Appendix A for: Estimating Biodiversity Impact from Agricultural Statistics: an application to Food Quality Schemes in France

Sarah HUET, Clélia SIRAMI, Abdoul DIALLO, Julie REGOLO, Ainhoa IHASUSTA, Ludovic ARNAUD and Valentin BELLASSEN

1. Supplementary Method

1.1 Landscape variables

To take into account the landscape dimension in the BVIAS model applied to environmental labeling, we estimate three parameters: hedge density, agricultural plot size, crop diversity and ground cover. To calculate the variables necessary for estimating these parameters, we first extract from the 2020 French LPIS (IGN, 2020) all plots from farms registered our 2020 FADN-AC-FQS database.

To extract these plots, we use the PACAGE plot numbers (identification number of the plot in the Common Agricultural Policy registration) as primary join key. We face three situations: - The farm has the same PACAGE number in FADN and AC: then the LPIS plots with this PACAGE number are associated with that FADN-AC-FQS farm; - The PACAGE numbers differ between FADN and AC and only one of them corresponds to a LPIS PACAGE number: the LPIS plots of this PACAGE number are associated with that FADN-AC-FQS farm - No PACAGE number is registered in either the FADN or the AC, or the PACAGE numbers differ between FADN and AC, and none of them corresponds to a LPIS PACAGE number: no plot is associated with a FADN-AC-FQS holding. Such cases are found for 443 and 318 metropolitan farms, respectively. These holdings are mainly concentrated in the OTEX viticulture, market gardening and pig and/or poultry farming.

This extraction results in a subsample of the LPIS that we intersect with the Hedges layer of the BD TOPO® database (IGN and ASP, 2020) to determine the variables required to estimate the three parameters.

1.1.1 Hedge density

Hedge density is the ratio of the sum in linear meters of hedge to the area of the holding (UAA). For the calculation of linear lengths, we use the same procedure as previously used in a similar work at the scale of the regions of France (Bamière et al., 2023).

Firstly, we need to expand the LPIS blocks. Hedges are not always inside, outside or at the edge of LPIS blocks (all three cases present). However, they rarely cut a block in half. The first step is to expand the LPIS blocks by a buffer zone of 10m. The length of 10 m for the buffer zone is determined visually. It seems that beyond 10 m, it is common to make mistakes when combining a hedge with two islets which are actually separated by a road. Conversely, below 10 m, it seems common to count as "no cultivation" hedges that are clearly on the edge of a field.

Secondly, we intersected the hedge lines with the expanded blocks to determine a border length. There are four cases for each piece of hedge from the intersection: 1. The piece of hedge does not intersect any enlarged block. The piece of hedge is considered to be lined with non-agricultural uses on both sides and twice its length is assigned to "No culture". 2. The piece of hedge intersects a single expanded block. The

piece of hedge is considered to be bordered by agricultural use on one side, and non-agricultural use on the other. Its length is assigned once to the crop/grassland type of the block and another time to "No crop". 3. The hedge piece intersects two blocks. The hedge piece is considered to be bordered on both sides by agricultural use. Its length is assigned once to the type of crop/grassland of each block. 4. The piece of hedge intersects more than two blocks. Its length is then affected more than twice. This case results in an aberration (hedge with more than three sides). To avoid this, the corresponding lengths are multiplied by the ratio of the initial length and the sum of the affected lengths so that the sum of the corrected lengths is exactly equal to the initial hedge piece length.

Then, to calculate the hedge density for each farm, we distinguish four plot categories for each variable (hedge length and UAA): - Total - Permanent grassland (including grasslands, moors and alpine pastures) - Arboriculture (including vines) - Other crops

1.1.2 Mean field size

The mean field size is the ratio of the UAA to the number of plots. We distinguish four plot categories for each variable (UAA and number of plots): - Total - Permanent grassland (including grasslands, moors and alpine pastures) - Arboriculture (including vines) - Other crops

1.1.3 Crop diversity

We calculate the Shannon index (Eq. 4), using the number of crops as the number of species and the surface area as the abundance, only for arable land use type at the farm scale:

$$H' = -\sum_{i=1}^{R} p_i \cdot \ln p_i \, (4)$$

where R is the total number of crops, p is the ratio of the area for the crop i on the total arable area.

1.1.4 Ground cover

The vegetation cover data is generated from the Sentinel-2 (S2) raster data (European Space Agency, 2022). The method is based on the calculation of the NDVI (Normalized Difference Vegetation Index) (Araya et al., 2018; Bockstaller et al., 2021) for the crop year 2020 (i.e., the period from 01/10/2019 to 30/09/2020, summing 366 days as 2020 is a bissextile year). A day is counted as "green" if the index is greater than a threshold of 0.3 that separates a bare soil state to a vegetated soil state (Araya et al., 2018). From the raster data, we calculated the average number of covered days using the area statistics algorithm in QGIS (Team, 2022) for four plot categories:

- Total
- Permanent grassland (including grasslands, moors and alpine pastures)
- Arboriculture (including vines)
- Other crops

The number of uncovered days is then calculated as: Number of uncovered days $= 366 -$ number of covered days.

1.2 Model calibration and validation according to literature

In order to calibrate and validate our model, we compare estimated biodiversity levels with in situ measurements identified in the literature. These in situ measurements serve as references and we adjust the limits of the intervals of the land use types (grassland and arable), the parameters of the normalization

functions of the input variables (cultural practices and landscape variables) and the weighting coefficients of the variables in order to minimize the distance between our results and the in situ measurements selected.

1.2.1 Levels of biodiversity in different land use types

To compare the levels of biodiversity of different land use types, we base our analysis on Gallego-Zamorano et al. (2022) who estimated the richness in plant species according to nine archetypes of land use in relation to the richness in plant species in a natural space. We define the equivalent of each archetype in our modelling framework (Table A1).

Table A1: Calibration by land use type. [1] For the minimal-intensity archetypes, the observed variability is due to the idiosyncrasy of biodiversity rather than the variability of the intensity of practices: we therefore choose the average as the limit. [2] For high-intensity archetypes, variability is due to the variability of intensity of practices rather than the idiosyncrasies of biodiversity: we therefore choose the lower limit as maximum intensity of practice.

Land use archetypes (Gallego-Zamorano et al., 2022)	Equivalent in our modeling framework	Plant species richness relative to primary vegetation
Pasture, Minimal-intensity use (P-M)	BV_{LU} upper bound for grasslands ($BV_{LU,P,ha,max}$)	0,92 (P-M average)[1]
Pasture, Light- and high-intensity use (P-LH)	Pastures with all variables at the median of our grassland sample	0,57 (95% CI: 0,44-0,76)
Pasture, Light- and high-intensity use (P-LH)	BV_{LU} lower bound for grasslands ($BV_{LU,P,ha,min}$)	0,44 (P-LH lower bound)[2]
Cropland, Minimal-intensity use (C-M)	BV_{LU} upper bound for croplands ($BV_{LU,C,ha,max}$)	0,52 (C-M average)[1]
Cropland, High-intensity use (C-H)	Croplands with all variables at the median of our arable sample	0,32 (95% CI: 0,23-0,44)
Cropland, High-intensity use (C-H)	BV_{LU} lower bound for grasslands ($BV_{LU,C,ha,min}$)	0,23 (C-H lower bound)[2]

1.2.2 Effect of practices on biodiversity level

1.2.2.1 Within-field practices

In the absence of a meta-analysis of the effect of pesticides on taxa and "essential" functional groups (pollinators, vertebrates), we use aquatic invertebrates as a benchmark whose diversity is reduced by *27-42*% in an environment with high pesticide use (Beketov et al., 2013). This is consistent with the use of ecofreshwater toxicity for weighting estimated inputs of pesticides in our model (see Section 2.4.3).

We consider that within-field practices of nitrogen fertilization and pesticide application have comparable effects on biodiversity. Pesticides are as often implicated in the causes of biodiversity loss as fertilization (Sánchez-Bayo and Wyckhuys, 2019).

Table A2: Calibration by effect size. [1] n our modelling framework, the best translation of a landscape

with few pesticides as described in Beketov et al. (2013) is an average landscape with half-grassland and half-crop with few pesticides. The best expression of a landscape with an intensive use of pesticides is a landscape composed only of the most pesticide intensive crops. [2] Nitrogen fertilization impact assumed to be equal to pesticide impact, according to Sánchez-Bayo and Wyckhuys (2019). [3] Ratio of predictions after matching procedure for structural characteristics of farms (simple matching).

Archetypes	Taxa considered	Geographical area	Equivalent in our modeling framework	Expected result from literature
Low VS maximum use of pesticides (Beketov et al., 2013)	Stream invertebrates	Germany, France and Australia	Equation 5 [1]	0.35
Few VS maximum nitrogen inputs (Sánchez-Bayo and Wyckhuys, 2019)	Entomofauna	Worldwide	Equation 6	0.352
High crop diversity with 3 crop type sampled VS low crop diversity with 1 crop type sampled (Sirami et al., 2019)	Plants, bees, butterflies, hoverflies, carabids, spiders, and birds	France, United Kingdom, Germany, Spain and Canada	Equation 7	0.18
Small VS large mean field size (Sirami et al., 2019)			Equation 9	0.05
High VS low density of small woody features (Vallé et al., 2023)	Birds, bats and bush-crickets	France	Equation 8	0.20
Organic VS conventional wheat & grassland (Tuck et al., 2014)	Arthropods, birds, microbes and plants	Europe and North America	Equation 10 [3]	0.30 (wheat) & 0.15 (grassland)

$$\frac{BV_{loc,P,ha}\left(\text{pesticide use } 5^{th} \text{ percentile, other variable medians}\right)+BV_{loc,C,ha}\left(\text{pesticide use } 5^{th} \text{ percentile, other varia}\right)}{2}$$

$$\frac{BV_{loc,P,ha}\left(\text{nitrogen input } 5^{th} \text{ percentile, other variable medians}\right)+BV_{loc,C,ha}\left(\text{nitrogen input } 5^{th} \text{ percentile, other v}\right)}{2}$$

$$\frac{BV_{loc,P,ha}\left(\text{nitrogen input } 5^{th} \text{ percentile, other variable medians}\right)+BV_{loc,C,ha}\left(\text{ni}\right)}{2}$$

$$\frac{BV_{loc,ha}\left(\text{Shannon index } 75^{th} \text{ percentile of farms with 3 crops, other variable medians}\right)-BV_{loc,ha}\left(\text{Shannon index}\right)}{BV_{loc,ha}\left(\text{Shannon index } 75^{th} \text{ percentile of farms with 3 crops, other v}\right)}$$

$$\frac{BV_{loc,ha}(\text{hedge density }75^{th}\text{ percentile, other variable medians}) - BV_{loc,ha}(\text{hedge density }25^{th}\text{ percentile, other va}}{BV_{loc,ha}(\text{hedge density }75^{th}\text{ percentile, other variable medians})}$$

$$\frac{BV_{loc,ha}(\text{mean field size }25^{th}\text{ percentile, other variable medians}) - BV_{loc,ha}(\text{mean field size }75^{th}\text{ percentile, other}}{BV_{loc,ha}(\text{mean field size }25^{th}\text{ percentile, other variable medians})}$$

$$\frac{BV_{loc,ha}(\text{organic soft wheat}) - BV_{loc,ha}(\text{conventional soft wheat})}{BV_{loc,ha}(\text{organic soft wheat})}(10)$$

1.2.2.2 Landscape effects

To compare the effects of mean field size and crop diversity, we relied on Sirami et al. (2019), one of the only studies that provides effect sizes for these two variables (Table A2).

To compare the effects of hedges we use Vallé et al. (2023) which estimate the effect of small woody features (Table A2). To retrieve the average percentile corresponding to the breaking point and the average effect size of the predicted species diversity (Simpson index) of the three taxa, we used WebPlotDigitizer (version 5.2) to calculate the area under the density curve and retrieve breaking point Simpson index value (see sheet "valle_2023_WebPlotDigitizer" in Appendix C).

1.2.2.3 Comparison of labels

The effect of organic farming on biodiversity has been examined by numerous studies whose results have been aggregated in a meta-analysis comparing in situ measurements of species richness between organic and conventional agriculture (Tuck et al., 2014). We take the observed differences to calibrate the model (Table A2).

1.2.2.4 The case of grassland

In our modelling framework, since grasslands are not ploughed nor receive pesticides, we do not use these variables to calibrate the set of grassland parameters. Similarly, the crop diversity does not apply to grassland. We also considered that the effects of hedges and parcel size are negligible for grasslands. These five variables will therefore be weighted to zero for the calculation of the BVIAS of grasslands.

1.2.3 Model calibration

First, the range of possible biodiversity values (in which BV_{LU} is projected for normalization; see Section 2.1) for each land use type (arable and grassland) are defined as in Gallego-Zamorano et al. (2022) (Table A1). Secondly, the set of response function parameters ($BVC_{i,l,v}$, see Section 2.1) and the aggregation weights for each type of land use are calibrated (separately for arable and grassland). The calibration pursues the objective of minimizing the fourth power of the distance between our predicted values and the target values of Table A1 (average croplands and grasslands) and Table A2:

- Initially, the functional forms of the responses are taken from Lindner et al. (2019). Unlike Lindner's method Lindner et al. (2022) that aggregates variables by averaging them, the variable weights used for their aggregation are assigned manually such that their value corresponds to the effect size of variables listed in the literature (Table 6), with a minimum weight of 0.05 (for the negative effect sizes and the variables not listed in the literature).

- Then, both the functional forms of the responses (all constants except epsilon and gamma, which modulates the response range, and beta which define the variation direction; see code available

online) and the aggregation weights are calibrated using the basic optim function of R (Team, 2023)

- Note that the effect size of each variable (or its relative importance) results from the combination of both weights and functional forms. In principle, a very "responsive" functional form combined with low weight can confer the same importance as a very "flat" functional form combined with high weight.

2 Supplementary Tables

2.1 Descriptive statistics

*Table A3: **Descriptive statistics of the input variables for the arable land use type***

Practice	Input variable	Value	Mean	sd	Min	25th percentile	Median	75th percentile	Max
Hedge Density (linear m / ha)	A.2.1	Raw	51.06	53.92	0	15.82	34.41	70.52	2,363.62
-	-	Cut at 95th percentile	49.16	43.54	0	15.82	34.41	70.52	162.72
-	-	Normalized	0.3	0.27	0	0.1	0.21	0.43	1
Mean Field Size (ha)	A.2.2	Raw	3.22	2.04	0.01	1.82	2.79	4.1	22.51
-	-	Cut at 95th percentile	3.09	1.66	0.01	1.82	2.79	4.1	6.64
-	-	Normalized	0.47	0.25	0	0.27	0.42	0.62	1
Tillage (L diesel / ha)	A.3.1	Raw	64.05	69.25	0	21.62	49.92	86.4	2,012.82
-	-	Cut at 95th percentile	59.53	49.09	0	21.62	49.92	86.4	177.61
-	-	Normalized	0.34	0.28	0	0.12	0.28	0.49	1
Soil Cover (Number of uncovered day)	A.3.2	Raw	80.52	41.02	0	48.56	80.95	109.66	309.43
-	-	Cut at 95th percentile	79.43	38.73	0	48.56	80.95	109.66	145.43
-	-	Normalized	0.55	0.27	0	0.33	0.56	0.75	1

Crop Diversity (Shannon Index)	A.3.3	Raw	3.57	1.39	1	2.57	3.49	4.46	9.89
-	-	Cut at 95th percentile	3.5	1.24	1	2.57	3.49	4.46	5.51
-	-	Normalized	0.63	0.23	0.18	0.47	0.63	0.81	1
Share of mineral fertilization (%)	A.4.3	Raw	0.47	0.33	0	0.17	0.45	0.78	1
-	-	Cut at 95th percentile	0.47	0.33	0	0.17	0.45	0.78	0.95
-	-	Normalized	0.49	0.35	0	0.18	0.48	0.83	1
Nitrogen fertilization (kg N / ha)	A.4.5	Raw	172.03	134.46	0	76.63	173.56	242.16	5,585.04
-	-	Cut at 95th percentile	167.06	109.17	0	76.63	173.56	242.16	382.79
-	-	Normalized	0.44	0.29	0	0.2	0.45	0.63	1
Pesticides (€~TFI~UDNu / ha)	A.5.1	Raw	150.09	204.57	0	54.85	108.16	178.85	10,078.98
-	-	Cut at 95th percentile	130.93	104.87	0	54.85	108.16	178.85	414.78
-	-	Normalized	0.32	0.25	0	0.13	0.26	0.43	1

*Table A4: **Descriptive statistics of the input variables for the grassland land use type***

Practice	Input variable	Value	Mean	sd	Min	25th percentile	Median	75th percentile	Max
Share of mineral fertilization (%)	A.4.3	Raw	0.22	0.18	0	0.07	0.2	0.31	1
-	-	Cut at 95th percentile	0.21	0.16	0	0.07	0.2	0.31	0.56
-	-	Normalized	0.38	0.29	0	0.13	0.36	0.55	1

Nitrogen fertilization (kg N / ha)	A.4.5	Raw	106.26	192.04	3.52	83.04	102.12	118.43	8,773.87
-	-	Cut at 95th percentile	97.95	29.12	3.52	83.04	102.12	118.43	142.72
-	-	Normalized	0.69	0.2	0.02	0.58	0.72	0.83	1

2.2 Optimization results

Table A5: **Distances between BVIAS results and the literature values**, *before and after optimization*

Archetypes	Land use type	Literature values	Biodiversity Value before optimization	Distance before optimization	Biodiversity Value after optimization	Distance after optimization
Median crop	arable	0.32	0.34	1.525×10^{-7}	0.271	5.885×10^{-6}
Hedge Density	arable	0.2	0.082	1.946×10^{-4}	0.116	4.935×10^{-5}
Mean Field Size	arable	0.05	0.003	4.993×10^{-6}	0.018	1.066×10^{-6}
Crop diversity	arable	0.18	0.059	2.113×10^{-4}	0.069	1.511×10^{-4}
Fertilization	arable	0.35	0.196	5.651×10^{-4}	0.264	5.453×10^{-5}
Pesticide	arable	0.35	0.259	7.003×10^{-5}	0.42	2.469×10^{-5}
Organic wheat / Conventional wheat	arable	0.296	0.206	6.611×10^{-5}	0.22	3.226×10^{-5}
Median grassland	grassland	0.57	0.491	3.920×10^{-5}	0.567	5.926×10^{-11}
Fertilization	grassland	0.35	0.221	2.796×10^{-4}	0.348	1.946×10^{-11}
Organic wheat / Conventional wheat	grassland	0.145	0.152	1.498×10^{-9}	0.221	3.205×10^{-5}
Mean distance	-	-	-	1.431×10^{-4}	-	3.509×10^{-5}

Table A6: **Distances between BVIAS results and the literature values**, *before and after optimization*

Land use type	Input variable	Description	Initial weights	Optimized weights
arable	A.2.1	Hedge density (linear meter / ha)	0.2	0.396
arable	A.2.2	Mean Field Size (ha)	0.05	0.05
arable	A.3.1	Tillage (L diesel / ha)	0.05	0.05

arable	A.3.2	Ground cover (number of uncovered day)	0.05	0.05
arable	A.3.3	Crop Diversity (Shannon Index)	0.18	0.382
arable	A.4.3	Share of mineral fertilizer (%)	0.05	0.05
arable	A.4.5	Fertilization (kg N / ha)	0.35	0.986
arable	A.5.1	Pesticides (€~TFI~UDNu / ha)	0.35	0.979
grassland	A.4.3	Share of mineral fertilizer (%)	0.25	0.604
grassland	A.4.5	Fertilization (kg N / ha)	0.75	0.766

2.3 Practices

2.3.1 Crops

*Table A7: **Differences in crop cultivation practices between certified and conventional cereal production.** Significant differences before (Tukey HSD test, p-value ≤ 0.05) and after matching (paired t-test, Bonferroni adjusted p-value ≤ 0.05) are written in bold.*

See "crops" sheet in Appendix D

2.3.2 Husbandry

*Table A8: **Differences in husbandry practices between certified and conventional milk production.** Significant differences before (Tukey HSD test, p-value ≤ 0.05) and after matching (paired t-test, Bonferroni adjusted p-value ≤ 0.05) are written in bold.*

See "husbandry" sheet in Appendix D

2.3.3 Feed crops

*Table A9: **Differences in livestock feed production practices between certified and conventional milk production.** Significant differences before (Tukey HSD test, p-value ≤ 0.05) and after matching (paired t-test, Bonferroni adjusted p-value ≤ 0.05) are written in bold.*

See "feed" sheet in Appendix D

9.2.4 Nutritional values

Table A10: Nutritional values as retrieve from Agence nationale de sécurité sanitaire (2020).

Product	Description in the Ciqual data base	Energy, Regulation EU No 1169/2011 (kcal/100g)	Protein (g/100g)
Wheat	Wheat flour, type 110	343	9,61

Product	Description in the Ciqual data base	Energy, Regulation EU No 1169/2011 (kcal/100g)	Protein (g/100g)
Milk	Milk, whole, pasteurised	56,5	3,3

3. Section 3: Supplementary Figures
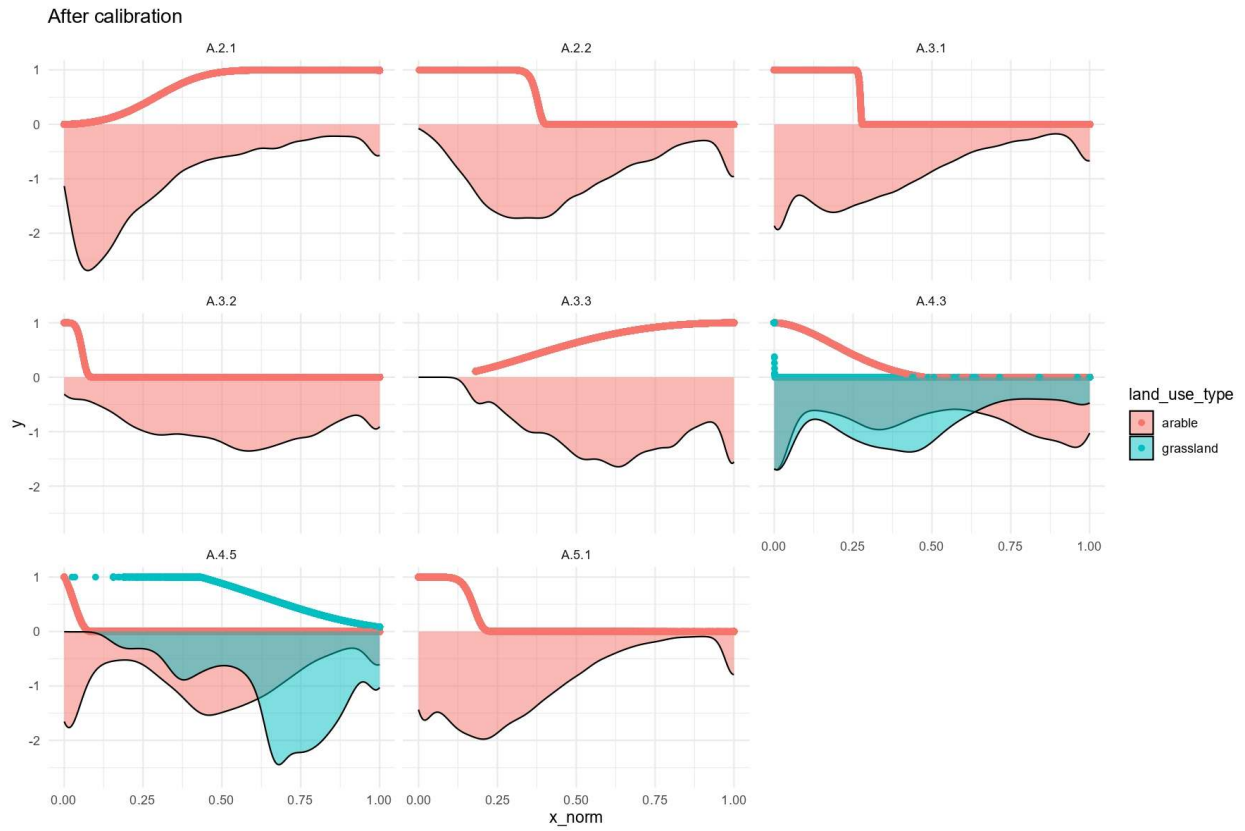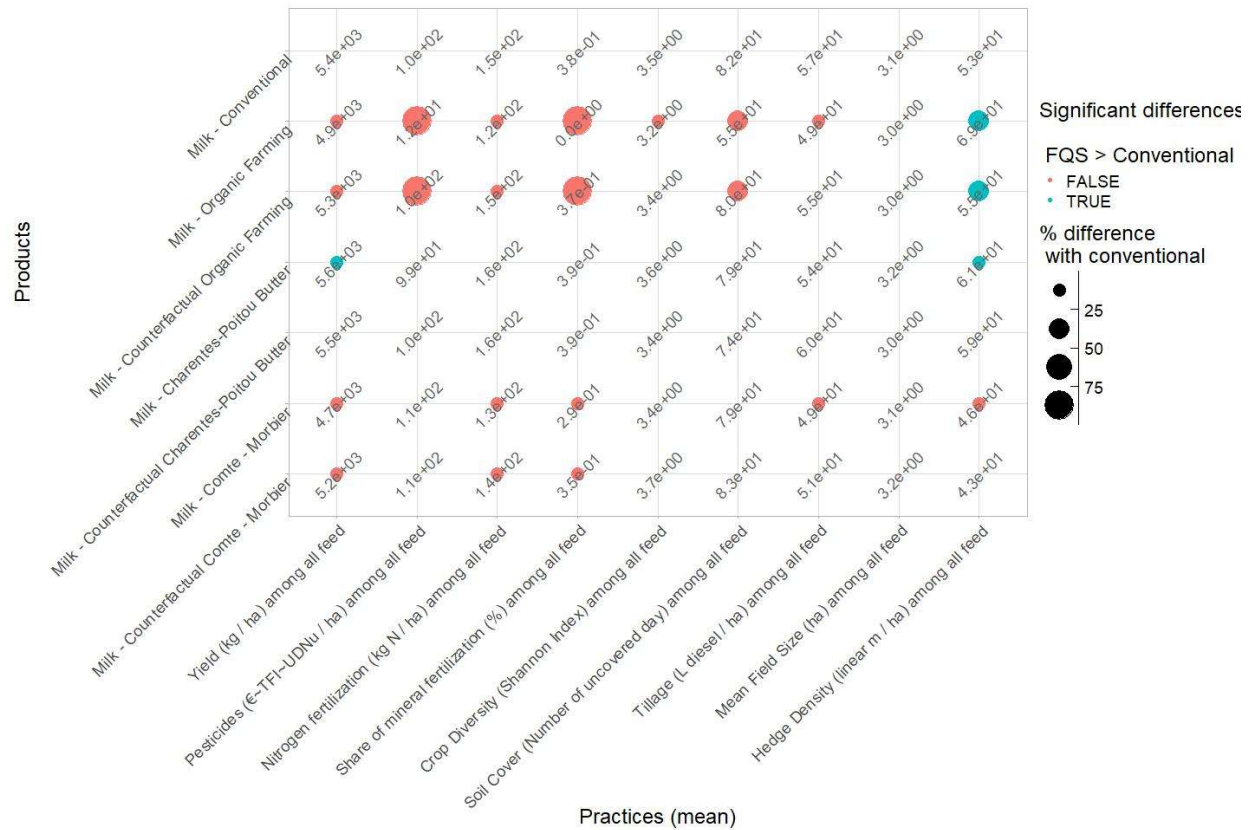

Distribution of input variables for arable land use type

Distribution of input variables for grassland land use type
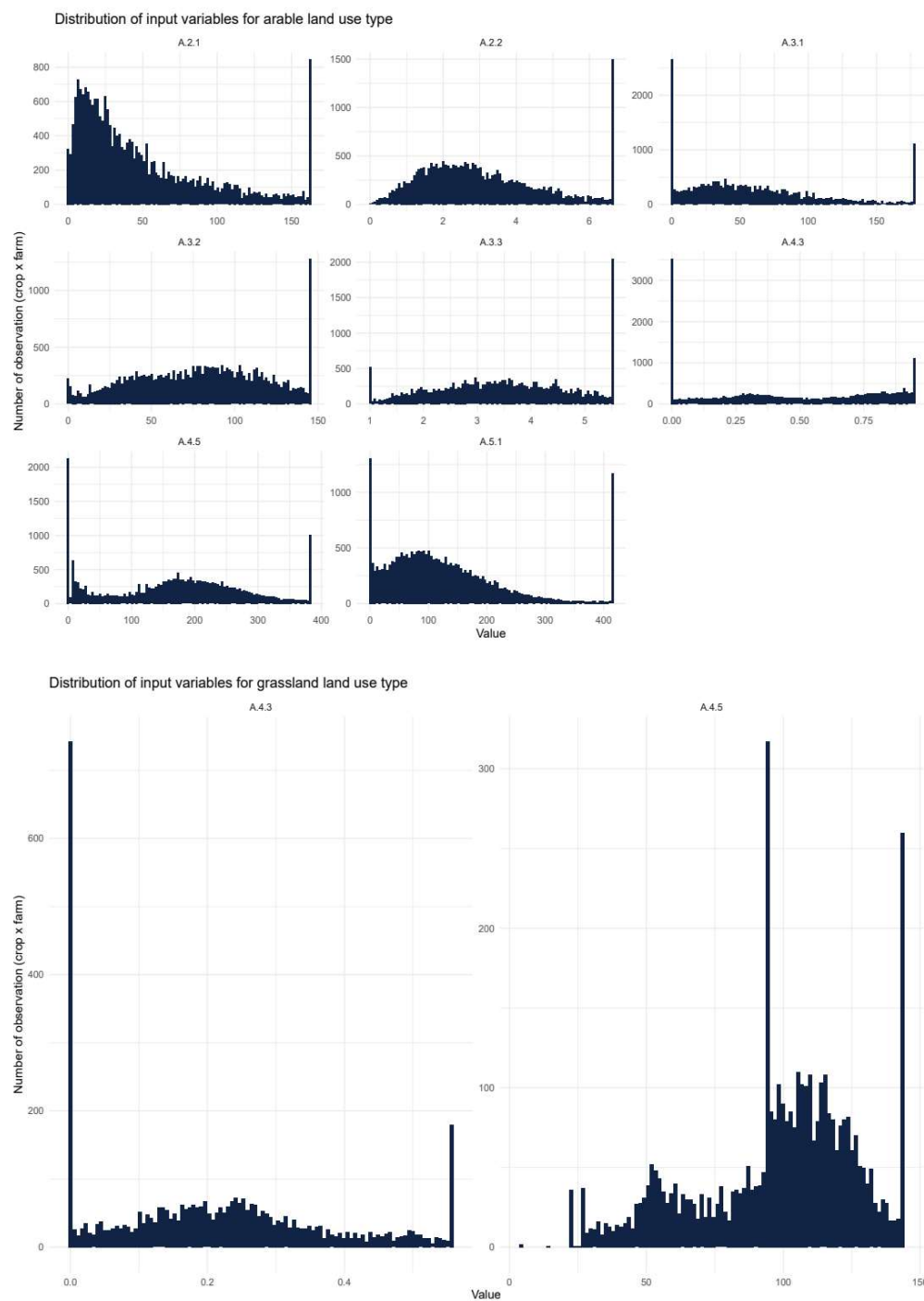
*Figure A1: **Distribution of input variables** for arable and grassland land use types*


*Figure A2: **Biodiversity value contribution** as function of the normalized input variables and their distribution, by land use type*
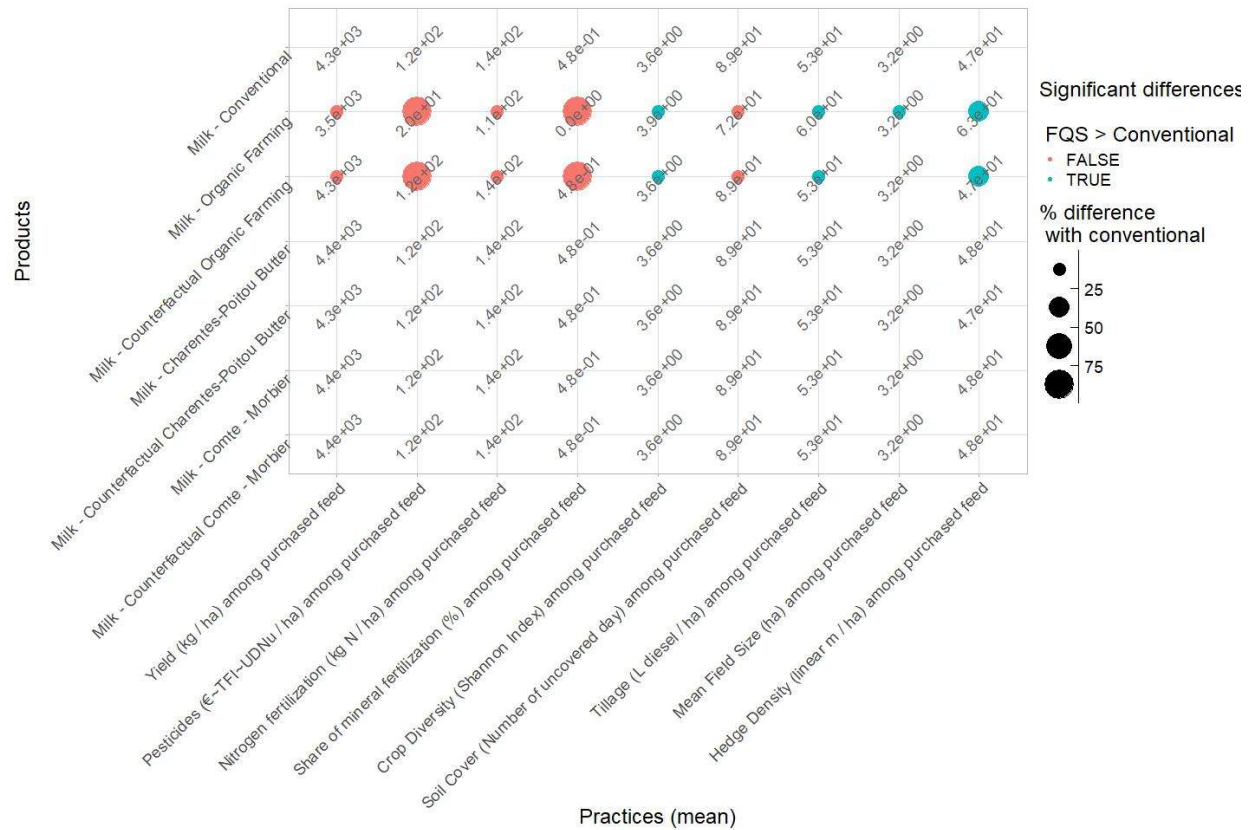
*Figure A3: **Feed production practices (average of all feed) for some FQSs and conventional milk productions.** The mean is displayed with a dot for significant differences between FQS and either conventional (dot displayed on FQS row; Tukey HSD p-value ≤ 0.05) or counterfactual (dot displayed on counterfactual row; paired t-test, Bonferroni adjusted p-value ≤ 0.05). The dot color depends on whether the comparison between the FQS and either conventional or counterfactual is positive (blue dot) or negative (red dot). The dot size depends on the proportion of the difference between the FQS and either conventional or counterfactual. The full results for all FQSs and products studied are presented in Table A9.*
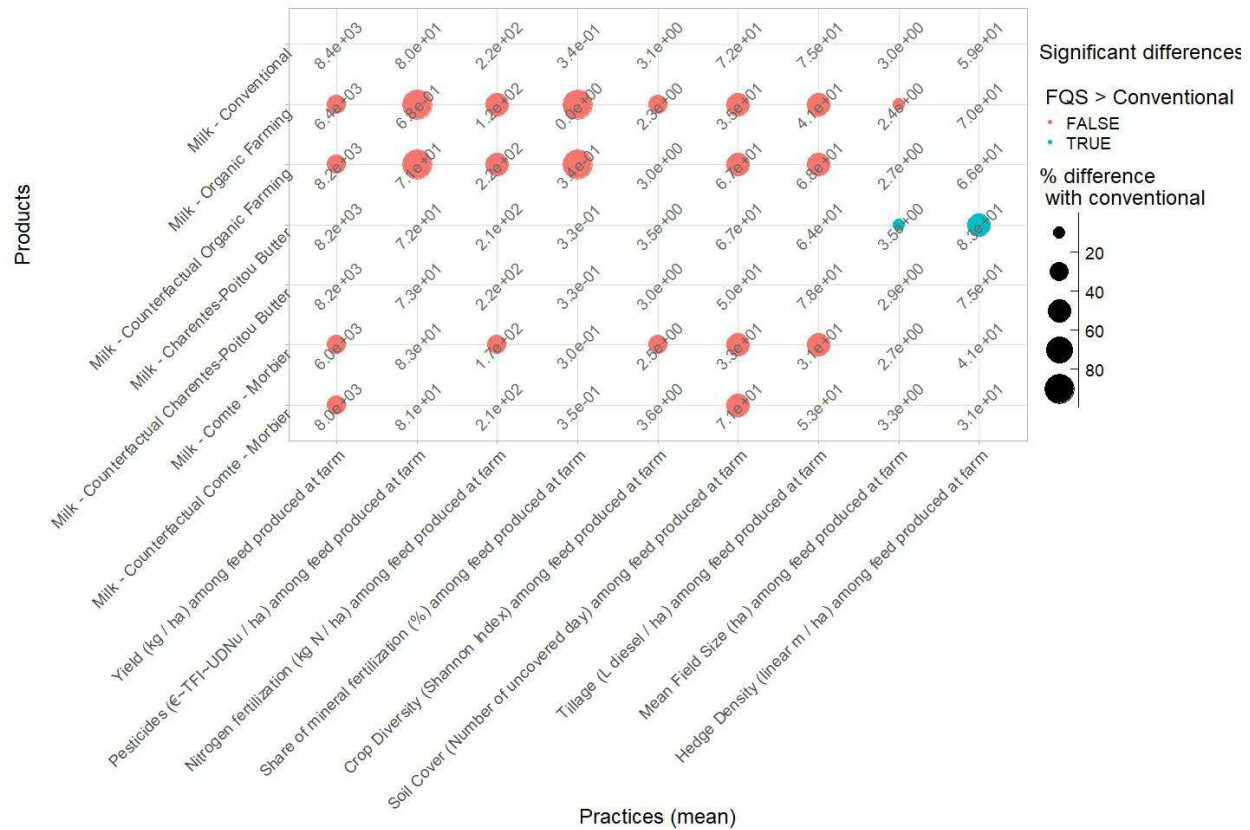
*Figure A4: **Feed production practices (average of feed produced at farm) for some FQSs and conventional milk productions.** The mean is displayed with a dot for significant differences between FQS and either conventional (dot displayed on FQS row; Tukey HSD p-value ≤ 0.05) or counterfactual (dot displayed on counterfactual row; paired t-test, Bonferroni adjusted p-value ≤ 0.05). The dot color depends on whether the comparison between the FQS and either conventional or counterfactual is positive (blue dot) or negative (red dot). The dot size depends on the proportion of the difference between the FQS and either conventional or counterfactual. The full results for all FQSs and products studied are presented in Table A9.*
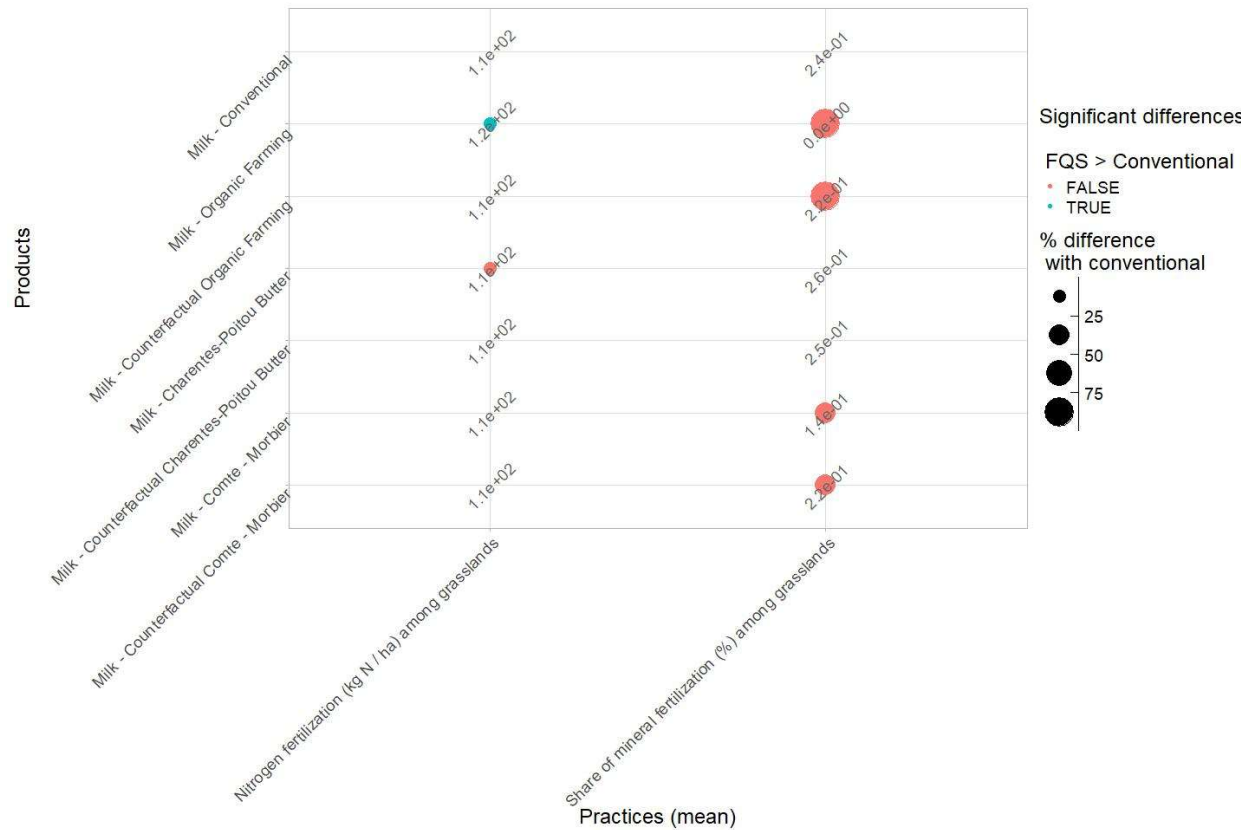
*Figure A5: **Feed production practices (average of grasslands) for some FQSs and conventional milk productions.** The mean is displayed with a dot for significant differences between FQS and either conventional (dot displayed on FQS row; Tukey HSD p-value ≤ 0.05) or counterfactual (dot displayed on counterfactual row; paired t-test, Bonferroni adjusted p-value ≤ 0.05). The dot color depends on whether the comparison between the FQS and either conventional or counterfactual is positive (blue dot) or negative (red dot). The dot size depends on the proportion of the difference between the FQS and either conventional or counterfactual. The full results for all FQSs and products studied are presented in Table A9.*

*Figure A6: **Feed production practices (average of purchased feed) for some FQSs and conventional milk productions.** The mean is displayed with a dot for significant differences between FQS and either conventional (dot displayed on FQS row; Tukey HSD p-value ≤ 0.05) or counterfactual (dot displayed on counterfactual row; paired t-test, Bonferroni adjusted p-value ≤ 0.05). The dot color depends on whether the comparison between the FQS and either conventional or counterfactual is positive (blue dot) or negative (red dot). The dot size depends on the proportion of the difference between the FQS and either conventional or counterfactual. The full results for all FQSs and products studied are presented in Table A9.*