA, and DEA [152, 358], leading to the accumulation of both compounds. However, the accumulation of DIA and DEA occurred largely at concentrations below detection and the overall dynamics of AT and its metabolites was comparable to the other models within this group. Models within this group are not distinguishable based on AT, metabolite observations (HY, DIA, and DEA, CA), and CO<sub>2</sub>. Because each model variant within M|1-4 represents a specific variant of potentially active AT degradation pathways, additional measurements of microbial biomass might improve model discrimination and thus allow identification of the corresponding active AT degradation pathways, as we discuss below 9.4.3.

#### **9.4.2 Which experimental designs provide the most informative data for model discrimination?**

We analyzed the discrimination of the three identified model groups in detail by calculating EDs between i) M|1-4 and M6, ii) M5, and M6, iii) M|1-4 and M5. In this analysis, comparison between M5 and M6 with group M|1-4 was done by treating the group M|1-4 as a single model by adding up the individual EDs from models M1 to M4 (group M|1-4) to M5 and M6 and normalizing it by the maximum added ED of the candidate designs. We performed a multi-objective Pareto analysis (see M&M) to determine the designs that maximize group discrimination, i.e., those experimental designs which lead to a maximum ED of one objective, while simultaneously

minimizing the ED decrease of the objectives (non-dominated Pareto optimal solutions).

### Optimal designs and measured pools:

From the Pareto analysis, six experimental designs out of 189 were identified as optimal (Figure 15): (i) measurement of DEA, DIA, and CA with short length, (ii) measurement of CA in mid length, (iii) measurement of DEA, DIA, and CA with mid length, (iv) measurement of CA in long length, (v) measurement of DIA and CA with long length, and (vi) measurement of DEA, DIA, and CA with long length.

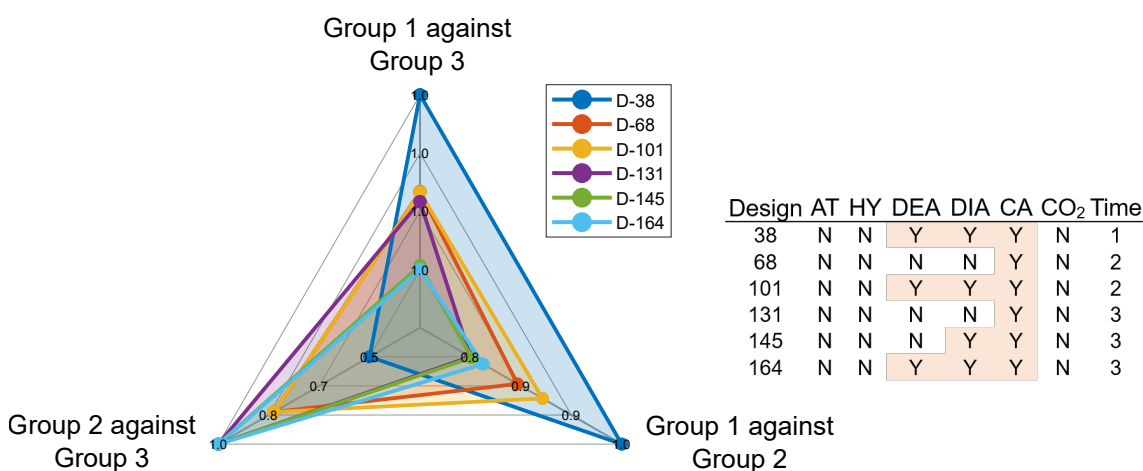


Figure 15: Optimal designs from the Pareto front based only on AT, metabolite (HY, DEA, DIA) and sinks (CA, CO<sub>2</sub>) measurements (Y = measured, N = not measured). Time column's values are: (1) short-term, every day sampling until day 25, (2) middle-term, every two days sampling until day 50, (3) long-term, every four days sampling until day 100

None of the 6 optimal designs include measurements of AT, HY or CO<sub>2</sub>. Thus, these chemical pools do not provide informative data for model discrimination (related to the similarities of their simulated time-series for all model versions in Figures S52-S57). Including these measurements incorporates noise into the ED measure, confounding model discrimination. For example, as shown in Table 12, addition of the AT output channel increases the noise terms with negligible increase in the comparison term of the energy distance. We can, therefore, conclude that to understand the fate of AT in real systems, one should prioritize information about the

intermediate metabolites (DEA, DIA, CA). Unfortunately, most experiments on AT to date have mainly measured AT [359, 360].

Table 12: Single elements of the energy distance (ED) between model M1 (X) and M6 (Y)

Design	AT	HY	DEA	DIA	CA	CO <sub>2</sub>	Time	$2 \cdot E\ X - Y\ $	$-E\ X - X'\ $	$-E\ Y - Y'\ $	$\sqrt{ED(X, Y)}$
38	N	N	Y	Y	Y	N	1	339.1	8.0	171.9	12.6
48	Y	N	Y	Y	Y	N	1	341.5	19.6	173.1	12.2
131	N	N	N	N	Y	N	3	198.6	13.8	35.6	12.2
136	Y	N	N	N	Y	N	3	211.6	38.9	45.5	11.3

Non-dominated designs (38, 131) and equivalent designs, including AT (48, 136). As expected, by including AT measurements, the increase in noise outweighs the increase in comparison, so that in total, the value of ED decreases.

CA seems to provide the most informative data among the intermediate metabolites because it is mainly accumulated in all selected optimal designs (principal end product of AT degradation in our simulations). This well resembles field observations, where further degradation of CA to CO<sub>2</sub> occurs only under low concentrations of nitrogen (N) in soils [137, 139]. Additionally, DIA and DEA, products of the dealkylation of AT [146, 344], occur in four of six non-dominated designs. The power of these two pools to enhance model discrimination is expected due to the different DIA and DEA dynamics simulated by the models. For example, DIA and DEA are hardly produced (under the detection limit of  $10^{-7}$  mg cm<sup>-3</sup>) by models in Group M1-4. In contrast, M6 and M5 simulate detectable DIA and DEA concentrations (Figure S52-S57), leading to differences in the ED.

### Optimal designs and sampling frequencies:

In our proposed designs, we incorporated short-, middle- and long-term experiments with a total of 25 equally-spaced samples. The non-dominated designs (Figure 15) include short-, middle- and long-term experiments. The first-order decay model can better be distinguished from the other clusters in short-term designs (Design 38 in Figure 15). This could be related to the tendency of these models to quickly reach steady-state in the simulations. In contrast, the

Monod-based models are better distinguishable in middle- and long-term experiments, probably because these clusters produce distinctive endpoints of the metabolites that differ according to the associated pathway.

### **9.4.3 Can measuring pools not commonly measured improve model discrimination?**

In the analysis in sections 3.1 and 3.2, we considered experimental designs that involve measurements of AT, the metabolites (HY, DIA, and DEA), and the sink pools CA and CO<sub>2</sub>. These correspond to measurements that could be collected in lab dissipation experiments [346]. Next, we examine the potential of less typical measurements: biomass and the NI and NE metabolite pools. We begin by defining three new sets of candidate designs incorporating along with the previous set of candidates (Section 9.4.1) i) measurements of total biomass and/or biomass of guild D (255 possible combinations of chemicals and biomass and three sampling frequencies giving 765 designs), ii) the metabolite NI (127 possible combinations of chemicals and three sampling frequencies giving 381 designs), and iii) the metabolite NE (127 possible combinations of chemicals and three sampling frequencies giving 381 designs). Because M6 and M5 can already be differentiated based on the chemical measurements, we focused only on distinguishing models within Group M1-4.

#### **Role of biomass measurements**

We observed that by adding the biomass information, the energy distances increased, allowing model discrimination (Figure 16) of all models within Group M1-4. These results highlight the importance of biomass measurements to identify the active AT degradation pathway in soil.

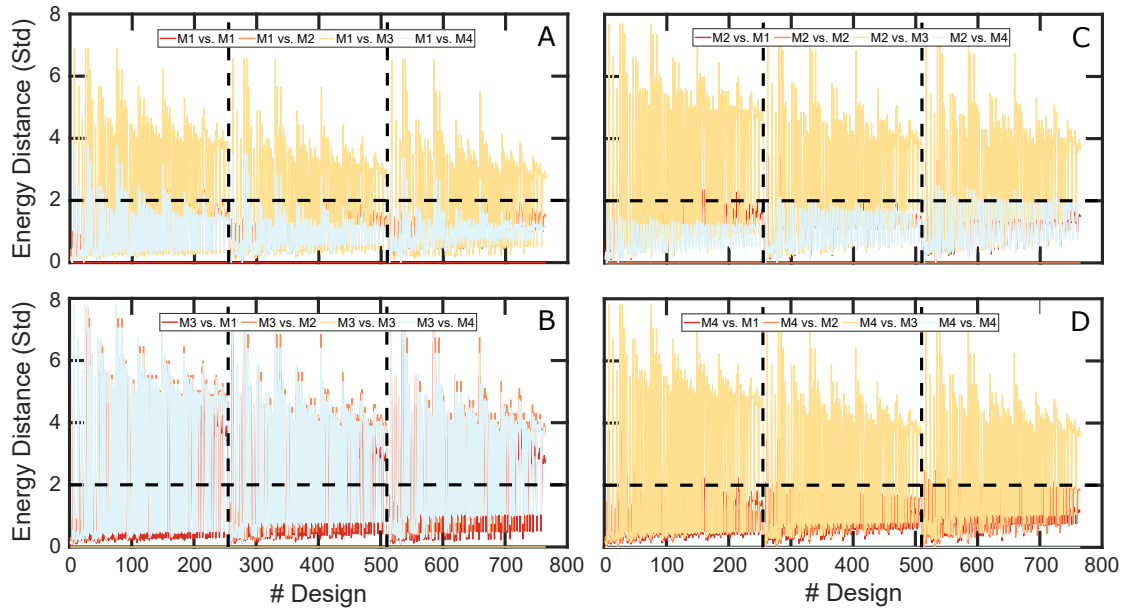


Figure 16: Pairwise energy distances (expressed in standard deviation units) over the candidate designs including biomass measurements. A. Model M1 against other models; B. Model M3 against other models; C. Model M2 against other models; D. Model M4 against other models. Vertical lines represent transition from one sampling frequency/duration to another: left is daily over 25 days; middle is every second day over 50; right is every four days over 100. The horizontal line represents the selected minimum energy distance threshold for model discrimination (distance of two standard deviations)

Next, we applied a Pareto analysis to the 765 biomass-including designs based on the new goal to maximize the distance between models in Groups M|1-4. As objectives we chose energy distance maximization between: i) M4 vs. M1, ii) M4 vs. M2, iii) M4 vs M3, iv) M3 vs. M1, v) M3 vs. M2, vi) M2 vs M1. We identified 15 out of the 765 designs as optimal [non-dominated] (Figure 17). The optimal designs showed that measurements of total biomass do not contribute to model discrimination. However, not surprisingly, measurement of guild D biomass helps to distinguish among the Monod models. Unfortunately, distinguishing specific guilds is challenging. Guild membership could be estimated from specific genes responsible for AT degradation, as done by Pagel et al. [43] for the herbicide MCPA. However, some degraders contain genes from multiple guilds, leading to an overestimation of the degrader biomass. Therefore, total biomass

is commonly prioritized over the biomass of specific degraders. We expect that advances in molecular biology will provide the tools to make a more accurate quantitative identification of particular degraders, and thus, pathway identification possible.

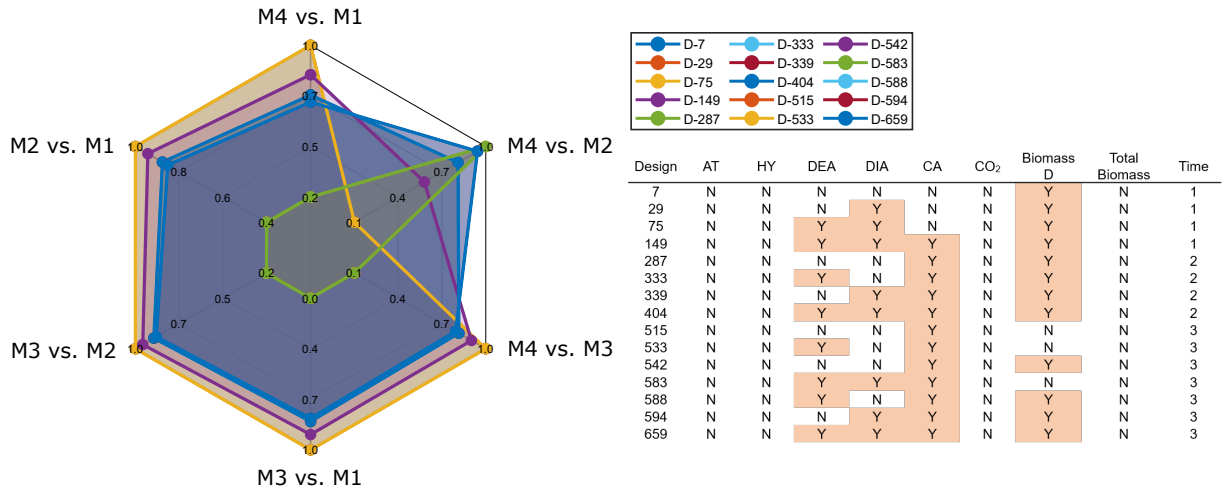


Figure 17: Optimal designs including biomass measurements (Biomass of guild D and total Biomass). Y = measured, N = not measured. Time column's values are: (1) short-term, every day sampling until day 25, (2) middle-term, every two days sampling until day 50, (3) long-term, every four days sampling until day 100

#### 9.4.4 Role of NI-NE measurements:

NI and NE metabolites appear only in models M1, M2, M3 and M1, M2 and M4, respectively. From Figure 18 (A, B, C), it becomes clear that NI measurements can only help to distinguish model M3 from models M1 and M2, especially in designs that include middle and longer sampling frequencies. Likewise, including NE measurements can only help to differentiate M4 from M1 and M2 (Figure 18 D, E, F).



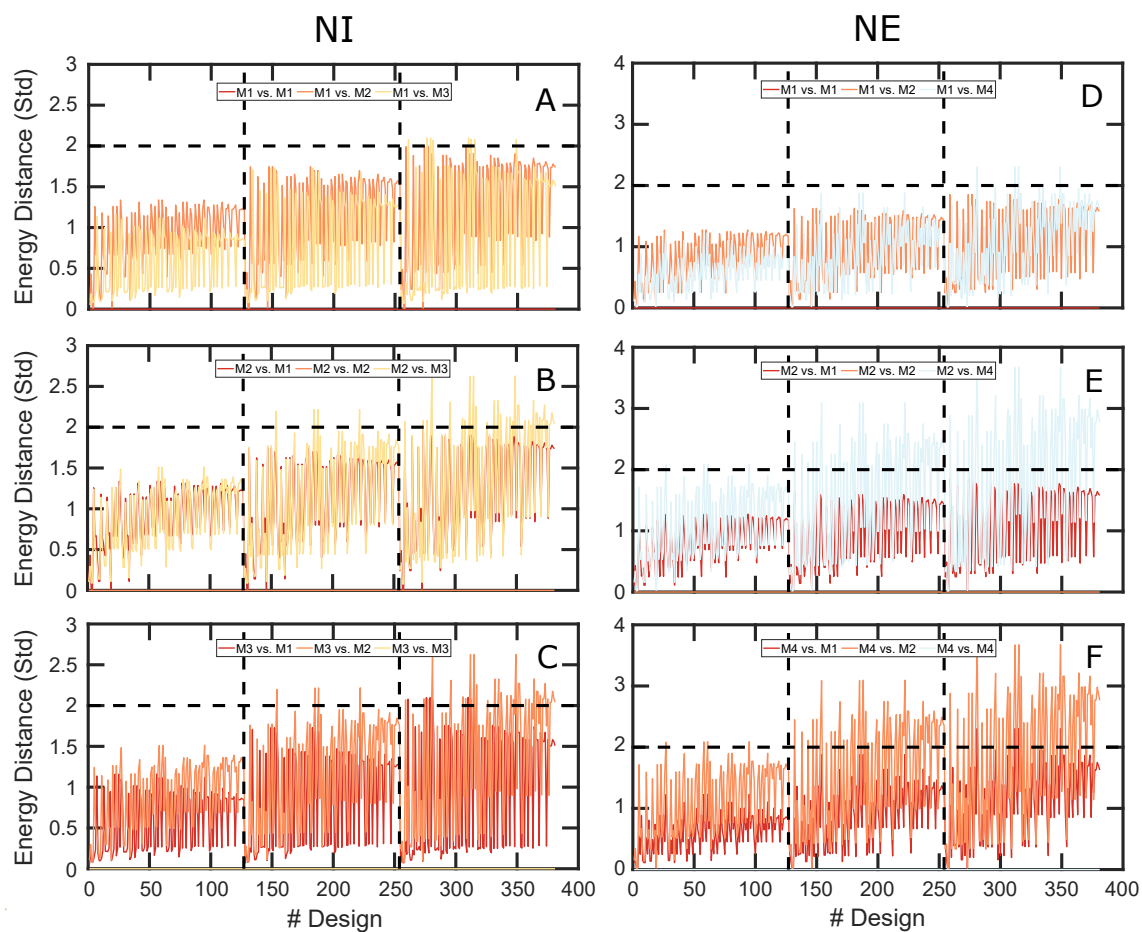


Figure 18: Pairwise energy distances (expressed in standard deviation units) over the candidate designs including i) NI: A. Model M1 against other models; B. Model M2 against other models; C. Model M3 against other models; and ii) NE: D. Model M1 against other models; E. Model M2 against other models; F. M4 against other models. Vertical lines represent transition from one sampling frequency/duration to another: left is daily over 25 days; middle is every second day over 50; right is every four days over 100. The horizontal line represents the selected minimum energy distance threshold for model discrimination (distance of two standard deviations)

Because adding NI and NE metabolites observations does not contribute to model discrimination between model M1 and M2, when applying the multi-objective Pareto analysis to NI- and NE-containing designs, we set objectives to maximize the difference i) M3 vs. M1, ii) M3 vs. M2, for the designs including NI; and i) M4 vs. M1, ii) M4 vs. M2 for the designs

including NE (381 in each case). Two designs were determined as the non-dominated designs for each analysis, and therefore, instead of in spiderplots, the results are presented in Table 13.

These designs include NI and NE observation. These two metabolites are rarely found in soils [361], probably because of their faster degradation rates mediated by the gene *atzC* [342]. Regarding sampling frequencies and duration of experiments, long-term experiments proved more helpful for model discrimination when including NI and NE measurements.

Table 13: Optimal designs after multiobjective pareto analysis of candidate designs including NI and NE in addition to AT, metabolites (HY, DIA, DEA) and sinks (CA, CO<sub>2</sub>)

Design	AT	HY	DEA	DIA	CA	CO <sub>2</sub>	NI	Time
310	N	N	Y	Y	N	N	Y	3
349	N	N	Y	Y	Y	N	Y	3
Design	AT	HY	DEA	DIA	CA	CO <sub>2</sub>	NE	Time
315	N	N	N	Y	Y	N	Y	3
349	N	N	Y	Y	Y	N	Y	3

Time column's values is: (3) long-term, every four days sampling until day 100

#### 9.4.5 Implications for biogeochemical modeling and data integration

In this work, we applied a prospective optimal design (OD) of experiments to find experimental sampling strategies that allow for discrimination among competing atrazine (AT) degradation models and the corresponding degradation pathways. Our method is reliable (Figure S58), and it can be performed prior to the execution of the experiment. Applying the Bayesian constrained-based parameter search algorithm (CBS-MCMC) for efficiently sampling the viable parameter set dramatically reduced the computational demand. The CBS-MCMC method is widely applicable to other biogeochemical models and provides a powerful tool to leverage expert knowledge for constructing robust prior parameter distributions for model sensitivity analysis or calibration.

In our study, we observed that the five proposed Monod models could be reduced to two groups, according to their predominant features. On the other hand, the first-order decay

model could only replace complex model formulations when looking at the AT degradation dynamics. However, considering the intermediate metabolites (DIA and DEA), and especially the sink pool CA is an integral part of understanding the complete degradation pathway of AT and of adequate model selection as well. This non-intuitive result is the consequence of the OD objective, which is, in our case, model discrimination. Furthermore, information on particular pesticide degraders showed potential to improve model and pathway identification. Thus, experimental measurements of specific guilds should be prioritized in the future.

For a practical application of our results towards the identification of the active AT degradation pathway in the system, we recommend using the following protocol:

1. include, but not limit to, the optimal design setup in the sampling strategy for the planned experiment. For better model discrimination, the best design should include measurements of specific biomass degraders (examples in Figure 17)
2. carry out model calibration for all available models or model groups against the collected data and the optimal design data
3. the best model for any particular case should best fit the data, and the differences among models or model clusters should become more predominant at the OD data points
4. the models or model groups that deviate the most from the data can be rejected (model invalidation) [336] as they are not valid representations of the system to be studied

The application of prospective OD of experiments requires that models use correct process descriptions. Therefore, the candidate model formulations must be carefully selected to ensure that the best possible representation is used [112]. As long as such valid process models are available - as in this study for atrazine degradation - model-based prospective OD will maximize the knowledge gain on soil systems from laboratory and field experiments.

## 10 General Discussion

Considering that many biodegradable pesticides can persist in soils [55], the objective of my thesis is to elucidate some of the key processes that control pesticide degradation, especially at low concentrations. I analyzed biological processes that can limit degradation, assuming an ideal scenario in which bacterial degraders and the specific pesticide are colocalized (Figure 2). This simplification helped me to:

- (i) evaluate complex bacterial processes (genetic and biophysical constraints) without the influence of soil heterogeneity [362], water transport processes [234] (dispersion, advection), or competing carbon sources [43, 363], all known to also affect degradation rates.
- (ii) easily validate model assumptions with data from lab experiments.

I used three pesticides as model pesticides (MCPA/2,4-D and atrazine) belonging to two pesticide classes (Chlorophenoxy herbicides and triazines, respectively) to explore different degradation mechanisms that could be transferable to other pesticides with similar properties. In section 5.2, I explained the characteristics of these pesticides and their role in my thesis. In the following subsections, I will discuss the main outcomes of my thesis and provide some insights for future directions of research in the field.

### 10.1 Improving process understanding and prediction of pesticide degradation in soils

I used a mechanistic, process-based modeling approach to evaluate pesticide degradation in soils. Among the different processes explored in my thesis, I could derive some mechanisms that control pesticide degradation and drive pesticide persistence in soils:

1. **Regulated gene expression (Research Question R1):** In the first study (Paper 1), and based on the need to improve the description of bacteria-driven pesticide degradation

in mechanistic models. [73], I explicitly incorporated regulated gene expression. This improved the description of the MCPA/2,4-D degradation in soils showing that regulated gene expression is a control for pesticide degradation.

2. **Mass-transfer across the membrane (Research Question R2):** In the second study (Paper 2), comparable to Sun et al. [364], I could confirm that mass-transfer limits degradation in the retentostat system, where low concentrations that lead to starvation are common [63]. However, by looking at long-term predictions, I found evidence that mass-transfer is unlikely to control pesticide persistence because the model simulations including mass-transfer led to complete degradation of atrazine, which is often not observed in the field. Contrasting results were found in water systems by Sun et al. [364], where mass-transfer of 2,6-dichlorobenzamide (BAM) across the cell membrane limits its degradation below  $600 \mu\text{g L}^{-1}$ .
3. **Bioenergetic constraints (Research Question R2):** According to LaRowe & Van Cappellen [64], the energy produced from catabolism of some pesticides might be insufficient to satisfy the metabolic needs of degraders. To validate this statement, I used transition state theory [242, 365] to model bacterial growth based on the use of pesticides as the sole carbon and energy source. Applying this concept to the degradation of atrazine, I found evidence that bioenergetic mechanisms are unlikely to control the persistence of hydroxyatrazine (the first metabolite of atrazine degradation). Additionally and, when the pesticide is degraded through mechanisms without energy gain, transition state theory is not sufficient to explain pesticide persistence. Hence, further research should prioritize energetic demands of biological transformations that do not involve energy gain, like atrazine degradation to hydroxyatrazine.
4. **Environmental factors: temperature and soil moisture (Research Question R3):** Pesticide dissipation experiments pointed out that low temperatures in combination with dry soil conditions, increase pesticide persistence in soils [366–368]. In the third study

(Paper 3), I assessed these findings through a mechanistic gene-centric model for MCPA degradation to evaluate a laboratory experiment. Although temperature, soil moisture, and pesticide concentration do play a role in the overall fate of the pesticide in soil, I found evidence that they mainly influence pesticide-derived carbon allocation. This conclusion is based on the increased incorporation of pesticide-derived C into the biomass, reflected by higher carbon use efficiencies under limiting conditions of temperature, soil moisture, and substrate concentration (10°C, pF 3.5 and 1 mg kg<sup>-1</sup>). Interestingly, model simulation results, as well as the observations, pointed to stress-induced mechanisms to survival. Under the evaluated limiting conditions, growth is not prioritized, and the higher C incorporation into the biomass might be related to the formation of substances to cope with stress [369]. Therefore, I could conclude that environmental factors play a role in pesticide-derived C allocation but are unlikely to control pesticide persistence. Even under limiting conditions, degradation of pesticides occurs. To confirm these findings, different and more persistent pesticides than atrazine should be prioritized in further studies.

With my models, I was able to identify some key processes relevant for pesticide degradation. Incorporating them into biogeochemical models improved predictions of pesticide degradation.

The full gene-centric model developed in the first study (Paper 1), although performing equally well as traditional Monod-kinetics models, showed to better represent bacterial dynamics in soil (gene and transcript data). This way, the gene-centric models allow us to study the relationship between process rates and functional genes. This relationship is typically assumed to be linear [128, 219], but my results challenge this assumption by showing a hysteretic relationship between these two variables. In conclusion, process rates cannot be directly derived from gene transcripts. Similar results were found by Störiko et al. [370] for bacterial denitrification in water systems. Additionally, using a gene-centric model capable of describing bacterial dynamics more accurately allowed me to derive better process descriptions of pesticide degradation, such as carbon allocation of pesticides in Paper 3. Despite the high model uncertainty, I could verify that

long-term predictions of hydroxyatrazine degradation after 30 years are better described with traditional Monod-kinetics in combination with retentostat data and leaching measurements.

It is important to keep in mind that working with complex, mechanistic models poses several challenges: the limited availability of data to validate model assumptions [75], the task of finding the most suitable model representation for our problem (“true model”) [109], and of model equifinality [73]. Section 10.2 will shed light on these challenges and outline possible solutions.

## **10.2 Challenges in mechanistic model development and how to overcome them**

“Essentially, all models are wrong, but some are useful” [371]. In my opinion, and based on my own experiences throughout my Ph.D., this expression is valid for probably all model applications. By working with mechanistic models, I found two main obstacles to model development: i) the nature of mechanistic models and ii) the complexity of the systems to model. This led me to explore the possibilities to overcome such obstacles.

1. **Nature of mechanistic models:** Marschmann et al. [73] thoroughly discussed the sloppiness of biogeochemical models, stating that “most model parameters cannot be derived from data”, making the models difficult to use for predictions or understanding systems behavior. Therefore, Marschmann et al. [73] applied a sophisticated data-driven method to derive less complex formulations (parsimonious model), whose parameters can finally be inferred from data. As an alternative to Marschmann et al. [73], I used a data-driven model reduction based on local sensitivity and identifiability scores in combination with information criterion indices [101] to find a potential parsimonious model formulation. Despite the advantages and disadvantages of both methods, they are suitable to obtain parsimonious model formulations and reveal structural model weaknesses. Currently, the main drawback is that experimental techniques cannot produce sufficient informative data for comprehensive model reduction procedures, which are necessary for further model development [75].

2. **Complexity of the systems to model:** In addition to the sloppy nature of most mechanistic models, our inability to comprehend complex systems directly translates to limitations in model development. This limited knowledge and understanding of all the processes involved in a particular problem [112], stems from our inability to measure all elements connected to it [372, 373]. For example, for atrazine degradation, many bacterial guilds are involved, but only a few have been identified. This leads to competing model formulations, each representing valid degradation pathways, making the selection of the “true model”, representing the active degradation pathway, a non-trivial problem [109]. This problem can only be tackled with more and better data; therefore, determining what quality and quantity of data are needed plays an important role in developing more robust models. In my thesis, I performed a prospective optimal design of experiments (before the possible execution of the experiments) to determine the most informative data to discriminate competing models of atrazine degradation in soil (Paper 4). Assuming that the “true model” is in our set of models, optimal design becomes a powerful tool to overcome equifinality, which arises when there are different valid model parametrizations for a single model or several competing model formulations for a single problem. Additionally, optimal design can be used to generate informative data to reduce uncertainty in model predictions [111, 112], and to strengthen model calibration (through better parameter identification).

### **10.3 Towards a predictive model: applicability and further extensions of the work**

The model formulations presented in this thesis are highly detailed in their formulation and impractical to be used on large scale [374]. It is important to consider that one of the aims of modeling development is to end up with a “predictive model”, a model with the predictive capabilities for practical applications, such as environmental fate modeling.

As the name suggests, predictive models are tools used to predict the behavior of chemicals,



such as pesticides, in all compartments of the environment (soil, water, air, and organisms) [374]. Guided by the principle of parsimony [374], a predictive model should provide a simple but still accurate description of the studied system, with modest dataset requirements for its functioning and a low computational effort to produce the simulations [375]. Such simple predictive models are then used to estimate the degradation endpoints of parent compounds [65], supporting environmental risk assessment [375] and decision-making [374, 376]. In the EU, predictive models are commonly based on first-order decay, ranging from simple first-order models to bi-phasic models (Gustafson and Holden model), that are able to incorporate the effects of soil heterogeneity into degradation dynamics.

Based on my work, I briefly recall the structure of my model development, as shown in Figure 19, including potential steps towards develop a predictive model. I suggest that a mechanistic process-based model can also provide the basis for developing a predictive model. The process to reveal the relevant mechanisms for a predictive model includes several, non-trivial steps and starts with the conceptualization of a theoretical model (Step 1 in Figure 19), which is formulated from literature and expert knowledge. The individual model assumptions of the theoretical model need to be validated against experimental data (Step 2 Figure 19), which requires laboratory data. For further model development (Step 3 Figure 19), Steps 2 and 3 create a feedback loop of data integration modeling, in which the theoretical model, through optimal design of experiments, can be used to identify the type of experiments needed to maximize data gain [112]. With better data, each iteration of calibration, validation, model reduction, and model selection further refines the model. After assembling the mechanistic process-based model, an upscaling process is necessary for the final development of a predictive model (Figure 19, Step 4).

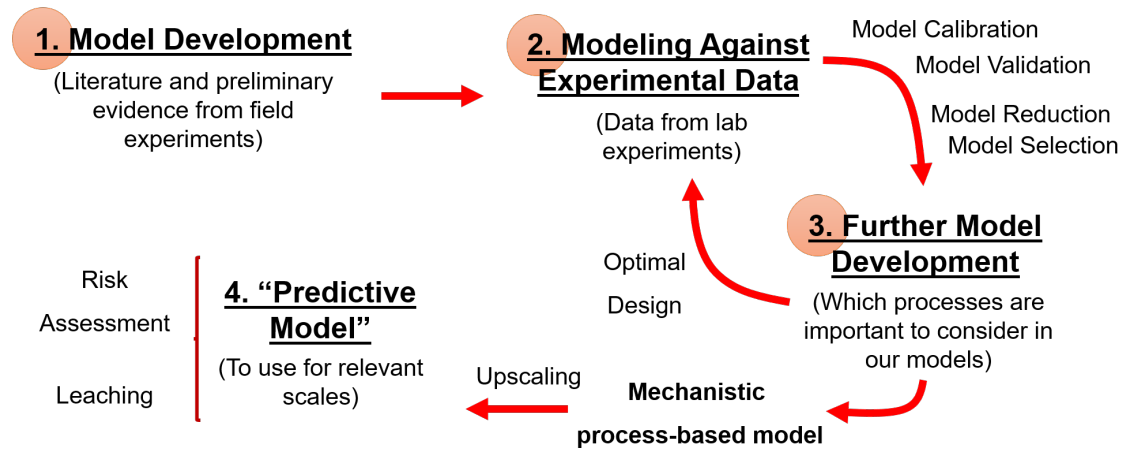


Figure 19: Modeling microbial regulation of pesticide turnover. Pipeline to develop a predictive model. Steps included in this thesis are marked with red circles

A good starting point of assembling and upscaling a mechanistic process-based model for pesticide degradation could be to improve the mechanistic PECCAD [43, 78] model. The mechanisms identified in my thesis to control pesticide degradation, especially under low concentrations (Section 10.1, Papers 1 to 3), can easily and directly be integrated in the PECCAD model, complementing its descriptions of pesticide degradation in soils. For example, the twofold regulation of microbial growth, as described in Papers 1 and 3 by the Hill function (for regulated gene expression) and Monod kinetics (for the substrate dependency), can be easily used to describe bacterial growth in the PECCAD model. Depending on the pesticide class, Monod kinetics could be used to represent substrate dependency of growth because we found that it can better describe pesticide persistence. Finally, the temperature function used in Paper 3 to account for the effect of temperature on pesticide degradation can be coupled to all relevant processes (growth, maintenance, death rates). Thus, the updated PECCAD model would offer a comprehensive representation of the processes relevant for pesticide degradation in soils, and at the same time, serve as a benchmark to compare simulations produced by simplified expressions with.

Literature offers several upscaling methods that could be utilized for pesticide degradation

models. When upscaling the updated PECCAD model, it should be possible to simulate effects of microbial and pesticide heterogeneities, variability of soil properties, and other environmental conditions on pesticide degradation in soils, all variable in time and space [377]. Chakrawal et al., [378] distinguished three approaches for the “upscaling of decomposition kinetics for carbon cycling models”, all of them applicable to pesticide degradation models: i) numerical spatial averaging, which divides the domain into grids and solves the mechanistic model for each grid. It is a computationally expensive method with a high demand of input data [379], ii) the effective parameter approach, which consists in deriving “effective parameters” to link degradation rates across different scales [380], and iii) analytical upscaling, which relies on a spatial averaging of the kinetic equations of small scales. However, the lack of the so-called “closure term for integration” [378] makes the method still dependent on small scales, and numerical integration similar to numerical spatial averaging is still needed. Out of the three, the analytical upscaling could have the greatest potential for a successful upscaling of small-scale pesticide degradation models if the closure integration problem could be solved.

#### **10.4 Research perspectives: Deep neural networks**

Machine learning techniques, represented by deep neural networks (DNN), have recently begun to be explored as an alternative approach to develop predictive models [381] and have the potential for pesticide degradation modeling. Machine learning can detect statistical relationships between input and output along multiple spatial and time scales [381]. These tools are currently used just for predictions, as done by Sigmund et al. [251], who predicted Freundlich isotherm sorption parameters for carbonaceous substances with a deep neural network. However, I believe deep neural networks could be combined with mechanistic modeling approaches [381], for example, to assist in the identification of the processes responsible for pesticide degradation. Also considering the advances in molecular biology, DNNs could be applied to large omics datasets. A systematic analysis of omics information (metagenomics, metatranscriptomics, metaproteomics, and metabolomics) could help to elucidate the “structure

and function of microbial communities” [382] involved in pesticide degradation in soils. A first idea of how this combination could look in a model development process is presented in Figure 20.

Adapted from Baker et al. [381], a DNN could be used with omics datasets to extract patterns from omics data that will help us to reduce complexity in the development of a predictive model. For example, microorganisms that use similar carbon sources or carry out similar degradation processes in soils could be grouped as a single super bacteria. However, machine learning can be prone to overfitting, which would result in a low predictive capability of the neural network. To avoid that, the mechanistic model would serve as a surrogate model that could capture known relationships and provide them to the DNN [383]. In a next step, the DNN could even be used to extract model parameters [251] or approximate the analytical solution of biokinetic equations [384]. For that, a training step of the DNN should be performed, combining the original data with the insights learned from the mechanistic model. After the validation, the neural network could be used for predictions. With such a modeling approach, the DNN could potentially bridge the gap between different scales, providing us with a promising direction for future research in the modeling of (microbial) regulation of pesticide turnover in soils.

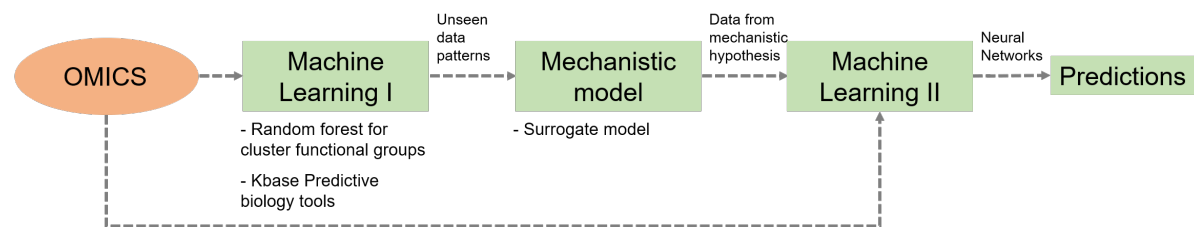


Figure 20: Theoretical pipeline to derive large scale predictions combining machine learning techniques and mechanistic modeling approaches. Adapted from Baker et al. [381]. Further information on Kbase tools can be found in the Kbase website

## 11 Appendices

### 11.1 Supplementary Information for Chapter 6 (Paper 1)

#### 11.1.1 Model Formulations for reduced Models

The full version of the governing differential equations for our proposed models are the following:

**V1 :**

$$\frac{dC_B^a}{dt} = \underbrace{\mu_P \cdot C_B^a}_{r_{\text{growth}}} - \underbrace{C_B^a \cdot a_a}_{r_{\text{decay-Active}}} \quad (1)$$

$$\frac{dC_B^d}{dt} = \underbrace{C_B^a \cdot a_a}_{r_{\text{decay-Active}}} - \underbrace{C_B^d \cdot a_d}_{r_{\text{decay-Dead}}} \quad (2)$$

$$\frac{dC_P^L}{dt} = \frac{-\mu_P \cdot C_B^a \cdot \left(\frac{1}{Y_P}\right) \left(\frac{\rho}{\theta}\right)}{\underbrace{\left(1 + \frac{\rho}{\theta} \cdot K_{FP} \cdot n_{FP} \cdot (C_P^L)^{(n_{FP}-1)}\right)}_{r_{\text{uptake}}}} \quad (3)$$

$$\frac{dCO_2}{dt} = \underbrace{\mu_P \cdot C_B^a \cdot \left(\frac{1 - Y_P}{Y_P}\right)}_{r_{\text{respiration}}} + \underbrace{C_B^d \cdot a_d \cdot a_{CO_2}}_{r_{\text{cell-decomposition}}} \quad (4)$$

$$\widehat{mRNA} = \underbrace{f_T \cdot \left(\frac{(C_P^L)^{n_H}}{(K_G)^{n_H} + (C_P^L)^{n_H}}\right)}_{r_{\text{transcription}}} \cdot C_B^a \quad (5)$$

Where the degradation rate coefficient  $\mu_P$  is:

$$\mu_P = \mu_{\text{max}} \cdot \left( \frac{\left( \frac{(C_P^L)^{(n_H+1)}}{(K_G)^{n_H} + (C_P^L)^{n_H}} \right)}{K_M + C_P^L} \right) \quad (6)$$

V2 :

$$\frac{dC_B^a}{dt} = \underbrace{\mu_P \cdot C_B^a}_{r_{\text{growth}}} - \underbrace{C_B^a \cdot a_a}_{r_{\text{decay-Active}}} + \underbrace{\tau \cdot k_r \cdot C_B^i}_{r_{\text{activation}}} - \underbrace{(1 - \tau) \cdot k_d \cdot C_B^a}_{r_{\text{uptake}}} \quad (7)$$

$$\frac{dC_B^i}{dt} = \underbrace{(1 - \tau) \cdot k_d \cdot C_B^a}_{r_{\text{deactivation}}} - \underbrace{\tau \cdot k_r \cdot C_B^i}_{r_{\text{activation}}} - \underbrace{C_B^i \cdot a_i}_{r_{\text{decay-Inactive}}} \quad (8)$$

$$\frac{dC_P^L}{dt} = \frac{-\mu_P \cdot C_B^a \cdot \left(\frac{1}{Y_P}\right) \left(\frac{\rho}{\theta}\right)}{\underbrace{\left(1 + \frac{\rho}{\theta} \cdot K_{FP} \cdot n_{FP} \cdot (C_P^L)^{(n_{FP}-1)}\right)}_{r_{\text{uptake}}}} \quad (9)$$

$$\frac{dCO_2}{dt} = \underbrace{\mu_P \cdot C_B^a \cdot \left(\frac{1 - Y_P}{Y_P}\right)}_{r_{\text{respiration}}} + \underbrace{(C_B^a \cdot a_a + C_B^i \cdot a_i) \cdot a_{CO_2}}_{r_{\text{cell-Decomposition}}} \quad (10)$$

$$\widehat{mRNA} = \underbrace{f_T \cdot \left(\frac{(C_P^L)^{n_H}}{(K_G)^{n_H} + (C_P^L)^{n_H}}\right)}_{r_{\text{transcription}}} \cdot C_B^a \quad (11)$$

Where the degradation rate coefficient  $\mu_P$  is:

$$\mu_P = \mu_{max} \cdot \left( \frac{\left( \frac{(C_P^L)^{(n_H+1)}}{(K_G)^{n_H} + (C_P^L)^{n_H}} \right)}{K_M + C_P^L} \right) \quad (12)$$

And the switch function  $\tau$  is:

$$\tau = \frac{1}{\exp\left(\frac{C_T - C_P^L}{n \cdot C_T}\right) + 1} \quad (13)$$

**V3 :**

$$\frac{dC_B^a}{dt} = \underbrace{\mu_P \cdot C_B^a}_{r_{\text{growth}}} - \underbrace{C_B^a \cdot a_a}_{r_{\text{decay-Active}}} \quad (14)$$

$$\frac{dC_P^L}{dt} = \frac{-\mu_P \cdot C_B^a \cdot \left(\frac{1}{Y_P}\right) \left(\frac{\rho}{\theta}\right)}{\underbrace{\left(1 + \frac{\rho}{\theta} \cdot K_{FP} \cdot n_{FP} \cdot (C_P^L)^{(n_{FP}-1)}\right)}_{r_{\text{uptake}}}} \quad (15)$$

$$\frac{dCO_2}{dt} = \underbrace{\mu_P \cdot C_B^a \cdot \left(\frac{1 - Y_P}{Y_P}\right)}_{r_{\text{respiration}}} + \underbrace{C_B^a \cdot a_a \cdot a_{CO_2}}_{r_{\text{cell-Decomposition}}} \quad (16)$$

$$\widehat{mRNA} = \underbrace{f_T \cdot \left(\frac{(C_P^L)^{n_H}}{(K_G)^{n_H} + (C_P^L)^{n_H}}\right)}_{r_{\text{transcription}}} \cdot C_B^a \quad (17)$$

Where the degradation rate coefficient  $\mu_P$  is:

$$\mu_P = \mu_{\text{max}} \cdot \left(\frac{(C_P^L)^{n_H}}{(K_G)^{n_H} + (C_P^L)^{n_H}}\right) \quad (18)$$

**V4 :**

$$\frac{dC_B^a}{dt} = \underbrace{\mu_P \cdot C_B^a}_{r_{\text{growth}}} - \underbrace{C_B^a \cdot a_a}_{r_{\text{decay-Active}}} \quad (19)$$

$$\frac{dC_P^L}{dt} = \frac{-\mu_P \cdot C_B^a \cdot \left(\frac{1}{Y_P}\right) \left(\frac{\rho}{\theta}\right)}{\underbrace{\left(1 + \frac{\rho}{\theta} \cdot K_{FP} \cdot n_{FP} \cdot (C_P^L)^{(n_{FP}-1)}\right)}_{r_{\text{uptake}}}} \quad (20)$$

$$\frac{dCO_2}{dt} = \underbrace{\mu_P \cdot C_B^a \cdot \left(\frac{1 - Y_P}{Y_P}\right)}_{r_{\text{respiration}}} + \underbrace{C_B^a \cdot a_a \cdot a_{CO_2}}_{r_{\text{cell-Decomposition}}} \quad (21)$$

$$mRNA = \underbrace{f_T \cdot C_B^a}_{r_{\text{transcription}}} \quad (22)$$

Where the degradation rate coefficient  $\mu_P$  is:

$$\mu_P = \mu_{max} \cdot \left( \frac{C_P^L}{K_M + C_P^L} \right) \quad (23)$$

V4' :

$$\frac{dC_B^a}{dt} = \underbrace{\mu_P \cdot C_B^a}_{r_{growth}} - \underbrace{C_B^a \cdot a_a}_{r_{decay-Active}} \quad (24)$$

$$\frac{dC_P^L}{dt} = \frac{-\mu_P \cdot C_B^a \cdot \left( \frac{1}{Y_P} \right) \left( \frac{\rho}{\theta} \right)}{\underbrace{\left( 1 + \frac{\rho}{\theta} \cdot K_{FP} \cdot n_{FP} \cdot (C_P^L)^{(n_{FP}-1)} \right)}_{r_{uptake}}} \quad (25)$$

$$\frac{dCO_2}{dt} = \underbrace{\mu_P \cdot C_B^a \cdot \left( \frac{1 - Y_P}{Y_P} \right)}_{r_{respiration}} + \underbrace{C_B^a \cdot a_a \cdot a_{CO_2}}_{r_{cell-Decomposition}} \quad (26)$$

Where the degradation rate coefficient  $\mu_P$  is:

$$\mu_P = \mu_{max} \cdot \left( \frac{C_P^L}{K_M + C_P^L} \right) \quad (27)$$



### 11.1.2 Methods: Sampling data points

Table S1: List of sampling points batch degradation experiments

2,4-D				MCPA			
Days	Mineralization	<i>tfdA</i> transcripts	<i>tfdA</i> genes	Days	Mineralization	<i>tfdA</i> transcripts	<i>tfdA</i> genes
2	X	X	X	1	X		X
6	X			4	X		
8	X	X	X	7	X	X	X
9	X			11	X		
10.6	X			14.4	X		
11.8	X			17.4	X	X	X
12.8	X	X	X	22	X		
14	X			27	X		
14.7	X	X	X	33	X	X	X
15.8	X	X	X	respike			
16.6	X			33	X		
17.6	X	X	X	33	X	X	X
19.7	X	X	X	33.1	X	X	X
22.7	X			33.4	X	X	X
24.7	X	X	X	34	X	X	X
respike				34.5	X	X	X
24.9	X			36	X	X	X
25.1	X	X	X	39	X	X	X
25.3	X			43	X	X	X
25.7	X	X	X	46	X	X	X
26.2	X	X	X	53	X	X	X
26.7	X	X	X	67	X		
27.1	X	X	X				
27.7	X						
28.6	X	X	X				
29.7	X						
33.6	X	X	X				

### 11.1.3 Methods: Local and Global Sensitivity and uncertainty analysis

#### Local and Global Sensitivity

#### Sensitivity and uncertainty analysis

1. Sensitivity coefficient (SC): Local sensitivity analysis evaluates the changes on model outputs based on small changes on the model parameters, and is calculated as the “first derivatives of model outputs with respect to model parameters” [200]. This focused analysis leads to a straightforward interpretation of the results [199]. We solved the sensitivity equations for the state variables and the parameters to obtain exact time-series of the sensitivity:

$$S_i = \frac{\delta Y_i}{\delta p_j} \approx \frac{Y_i(p + \Delta p) - Y_i(p)}{\Delta p} \quad (28)$$

Where  $Y_i$  represents *ith* the model output, and  $p_j$  the *jth* parameter. We calculated a dimensionless relative sensitivity coefficient [101]:

$$\tilde{S}_i = \frac{(Y_i(p + \Delta p) - Y_i(p)) \cdot p}{\Delta p \cdot Y_i(p + \Delta p)} \quad (29)$$

These values are calculated per time point, so the overall sensitivity measure per parameter is calculated as follows:

$$\tilde{S}_j = \sqrt{\sum_i \tilde{S}_i^2} \quad (30)$$

The distinction between high and low leverage parameters is arbitrary and based on the obtained coefficients.

2. Identifiability scores (IS): We followed the orthogonalization method from [100] to evaluate the estimability of our parameters based on the experimental data available [128]. We first calculated the sensitivity coefficient matrix  $Z$  for each parameter  $j$  for each measured point  $i$  based on the relative sensitivity coefficients  $\tilde{S}_j$  [101]:

$$Z = \begin{pmatrix} \tilde{S}_{1,1} & \tilde{S}_{1,2} & \cdots & \tilde{S}_{1,j} \\ \tilde{S}_{2,1} & \ddots & \tilde{S}_{2,j-1} & \vdots \\ \vdots & \tilde{S}_{i-1,2} & \ddots & 0 \\ \tilde{S}_{i,1} & \cdots & \tilde{S}_{i,j-1} & \tilde{S}_{i,j} \end{pmatrix} \quad (31)$$

From this matrix, we selected the parameter (*j*th column)  $X_1$  with the highest square root of sum of squared of the column elements. This parameter is defined as the most identifiable parameter. To account for the influences of the additional parameters, [100] suggested to project each column of  $Z$  onto the column  $X_1$ , and collect the residual in a residual matrix  $Res$  with the same dimensions as the  $Z$  matrix:

$$Res = Z - X_1 \cdot (X_1^T \cdot X_1)^{-1} \cdot X_1^T \cdot Z \quad (32)$$

This process is repeated for the second most identifiable parameter based on the second highest square root of sum of squared of the column elements, and the process continues until all the parameters have been classified.

The distinction between high and low identifiability of parameters is arbitrary and based on the obtained scores.

3. Percentage error of the estimation (PE): We determined the Cramer-Rao inequality estimator as an alternative confidence interval for our parameters [385]. We calculated a second Fisher information matrix  $FIM$  weighted by the covariance matrix:

$$FIM = S^T \cdot W \cdot S \quad (33)$$

Where  $W$  is the inverse of the covariance matrix [101]. The 95% confidence interval [385]:

$$CI = 1.96 \cdot \sqrt{FIM^{-1}} \quad (34)$$

We reported the error per parameter:

$$\%Error = \frac{CI \cdot 100\%}{pj} \quad (35)$$

Percentage errors higher than 100% represent poorly characterized parameter estimations.

4. Correlation matrix: We calculated the correlation of our parameter using the covariance matrix  $cov(J^T \cdot J)^{-1}$ :

$$corr_{i,j} = \frac{cov_{i,j}}{\sqrt{cov_{i,i} \cdot cov_{j,j}}} \quad (36)$$

#### Global sensitivity analysis

Unlike a local sensitivity analysis, which only evaluates the impact of small changes on model outputs, a global sensitivity analysis evaluates changes within the entire parameter space [239], varying all the parameters together [105] to have a more robust information of the importance of the model parameters regarding the model outputs [201]. The Morris method [106] or elementary effects method is a partially global and inexpensive method for screening the important parameters for a high dimensional problem [103, 105]. This method generates two sensitivity measures: the mean of the elementary effects or  $\mu^*$ , and the standard deviation of the elementary effects or  $\sigma$  [201], calculated by averaging continuous local sensitivities over the parameter space [103, 107].  $\mu^*$  describes the overall impact of the individual parameter on the model output, whereas  $\sigma$  estimates interactions with other parameters [103]. This method is not a global sensitivity analysis method, but it is more detailed than just a local sensitivity analysis.  $\mu^*$  and  $\sigma$  are analyzed together defining parameters with a negligible effect on the model outputs (small  $\mu^*$  and  $\sigma$ ), parameters with a linear effect ( $\mu^*$  higher than  $\sigma$ ), and non-linear effects or parameter interactions ( $\sigma$  higher than  $\mu^*$ , and both higher values) [73, 103].

## 11.1.4 Additional Results

## Model Calibration on MCPA data

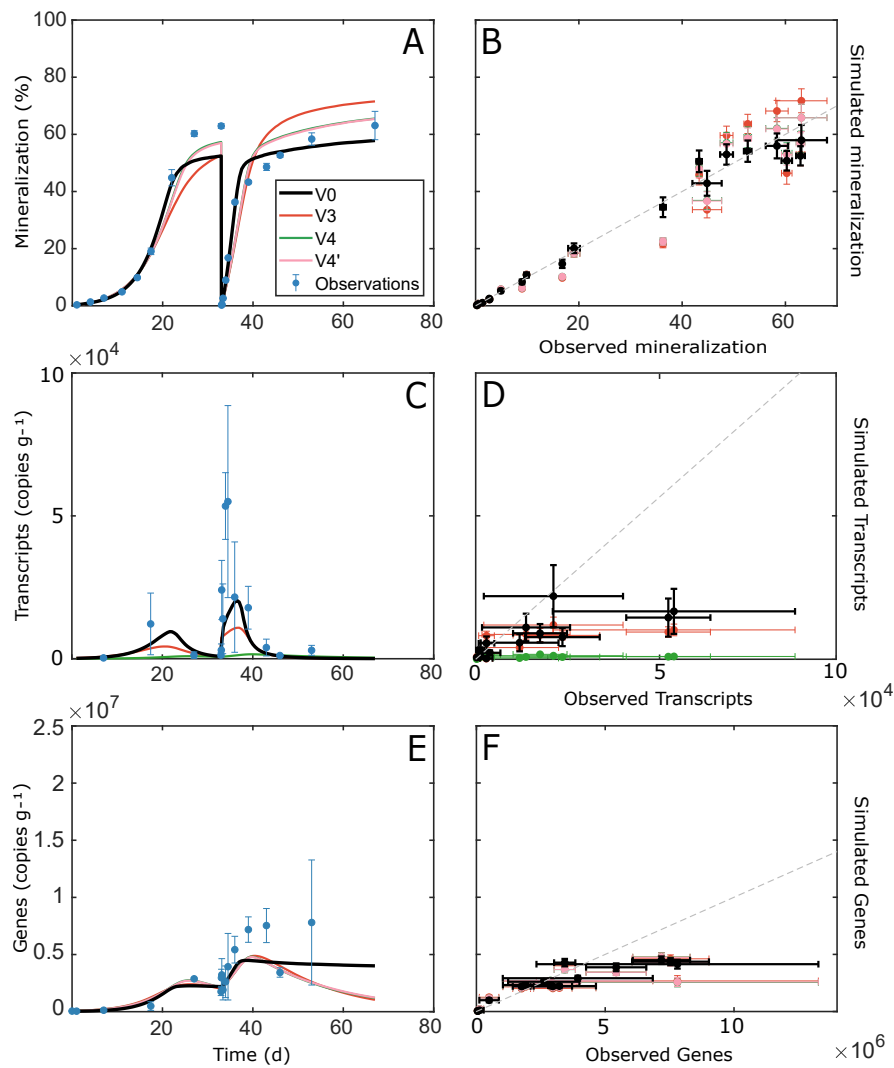


Figure S1: Model calibration on MCPA data. Time series of pesticide mineralization (A, B), *tfdA* mRNA copies (transcripts) per g of soil (C, D), *tfdA* DNA gene copies per g of soil (E, F). Error bars show the standard deviation of the data and of the simulations (based on MCMC ensembles, see Material and Methods 6.4.2)

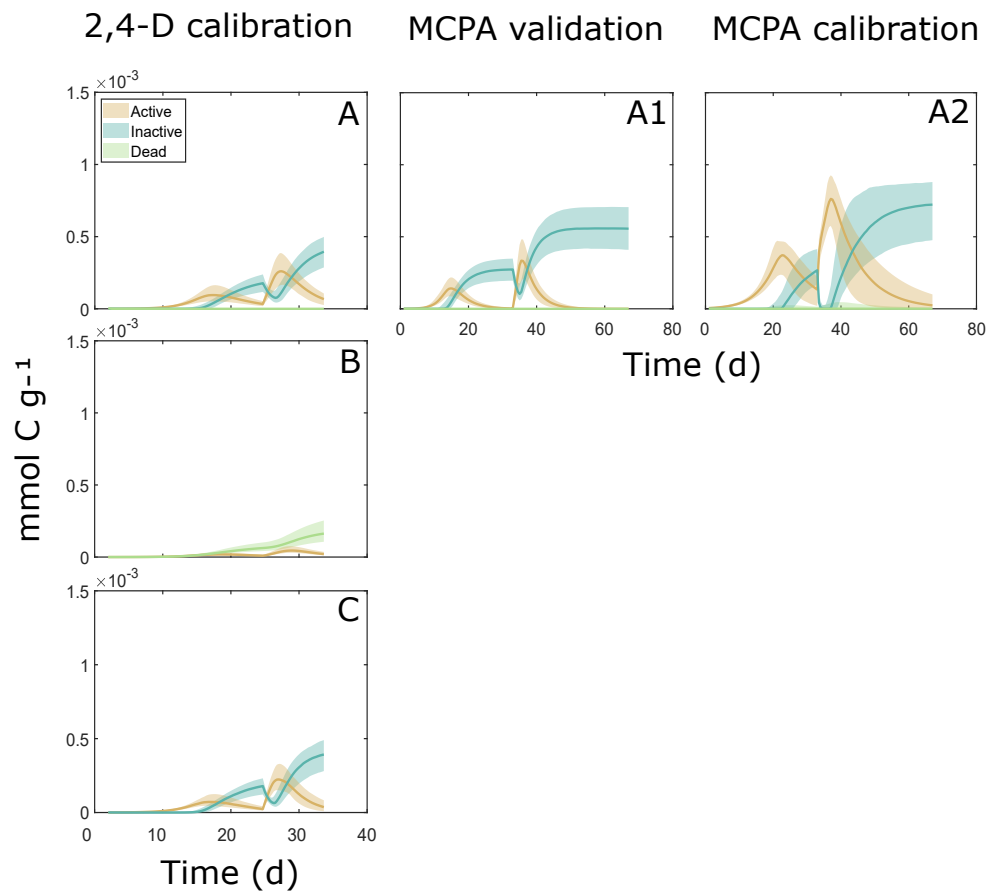
***tfdA* gene dynamics for both pesticides (2,4-D and MCPA)**

Figure S2: *tfdA* gene dynamics for V0 (A, A1, A2), V1 (B), V2 (C). V1 and V2 only include dynamics of 2,4-D calibration dataset. Panels showed the mean value of the corresponding dynamics (bold line) and the 95% confidence interval (based on MCMC ensembles, see Material and Methods 6.4.2)

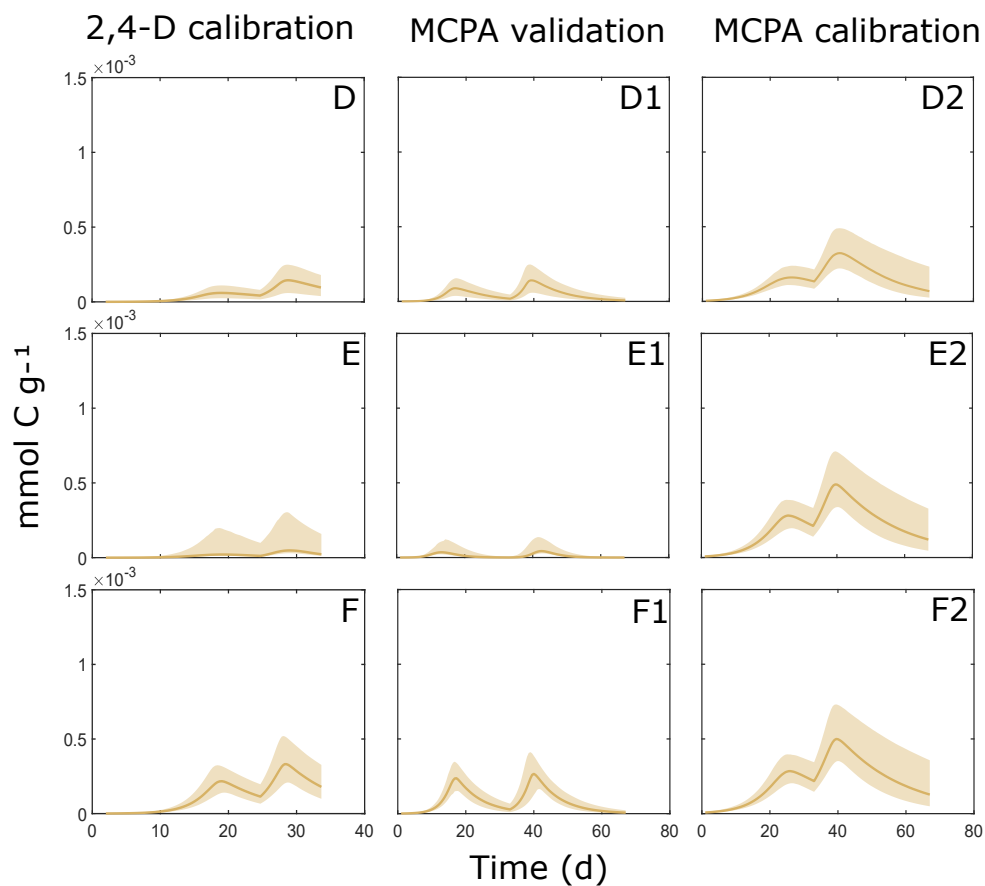


Figure S3: *tfdA* gene dynamics for V3 (D, D1, D2), V4 (E, E1, E2), V4' (F, F1, F2). Panels showed the mean value of the corresponding dynamics (bold line) and the 95% confidence interval (based on MCMC ensembles, see Material and Methods 6.4.2).

### 11.1.5 Residual pesticide (2,4-D and MCPA) concentration in soil

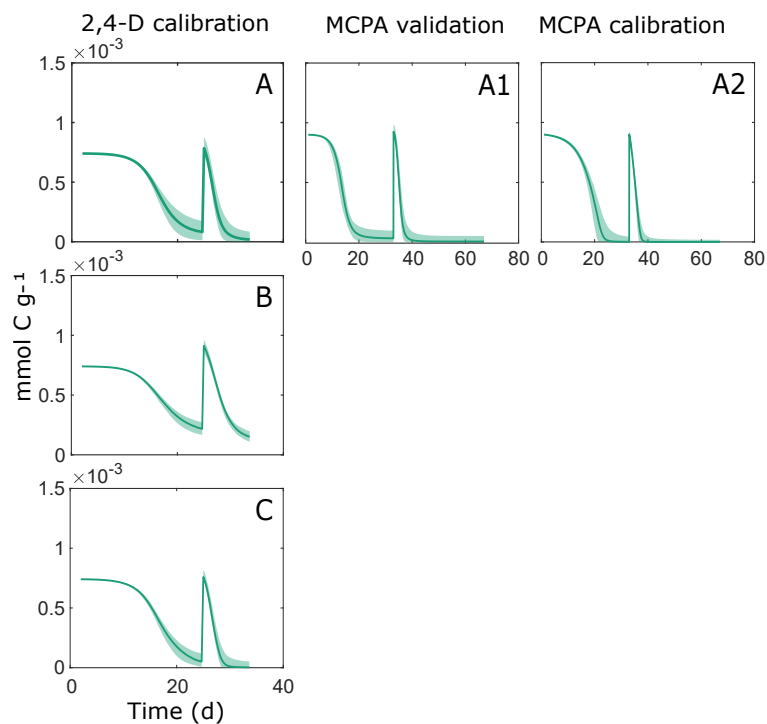


Figure S4: Total residual pesticides in soils for 2,4-D and MCPA for models: V0 (A, A1, A2), V1 (B), V2 (C): Panels showed the mean value of the corresponding dynamics (bold line) and the 95% confidence interval (based on MCMC ensembles, see Material and Methods 6.4.2)



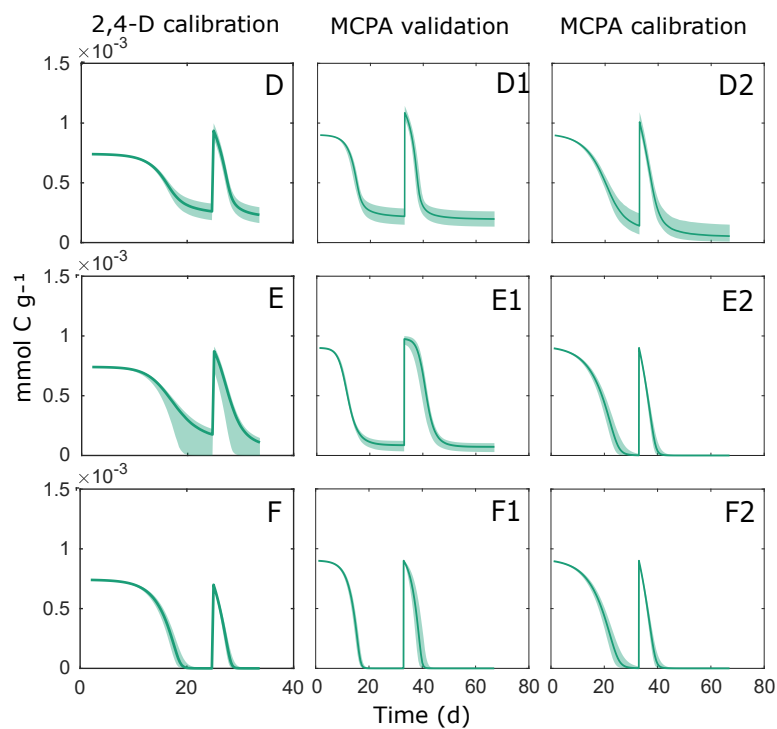


Figure S5: Total residual pesticides in soils for 2,4-D and MCPA for models: V3 (D, D1, D2), V4 (E, E1, E2), V4' (F, F1, F2). Panels showed the mean value of the corresponding dynamics (bold line) and the 95% confidence interval (based on MCMC ensembles, see Material and Methods 6.4.2)

### 11.1.6 Local Sensitivity Analysis of reduced model variants

Table S2: Uncertainty analysis for V1 for 2,4 D data.

Parameter	Best Fit	SC	IS	PE	MV	SD
$a_r$	0.9	$1.1 \cdot 10^3$	$1.1 \cdot 10^3$	32.6	0.4	0.2
$\mu_{max}$	0.6	477.0	60.2	65.2	101.9	45.4
$n_{FP}$	1.0	94.0	14.9	$1.2 \cdot 10^3$	0.9	0.1
$Y_P$	0.4	49.8	31.3	321.6	0.1	0.04
$f_1$	$7.9 \cdot 10^{-11}$	39.4	1.4	376.7	$3.2 \cdot 10^{-11}$	1.4
$K_G$	$9.3 \cdot 10^{-4}$	32.0	0.02	$3.7 \cdot 10^3$	0.05	42.4
$a_a$	0.1	27.6	4.8	54.9	0.3	1.2
$n_H$	4.8	16.0	0.9	184.1	1.6	0.7
$f_T$	0.02	13.0	6.4	271.3	4.0	51.1
$K_M$	$2.3 \cdot 10^{-4}$	7.3	$3.9 \cdot 10^{-3}$	$4.0 \cdot 10^3$	$2.8 \cdot 10^{-6}$	28.4
$K_{FP}$	0.1	2.6	$1.0 \cdot 10^{-5}$	$2.1 \cdot 10^4$	0.1	3.8
$a_{CO_2}$	0.9	1.5	0.4	367.7	0.5	0.2
$a_s$	67.2	$5.2 \cdot 10^{-5}$	$3.4 \cdot 10^{-8}$	$1.6 \cdot 10^5$	$3.1 \cdot 10^{-4}$	7.4

SC = Sensitivity coefficient, IS = Identifiability score, PE = Percentage Error, MV and SD = mean value and standard deviation of the estimation from  $DREAM_{(ZS)}$ . Parameters highlighted in yellow are suggested to be reduced.

Table S3: Uncertainty analysis for V2 for 2,4 D data.

Parameter	Best Fit	SC	IS	PE	MV	SD
$n_{FP}$	0.8	$1.9 \cdot 10^3$	$8.6 \cdot 10^3$	87.2	0.9	0.06
$\mu_{max}$	0.5	514.6	228.2	35.9	0.5	1.2
$K_G$	$2.2 \cdot 10^{-4}$	51.1	0.1	$6.0 \cdot 10^3$	$2.1 \cdot 10^{-7}$	159.5
$K_{FP}$	0.4	48.8	$4.6 \cdot 10^{-4}$	$6.0 \cdot 10^3$	0.2	4.6
$Y_P$	0.2	44.1	17.5	$1.1 \cdot 10^3$	0.3	0.05
$a_r$	0.1	40.4	6.5	51.6	$4.8 \cdot 10^{-4}$	10.9
$f_I$	$2.4 \cdot 10^{-11}$	40.3	0.0	$1.3 \cdot 10^3$	$8.3 \cdot 10^{-11}$	1.3
$n_H$	5.2	19.5	1.5	192.4	4.7	3
$f_T$	0.02	13.0	7.9	53.9	0.01	1.3
$a_{CO2}$	0.6	0.04	$5.0 \cdot 10^{-3}$	$1.3 \cdot 10^3$	0.5	0.2
$K_M$	$1.2 \cdot 10^{-6}$	$5.5 \cdot 10^{-3}$	$2.7 \cdot 10^{-5}$	$1.3 \cdot 10^4$	$6.3 \cdot 10^{-5}$	10.2
$k_r$	30.7	$2.7 \cdot 10^{-3}$	$1.2 \cdot 10^{-4}$	$1.8 \cdot 10^4$	0.7	1.3
$a_i$	$2.8 \cdot 10^{-5}$	$1.3 \cdot 10^{-11}$	$2.1 \cdot 10^{-14}$	$8.0 \cdot 10^8$	$2.1 \cdot 10^{-5}$	24.9
$C_T$	$5.4 \cdot 10^{-6}$	0.0	0.0	0.0	$3.8 \cdot 10^{-4}$	2.9
$k_d$	0.6	0.0	0.0	0.0	0.4	1.3

SC = Sensitivity coefficient, IS = Identifiability score, PE = Percentage Error, MV and SD = mean value and standard deviation of the estimation from  $DREAM_{(ZS)}$ . Parameters highlighted in yellow are suggested to be reduced.

Table S4: Uncertainty analysis for V3 for 2,4 D data.

Parameter	Best Fit	SC	IS	PE	MV	SD
$n_{FP}$	0.8	$2.0 \cdot 10^3$	$2.0 \cdot 10^3$	104.9	0.9	0.1
$\mu_{max}$	0.5	517.5	231.6	35.8	0.5	1.1
$K_{FP}$	0.5	52.9	$5.6 \cdot 10^{-4}$	$4.6 \cdot 10^3$	0.1	4.0
$K_G$	$1.8 \cdot 10^{-4}$	51.2	0.1	$5.1 \cdot 10^3$	$4.6 \cdot 10^{-4}$	2.3
$Y_P$	0.2	43.9	17.5	968.8	0.2	0.1
$a_a$	0.1	40.6	6.5	54.4	0.1	1.2
$f_1$	$2.3 \cdot 10^{-11}$	40.4	0.02	$1.1 \cdot 10^3$	$3.5 \cdot 10^{-11}$	1.6
$n_H$	5.3	19.5	1.5	139.4	5.7	1.3
$f_T$	0.02	13.0	7.9	28.3	0.02	1.2
$a_{CO_2}$	0.6	0.04	$4.7 \cdot 10^{-3}$	$1.2 \cdot 10^3$	0.5	0.2

SC = Sensitivity coefficient, IS = Identifiability score, PE = Percentage Error, MV and SD = mean value and standard deviation of the estimation from  $DREAM_{(ZS)}$ . Parameters highlighted in yellow are suggested to be reduced.

Table S5: Uncertainty analysis for V4 for 2,4 D data.

Parameter	Best Fit	SC	IS	PE	MV	SD
$n_{FP}$	0.8	$1.2 \cdot 10^4$	$1.2 \cdot 10^4$	113.1	0.9	0.1
$\mu_{max}$	525.4	$1.1 \cdot 10^3$	$8.8 \cdot 10^{-3}$	758.4	23.8	64.3
$K_M$	0.7	987.2	6.9	$1.8 \cdot 10^3$	$9.4 \cdot 10^{-3}$	209.4
$K_{FP}$	0.1	323.5	$1.4 \cdot 10^{-4}$	$3.6 \cdot 10^3$	0.2	3.6
$a_a$	0.4	179.9	36.9	44.0	0.2	1.8
$Y_P$	0.1	42.2	5.6	860.7	0.3	0.2
$f_1$	$9.6 \cdot 10^{-12}$	40.4	0.0	971.2	$4.4 \cdot 10^{-11}$	2.9
$f_T$	$5.8 \cdot 10^{-3}$	13.0	7.6	35.6	$4.8 \cdot 10^{-3}$	1.2
$a_{CO_2}$	0.4	0.0	$1.1 \cdot 10^{-3}$	$1.4 \cdot 10^3$	0.5	0.2

SC = Sensitivity coefficient, IS = Identifiability score, PE = Percentage Error, MV and SD = mean value and standard deviation of the estimation from  $DREAM_{(ZS)}$ . Parameters highlighted in yellow are suggested to be reduced.

Table S6: Uncertainty analysis for V4' for 2,4 D data.

Parameter	Best Fit	SC	IS	PE	MV	SD
$\mu_{max}$	0.5	388.8	388.8	225.9	0.6	1.1
$Y_P$	0.5	49.1	27.5	116.8	0.5	0.1
$n_{FP}$	0.9	27.2	0.2	$1.7 \cdot 10^4$	0.9	0.1
$f_1$	$1.1 \cdot 10^{-10}$	25.1	0.7	215.7	$1.1 \cdot 10^{-10}$	1.4
$a_a$	0.1	22.3	8.9	142.4	0.1	1.2
$K_M$	$1.8 \cdot 10^{-4}$	5.3	$1.4 \cdot 10^{-4}$	$1.1 \cdot 10^5$	$9.0 \cdot 10^{-5}$	3.9
$K_{FP}$	0.1	0.8	$1.6 \cdot 10^{-8}$	$3.9 \cdot 10^5$	0.2	4.0
$a_{CO2}$	0.5	0.4	0.0	320.9	0.4	0.2

SC = Sensitivity coefficient, IS = Identifiability score, PE = Percentage Error, MV and SD = mean value and standard deviation of the estimation from DREAM<sub>(ZS)</sub>. Parameters highlighted in yellow are suggested to be reduced.

### 11.1.7 Global Sensitivity Analysis Results - Morris Method

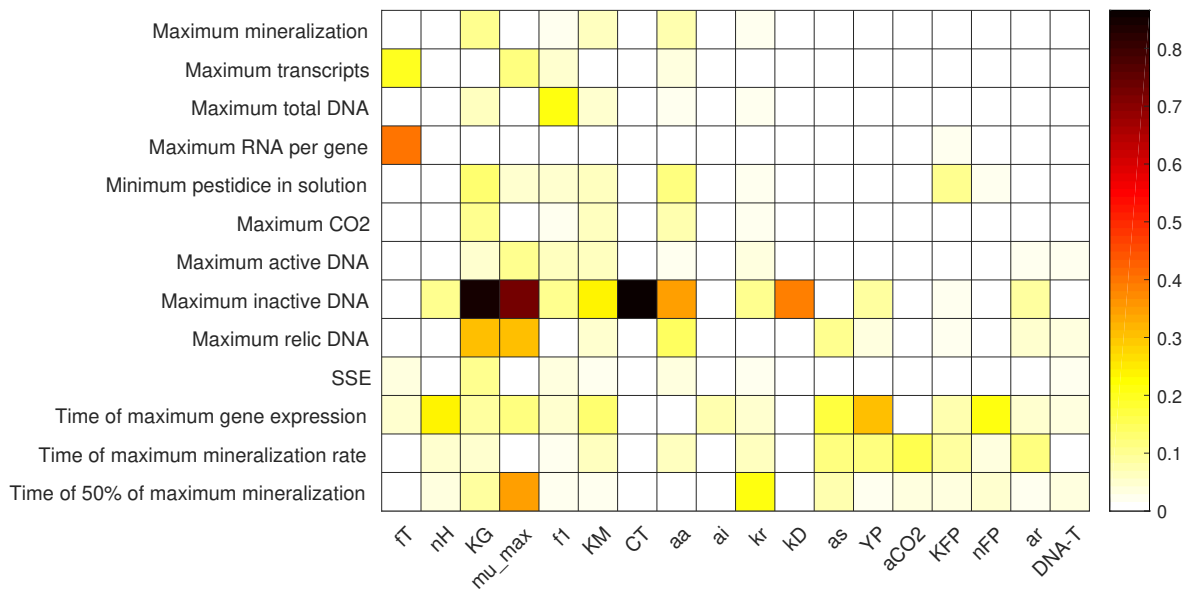


Figure S6: Morris Method results for V0. X axis shows the model parameters and Y axis shows the outputs chosen for the Morris analysis

### 11.1.8 Prior and posterior distribution of the model parameters on 2,4-D and MCPA data

data

Posterior distribution of the parameters of V0 for 2,4-D and MCPA data [117]

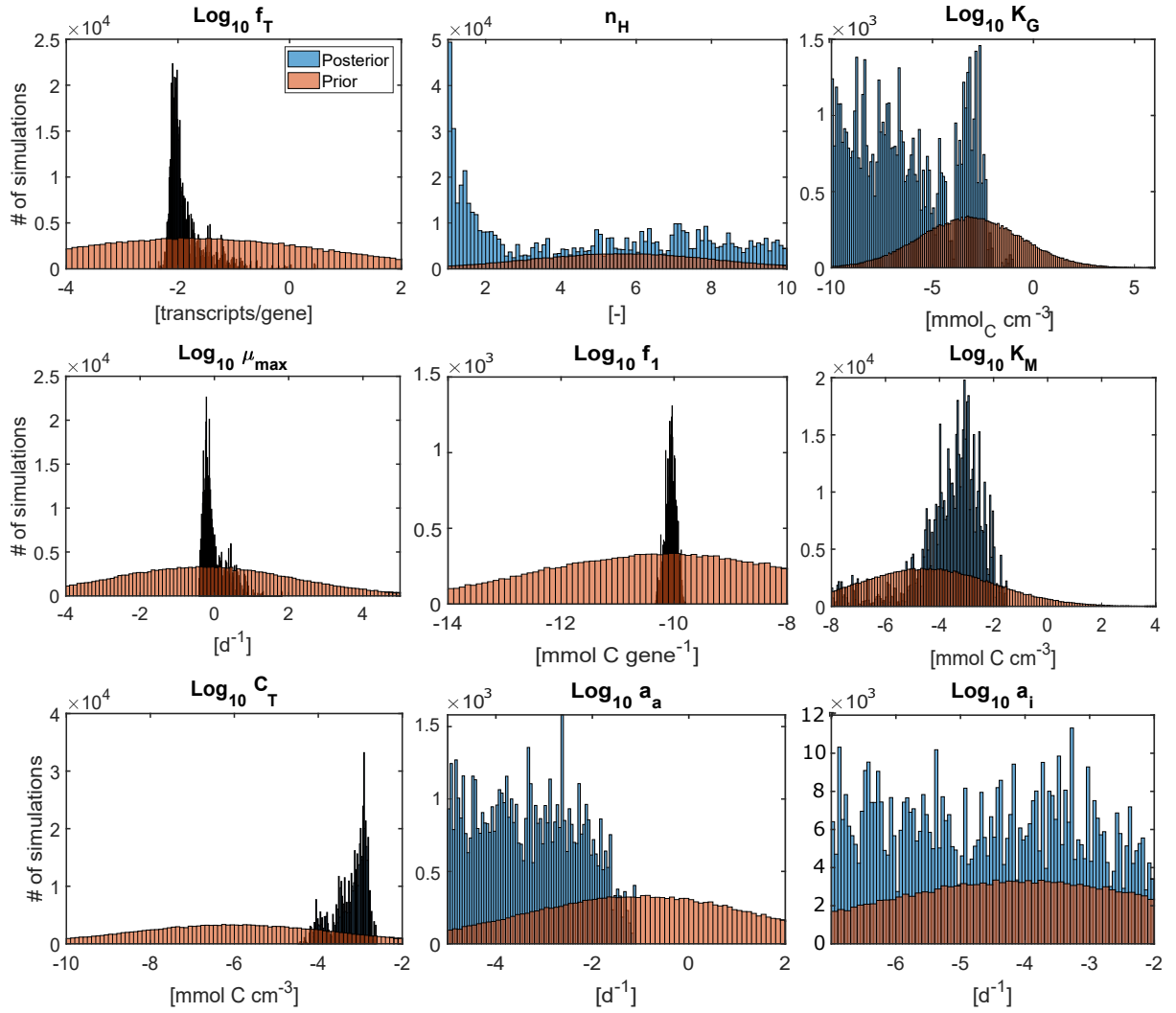


Figure S7: Prior and posterior distribution of parameters (1-9) for V0 with 2,4-D data

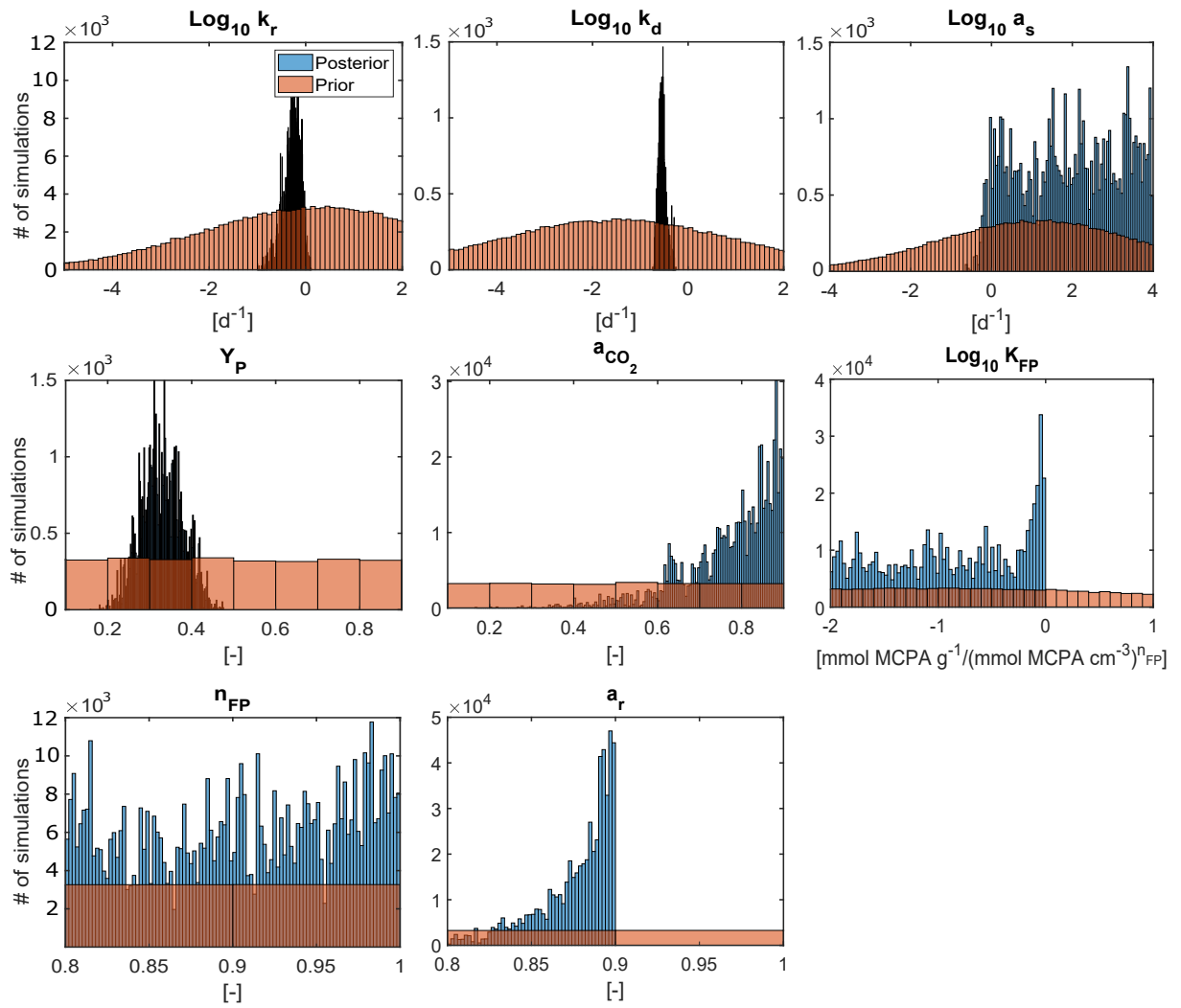


Figure S8: Prior and posterior distribution of parameters (10-17) for V0 with 2,4-D data

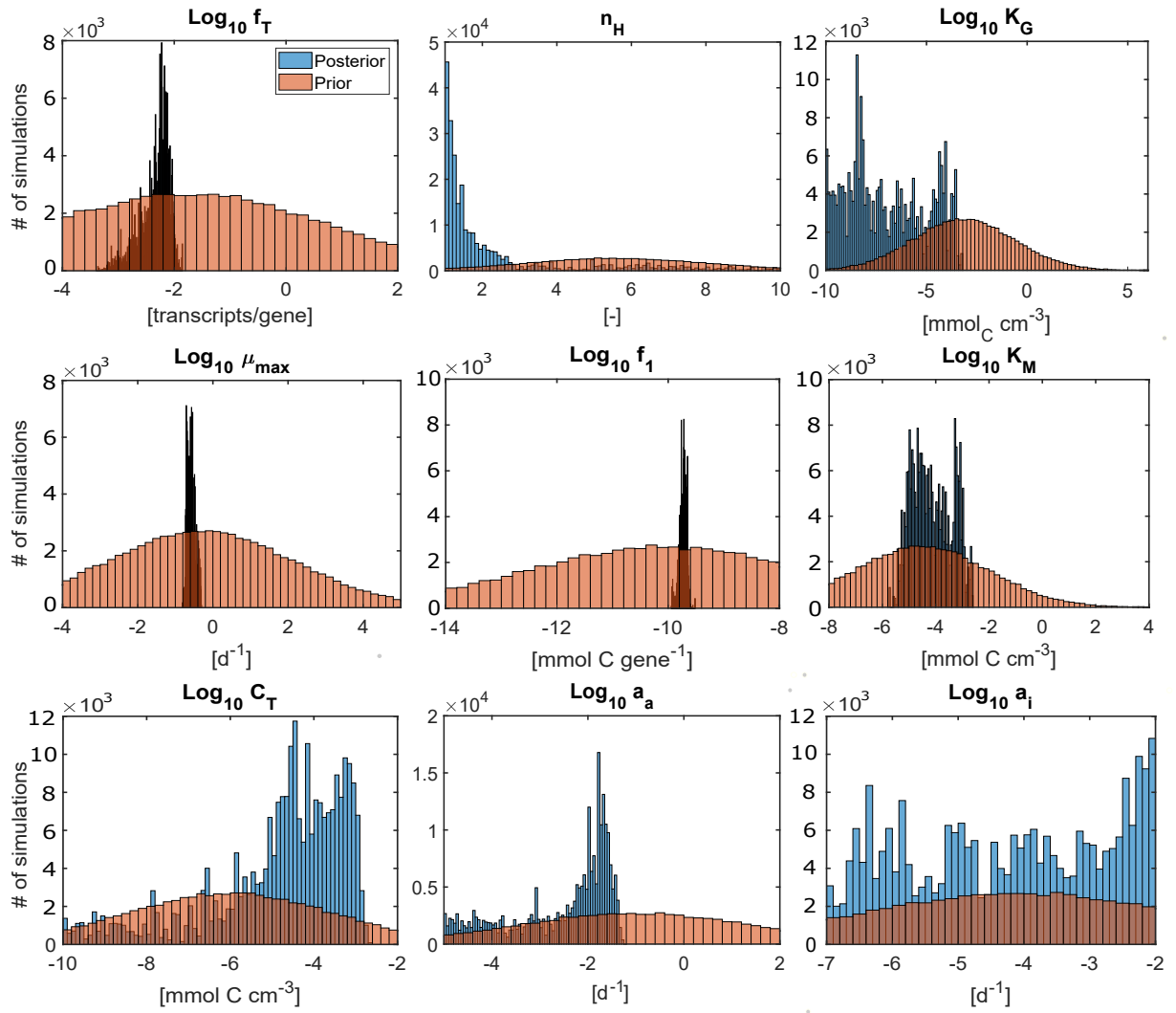


Figure S9: Prior and posterior distribution of parameters (1-9) for V0 with MCPA data



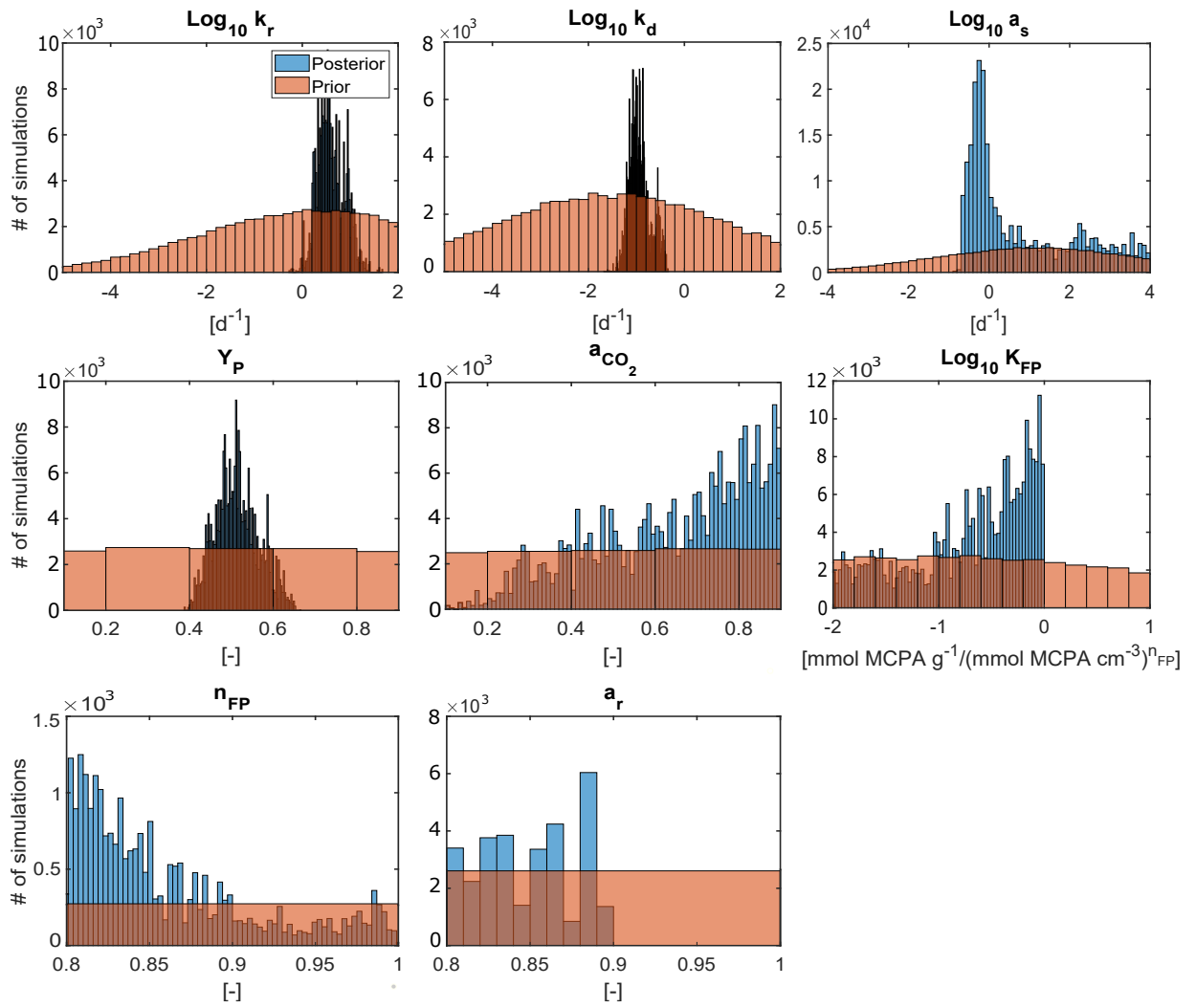


Figure S10: Prior and posterior distribution of parameters (10-17) for V0 with MCPA data

## Posterior distribution of V3 for 2,4-D and MCPA data

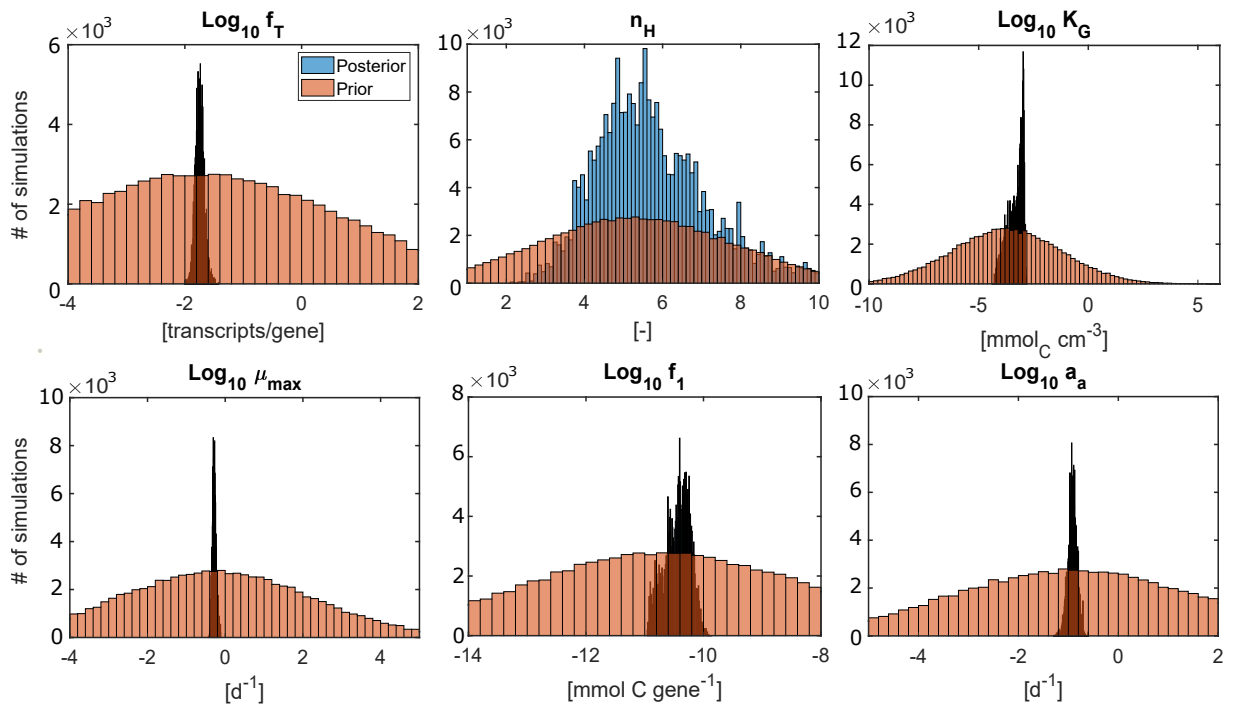


Figure S11: Prior and posterior distribution of parameters (1-6) for V3 with 2,4-D data

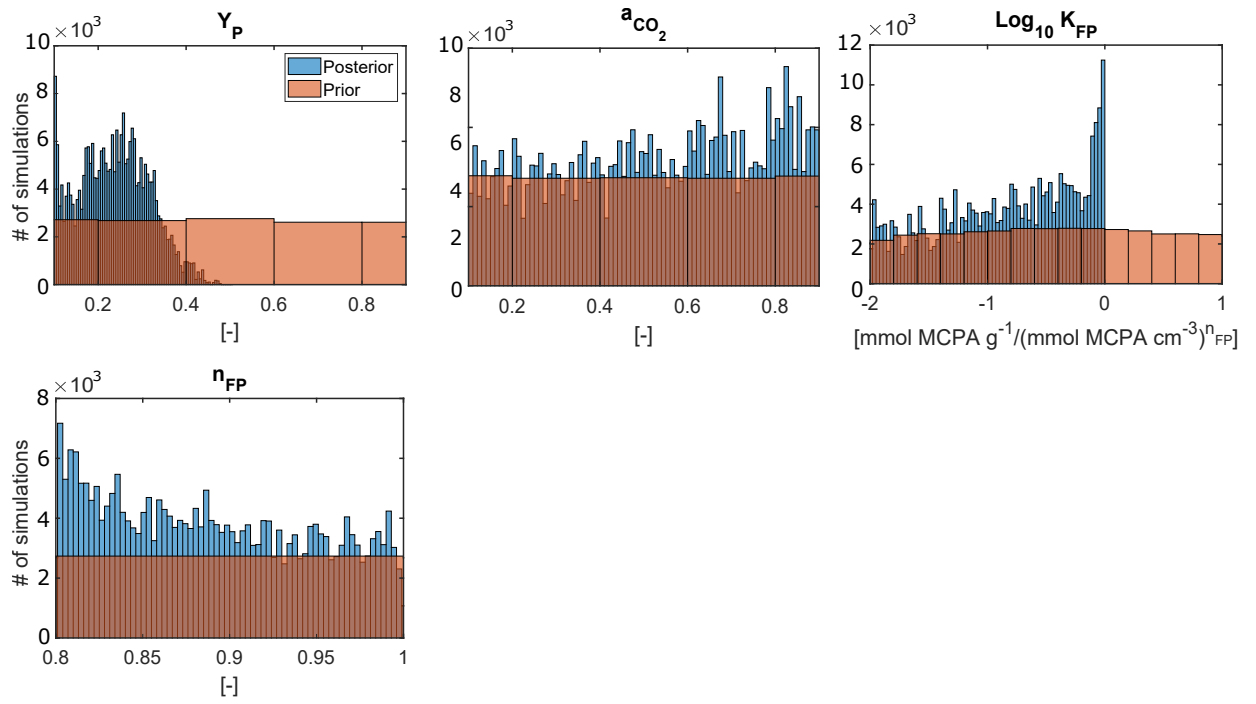


Figure S12: Prior and posterior distribution of parameters (7-10) for V3 with 2,4-D data

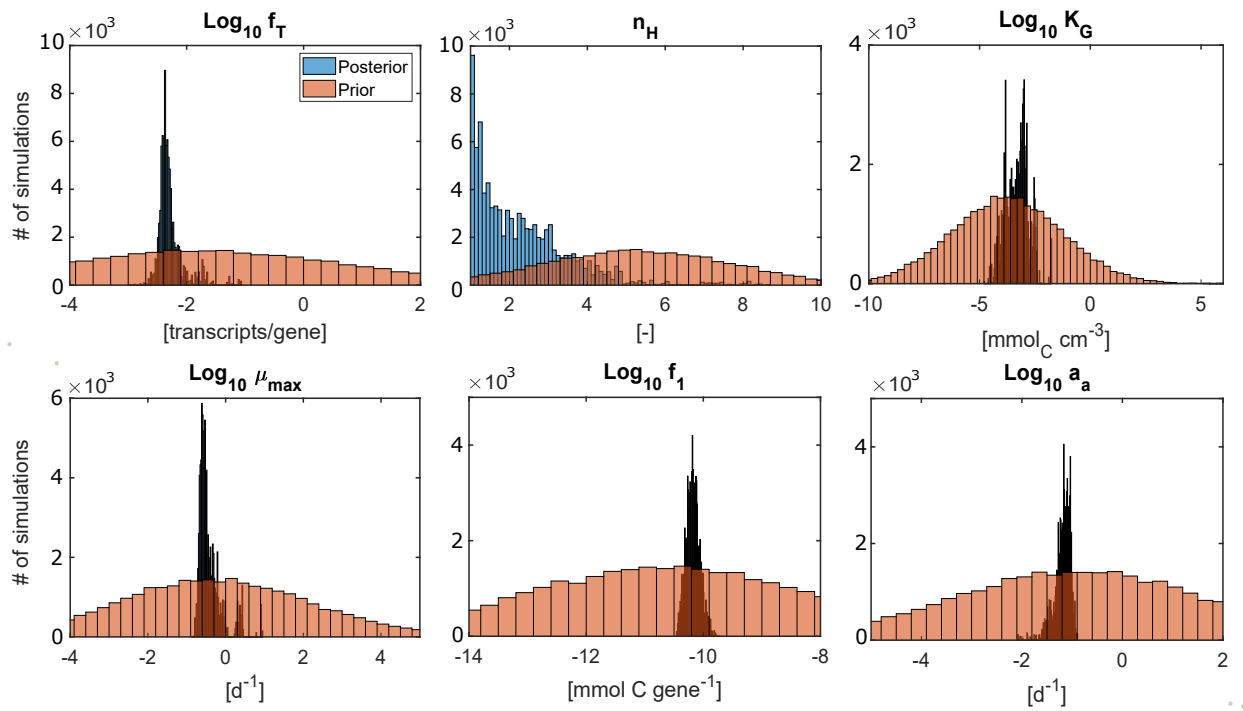


Figure S13: Prior and posterior distribution of parameters (1-6) for V3 with with MCPA data

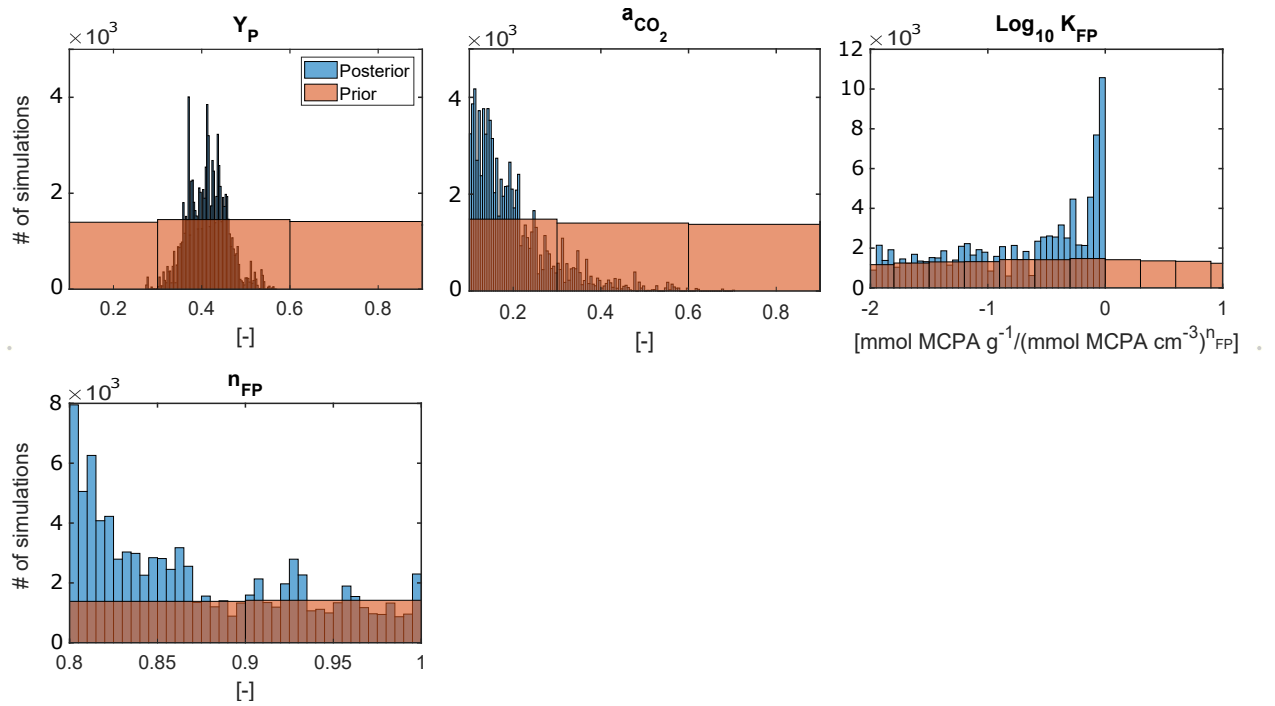


Figure S14: Prior and posterior distribution of parameters (7-10) for V3 with with MCPA data

Posterior distribution of the parameters of V4 for 2,4-D and MCPA data

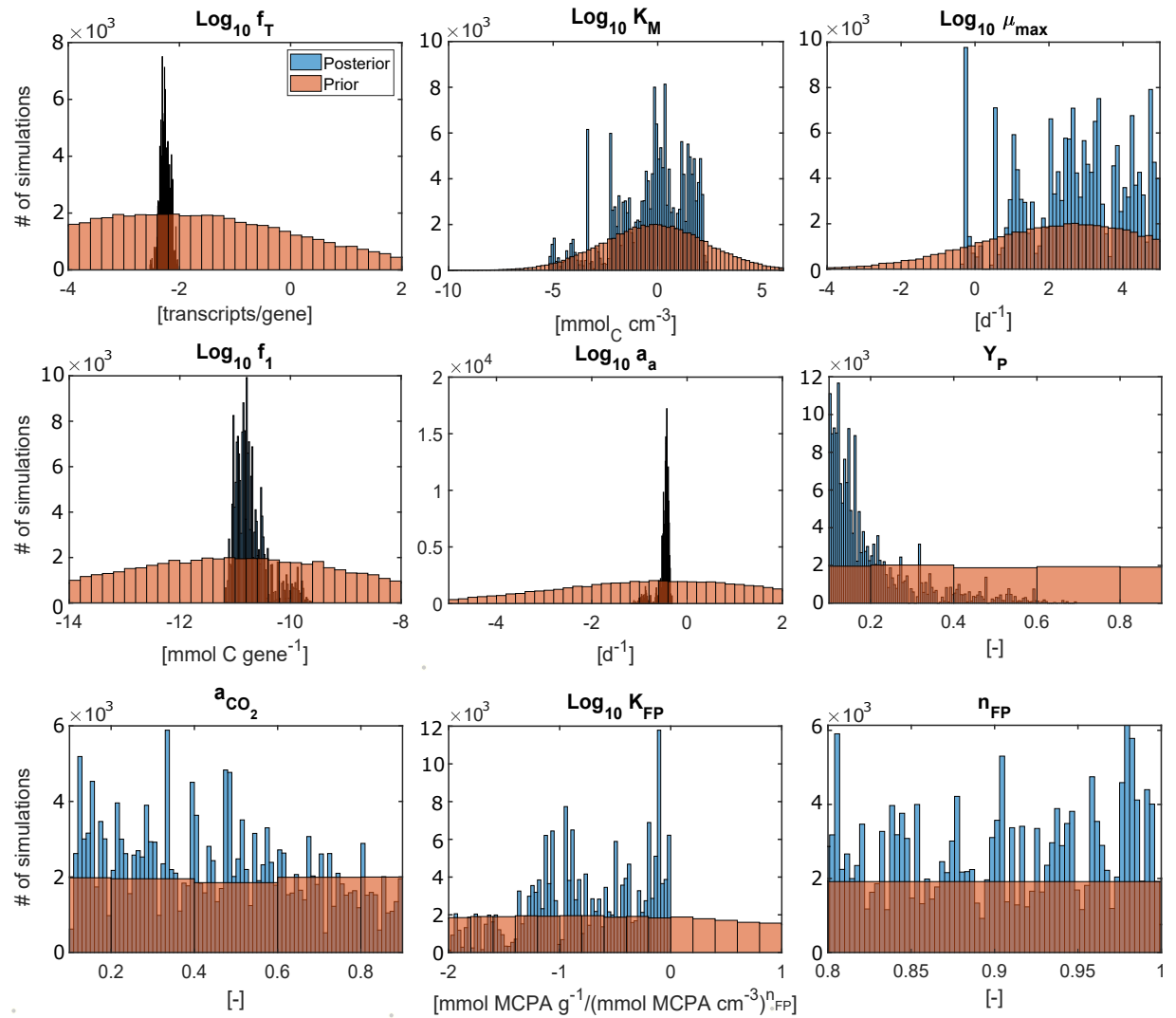


Figure S15: Prior and posterior distribution of parameters for V4 with 2,4-D data

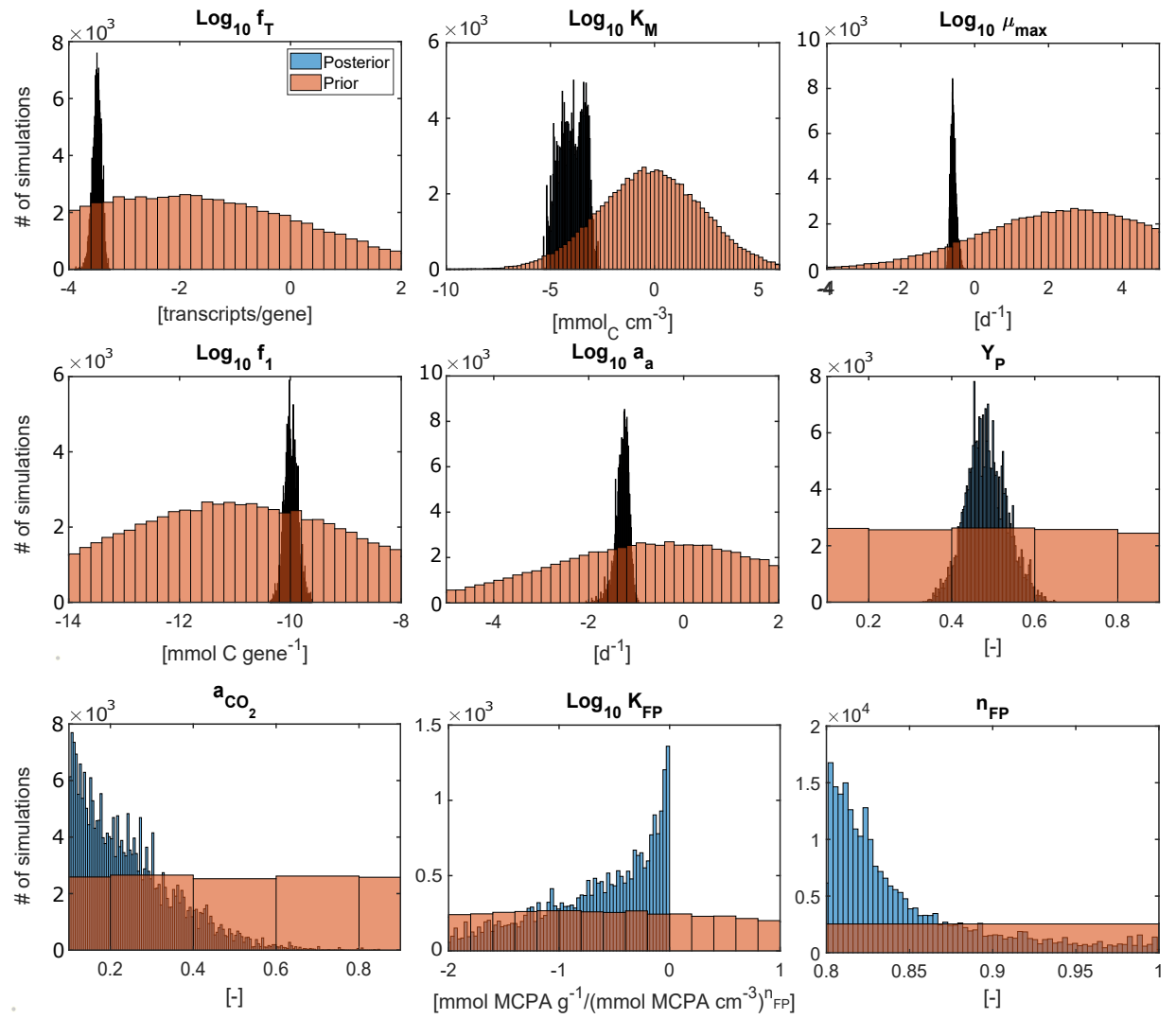


Figure S16: Prior and posterior distribution of parameters for V4 with MCPA data

Posterior distribution of V4' for 2,4-D and MCPA data

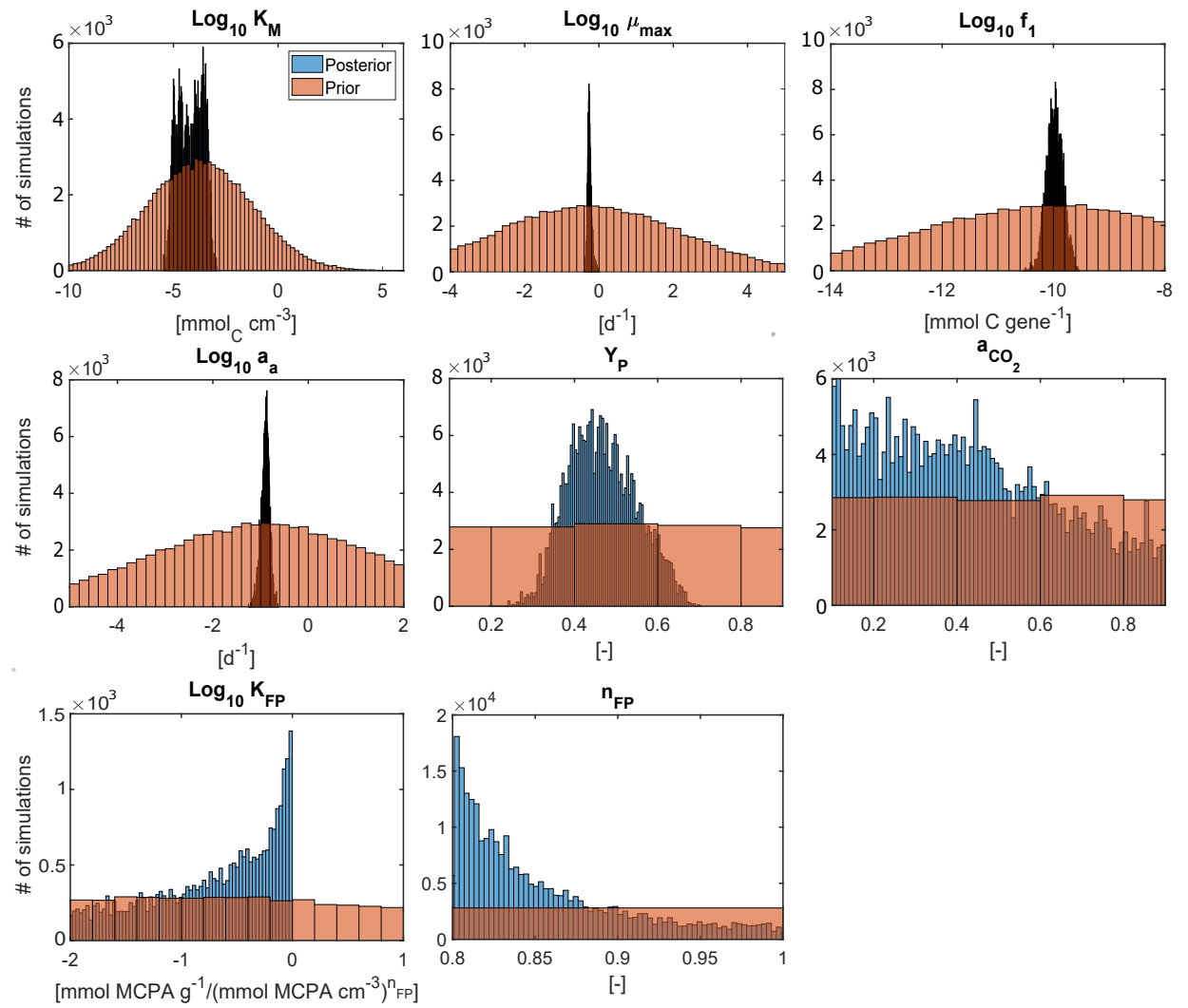


Figure S17: Prior and posterior distribution of parameters for V4' with 2,4-D data



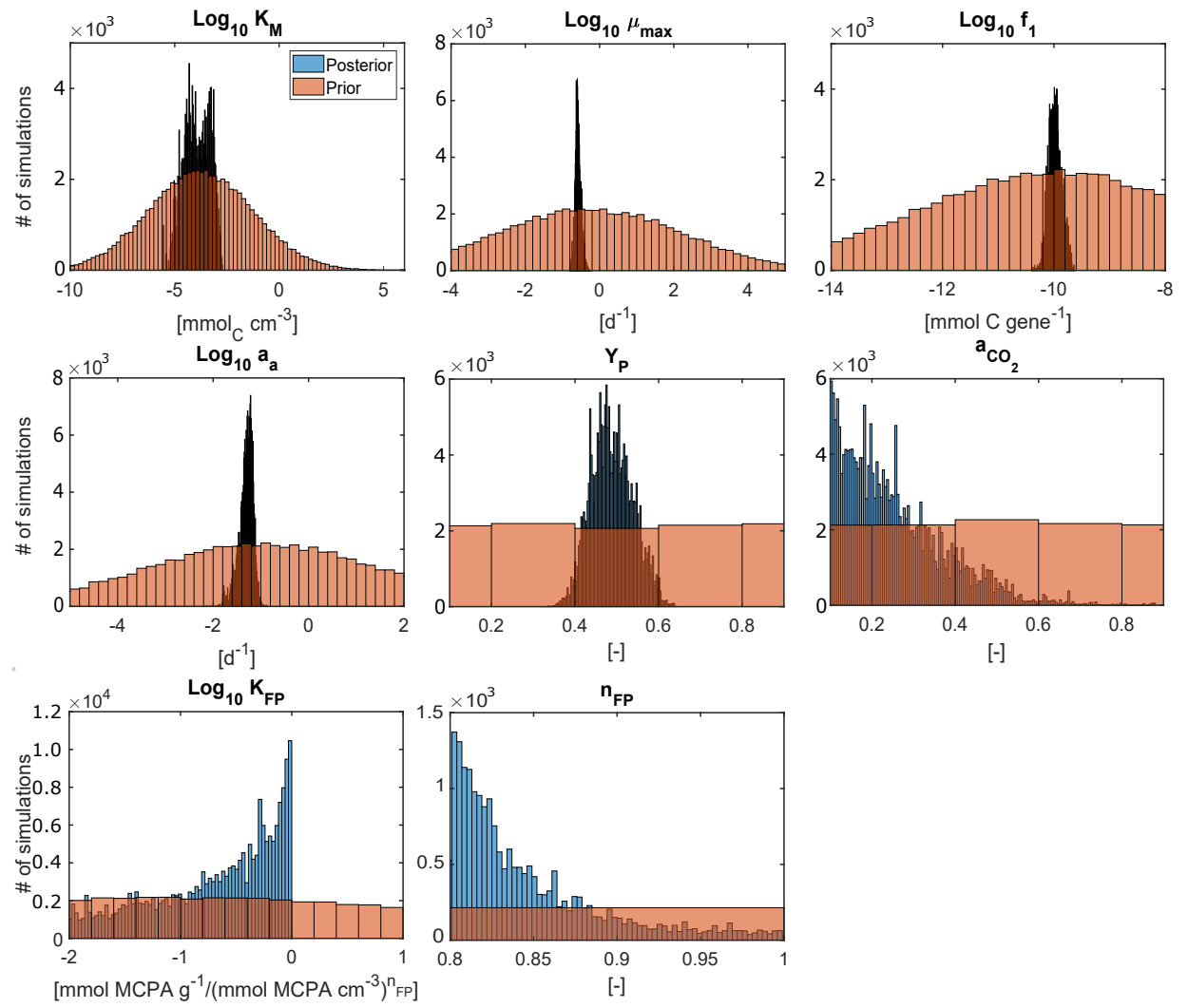


Figure S18: Prior and posterior distribution of parameters for V4' with MCPA data

Posterior distribution of the parameters of V1 for 2,4-D

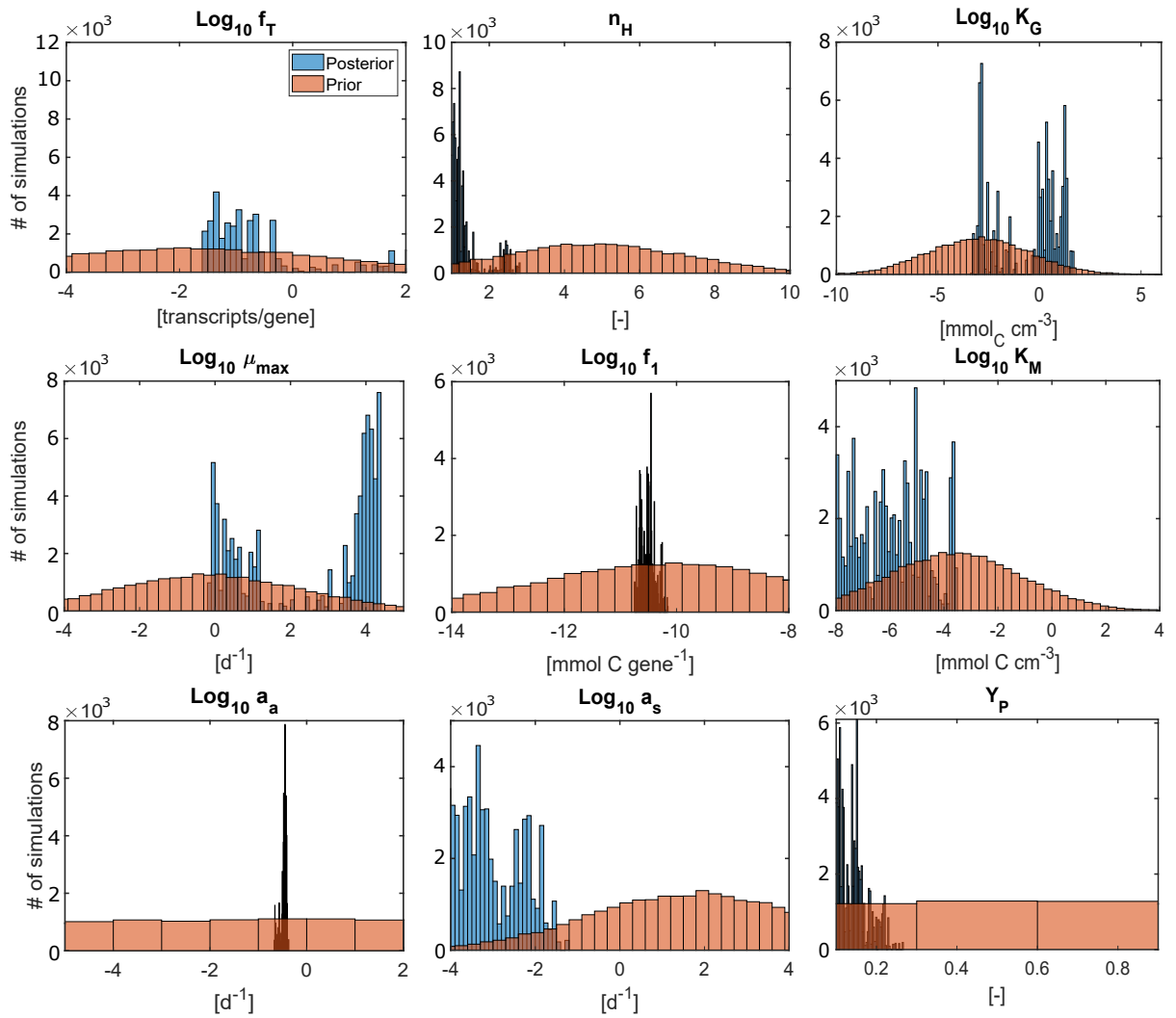


Figure S19: Prior and posterior distribution of parameters (1-9) for V1 with 2,4-D data

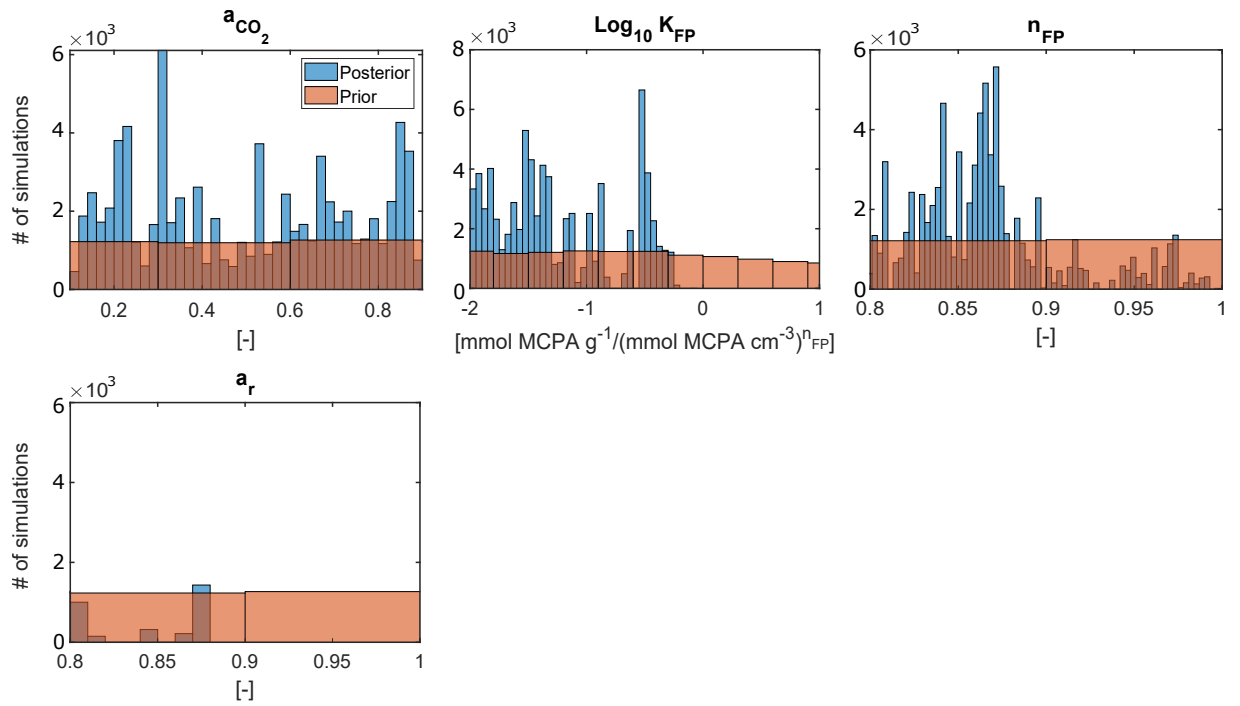


Figure S20: Prior and posterior distribution of parameters (10-13) for V1 with 2,4-D data

Posterior distribution of the parameters of V2 for 2,4-D

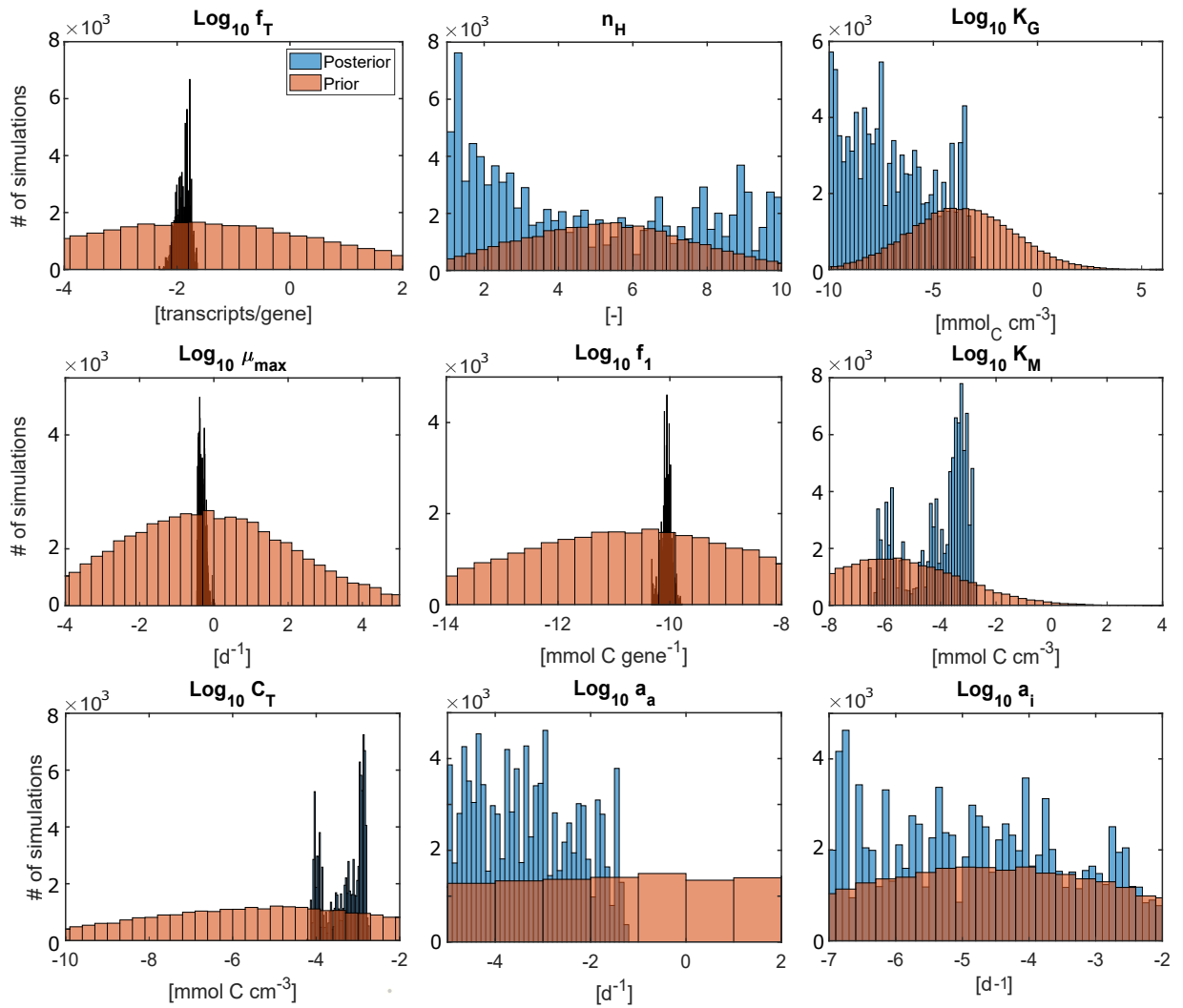


Figure S21: Prior and posterior distribution of parameters (1-9) for V2 with 2,4-D data

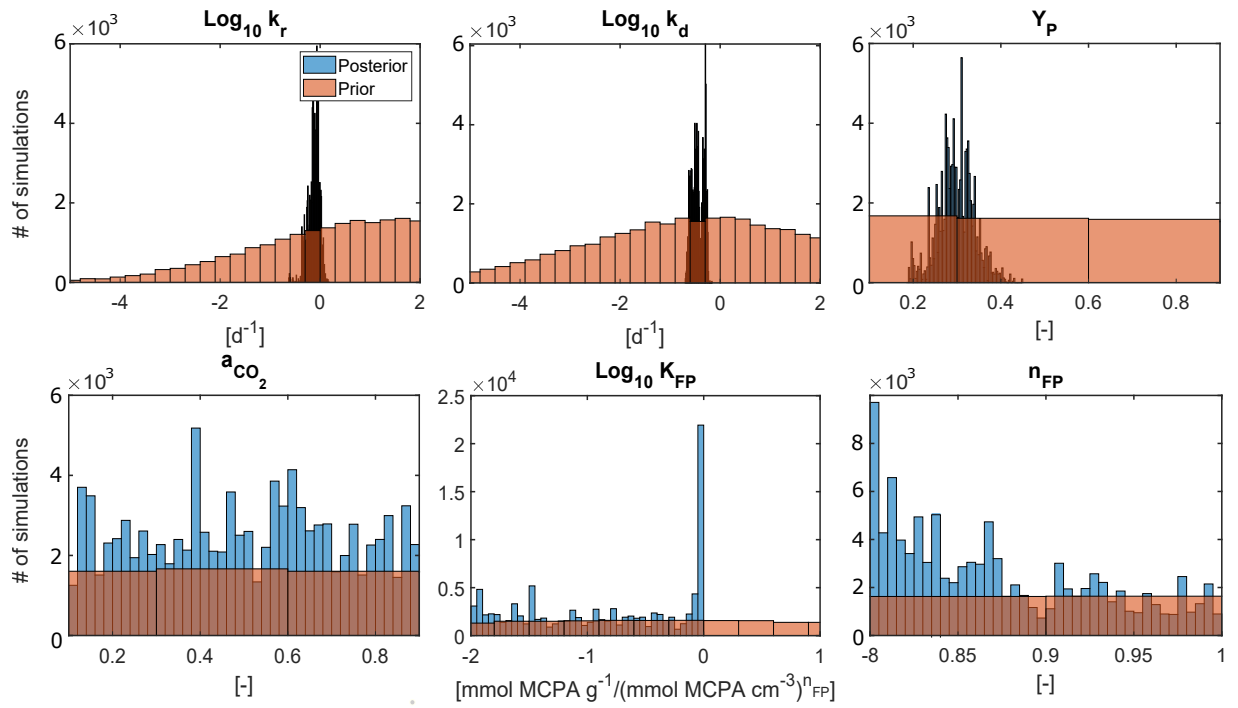


Figure S22: Prior and posterior distribution of parameters (10-15) for V2 with 2,4-D data

11.1.9 Correlation matrix of model variants

	$f_T$	$n_H$	$K_G$	$\mu_{max}$	$f_1$	$K_M$	$C_T$	$a_a$	$a_i$	$k_r$	$k_d$	$a_s$	$Y_P$	$a_{CO_2}$	$K_{FP}$	$n_{FP}$	$a_r$
$f_T$	1.0	-0.6	0.7	0.3	-0.4	-0.7	0.1	-0.1	0.0	-0.4	0.0	0.0	-0.4	0.1	0.1	0.1	0.1
$n_H$	-0.6	1.0	-0.6	0.0	0.3	0.4	-0.1	0.0	0.0	0.3	0.2	0.1	0.2	0.0	0.0	-0.1	-0.1
$K_G$	0.7	-0.6	1.0	0.1	-0.4	-0.5	0.2	0.0	0.0	-0.4	-0.3	-0.1	-0.3	0.1	-0.1	0.2	0.2
$\mu_{max}$	0.3	0.0	0.1	1.0	-0.1	0.2	0.2	0.1	-0.1	-0.5	-0.1	0.1	0.0	0.1	-0.1	0.2	0.3
$f_1$	-0.4	0.3	-0.4	-0.1	1.0	0.4	-0.1	0.1	0.0	0.4	-0.2	0.0	0.9	-0.2	0.0	0.0	0.1
$K_M$	-0.7	0.4	-0.5	0.2	0.4	1.0	0.2	0.1	0.0	0.2	0.0	0.1	0.4	0.0	-0.3	0.0	0.1
$C_T$	0.1	-0.1	0.2	0.2	-0.1	0.2	1.0	0.2	0.0	-0.3	-0.1	-0.1	0.0	0.2	-0.8	0.3	0.3
$a_a$	-0.1	0.0	0.0	0.1	0.1	0.1	0.2	1.0	0.0	0.0	-0.1	-0.1	0.2	0.0	-0.1	0.2	0.1
$a_i$	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
$k_r$	-0.4	0.3	-0.4	-0.5	0.4	0.2	-0.3	0.0	0.0	1.0	0.5	0.0	0.3	-0.1	0.1	-0.3	-0.2
$k_d$	0.0	0.2	-0.3	-0.1	-0.2	0.0	-0.1	-0.1	0.0	0.5	1.0	0.0	-0.2	0.1	0.0	-0.3	-0.2
$a_s$	0.0	0.1	-0.1	0.1	0.0	0.1	-0.1	-0.1	0.0	0.0	0.0	1.0	-0.1	0.0	0.1	0.0	0.0
$Y_P$	-0.4	0.2	-0.3	0.0	0.9	0.4	0.0	0.2	0.0	0.3	-0.2	-0.1	1.0	-0.1	-0.1	0.1	0.2
$a_{CO_2}$	0.1	0.0	0.1	0.1	-0.2	0.0	0.2	0.0	0.0	-0.1	0.1	0.0	-0.1	1.0	-0.1	0.1	0.1
$K_{FP}$	0.1	0.0	-0.1	-0.1	0.0	-0.3	-0.8	-0.1	0.0	0.1	0.0	0.1	-0.1	-0.1	1.0	0.1	-0.3
$n_{FP}$	0.1	-0.1	0.2	0.2	0.0	0.0	0.3	0.2	0.1	-0.3	-0.3	0.0	0.1	0.1	0.1	1.0	0.2
$a_r$	0.1	-0.1	0.2	0.3	0.1	0.1	0.3	0.1	0.0	-0.2	-0.2	0.0	0.2	0.1	-0.3	0.2	1.0

Figure S23: Correlation matrix for full model V0.

	$f_T$	$n_H$	$K_G$	$\mu_{max}$	$f_1$	$K_M$	$a_a$	$a_s$	$Y_P$	$a_{CO_2}$	$K_{FP}$	$n_{FP}$	$a_r$
$f_T$	1.0	-0.6	1.0	1.0	0.0	-0.3	0.6	-0.1	-0.1	-0.1	-0.1	0.0	0.2
$n_H$	-0.6	1.0	-0.6	-0.6	0.1	0.4	-0.8	0.1	0.1	0.0	0.2	0.1	0.1
$K_G$	1.0	-0.6	1.0	1.0	0.1	-0.2	0.6	-0.1	0.0	-0.1	-0.3	0.0	0.2
$\mu_{max}$	1.0	-0.6	1.0	1.0	0.0	-0.3	0.6	-0.1	0.0	-0.1	-0.1	0.0	0.3
$f_1$	0.0	0.1	0.1	0.0	1.0	0.2	-0.5	-0.2	1.0	0.1	-0.2	0.4	0.1
$K_M$	-0.3	0.4	-0.2	-0.3	0.2	1.0	-0.4	0.0	0.2	0.1	0.0	0.1	0.0
$a_a$	0.6	-0.8	0.6	0.6	-0.5	-0.4	1.0	0.0	-0.5	-0.1	-0.2	-0.2	0.1
$a_s$	-0.1	0.1	-0.1	-0.1	-0.2	0.0	0.0	1.0	-0.1	-0.1	0.0	-0.1	0.0
$Y_P$	-0.1	0.1	0.0	0.0	1.0	0.2	-0.5	-0.1	1.0	0.1	-0.1	0.4	0.1
$a_{CO_2}$	-0.1	0.0	-0.1	-0.1	0.1	0.1	-0.1	-0.1	0.1	1.0	-0.1	0.0	-0.1
$K_{FP}$	-0.1	0.2	-0.3	-0.1	-0.2	0.0	-0.2	0.0	-0.1	-0.1	1.0	0.1	-0.2
$n_{FP}$	0.0	0.1	0.0	0.0	0.4	0.1	-0.2	-0.1	0.4	0.0	0.1	1.0	0.1
$a_r$	0.2	0.1	0.2	0.3	0.1	0.0	0.1	0.0	0.1	-0.1	-0.2	0.1	1.0

Figure S24: Correlation matrix for the model variant V1.

	$f_T$	$n_H$	$K_G$	$\mu_{max}$	$f_1$	$K_M$	$C_T$	$a_a$	$a_i$	$k_r$	$k_d$	$Y_P$	$a_{CO_2}$	$K_{FP}$	$n_{FP}$
$f_T$	1.0	-0.2	0.2	-0.5	-0.2	-0.7	-0.4	0.1	-0.1	0.1	0.4	-0.3	-0.1	0.4	-0.4
$n_H$	-0.2	1.0	-0.5	-0.2	0.4	-0.1	-0.1	0.2	0.0	0.3	0.3	0.4	0.1	0.1	0.0
$K_G$	0.2	-0.5	1.0	0.2	-0.4	0.2	0.1	-0.2	0.0	-0.4	-0.4	-0.4	0.0	-0.1	0.1
$\mu_{max}$	-0.5	-0.2	0.2	1.0	-0.2	0.7	0.2	-0.1	0.1	-0.5	-0.7	-0.1	0.2	-0.2	0.1
$f_1$	-0.2	0.4	-0.4	-0.2	1.0	0.2	0.3	0.0	0.0	0.6	0.2	1.0	-0.1	-0.2	0.3
$K_M$	-0.7	-0.1	0.2	0.7	0.2	1.0	0.8	-0.2	0.1	-0.2	-0.6	0.1	0.1	-0.7	0.5
$C_T$	-0.4	-0.1	0.1	0.2	0.3	0.8	1.0	-0.1	0.0	0.0	-0.3	0.3	-0.1	-0.9	0.6
$a_a$	0.1	0.2	-0.2	-0.1	0.0	-0.2	-0.1	1.0	0.0	0.2	0.2	0.1	0.0	0.1	-0.1
$a_i$	-0.1	0.0	0.0	0.1	0.0	0.1	0.0	0.0	1.0	0.0	-0.1	0.0	0.0	0.0	0.0
$k_r$	0.1	0.3	-0.4	-0.5	0.6	-0.2	0.0	0.2	0.0	1.0	0.6	0.6	-0.1	0.1	0.0
$k_d$	0.4	0.3	-0.4	-0.7	0.2	-0.6	-0.3	0.2	-0.1	0.6	1.0	0.2	-0.1	0.3	-0.2
$Y_P$	-0.3	0.4	-0.4	-0.1	1.0	0.1	0.3	0.1	0.0	0.6	0.2	1.0	-0.1	-0.2	0.2
$a_{CO_2}$	-0.1	0.1	0.0	0.2	-0.1	0.1	-0.1	0.0	0.0	-0.1	-0.1	-0.1	1.0	0.1	0.0
$K_{FP}$	0.4	0.1	-0.1	-0.2	-0.2	-0.7	-0.9	0.1	0.0	0.1	0.3	-0.2	0.1	1.0	-0.3
$n_{FP}$	-0.4	0.0	0.1	0.1	0.3	0.5	0.6	-0.1	0.0	0.0	-0.2	0.2	0.0	-0.3	1.0

Figure S25: Correlation matrix for the model variant V2.

	$f_T$	$n_H$	$K_G$	$\mu_{max}$	$f_1$	$a_a$	$Y_P$	$a_{CO_2}$	$K_{FP}$	$n_{FP}$
$f_T$	1.0	-0.2	0.2	0.3	0.4	-0.1	0.3	0.2	0.0	0.2
$n_H$	-0.2	1.0	-0.1	-0.7	-0.1	-0.7	-0.2	0.0	0.1	-0.2
$K_G$	0.2	-0.1	1.0	0.2	0.4	0.1	0.4	0.1	-0.9	0.5
$\mu_{max}$	0.3	-0.7	0.2	1.0	-0.2	0.9	-0.1	0.0	-0.1	0.1
$f_1$	0.4	-0.1	0.4	-0.2	1.0	-0.4	1.0	0.3	-0.3	0.5
$a_a$	-0.1	-0.7	0.1	0.9	-0.4	1.0	-0.3	-0.1	0.0	0.1
$Y_P$	0.3	-0.2	0.4	-0.1	1.0	-0.3	1.0	0.3	-0.4	0.5
$a_{CO_2}$	0.2	0.0	0.1	0.0	0.3	-0.1	0.3	1.0	0.0	0.0
$K_{FP}$	0.0	0.1	-0.9	-0.1	-0.3	0.0	-0.4	0.0	1.0	-0.2
$n_{FP}$	0.2	-0.2	0.5	0.1	0.5	0.1	0.5	0.0	-0.2	1.0

Figure S26: Correlation matrix for the model variant V3.

	$f_T$	$K_M$	$\mu_{max}$	$f_1$	$a_a$	$Y_P$	$a_{CO_2}$	$K_{FP}$	$\eta_{FP}$
$f_T$	1.0	0.6	0.6	-0.6	0.7	-0.6	0.0	-0.1	0.3
$K_M$	0.6	1.0	1.0	-0.8	0.9	-0.8	-0.1	-0.3	0.5
$\mu_{max}$	0.6	1.0	1.0	-0.8	0.8	-0.8	-0.1	-0.2	0.5
$f_1$	-0.6	-0.8	-0.8	1.0	-1.0	1.0	0.2	0.1	-0.5
$a_a$	0.7	0.9	0.8	-1.0	1.0	-1.0	-0.2	-0.1	0.5
$Y_P$	-0.6	-0.8	-0.8	1.0	-1.0	1.0	0.3	0.1	-0.5
$a_{CO_2}$	0.0	-0.1	-0.1	0.2	-0.2	0.3	1.0	0.1	-0.2
$K_{FP}$	-0.1	-0.3	-0.2	0.1	-0.1	0.1	0.1	1.0	0.0
$\eta_{FP}$	0.3	0.5	0.5	-0.5	0.5	-0.5	-0.2	0.0	1.0

Figure S27: Correlation matrix for the model variant V4.

	$K_M$	$\mu_{max}$	$f_1$	$a_a$	$Y_P$	$a_{CO_2}$	$K_{FP}$	$\eta_{FP}$
$K_M$	1.0	0.7	-0.2	0.3	-0.1	-0.1	-0.8	0.6
$\mu_{max}$	0.7	1.0	-0.5	0.6	-0.4	-0.3	-0.3	0.6
$f_1$	-0.2	-0.5	1.0	-0.9	0.9	0.7	0.1	-0.2
$a_a$	0.3	0.6	-0.9	1.0	-0.9	-0.6	-0.1	0.3
$Y_P$	-0.1	-0.4	0.9	-0.9	1.0	0.9	0.1	-0.2
$a_{CO_2}$	-0.1	-0.3	0.7	-0.6	0.9	1.0	0.0	-0.1
$K_{FP}$	-0.8	-0.3	0.1	-0.1	0.1	0.0	1.0	-0.1
$\eta_{FP}$	0.6	0.6	-0.2	0.3	-0.2	-0.1	-0.1	1.0

Figure S28: Correlation matrix for the model variant V4'.



## 11.2 Supplementary Information for Chapter 7 (Paper 2)

### 11.2.1 Illustration of the degradation rate vs. Substrate concentration for both model variants M and T:

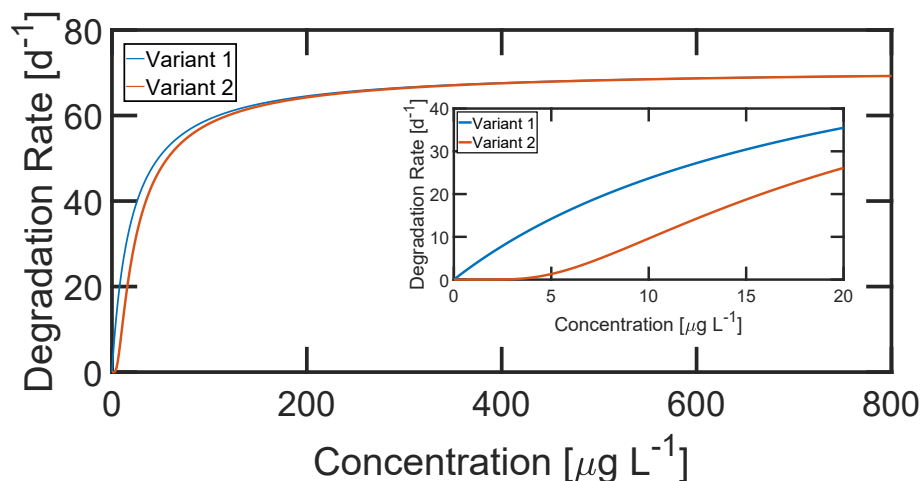


Figure S29: Illustration of the degradation rate vs. Substrate concentration for both model variants M (1) and T (2) at high concentrations and low concentration (inset), for  $k_{HY} = 71 \text{ d}^{-1}$  and  $K_M^{HY} = 20 \mu\text{g L}^{-1}$

### 11.2.2 Determination of correction factor:

In this section, we show the determination of the correction factor for the inner pools. For the atrazine (AT) pool, we started with a mass-based formulation of each pools. We defined  $V_R$  as the total volume of the reactor and  $V_B$  as the total volume of bacteria. Due to the fact that  $V_B$  is negligible compared to  $V_R$ , we assumed that the total volume equals the  $V_R$ . We referred the concentration of cells  $C_B$  to the total volume or  $V_R$ , so that the total mass of cells ( $M_B$ ) is  $M_B = C_B \cdot V_R$ . In turn, the concentration of AT inside the cell are referred to  $V_B$ , so that the mass of AT ( $M_{AT}^i$ ) equals:  $M_{AT}^i = C_{AT}^i \cdot V_B$ . In the same manner, correction factors for the metabolite hydroxyatrazine and the different AT isotopologues can be calculated.

1. First step: Mass-based model formulation of AT:

$$\frac{dM_{AT}^i}{dt} = \underbrace{r_e \cdot M_B \cdot \left( \frac{M_{AT}^o}{V_R} - \frac{M_{AT}^i}{V_B} \right)}_{r_{mass-transfer}} - \underbrace{\frac{k_1 \cdot M_B \cdot \frac{M_{AT}^i}{V_B}}{K_1 + \frac{M_{AT}^i}{V_B}}}_{r_{degradation}^{AT}} \quad (37)$$

2. Second step: Deriving concentration formulation for AT:

$$\frac{dM_{AT}^i}{dt} = r_e \cdot M_B \cdot \underbrace{\left( \frac{M_{AT}^o}{V_R} - \frac{M_{AT}^i}{V_B} \right)}_{r_{mass-transfer}} - \underbrace{\frac{k_1 \cdot M_B \cdot \frac{M_{AT}^i}{V_B}}{K_1 + \frac{M_{AT}^i}{V_B}}}_{r_{degradation}^{AT}}$$

$$\frac{d(C_{AT}^i \cdot V_B)}{dt} = \underbrace{r_e \cdot C_B \cdot V_R \cdot (C_{AT}^o - C_{AT}^i)}_{r_{mass-transfer}} - \underbrace{\frac{k_1 \cdot C_B \cdot V_R \cdot C_{AT}^i}{K_1 + C_{AT}^i}}_{r_{degradation}^{AT}} \quad (38)$$

Applying the product rule to the left hand side of S2:

$$V_B \cdot \frac{dC_{AT}^i}{dt} + C_{AT}^i \cdot \frac{dV_B}{dt} = \underbrace{r_e \cdot C_B \cdot V_R \cdot (C_{AT}^o - C_{AT}^i)}_{r_{mass-transfer}} - \underbrace{\frac{k_1 \cdot C_B \cdot V_R \cdot C_{AT}^i}{K_1 + C_{AT}^i}}_{r_{degradation}^{AT}} \quad (39)$$

Solving for  $\frac{dC_{AT}^i}{dt}$ :

$$V_B \cdot \frac{dC_{AT}^i}{dt} = \underbrace{r_e \cdot C_B \cdot V_R \cdot (C_{AT}^o - C_{AT}^i)}_{r_{mass-transfer}} - \underbrace{\frac{k_1 \cdot C_B \cdot V_R \cdot C_{AT}^i}{K_1 + C_{AT}^i}}_{r_{degradation}^{AT}} - C_{AT}^i \cdot \frac{dV_B}{dt} \quad (40)$$

Dividing by  $V_B$ :

$$\frac{dC_{AT}^i}{dt} = \underbrace{\frac{r_e \cdot C_B \cdot V_R \cdot (C_{AT}^o - C_{AT}^i)}{V_B}}_{r_{mass-transfer}} - \underbrace{\frac{k_1 \cdot C_B \cdot V_R \cdot C_{AT}^i}{(K_1 + C_{AT}^i) \cdot V_B}}_{r_{degradation}^{AT}} - \underbrace{\frac{C_{AT}^i}{V_B} \cdot \frac{dV_B}{dt}}_{r_{correction}^{AT}} \quad (41)$$

Replacing:  $V_B = \frac{V_B^u \cdot C_B \cdot V_R}{f_{cell}}$ , and  $\frac{dV_B}{dt} = \frac{V_B^u \cdot V_R}{f_{cell}} \cdot \frac{dC_B}{dt}$ :

$$\frac{dC_{AT}^i}{dt} = \underbrace{\frac{r_e \cdot C_B \cdot V_R \cdot (C_{AT}^o - C_{AT}^i)}{\frac{V_B^u \cdot C_B \cdot V_R}{f_{cell}}}}_{r_{mass-transfer}} - \underbrace{\frac{k_1 \cdot C_B \cdot V_R \cdot C_{AT}^i}{(K_1 + C_{AT}^i) \cdot \frac{V_B^u \cdot C_B \cdot V_R}{f_{cell}}}}_{r_{degradation}^{AT}} \quad (42)$$

$$- \underbrace{\frac{C_{AT}^i}{\frac{V_B^u \cdot C_B \cdot V_R}{f_{cell}}} \cdot \frac{V_B^u \cdot V_R}{f_{cell}} \cdot \frac{dC_B}{dt}}_{r_{correction}^{AT}} \quad (43)$$

where:  $f_{cell}$  is a conversion factor from cells to carbon,  $V_u$  [L] is the volume of a single bacterium set to  $1 \cdot 10^{-15}$  [63].

3. Third step: Final concentration formulation for AT:

$$\frac{dC_{AT}^i}{dt} = \underbrace{r_e \cdot (C_{AT}^o - C_{AT}^i)}_{r_{mass-transfer}} \cdot \frac{f_{cell}}{V_B^u} - \underbrace{\frac{k_1 \cdot C_{AT}^i}{(K_1 + C_{AT}^i)} \cdot \frac{f_{cell}}{V_B^u}}_{r_{degradation}^{AT}} \quad (44)$$

$$- \underbrace{\frac{C_{AT}^i}{C_B} \cdot \frac{dC_B}{dt}}_{r_{correction}^{AT}} \quad (45)$$

### 11.2.3 Determination of standard deviation of data $\sigma_i^2$ (eq. 74 of the main text)

Since we only had two replicates (two reactors) for each observation in both systems (chemostat and retentostat), as described by [63, 238], we first calculated the mean, standard deviation and coefficient of variation per observation type and dilution rate, using both replicates at steady

state. Because for some dilution rates, only one repetition was measured, we determined a mean coefficient of variation for AT and HY measurements, and one for biomass to recalculate the standard deviation of each dilution rate and observation type (eqs. 36 and 43 in the main paper):

Table S7: Recalculated standard deviations for chemostat and retentostat calibration based on [63, 238]

Observation type	Dilution rate	Mean	Original		Recalculated	
			Std	Cv	Std	Cv
AT [ $\mu\text{g L}^{-1}$ ]	C1 = 0.023	79.3	1.9	0.02	3.3	0.04
	C2 = 0.032	96.0	0.0	0.00	3.9	0.04
	C3 = 0.048	235.5	4.9	0.02	9.7	0.04
	C4 = 0.056	295.0			12.1	0.04
	C5 = 0.068	446.0	5.7	0.01	18.3	0.04
HY [ $\mu\text{g L}^{-1}$ ]	C1 = 0.023	241.0	7.07	0.03	9.9	0.04
	C2 = 0.032	352.0	38.2	0.11	14.4	0.04
	C3 = 0.048	414.5	24.7	0.06	17.0	0.04
	C4 = 0.056	490.0			20.1	0.04
	C5 = 0.068	714.0	35.4	0.05	29.3	0.04
Cells [ $\text{cell L}^{-1}$ ]	C1 = 0.023	$2.3 \cdot 10^{10}$	$2.2 \cdot 10^9$	0.10	$1.6 \cdot 10^9$	0.04
	C2 = 0.032	$2.4 \cdot 10^{10}$	$2.3 \cdot 10^9$	0.10	$1.7 \cdot 10^9$	0.07
	C3 = 0.048	$2.6 \cdot 10^{10}$	$7.1 \cdot 10^7$	0.00	$1.8 \cdot 10^9$	0.07
	C4 = 0.048	$2.7 \cdot 10^{10}$	$8.5 \cdot 10^8$	0.03	$1.9 \cdot 10^9$	0.07
	C5 = 0.068	$2.1 \cdot 10^{10}$	$8.5 \cdot 10^8$	0.04	$1.5 \cdot 10^9$	0.07
AT [ $\mu\text{g L}^{-1}$ ]	R1 = 0.02	12.9	0.6	0.05	0.5	0.04
HY [ $\mu\text{g L}^{-1}$ ]	R1 = 0.02	12.5	0.7	0.06	0.5	0.04
Cells [ $\text{cell L}^{-1}$ ]	R1 = 0.02	$2.5 \cdot 10^{11}$	$3.9 \cdot 10^{10}$	0.16	$1.8 \cdot 10^{10}$	0.07

Std = Standard deviation, C.V = coefficient of variation defined by the standard deviation divided by the mean. Blank cells means that one one repetition was taken

**11.2.4 Soil Observations in Poltringen and Tailfingen, Germany****Pesticide Inventory**

Table S8: Pesticide Inventory - Atrazine and Hydroxyatrazine in soils

Sample	Atrazine [ $\mu\text{g kg}^{-1}$ ]	Atrazine-2-hydroxy [ $\mu\text{g kg}^{-1}$ ]
P 0- 30 a	0.3	2.2
P 0- 30 b	0.5	1.5
P 0- 30 c	0.2	2.7
Mean	0.3	2.1
T 0- 30 a	0.7	1.7
T 0- 30 b	0.6	2.0
T 0- 30 c	0.6	2.1
Mean	0.6	1.9

## Sorption test results

Table S9: Sorption test results

AT conc. [ $\mu\text{g/L}$ ]	Poltringen		Tailfingen	
	Cw* [ $\mu\text{g/L}$ ]	Cs* [ $\mu\text{g/Kg}$ ]	Cw* [ $\mu\text{g/L}$ ]	Cs* [ $\mu\text{g/Kg}$ ]
0.06	0.027	0.43	0.033	0.67
	0.026	0.43	0.032	0.67
	0.028	0.43	0.030	0.68
0.4	0.260	0.66	0.269	0.89
	0.268	0.65	0.258	0.91
	0.299	0.58	0.268	0.89
4.0	2.344	4.18	3.881	1.34
	2.387	4.06	2.537	4.03
	2.452	3.95	2.292	4.52
36.0	22.688	27.57	21.109	30.61
	22.686	27.26	22.237	28.27
	22.575	27.48	28.575	15.78
420.0	246.709	346.73	214.333	407.25
	228.472	379.72	276.396	285.26
	370.428		313.689	209.25
2060.0	2364.020		1225.239	1680.72
	1631.287	900.84	1445.130	1241.42
	1274.485	1590.91	1353.376	1418.81

\* equilibrium concentrations after 13 days. AT = atrazine

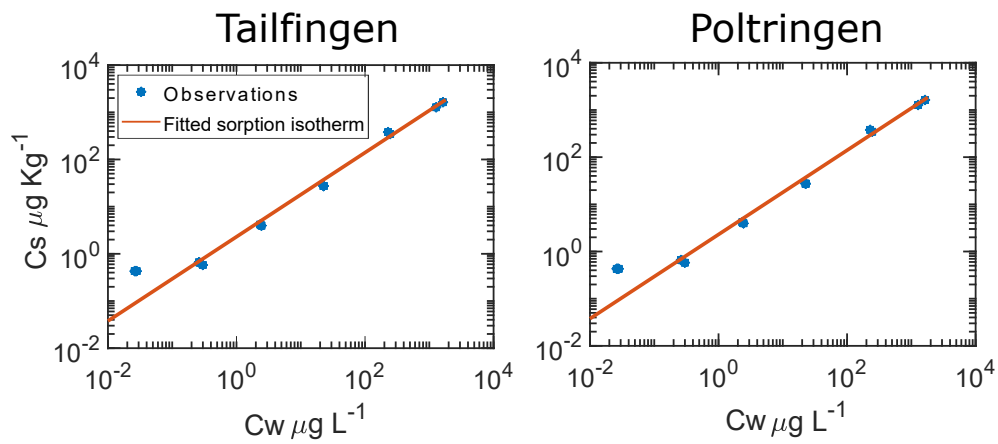
**Determination of sorption parameters**

Figure S30: Sorption fitting for both study sites

**11.2.5 Seepage pesticide (atrazine and hydroxyatrazine) concentration results**

Concentrations of atrazine and hydroxyatrazine at the two location were on average 0.1 ng L<sup>-1</sup> and 10 ng L<sup>-1</sup>, respectively.

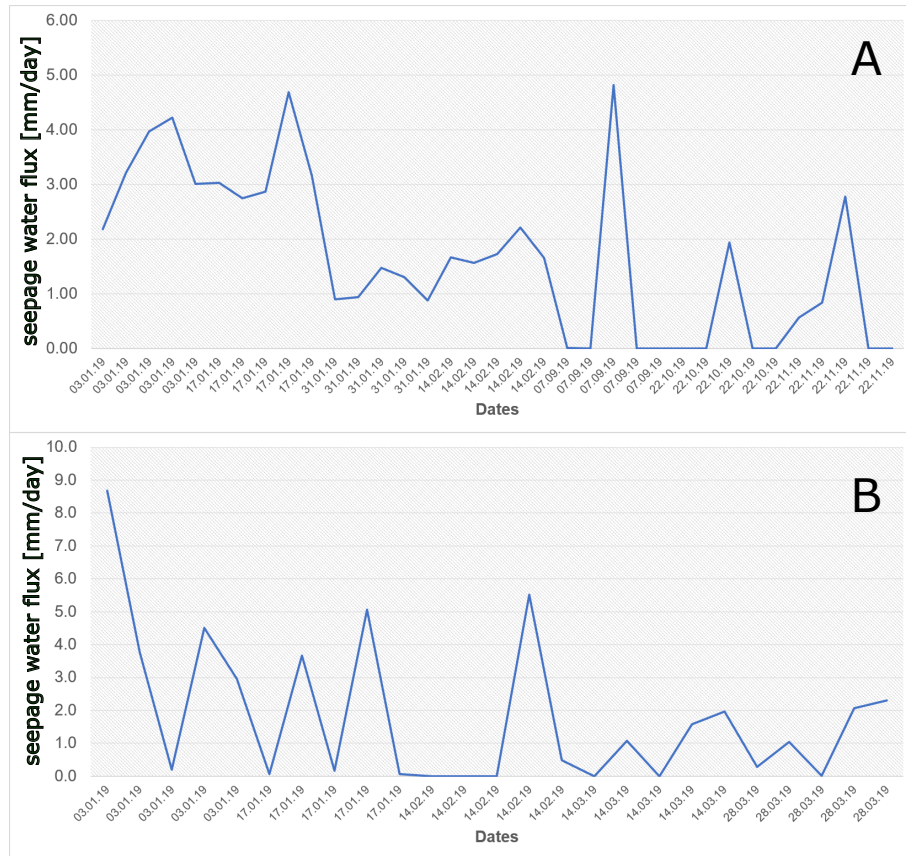


Figure S31: Seepage water flux at A. Poltringen and B. Tailfingen at 50cm depth measured at 2019.



## 11.2.6 Calibration to chemostat and retentostat data of alternative Monod-Model

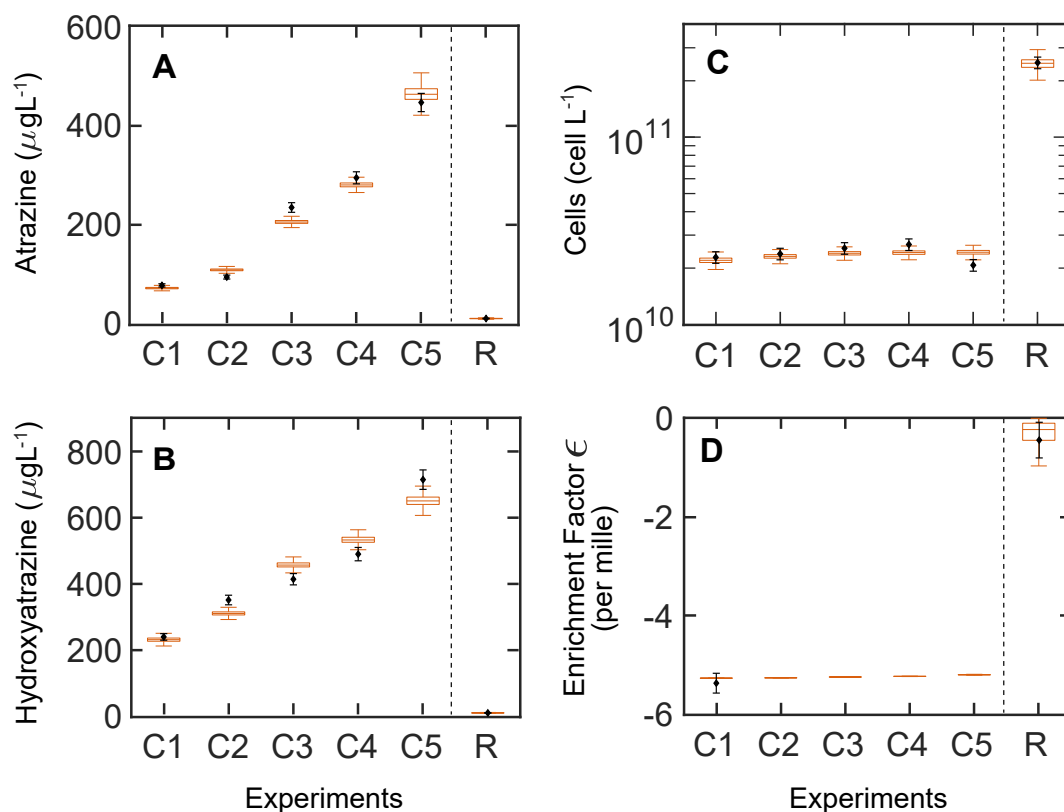


Figure S32: Simulations (boxplots) using model variant M (including Monod-kinetics) and measurements (blank diamonds + estimated standard deviation). **A-C.** Steady-state concentrations for the chemostat (for five dilution rates: C1-C5: 0.023, 0.032, 0.048, 0.056, 0.068  $\text{h}^{-1}$ ) and the retentostat (dilution rate: R: 0.020  $\text{h}^{-1}$ ). The middle line in the boxplot is the median of the ensemble outputs from the MCMC (see M&M 7.4); boxes represent 25% and 75% percentiles, and whiskers corresponds to  $\pm 1.5 \times \text{IQR}$  (interquartile range). **D.** Enrichment factors ( $\epsilon$ ), reported only for the lowest dilution rate of the chemostat (C1) and retentostat (R), but simulated for C2-C5.

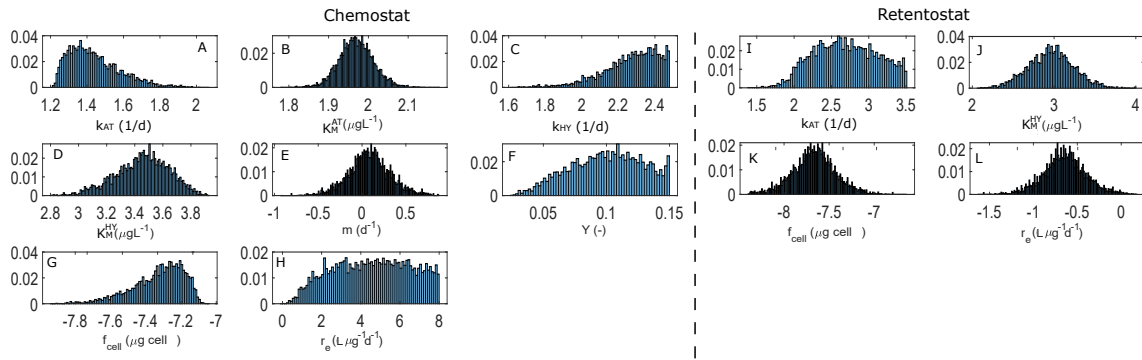


Figure S33: Posterior distributions for calibrations with chemostat (8 parameters) and retentostat (4 parameters) data. All parameters are expressed in log scale with the exception of the parameter  $Y$ .

### 11.2.7 Correlation tables

Table S11: Parameter correlation for model variant T data using MCMC ensemble

	$k_{AT}$	$K_M^{AT}$	$k_{HY}$	$K_M^{HY}$	$m$	$Y$	$f_{cell}$	$r_e$	$k_{AT}$	$K_M^{HY}$	$f_{cell}$	$r_e$
$k_{AT}$	1.00	0.12	1.00	0.01	0.42	-1.00	-1.00	0.07	0.24	0.18	-0.44	0.44
$K_M^{AT}$	0.12	1.00	0.07	0.14	-0.42	-0.08	-0.05	0.03	-0.17	0.52	0.42	-0.42
$k_{HY}$	1.00	0.07	1.00	0.09	0.43	-1.00	-1.00	0.07	0.24	0.18	-0.45	0.45
$K_M^{HY}$	0.01	0.14	0.09	1.00	-0.26	-0.01	0.00	-0.03	-0.12	0.36	0.26	-0.26
$m$	0.42	-0.42	0.43	-0.26	1.00	-0.43	-0.45	0.07	0.47	-0.78	-0.99	1.00
$Y$	-1.00	-0.08	-1.00	-0.01	-0.43	1.00	1.00	-0.07	-0.24	-0.18	0.45	-0.45
$f_{cell}$	-1.00	-0.05	-1.00	0.00	-0.45	1.00	1.00	-0.07	-0.25	-0.15	0.47	-0.47
$r_e$	0.07	0.03	0.07	-0.03	0.07	-0.07	-0.07	1.00	0.04	-0.03	-0.07	0.07
$k_{AT}$	0.24	-0.17	0.24	-0.12	0.47	-0.24	-0.25	0.04	1.00	-0.35	-0.47	0.44
$K_M^{HY}$	0.18	0.52	0.18	0.36	-0.78	-0.18	-0.15	-0.03	-0.35	1.00	0.77	-0.77
$f_{cell}$	-0.44	0.42	-0.45	0.26	-0.99	0.45	0.47	-0.07	-0.47	0.77	1.00	-1.00
$r_e$	0.44	-0.42	0.45	-0.26	1.00	-0.45	-0.47	0.07	0.44	-0.77	-1.00	1.00

Diagonal highlighted in light green; highly correlated parameters ( $\sigma^2 > 0.80$ ) highlighted in light blue.

Table S10: Calibrated parameter values of the model variant M based on a MCMC ensemble for Chemostat and Retentostat fits

Parameters	Chemostat				Retentostat			
	MAP	Mean	Std	Range	MAP	Mean	Std	Range
$k_{AT}$ [ $\text{d}^{-1}$ ]	22.8	27.5	1.4	[16.0 – 115.7]	214.8	489.9	2.6	[25.7 – 3159.2]
$K_M^{HY}$ [ $\mu\text{g L}^{-1}$ ]	$5.1 \cdot 10^3$	$2.8 \cdot 10^3$	1.5	[642.7 – $8.4 \cdot 10^3$ ]	$1.6 \cdot 10^3$	920.4	1.9	[116.1 – $9.6 \cdot 10^3$ ]
$f_{cell}$ [ $\mu\text{g cell}^{-1}$ ]	$5.6 \cdot 10^{-8}$	$4.7 \cdot 10^{-8}$	1.4	[ $1.2 \cdot 10^{-8}$ – $9.0 \cdot 10^{-8}$ ]	$2.5 \cdot 10^{-8}$	$2.2 \cdot 10^{-8}$	1.7	[ $4.0 \cdot 10^{-9}$ – $2.3 \cdot 10^{-7}$ ]
$r_e$ [ $\text{L d}^{-1} \mu\text{g}^{-1}$ ]	$8.8 \cdot 10^3$	$3.7 \cdot 10^4$	88.7	[ $0.9 - 1.0 \cdot 10^6$ ]	0.2	0.2	1.8	[0.02 – 1.5]
$K_M^{AT}$ [ $\mu\text{g L}^{-1}$ ]	92.9	93.2	1.1	[60.3 – 149.8]				
$k_{HY}$ [ $\text{d}^{-1}$ ]	263.9	185.9	1.4	[41.4 – 300.0]				
$m$ [ $\text{d}^{-1}$ ]	1.0	1.2	1.7	[0.1 – 7.4]				
$Y$ [-]	0.1	0.1	1.4	[0.02 – 0.1]				

Maximum a-posteriori (MAP), mean, standard deviation (Std) and range of the MCMC ensemble for both engineered systems.

Table S12: Parameter correlation for model variant M data using MCMC ensemble

	$k_{AT}$	$K_M^{AT}$	$k_{HY}$	$K_M^{HY}$	$m$	$Y$	$f_{cell}$	$r_e$		$k_{AT}$	$K_M^{HY}$	$f_{cell}$	$r_e$
$k_{AT}$	1.00	-0.09	0.34	-0.52	0.70	-1.00	-0.98	0.07	0.23	-0.42	-0.69	0.69	
$K_M^{AT}$	-0.09	1.00	0.15	0.33	-0.65	0.15	0.24	0.01	-0.19	0.64	0.64	-0.64	
$k_{HY}$	0.34	0.15	1.00	0.61	0.06	-0.35	-0.31	-0.04	0.02	0.50	-0.06	0.06	
$K_M^{HY}$	-0.52	0.33	0.61	1.00	-0.60	0.53	0.56	-0.09	-0.20	0.85	0.59	-0.59	
$m$	0.70	-0.65	0.06	-0.60	1.00	-0.70	-0.77	0.07	0.31	-0.83	-0.99	0.99	
$Y$	-1.00	0.15	-0.35	0.53	-0.70	1.00	0.99	-0.06	-0.24	0.42	0.70	-0.70	
$f_{cell}$	-0.98	0.24	-0.31	0.56	-0.77	0.99	1.00	-0.07	-0.26	0.50	0.77	-0.77	
$r_e$	0.07	0.01	-0.04	-0.09	0.07	-0.06	-0.07	1.00	0.03	-0.08	-0.07	0.07	
$k_{AT}$	0.23	-0.19	0.02	-0.20	0.31	-0.24	-0.26	0.03	1.00	-0.26	-0.31	0.23	
$K_M^{HY}$	-0.42	0.64	0.50	0.85	-0.83	0.42	0.50	-0.08	-0.26	1.00	0.83	-0.83	
$f_{cell}$	-0.69	0.64	-0.06	0.59	-0.99	0.70	0.77	-0.07	-0.31	0.83	1.00	-0.98	
$r_e$	0.69	-0.64	0.06	-0.59	0.99	-0.70	-0.77	0.07	0.23	-0.83	-0.98	1.00	

Diagonal highlighted in light green; highly correlated parameters ( $\sigma^2 > 0.80$ ) highlighted in light blue.

## 11.2.8 Global sensitivity results

### Morris Method

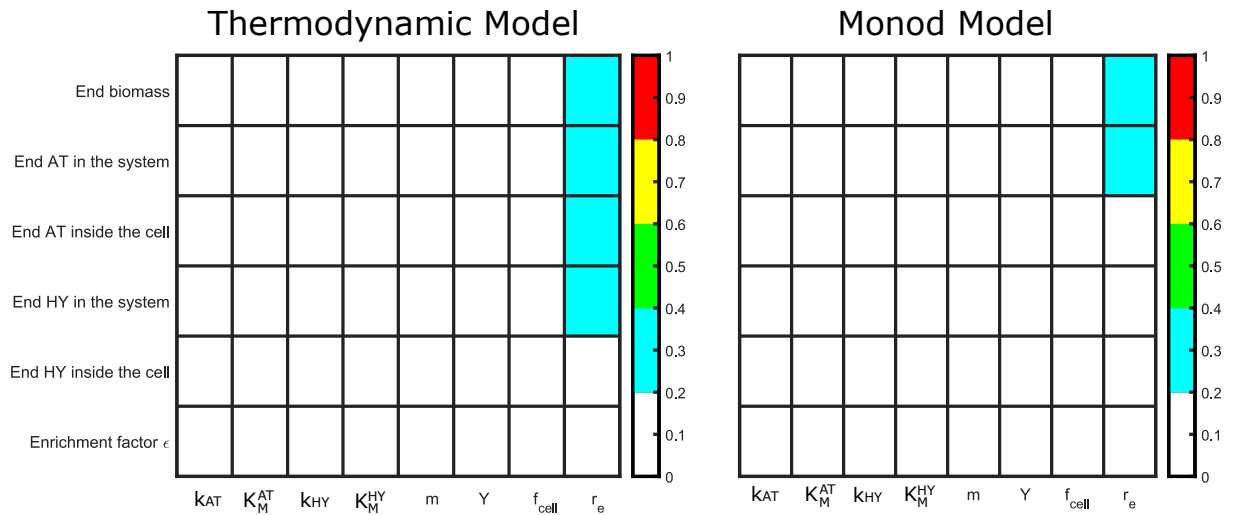


Figure S34:  $l_2$  – norm values from Morris Method [73, 105] for thermodynamic and Monod Model

## Sobol indices

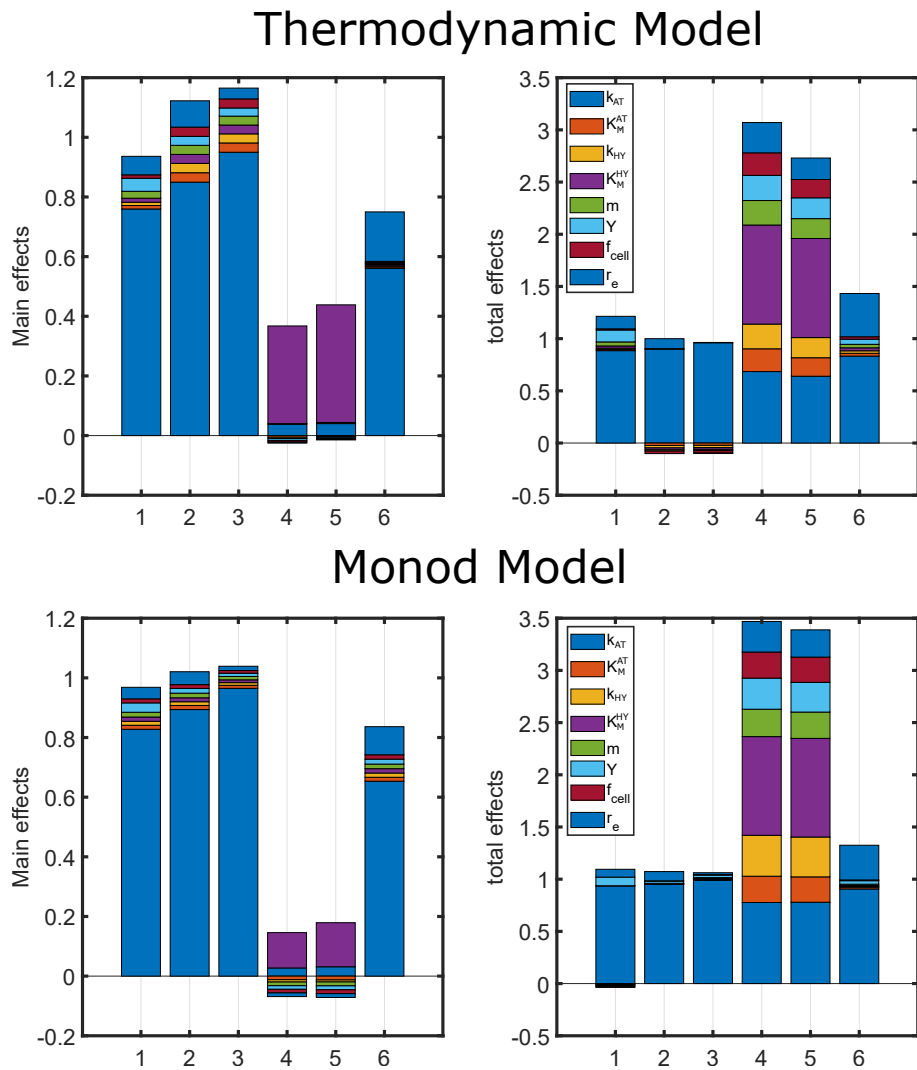


Figure S35: Sobol Indices [202]: 1: Final Biomass, 2 Final AT in the system, 3 Final AT inside the cell, 4: Final HY in the system, 5: Final HY inside the cell, 6: Enrichment Factor  $\epsilon$  fit.

Table S13: Local sensitivity scores for soil model variants of residual AT in soils

Parameters	Variant M-NL		Variant M-L		Variant T-NL		Variant T-L	
	Chemostat	Retentostat	Chemostat	Retentostat	Chemostat	Retentostat	Chemostat	Retentostat
$k_{AT}$	407.6	2.8	440.4	2.8	406.0	2.7	323.5	2.6
$k_{AT}^M$	399.3	2.9	392.3	2.9	386.3	2.7	296.6	2.6
$k_{HY}$	9.5	0.01	25.6	0.03	32.4	0.1	0.2	0.1
$k_{HY}^M$	9.5	0.01	25.2	0.03	819.7	2.6	8.2	2.7
$m$	126.8	105.1	115.0	106.5	77.8	229.5	0.2	221.1
$Y$	$1.3 \cdot 10^{-7}$	$1.7 \cdot 10^{-7}$	$1.9 \cdot 10^{-4}$	$2.7 \cdot 10^{-8}$	$7.1 \cdot 10^{-5}$	$2.2 \cdot 10^{-6}$	$1.3 \cdot 10^{-8}$	$2.9 \cdot 10^{-4}$
$f_{cell}$	$6.6 \cdot 10^{-9}$	$2.3 \cdot 10^{-5}$	$1.5 \cdot 10^{-7}$	$8.6 \cdot 10^{-7}$	$8.6 \cdot 10^{-8}$	$1.1 \cdot 10^{-7}$	$3.9 \cdot 10^{-8}$	$2.3 \cdot 10^{-8}$
$r_e$	$1.3 \cdot 10^{-6}$	320.6	$2.3 \cdot 10^{-7}$	321.8	$3.6 \cdot 10^{-6}$	299.0	$1.5 \cdot 10^{-6}$	286.6
$K_{AT}$	508.6	547.2	495.5	589.9	475.2	496.1	442.5	538.5
$n_{AT}$	$1.2 \cdot 10^4$	$2.1 \cdot 10^4$	$1.2 \cdot 10^4$	$2.3 \cdot 10^4$	$5.7 \cdot 10^3$	$1.1 \cdot 10^4$	$3.5 \cdot 10^3$	$1.3 \cdot 10^4$
$K_{HY}$	12.5	0.05	20.5	0.1	$1.1 \cdot 10^3$	3.6	7.2	3.7
$n_{HY}$	62.6	$5.2 \cdot 10^{-4}$	105.0	0.03	$1.2 \cdot 10^4$	4.8	52.2	5.1
$M$	157.2	98.0	245.7	98.3	297.6	22.9	322.9	21.7
AT - input	95.3	95.6	70.1	95.9	$1.1 \cdot 10^3$	245.1	1.4	238.6
$v_v$	-	-	0.0	0.0	-	-	1.5	1.3

Table S14: Local sensitivity scores for soil model variants of residual HY in soils

Parameters	Variant M=NL		Variant M=L		Variant T=NL		Variant T=L	
	Chemostat	Retentostat	Chemostat	Retentostat	Chemostat	Retentostat	Chemostat	Retentostat
$k_{AT}$	0.2	$4.4 \cdot 10^{-5}$	0.2	$4.1 \cdot 10^{-5}$	$5.8 \cdot 10^{-3}$	$5.0 \cdot 10^{-6}$	0.3	$7.0 \cdot 10^{-6}$
$K_M^{AT}$	0.03	$1.8 \cdot 10^{-5}$	0.04	$1.7 \cdot 10^{-5}$	$2.0 \cdot 10^{-3}$	$1.1 \cdot 10^{-6}$	0.1	$1.5 \cdot 10^{-6}$
$k_{HY}$	51.4	51.40	42.4	50.8	0.3	0.5	$1.1 \cdot 10^{-3}$	0.6
$K_M^{HY}$	50.6	48.0	41.6	47.7	6.8	12.8	0.1	12.1
$m$	12.6	38.8	7.9	38.7	0.04	0.4	$2.2 \cdot 10^{-6}$	0.6
$Y$	$4.4 \cdot 10^{-8}$	$5.1 \cdot 10^{-10}$	$8.8 \cdot 10^{-7}$	$7.9 \cdot 10^{-6}$	$3.3 \cdot 10^{-8}$	$7.5 \cdot 10^{-12}$	$8.6 \cdot 10^{-9}$	$8.3 \cdot 10^{-6}$
$f_{cell}$	$6.0 \cdot 10^{-13}$	$5.0 \cdot 10^{-15}$	$5.7 \cdot 10^{-12}$	$8.9 \cdot 10^{-15}$	$3.1 \cdot 10^{-14}$	$1.1 \cdot 10^{-11}$	$7.7 \cdot 10^{-13}$	$2.5 \cdot 10^{-14}$
$r_e$	$4.2 \cdot 10^{-9}$	25.6	$1.5 \cdot 10^{-9}$	25.5	$2.4 \cdot 10^{-12}$	0.02	$1.5 \cdot 10^{-10}$	0.01
$K_{AT}$	0.03	$9.9 \cdot 10^{-4}$	0.05	$1.1 \cdot 10^{-3}$	$2.2 \cdot 10^{-3}$	$4.1 \cdot 10^{-6}$	0.2	$8.4 \cdot 10^{-6}$
$n_{AT}$	0.3	$1.8 \cdot 10^{-4}$	0.5	$3.1 \cdot 10^{-4}$	$4.3 \cdot 10^{-3}$	0.01	2.4	0.02
$K_{HY}$	66.5	197.7	77.2	227.0	8.7	18.3	5.5	32.9
$n_{HY}$	327.8	2.7	311.3	0.1	80.9	23.8	39.4	37.3
$M$	16.5	37.0	17.4	36.7	0.2	0.04	0.2	$7.7 \cdot 10^{-3}$
AT - input	0.5	17.7	$8.4 \cdot 10^{-4}$	17.8	0.9	0.2	8.7	0.2
$v_v$	-	-	0.1	0.1	-	-	6.4	1.9

## 11.2.9 Local sensitivity analysis

## 11.2.10 Alternative model structure for engineered systems

## Model description

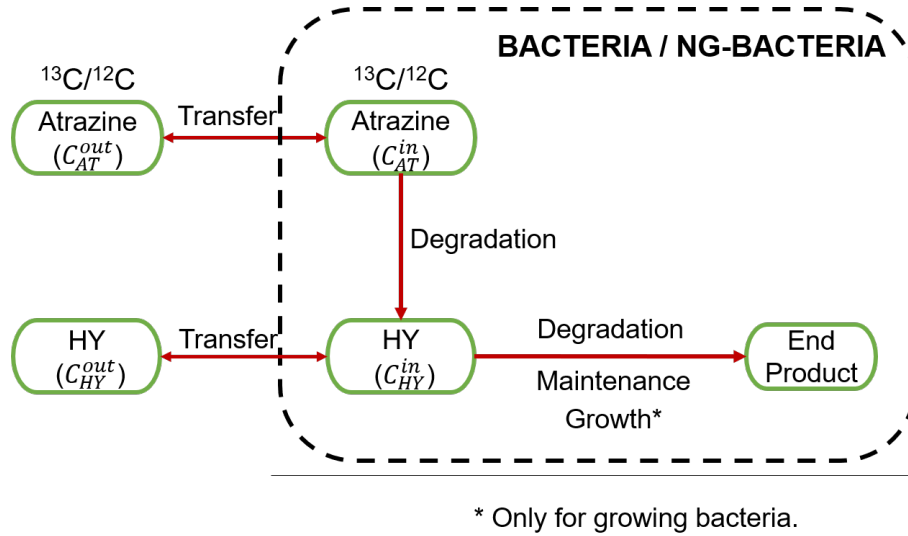


Figure S36: Alternative model structure for engineered systems, including a pool of growing and non-growing bacteria.

## Governing equations

1. Growing bacteria ( $C_B$ )

$$\frac{dC_B}{dt} = C_B \cdot \left( \underbrace{\frac{k_{HY} \cdot C_{HY_g}^i}{K_{HY} + C_{HY_g}^i} \cdot Y}_{\text{growth}} - \overbrace{\left( 1 - \frac{C_{HY_g}^i}{K_{HY} + C_{HY_g}^i} \right) \cdot m \cdot Y}_{\text{maintenance}} - \underbrace{r_D \cdot \beta}_{\text{fraction filtered}} - \overbrace{(1 - \tau_g) \cdot k_d}_{\text{deactivation}} \right) + \underbrace{\tau_{ng} \cdot k_r \cdot C_{ng}}_{\text{activation}} \quad (46)$$



2. Non-growing bacteria ( $C_{ng}$ )

$$\frac{dC_{ng}}{dt} = C_{ng} \cdot \left[ \overbrace{\left( 1 - \frac{C_{HY_{ng}}^i}{K_{HY} + C_{HY_{ng}}^i} \right) \cdot m \cdot Y}^{\text{maintenance}} \underbrace{- r_D \cdot \beta}_{\text{fraction filtered}} \overbrace{- \tau_{ng} \cdot k_r}_{\text{activation}} \right] + \underbrace{(1 - \tau_g) \cdot k_d \cdot C_B}_{\text{deactivation}} \quad (47)$$

3. Light AT isotopologue inside growing bacteria ( $C_{AT_{ig}}^l$ )

$$\begin{aligned} \frac{dC_{AT_{ig}}^l}{dt} = & \underbrace{- \frac{k_{AT} \cdot C_{AT_{ig}}^l}{K_{AT} + C_{AT_{ig}}^l + C_{AT_{ig}}^h} \cdot \frac{f_{cell}}{V_B^u}}_{\text{degradation}} + \underbrace{re \cdot \frac{f_{cell}}{V_B^u} \cdot (C_{AT_o}^l - C_{AT_{ig}}^l)}_{\text{mass-transfer}} \\ & \underbrace{-(1 - \tau_g) \cdot k_d \cdot C_{AT_{ig}}^l}_{\text{deactivation}} + \underbrace{\frac{\tau_{ng} \cdot k_r \cdot C_{AT_{ing}}^l \cdot C_{ng}}{C_B}}_{\text{activation}} - \underbrace{\frac{C_{AT_{ig}}^l}{C_B} \cdot \frac{dC_B}{dt}}_{\text{correction-factor}} \end{aligned} \quad (48)$$

4. Heavy AT isotopologue inside growing bacteria ( $C_{AT_{ig}}^h$ )

$$\begin{aligned} \frac{dC_{AT_{ig}}^h}{dt} = & \underbrace{- \frac{\alpha \cdot k_{AT} \cdot C_{AT_{ig}}^h}{K_{AT} + C_{AT_{ig}}^l + C_{AT_{ig}}^h} \cdot \frac{f_{cell}}{V_B^u}}_{\text{degradation}} + \underbrace{re \cdot (C_{AT_o}^h - C_{AT_{ig}}^h) \cdot \frac{f_{cell}}{V_B^u}}_{\text{mass-transfer}} \\ & \underbrace{-(1 - \tau_g) \cdot k_d \cdot C_{AT_{ig}}^h}_{\text{deactivation}} + \underbrace{\frac{\tau_{ng} \cdot k_r \cdot C_{AT_{ing}}^h \cdot C_{ng}}{C_B}}_{\text{activation}} - \underbrace{\frac{C_{AT_{ig}}^h}{C_B} \cdot \left( \frac{dC_B}{dt} \right)}_{\text{correction-factor}} \end{aligned} \quad (49)$$

5. Light AT isotopologue inside non-growing bacteria ( $C_{AT_{ing}}^l$ )

$$\begin{aligned}
\frac{dC_{AT_{ing}}^l}{dt} = & \underbrace{-\frac{k_{AT} \cdot C_{AT_{ing}}^l}{K_{AT} + C_{AT_{ing}}^l + C_{AT_{ing}}^h} \cdot \frac{f_{cell}}{V_B^u}}_{\text{degradation}} + \underbrace{re \cdot (C_{AT_o}^l - C_{AT_{ing}}^l) \cdot \frac{f_{cell}}{V_B^u}}_{\text{mass-transfer}} \\
& + \underbrace{\frac{(1 - \tau_g) \cdot k_d \cdot C_{AT_{ig}}^l \cdot C_B}{C_{ng}}}_{\text{deactivation}} - \underbrace{\tau_{ng} \cdot k_r \cdot C_{AT_{ing}}^l}_{\text{activation}} - \underbrace{\frac{C_{AT_{ing}}^l}{C_{ng}} \cdot \left(\frac{dC_{ng}}{dt}\right)}_{\text{correction-factor}} \quad (50)
\end{aligned}$$

6. Heavy AT isotopologue inside non-growing bacteria ( $C_{AT_{ing}}^h$ )

$$\begin{aligned}
\frac{dC_{AT_{ing}}^h}{dt} = & \underbrace{-\frac{\alpha \cdot k_{AT} \cdot C_{AT_{ing}}^h}{K_{AT} + C_{AT_{ing}}^l + C_{AT_{ing}}^h} \cdot \frac{f_{cell}}{V_B^u}}_{\text{degradation}} + \underbrace{re \cdot (C_{AT_o}^h - C_{AT_{ing}}^h) \cdot \frac{f_{cell}}{V_B^u}}_{\text{mass-transfer}} \\
& + \underbrace{\frac{(1 - \tau_g) \cdot k_d \cdot C_{AT_{ig}}^h \cdot C_B}{C_{ng}}}_{\text{deactivation}} - \underbrace{\tau_{ng} \cdot k_r \cdot C_{AT_{ing}}^h}_{\text{activation}} - \underbrace{\frac{C_{AT_{ing}}^h}{C_{ng}} \cdot \left(\frac{dC_{ng}}{dt}\right)}_{\text{correction-factor}} \quad (51)
\end{aligned}$$

7. Light AT isotopologue outside bacteria ( $C_{AT_o}^l$ )

$$\begin{aligned}
\frac{dC_{AT_o}^l}{dt} = & \underbrace{r_D \cdot (C_I^l - C_{AT_o}^l)}_{\text{input-output-of-system}} - \underbrace{re \cdot (C_{AT_o}^l - C_{AT_{ig}}^l) \cdot C_B}_{\text{mass-transfer-g}} - \underbrace{re \cdot (C_{AT_o}^l - C_{AT_{ing}}^l) \cdot C_{ng}}_{\text{mass-transfer-ng}} \quad (52)
\end{aligned}$$

8. Heavy AT isotopologue outside bacteria ( $C_{AT_o}^h$ )

$$\begin{aligned}
\frac{dC_{AT_o}^h}{dt} = & \underbrace{r_D \cdot (C_I^h - C_{AT_o}^h)}_{\text{input-output-of-system}} - \underbrace{re \cdot (C_{AT_o}^h - C_{AT_{ig}}^h) \cdot C_B}_{\text{mass-transfer-g}} - \underbrace{re \cdot (C_{AT_o}^h - C_{AT_{ing}}^h) \cdot C_{ng}}_{\text{mass-transfer-ng}} \quad (53)
\end{aligned}$$

9. HY inside growing bacteria ( $C_{HY_g}^i$ )

$$\begin{aligned}
\frac{dC_{HY_g}^i}{dt} = & \underbrace{\frac{k_{AT} \cdot C_{AT_{ig}}^l}{K_{AT} + C_{AT_{ig}}^l + C_{AT_{ig}}^h} \cdot \frac{f_{cell}}{v_B^u} + \frac{\alpha \cdot k_{AT} \cdot C_{AT_{ig}}^h}{K_{AT} + C_{AT_{ig}}^l + C_{AT_{ig}}^h} \cdot \frac{f_{cell}}{v_B^u}}_{\text{Input-AT}} \\
& - \underbrace{\frac{k_{HY} \cdot C_{HY_g}^i}{K_{HY} + C_{HY_g}^i} \cdot \frac{f_{cell}}{v_B^u}}_{\text{Degradation}} + \underbrace{re \cdot (C_{HY_o} - C_{HY_g}^i) \cdot \frac{f_{cell}}{v_B^u}}_{\text{Mass-Transfer-HY}} \\
& - \underbrace{(1 - \tau_g) \cdot k_d \cdot C_{HY_g}^i}_{\text{deactivation}} + \underbrace{\frac{\tau_{ng} \cdot k_r \cdot C_{HY_{ng}}^i \cdot C_B}{C_B}}_{\text{activation}} \\
& - \underbrace{\left( \frac{m \cdot C_{HY_g}^i}{K_{HY} + C_{HY_g}^i} \right) \cdot \frac{f_{cell}}{v_B^u}}_{\text{Maintenance}} - \underbrace{\frac{C_{HY_g}^i}{C_B} \cdot \left( \frac{dC_B}{dt} \right)}_{\text{Correction-Factor}} \quad (54)
\end{aligned}$$

10. HY inside non-growing bacteria ( $C_{HY_{ng}}^i$ )

$$\begin{aligned}
\frac{dC_{HY_{ng}}^i}{dt} = & \underbrace{\frac{k_{AT} \cdot C_{AT_{ing}}^l}{K_{AT} + C_{AT_{ing}}^l + C_{AT_{ing}}^h} \cdot \frac{f_{cell}}{v_B^u} + \frac{\alpha \cdot k_{AT} \cdot C_{AT_{ing}}^h}{K_{AT} + C_{AT_{ing}}^l + C_{AT_{ing}}^h} \cdot \frac{f_{cell}}{v_B^u}}_{\text{Input-AT}} \\
& - \underbrace{\frac{k_{HY} \cdot C_{HY_{ng}}^i}{K_{HY} + C_{HY_{ng}}^i} \cdot \frac{f_{cell}}{v_B^u}}_{\text{Degradation}} + \underbrace{re \cdot (C_{HY_o} - C_{HY_{ng}}^i) \cdot \frac{f_{cell}}{v_B^u}}_{\text{Mass-Transfer-HY}} \\
& + \underbrace{\frac{(1 - \tau_g) \cdot k_d \cdot C_{HY_g}^i \cdot C_B}{C_{ng}}}_{\text{deactivation}} - \underbrace{\tau_{ng} \cdot k_r \cdot C_{HY_{ng}}^i}_{\text{activation}} \\
& - \underbrace{\left( \frac{m \cdot C_{HY_{ng}}^i}{K_{HY} + C_{HY_{ng}}^i} \right) \cdot \frac{f_{cell}}{v_B^u}}_{\text{Maintenance}} - \underbrace{\frac{C_{HY_{ng}}^i}{C_{ng}} \cdot \left( \frac{dC_{ng}}{dt} \right)}_{\text{Correction-Factor}} \quad (55)
\end{aligned}$$

11. HY outside bacteria ( $C_{HY_o}$ )

$$\frac{d(C_{HY_o})}{dt} = \underbrace{-re \cdot (C_{HY_o} - C_{HY_g}^i) \cdot C_B}_{\text{Mass-Transfer-HY-g}} - \underbrace{re \cdot (C_{HY_o} - C_{HY_{ng}}^i) \cdot C_{ng}}_{\text{Mass-Transfer-HY-ng}} + \overbrace{-r_D \cdot C_{HY_o}}^{\text{Fraction-leaving-system}} \quad (56)$$

**Combined calibration of retentostat and chemostat results using Simulated annealing from Matlab**

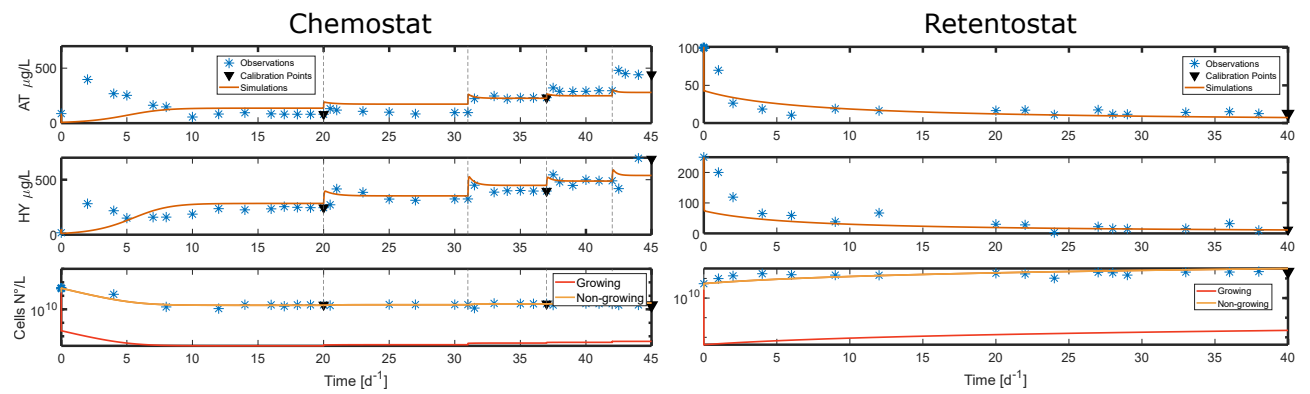


Figure S37: Chemostat and Retentostat fit: Atrazine, hydroxyatrazine and biomass

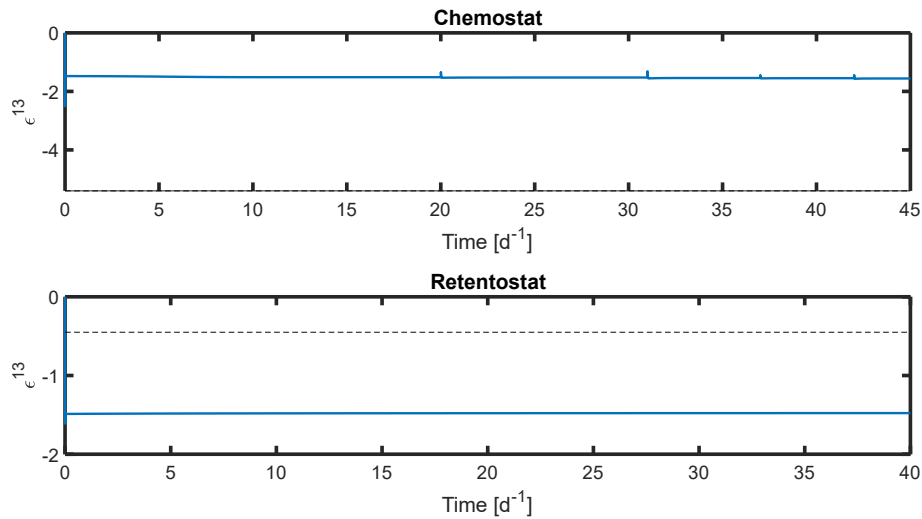


Figure S38: Chemostat and Retentostat fit: Enrichment factor. Dashed line indicate the enrichment value for chemostat -5.4‰ and retentostat -0.45‰

### 11.3 Supplementary Information for Chapter 8 (Paper 3)

#### 11.3.1 Main characteristics of the sampled soil

Table S15: Chemical and physical soil properties

Soil Horizon	Depth	pH	C <sub>org</sub>	Nitrogen	Phosphate	Sand	Silt	Clay
Ap	[cm]	[CaCl <sub>2</sub> ]	[mg g <sup>-1</sup> ]	[mg g <sup>-1</sup> ]	[mg g <sup>-1</sup> ]	[%]	[%]	[%]
	0-5	6.48	18.4	2.1	1.038	2.26	72.04	23.8

#### 11.3.2 MCPA degrader abundance and activity

Table S16: Reverse transcription

Reaction	Mixture	Temperature profile
reaction 1	11 µl DNA digestion sample	5 min at 65°C
(+ and - sample)	1 µl random primer 1 µl dNTPs	>= 1 min at 4°C
reaction 2	• Preparation (+ samples):	5 min at 25°C
	4 µl 5x First Strand Buffer	60 min at 50°C
	1 µl 0.1 M DTT (100mM)	15 min at 70°C
	1 µl RNase OUT	Cool down at 4°C
	1 µl reverse transcriptase (200 U/l)	
	13 µl reaction 1	
	• Preparation (- samples):	
	4 µl 5x First Strand Buffer	
	1 µl 0.1 M DTT (100mM)	
	1 µl RNase OUT	
	1 µl DEPC water	
	13 µl reaction 1	

Table S17: Digestion

Procedure	Reaction mixture	Temperature profile
Incubation	20 $\mu$ L RNA sample	30 min at 37 °C
	2.4 $\mu$ L 10x Turbo DNA buffer	
	1.6 $\mu$ L Turbo DNase	
DNase Inactivation	0.16 Vol Inactivation reagent	5 min at room temperature

### 11.3.3 Model parameters

Table S19: Model Parameters

Hierarchy	Parameters	Units	Description	Range
C	$f_T$	transcripts gene <sup>-1</sup>	Conversion factor transcripts per gene	[10 <sup>-3</sup> –10 <sup>3</sup> ]
C	$n_H$	-	Hill exponent	[1-10]
	$K_G$	mmol cm <sup>-3</sup>	Hill constant	[10 <sup>-10</sup> –10 <sup>3</sup> ]
C	$\mu_{max}$	d <sup>-1</sup>	Maximum growth rate coefficient	[0.1–5]
W	$a_a$	d <sup>-1</sup>	Decay rate coefficient	[10 <sup>-3</sup> –0.1]
	$Y_P$	-	Yield coefficient	[0.1–0.9]
	$m$	d <sup>-1</sup>	Maintenance coefficient	[10 <sup>-5</sup> –0.1]
	$K_M$	mmol cm <sup>-3</sup>	Monod constant	[10 <sup>-5</sup> –10 <sup>3</sup> ]
W	$a_{CO_2}$	d <sup>-1</sup>	Decay rate coefficient of the NER	[10 <sup>-5</sup> –0.1]
	$f_1$	mmol gene <sup>-1</sup>	Conversion factor cell to carbon	[10 <sup>-12</sup> –10 <sup>-9</sup> ]
	$Q_{10}$	-	Temperature function constant	[1.1–3]
S	$C_B^g$	gene g <sup>-1</sup>	Initial biomass	[10 <sup>4</sup> –10 <sup>6</sup> ]

C = concentration specific; W = water level specific; S = soil/sample specific

Table S18: Primer sequence

Target sequence	Primer*	qPCR conditions	References
16S rRNA genes	341F: CCT ACG GGA GGC AGC AG 515R: ATT ACC GCG GCT GCT GGC A	600 s at 95°C, Cycle (35): 15 s at 95°C, 30 s at 60°C, 30 s at 72°C, 30 s at 75°C (m.o.f) melting curve	López-Gutiérrez et al. [386]
<i>tfdA</i>	F: GAG CAC TAC GCR CTG AAY TCC CG R: GTC GCG TGC TCG AGA AG	600 s at 95°C, Cycle (40): 15 s at 95°C, 30 s at 64°C, 30 s at 72°C, 30 s at 81°C (m.o.f) melting curve	Bællum et al. [171]
<i>cadA</i>	F: AAG CTG CAR TTT GAR AAY GG R: MGG ATT GAA ATC CTG RTA	600 s at 95°C, Cycle (40): 15 s at 95°C, 30 s at 55°C, 30 s at 72°C, 30 s at 80°C (m.o.f) melting curve	Kitagawa et al. [174]
m.o.f - measurement of fluorescence			

Table S20: Calibrated model parameters

Hierarchy	Parameters	1 mg		20 mg	
		pF = 1.8	pF = 3.5	pF = 1.8	pF = 3.5
C.S	$f_T$	0.07	0.07	0.2	0.2
C.S	$n_H$	8.1	8.1	7.6	7.6
	$K_G$	$6.04 \cdot 10^{-7}$	$2.80 \cdot 10^{-8}$	$2.40 \cdot 10^{-7}$	$1.60 \cdot 10^{-8}$
C.S	$\mu_{max}$	1.8	1.8	0.2	0.2
WLS	$a_a$	0.01	0.009	0.01	0.009
	$Y_P$	0.7	0.87	0.6	0.7
	$m$	0.005	0.008	0.03	0.009
	$K_M$	0.0004	0.0001	0.0007	0.0002
WLS	$a_{CO_2}$	0.0003	0.0005	0.0003	0.0005
	$f_1$	$5.0 \cdot 10^{-11}$	$1.4 \cdot 10^{-11}$	$6.1 \cdot 10^{-10}$	$5.6 \cdot 10^{-10}$
	$Q_{10}$	1.5	1.4	2.03	2.07
S.S	<i>Initial</i>	54875.2	54875.2	54875.2	54875.2
	<i>Biomass</i>				

C.S = concentration specific; WLS = water level specific; S.S = soil/sample specific

Model codes available under: DOI: 10.5281/zenodo.5081655.



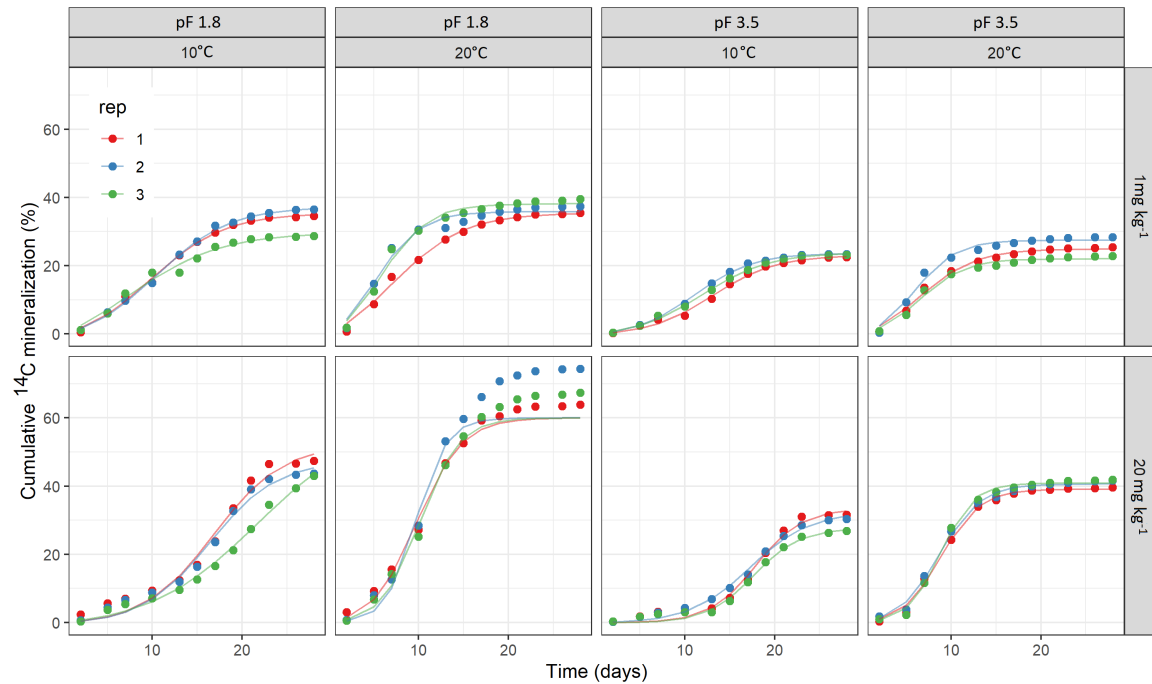
11.3.4  $^{14}\text{C}$  incorporation to  $^{14}\text{CO}_2$  respiration

Figure S39: Cumulative  $^{14}\text{CO}_2$  mineralization of two MCPA concentrations as a function of soil temperature and soil moisture over time. Mineralization of MCPA is represented by the percentage of initial  $^{14}\text{C}$ -MCPA. Curves were fitted to the data points via a logistic model.

Table S21: Contrast of the estimated marginal means of mineralization on day 28 as a function of temperature (the contrast function setting interaction = "tukey")

temp_treatment_tukey	Estimate	SE	df	t.ratio	p.value
1 - 20°C	-10.5	1.2	16	-8.7	$1.9 \cdot 10^{-7}$

Table S22: Interaction contrast of the estimated marginal means of mineralization on day 28 as a function of temperature and MCPA concentration (the contrast function setting interaction = "tukey")

temp_treatment_tukey	Concentration	Estimate	SE	df	t.ratio	p.value
1 - 20°C	1	-3.3	1.7	16	-1.9	0.07
1 - 20°C	20	-17.7	1.7	16	-10.3	$1.7 \cdot 10^{-8}$

Table S23: Contrast of the estimated marginal means of mineralization on day 28 as a function of soil moisture (the contrast function setting interaction = "tukey")

water_treatment_tukey	Estimate	SE	df	t.ratio	p.value
pF_1.8 - pF_3.5	16.2	1.2	16	13.4	$4.2 \cdot 10^{-10}$

Table S24: Interaction contrast of the estimated marginal means of mineralization on day 28 as a function of soil moisture and MCPA concentration (the contrast function setting interaction = "tukey")

water_treatment_tukey	Concentration	Estimate	SE	df	t.ratio	p.value
pF_1.8 - pF_3.5	1	11.1	1.7	16	6.5	$7.8 \cdot 10^{-6}$
pF_1.8 - pF_3.5	20	21.3	1.7	16	12.4	$1.2 \cdot 10^{-9}$

Table S25: Interaction contrast of the estimated marginal means of *tfdA* copies  $g^{-1}$  as a function of temperature, MCPA concentration, time and soil moisture (the contrast function setting interaction = "tukey")

Temp_tukey	MCPA	Day	pF	Estimate	SE	df	t.ratio	p.value
1 - 20°C	0	0	1.8	$-1.1 \cdot 10^{-12}$	91.8	24	$-1.1 \cdot 10^{-14}$	1.00
1 - 20°C	1,000	0	1.8	$2.6 \cdot 10^{-13}$	91.8	24	$2.9 \cdot 10^{-15}$	1.00
1 - 20°C	20,000	0	1.8	$6.6 \cdot 10^{-13}$	91.8	24	$7.2 \cdot 10^{-15}$	1.00
1 - 20°C	0	6	1.8	-51.7	91.8	24	-0.6	0.58
1 - 20°C	1,000	6	1.8	-180.1	91.8	24	-2.0	0.06
1 - 20°C	20,000	6	1.8	-85.2	91.8	24	-0.9	0.36
1 - 20°C	0	10	1.8	-138.2	91.8	24	-1.5	0.15
1 - 20°C	1,000	10	1.8	-130.8	91.8	24	-1.4	0.17
1 - 20°C	20,000	10	1.8	$-1.2 \cdot 10^{-3}$	91.8	24	-12.8	$3.0 \cdot 10^{-12}$
1 - 20°C	0	15	1.8	277.6	91.8	24	3.0	0.01
1 - 20°C	1,000	15	1.8	98.0	91.8	24	1.1	0.30
1 - 20°C	20,000	15	1.8	-915.8	91.8	24	-10.0	$5.2 \cdot 10^{-10}$
1 - 20°C	0	26	1.8	-2.3	91.8	24	0.0	0.98
1 - 20°C	1,000	26	1.8	4.9	91.8	24	0.1	0.96
1 - 20°C	20,000	26	1.8	-0.4	91.8	24	0.0	1.00
1 - 20°C	0	0	3.5	$-2.9 \cdot 10^{-15}$	91.8	24	$-3.2 \cdot 10^{-15}$	1.00
1 - 20°C	1,000	0	3.5	$1.2 \cdot 10^{-13}$	91.8	24	$1.3 \cdot 10^{-15}$	1.00
1 - 20°C	20,000	0	3.5	$1.2 \cdot 10^{-13}$	91.8	24	$1.3 \cdot 10^{-15}$	1.00
1 - 20°C	0	6	3.5	74.9	91.8	24	0.8	0.42
1 - 20°C	1,000	6	3.5	-33.4	91.8	24	-0.4	0.72
1 - 20°C	20,000	6	3.5	55.4	91.8	24	0.6	0.55
1 - 20°C	0	10	3.5	-18.7	91.8	24	-0.2	0.84
1 - 20°C	1,000	10	3.5	-55.6	91.8	24	-0.6	0.55
1 - 20°C	20,000	10	3.5	$-1.5 \cdot 10^{-3}$	91.8	24	-16.4	$1.6 \cdot 10^{-14}$
1 - 20°C	0	15	3.5	44.0	91.8	24	0.5	0.64
1 - 20°C	1,000	15	3.5	130.0	91.8	24	1.4	0.17
1 - 20°C	20,000	15	3.5	-580.9	91.8	24	-6.3	$1.5 \cdot 10^6$
1 - 20°C	0	26	3.5	-0.6	91.8	24	0.0	0.99
1 - 20°C	1,000	26	3.5	4.5	91.8	24	0.0	0.96
1 - 20°C	20,000	26	3.5	2.3 202	91.8	24	0.0	0.98

Table S26: Interaction contrast of the estimated marginal means of *tfdA* gene transcript as a function of temperature, MCPA concentration, time and soil moisture (the contrast function setting interaction = "tukey")

Temp_tukey	MCPA	Day	pF	Estimate	SE	df	t.ratio	p.value
1 - 20°C	0	0	1.8	$2.0 \cdot 10^{-10}$	$1.0 \cdot 10^5$	24	$2.0 \cdot 10^{-15}$	1.00
1 - 20°C	1,000	0	1.8	$9.5 \cdot 10^{-11}$	$1.0 \cdot 10^5$	24	$9.5 \cdot 10^{-16}$	1.00
1 - 20°C	20,000	0	1.8	$-1.1 \cdot 10^{-10}$	$1.0 \cdot 10^5$	24	$-1.1 \cdot 10^{-15}$	1.00
1 - 20°C	0	6	1.8	-346.1	$1.0 \cdot 10^5$	24	$-3.5 \cdot 10^{-3}$	1.00
1 - 20°C	1,000	6	1.8	$-3.1 \cdot 10^4$	$1.0 \cdot 10^5$	24	-0.3	0.76
1 - 20°C	20,000	6	1.8	$-2.0 \cdot 10^4$	$1.0 \cdot 10^5$	24	-0.2	0.84
1 - 20°C	0	10	1.8	-70.3	$1.0 \cdot 10^5$	24	$-7.0 \cdot 10^{-4}$	1.00
1 - 20°C	1,000	10	1.8	$1.9 \cdot 10^4$	$1.0 \cdot 10^5$	24	0.2	0.85
1 - 20°C	20,000	10	1.8	$-2.0 \cdot 10^6$	$1.0 \cdot 10^5$	24	-15.3	$6.8 \cdot 10^{-14}$
1 - 20°C	0	15	1.8	$5.0 \cdot 10^3$	$1.0 \cdot 10^5$	24	0.05	0.96
1 - 20°C	1,000	15	1.8	$1.6 \cdot 10^4$	$1.0 \cdot 10^5$	24	0.2	0.87
1 - 20°C	20,000	15	1.8	$-8.0 \cdot 10^4$	$1.0 \cdot 10^5$	24	-0.8	0.44
1 - 20°C	0	26	1.8	-519.9	$1.0 \cdot 10^5$	24	$-5.2 \cdot 10^{-3}$	1.00
1 - 20°C	1,000	26	1.8	18.2	$1.0 \cdot 10^5$	24	$1.8 \cdot 10^{-4}$	1.00
1 - 20°C	20,000	26	1.8	$1.5 \cdot 10^5$	$1.0 \cdot 10^5$	24	1.5	0.16
1 - 20°C	0	0	3.5	$1.5 \cdot 10^{-10}$	$1.0 \cdot 10^5$	24	$1.5 \cdot 10^{-15}$	1.00
1 - 20°C	1,000	0	3.5	$1.2 \cdot 10^{-10}$	$1.0 \cdot 10^5$	24	$1.2 \cdot 10^{-15}$	1.00
1 - 20°C	20,000	0	3.5	$3.9 \cdot 10^{-11}$	$1.0 \cdot 10^5$	24	$3.9 \cdot 10^{-16}$	1.00
1 - 20°C	0	6	3.5	$-2.6 \cdot 10^{-10}$	$1.0 \cdot 10^5$	24	$-2.6 \cdot 10^{-15}$	1.00
1 - 20°C	1,000	6	3.5	$7.1 \cdot 10^3$	$1.0 \cdot 10^5$	24	0.07	0.94
1 - 20°C	20,000	6	3.5	$-2.1 \cdot 10^4$	$1.0 \cdot 10^5$	24	-0.2	0.84
1 - 20°C	0	10	3.5	971.3	$1.0 \cdot 10^5$	24	$9.7 \cdot 10^{-3}$	0.99
1 - 20°C	1,000	10	3.5	$1.8 \cdot 10^4$	$1.0 \cdot 10^5$	24	0.2	0.86
1 - 20°C	20,000	10	3.5	$-1.4 \cdot 10^6$	$1.0 \cdot 10^5$	24	-14.1	$4.4 \cdot 10^{-13}$
1 - 20°C	0	15	3.5	-25.7	$1.0 \cdot 10^5$	24	$-2.6 \cdot 10^{-4}$	1.00
1 - 20°C	1,000	15	3.5	$2.6 \cdot 10^3$	$1.0 \cdot 10^5$	24	0.03	0.98
1 - 20°C	20,000	15	3.5	$-9.4 \cdot 10^3$	$1.0 \cdot 10^5$	24	-0.09	0.93
1 - 20°C	0	26	3.5	-10.6	$1.0 \cdot 10^5$	24	$-1.1 \cdot 10^{-4}$	1.00
1 - 20°C	1,000	26	3.5	12.4	$1.0 \cdot 10^5$	24	$1.2 \cdot 10^{-4}$	1.00
1 - 20°C	20,000	26	3.5	$4.7 \cdot 10^3$	$1.0 \cdot 10^5$	24	0.05	0.96

Table S27: Contrast of the estimated marginal means of  $^{14}\text{C}$  incorporation as a function of soil moisture (the contrast function setting interaction = "tukey")

Water_treatment_tukey	mcpa	Estimate	SE	df	t.ratio	p.value
pF 1.8 - pF 3.5	1 mg kg <sup>-1</sup>	-1.2	1.5	16	-0.8	0.4
pF 1.8 - pF 3.5	20 mg kg <sup>-1</sup>	2.7	1.5	16	1.8	0.1

Table S28: Contrast of the estimated marginal means of  $^{14}\text{C}$  incorporation as a function of temperature (the contrast function setting interaction = "tukey")

Temp_treatment_tukey	Day	estimate	SE	df	t.ratio	p.value
1 - 20°C	5	2.9	1.3	16	2.2	0.04
1 - 20°C	15	8.0	1.3	16	6.0	1.8 · 10 <sup>-5</sup>
1 - 20°C	28	6.5	1.4	16	4.8	2.0 · 10 <sup>-4</sup>

Table S29: Contrast of the estimated marginal means of  $CUE_M$  as a function of MCPA concentration (the contrast function setting interaction = "tukey")

Concentration_tukey	day	Estimate	SE	df	t.ratio	p.value
1 - 20 mg kg <sup>-1</sup>	5	-0.05	0.02	16	-2.1	0.05
1 - 20 mg kg <sup>-1</sup>	15	-0.06	0.03	16	-2.4	0.03
1 - 20 mg kg <sup>-1</sup>	28	0.05	0.02	16	2.1	0.05

Table S30: Contrast of the estimated marginal means of  $CUE_M$  as a function of temperature (the contrast function setting interaction = "tukey")

temp_treatment_tukey	day	Estimate	SE	df	t.ratio	p.value
1 - 20°C	5	0.2	0.0239	16	9.6	4.8 · 10 <sup>-8</sup>
1 - 20°C	15	0.4	0.0252	16	13.9	2.3 · 10 <sup>-10</sup>
1 - 20°C	28	0.1	0.0246	16	5.5	5.2 · 10 <sup>-5</sup>

Table S31: Contrast of the estimated marginal means of  $CUE_M$  as a function of soil moisture (the contrast function setting interaction = "tukey")

water_treatment_tukey	day	Estimate	SE	df	t.ratio	p.value
pF_1.8 - pF_3.5	5	-0.15	0.024	16	-6.3	$9.7 \cdot 10^{-6}$
pF_1.8 - pF_3.5	15	-0.12	0.025	16	-4.8	$2.2 \cdot 10^{-4}$
pF_1.8 - pF_3.5	28	-0.06	0.025	16	-2.5	0.02

### 11.3.5 $CUE_C$

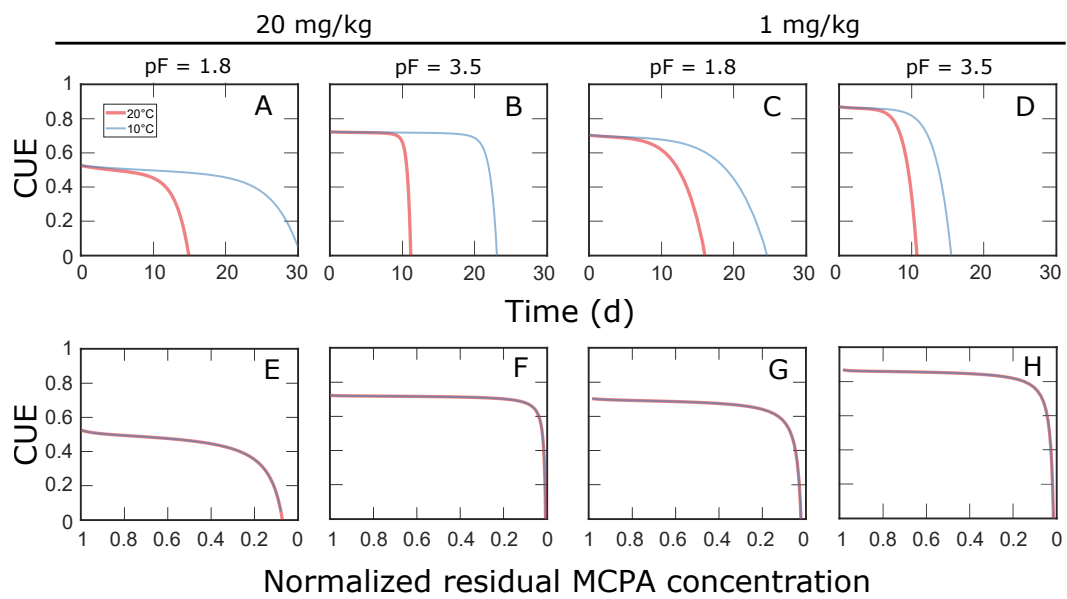


Figure S40:  $CUE_C$  (eq. 107) vs. time (d) showed in panels A to D, and  $CUE_C$  vs. normalized residual MCPA concentration in soils showed in panels E to H

## 11.4 Supplementary Information for Chapter 9 (Paper 4)

### 11.4.1 Model scenario descriptions

#### Model M1

The different bacterial guilds and the corresponding carbon sources are:

- $N_A$  (Guild A) = AT, HY, NE, NI, DEA, DIA.
- $N_B$  (Guild B) = HY, NI.
- $N_C$  (Guild C) = HY, NE, NI.
- $N_D$  (Guild D) = AT.

#### *Biokinetic functions*

##### Growth rate coefficient of Guild A

$$\mu_{A,AT} = \frac{\mu_{max}^{A,AT} \cdot AT^L}{\mu_{max}^{A,AT} + \mu_{N_A}} \quad (57)$$

$$\mu_{A,HY} = \frac{\mu_{max}^{A,HY} \cdot HY^L}{\mu_{max}^{A,HY} + \mu_{N_A}} \quad (58)$$

$$\mu_{A,NE} = \frac{\mu_{max}^{A,NE} \cdot NE^L}{\mu_{max}^{A,NE} + \mu_{N_A}} \quad (59)$$

$$\mu_{A,NI} = \frac{\mu_{max}^{A,NI} \cdot NI^L}{\mu_{max}^{A,NI} + \mu_{N_A}} \quad (60)$$

$$\mu_{A,DEA} = \frac{\mu_{max}^{A,DEA} \cdot DEA^L}{\mu_{max}^{A,DEA} + \mu_{N_A}} \quad (61)$$

$$\mu_{A,DIA} = \frac{\mu_{max}^{A,DIA} \cdot DIA^L}{\mu_{max}^{A,DIA} + \mu_{N_A}} \quad (62)$$

where:

$$\mu_{N_A} = AT^L + K_{A,AT} + HY^L + K_{A,HY} + NE^L + K_{A,NE} + NI^L + K_{A,NI} + DEA^L + K_{A,DEA} + DIA^L + K_{A,DIA} \quad (63)$$

#### Growth rate coefficient of Guild B

$$\mu_{B,HY} = \frac{\mu_{max}^{B,HY} \cdot HY^L}{\mu_{max}^{B,HY} + \mu_{N_B}} \quad (64)$$

$$\mu_{B,NI} = \frac{\mu_{max}^{B,NI} \cdot NI^L}{\mu_{max}^{B,NI} + \mu_{N_B}} \quad (65)$$

where:

$$\mu_{N_B} = HY^L + K_{B,HY} + NI^L + K_{B,NI} \quad (66)$$

#### Growth rate coefficient of Guild C

$$\mu_{C,HY} = \frac{\mu_{max}^{C,HY} \cdot HY^L}{\mu_{max}^{C,HY} + \mu_{N_C}} \quad (67)$$

$$\mu_{C,NE} = \frac{\mu_{max}^{C,NE} \cdot NE^L}{\mu_{max}^{C,NE} + \mu_{N_C}} \quad (68)$$

$$\mu_{C,NI} = \frac{\mu_{max}^{C,NI} \cdot NI^L}{\mu_{max}^{C,NI} + \mu_{N_C}} \quad (69)$$

where:

$$\mu_{N_C} = HY^L + K_{C,HY} + NE^L + K_{C,NE} + NI^L + K_{C,NI} \quad (70)$$

#### Growth rate coefficient of Guild D

$$\mu_{D,AT} = \frac{\mu_{max}^{D,AT} \cdot AT^L}{\mu_{max}^{D,AT} + \mu_{N_D}} \quad (71)$$



where:

$$\mu_{N_D} = AT^L + K_{D,AT} \quad (72)$$

Switch function for active/reactivation of Guild A

$$\tau_A = \frac{1}{\exp\left(\frac{N_{th}^A - (AT^L + HY^L + NE^L + NI^L + DEA^L + DIA^L)}{n \cdot N_{th}^A}\right) + 1} \quad (73)$$

Switch function for active/reactivation of Guild B

$$\tau_B = \frac{1}{\exp\left(\frac{N_{th}^B - (AT^L + HY^L + NI^L)}{n \cdot N_{th}^B}\right) + 1} \quad (74)$$

Switch function for active/reactivation of Guild C

$$\tau_C = \frac{1}{\exp\left(\frac{N_{th}^C - (HY^L + NE^L + NI^L)}{n \cdot N_{th}^C}\right) + 1} \quad (75)$$

Switch function for active/reactivation of Guild D

$$\tau_D = \frac{1}{\exp\left(\frac{N_{th}^D - (AT^L)}{n \cdot N_{th}^D}\right) + 1} \quad (76)$$

Dechlorination AT to HY by Guild B

$$K_{AT-HY}^o = \frac{k_{AT-HY} \cdot AT^L}{K_{AT-HY} + AT^L} \quad (77)$$

*Governing equations*

**Active Guild A**

$$\begin{aligned} \frac{dB_A^a}{dt} = & B_A^a \cdot (\mu_{A,AT} \cdot f_{AT-CA} + \mu_{A,HY} \cdot f_{HY-CA} + \mu_{A,NE} \cdot f_{NE-CA} + \mu_{A,NI} \cdot f_{NI-CA} + \\ & \mu_{A,DEA} \cdot f_{DEA-CA} + \mu_{A,DIA} \cdot f_{DIA-CA} - a_{a,A}) - (1 - \tau_A) \cdot k_A^d \cdot B_A^a + \tau_A \cdot k_A^a \cdot B_A^d \end{aligned} \quad (78)$$

**Dormant Guild A**

$$\frac{dB_A^d}{dt} = (1 - \tau_A) \cdot k_A^d \cdot B_A^a - \tau_A \cdot k_A^a \cdot B_A^d - B_A^d \cdot a_{d,A} \quad (79)$$

**Active Guild B**

$$\begin{aligned} \frac{dB_B^a}{dt} = B_B^a \cdot (\mu_{B,HY} \cdot f_{HY-NE} + \mu_{B,NI} \cdot f_{NI-CA} - a_{a,B}) - \\ (1 - \tau_B) \cdot k_B^d \cdot B_B^a + \tau_B \cdot k_B^a \cdot B_B^d \end{aligned} \quad (80)$$

**Dormant Guild B**

$$\frac{dB_B^d}{dt} = (1 - \tau_B) \cdot k_A^d \cdot B_B^a - \tau_B \cdot k_B^a \cdot B_B^d - B_B^d \cdot a_{d,B} \quad (81)$$

**Active Guild C**

$$\begin{aligned} \frac{dB_C^a}{dt} = B_C^a \cdot (\mu_{C,HY} \cdot f_{HY-NI} + \mu_{C,NE} \cdot f_{NE-CA} + \mu_{C,NI} \cdot f_{NI-CA} - a_{a,C}) - \\ (1 - \tau_C) \cdot k_C^d \cdot B_C^a + \tau_C \cdot k_C^a \cdot B_C^d \end{aligned} \quad (82)$$

**Dormant Guild C**

$$\frac{dB_C^d}{dt} = (1 - \tau_C) \cdot k_A^d \cdot B_C^a - \tau_C \cdot k_C^a \cdot B_C^d - B_C^d \cdot a_{d,C} \quad (83)$$

**Active Guild D**

$$\frac{dB_D^a}{dt} = B_D^a \cdot (\mu_{D,AT} \cdot f_D^o - a_{a,D}) - (1 - \tau_D) \cdot k_D^d \cdot B_D^a + \tau_D \cdot k_D^a \cdot B_D^d \quad (84)$$

**Dormant Guild D**

$$\frac{dB_D^d}{dt} = (1 - \tau_D) \cdot k_D^d \cdot B_D^a - \tau_D \cdot k_D^a \cdot B_D^d - B_D^d \cdot a_{d,D} \quad (85)$$

**Atrazine (AT) in Solution**

$$\begin{aligned} \frac{dAT^L}{dt} = - \left( B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,AT} \cdot \left( \frac{f_{AT-CA}}{Y_{A,AT}} - (1 - f_{AT-CA}) \right) + K_{AT-HY}^o \cdot \frac{\rho}{\theta} \cdot B_B^a + k_o \cdot AT^L + \right. \\ \left. B_D^a \cdot \frac{\rho}{\theta} \cdot \mu_{D,AT} \cdot \left( \frac{f_D^o}{Y_{D,AT}} - (1 - f_D^o) \right) \right) \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{AT} \cdot n_{FP}^{AT} \cdot (AT^L)^{(n_{FP}^{AT}-1)} \right)^{(-1)} \end{aligned} \quad (86)$$

**Hydroxyatrazine (HY) in Solution**

$$\begin{aligned} \frac{dHY^L}{dt} = & \left( B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,AT} \cdot (1 - f_H) \cdot (1 - f_{AT-CA}) + K_{AT-HY}^o \cdot \frac{\rho}{\theta} \cdot B_B^a + k_o \cdot AT^L \right. \\ & \left. - B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,HY} \cdot \left( \frac{f_{HY-CA}}{Y_{A,HY}} - (1 - f_{HY-CA}) \right) \right) - B_B^a \cdot \frac{\rho}{\theta} \cdot \mu_{B,HY} \cdot \left( \frac{f_{HY-NE}}{Y_{B,HY}} - (1 - f_{HY-NE}) \right) \\ & - B_C^a \cdot \frac{\rho}{\theta} \cdot \mu_{C,HY} \cdot \left( \frac{f_{HY-NI}}{Y_{C,HY}} - (1 - f_{HY-NI}) \right) \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{HY} \cdot n_{FP}^{HY} \cdot (HY^L)^{(n_{FP}^{HY}-1)} \right)^{(-1)} \end{aligned} \quad (87)$$

**N-Ethylammelide (NE) in Solution**

$$\begin{aligned} \frac{dNE^L}{dt} = & \left( B_B^a \cdot \frac{\rho}{\theta} \cdot \mu_{B,HY} \cdot (1 - f_{HY-NE}) - B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,NE} \cdot \left( \frac{f_{NE-CA}}{Y_{A,NE}} - (1 - f_{NE-CA}) \right) \right) \\ & - B_C^a \cdot \frac{\rho}{\theta} \cdot \mu_{C,NE} \cdot \left( \frac{f_{NE-CA}}{Y_{C,NE}} - (1 - f_{NE-CA}) \right) \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{NE} \cdot n_{FP}^{NE} \cdot (NE^L)^{(n_{FP}^{NE}-1)} \right)^{(-1)} \end{aligned} \quad (88)$$

**N-Isopropylammelide (NI) in Solution**

$$\begin{aligned} \frac{dNI^L}{dt} = & \left( B_C^a \cdot \frac{\rho}{\theta} \cdot \mu_{C,HY} \cdot (1 - f_{HY-NI}) - B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,NI} \cdot \left( \frac{f_{NI-CA}}{Y_{A,NI}} - (1 - f_{NI-CA}) \right) \right) \\ & - B_B^a \cdot \frac{\rho}{\theta} \cdot \mu_{B,NI} \cdot \left( \frac{f_{NI-CA}}{Y_{B,NI}} - (1 - f_{NI-CA}) \right) - B_C^a \cdot \frac{\rho}{\theta} \cdot \mu_{C,NI} \cdot \left( \frac{f_{NI-CA}}{Y_{C,NI}} - (1 - f_{NI-CA}) \right) \\ & \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{NI} \cdot n_{FP}^{NI} \cdot (NI^L)^{(n_{FP}^{NI}-1)} \right)^{(-1)} \end{aligned} \quad (89)$$

**Deethylatrazine (DEA) in Solution**

$$\begin{aligned} \frac{dDEA^L}{dt} = & \left( B_D^a \cdot \frac{\rho}{\theta} \cdot \mu_{D,AT} \cdot (1 - f_D^o) \cdot (1 - f_D) - B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,DEA} \cdot \left( \frac{f_{DEA-CA}}{Y_{A,DEA}} - (1 - f_{DEA-CA}) \right) \right) \\ & \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{DEA} \cdot n_{FP}^{DEA} \cdot (DEA^L)^{(n_{FP}^{DEA}-1)} \right)^{(-1)} \end{aligned} \quad (90)$$

**Deisopropylatrazine (DIA) in Solution**

$$\begin{aligned} \frac{dDIA^L}{dt} = & \left( B_D^a \cdot \frac{\rho}{\theta} \cdot \mu_{D,AT} \cdot (1 - f_D^o) \cdot f_D - B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,DIA} \cdot \left( \frac{f_{DIA-CA}}{Y_{A,DIA}} - (1 - f_{DIA-CA}) \right) \right) \\ & \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{DIA} \cdot n_{FP}^{DIA} \cdot (DIA^L)^{(n_{FP}^{DIA}-1)} \right)^{(-1)} \end{aligned} \quad (91)$$

**Cyanuric Acid (CA) in Solution**

$$\begin{aligned}
\frac{dCA^L}{dt} = & \left( B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,AT} \cdot (1 - f_{AT-CA}) \cdot f_H + B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,HY} \cdot (1 - f_{HY-CA}) \right) \\
& + B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,NE} \cdot (1 - f_{NE-CA}) + B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,NI} \cdot (1 - f_{NI-CA}) \\
& + B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,DEA} \cdot (1 - f_{DEA-CA}) + B_A^a \cdot \frac{\rho}{\theta} \cdot \mu_{A,DIA} \cdot (1 - f_{DIA-CA}) \\
& + B_B^a \cdot \frac{\rho}{\theta} \cdot \mu_{B,NI} \cdot (1 - f_{NI-CA}) + B_C^a \cdot \frac{\rho}{\theta} \cdot \mu_{C,NI} \cdot (1 - f_{NI-CA}) \\
& + B_C^a \cdot \frac{\rho}{\theta} \cdot \mu_{C,NE} \cdot (1 - f_{NE-CA} - CA^L \cdot d_{CA-CO_2} \cdot \frac{K_I}{NO_3 + K_I}) \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{CA} \cdot n_{FP}^{CA} \cdot (CA^L)^{(n_{FP}^{CA}-1)} \right)^{(-1)}
\end{aligned} \tag{92}$$

**CO<sub>2</sub>**

$$\begin{aligned}
\frac{dCO_2}{dt} = & CA^L \cdot d_{CA-CO_2} \cdot \frac{\rho}{\theta} \cdot \frac{K_I}{NO_3 + K_I} + B_A^a \cdot \mu_{A,AT} \cdot f_{AT-CA} \cdot \left( \frac{1 - Y_{A,AT}}{Y_{A,AT}} \right) \\
& + B_D^a \cdot \mu_{D,AT} \cdot f_D^o \cdot \left( \frac{1 - Y_{D,AT}}{Y_{D,AT}} \right) + B_A^a \cdot \mu_{A,HY} \cdot f_{HY-CA} \cdot \left( \frac{1 - Y_{A,HY}}{Y_{A,HY}} \right) \\
& + B_A^a \cdot \mu_{A,NE} \cdot f_{NE-CA} \cdot \left( \frac{1 - Y_{A,NE}}{Y_{A,NE}} \right) + B_A^a \cdot \mu_{A,NI} \cdot f_{NI-CA} \cdot \left( \frac{1 - Y_{A,NI}}{Y_{A,NI}} \right) \\
& + B_A^a \cdot \mu_{A,DEA} \cdot f_{DEA-CA} \cdot \left( \frac{1 - Y_{A,DEA}}{Y_{A,DEA}} \right) + B_A^a \cdot \mu_{A,DIA} \cdot f_{DIA-CA} \cdot \left( \frac{1 - Y_{A,DIA}}{Y_{A,DIA}} \right) \\
& + B_B^a \cdot \mu_{B,NI} \cdot f_{NI-CA} \cdot \left( \frac{1 - Y_{B,NI}}{Y_{B,NI}} \right) + B_B^a \cdot \mu_{B,HY} \cdot f_{HY-NE} \cdot \left( \frac{1 - Y_{B,HY}}{Y_{B,HY}} \right) \\
& + B_C^a \cdot \mu_{C,NE} \cdot f_{NE-CA} \cdot \left( \frac{1 - Y_{C,NE}}{Y_{C,NE}} \right) + B_C^a \cdot \mu_{C,NI} \cdot f_{NI-CA} \cdot \left( \frac{1 - Y_{C,NI}}{Y_{C,NI}} \right) \\
& + B_C^a \cdot \mu_{C,HY} \cdot f_{HY-NI} \cdot \left( \frac{1 - Y_{C,HY}}{Y_{C,HY}} \right) + (B_A^a \cdot a_{a,A} + B_A^d \cdot a_{d,A} + B_B^a \cdot a_{a,B} + B_B^d \cdot a_{d,B} \\
& + B_C^a \cdot a_{a,C} + B_C^d \cdot a_{d,C} + B_D^a \cdot a_{a,D} + B_D^d \cdot a_{d,D}) \cdot f_R
\end{aligned} \tag{93}$$

**Dissolved organic carbon(DOC)**

$$\frac{dDOC}{dt} = (B_A^a \cdot a_{a,A} + B_A^d \cdot a_{d,A} + B_B^a \cdot a_{a,B} + B_B^d \cdot a_{d,B} + B_C^a \cdot a_{a,C} + B_C^d \cdot a_{d,C} + B_D^a \cdot a_{a,D} + B_D^d \cdot a_{d,D}) \cdot (1 - f_R) \tag{94}$$

**Model M2**

The different bacterial guilds and the corresponding carbon sources are:

- $N_B$  (Guild B) = HY, NI.

- $N_C$  (Guild C) = HY, NE, NI.
- $N_D$  (Guild D) = AT.

**Model M3**

The different bacterial guilds and the corresponding carbon sources are:

- $N_A$  (Guild A) = AT, HY, NI, DEA, DIA.
- $N_C$  (Guild C) = HY, NI.
- $N_D$  (Guild D) = AT.

**Model M4**

The different bacterial guilds and the corresponding carbon sources are:

- $N_A$  (Guild A) = AT, HY, NE, DEA, DIA.
- $N_B$  (Guild B) = HY.
- $N_D$  (Guild D) = AT.

**Model M5**

The different bacterial guilds and the corresponding carbon sources are:

- $N_A$  (Guild A) = AT, HY, DEA, DIA.
- $N_D$  (Guild D) = AT.

**Model M6**

First order decay model on AT, main metabolites (AT, HY, DEA, DIA), and sinks (CA, CO<sub>2</sub>)

*Governing equations*

**Atrazine (AT) in Solution**

$$\frac{dAT^L}{dt} = - (AT^L \cdot (d_{AT-HY} + d_{AT-DD})) \cdot \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{AT} \cdot n_{FP}^{AT} \cdot (AT^L)^{(n_{FP}^{AT}-1)} \right)^{(-1)} \quad (95)$$

**Hydroxyatrazine (HY) in Solution**

$$\frac{dHY^L}{dt} = (AT^L \cdot d_{AT-HY} - HY^L \cdot d_{HY-CA}) \cdot \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{HY} \cdot n_{FP}^{HY} \cdot (HY^L)^{(n_{FP}^{HY}-1)} \right)^{(-1)} \quad (96)$$

**Deethylatrazine (DEA) in Solution**

$$\frac{dDEA^L}{dt} = (f_D \cdot AT^L \cdot d_{AT-DD} - DEA^L \cdot d_{DEA-CA}) \cdot \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{DEA} \cdot n_{FP}^{DEA} \cdot (DEA^L)^{(n_{FP}^{DEA}-1)} \right)^{(-1)} \quad (97)$$

**Deisopropylatrazine (DIA) in Solution**

$$\frac{dDIA^L}{dt} = ((1 - f_D) \cdot AT^L \cdot d_{AT-DD} - DIA^L \cdot d_{DIA-CA}) \cdot \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{DIA} \cdot n_{FP}^{DIA} \cdot (DIA^L)^{(n_{FP}^{DIA}-1)} \right)^{(-1)} \quad (98)$$

**Cyanuric Acid (CA) in Solution**

$$\begin{aligned} \frac{dCA^L}{dt} = & (HY^L \cdot d_{HY-CA} \cdot f_{HY-CA} + DEA^L \cdot d_{DEA-CA} \cdot f_{DEA-CA} + DIA^L \cdot d_{DIA-CA} \cdot f_{DIA-CA}) \\ & - CA^L \cdot d_{CA-CO_2} \cdot \frac{K_I}{NO_3 + K_I} \Big) \cdot \left( 1 + \frac{\rho}{\theta} \cdot K_{FP}^{CA} \cdot n_{FP}^{CA} \cdot (CA^L)^{(n_{FP}^{CA}-1)} \right)^{(-1)} \end{aligned} \quad (99)$$

CO<sub>2</sub>

$$\begin{aligned} \frac{dCO_2}{dt} = & CA^L \cdot \frac{\theta}{\rho} \cdot d_{CA-CO_2} \cdot \frac{\theta}{\rho} \cdot \frac{K_I}{NO_3 + K_I} + HY^L \cdot \frac{\theta}{\rho} \cdot d_{HY-CA} \cdot (1 - f_{HY-CA}) \\ & + DEA^L \cdot \frac{\theta}{\rho} \cdot d_{DEA-CA} \cdot (1 - f_{DEA-CA}) + DIA^L \cdot \frac{\theta}{\rho} \cdot d_{DIA-CA} \cdot (1 - f_{DIA-CA}) \end{aligned} \quad (100)$$

## 11.4.2 Modified constraint-based parameter search algorithm - Sampling results

### Sampling histograms

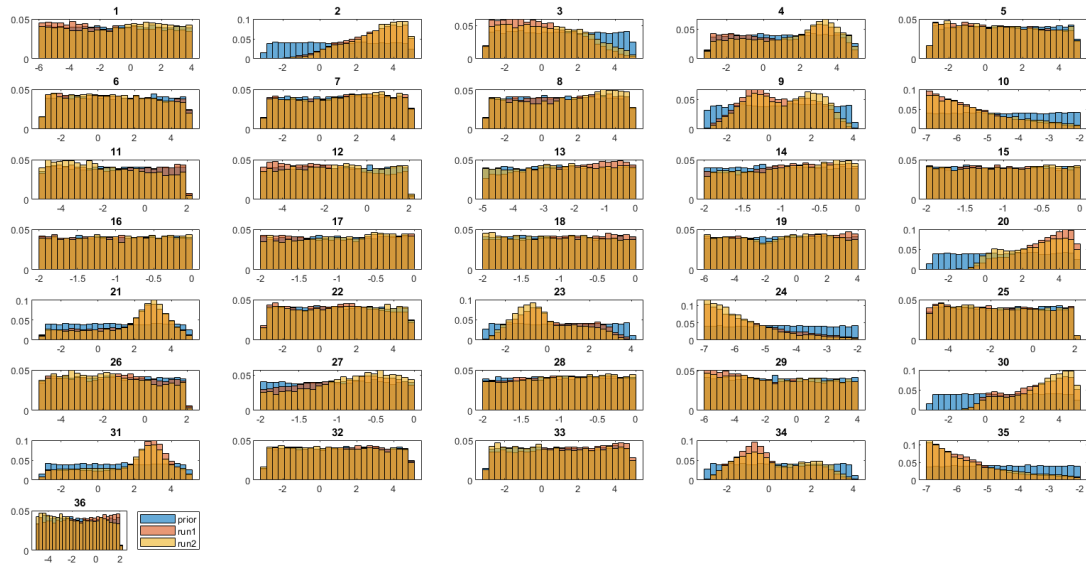


Figure S41: Sampled parameters for model scenario M1 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results

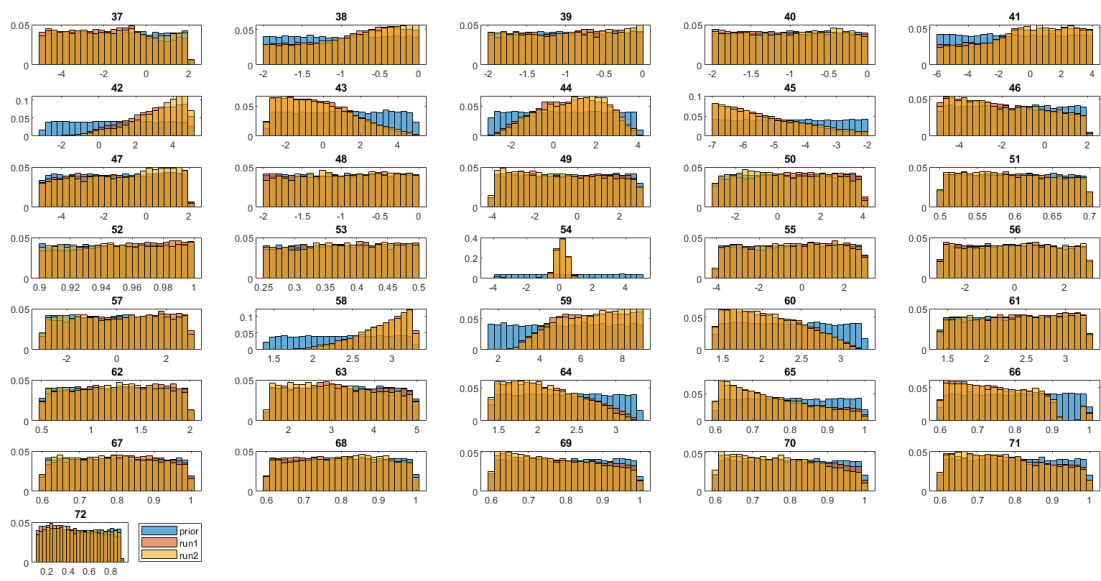


Figure S42: Sampled parameters (continuation) for model scenario M1 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results



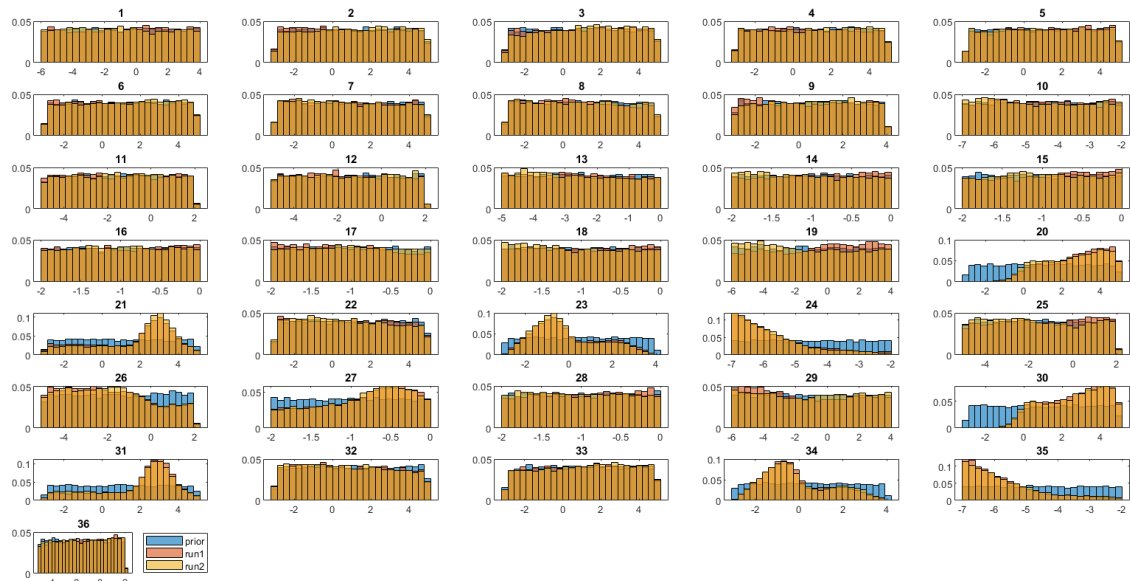


Figure S43: Sampled parameters for model scenario M2 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results

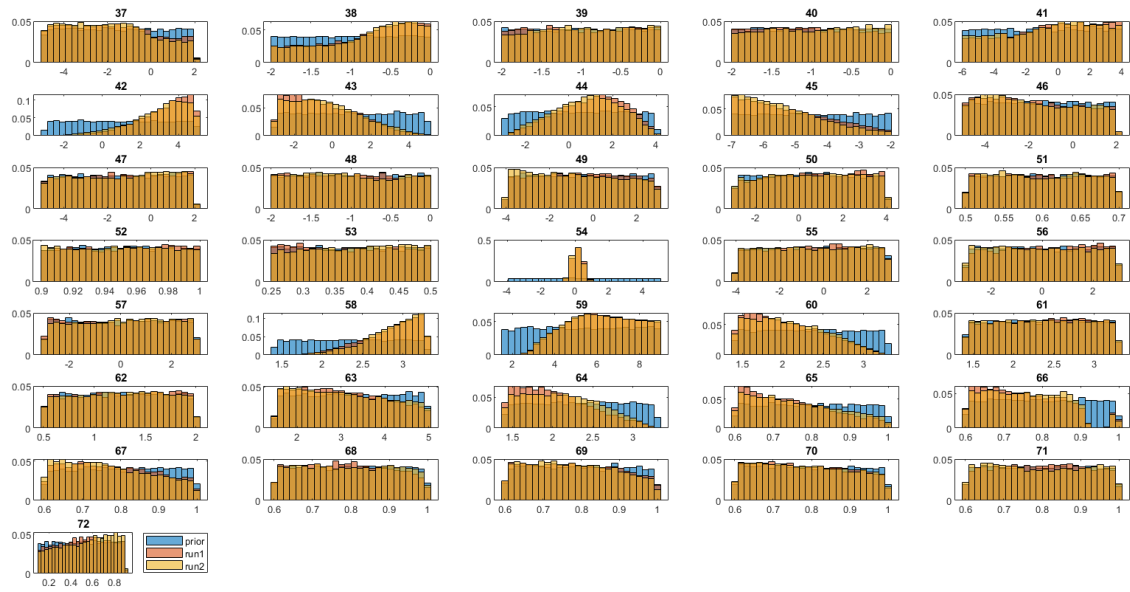


Figure S44: Sampled parameters (continuation) for model scenario M2 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results

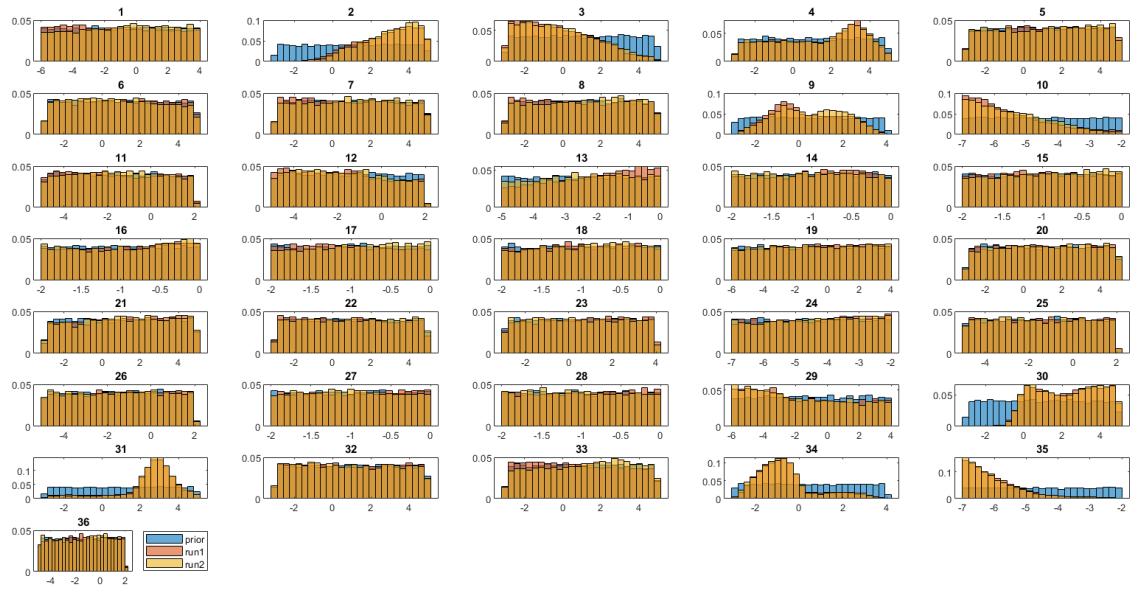


Figure S45: Sampled parameters for model scenario M3 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results

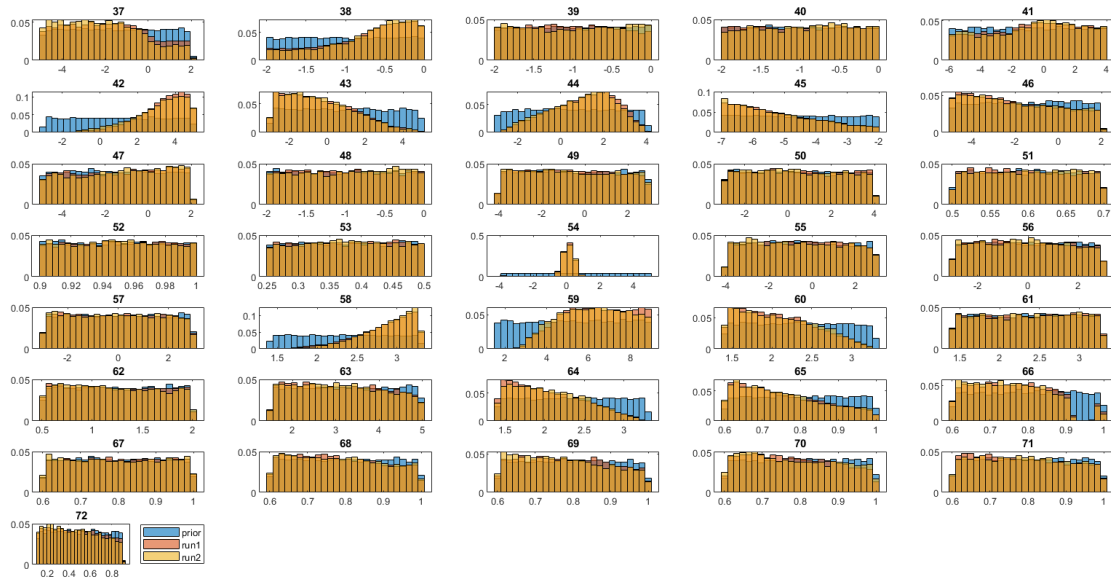


Figure S46: Sampled parameters (continuation) for model scenario M3 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results

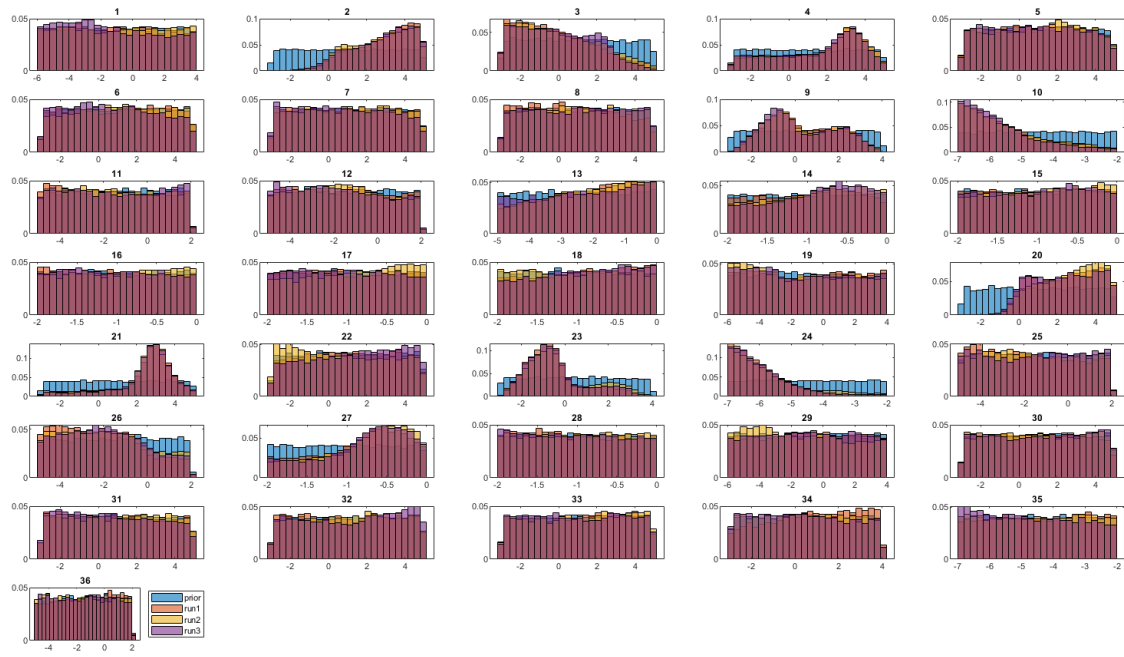


Figure S47: Sampled parameters for model scenario M4 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of three independent runs of the sampling algorithm (pink, yellow and purple). Both independent runs show similar sampling results

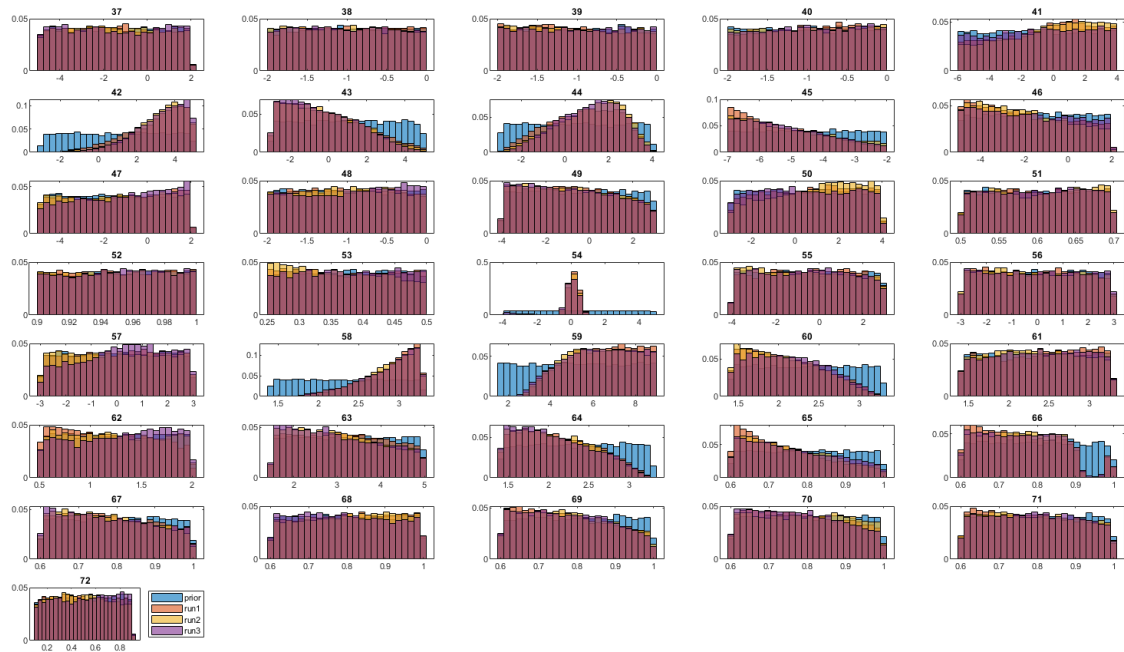


Figure S48: Sampled parameters (continuation) for model scenario M4 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of three independent runs of the sampling algorithm (pink, yellow and purple). Both independent runs show similar sampling results

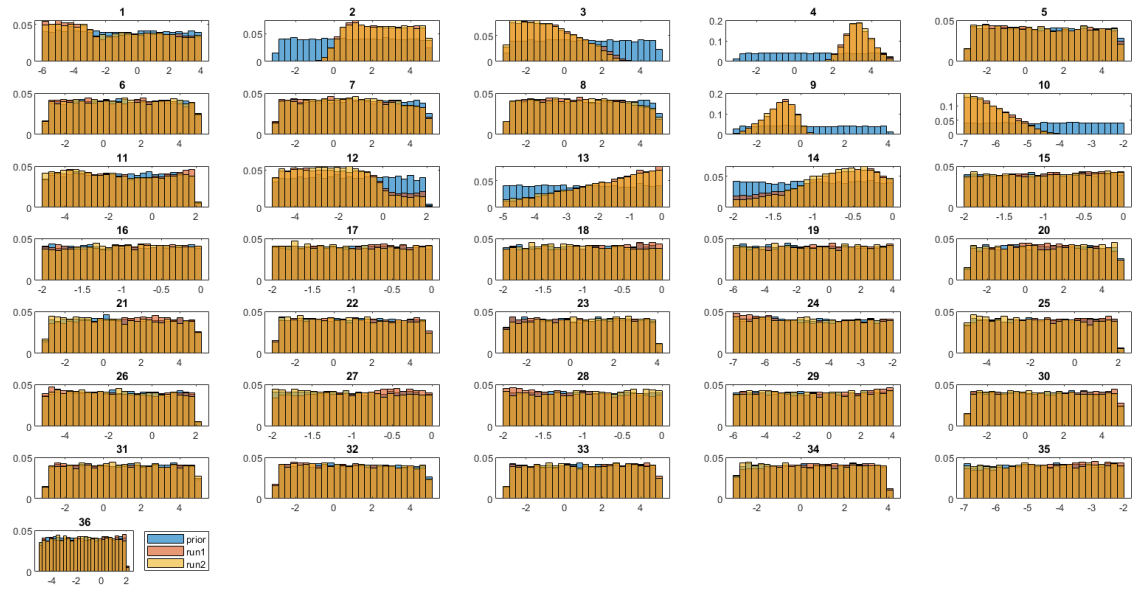


Figure S49: Sampled parameters for model scenario M5 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results

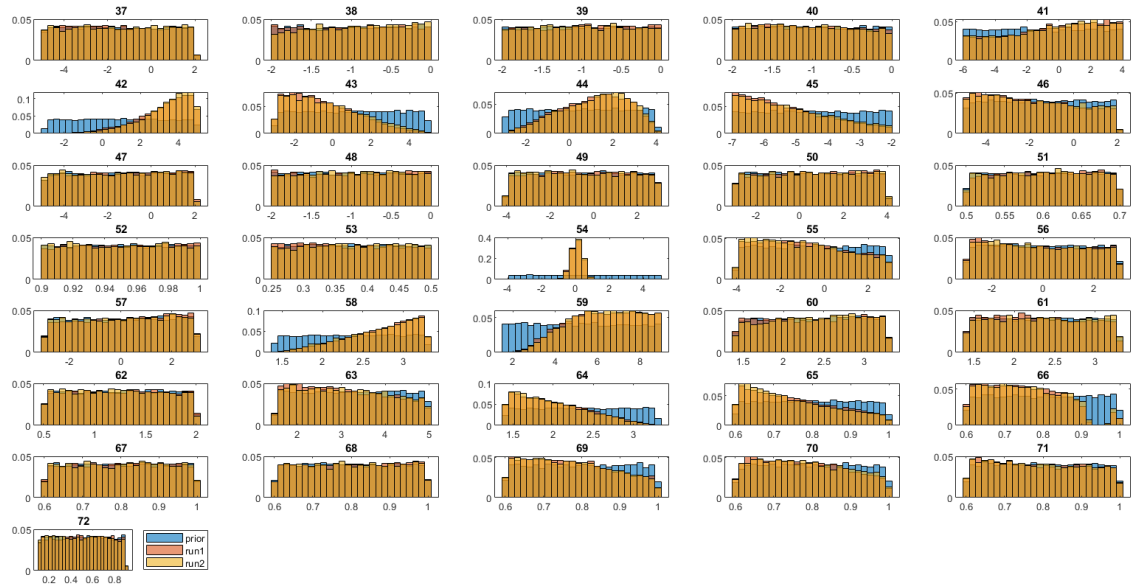


Figure S50: Sampled parameters (continuation) for model scenario M5 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results

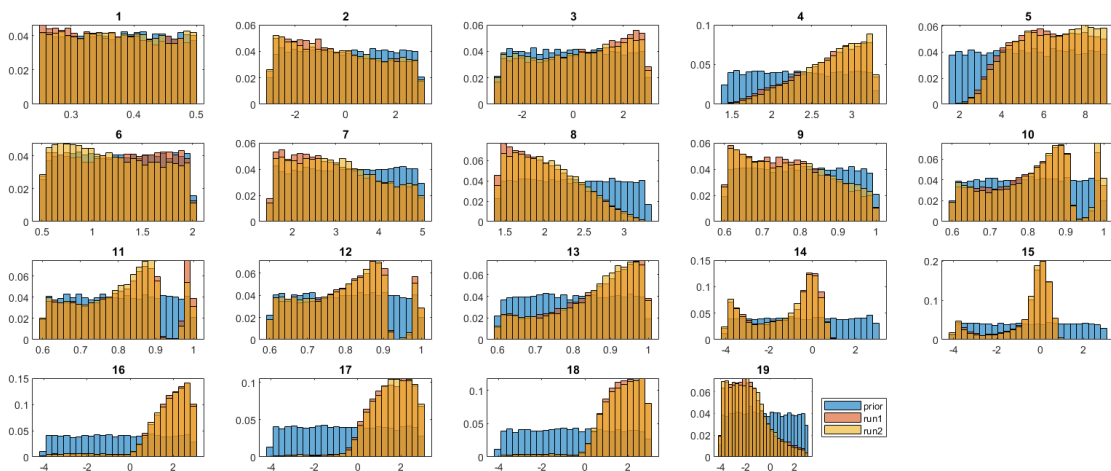


Figure S51: Sampled parameters for model scenario M6 after applying the modified constraint-based parameter search algorithm. We showed prior parameter distribution (blue) and the results of two independent runs of the sampling algorithm (pink and yellow). Both independent runs show similar sampling results



## Model outputs - Spaghetti plots

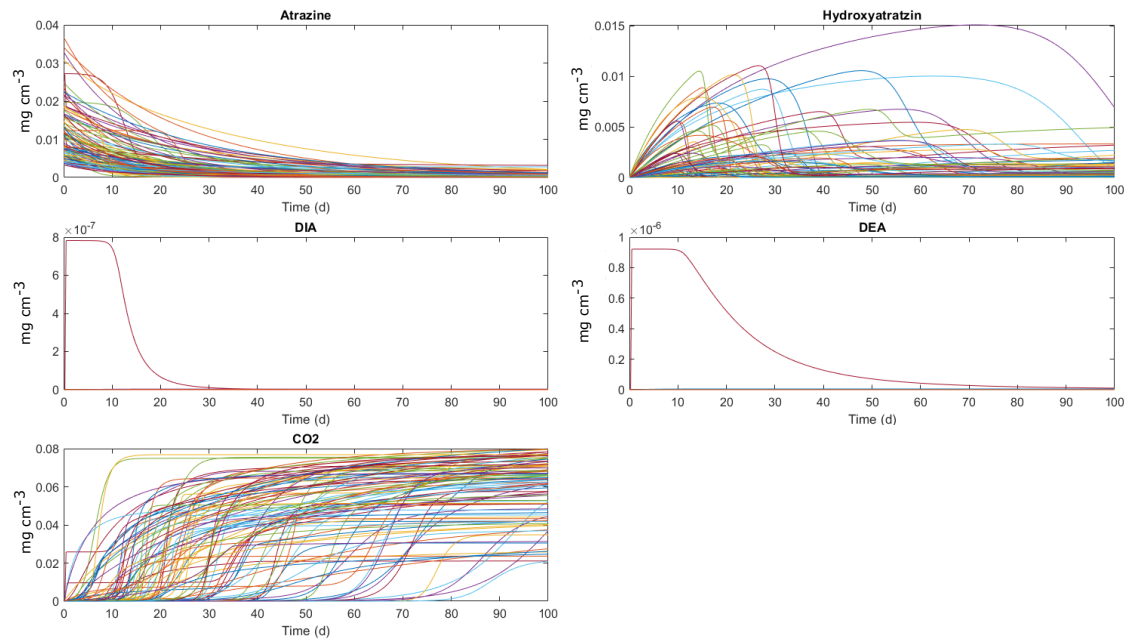


Figure S52: Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO<sub>2</sub>)) presented as spaghetti plots for model scenario M1. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm. Pools of DIA and DEA lie under the detection limit of  $1 \cdot 10^{-7} \text{ mg cm}^{-3}$

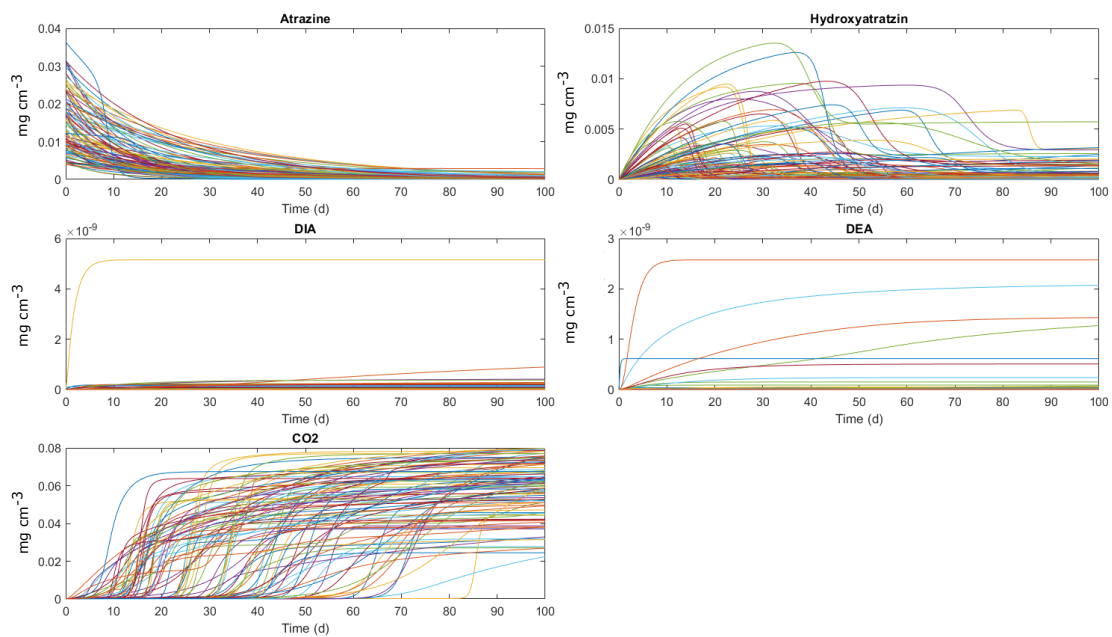


Figure S53: Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO<sub>2</sub>)) presented as spaghetti plots for model scenario M2. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm. Pools of DIA and DEA lie under the detection limit of  $1 \cdot 10^{-7} \text{ mg cm}^{-3}$

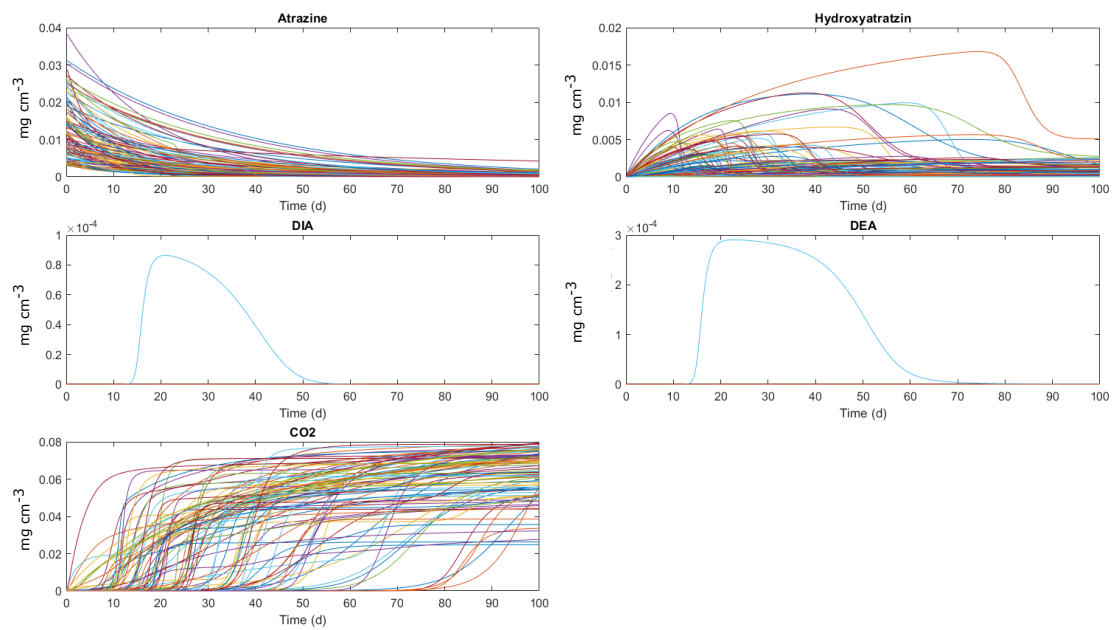


Figure S54: Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO<sub>2</sub>)) presented as spaghetti plots for model scenario M3. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm. Pools of DIA and DEA lie under the detection limit of  $1 \cdot 10^{-7} \text{ mg cm}^{-3}$

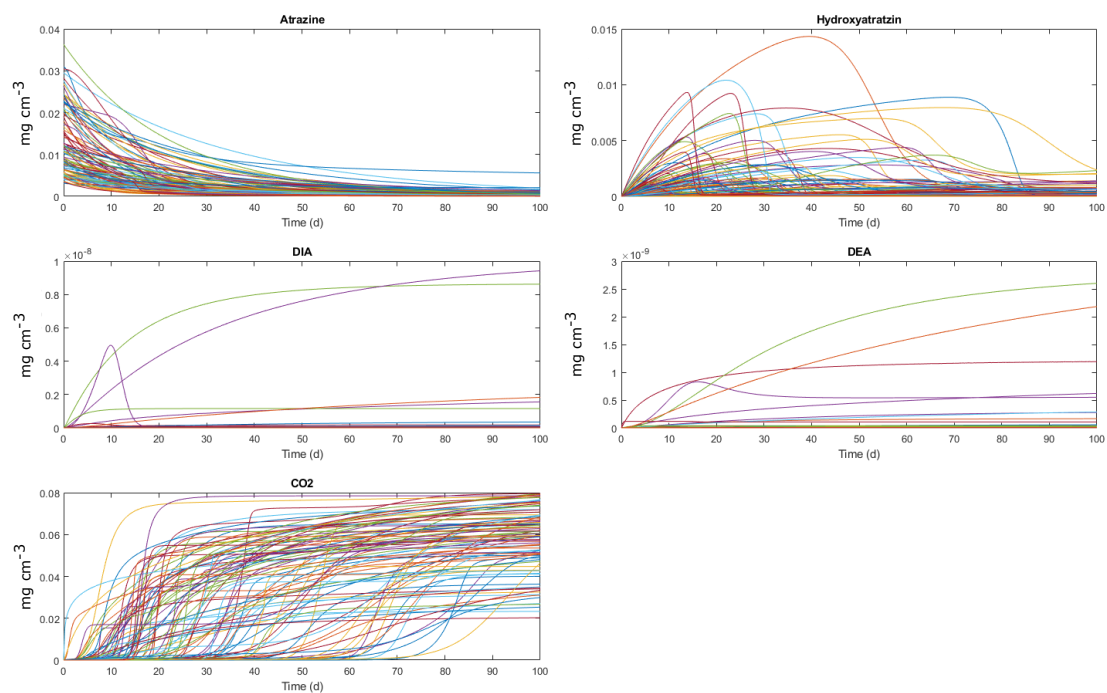


Figure S55: Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA,  $\text{CO}_2$ )) presented as spaghetti plots for model scenario M4. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm. Pools of DIA and DEA lie under the detection limit of  $1 \cdot 10^{-7} \text{ mg cm}^{-3}$

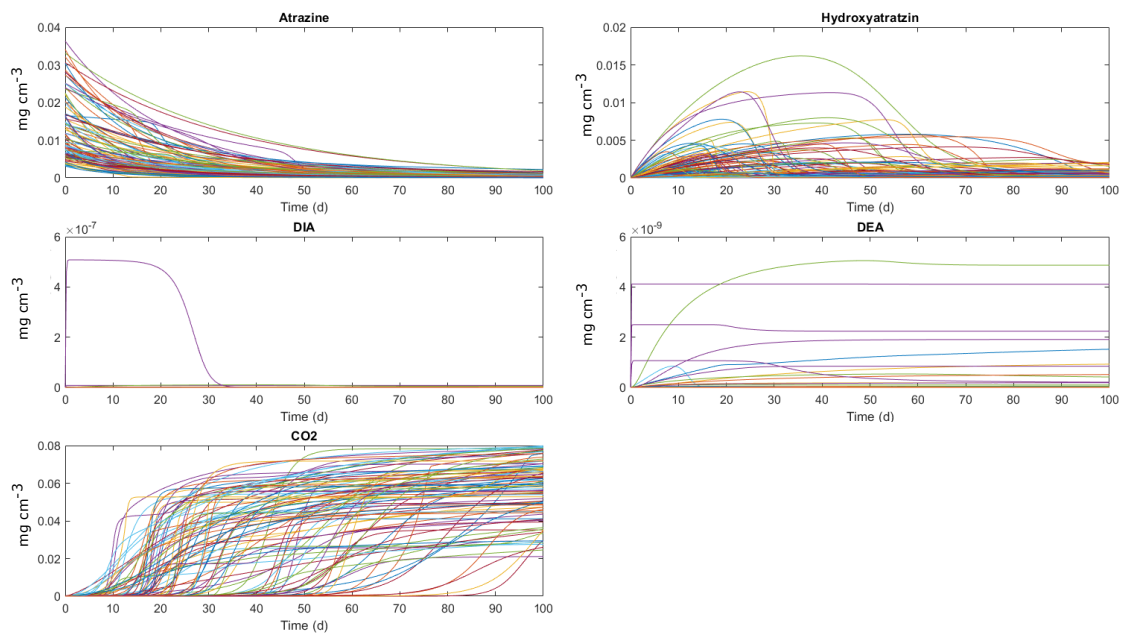


Figure S56: Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO<sub>2</sub>)) presented as spaghetti plots for model scenario M5. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm

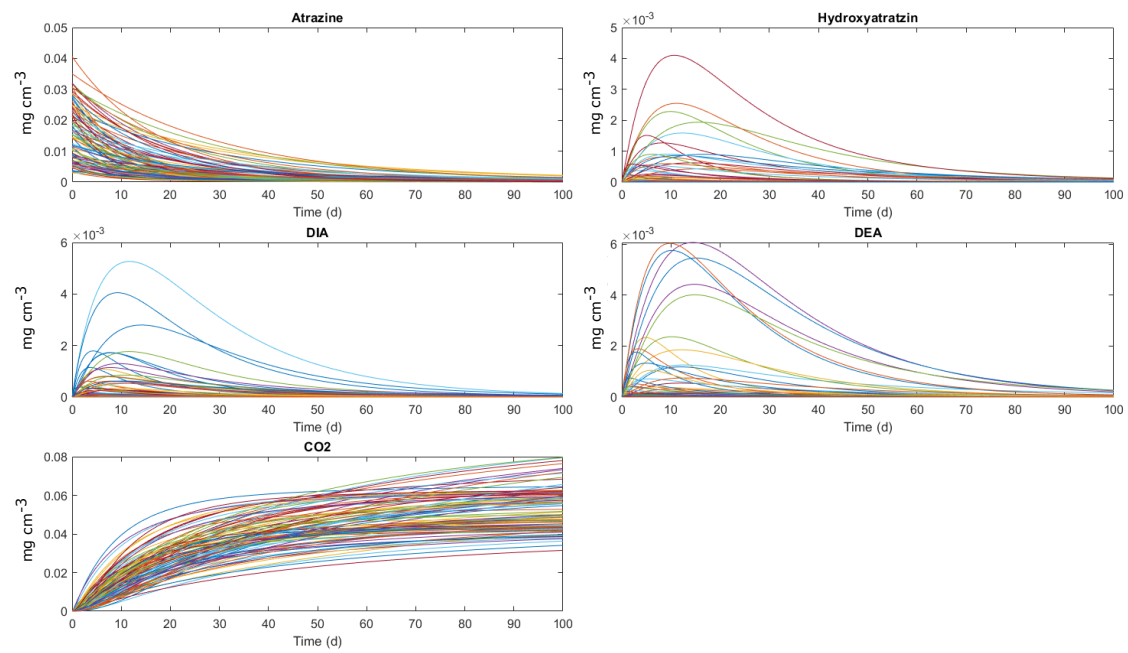


Figure S57: Model outputs (AT, metabolites (HY, DEA, DIA), and sinks (CA, CO<sub>2</sub>)) presented as spaghetti plots for model scenario M6. We showed 100 randomly selected outputs out of the 30,000 produced from the modified constraint-based parameter search algorithm

### 11.4.3 Energy distance (ED) robustness results

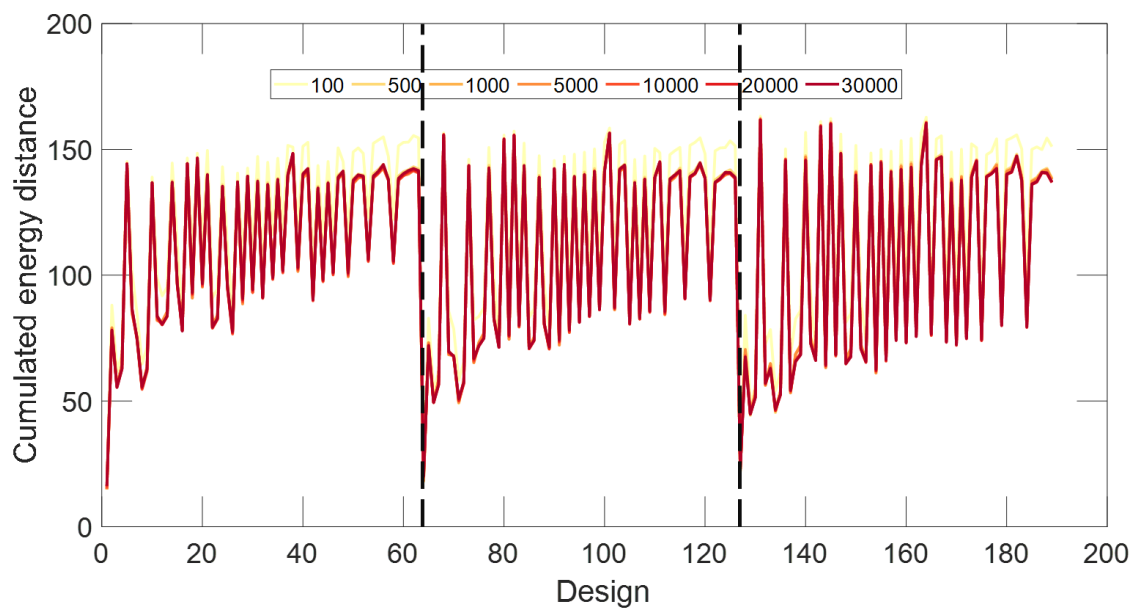


Figure S58: Robustness test. Energy distance values per design candidates, including AT, metabolites (HY, DEA, DIA), and sinks (CA, CO<sub>2</sub>) for the different ensemble sizes. Ensemble size of 100 differs from the bigger ensemble sizes, but an ensemble size of 1,000, 5,000, 10,000 seems to produce similar results

---

## References

- (1) Bonner, M. R.; Alavanja, M. C. R. Pesticides, human health, and food security. *Food and Energy Security* **2017**, *6*, 89–93, DOI: 10.1002/fes3.112.
- (2) Cooper, J.; Dobson, H. The benefits of pesticides to mankind and the environment. *Crop Protection* **2007**, *26*, 1337–1348, DOI: 10.1016/j.cropro.2007.03.022.
- (3) Damalas, C. A. Understanding benefits and risks of pesticide use. *Scientific Research and Essay* **2009**, *4*, 945–949.
- (4) Müller, K.; Deurer, M.; Northcott, G.; Clothier, B. E. Concept for assessing product-related pesticide. *New Zealand Plant Protection* **2010**, *46*, 39–46.
- (5) Popp, J.; Pet, K.; Nagy, J. Pesticide productivity and food security . A review. *Agronomy for Sustainable Development* **2013**, *33*, 243–255, DOI: 10.1007/s13593-012-0105-x.
- (6) Nicolopoulou-Stamati, P.; Maipas, S.; Kotampasi, C.; Stamatis, P.; Hens, L. Chemical Pesticides and Human Health: The Urgent Need for a New Concept in Agriculture, 2016.
- (7) Hassaan, M. A.; El Nemr, A. Pesticides pollution: Classifications, human health impact, extraction and treatment techniques. *The Egyptian Journal of Aquatic Research* **2020**, *46*, 207–220, DOI: 10.1016/j.ejar.2020.08.007.
- (8) World Health Organization *The WHO Recommended Classification of Pesticides by Hazard and Guidelines to Classification 2019*; techreport; Geneva: World Health Organization, 2019, **page** 92.
- (9) Food and Agriculture Organization of the United Nations (FAO) Pesticides Use, 2020.
- (10) Sharma, A.; Kumar, V.; Shahzad, B.; Tanveer, M.; Sidhu, G. P. S.; Handa, N.; Kohli, S. K.; Yadav, P.; Bali, A. S.; Parihar, R. D.; Dar, O. I.; Singh, K.; Jasrotia, S.; Bakshi, P.; Ramakrishnan, M.; Kumar, S.; Bhardwaj, R.; Thukral, A. K. Worldwide pesticide usage and its impacts on ecosystem. *SN Applied Sciences* **2019**, *1*, 1446, DOI: 10.1007/s42452-019-1485-1.



- 
- (11) Hedlund, J.; Longo, S. B.; York, R. Agriculture, Pesticide Use, and Economic Development: A Global Examination (1990–2014). *Rural Sociology* **2020**, *85*, 519–544, DOI: 10.1111/ruso.12303.
- (12) Schreinemachers, P.; Tipraqsa, P. Agricultural pesticides and land use intensification in high, middle and low income countries. *Food Policy* **2012**, *37*, 616–626, DOI: 10.1016/j.foodpol.2012.06.003.
- (13) Nwani, C. D.; Lakra, W. S.; Nagpure, N. S.; Kumar, R.; Kushwaha, B.; Srivastava, S. K. Toxicity of the herbicide atrazine: effects on lipid peroxidation and activities of antioxidant enzymes in the freshwater fish *Channa punctatus* (Bloch). *International journal of environmental research and public health* **2010**, *7*, 3298–3312, DOI: 10.3390/ijerph7083298.
- (14) Zheng, L.; Zhang, Y.; Yan, Z.; Zhang, J.; Li, L.; Zhu, Y.; Zhang, Y.; Zheng, X.; Wu, J.; Liu, Z. Derivation of predicted no-effect concentration and ecological risk for atrazine better based on reproductive fitness. *Ecotoxicology and Environmental Safety* **2017**, *142*, 464–470, DOI: 10.1016/j.ecoenv.2017.04.006.
- (15) Svingen, T.; Christiansen, S.; Taxvig, C.; Vinggaard, A. M. in Skinner, M. K. B. T. -. E. o. R. ( E. editor; Academic Press: Oxford, 2018, pages 624–628, DOI: 10.1016/B978-0-12-801238-3.64383-0.
- (16) Graymore, M.; Stagnitti, F.; Allinson, G. Impacts of atrazine in aquatic ecosystems. *Environment International* **2001**, *26*, 483–495, DOI: 10.1016/S0160-4120(01)00031-9.
- (17) Carson 1907-1964, R., *Silent spring*; 40th anniversary edition. Boston : Houghton Mifflin, 2002. ©1962: 0002.
- (18) Zhang, J.; Zhang, J.; Liu, R.; Gan, J.; Liu, J.; Liu, W. Endocrine-Disrupting Effects of Pesticides through Interference with Human Glucocorticoid Receptor. *Environmental Science & Technology* **2016**, *50*, 435–443, DOI: 10.1021/acs.est.5b03731.

- 
- (19) Ji, C.; Song, Q.; Chen, Y.; Zhou, Z.; Wang, P.; Liu, J.; Sun, Z.; Zhao, M. The potential endocrine disruption of pesticide transformation products (TPs): The blind spot of pesticide risk assessment. *Environment International* **2020**, *137*, 105490, DOI: 10.1016/j.envint.2020.105490.
- (20) Ghafouri-Khosrowshahi, A.; Ranjbar, A.; Mousavi, L.; Nili-Ahmadabadi, H.; Ghaffari, F.; Zeinvand-Lorestani, H.; Nili-Ahmadabadi, A. Chronic exposure to organophosphate pesticides as an important challenge in promoting reproductive health: A comparative study. *Journal of education and health promotion* **2019**, *8*, 149, DOI: 10.4103/jehp.jehp\_148\_19.
- (21) Sánchez, O. F.; Lin, L.; Bryan, C. J.; Xie, J.; Freeman, J. L.; Yuan, C. Profiling epigenetic changes in human cell line induced by atrazine exposure. *Environmental Pollution* **2020**, *258*, 113712, DOI: 10.1016/j.envpol.2019.113712.
- (22) VoPham, T.; Bertrand, K. A.; Hart, J. E.; Laden, F.; Brooks, M. M.; Yuan, J.-M.; Talbott, E. O.; Ruddell, D.; Chang, C.-C. H.; Weissfeld, J. L. Pesticide exposure and liver cancer: a review. *Cancer causes & control: CCC* **2017**, *28*, 177–190, DOI: 10.1007/s10552-017-0854-6.
- (23) Navarrete-Meneses, M. d. P.; Pérez-Vera, P. Pyrethroid pesticide exposure and hematological cancer: epidemiological, biological and molecular evidence. *Reviews on Environmental Health* **2019**, *34*, 197–210, DOI: 10.1515/reveh-2018-0070.
- (24) Pietrzak, D.; Kania, J.; Malina, G.; Kmiecik, E.; Wator, K. Pesticides from the EU First and Second Watch Lists in the Water Environment. *CLEAN – Soil, Air, Water* **2019**, *47*, 1800376, DOI: 10.1002/c1en.201800376.
- (25) Jablonowski, N. D.; Schäffer, A. Still present after all these years : persistence plus potential toxicity raise questions about the use of atrazine. *Environmental Science and Pollution Research* **2011**, *18*, 328–331, DOI: 10.1007/s11356-010-0431-y.
- (26) Singh, A.; Dhiman, N.; Kar, A. K.; Singh, D.; Purohit, M. P.; Ghosh, D.; Patnaik, S. Advances in controlled release pesticide formulations: Prospects to safer integrated

- pest management and sustainable agriculture. *Journal of Hazardous Materials* **2020**, 385, 121525, DOI: 10.1016/j.jhazmat.2019.121525.
- (27) Mfarrej, M. F. B.; Rara, F. M. Competitive, Sustainable Natural Pesticides. *Acta Ecologica Sinica* **2019**, 39, 145–151, DOI: 10.1016/j.chnaes.2018.08.005.
- (28) M. Domingues, P.; Santos, L. Essential oil of pennyroyal (*Mentha pulegium*): Composition and applications as alternatives to pesticides—New tendencies. *Industrial Crops and Products* **2019**, 139, 111534, DOI: 10.1016/j.indcrop.2019.111534.
- (29) Nehra, M.; Dilbaghi, N.; Marrazza, G.; Kaushik, A.; Sonne, C.; Kim, K.-H.; Kumar, S. Emerging nanobiotechnology in agriculture for the management of pesticide residues. *Journal of Hazardous Materials* **2021**, 401, 123369, DOI: 10.1016/j.jhazmat.2020.123369.
- (30) Sarmah, A. K.; Müller, K.; Ahmad, R. Fate and behaviour of pesticides in the agroecosystem—a review with a New Zealand perspective. *Soil Research* **2004**, 42, 125–154.
- (31) Gavrilesco, M. Fate of Pesticides in the Environment and its Bioremediation. *Engineering in Life Sciences* **2005**, 5, 497–526, DOI: 10.1002/elsc.200520098.
- (32) Das, S.; Hageman, K. J.; Taylor, M.; Michelsen-Heath, S.; Stewart, I. Fate of the organophosphate insecticide, chlorpyrifos, in leaves, soil, and air following application. *Chemosphere* **2020**, 243, 125194, DOI: 10.1016/j.chemosphere.2019.125194.
- (33) Potter, T. L.; Coffin, A. W. Assessing pesticide wet deposition risk within a small agricultural watershed in the Southeastern Coastal Plain (USA). *Science of The Total Environment* **2017**, 580, 158–167, DOI: 10.1016/j.scitotenv.2016.11.020.
- (34) Eker, G.; Tasdemir, Y. Atmospheric Deposition of Organochlorine Pesticides (OCPs): Species, Levels, Diurnal and Seasonal Fluctuations, Transfer Velocities. *Archives of Environmental Contamination and Toxicology* **2018**, 75, 625–633, DOI: 10.1007/s00244-018-0560-8.

- 
- (35) Chen, C.; Guo, W.; Ngo, H. H. Pesticides in stormwater runoff—A mini review. *Frontiers of Environmental Science & Engineering* **2019**, *13*, 72, DOI: 10.1007/s11783-019-1150-3.
- (36) Akay Demir, A. E.; Dilek, F. B.; Yetis, U. A new screening index for pesticides leachability to groundwater. *Journal of Environmental Management* **2019**, *231*, 1193–1202, DOI: 10.1016/j.jenvman.2018.11.007.
- (37) Li, Z. Spatiotemporal pattern models for bioaccumulation of pesticides in common herbaceous and woody plants. *Journal of Environmental Management* **2020**, *276*, 111334, DOI: 10.1016/j.jenvman.2020.111334.
- (38) Clasen, B.; Loro, V. L.; Murussi, C. R.; Tiecher, T. L.; Moraes, B.; Zanella, R. Bioaccumulation and oxidative stress caused by pesticides in *Cyprinus carpio* reared in a rice-fish system. *Science of The Total Environment* **2018**, *626*, 737–743, DOI: 10.1016/j.scitotenv.2018.01.154.
- (39) Bouhala, A.; Lahmar, H.; Benamira, M.; Moussi, A.; Trari, M. Photodegradation of Organophosphorus Pesticides in Honey Medium by Solar Light Irradiation. *Bulletin of Environmental Contamination and Toxicology* **2020**, *104*, 792–798, DOI: 10.1007/s00128-020-02858-1.
- (40) Masbou, J.; Drouin, G.; Payraudeau, S.; Imfeld, G. Carbon and nitrogen stable isotope fractionation during abiotic hydrolysis of pesticides. *Chemosphere* **2018**, *213*, 368–376, DOI: 10.1016/j.chemosphere.2018.09.056.
- (41) Wolfe, N. L.; Mingelgrin, U.; Miller, G. C. Abiotic Transformations in Water, Sediments, and Soil, 1990, DOI: doi:10.2136/sssabookser2.c5.
- (42) Hayatsu, M.; Tago, K.; Fukui, M.; Sekiya, E. in *Environmental Fate and Safety Management of Agrochemicals* ACS Symposium Series, **volume** 899; American Chemical Society: 2005, **pages** 8–82, DOI: doi:10.1021/bk-2005-0899.ch008.

- (43) Pagel, H.; Poll, C.; Ingwersen, J.; Kandeler, E.; Streck, T. Modeling coupled pesticide degradation and organic matter turnover: From gene abundance to process rates. *Soil Biology and Biochemistry* **2016**, *103*, 349–364, DOI: 10.1016/j.soilbio.2016.09.014.
- (44) Cycoń, M.; Mroziak, A.; Piotrowska-Seget, Z. Bioaugmentation as a strategy for the remediation of pesticide-polluted soil: A review. *Chemosphere* **2017**, *172*, 52–71, DOI: 10.1016/j.chemosphere.2016.12.129.
- (45) Madureira Barroso, G.; dos Santos, J. B.; de Oliveira, I. T.; Rocha Nunes, T. K. M.; Alves Ferreira, E.; Marinho Pereira, I.; Valadão Silva, D.; de Freitas Souza, M. Tolerance of Bradyrhizobium sp. BR 3901 to herbicides and their ability to use these pesticides as a nutritional source. *Ecological Indicators* **2020**, *119*, 106783, DOI: 10.1016/j.ecolind.2020.106783.
- (46) Ambrosini, R.; Ferrario, C.; Pittino, F.; Tagliaferri, I.; Gandolfi, I.; Bestetti, G.; Azzoni, R. S.; Diolaiuti, G. A.; Smiraglia, C.; Franzetti, A.; Villa, S. **in** *EGU General Assembly Conference Abstracts AA*(Department of Earth **andmore**, 2017, **page** 6142.
- (47) Hassen, W.; Neifar, M.; Cherif, H.; Mahjoubi, M.; Souissi, Y.; Raddadi, N.; Fava, F.; Cherif, A. Assessment of genetic diversity and bioremediation potential of pseudomonads isolated from pesticide-contaminated artichoke farm soils. *3 Biotech* **2018**, *8*, 263, DOI: 10.1007/s13205-018-1256-5.
- (48) Hussain, S.; Siddique, T.; Saleem, M.; Arshad, M.; Khalid, A. B. T. -. A. i. A. **in** *Academic Press* 2009; **volume** 102, **pages** 159–200, DOI: 10.1016/S0065-2113(09)01005-0.
- (49) Udiković-Kolić, N.; Scott, C.; Martin-Laurent, F. Evolution of atrazine-degrading capabilities in the environment. *Applied Microbiology and Biotechnology* **2012**, *96*, 1175–1189, DOI: 10.1007/s00253-012-4495-0.

- 
- (50) Robertson, B. K.; Alexander, M. Growth-linked and cometabolic biodegradation: Possible reason for occurrence or absence of accelerated pesticide biodegradation. *Pesticide Science* **1994**, *41*, 311–318, DOI: 10.1002/ps.2780410405.
- (51) Nzila, A. Update on the cometabolism of organic pollutants by bacteria. *Environmental Pollution* **2013**, *178*, 474–482, DOI: 10.1016/j.envpol.2013.03.042.
- (52) Barriada-Pereira, M.; González-Castro, M. J.; Muniategui-Lorenzo, S.; López-Mahía, P.; Prada-Rodríguez, D.; Fernández-Fernández, E. Organochlorine pesticides accumulation and degradation products in vegetation samples of a contaminated area in Galicia (NW Spain). *Chemosphere* **2005**, *58*, 1571–1578, DOI: 10.1016/j.chemosphere.2004.10.016.
- (53) González-Rodríguez, R. M.; Rial-Otero, R.; Cancho-Grande, B.; Simal-Gándara, J. Occurrence of fungicide and insecticide residues in trade samples of leafy vegetables. *Food Chemistry* **2008**, *107*, 1342–1347, DOI: 10.1016/j.foodchem.2007.09.045.
- (54) Liu, T.; Yuan, C.; Gao, Y.; Luo, J.; Yang, S.; Liu, S.; Zhang, R.; Zou, N. Exogenous salicylic acid mitigates the accumulation of some pesticides in cucumber seedlings under different cultivation methods. *Ecotoxicology and Environmental Safety* **2020**, *198*, 110680, DOI: 10.1016/j.ecoenv.2020.110680.
- (55) Silva, V.; Mol, H. G. J.; Zomer, P.; Tienstra, M.; Ritsema, C. J.; Geissen, V. Pesticide residues in European agricultural soils – A hidden reality unfolded. *Science of The Total Environment* **2019**, *653*, 1532–1545, DOI: 10.1016/j.scitotenv.2018.10.441.
- (56) Jablonowski, N. D.; Hamacher, G.; Martinazzo, R.; Langen, U.; Köppchen, S.; Hofmann, D.; Burauel, P. Metabolism and Persistence of Atrazine in Several Field Soils with Different Atrazine Application Histories. *Journal of Agricultural and Food Chemistry* **2010**, *58*, 12869–12877, DOI: 10.1021/jf103577j.

- 
- (57) Vonberg, D.; Vanderborght, J.; Cremer, N.; Pütz, T.; Herbst, M.; Vereecken, H. 20 years of long-term atrazine monitoring in a shallow aquifer in western Germany. *Water* **2014**, *50*, 294e306 Available, DOI: 10.1016/j.watres.2013.10.032.
- (58) Siek, M.; Paszko, T. Factors affecting coupled degradation and time-dependent sorption processes of tebuconazole in mineral soil profiles. *Science of The Total Environment* **2019**, *690*, 1035–1047, DOI: 10.1016/j.scitotenv.2019.06.409.
- (59) Mudhoo, A.; Garg, V. K. Sorption, Transport and Transformation of Atrazine in Soils, Minerals and Composts: A Review. *Pedosphere* **2011**, *21*, 11–25, DOI: 10.1016/S1002-0160(10)60074-4.
- (60) Shi, A.; Chakrawal, A.; Manzoni, S.; Fischer, B. M. C.; Nunan, N.; Herrmann, A. M. Substrate spatial heterogeneity reduces soil microbial activity. *Soil Biology and Biochemistry* **2021**, *152*, 108068, DOI: 10.1016/j.soilbio.2020.108068.
- (61) Coche, A.; Babey, T.; Rapaport, A.; Gonod, L. V.; Garnier, P.; Nunan, N.; de Dreuzy, J.-R. Bacterial density as an unexpected factor regulating decomposition by soil oligotrophs. *bioRxiv* **2020**, 2020.11.16.384735, DOI: 10.1101/2020.11.16.384735.
- (62) Ehrl, B. N.; Kundu, K.; Gharasoo, M.; Marozava, S.; Elsner, M. Rate-Limiting Mass Transfer in Micropollutant Degradation Revealed by Isotope Fractionation in Chemostat. *Environmental Science & Technology* **2019**, *53*, 1197–1205, DOI: 10.1021/acs.est.8b05175.
- (63) Kundu, K.; Marozava, S.; Ehrl, B.; Merl-Pham, J.; Griebler, C.; Elsner, M. Defining lower limits of biodegradation: atrazine degradation regulated by mass transfer and maintenance demand in *Arthrobacter aurescens* TC1. *The ISME Journal* **2019**, *13*, 2236–2251, DOI: 10.1038/s41396-019-0430-z.
- (64) LaRowe, D. E.; Van Cappellen, P. Degradation of natural organic matter: A thermodynamic analysis. *Geochimica et Cosmochimica Acta* **2011**, *75*, 2030–2042, DOI: 10.1016/j.gca.2011.01.020.

- 
- (65) FOCUS Generic guidance for estimating persistence and degradation kinetics from environmental fate studies on pesticides in EU registration. *Focus* **2011**, 1–436.
- (66) Basile, A.; Ciollaro, G.; Coppola, A. Hysteresis in soil water characteristics as a key to interpreting comparisons of laboratory and field measured hydraulic properties. *Water Resources Research* **2003**, *39*, DOI: 10.1029/2003WR002432.
- (67) Herrero-Hernández, E.; Marín-Benito, J. M.; Andrades, M. S.; Sánchez-Martín, M. J.; Rodríguez-Cruz, M. S. Field versus laboratory experiments to evaluate the fate of azoxystrobin in an amended vineyard soil. *Journal of Environmental Management* **2015**, *163*, 78–86, DOI: 10.1016/j.jenvman.2015.08.010.
- (68) Limpens, J.; Granath, G.; Aerts, R.; Heijmans, M. M. P. D.; Sheppard, L. J.; Bragazza, L.; Williams, B. L.; Rydin, H.; Bubier, J.; Moore, T.; Rochefort, L.; Mitchell, E. A. D.; Buttler, A.; van den Berg, L. J. L.; Gunnarsson, U.; Francez, A. -.; Gerdol, R.; Thormann, M.; Grosvernier, P.; Wiedermann, M. M.; Nilsson, M. B.; Hoosbeek, M. R.; Bayley, S.; Nordbakken, J. -.; Paulissen, M. P. C. P.; Hotes, S.; Breeuwer, A.; Ilomets, M.; Tomassen, H. B. M.; Leith, I.; Xu, B. Glasshouse vs field experiments: do they yield ecologically similar results for assessing N impacts on peat mosses? *New Phytologist* **2012**, *195*, 408–418, DOI: 10.1111/j.1469-8137.2012.04157.x.
- (69) Brüggemann, J.; Bizer, K. Laboratory experiments in innovation research: a methodological overview and a review of the current literature. *Journal of Innovation and Entrepreneurship* **2016**, *5*, 24, DOI: 10.1186/s13731-016-0053-9.
- (70) Kuhn, N. J.; Greenwood, P.; Fister, W. in Thornbush, M. J., Allen, C. D., Fitzpatrick, F. A. B. T. -.; D. i. E. S. P. **editors**; Elsevier: 2014; **volume** 18, **pages** 175–200, DOI: 10.1016/B978-0-444-63402-3.00011-X.
- (71) Almquist, J. Kinetic Models in Life Science – Contributions to Methods and Applications, phdthesis, CHALMERS UNIVERSITY OF TECHNOLOGY - Sweden, 2017, **page** 195.



- (72) Jablonowski, N. D.; Linden, A.; Köppchen, S.; Thiele, B.; Hofmann, D.; Mittelstaedt, W.; Pütz, T.; Burauel, P. Long-term persistence of various <sup>14</sup>C-labeled pesticides in soils. *Environmental Pollution* **2012**, *168*, 29–36, DOI: 10.1016/j.envpol.2012.04.022.
- (73) Marschmann, G. L.; Pagel, H.; Kügler, P.; Streck, T. Equifinality, sloppiness, and emergent structures of mechanistic soil biogeochemical models. *Environmental Modelling and Software* **2019**, *122*, 104518, DOI: 10.1016/j.envsoft.2019.104518.
- (74) Richter, O.; Diekkrüger, B.; Nörtersheuser, P., *Environmental Fate Modelling of Pesticides: From the Laboratory to the Field Scale*; John Wiley & Sons: Weinheim, 1996, **page** 293, DOI: 10.1002/9783527614790.
- (75) Li, M.; Qian, W. J.; Gao, Y.; Shi, L.; Liu, C. Functional Enzyme-Based Approach for Linking Microbial Community Functions with Biogeochemical Process Kinetics. *Environmental Science and Technology* **2017**, *51*, 11848–11857, DOI: 10.1021/acs.est.7b03158.
- (76) Kreft, J.-U.; Plugge, C. M.; Prats, C.; Leveau, J. H. J.; Zhang, W.; Hellweger, F. L. From Genes to Ecosystems in Microbiology: Modeling Approaches and the Importance of Individuality. *Frontiers in Microbiology* **2017**, *8*, 2299, DOI: 10.3389/fmicb.2017.02299.
- (77) Xu, X.; Zarecki, R.; Medina, S.; Ofaim, S.; Liu, X.; Chen, C.; Hu, S.; Brom, D.; Gat, D.; Porob, S.; Eizenberg, H.; Ronen, Z.; Jiang, J.; Freilich, S. Modeling microbial communities from atrazine contaminated soils promotes the development of biostimulation solutions. *The ISME Journal* **2019**, *13*, 494–508, DOI: 10.1038/s41396-018-0288-5.
- (78) Pagel, H.; Ingwersen, J.; Poll, C.; Kandeler, E.; Streck, T. Micro-scale modeling of pesticide degradation coupled to carbon turnover in the detritusphere: Model description and sensitivity analysis. *Biogeochemistry* **2014**, *117*, 185–204, DOI: 10.1007/s10533-013-9851-3.
- (79) Grovermann, C.; Schreinemachers, P.; Riwthong, S.; Berger, T. ‘Smart’ policies to reduce pesticide use and avoid income trade-offs: An agent-based model applied to Thai

- agriculture. *Ecological Economics* **2017**, *132*, 91–103, DOI: 10 . 1016 / j . ecolecon . 2016 . 09 . 031.
- (80) Misra, B. B.; Langefeld, C.; Olivier, M.; Cox, L. A. Integrated omics: tools, advances and future approaches. *Journal of Molecular Endocrinology* **2019**, *62*, R21–R45, DOI: 10 . 1530 / JME - 18 - 0055.
- (81) O'Donnell, S. T.; Ross, R. P.; Stanton, C. The Progress of Multi-Omics Technologies: Determining Function in Lactic Acid Bacteria Using a Systems Level Approach, 2020.
- (82) Reed, D. C.; Algar, C. K.; Huber, J. A.; Dick, G. J. Gene-centric approach to integrating environmental genomics and biogeochemical models. *Proceedings of the National Academy of Sciences* **2014**, *111*, 1879–1884, DOI: 10 . 1073 / pnas . 1313713111.
- (83) Louca, S.; Hawley, A. K.; Katsev, S.; Torres-Beltran, M.; Bhatia, M. P.; Kheirandish, S.; Michiels, C. C.; Capelle, D.; Lavik, G.; Doebeli, M.; Crowe, S. A.; Hallam, S. J. Integrating biogeochemistry with multiomic sequence information in a model oxygen minimum zone. *Proceedings of the National Academy of Sciences* **2016**, *113*, E5925–E5933, DOI: 10 . 1073 / pnas . 1602897113.
- (84) Song, H.-S.; Thomas, D. G.; Stegen, J. C.; Li, M.; Liu, C.; Song, X.; Chen, X.; Fredrickson, J. K.; Zachara, J. M.; Scheibe, T. D. Regulation-Structured Dynamic Metabolic Model Provides a Potential Mechanism for Delayed Enzyme Response in Denitrification Process. *Frontiers in microbiology* **2017**, *8*, 1866, DOI: 10 . 3389 / fmicb . 2017 . 01866.
- (85) Song, H.-S.; Stegen, J. C.; Graham, E. B.; Lee, J.-Y.; Garayburu-Caruso, V. A.; Nelson, W. C.; Chen, X.; Moulton, J. D.; Scheibe, T. D. Representing Organic Matter Thermodynamics in Biogeochemical Reactions via Substrate-Explicit Modeling, 2020.
- (86) McClure, R. S.; Lee, J.-Y.; Chowdhury, T. R.; Bottos, E. M.; White, R. A.; Kim, Y.-M.; Nicora, C. D.; Metz, T. O.; Hofmockel, K. S.; Jansson, J. K.; Song, H.-S. Integrated network modeling approach defines key metabolic responses of soil microbiomes to perturbations. *Scientific Reports* **2020**, *10*, 10882, DOI: 10 . 1038 / s41598 - 020 - 67878 - 7.

- 
- (87) Jensen, A.; Ford, W.; Fox, J.; Husic, A. Improving In-Stream Nutrient Routines in Water Quality Models Using Stable Isotope Tracers: A Review and Synthesis. *Transactions of the ASABE* **2018**, *61*, 139–157, DOI: 10.13031/trans.12545.
- (88) Ingwersen, J.; Poll, C.; Streck, T.; Kandeler, E. Micro-scale modelling of carbon turnover driven by microbial succession at a biogeochemical interface. *Soil Biology and Biochemistry* **2008**, *40*, 872–886, DOI: 10.1016/j.soilbio.2007.10.018.
- (89) Husic, A.; Fox, J.; Adams, E.; Pollock, E.; Ford, W.; Agouridis, C.; Backus, J. Quantification of nitrate fate in a karst conduit using stable isotopes and numerical modeling. *Water Research* **2020**, *170*, 115348, DOI: 10.1016/j.watres.2019.115348.
- (90) Schwartz, E.; Hayer, M.; Hungate, B. A.; Koch, B. J.; McHugh, T. A.; Mercurio, W.; Morrissey, E. M.; Soldanova, K. Stable isotope probing with  $^{18}\text{O}$ -water to investigate microbial growth and death in environmental samples. *Current Opinion in Biotechnology* **2016**, *41*, 14–18, DOI: 10.1016/j.copbio.2016.03.003.
- (91) Sprenger, M.; Tetzlaff, D.; Buttle, J.; Laudon, H.; Leistert, H.; Mitchell, C. P. J.; Snelgrove, J.; Weiler, M.; Soulsby, C. Measuring and Modeling Stable Isotopes of Mobile and Bulk Soil Water. *Vadose Zone Journal* **2018**, *17*, 170149, DOI: 10.2136/vzj2017.08.0149.
- (92) Denk, T. R. A.; Mohn, J.; Decock, C.; Lewicka-Szczebak, D.; Harris, E.; Butterbach-Bahl, K.; Kiese, R.; Wolf, B. The nitrogen cycle: A review of isotope effects and isotope modeling approaches. *Soil Biology and Biochemistry* **2017**, *105*, 121–137, DOI: 10.1016/j.soilbio.2016.11.015.
- (93) Ehrl, B. N.; Gharasoo, M.; Elsner, M. Isotope Fractionation Pinpoints Membrane Permeability as a Barrier to Atrazine Biodegradation in Gram-negative *Pseudomonas* sp. Nea-C. *Environmental Science & Technology* **2018**, *52*, 4137–4144, DOI: 10.1021/acs.est.7b06599.
- (94) Benes, B.; Guan, K.; Lang, M.; Long, S. P.; Lynch, J. P.; Marshall-Colón, A.; Peng, B.; Schnable, J.; Sweetlove, L. J.; Turk, M. J. Multiscale computational models can guide

- experimentation and targeted measurements for crop improvement. *The Plant Journal* **2020**, *103*, 21–31, DOI: 10.1111/tpj.14722.
- (95) Wagena, M. B.; Bhatt, G.; Buell, E.; Sommerlot, A. R.; Fuka, D. R.; Easton, Z. M. Quantifying model uncertainty using Bayesian multi-model ensembles. *Environmental Modelling & Software* **2019**, *117*, 89–99, DOI: 10.1016/j.envsoft.2019.03.013.
- (96) Renard, B.; Kavetski, D.; Kuczera, G.; Thyer, M.; Franks, S. W. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research* **2010**, *46*, DOI: 10.1029/2009WR008328.
- (97) La Cecilia, D. B. T. Comprehensive modeling of agrochemicals biodegradation in soil: A multidisciplinary approach to make informed choices to protect human health and the environment, phdthesis, 2019, **page** 198.
- (98) Chow, R.; Scheidegger, R.; Doppler, T.; Dietzel, A.; Fencia, F.; Stamm, C. A review of long-term pesticide monitoring studies to assess surface water quality trends. *Water Research X* **2020**, 100064, DOI: 10.1016/j.wroa.2020.100064.
- (99) Guthke (geb. Schöniger), A. Bayesian Assessment of Conceptual Uncertainty in Hydrosystem Modeling, phdthesis, 2016, **page** 144, DOI: 10.15496/publikation-9870.
- (100) Yao, K. Z.; Shaw, B. M.; Kou, B.; Mcauley, K. B.; Bacon, D. W. Modeling Ethylene / Butene Copolymerization with Multi-site Catalysts : Parameter Estimability and Experimental Design. *Polymer Reaction Engineering* **2003**, *11*, 563–588, DOI: 10.1081/PRE-120024426.
- (101) Malwade, A.; Nguyen, A.; Sadat-Mousavi, P.; Ingalls, B. P. Predictive modeling of a batch filter mating process. *Frontiers in Microbiology* **2017**, *8*, 1–11, DOI: 10.3389/fmicb.2017.00461.
- (102) Saltelli, A. Sensitivity Analysis for Importance Assessment. *Risk Analysis* **2002**, *22*, 579–590, DOI: 10.1111/0272-4332.00040.

- 
- (103) Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S., *Global Sensitivity Analysis. The Primer*; Wiley: West Sussex, 2008, **page** 305.
- (104) Hoops, S.; Hontecillas, R.; Abedi, V.; Leber, A.; Philipson, C.; Carbo, A.; Bassaganya-Riera, J. **in** *Computational Immunology, Models and Tools* Bassaganya-Riera, J. B. T. -. C. I., **editor**; Academic Press: 2016; **chapter** Chapter 5, **pages** 63–78, DOI: 10.1016/B978-0-12-803697-6.00005-9.
- (105) Link, K. G.; Stobb, M. T.; Paola, J. D.; Neeves, K. B.; Fogelson, L.; Sindi, S. S.; Leiderman, K. A local and global sensitivity analysis of a mathematical model of coagulation and platelet deposition under flow. *PLoS ONE* **2018**, *13*, 1–38, DOI: 10.1371/journal.pone.0200917.
- (106) Morris, M. D. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics* **1991**, *33*, 161–174.
- (107) Campolongo, F.; Cariboni, J.; Saltelli, A. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software* **2007**, *22*, 1509–1518, DOI: 10.1016/j.envsoft.2006.10.004.
- (108) López-Fidalgo, J.; Tommasi, C. **in** *The Mathematics of the Uncertain* Gil, E., Gil, E., Gil, J., Gil, M. Á., **editors**; Springer International Publishing: Cham, 2018, **pages** 253–263, DOI: 10.1007/978-3-319-73848-2\_24.
- (109) Höge, M.; Wöhling, T.; Nowak, W. A Primer for Model Selection: The Decisive Role of Model Complexity. *Water Resources Research* **2018**, *54*, 1688–1715, DOI: 10.1002/2017WR021902.
- (110) Diggle, P.; Lophaven, S. Bayesian Geostatistical Design. *Scandinavian Journal of Statistics* **2006**, *33*, 53–64.
- (111) Allgeier, J.; González-Nicolás, A.; Erdal, D.; Nowak, W.; Cirpka, O. A. A Stochastic Framework to Optimize Monitoring Strategies for Delineating Groundwater Divides, 2020, DOI: 10.3389/feart.2020.554845.

- 
- (112) Nowak, W.; Guthke, A. Entropy-Based Experimental Design for Optimal Model Discrimination in the Geosciences. *Entropy* **2016**, *18*, 409, DOI: 10.3390/e18110409.
- (113) Leube, P. C.; Geiges, A.; Nowak, W. Bayesian assessment of the expected data impact on prediction confidence in optimal sampling design. *Water Resources Research* **2012**, *48*, DOI: 10.1029/2010WR010137.
- (114) Székely, G. J.; Rizzo, M. L. Energy statistics: A class of statistics based on distances. *Journal of Statistical Planning and Inference* **2013**, *143*, 1249–1272, DOI: 10.1016/j.jspi.2013.03.018.
- (115) Mélykúti, B.; August, E.; Papachristodoulou, A.; El-Samad, H. Discriminating between rival biochemical network models: three approaches to optimal experiment design. *BMC Systems Biology* **2010**, *4*, 38, DOI: 10.1186/1752-0509-4-38.
- (116) Weiss, S. M.; Indurkha, N. Rule-Based Machine Learning Methods for Functional Prediction. *J. Artif. Int. Res.* **1995**, *3*, 383–403.
- (117) Vrugt, J. A. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environmental Modelling and Software* **2016**, *75*, 273–316, DOI: 10.1016/j.envsoft.2015.08.013.
- (118) Gharari, S.; Shafiei, M.; Hrachowitz, M.; Kumar, R.; Fenicia, F.; Gupta, H. V.; Savenije, H. H. G. A constraint-based search algorithm for parameter identification of environmental models. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 4861–4870, DOI: 10.5194/hess-18-4861-2014.
- (119) Grossmann, K. Mediation of Herbicide Effects by Hormone Interactions. *Journal of Plant Growth Regulation* **2003**, *22*, 109–122, DOI: 10.1007/s00344-003-0020-0.
- (120) Grossmann, K. Auxin herbicides: current status of mechanism and mode of action. *Pest Management Science* **2010**, *66*, 113–120, DOI: 10.1002/ps.1860.

- (121) Caux, P.-Y.; Kent, R. A.; Bergeron, V.; Fan, G. T.; MacDonald, D. D. Environmental fate and effects of MCPA: A Canadian perspective. *Critical Reviews in Environmental Science and Technology* **1995**, *25*, 313–376, DOI: 10.1080/10643389509388483.
- (122) Tiedje, J. M.; Duxbury, J. M.; Alexander, M.; Dawson, J. E. 2,4-D Metabolism: Pathway of Degradation of Chlorocatechols by *Arthrobacter* sp. *Journal of Agricultural and Food Chemistry* **1969**, *17*, 1021–1026, DOI: 10.1021/jf60165a037.
- (123) Kilpi, S.; Backström, V.; Korhola, M. Degradation of 2-Methyl-4-Chlorophenoxyacetic acid (MCPA), 2,4-Dichlorophenoxyacetic acid (2,4-D), Benzoic acid and Salicylic acid by *Pseudomonas* sp. HV3. *FEMS Microbiology Letters* **1980**, *8*, 177–182.
- (124) Fu, F. F.; Xiao, L. X.; Wang, W.; Xu, X. Q.; Xu, L. J.; Qi, G. M.; Chen, G. N. Study on the degradation of 2,4-dichlorophenoxyacetic acid (2,4-D) and 2-methyl-4-chlorophenoxyacetic sodium (MCPA sodium) in natural agriculture-soils of Fuzhou, China using capillary electrophoresis. *Science of the Total Environment* **2009**, *407*, 1998–2003, DOI: 10.1016/j.scitotenv.2008.11.023.
- (125) Evans, W. C.; Smith, B. S. W.; Fernley, H. N.; Davies, J. I. Bacterial metabolism of 2,4-dichlorophenoxyacetate. *Biochemical Journal* **2015**, *122*, 543–551, DOI: 10.1042/bj1220543.
- (126) Leveau, J. H. J.; König, F.; Fuchslin, H.; Werlen, C.; Van Der Meer, J. R. Dynamics of multigene expression during catabolic adaptation of *Ralstonia eutropha* JMP134 (pJP4) to the herbicide 2,4-dichlorophenoxyacetate. *Molecular Microbiology* **1999**, *33*, 396–406.
- (127) Itoh, K.; Kanda, R.; Sumita, Y.; Kim, H.; Kamagata, Y.; Suyama, K.; Yamamoto, H.; Hausinger, R. P.; Tiedje, J. M. *tfdA*-like genes in 2,4-dichlorophenoxyacetic acid-degrading bacteria belonging to the *Bradyrhizobium*-*Agromonas*-*Nitrobacter*-*Afipia* cluster in  $\alpha$ -Proteobacteria. *Applied and Environmental Microbiology* **2002**, *68*, 3449–3454, DOI: 10.1128/AEM.68.7.3449-3454.2002.

- 
- (128) Bælum, J.; Nicolaisen, M. H.; Holben, W. E.; Strobel, B. W.; Sørensen, J.; Jacobsen, C. S. Direct analysis of *tfdA* gene expression by indigenous bacteria in phenoxy acid amended agricultural soil. *ISME Journal* **2008**, *2*, 677–687, DOI: 10.1038/ismej.2008.21.
- (129) Itoh, K.; Tashiro, Y.; Uobe, K.; Kamagata, Y.; Suyama, K.; Yamamoto, H. Root Nodule Bradyrhizobium spp. Harbor *tfdA* and *cadA*, Homologous with Genes Encoding 2,4-Dichlorophenoxyacetic Acid-Degrading Proteins. *Applied and Environmental Microbiology* **2004**, *70*, 2110–2118, DOI: 10.1128/AEM.70.4.2110-2118.2004.
- (130) Liu, Y. J.; Liu, S. J.; Drake, H. L.; Horn, M. A. Consumers of 4-chloro-2-methylphenoxyacetic acid from agricultural soil and rhizosphere harbor *cadA*, *r/sdpA*, and *tfdA*-like gene encoding oxygenases. *FEMS Microbiology Ecology* **2013**, *86*, 114–129, DOI: 10.1111/1574-6941.12144.
- (131) Ditterich, F.; Poll, C.; Pagel, H.; Babin, D.; Smalla, K.; Horn, M. A.; Streck, T.; Kandeler, E. Succession of bacterial and fungal 4-chloro-2-methylphenoxyacetic acid degraders at the soil – litter interface. *FEMS Microbiology Ecology* **2013**, *86*, 85–100, DOI: 10.1111/1574-6941.12131.
- (132) López-Muñoz, M. J.; Aguado, J.; Revilla, A. Photocatalytic removal of s-triazines: Evaluation of operational parameters. *Catalysis Today* **2011**, *161*, 153–162, DOI: 10.1016/j.cattod.2010.10.076.
- (133) Yanze-Kontchou, C.; Gschwind, N. Mineralization of the herbicide atrazine as a carbon source by a *Pseudomonas* strain. *Applied and Environmental Microbiology* **1994**, *60*, 4297 LP –4302.
- (134) Mandelbaum, R. T.; Allan, D. L.; Wackett, L. P. Isolation and characterization of a *Pseudomonas* sp. that mineralizes the s-triazine herbicide atrazine. *Applied and Environmental Microbiology* **1995**, *61*, 1451–1457, DOI: 10.1080/10408399209527575.



- (135) Fernandes, A. F. T.; da Silva, M. B. P.; Martins, V. V.; Miranda, C. E. S.; Stehling, E. G. Isolation and characterization of a *Pseudomonas aeruginosa* from a virgin Brazilian Amazon region with potential to degrade atrazine. *Environmental Science and Pollution Research* **2014**, *21*, 13974–13978, DOI: 10.1007/s11356-014-3316-7.
- (136) Mongodin, E. F.; Shapir, N.; Daugherty, S. C.; DeBoy, R. T.; Emerson, J. B.; Shvartzbeyn, A.; Radune, D.; Vamathevan, J.; Riggs, F.; Grinberg, V.; Khouri, H.; Wackett, L. P.; Nelson, K. E.; Sadowsky, M. J. Secrets of Soil Survival Revealed by the Genome Sequence of *Arthrobacter aurescens* TC1. *PLOS Genetics* **2006**, *2*, e214.
- (137) García-González, V.; Govantes, F.; Porrúa, O.; Santero, E. Regulation of the *Pseudomonas* sp. strain ADP cyanuric acid degradation operon. *Journal of bacteriology* **2005**, *187*, 155–167, DOI: 10.1128/JB.187.1.155-167.2005.
- (138) Galíndez-Nájera, S. P.; Llamas-Martínez, M. A.; Ruiz-Ordaz, N.; Juárez-Ramírez, C.; Mondragón-Parada, M. E.; Ahuatzí-Chacón, D.; Galíndez-Mayer, J. Cyanuric acid biodegradation by a mixed bacterial culture of *Agrobacterium tumefaciens* and *Acinetobacter* sp. in a packed bed biofilm reactor. *Journal of industrial microbiology & biotechnology* **2009**, *36*, 275–284, DOI: 10.1007/s10295-008-0496-5.
- (139) Dutta, A.; Vasudevan, V.; Nain, L.; Singh, N. Characterization of bacterial diversity in an atrazine degrading enrichment culture and degradation of atrazine, cyanuric acid and biuret in industrial wastewater. *Journal of Environmental Science and Health, Part B* **2016**, *51*, 24–34, DOI: 10.1080/03601234.2015.1080487.
- (140) Topp, E.; Zhu, H.; Nour, S. M.; Houot, S.; Lewis, M.; Cuppels, D. Characterization of an atrazine-degrading *Pseudaminobacter* sp. isolated from Canadian and French agricultural soils. *Applied and environmental microbiology* **2000**, *66*, 2773–2782, DOI: 10.1128/aem.66.7.2773-2782.2000.

- (141) Smith, D.; Alvey, S.; Crowley, D. E. Cooperative catabolic pathways within an atrazine-degrading enrichment culture isolated from soil. *FEMS Microbiology Ecology* **2005**, *53*, 265–273, DOI: 10.1016/j.femsec.2004.12.011.
- (142) Smith, D.; Crowley, D. E. Contribution of ethylamine degrading bacteria to atrazine degradation in soils. *FEMS Microbiology Ecology* **2006**, *58*, 271–277, DOI: 10.1111/j.1574-6941.2006.00168.x.
- (143) Seybold, C. A.; Mersie, W. Adsorption and Desorption of Atrazine, Deethylatrazine, Deisopropylatrazine, Hydroxyatrazine, and Metolachlor in Two Soils from Virginia. *Journal of Environmental Quality* **1996**, *25*, 1179–1185, DOI: 10.2134/jeq1996.00472425002500060002x.
- (144) Lerch Robert N.; Thurman E. Michael; Blanchard, P. E. Hydroxyatrazine in soils and sediments. *Environmental Toxicology and Chemistry* **1999**, *18*, 2161–2168, DOI: 10.1002/etc.5620181007.
- (145) Solomon, R. D. J.; Kumar, A.; Satheeja Santhi, V. Atrazine biodegradation efficiency, metabolite detection, and trzD gene expression by enrichment bacterial cultures from agricultural soil. *Journal of Zhejiang University. Science. B* **2013**, *14*, 1162–1172, DOI: 10.1631/jzus.B1300001.
- (146) Kolekar, P. D.; Phugare, S. S.; Jadhav, J. P. Biodegradation of atrazine by *Rhodococcus* sp. BCH2 to N-isopropylammelide with subsequent assessment of toxicity of biodegraded metabolites. *Environmental science and pollution research international* **2014**, *21*, 2334–2345, DOI: 10.1007/s11356-013-2151-6.
- (147) Brand, W. A.; Assonov, S. S.; Coplen, T. B. Correction for the  $^{17}\text{O}$  interference in  $\delta(^{13}\text{C})$  measurements when analyzing  $\text{CO}_2$  with stable isotope mass spectrometry (IUPAC Technical Report). *Pure and Applied Chemistry* **2010**, *82*, 1719–1733, DOI: 10.1351/PAC-REP-09-01-05.

- 
- (148) Nossent, J.; Elsen, P.; Bauwens, W. Sobol' sensitivity analysis of a complex environmental model. *Environmental Modelling & Software* **2011**, *26*, 1515–1525, DOI: 10.1016/j.envsoft.2011.08.010.
- (149) Arias-Estévez, M.; López-Periago, E.; Martínez-Carballo, E.; Simal-Gándara, J.; Mejuto, J.-C.; García-Río, L. The mobility and degradation of pesticides in soils and the pollution of groundwater resources. *Agriculture, Ecosystems and Environment* **2008**, *123*, 247–260, DOI: 10.1016/j.agee.2007.07.011.
- (150) Örneby, K.; Jonsson, A.; Stenström, J. A new concept for reduction of diffuse contamination by simultaneous application of pesticide and pesticide-degrading microorganisms. *Biodegradation* **2010**, *21*, 21–29, DOI: 10.1007/s10532-009-9278-7.
- (151) Vryzas, Z. ScienceDirect Pesticide fate in soil-sediment-water environment in relation to contamination preventing actions. *Current Opinion in Environmental Science & Health* **2018**, *4*, 5–9, DOI: 10.1016/j.coesh.2018.03.001.
- (152) Huang, Y.; Xiao, L.; Li, F.; Xiao, M.; Lin, D.; Long, X.; Wu, Z. Microbial Degradation of Pesticide Residues and an Emphasis on the Degradation of Cypermethrin and 3-phenoxy Benzoic Acid: A Review. *Molecules* **2018**, *23*, 2313, DOI: 10.3390/molecules23092313.
- (153) Kästner, M.; Nowak, K. M.; Miltner, A.; Trapp, S.; Schäffer, A. Classification and modelling of nonextractable residue (NER) formation of xenobiotics in soil - A synthesis. *Critical Reviews in Environmental Science and Technology* **2014**, *44*, 2107–2171, DOI: 10.1080/10643389.2013.828270.
- (154) Fenner, K.; Canonica, S.; Wackett, L. P.; Elsner, M. Evaluating Pesticide Degradation in the Environment : Blind Spots and Emerging Opportunities. *Science* **2013**, *341*, 752–758, DOI: 10.1126/science.1236281.

- 
- (155) Li, L.; Maher, K.; Navarre-sitchler, A.; Druhan, J.; Meile, C.; Lawrence, C.; Moore, J.; Perdrial, J.; Sullivan, P.; Thompson, A.; Jin, L.; Bolton, E. W.; Brantley, S. L.; Dietrich, W. E.; Mayer, K. U.; Steefel, C. I.; Valocchi, A.; Zachara, J.; Kocar, B.; Mcintosh, J.; Tutolo, B. M.; Kumar, M.; Sonnenthal, E.; Bao, C.; Beisman, J. Expanding the role of reactive transport models in critical zone processes. *Earth-Science Reviews* **2017**, *165*, 280–301, DOI: 10.1016/j.earscirev.2016.09.001.
- (156) Thullner, M.; Regnier, P. Microbial Controls on the Biogeochemical Dynamics in the Subsurface. *Reviews in Mineralogy & Geochemistry* **2019**, *85*, 265–302.
- (157) Song, H.-S.; Liu, C. Dynamic Metabolic Modeling of Denitrifying Bacterial Growth: The Cybernetic Approach. *Industrial & Engineering Chemistry Research* **2015**, *54*, 10221–10227, DOI: 10.1021/acs.iecr.5b01615.
- (158) Pagel, H.; Ingwersen, J.; Poll, C.; Kandeler, E.; Streck, T. Micro-scale modeling of pesticide degradation coupled to carbon turnover in the detritusphere: Model description and sensitivity analysis. *Biogeochemistry* **2014**, *117*, 185–204, DOI: 10.1007/s10533-013-9851-3.
- (159) Duah Yentumi, S.; Kuwatsuka, S. Microbial degradation of benthocarb, MCPA and 2,4-d herbicides in perfused soils amended with organic matter and chemical fertilizers. *Soil Science and Plant Nutrition* **1982**, *28*, 19–26, DOI: 10.1080/00380768.1982.10432368.
- (160) Chaudhry, G. R.; Huang, G. H. Isolation and characterization of a new plasmid from a *Flavobacterium* sp. which carries the genes for degradation of 2,4-dichlorophenoxyacetate. *Journal of Bacteriology* **1988**, *170*, 3897–3902, DOI: 10.1128/jb.170.9.3897-3902.1988.
- (161) Pieper, D. H.; Reineke, W.; Engesser, K.-h.; Knackmuss, H.-j. Metabolism of 2,4-dichlorophenoxyacetic acid, 4-chloro-2-methylphenoxyacetic acid and 2-

- methylphenoxyacetic acid by *Alcaligenes eutrophus* JMP 134. *Archives of Microbiology* **1988**, *150*, 95–102.
- (162) Nielsen, T. K.; Xu, Z.; Gözdereliler, E.; Aamand, J.; Hansen, L. H.; Sørensen, S. R. Novel insight into the genetic context of the *cadAB* genes from a 4-chloro-2-methylphenoxyacetic acid-degrading *Sphingomonas*. *PLoS ONE* **2013**, *8*, DOI: 10.1371/journal.pone.0083346.
- (163) Lennon, J. T.; Jones, S. E. Microbial seed banks : the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology* **2011**, *9*, 119–130, DOI: 10.1038/nrmicro2504.
- (164) Blagodatskaya, E.; Kuzyakov, Y. Active microorganisms in soil: Critical review of estimation criteria and approaches. *Soil Biology and Biochemistry* **2013**, *67*, 192–211, DOI: 10.1016/j.soilbio.2013.08.024.
- (165) Carini, P.; Marsden, P. J.; Leff, J. W.; Morgan, E. E.; Strickland, M. S.; Fierer, N. Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nature Microbiology* **2016**, *2*, 1–6, DOI: 10.1038/nmicrobiol.2016.242.
- (166) Lennon, J.; Muscarella, M.; Placella, S.; Lehmkuhl, B. How, When, and Where Relic DNA Affects Microbial Diversity. **2018**, *9*, 1–14.
- (167) Couradeau, E.; Sasse, J.; Goudeau, D.; Nath, N.; Hazen, T. C.; Bowen, B. P.; Malmstrom, R. R.; Northen, T. R. Study of Oak Ridge soils using BONCAT-FACS-Seq reveals that a large fraction of the soil microbiome is active. *bioRxiv* **2018**, DOI: 10.1101/404087.
- (168) Maghrebi, M.; Jankovic, I.; Fiori, A.; Dagan, G. Effective retardation factor for transport of reactive solutes in highly heterogeneous porous formations. *Water Resources Research* **2013**, *49*, 8600–8604, DOI: 10.1002/2013WR014429.
- (169) Stolpovsky, K.; Martinez-Lavanchy, P.; Heipieper, H. J.; Van Cappellen, P.; Thullner, M. Incorporating dormancy in dynamic microbial community models. *Ecological Modelling* **2011**, *222*, 3092–3102, DOI: 10.1016/j.ecolmodel.2011.07.006.

- (170) Mellage, A.; Eckert, D.; Grösbacher, M.; Inan, A. Z.; Cirpka, O. A.; Griebler, C. Dynamics of Suspended and Attached Aerobic Toluene Degraders in Small-Scale Flow-through Sediment Systems under Growth and Starvation Conditions. *Environmental Science and Technology* **2015**, *49*, 7161–7169, DOI: 10.1021/es5058538.
- (171) Baelum, J.; Henriksen, T.; Hansen, H. C. B.; Jacobsen, C. S. Degradation of 4-chloro-2-methylphenoxyacetic acid in top- and subsoil is quantitatively linked to the class III tfdA gene. *Applied and environmental microbiology* **2006**, *72*, 1476–1486, DOI: 10.1128/AEM.72.2.1476-1486.2006.
- (172) Paulin, M. M.; Nicolaisen, M. H.; Sørensen, J. Abundance and expression of enantioselective rdpA and sdpA dioxygenase genes during degradation of the racemic herbicide (R,S)-2-(2,4-dichlorophenoxy) propionate in soil. *Applied and Environmental Microbiology* **2010**, *76*, 2873–2883, DOI: 10.1128/AEM.02270-09.
- (173) Lee, T. H.; Kurata, S.; Nakatsu, C. H.; Kamagata, Y. Molecular analysis of bacterial community based on 16S rDNA and functional genes in activated sludge enriched with 2,4-dichlorophenoxyacetic acid (2,4-D) under different cultural conditions. *Microbial ecology* **2005**, *49*, 151–162, DOI: 10.1007/s00248-003-1035-6.
- (174) Kitagawa, W.; Takami, S.; Miyauchi, K.; Masai, E.; Kamagata, Y.; Tiedje, J. M.; Fukuda, M. Novel 2,4-Dichlorophenoxyacetic Acid Degradation Genes from Oligotrophic *Bradyrhizobium* sp. Strain HW13 Isolated from a Pristine Environment. *Journal of Bacteriology* **2002**, *184*, 509 LP –518, DOI: 10.1128/JB.184.2.509-518.2002.
- (175) Müller, T. A.; Fleischmann, T.; van der Meer, J. R.; Kohler, H.-P. E. Purification and Characterization of Two Enantioselective  $\alpha$ -Ketoglutarate-Dependent Dioxygenases, RdpA and SdpA, from *Sphingomonas herbicidovorans* MH. *Applied and Environmental Microbiology* **2006**, *72*, 4853 LP –4861, DOI: 10.1128/AEM.02758-05.
- (176) Rodríguez-Cruz, M. S.; Baelum, J.; Shaw, L. J.; Sørensen, S. R.; Shi, S.; Aspray, T.; Jacobsen, C. S.; Bending, G. D. Biodegradation of the herbicide mecoprop-p with soil depth and

- its relationship with class III *tfdA* genes. *Soil Biology and Biochemistry* **2010**, *42*, 32–39, DOI: 10.1016/j.soilbio.2009.09.018.
- (177) Ingalls, B. P., *Mathematical Modeling in Systems Biology: An Introduction*; The MIT Press; MIT Press: 2013.
- (178) Bhasi, K.; Forrest, A.; Ramanathan, M. SPLINDID: A semi-parametric, model-based method for obtaining transcription rates and gene regulation parameters from genomic and proteomic expression profiles. *Bioinformatics* **2005**, *21*, 3873–3879, DOI: 10.1093/bioinformatics/bti624.
- (179) Nowak, K. M.; Miltner, A.; Gehre, M.; Schäffer, A.; Kastner, M. Formation and fate of bound residues from microbial biomass during 2, 4-D degradation in soil. *Environmental Science and Technology* **2011**, *45*, 999–1006, DOI: 10.1021/es103097f.
- (180) Senn, H.; Lendenmann, U.; Snozzi, M.; Hamer, G.; Egli, T. The growth of *Escherichia coli* in glucose-limited chemostat cultures: a re-examination of the kinetics. *Biochimica et Biophysica Acta* **1994**, *1201*, 424–436.
- (181) Blagodatsky, S. A.; Richter, O. Microbial growth in soil and nitrogen turnover: a theoretical model considering the activity state of microorganisms. *Soil Biology and Biochemistry* **1998**, *30*, 1743–1755, DOI: 10.1016/S0038-0717(98)00028-5.
- (182) Thorstensen, C. W.; Lode, O.; Eklo, O. M.; Christiansen, A. Sorption of Bentazone, Dichlorprop, MCPA, and Propiconazole in Reference Soils from Norway. *Journal of Environmental Quality* **2001**, *30*, 2046–2052, DOI: 10.2134/jeq2001.2046.
- (183) Vrede, K.; Heldal, M.; Norland, S.; Bratbak, G. Elemental Composition (C, N, P) and Cell Volume of Exponentially Growing and. *Applied and Environmental Microbiology* **2002**, *68*, 2965–2971, DOI: 10.1128/AEM.68.6.2965.
- (184) Westendorf, A.; Benndorf, D.; Pribyl, T.; Harms, H.; Müller, R. H. Kinetic traits and enzyme form patterns of (R)-2-(2,4-dichlorophenoxy)propionate/ $\alpha$ -ketoglutarate dioxygenase

- (RdpA) after expression in different bacterial strains. *Engineering in Life Sciences* **2006**, *6*, 552–559, DOI: 10.1002/elsc.200620165.
- (185) Prinz, H. Hill coefficients, dose-response curves and allosteric mechanisms. *Journal of chemical biology* **2010**, *3*, 37–44, DOI: 10.1007/s12154-009-0029-3.
- (186) Füchslin, H. P.; Schneider, C.; Egli, T. In glucose-limited continuous culture the minimum substrate concentration for growth,  $s_{min}$ , is crucial in the competition between the enterobacterium *Escherichia coli* and *Chelatobacter heintzii*, an environmentally abundant bacterium. *International Society for Microbial Ecology* **2012**, *6*, 777–789, DOI: 10.1038/ismej.2011.143.
- (187) Taniguchi, Y.; Choi, P. J.; Li, G.-W.; Chen, H.; Babu, M.; Hearn, J.; Emili, A.; Xie, X. S. Quantifying *E. coli* Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science* **2010**, *329*, 533–538, DOI: 10.1126/science.1188308.
- (188) Fuchs, G.; Voichek, Y.; Benjamin, S.; Gilad, S.; Amit, I.; Oren, M. 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *Genome Biology* **2014**, *15*, R69, DOI: 10.1186/gb-2014-15-5-r69.
- (189) Laalami, S.; Zig, L.; Putzer, H. Initiation of mRNA decay in bacteria. *Cellular and molecular life sciences : CMLS* **2014**, *71*, 1799–1828, DOI: 10.1007/s00018-013-1472-4.
- (190) Mellage, A.; Eckert, D.; Grösbacher, M.; Inan, A. Z.; Cirpka, O. A.; Griebler, C. Dynamics of Suspended and Attached Aerobic Toluene Degraders in Small-Scale Flow-through Sediment Systems under Growth and Starvation Conditions. *Environmental Science and Technology* **2015**, *49*, 7161–7169, DOI: 10.1021/es5058538.
- (191) Ashyraliyev, M.; Fomekong-Nanfack, Y.; Kaandorp, J. A.; Blom, J. G. Systems biology: Parameter estimation for biochemical models. *FEBS Journal* **2009**, *276*, 886–902, DOI: 10.1111/j.1742-4658.2008.06844.x.
- (192) Gelman, A.; Rubin, D. B. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* **1992**, *7*, 457–472, DOI: 10.1214/ss/1177011136.



- 
- (193) Akaike, H. A New Look at the Statistical Model Identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723, DOI: 10.1109/TAC.1974.1100705.
- (194) Schwarz, G. Estimating the Dimension of a Model. *The Annals of Statistics* **1978**, *6*, 461–464, DOI: 10.1214/aos/1176344136.
- (195) Emmert-Streib, F.; Dehmer, M. Evaluation of Regression Models: Model Assessment, Model Selection and Generalization Error, 2019, DOI: 10.3390/make1010032.
- (196) Fryer, J.; Kirkland, K. Field experiments to investigate long-term effects of repeated applications of MCPA, tri-allate, simazine and linuron: Report after 6 years. *Weed Research* **1970**, *10*, 133–158, DOI: 10.1111/j.1365-3180.1970.tb00934.x.
- (197) Boivin, A.; Amellal, S.; Schiavon, M.; Van Genuchten, M. T. 2,4-Dichlorophenoxyacetic acid (2,4-D) sorption and degradation dynamics in three agricultural soils. *Environmental Pollution* **2005**, *138*, 92–99, DOI: 10.1016/j.envpol.2005.02.016.
- (198) Mierzejewska, E.; Baran, A.; Tankiewicz, M.; Urbaniak, M. Removal and ecotoxicity of 2,4-D and MCPA in microbial cultures enriched with structurally-similar plant secondary metabolites. *Water (Switzerland)* **2019**, *11*, 1–16, DOI: 10.3390/w11071451.
- (199) Ingalls, B. Sensitivity Analysis : from model parameters to system behaviour Local Sensitivity Analysis : Derivation. *Essays in biochemistry* **2008**, *45*, 177–193.
- (200) Zi, Z. Sensitivity analysis approaches applied to systems biology models. *IET Systems Biology* **2011**, *5*, 336–346, DOI: 10.1049/iet-syb.2011.0015.
- (201) Pianosi, F.; Sarrazin, F.; Wagener, T. A Matlab toolbox for Global Sensitivity Analysis. *Environmental Modelling and Software* **2015**, *70*, 80–85, DOI: 10.1016/j.envsoft.2015.04.009.
- (202) Pianosi, F.; Wagener, T. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling and Software* **2015**, *67*, 1–11, DOI: 10.1016/j.envsoft.2015.01.004.

- 
- (203) Jablonowski, N. D.; Köppchen, S.; Hofmann, D.; Schäffer, A.; Burauel, P. Persistence of <sup>14</sup>C-labeled atrazine and its residues in a field lysimeter soil after 22 years. *Environmental Pollution* **2009**, *157*, 2126–2131, DOI: 10.1016/j.envpol.2009.02.004.
- (204) Jensen, P. H.; Hansen, H. C. B.; Rasmussen, J.; Jacobsen, O. S. Sorption-controlled degradation kinetics of MCPA in soil. *Environmental Science and Technology* **2004**, *38*, 6662–6668, DOI: 10.1021/es0494095.
- (205) Khan, S. U. Equilibrium and kinetic studies of the adsorption of 2, 4-D and picloram on humic acid. *Canadian Journal of Soil Science* **1973**, *53*, 429–434, DOI: 10.4141/cjss73-060.
- (206) Benoit, P.; Barriuso, E.; Houot, S.; Calvet, R. Influence of the nature of soil organic matter on the sorption-desorption of 4-chlorophenol, 2,4-dichlorophenol and the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D). *European Journal of Soil Science* **1996**, *47*, 567–578, DOI: 10.1111/j.1365-2389.1996.tb01856.x.
- (207) Hameed, B. H.; Salman, J. M.; Ahmad, A. L. Adsorption isotherm and kinetic modeling of 2,4-D pesticide on activated carbon derived from date stones. *Journal of Hazardous Materials* **2009**, *163*, 121–126, DOI: 10.1016/j.jhazmat.2008.06.069.
- (208) Spuler, M. J.; Briceño, G.; Duprat, F.; Jorquera, M.; Céspedes, C.; Palma, G. Sorption Kinetics of 2,4-D and Diuron Herbicides in a Urea-Fertilized Andisol. *Journal of Soil Science and Plant Nutrition* **2019**, *19*, 313–320, DOI: 10.1007/s42729-019-00031-0.
- (209) Jones, S. E.; Lennon, J. T. Dormancy contributes to the maintenance of microbial diversity. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107*, 5881–5886, DOI: 10.1073/pnas.0912765107.
- (210) Cycoń, M.; Lewandowska, A.; Piotrowska-Seget, Z. Comparison of mineralization dynamics of 2,4-Dichlorophenoxyacetic acid (2,4-D) and 4-Chloro-2-Methylphenoxyacetic acid (MCPA) in soils of different textures. *Polish Journal of Environmental Studies* **2011**, *20*, 293–301.

- (211) Helweg, A. Degradation and adsorption of  $^{14}\text{C}$ -MCPA in soil—influence of concentration, temperature and moisture content on degradation. *Weed Research* **1987**, *27*, 287–296, DOI: 10.1111/j.1365-3180.1987.tb00765.x.
- (212) Crespín, M. A.; Gallego, M.; Valcárcel, M.; González, J. L. Study of the Degradation of the Herbicides 2,4-D and MCPA at Different Depths in Contaminated Agricultural Soil. *Environmental Science & Technology* **2001**, *35*, 4265–4270, DOI: 10.1021/es0107226.
- (213) López-Piñeiro, A.; Peña, D.; Albarrán, A.; Sánchez-Llerena, J.; Becerra, D. Behavior of MCPA in four intensive cropping soils amended with fresh, composted, and aged olive mill waste. *Journal of Contaminant Hydrology* **2013**, *152*, 137–146, DOI: 10.1016/j.jconhyd.2013.07.003.
- (214) Saleh, O.; Pagel, H.; Enowashu, E.; Devers, M.; Martin-Laurent, F.; Streck, T.; Kandler, E.; Poll, C. Evidence for the importance of litter as a co-substrate for MCPA dissipation in an agricultural soil. *Environmental Science and Pollution Research* **2016**, *23*, 4164–4175, DOI: 10.1007/s11356-015-4633-1.
- (215) Riefer, P.; Klausmeyer, T.; Schmidt, B.; Schäffer, A.; Schwarzbauer, J. Distribution and incorporation mode of the herbicide MCPA in soil derived organo-clay complexes. *Journal of Environmental Science and Health, Part B* **2017**, *52*, 584–599, DOI: 10.1080/03601234.2017.1318639.
- (216) Botero, L. R.; Mougin, C.; Peñuela, G.; Barriuso, E. Formation of 2,4-D bound residues in soils: New insights into microbial metabolism. *Science of The Total Environment* **2017**, *584-585*, 715–722, DOI: 10.1016/j.scitotenv.2017.01.105.
- (217) Sayler, G. S.; Fleming, J. T.; Nivens, D. E. Gene expression monitoring in soils by mRNA analysis and gene lux fusions. *Current Opinion in Biotechnology* **2001**, *12*, 455–460, DOI: 10.1016/s0958-1669(00)00245-7.
- (218) Afzal, M.; Yousaf, S.; Reichenauer, T. G.; Kuffner, M.; Sessitsch, A. Soil type affects plant colonization, activity and catabolic gene expression of inoculated bacterial strains during

- phytoremediation of diesel. *Journal of Hazardous Materials* **2011**, *186*, 1568–1575, DOI: 10.1016/j.jhazmat.2010.12.040.
- (219) Monard, C.; Martin-Laurent, F.; Lima, O.; Devers-Lamrani, M.; Binet, F. Estimating the biodegradation of pesticide in soils by monitoring pesticide-degrading gene expression. *Biodegradation* **2013**, *24*, 203–213, DOI: 10.1007/s10532-012-9574-5.
- (220) Beven, K.; Freer, J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* **2001**, *249*, 11–29, DOI: 10.1016/S0022-1694(01)00421-8.
- (221) Snowden, T. J.; Tindall, M. J.; Graaf, P. H. V. D. Methods of Model Reduction for Large-Scale Biological Systems : A Survey of Current Methods and Trends. *Bulletin of Mathematical Biology* **2017**, *79*, 1449–1486, DOI: 10.1007/s11538-017-0277-2.
- (222) Roser, M. Pesticides. *Our World in Data* **2019**, <https://ourworldindata.org/pesticides>.
- (223) Craven, A.; Hoy, S. Pesticide persistence and bound residues in soil—regulatory significance. *Environmental Pollution* **2005**, *133*, 5–9, DOI: 10.1016/j.envpol.2004.04.010.
- (224) Bethsass, J.; Colangelo, A. European Union Bans Atrazine, While the United States Negotiates Continued Use. *International Journal of Occupational and Environmental Health* **2006**, *12*, 260–267, DOI: 10.1179/oeh.2006.12.3.260.
- (225) Singh, S.; Kumar, V.; Chauhan, A.; Datta, S.; Wani, A. B.; Singh, N.; Singh, J. Toxicity, degradation and analysis of the herbicide atrazine. *Environmental Chemistry Letters* **2018**, *16*, 211–237, DOI: 10.1007/s10311-017-0665-8.
- (226) De Albuquerque, F. P.; de Oliveira, J. L.; Moschini-Carlos, V.; Fraceto, L. F. An overview of the potential impacts of atrazine in aquatic environments: Perspectives for tailored solutions based on nanotechnology. *Science of The Total Environment* **2020**, *700*, 134868, DOI: 10.1016/j.scitotenv.2019.134868.

- (227) Shipitalo, M. J.; Owens, L. B. Atrazine, deethylatrazine, and deisopropylatrazine in surface runoff from conservation tilled watersheds. *Environmental science technology* **2003**, *37*, 944–950, DOI: 10.1021/es020870b.
- (228) Ehrl, B. N.; Mogusu, E. O.; Kim, K.; Hofstetter, H.; Pedersen, J. A.; Elsner, M. High Permeation Rates in Liposome Systems Explain Rapid Glyphosate Biodegradation Associated with Strong Isotope Fractionation. *Environmental Science & Technology* **2018**, *52*, 7259–7268, DOI: 10.1021/acs.est.8b01004.
- (229) Kiss, A.; Rapi, S.; Csutorás, C. GC / MS studies on revealing products and reaction mechanism of photodegradation of pesticides. *Microchemical Journal* **2007**, *85*, 13–20, DOI: 10.1016/j.microc.2006.06.017.
- (230) Kiss, A.; Virág, D. Interpretation and modelling of environmental behaviour of diverse pesticides by revealing photodecomposition mechanisms. *Microchemical Journal* **2009**, *92*, 119–122, DOI: 10.1016/j.microc.2008.12.002.
- (231) Novak, J. M.; Jayachandran, K.; Moorman, T. B.; Weber, J. B. Sorption and Binding of Organic Compounds in Soils and Their Relation to Bioavailability, 1995, DOI: doi : 10.2136/sssaspepub43.c2.
- (232) Guo, L.; Jury, W. A.; Wagenet, R. J.; Flury, M. Dependence of pesticide degradation on sorption: nonequilibrium model and application to soil reactors. *Journal of Contaminant Hydrology* **2000**, *43*, 45–62, DOI: 10.1016/S0169-7722(99)00097-2.
- (233) Dechesne, A.; Owsianiak, M.; Bazire, A.; Grundmann, G. L.; Binning, P. J.; Smets, B. F. Biodegradation in a Partially Saturated Sand Matrix: Compounding Effects of Water Content, Bacterial Spatial Distribution, and Motility. *Environmental Science & Technology* **2010**, *44*, 2386–2392, DOI: 10.1021/es902760y.
- (234) Pinheiro, M.; Pagel, H.; Poll, C.; Ditterich, F.; Garnier, P.; Streck, T.; Kandeler, E.; Vieublé Gonod, L. Water flow drives small scale biogeography of pesticides and bacterial pesticide

- degraders - A microcosm study using 2,4-D as a model compound. *Soil Biology and Biochemistry* **2018**, *127*, 137–147, DOI: 10.1016/j.soilbio.2018.09.024.
- (235) McKelvie, J. R.; Mackay, D. M.; de Sieyes, N. R.; Lacrampe-Couloume, G.; Sherwood Lollar, B. Quantifying MTBE biodegradation in the Vandenberg Air Force Base ethanol release study using stable carbon isotopes. *Journal of Contaminant Hydrology* **2007**, *94*, 157–165, DOI: 10.1016/j.jconhyd.2007.05.008.
- (236) Thullner, M.; Kampara, M.; Richnow, H. H.; Harms, H.; Wick, L. Y. Impact of Bioavailability Restrictions on Microbially Induced Stable Isotope Fractionation. 1. Theoretical Calculation. *Environmental Science & Technology* **2008**, *42*, 6544–6551, DOI: 10.1021/es702782c.
- (237) Braeckevelt, M.; Fischer, A.; Kästner, M. Field applicability of Compound-Specific Isotope Analysis (CSIA) for characterization and quantification of in situ contaminant degradation in aquifers. *Applied Microbiology and Biotechnology* **2012**, *94*, 1401–1421, DOI: 10.1007/s00253-012-4077-1.
- (238) Kundu, K.; Weber, N.; Griebler, C.; Elsner, M. Phenotypic heterogeneity as key factor for growth and survival under oligotrophic conditions. *Environmental Microbiology* **2020**, *n/a*, DOI: 10.1111/1462-2920.15106.
- (239) Gharasoo, M.; Ehrl, B. N.; Cirpka, O. A.; Elsner, M. Modeling of Contaminant Biodegradation and Compound-Specific Isotope Fractionation in Chemostats at Low Dilution Rates. *Environmental Science & Technology* **2018**, *53*, 1186–1196, DOI: 10.1021/acs.est.8b02498.
- (240) Marozava, S.; Meyer, A. H.; Pérez-de-Mora, A.; Gharasoo, M.; Zhuo, L.; Wang, H.; Cirpka, O. A.; Meckenstock, R. U.; Elsner, M. Mass Transfer Limitation during Slow Anaerobic Biodegradation of 2-Methylnaphthalene. *Environmental Science & Technology* **2019**, *53*, 9481–9490, DOI: 10.1021/acs.est.9b01152.

- 
- (241) Desmond-Le Quéméner, E.; Bouchez, T. A thermodynamic theory of microbial growth. *The ISME journal* **2014**, *8*, 1747–1751, DOI: 10.1038/ismej.2014.7.
- (242) Ugalde-Salas, P.; Desmond-Le Quéméner, E.; Harmand, J.; Rapaport, A.; Bouchez, T. Insights from Microbial Transition State Theory on Monod's Affinity Constant. *Scientific Reports* **2020**, *10*, 5323, DOI: 10.1038/s41598-020-62213-6.
- (243) Pirt, S. J. Maintenance energy: a general model for energy-limited and energy-sufficient growth. *Archives of Microbiology* **1982**, *133*, 300–302, DOI: 10.1007/BF00521294.
- (244) Klier, C.; Grundmann, S.; Gayler, S.; Priesack, E. Modelling the Environmental Fate of the Herbicide Glyphosate in Soil Lysimeters. *Water, Air, & Soil Pollution: Focus* **2008**, *8*, 187–207, DOI: 10.1007/s11267-007-9171-5.
- (245) Ercan, O.; Bisschops, M. M. M.; Overkamp, W.; Jørgensen, T. R.; Ram, A. F.; Smid, E. J.; Pronk, J. T.; Kuipers, O. P.; Daran-Lapujade, P.; Kleerebezem, M. Physiological and Transcriptional Responses of Different Industrial Microbes at Near-Zero Specific Growth Rates. *Applied and Environmental Microbiology* **2015**, *81*, byeditorMüller, V., 5662 LP–5670, DOI: 10.1128/AEM.00944-15.
- (246) Devers, M.; Soulas, G.; Martin-Laurent, F. Real-time reverse transcription PCR analysis of expression of atrazine catabolism genes in two bacterial strains isolated from soil. *Journal of Microbiological Methods* **2004**, *56*, 3–15, DOI: 10.1016/j.mimet.2003.08.015.
- (247) Carmo, A. M.; Hundal, L. S.; Thompson, M. L. Sorption of Hydrophobic Organic Compounds by Soil Materials: Application of Unit Equivalent Freundlich Coefficients. *Environmental Science & Technology* **2000**, *34*, 4363–4369, DOI: 10.1021/es000968v.
- (248) Kleineidam, S.; Schüth, C.; Grathwohl, P. Solubility-Normalized Combined Adsorption-Partitioning Sorption Isotherms for Organic Pollutants. *Environmental Science & Technology* **2002**, *36*, 4689–4697, DOI: 10.1021/es010293b.

- (249) Vryzas, Z.; Papadopoulou-Mourkidou, E.; Soulios, G.; Prodromou, K. Kinetics and adsorption of metolachlor and atrazine and the conversion products (deethylatrazine, deisopropylatrazine, hydroxyatrazine) in the soil profile of a river basin. *European Journal of Soil Science* **2007**, *58*, 1186–1199, DOI: 10.1111/j.1365-2389.2007.00913.x.
- (250) Krutz, L. J.; Shaner, D. L.; Zablotowicz, R. M. Enhanced degradation and soil depth effects on the fate of atrazine and major metabolites in Colorado and Mississippi soils. *Journal of environmental quality* **2010**, *39*, 1369–1377, DOI: 10.2134/jeq2009.0197.
- (251) Sigmund, G.; Gharasoo, M.; Hüffer, T.; Hofmann, T. Deep Learning Neural Network Approach for Predicting the Sorption of Ionizable and Polar Organic Pollutants to a Wide Range of Carbonaceous Materials. *Environmental Science and Technology* **2020**, *54*, 4583–4591, DOI: 10.1021/acs.est.9b06287.
- (252) Lever, M. A.; Rogers, K. L.; Lloyd, K. G.; Overmann, J.; Schink, B.; Thauer, R. K.; Hoehler, T. M.; Jørgensen, B. B. Life under extreme energy limitation: a synthesis of laboratory- and field-based investigations. *FEMS Microbiology Reviews* **2015**, *39*, 688–728, DOI: 10.1093/femsre/fuv020.
- (253) Hochman, D.; Dor, M.; Mishael, Y. Diverse effects of wetting and drying cycles on soil aggregation: Implications on pesticide leaching. *Chemosphere* **2021**, *263*, 127910, DOI: 10.1016/j.chemosphere.2020.127910.
- (254) Van Genuchten, M. T.; Wagenet, R. J. Two-Site/Two-Region Models for Pesticide Transport and Degradation: Theoretical Development and Analytical Solutions. *Soil Science Society of America Journal* **1989**, *53*, 1303–1310, DOI: 10.2136/sssaj1989.03615995005300050001x.
- (255) Streck, T.; Poletika, N. N.; Jury, W. A.; Farmer, W. J. Description of Simazine Transport with Rate-Limited, Two-Stage, Linear and Nonlinear Sorption. *Water Resources Research* **1995**, *31*, 811–822, DOI: 10.1029/94WR02822.



- 
- (256) Altfelder, S.; Streck, T. Capability and limitations of first-order and diffusion approaches to describe long-term sorption of chlortoluron in soil. *Journal of Contaminant Hydrology* **2006**, *86*, 279–298, DOI: 10.1016/j.jconhyd.2006.04.002.
- (257) Villaverde, J.; van Beinum, W.; Beulke, S.; Brown, C. D. The Kinetics of Sorption by Retarded Diffusion into Soil Aggregate Pores. *Environmental Science & Technology* **2009**, *43*, 8227–8232, DOI: 10.1021/es9015052.
- (258) Yu, H.; Liu, Y.; Shu, X.; Fang, H.; Sun, X.; Pan, Y.; Ma, L. Equilibrium, kinetic and thermodynamic studies on the adsorption of atrazine in soils of the water fluctuation zone in the Three-Gorges Reservoir. *Environmental Sciences Europe* **2020**, *32*, 27, DOI: 10.1186/s12302-020-00303-y.
- (259) Babey, T.; Vieublé-gonod, L.; Rapaport, A.; Pinheiro, M.; Garnier, P.; Dreuzy, J.-r. D. Spatiotemporal simulations of 2, 4-D pesticide degradation by microorganisms in 3D soil-core experiments. *Ecological Modelling* **2017**, *344*, 48–61, DOI: 10.1016/j.ecolmodel.2016.11.006.
- (260) König, S.; Vogel, H.-J.; Harms, H.; Worrlich, A. Physical, Chemical and Biological Effects on Soil Bacterial Dynamics in Microscale Models. *Frontiers in Ecology and Evolution* **2020**, *8*, 53, DOI: 10.3389/fevo.2020.00053.
- (261) Pagel, H.; Kriesche, B.; Uksa, M.; Poll, C.; Kandeler, E.; Schmidt, V.; Streck, T. Spatial Control of Carbon Dynamics in Soil by Microbial Decomposer Communities, 2020.
- (262) Pot, V.; Portell, X.; Otten, W.; Garnier, P.; Monga, O.; Baveye, P. C. Accounting for soil architecture and microbial dynamics in microscale models: Current practices in soil science and the path ahead. *European Journal of Soil Science* **2021**, *n/a*, DOI: 10.1111/ejss.13142.
- (263) Helweg, A. Degradation and adsorption of <sup>14</sup>C-mecoprop (MCP) in surface soils and in subsoil. Influence of temperature, moisture content, sterilization and concentration on

- degradation. *Science of The Total Environment* **1993**, *132*, 229–241, DOI: 10.1016/0048-9697(93)90134-R.
- (264) Schroll, R.; Becher, H. H.; Dörfler, U.; Gayler, S.; Grundmann, S.; Hartmann, H. P.; Ruoss, J. Quantifying the effect of soil moisture on the aerobic microbial mineralization of selected pesticides in different soils. *Environmental Science & Technology* **2006**, *40*, 3305–3312, DOI: 10.1021/es052205j.
- (265) Bouseba, B.; Zertal, A.; Beguet, J.; Rouard, N.; Devers, M.; Martin, C.; Martin-Laurent, F. Evidence for 2,4-D mineralisation in Mediterranean soils: impact of moisture content and temperature. *Pest Management Science* **2009**, *65*, 1021–1029, DOI: 10.1002/ps.1789.
- (266) Allison, S. D.; Wallenstein, M. D.; Bradford, M. A. Soil-carbon response to warming dependent on microbial physiology. *Nature Geoscience* **2010**, *3*, 336–340, DOI: 10.1038/ngeo846.
- (267) Thirunarayanan, K.; Zimdahl, R. L.; Smika, D. E. Chlorsulfuron Adsorption and Degradation in Soil. *Weed Science* **1985**, *33*, 558–563, DOI: 10.1017/S0043174500082849.
- (268) Choi, J. S.; Fermanian, T. W.; Wehner, D. J.; Spomer, L. A. Effect of Temperature, Moisture, and Soil Texture on DCPA Degradation. *Agronomy Journal* **1988**, *80*, 108, DOI: 10.2134/agronj1988.00021962008000010024x.
- (269) Muskus, A. M.; Krauss, M.; Miltner, A.; Hamer, U.; Nowak, K. M. Degradation of glyphosate in a Colombian soil is influenced by temperature, total organic carbon content and pH. *Environmental Pollution* **2020**, *259*, 113767, DOI: 10.1016/j.envpol.2019.113767.
- (270) Muskus, A. M.; Krauss, M.; Miltner, A.; Hamer, U.; Nowak, K. M. Effect of temperature, pH and total organic carbon variations on microbial turnover of <sup>13</sup>C/<sup>15</sup>N-glyphosate in agricultural soil. *Science of The Total Environment* **2019**, *658*, 697–707, DOI: 10.1016/j.scitotenv.2018.12.195.

- (271) Schimel, J. P.; Schaeffer, S. M. Microbial control over carbon cycling in soil. *Frontiers in microbiology* **2012**, *3*, 348, DOI: 10.3389/fmicb.2012.00348PM-23055998.
- (272) Chowdhury, A.; Pradhan, S.; Saha, M.; Sanyal, N. Impact of pesticides on soil microbiological parameters and possible bioremediation strategies. *Indian Journal of Microbiology* **2008**, *48*, 114–127, DOI: 10.1007/s12088-008-0011-8.
- (273) Pinheiro, M.; Garnier, P.; Beguet, J.; Martin Laurent, F.; Vieublé Gonod, L. The millimetre-scale distribution of 2,4-D and its degraders drives the fate of 2,4-D at the soil core scale. *Soil Biology and Biochemistry* **2015**, *88*, 90–100, DOI: 10.1016/j.soilbio.2015.05.008.
- (274) Jury, W. A.; Spencer, W. F.; Farmer, W. J. Behavior Assessment Model for Trace Organics in Soil: I. Model Description. *Journal of Environmental Quality* **1987**, *16*, 448, DOI: 10.2134/jeq1987.00472425001600040027x.
- (275) Manzoni, S.; Porporato, A. Soil carbon and nitrogen mineralization: Theory and models across scales. *Soil Biology and Biochemistry* **2009**, *41*, 1355–1379, DOI: 10.1016/j.soilbio.2009.02.031.
- (276) Hamamoto, S.; Moldrup, P.; Kawamoto, K.; Komatsu, T. Excluded-volume expansion of Archie's law for gas and solute diffusivities and electrical and thermal conductivities in variably saturated porous media. *Water Resources Research* **2010**, *46*, DOI: 10.1029/2009WR008424.
- (277) Moyano, F. E.; Manzoni, S.; Chenu, C. Responses of soil heterotrophic respiration to moisture availability: An exploration of processes and models. *Soil Biology and Biochemistry* **2013**, *59*, 72–85, DOI: 10.1016/j.soilbio.2013.01.002.
- (278) Griffin, D. M. **in** *Water Potential Relations in Soil Microbiology* Parr, J. F., Gardner, W. R., Elliott, L. F., **editors**; SSSA Special Publications TS - CrossRef; SSSA: [S.l.], 1981, **pages** 141–151, DOI: 10.2136/sssaspecpub9.c5.

- (279) Audus, L. J. The decomposition of 2:4-dichlorophenoxyacetic acid and 2-methyl-4-chlorophenoxyacetic acid in the soil. *Journal of the Science of Food and Agriculture* **1952**, 3, 268–274, DOI: 10.1002/jsfa.2740030607.
- (280) Parr, J. F.; Gardner, W. R.; Elliott, L. F., *Water Potential Relations in Soil Microbiology*; Parr, J. F., Gardner, W. R., Elliott, L. F., **editors**; SSSA: [S.I.], 1981, DOI: 10.2136/sssaspecpub9.
- (281) Griffin, D. M. Soil water in the ecology of fungi. *Annual Review of Phytopathology* **1969**, 289–310, DOI: 10.1146/annurev.py.07.090169.001445.
- (282) Schimel, J.; Balser, T. C.; Wallenstein, M. Microbial stress-response physiology and its implications for ecosystem function. *Ecology* **2007**, 88, 1386–1394, DOI: 10.1890/06-0219.
- (283) Zeglin, L. H.; Bottomley, P. J.; Jumpponen, A.; Rice, C. W.; Arango, M.; Lindsley, A.; McGowan, A.; Mfombep, P.; Myrold, D. D. Altered precipitation regime affects the function and composition of soil microbial communities on multiple time scales. *Ecology* **2013**, 94, 2334–2345, DOI: 10.1890/12-2018.1.
- (284) Manzoni, S.; Schaeffer, S. M.; Katul, G.; Porporato, A.; Schimel, J. P. A theoretical analysis of microbial eco-physiological and diffusion limitations to carbon cycling in drying soils. *Soil Biology and Biochemistry* **2014**, 73, 69–83, DOI: 10.1016/j.soilbio.2014.02.008.
- (285) Schoen, S.; Winterlin, W. The effects of various soil factors and amendments on the degradation of pesticide mixtures. *Journal of Environmental Science and Health, Part B* **1987**, 22, 347–377, DOI: 10.1080/03601238709372561.
- (286) Karanasios, E.; Karpouzas, D. G.; Tsiropoulos, N. G. Key parameters and practices controlling pesticide degradation efficiency of biobed substrates. *Journal of Environmental Science and Health, Part B* **2012**, 47, 589–598, DOI: 10.1080/03601234.2012.665753.

- (287) Helweg, A.; Fomsgaard, I. S.; Reffstrup, T. K.; Srensen, H. Degradation of Mecoprop and Isoproturon in Soil Influence of Initial Concentration. *International Journal of Environmental Analytical Chemistry* **1998**, *70*, 133–148, DOI: 10 . 1080 / 03067319808032610.
- (288) Fomsgaard, I. S.; Kristensen, K. Influence of microbial activity, organic carbon content, soil texture and soil depth on mineralisation rates of low concentrations of <sup>14</sup>C-mecoprop—development of a predictive model. *Ecological Modelling* **1999**, *122*, 45–68, DOI: 10 . 1016/S0304-3800(99)00118-0.
- (289) Wirsching, J.; Pagel, H.; Ditterich, F.; Uksa, M.; Werneburg, M.; Zwiener, C.; Berner, D.; Kandeler, E.; Poll, C. Biodegradation of Pesticides at the Limit. *Frontiers in Microbiology* **2020**, *11*, 1476, DOI: 10.3389/fmicb.2020.02107.
- (290) Patzko, T. Degradation of MCPA in soil horizons of polish agricultural soils. *Polish Journal of Environmental Studies* **2009**, 1083–1091.
- (291) Fredslund, L.; Vinther, F. P.; Brinch, U. C.; Elsgaard, L.; Rosenberg, P.; Jacobsen, C. S. Spatial Variation in 2-Methyl-4-chlorophenoxyacetic Acid Mineralization and Sorption in a Sandy Soil at Field Level. *Journal of Environmental Quality* **2008**, *37*, 1918–1928, DOI: 10 . 2134/jeq2006 . 0208.
- (292) Oh, K.-H.; Ahn, S.-K.; Yoon, K.-H.; Kim, Y.-S. Biodegradation of the phenoxy herbicide MCPA by microbial consortia isolated from a rice field. *Bulletin of Environmental Contamination and Toxicology* **1995**, *55*, DOI: 10 . 1007/BF00196033.
- (293) Duo-Sen, L.; Shui-Ming, Z. Kinetic model for degradative processes of pesticides in soil. *Ecological Modelling* **1987**, *37*, 131–138, DOI: 10 . 1016/0304-3800(87)90021-4.
- (294) Vance, E. D.; Brookes, P. C.; Jenkinson, D. S. An extraction method for measuring soil microbial biomass C. *Soil Biology and Biochemistry* **1987**, *19*, 703–707, DOI: 10 . 1016/0038-0717(87)90052-6.

- (295) Poll, C.; Pagel, H.; Devers-Lamrani, M.; Martin-Laurent, F.; Ingwersen, J.; Streck, T.; Kandeler, E. Regulation of bacterial and fungal MCPA degradation at the soil-litter interface. *Soil Biology and Biochemistry* **2010**, *42*, 1879–1887, DOI: 10.1016/j.soilbio.2010.07.013.
- (296) Joergensen, R. G. The fumigation-extraction method to estimate soil microbial biomass. *Soil Biology and Biochemistry* **1996**, *28*, 25–31, DOI: 10.1016/0038-0717(95)00102-6.
- (297) Chavez Rodriguez, L.; Ingalls, B.; Schwarz, E.; Streck, T.; Uksa, M.; Pagel, H. Gene-Centric Model Approaches for Accurate Prediction of Pesticide Biodegradation in Soils. *Environmental Science & Technology* **2020**, *54*, 13638–13650, DOI: 10.1021/acs.est.0c03315.
- (298) Sierra, C. A.; Trumbore, S. E.; Davidson, E. A.; Vicca, S.; Janssens, I. Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture. *Journal of Advances in Modeling Earth Systems* **2015**, *7*, 335–356, DOI: 10.1002/2014MS000358.
- (299) Wang, G.; Post, W. M. A theoretical reassessment of microbial maintenance and implications for microbial ecology modeling. *FEMS Microbiology Ecology* **2012**, *81*, 610–617, DOI: 10.1111/j.1574-6941.2012.01389.x.
- (300) Gawlik, B. M.; Lamberty, A.; Pauwels, J.; Blum, W. E. H.; Mentler, A.; Bussian, B.; Eklo, O.; Fox, K.; Kördel, W.; Hennecke, D.; Maurer, T.; Perrin-Ganier, C.; Pflugmacher, J.; Romero-Taboada, E.; Szabo, G.; Muntau, H. Certification of the European reference soil set (IRMM-443—EUROSOILS). Part I. Adsorption coefficients for atrazine, 2,4-D and lindane. *Science of The Total Environment* **2003**, *312*, 23–31, DOI: 10.1016/S0048-9697(03)00193-1.
- (301) Sinsabaugh, R. L.; Manzoni, S.; Moorhead, D. L.; Richter, A. Carbon use efficiency of microbial communities. *Ecology Letters* **2013**, *16*, 930–939, DOI: 10.1111/ele.12113.

- 
- (302) Geyer, K. M.; Kyker-Snowman, E.; Grandy, A. S.; Frey, S. D. Microbial carbon use efficiency: accounting for population, community, and ecosystem-scale controls over the fate of metabolized organic matter. *Biogeochemistry* **2016**, *127*, 173–188, DOI: 10.1007/s10533-016-0191-y.
- (303) Manzoni, S.; Čapek, P.; Porada, P.; Thurner, M.; Winterdahl, M.; Beer, C.; Brüchert, V.; Frouz, J.; Herrmann, A. M.; Lindahl, B. D.; Lyon, S. W.; Šantrůčková, H.; Vico, G.; Way, D. Reviews and syntheses: Carbon use efficiency from organisms to ecosystems – definitions, theories, and empirical evidence. *Biogeosciences* **2018**, *15*, 5929–5949, DOI: 10.5194/bg-15-5929-2018.
- (304) Pinheiro, J.; Bates, D.; DebRoy, S.; Sarkar, D. Linear and Nonlinear Mixed Effects Models. **2020**.
- (305) Lenth, R.; Buerkner, P.; Herve, M.; Love, J.; Riebl, H.; Singmann, H. emmeans, 2020.
- (306) Manzoni, S.; Schimel, J. P.; Porporato, A. Responses of soil microbial communities to water stress: results from a meta-analysis. *Ecology* **2012**, *93*, 930–938, DOI: 10.1890/11-0026.1.
- (307) Nowak, K. M.; Miltner, A.; Poll, C.; Kandeler, E.; Streck, T.; Pagel, H. Plant litter enhances degradation of the herbicide MCPA and increases formation of biogenic non-extractable residues in soil. *Environment international* **2020**, *142*, 105867, DOI: 10.1016/j.envint.2020.105867.
- (308) Davidson, E. A.; Janssens, I. A. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature* **2006**, *440*, 165–173, DOI: 10.1038/nature04514.
- (309) Sparling, G. P.; West, A. W.; Reynolds, J. Influence of soil moisture regime on the respiration response of soils subjected to osmotic stress. *Soil Research* **1989**, *27*, 161–168, DOI: 10.1071/SR9890161.

- 
- (310) Ilstedt, U.; Nordgren, A.; Malmer, A. Optimum soil water for soil respiration before and after amendment with glucose in humid tropical acrisols and a boreal mor layer. *Soil Biology and Biochemistry* **2000**, *32*, 1591–1599, DOI: 10.1016/S0038-0717(00)00073-0.
- (311) Papendick, R. I.; Campbell, G. S. Theory and Measurement of Water Potential, 1981, DOI: 10.2136/sssaspecpub9.c1.
- (312) Jacob, B.; Trine, H.; Bruun, H. H. C.; Suhr, J. C. Degradation of 4-Chloro-2-Methylphenoxyacetic Acid in Top- and Subsoil Is Quantitatively Linked to the Class III tfdA Gene. *Applied and Environmental Microbiology* **2006**, *72*, 1476–1486, DOI: 10.1128/AEM.72.2.1476-1486.2006.
- (313) Hiller, E.; Bartal, M.; Milička, J.; Čerňanský, S. Environmental fate of the herbicide MCPA in two soils as affected by the presence of wheat ash. *Water, Air, and Soil Pollution* **2009**, *197*, 395–402, DOI: 10.1007/s11270-008-9820-y.
- (314) Peña, D.; López-Piñeiro, A.; Albarrán, Á.; Becerra, D.; Sánchez-Llerena, J. Environmental fate of the herbicide MCPA in agricultural soils amended with fresh and aged de-oiled two-phase olive mill waste. *Environmental Science and Pollution Research* **2015**, *22*, 13915–13925, DOI: 10.1007/s11356-015-4622-4.
- (315) Vieublé Gonod, L. Variabilité spatiale de la minéralisation de substrats carbonés (2,4-D, leucine, lysine) dans la matrice solide du sol, Theses, Université Claude Bernard Lyon 1, 2002, 250 p.
- (316) Gözdereliler, E.; Boon, N.; Aamand, J.; De Roy, K.; Granitsiotis, M. S.; Albrechtsen, H.-J.; Sørensen, S. R. Comparing metabolic functionalities, community structures, and dynamics of herbicide-degrading communities cultivated with different substrate concentrations. *Applied and environmental microbiology* **2013**, *79*, 367–375, DOI: 10.1128/AEM.02536-12.



- (317) Geyer, K. M.; Dijkstra, P.; Sinsabaugh, R.; Frey, S. D. Clarifying the interpretation of carbon use efficiency in soil through methods comparison. *Soil Biology and Biochemistry* **2019**, *128*, 79–88, DOI: 10.1016/j.soilbio.2018.09.036.
- (318) Manzoni, S.; Čapek, P.; Mooshammer, M.; Lindahl, B. D.; Richter, A.; Šantrůčková, H. Optimal metabolic regulation along resource stoichiometry gradients. *Ecology Letters* **2017**, *20*, 1182–1191, DOI: 10.1111/ele.12815.
- (319) Frey, S. D.; Lee, J.; Melillo, J. M.; Six, J. The temperature response of soil microbial efficiency and its feedback to climate. *Nature Climate Change* **2013**, *3*, 395–398, DOI: 10.1038/nclimate1796.
- (320) Domeignoz-Horta, L. A.; Pold, G.; Liu, X.-J. A.; Frey, S. D.; Melillo, J. M.; DeAngelis, K. M. Microbial diversity drives carbon use efficiency in a model soil. *Nature Communications* **2020**, *11*, 3684, DOI: 10.1038/s41467-020-17502-z.
- (321) Dacal, M.; Delgado-Baquerizo, M.; Barquero, J.; Berhe, A. A.; Gallardo, A.; Maestre, F. T.; García-Palacios, P. Temperature Increases Soil Respiration Across Ecosystem Types and Soil Development, But Soil Properties Determine the Magnitude of This Effect. *Ecosystems* **2021**, DOI: 10.1007/s10021-021-00648-2.
- (322) Pold, G.; Domeignoz-Horta, L. A.; Morrison, E. W.; Frey, S. D.; Sistla, S. A.; DeAngelis, K. M. Carbon use efficiency and its temperature sensitivity covary in soil bacteria. *mBio* **2020**, *11*, **byeditor**Giovannoni, S. J., e02293–19, DOI: 10.1128/mBio.02293-19.
- (323) Laidler, K. J. The development of the Arrhenius equation. *Journal of Chemical Education* **1984**, *61*, 494, DOI: 10.1021/ed061p494.
- (324) Jones, D. L.; Olivera-Ardid, S.; Klumpp, E.; Knief, C.; Hill, P. W.; Lehndorff, E.; Bol, R. Moisture activation and carbon use efficiency of soil microbial communities along an aridity gradient in the Atacama Desert. *Soil Biology and Biochemistry* **2018**, *117*, 68–71, DOI: 10.1016/j.soilbio.2017.10.026.

- (325) Steinweg, J. M.; Plante, A. F.; Conant, R. T.; Paul, E. A.; Tanaka, D. L. Patterns of substrate utilization during long-term incubations at different temperatures. *Soil Biology and Biochemistry* **2008**, *40*, 2722–2728, DOI: 10.1016/j.soilbio.2008.07.002.
- (326) Parker, L. W.; Doxtader, K. G. Kinetics of the Microbial Degradation of 2,4-D in Soil. *Journal of Environmental Quality* **1983**, *12*, 553–558, DOI: 10.2134/jeq1983.00472425001200040024x.
- (327) Cattaneo, M. V.; Masson, C.; Greer, C. W. The influence of moisture on microbial transport, survival and 2,4-D biodegradation with a genetically marked *Burkholderia cepacia* in unsaturated soil columns. *Biodegradation* **1997**, *8*, 87–96, DOI: 10.1023/A:1008236401342.
- (328) Castillo, M. d. P.; Torstensson, L. Effect of Biobed Composition, Moisture, and Temperature on the Degradation of Pesticides. *Journal of Agricultural and Food Chemistry* **2007**, *55*, 5725–5733, DOI: 10.1021/jf0707637.
- (329) Russell, J. B.; Cook, G. M. Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiological Reviews* **1995**, *59*, 48–62, DOI: 10.1128/mr.59.1.48-62.1995.
- (330) Karlsson, A. S.; Lesch, M.; Weihermüller, L.; Thiele, B.; Disko, U.; Hofmann, D.; Vereecken, H.; Spielvogel, S. Pesticide contamination of the upper Elbe River and an adjacent floodplain area. *Journal of Soils and Sediments* **2020**, *20*, 2067–2081, DOI: 10.1007/s11368-020-02571-w.
- (331) Ngigi, A. N.; Getenga, Z. M.; Boga, H. I.; Ndalut, P. K. Biodegradation of s-triazine herbicide atrazine by *Enterobacter cloacae* and *Burkholderia cepacia* sp. from long-term treated sugarcane-cultivated soils in Kenya. *Journal of environmental science and health. Part. B, Pesticides, food contaminants, and agricultural wastes* **2012**, *47*, 769–778, DOI: 10.1080/03601234.2012.676364.

- (332) Meyer, A. H.; Dybala-Defratyka, A.; Alaimo, P. J.; Geronimo, I.; Sanchez, A. D.; Cramer, C. J.; Elsner, M. Cytochrome P450-catalyzed dealkylation of atrazine by *Rhodococcus* sp. strain NI86/21 involves hydrogen atom transfer rather than single electron transfer. *Dalton Transactions* **2014**, 43, 12175–12186, DOI: 10.1039/C4DT00891J.
- (333) Zhang, J. J.; Gao, S.; Xu, J. Y.; Lu, Y. C.; Lu, F. F.; Ma, L. Y.; Su, X. N.; Yang, H. Degrading and Phytoextracting Atrazine Residues in Rice (*Oryza sativa*) and Growth Media Intensified by a Phase II Mechanism Modulator. *Environmental Science & Technology* **2017**, 51, 11258–11268, DOI: 10.1021/acs.est.7b02346.
- (334) Arhonditsis, G. B.; Perhar, G.; Zhang, W.; Massos, E.; Shi, M.; Das, A. Addressing equifinality and uncertainty in eutrophication models. *Water Resources Research* **2008**, 44, DOI: 10.1029/2007WR005862.
- (335) Chen, B. H.; Asprey, S. P. On the Design of Optimally Informative Dynamic Experiments for Model Discrimination in Multiresponse Nonlinear Situations. *Industrial & Engineering Chemistry Research* **2003**, 42, 1379–1390, DOI: 10.1021/ie0203025.
- (336) Kremling, A.; Fischer, S.; Gadkar, K.; Doyle, F. J.; Sauter, T.; Bullinger, E.; Allgöwer, F.; Gilles, E. D. A benchmark for methods in reverse engineering and model discrimination: problem formulation and solutions. *Genome research* **2004**, 14, 1773–1785, DOI: 10.1101/gr.1226004.
- (337) Bekins, B. A.; Warren, E.; Godsy, E. M. A Comparison of Zero-Order, First-Order, and Monod Biotransformation Models. *Groundwater* **1998**, 36, 261–268, DOI: 10.1111/j.1745-6584.1998.tb01091.x.
- (338) Kumar, A.; Singh, N. Atrazine and its metabolites degradation in mineral salts medium and soil using an enrichment culture. *Environmental Monitoring and Assessment* **2016**, 188, 142, DOI: 10.1007/s10661-016-5144-3.

- (339) Deutch, C. E.; Bui, A. P.; Ho, T. Growth of *Paenarthrobacter aurescens* strain TC1 on atrazine and isopropylamine during osmotic stress. *Annals of Microbiology* **2018**, *68*, 569–577, DOI: 10.1007/s13213-018-1364-9.
- (340) Zhang, Y.; Jiang, Z.; Cao, B.; Hu, M.; Wang, Z.; Dong, X. Metabolic ability and gene characteristics of *Arthrobacter* sp. strain DNS10, the sole atrazine-degrading strain in a consortium isolated from black soil. *International Biodeterioration & Biodegradation* **2011**, *65*, 1140–1144, DOI: 10.1016/j.ibiod.2011.08.010.
- (341) Ma, L.; Chen, S.; Yuan, J.; Yang, P.; Liu, Y.; Stewart, K. Rapid biodegradation of atrazine by *Ensifer* sp. strain and its degradation genes. *International Biodeterioration & Biodegradation* **2017**, *116*, 133–140, DOI: 10.1016/j.ibiod.2016.10.022.
- (342) Fan, X.; Song, F. Bioremediation of atrazine: recent advances and promises. *Journal of Soils and Sediments* **2014**, *14*, 1727–1737, DOI: 10.1007/s11368-014-0921-5.
- (343) Shapir, N.; Mandelbaum, R. T. Atrazine Degradation in Subsurface Soil by Indigenous and Introduced Microorganisms. *Journal of Agricultural and Food Chemistry* **1997**, *45*, 4481–4486, DOI: 10.1021/jf970423t.
- (344) Huang, H.; Zhang, S.; Shan, X.-q.; Chen, B.-D.; Zhu, Y.-G.; Bell, J. N. B. Effect of arbuscular mycorrhizal fungus (*Glomus caledonium*) on the accumulation and metabolism of atrazine in maize (*Zea mays* L.) and atrazine dissipation in soil. *Environmental Pollution* **2007**, *146*, 452–457, DOI: 10.1016/j.envpol.2006.07.001.
- (345) Aukema, K.; Tassoulas, L.; Robinson, S.; Konopatski, J.; Bygd, M.; Wackett, L.; Hideaki, N. Cyanuric acid biodegradation via Biuret: Physiology, taxonomy, and geospatial distribution. *Applied and Environmental Microbiology* **2021**, *86*, e01964–19, DOI: 10.1128/AEM.01964-19.
- (346) Krutz, L. J.; Shaner, D. L.; Accinelli, C.; Zablutowicz, R. M.; Henry, W. B. Atrazine dissipation in s-triazine-adapted and nonadapted soil from Colorado and Mississippi:

- implications of enhanced degradation on atrazine fate and transport parameters. *Journal of environmental quality* **2008**, *37*, 848–857, DOI: 10.2134/jeq2007.0448.
- (347) La Cecilia, D.; Maggi, F. In-situ atrazine biodegradation dynamics in wheat (*Triticum*) crops under variable hydrologic regime. *Journal of Contaminant Hydrology* **2017**, *203*, 104–121, DOI: 10.1016/j.jconhyd.2017.05.004.
- (348) Kolekar, P. D.; Patil, S. M.; Suryavanshi, M. V.; Suryawanshi, S. S.; Khandare, R. V.; Govindwar, S. P.; Jadhav, J. P. Microcosm study of atrazine bioremediation by indigenous microorganisms and cytotoxicity of biodegraded metabolites. *Journal of Hazardous Materials* **2019**, *374*, 66–73, DOI: 10.1016/j.jhazmat.2019.01.023.
- (349) Henn, C.; Monteiro, D. A.; Boscolo, M.; da Silva, R.; Gomes, E. Biodegradation of atrazine and ligninolytic enzyme production by basidiomycete strains. *BMC Microbiology* **2020**, *20*, 266, DOI: 10.1186/s12866-020-01950-0.
- (350) Zhang, Y.; Tian, L.; Ma, R.; Wei, Y.; Yang, J.; Sun, Y.; Li, L. Effects of atrazine on the black community land function and role of microbial ecology. *IOP Conference Series: Earth and Environmental Science* **2020**, *446*, 32093, DOI: 10.1088/1755-1315/446/3/032093.
- (351) Abate, G.; Masini, J. C. Sorption of Atrazine, Propazine, Deethylatrazine, Deisopropylatrazine and Hydroxyatrazine onto Organovermiculite Gilberto. *Journal of the Brazilian Chemical Society* **2005**, *16*, 936–943.
- (352) Krutz, L. J.; Burke, I. C.; Reddy, K. N.; Zablotowicz, R. M.; Price, A. J.; Krutz, L. J.; Burke, I. C.; Reddy, K. N.; Zablotowicz, R. M.; Price, A. J. Enhanced Atrazine Degradation : Evidence for Reduced Residual Weed Control and a Method for Identifying Adapted Soils and Predicting Herbicide Persistence. *Weed Science* **2009**, *57*, 427–434, DOI: 10.1614/WS-09-0.
- (353) Zamora-Sillero, E.; Hafner, M.; Ibig, A.; Stelling, J.; Wagner, A. Efficient characterization of high-dimensional parameter spaces for systems biology. *BMC Systems Biology* **2011**, *5*, 142, DOI: 10.1186/1752-0509-5-142.

- (354) Polityko, E. Calculation of Pareto points, 2021.
- (355) Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* **2002**, *6*, 182–197, DOI: 10.1109/4235.996017.
- (356) Moses, Y. spider<sub>p</sub>lot, 2021.
- (357) De Paula, R. T.; Bergson, A.; de Queiroz, M.; Neves, A. A.; da Silva, A. A. Leaching and persistence of ametryn and atrazine in red & yellow latosol. *Journal of Environmental Science and Health* **2016**, *51*, 90–96.
- (358) Huang, H.; Zhang, C.; Rong, Q.; Li, C.; Mao, J.; Liu, Y.; Chen, J.; Liu, X. Effect of two organic amendments on atrazine degradation and microorganisms in soil. *Applied Soil Ecology* **2020**, *152*, 103564, DOI: 10.1016/j.apsoil.2020.103564.
- (359) Briceño, G.; Jorquera, M. A.; Demanet, R.; Mora, M. L.; Durán, N.; Palma, G. Effect of cow slurry amendment on atrazine dissipation and bacterial community structure in an agricultural Andisol. *Science of The Total Environment* **2010**, *408*, 2833–2839, DOI: 10.1016/j.scitotenv.2010.03.014.
- (360) Nousiainen, A. O.; Björklöf, K.; Sagarkar, S.; Mukherjee, S.; Purohit, H. J.; Kapley, A.; Jørgensen, K. S. Atrazine degradation in boreal nonagricultural subsoil and tropical agricultural soil. *Journal of Soils and Sediments* **2014**, *14*, 1179–1188, DOI: 10.1007/s11368-014-0868-6.
- (361) De Souza, M. L.; Newcombe, D.; Alvey, S.; Crowley, D. E.; Hay, A.; Sadowsky, M. J.; Wackett, L. P. Molecular Basis of a Bacterial Consortium: Interspecies Catabolism of Atrazine. *Applied and Environmental Microbiology* **2000**, *66*, 1252 LP –1252, DOI: 10.1128/AEM.66.3.1252-1252.2000.
- (362) Dechesne, A.; Badawi, N.; Aamand, J.; Smets, B. F. Fine scale spatial variability of microbial pesticide degradation in soil: scales, controlling factors, and implications. *Frontiers in Microbiology* **2014**, *5*, 667, DOI: 10.3389/fmicb.2014.00667.

- (363) Adam, I. K. U.; Miltner, A.; Kästner, M. Degradation of <sup>13</sup>C-labeled pyrene in soil-compost mixtures and fertilized soil. *Applied Microbiology and Biotechnology* **2015**, *99*, 9813–9824, DOI: 10.1007/s00253-015-6822-8.
- (364) Sun, F.; Mellage, A.; Gharasoo, M.; Melsbach, A.; Cao, X.; Zimmermann, R.; Griebler, C.; Thullner, M.; Cirpka, O. A.; Elsner, M. Mass-Transfer-Limited Biodegradation at Low Concentrations—Evidence from Reactive Transport Modeling of Isotope Profiles in a Bench-Scale Aquifer. *Environmental Science & Technology* **2021**, DOI: 10.1021/acs.est.0c08566.
- (365) Quéméner, E. D.-L.; Bouchez, T. A thermodynamic theory of microbial growth. *International Society for Microbial Ecology* **2014**, *8*, 1747–1751, DOI: 10.1038/ismej.2014.7.
- (366) Bento, C. P. M.; Yang, X.; Gort, G.; Xue, S.; van Dam, R.; Zomer, P.; Mol, H. G. J.; Ritsema, C. J.; Geissen, V. Persistence of glyphosate and aminomethylphosphonic acid in loess soil under different combinations of temperature, soil moisture and light/darkness. *Science of The Total Environment* **2016**, *572*, 301–311, DOI: 10.1016/j.scitotenv.2016.07.215.
- (367) Su, W.; Hao, H.; Wu, R.; Xu, H.; Xue, F.; Lu, C. Degradation of Mesotrione Affected by Environmental Conditions. *Bulletin of Environmental Contamination and Toxicology* **2017**, *98*, 212–217, DOI: 10.1007/s00128-016-1970-9.
- (368) Han, D.; Yan, D.; Wang, Q.; Fang, W.; Wang, X.; Li, J.; Wang, D.; Li, Y.; Ouyang, C.; Cao, A. Effects of soil type, temperature, moisture, application dose, fertilizer, and organic amendments on chemical properties and biodegradation of dimethyl disulfide in soil. *Land Degradation & Development* **2018**, *29*, 4282–4290, DOI: 10.1002/ldr.3177.
- (369) Csonka, L. N. Physiological and genetic responses of bacteria to osmotic stress. *Microbiological reviews* **1989**, *53*, 121–147.

- (370) Störiko, A.; Pagel, H.; Mellage, A.; Cirpka, O. A. Does It Pay Off to Explicitly Link Functional Gene Expression to Denitrification Rates in Reaction Models?, 2021, DOI: 10.3389/fmicb.2021.684146.
- (371) Box, G. E. P. Science and Statistics. *Journal of the American Statistical Association* **1976**, *71*, 791–799, DOI: 10.2307/2286841.
- (372) Beven, K. J. Uniqueness of place and process representations in hydrological modelling. *Hydrol. Earth Syst. Sci.* **2000**, *4*, 203–213, DOI: 10.5194/hess-4-203-2000.
- (373) Garcia-Cremades, M.; Melillo, N.; Troconiz, I. F.; Magni, P. Mechanistic Multiscale Pharmacokinetic Model for the Anticancer Drug 2',2'-difluorodeoxycytidine (Gemcitabine) in Pancreatic Cancer. *Clinical and Translational Science* **2020**, *13*, 608–617, DOI: 10.1111/cts.12747.
- (374) Suciu, N.; Tanaka, T.; Trevisan, M.; Schuhmacher, M.; Nadal, M.; Rovira, J.; Segui, X.; Casal, J.; Darbra, R. M.; Capri, E. **in** *The Handbook of Environmental Chemistry* Bilitewski, B., Darbra, R. M., Barceló, D., **editors**; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013, **pages** 47–71, DOI: 10.1007/698\_2012\_177.
- (375) MacLeod, M.; Scheringer, M.; McKone, T. E.; Hungerbühler, K. The State of Multimedia Mass-Balance Modeling in Environmental Science and Decision-Making. *Environmental Science & Technology* **2010**, *44*, 8360–8364, DOI: 10.1021/es100968w.
- (376) Council, N. R., *Models in Environmental Regulatory Decision Making*; The National Academies Press: Washington, DC, 2007, **page** 286, DOI: 10.17226/11972.
- (377) Ebrahimi, A.; Or, D. On Upscaling of Soil Microbial Processes and Biogeochemical Fluxes From Aggregates to Landscapes. *Journal of Geophysical Research: Biogeosciences* **2018**, *123*, 1526–1547, DOI: 10.1029/2017JG004347.
- (378) Chakrawal, A.; Herrmann, A. M.; Koestel, J.; Jarsjö, J.; Nunan, N.; Kätterer, T.; Manzoni, S. Dynamic upscaling of decomposition kinetics for carbon cycling models. *Geosci. Model Dev.* **2020**, *13*, 1399–1429, DOI: 10.5194/gmd-13-1399-2020.



- 
- (379) Allison, S. D. A trait-based approach for modelling microbial litter decomposition. *Ecology Letters* **2012**, *15*, 1058–1070, DOI: 10.1111/j.1461-0248.2012.01807.xM4.
- (380) Manzoni, S.; Porporato, A.; Schimel, J. P. Soil heterogeneity in lumped mineralization–immobilization models. *Soil Biology and Biochemistry* **2008**, *40*, 1137–1148, DOI: 10.1016/j.soilbio.2007.12.006.
- (381) Baker, R. E.; Peña, J.-M.; Jayamohan, J.; Jérusalem, A. Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biology Letters* **2018**, *14*, 20170660, DOI: 10.1098/rsbl.2017.0660.
- (382) Overy, D. P.; Bell, M. A.; Habtewold, J.; Helgason, B. L.; Gregorich, E. G. “Omics” Technologies for the Study of Soil Carbon Stabilization: A Review, 2021, DOI: 10.3389/fenvs.2021.617952.
- (383) Gaw, N.; Hawkins-Daarud, A.; Hu, L. S.; Yoon, H.; Wang, L.; Xu, Y.; Jackson, P. R.; Singleton, K. W.; Baxter, L. C.; Eschbacher, J.; Gonzales, A.; Nespodzany, A.; Smith, K.; Nakaji, P.; Mitchell, J. R.; Wu, T.; Swanson, K. R.; Li, J. Integration of machine learning and mechanistic models accurately predicts variation in cell density of glioblastoma using multiparametric MRI. *Scientific Reports* **2019**, *9*, 10063, DOI: 10.1038/s41598-019-46296-4.
- (384) Lagergren, J. H.; Nardini, J. T.; Baker, R. E.; Simpson, M. J.; Flores, K. B. Biologically-informed neural networks guide mechanistic modeling from sparse experimental data. *PLOS Computational Biology* **2020**, *16*, e1008462, DOI: 10.1371/journal.pcbi.1008462.
- (385) Gadkar, K.; Gunawan, R.; Doyle, F. Iterative approach to model identification of biological networks. *BMC Bioinformatics* **2005**, *6*, 155, DOI: 10.1186/1471-2105-6-155.
- (386) López-Gutiérrez, J. C.; Henry, S.; Hallet, S.; Martin-Laurent, F.; Catroux, G.; Philippot, L. Quantification of a novel group of nitrate-reducing bacteria in the environment by

real-time PCR. *Journal of Microbiological Methods* **2004**, 57, 399–407, DOI: 10.1016/j.mimet.2004.02.009.

## 12 Acknowledgements

My journey as a Ph.D. student started about three and a half years ago. Although it was challenging, with lots of ups and downs, it was totally worth it. During these years, I met and worked with wonderful and inspiring people that supported me and helped me to reach the finish line.

First of all, I want to express my deepest gratitude to my dear supervisors. To Prof. Dr. Thilo Streck, my principal supervisor, for giving me the opportunity to immerse myself in the world of mathematical modeling since the first year of my master's degree. Your interesting and often challenging questions helped me to improve my research and to find my way in science. My deepest thanks also go to Dr. Brian Ingalls from the University of Waterloo. Despite of being an expert in modeling and math, you valued each piece of knowledge and every idea that I brought to the discussion and made me feel confident every time. Thanks to your guidance, I was not only able to advance my research but also to overcome the imposter syndrome that has been haunting me over the years. To me you are not just a supervisor but a mentor. Last but not least, all my gratitude to Dr. Holger Pagel, a super talented postdoc, who, with endless patience, accompanied me since my first steps into the field of biogeochemical modeling. You supported me all the time, and far beyond the responsibilities of a supervisor. I owe you a lot, and I will try to make you proud of me.

I had the chance to work with outstanding researchers who collaborated with me as co-authors or colleagues, and who enriched my papers with their insightful comments. First of all, my deepest gratitude to Dr. Sebastian Gayler, who supervised my master thesis and who recommended me for the Ph.D. position. I will never forget all you have done for me. Also, my gratitude to Prof. Dr. Wolfgang Nowak, an inspiring researcher who always took time for me when I needed advice and who motivated me to keep going when I was experiencing the deepest moments in my Ph.D. Thank you so much for it. I am very thankful to Dr. Ana Gonzales-Nicolas. Ana, tú me inspiras mucho como científica y mujer, gracias por siempre estar presta a escucharme y brindarme un consejo.

I am very grateful to my fellows from the biogeophysics department of the University of Hohenheim, especially my dearest friends from the Meiereihof Dani, Hossein, Fasil, Michelle, Verena, Joachim, and also to Arne, Pascal, Lona, and Rashav. You are like family to me. All our experiences together, the walks to the canteen, the tasteless food, the long talks after lunch, and the “we are here for you” made me feel at home all the time. Special thanks to Hannes Wirsching, we had it hard with the paper, but we managed to finish it; well done. Also, special thanks to my good friend Erik Schwarz. Du bist so talentiert und ich kann kaum erwarten zu sehen, was die Zukunft dir bringen wird. Go for it! Guys, you are all the best, and I could not have wished for a better team to work with.

I also want to thank my fellow Ph.D. students from the RTG, who were special companions to me on this journey. The meetings every Thursday, the fall and spring schools, the game sessions after the lectures, all that we went through together will be treasured forever. Anna, muchas gracias por tu apoyo, eres muy especial para mí, te admiro mucho por tu inteligencia y por ser la hermosa persona que eres, te debo mucho. And Ishani, having you at my side made this journey easier, thank you so much. I also want to thank the amazing people from the applied math department at the University of Waterloo, Julien, Nate, Christian, and Saptarshi. You guys made my days in Canada unforgettable, and I hope our paths will cross again. Meeting and collaborating with these inspiring people was possible thanks to the Research Training Group 1829 “Integrated Hydrosystem Modeling” led by Prof. Dr.-Ing. Olaf A. Cirpka.

I want to thank the amazing people from my home country Peru, who inspired me but who are sadly not with me anymore to see me fulfilling this dream. My godfather Dr. Guillermo Aguirre Yato and Ing. Carlos Alberto Llerena Pinto. I am sure that wherever you are now, you will be proud of me. Last but not least, I want to thank my dad (Papi), my mom (Mami), and my little sister, Paula. You and our ‘babies’ are the reason why I wake up every day and work hard, trying to forget that I am far away from home. Because I have you by my side, I can endure all hardships and make you proud of me. I dedicate this thesis to you. And finally, and most importantly, all my gratitude to Matthias Mamsch. You know that I would not be here without

you. Coming to Germany was the best decision of my life because I could meet you. I admire you and love you, and I am so lucky to have you in my life. Danke mi amor.

## 13 Curriculum Vitae

### Luciana Chavez Rodriguez

Date of birth: 27.07.1990

Place of birth: Caraz, Ancash, Peru

<https://orcid.org/0000-0003-1510-6695>

#### EDUCATION

- 2018-2021 PhD student at the Institute of Soil Science and Land Evaluation, Biogeophysics (Prof. Thilo Streck), University of Hohenheim
- 2015-2017 Master degree in Environmental protection and agricultural food production, University of Hohenheim
- 2007-2014 Bachelor and Diploma degree in environmental engineering, faculty of science, Universidad Agraria la Molina, Peru

#### WORK EXPERIENCE

- 2019 – 2020 Lecturer of the computer exercise of the course Environmental modeling for the Master Programs ENVIROFOOD, ENVIRO and Landscape Ecology, University of Hohenheim, Germany, winter semester 2019-2020
- 2016-2017 Student assistant for the courses of Environmental Modelling and Spatial Data Analysis with GIS, University of Hohenheim, Germany
- 2015 Lab assistant at the Helmholtz-Zentrum für Umweltforschung - UFZ
- 2015 Lab assistant at the Universidad Agraria la Molina, Peru - Soils and plant Analysis Lab
- 2014 Internship at the Peruvian Society of Soil Science
- 2012 Internship at Consultores Asociados en Naturaleza y Desarrollo (CANDES)

---

Place, Date

Signature of the student

## 14 Declaration

### Annex 3

**Declaration in lieu of an oath on independent work according to Sec. 18(3) sentence 5 of the University of Hohenheim's Doctoral Regulations for the Faculties of Agricultural Sciences, Natural Sciences, and Business, Economics and Social Sciences**

1. The dissertation submitted on the topic

Modeling Microbial Regulation of Pesticide Turnover in Soils

is work done independently by me.

2. I only used the sources and aids listed and did not make use of any impermissible assistance from third parties. In particular, I marked all content taken word-for-word or paraphrased from other works.

3. I did not use the assistance of a commercial doctoral placement or advising agency.

4. I am aware of the importance of the declaration in lieu of oath and the criminal consequences of false or incomplete declarations in lieu of oath.

I confirm that the declaration above is correct. I declare in lieu of oath that I have declared only the truth to the best of my knowledge and have not omitted anything.

---

Place, Date

---

Signature of the student