

# Quantity based indicators fail to identify extreme pesticide risks

**Journal Article** 

Author(s): <u>Möhring, Niklas</u> (b); Gaba, Sabrina; <u>Finger, Robert</u> (b)

Publication date: 2019-01-01

Permanent link: https://doi.org/10.3929/ethz-b-000288799

Rights / license: Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

**Originally published in:** Science of The Total Environment 646, <u>https://doi.org/10.1016/j.scitotenv.2018.07.287</u>

This page was generated automatically upon download from the <u>ETH Zurich Research Collection</u>. For more information, please consult the <u>Terms of use</u>.

## **1** Quantity based indicators fail to identify extreme pesticide risks

## 2 Authors: Niklas Möhring<sup>a\*</sup>, Sabrina Gaba<sup>b</sup>, Robert Finger<sup>a</sup>

- <sup>a</sup> Agricultural Economics and Policy Group, ETH Zürich
- <sup>4</sup> <sup>b</sup> USC 1339, Centre d'Etudes Biologiques de Chizé, INRA, 79360 Villiers en
- 5 Bois, France
- 6 \*Corresponding author: <u>nmoehring@ethz.ch</u>
- 7

#### 8 Graphical Abstract



Differ by Indicator (I vs. II)

#### 13 Abstract.

14 As a matter of policy, minimizing human health and environmental risks associated with pesticide use is a major 15 challenge but necessary for improving agricultural sustainability. Efficient and effective policies that encourage 16 the use of less risky pesticides, such as pesticide taxes, necessitate a precise and realistic quantification of 17 potential adverse effects. Various indicators are currently utilized in policies and they focus mainly on a purely 18 quantitative dimension of the pesticides used, which can lead potentially to unfavorable outcomes of pesticide 19 policies. A unique dataset applied to pesticide use by Swiss farmers in winter wheat and potato production, 20 demonstrates that on average the two most important quantitative indicators show a significant correlation with 21 pesticide risks as expressed by the Danish Load Indicator. However, they have almost no explanatory power for 22 extreme risks (e.g. most intensive use patterns for pesticides with unfavorable toxicity profiles). Results remain 23 stable over a range of aggregation levels, from application- to farm-level indicators of pesticide use. These 24 findings render the commonly used, quantitative indicators ineffective to reduce potential environmental and 25 human health risks of pesticides and, in the worst case, lead to misinformed market-based pesticide policies

- 26 consequential to National Action Plans.
- 27 Keywords: Pesticides, Pesticide risks, Pesticide policies, Pesticide indicators

#### 28 1. Introduction

- 29 Current agricultural production systems often rely on an intensive use of pesticides and other
- 30 agrochemical inputs. Pesticides are tightly regulated in many countries, subject to rigorous testing
- 31 and highly conservative risk assessment paradigms. However, the use of pesticides can still present
- 32 potential risks to human and environmental health (Strange and Scott 2005; Damalas 2009; Beketov
- et al. 2013; Malaj et al. 2014; Stehle and Schulz 2015). The introduction of effective policies to reduce
- 34 adverse effects of pesticides, while maintaining production levels, is a major challenge on the way to
- 35 achieving improved sustainability in agriculture (Tilman et al. 2002). In the European Union, the US
- 36 and China, public policies have been established to address pesticide risks and stricter pesticide
- 37 policies are also being implemented in several other countries (Lefebvre et al. 2015; Pimentel and
- 38 Burgess 2014; Osteen and Fernandez-Cornejo 2013; Zhang and Wen 2008; Sun et al. 2012, MAAF
- 39 2015, Bundesrat 2017, Böcker et al., 2018). However, the effectiveness of current policies has been
- 40 questioned recently (e.g. Hossard et al. 2017, Finger, 2018). The focus of this research was to
- 41 evaluate pesticide use/risk indicators utilized for the purposes of informing market-based pesticide
- 42 policy and high-level reduction targets related to National Action Plans.
- Setting up policies which promote a reduction in the impacts of pesticides on the environment and
  human health is far from straightforward. Pesticides are highly heterogeneous with respect to
  properties, application regimes and their potential impact on the environment and human health.
- 46 For instance, in the EU alone, a range of 494 active substances for pesticides, with potentially
- different adverse effects, are currently authorized (EU 2017). The choice of suitable pesticide
- 48 indicators to quantify pesticide use is therefore essential to define efficient and effective policy
- 49 measures.
- 50 Currently, implemented indicators differ significantly. For example, there are simple, purely
- 51 quantitative indicators like the Quantity of Active Ingredients (QA) and the Treatment Frequency
- 52 Index (TFI), which abstract from inherent pesticide properties, to very detailed, risk-adjusted
- 53 indicators such as the Load Index (LI). QA is a simple measure of kilograms of pesticides used per
- area. TFI measures the intensity of applications, i.e. quantity applied per unit of cropped area in

55 relation to the recommended dosage (Coll and Wajnberg 2017). The LI indicator accounts for 56 application intensity as well as a broad range of potential environmental and health risks for each 57 pesticide (Miljøministeriet 2012; Kudsk et al. 2018). The indicator chosen differs across countries and 58 institutions. For example, France uses the QA and TFI indicators to set targets for pesticide policies 59 (MAA 2017). Furthermore, QA and TFI are applied as key indicators for pesticide use statistics by 60 institutions worldwide (Eurostat 2017; USDA 2017; JKI 2017; MAA 2017) and are standard indicators 61 for studies on the economics of pesticide use (Ghimire and Woodward 2013; Hossard et al. 2014; 62 Gaba et al. 2016; Perry et al. 2016; Lechenet et al. 2017). The Danish Load Index (LI) is currently the only risk-based indicator implemented in European pesticide policies which holistically assesses 63 64 potential environmental and health risks of pesticides on a product level. As with the purely 65 quantitative QA and TFI indicators, this allows pesticide risks to be upscaled along a gradient of 66 temporal and spatial resolution (Kudsk et al., 2018). Since 2013, it has been used in Denmark for the 67 assessment of policy targets and at the same time as the basis for pesticide taxation (Böcker and 68 Finger 2016; Kudsk et al. 2018).

69 However, it is hypothesized that large quantities of pesticides, as indicated by high QA or TFI 70 indicator values, may not inherently mean higher risks for human health and the environment (e.g. 71 high LI indicator values). For example, Kniss (2017) recently showed that herbicide use trends for 72 major crops in the US were reversed when the assessment was switched from quantity-based to 73 toxicity-based indicators. More importantly, the use of quantity-based indicators compared to risk-74 adjusted indicators may lead to major shifts in policy targets if indicators rank pesticide use 75 inconsistently. These policy targets include, for example, the reduction of temporal or spatial 76 "hotspots" and extreme application regimes (over a given cropping season), as extreme applications 77 are major contributors to the negative effects of pesticides on human health (Larsen et al. 2017) and 78 the environment (Releya and Hovermann 2006; Gordon et al. 2012; Bundschuh et al. 2013, Topping 79 and Elmeros 2016). Along these lines, Larsen et al. (2017) conclude that there is a need for the 80 implementation of pesticide policies that tackle extreme applications. Market-based policy measures 81 such as specific taxes, quotas or subsidies can complement regulatory frameworks and admission 82 procedures in achieving this target (Baumol and Oates, 1988). However, a misspecification of policy 83 targets may result in biased policy incentives and finally, adverse policy outcomes. Depending on the 84 degree of inconsistencies between indicators, indicator choice may therefore have severe 85 implications for policy outcomes. Although a well-informed policy discussion is of vital importance, 86 there is a lack of studies which quantify the extent of inconsistencies between pesticide use 87 indicators in a common, robust framework.

88 This research gap was addressed by investigating the consistency of pesticide use rankings between 89 the purely quantitative, but widely used QA and TFI indicators and the LI indicator. The focus goes 90 beyond average consistency by analyzing consistency for "extreme" application regimes as well as 91 temporal and spatial "hotspots". Throughout the article, we refer to extreme applications as the 92 most intensive use patterns and highest risk scenarios and profiles compared to all other 93 applications, i.e. the upper tails of the observed distributions of pesticide applications. In the 94 analysis, across-indicator consistency was tested using a unique panel dataset of pesticide 95 applications in real farming conditions in Switzerland for two major crops, potatoes and winter 96 wheat. These crops were chosen because potatoes are characterized by the highest average 97 pesticide use and winter wheat is the most abundant crop in European (and Swiss) arable crop 98 production. Pesticide application patterns of farmers, including choices made regarding the products

- 99 used, their concentration, spatial distribution of application, and their timing, are strongly
- 100 heterogeneous across farmers. A comparison of indicators on "real" application data was therefore
- 101 necessary to derive meaningful policy recommendations, especially regarding risks from the most
- 102 intensive pesticide use patterns. The analysis started off by comparing the structure of the indicators.
- 103 Then correlation coefficients were used to test the consistency of indicator rankings over the whole
- 104 distribution. Secondly, copulas were used to analyze tail dependence between pesticide use
- indicators. This meant focus could be placed on across-indicator consistency for extremes, i.e.
- 106 observations in the tails of the distribution. The detailed dataset allowed the robustness of results to
- 107 be assessed on different aggregation levels and for different pesticide types (i.e. all pesticides,
- 108 herbicides and fungicides).
- 109 Current pesticide policies aim to reduce negative environmental and health effects of pesticides.
- 110 Market-based policy measures achieve this target by incentivizing a change in the farmers'
- application behavior. The objective of this paper is to show that the choice of underlying pesticide
- 112 indicators can crucially shift policy targets and incentives, and in the worst-case lead to adverse
- 113 outcomes of pesticide policies. The analysis emphasizes that the comparison of indicators should not
- 114 merely be based on their average fit. In fact, quantity-based indicators are found to be unsuitable to
- 115 proxy high-risk situations.

#### 116 2. Background

#### 117 2.1. Current pesticide policies

118 Since 2012, European Union member states have to draw up National Action Plans (NAPs) for a "Sustainable use of pesticides" with the goal of reducing "risks and impacts of pesticide use on 119 120 human health and the environment" (Directive 2009/128/EC). The revision of existing, and 121 implementation of new pesticide policies is an ongoing process. In the EU, Directive 2009/128/EC 122 demands that EU member states review NAPs every five years and the EU commission has recently 123 announced a "REFIT" evaluation of all EU pesticide legislation (EC 2015). Furthermore, adverse 124 effects from pesticide use are at the top of the policy agenda in other countries like Switzerland, the 125 US or China (Bundesrat 2017, Pimentel and Burgess 2014; Osteen and Fernandez-Cornejo 2013; 126 Zhang and Wen 2008; Sun et al. 2012). Policy measures within the framework of the NAPs include 127 training and extension services, the promotion of Integrated Pest Management, inspection and 128 renewal of equipment, information and awareness raising, as well as market-based policy measures 129 like pesticide taxes (Böcker and Finger 2016). Market-based policies complement existing regulatory 130 frameworks and pesticide admission procedures (e.g. see regulations EC No. 1107/2009 and EC No. 131 396/2005 in the EU and the Federal Insecticide, Fungicide and Rodenticide Act and the Federal Food, 132 Drug and Cosmetic Act in the USA). Their mix provides an "optimal regulatory strategy" (Skevas et al. 133 2013, p.97) to reduce externalities from pesticide use on the environment and human health. The 134 analysis in this article focuses on the incentives of market-based pesticide policies and the definition 135 of pesticide reduction targets. See e.g. Storck et al. (2017), Coll and Wajnberg (2017) and EC SAM 136 (2018) for a discussion of current pesticide regulations and admission procedures.

#### 137 2.2. Utilized pesticide indicators

- 138 The design of pesticide policies, i.e. setting and reviewing targets and defining pesticide indicators, is
- 139 complex due to the heterogeneity of available pesticide products, production environments and
- 140 farming practices. Worldwide, many countries have reported national targets for the reduction of

141 impacts from pesticide use. In Europe, the majority of EU member states as well as Switzerland have 142 defined risk reduction targets - yet, nine EU countries also report pesticide use reduction targets. 143 However, a clear definition of measurable targets is less common. For example, only five EU member 144 states have so far set "high level, measurable reduction targets" according to EU DGHFS (2017). Out 145 of these five member states, four have defined risk reduction targets (Belgium, Denmark, Greece and 146 Germany) and one has defined a use reduction target (France). Risk indicators utilized in these 147 member states include the POCER indicator in Belgium (Vercruysse and Steurbaut 2002, Claeys et al. 148 2005, Houbraken et al. 2016), the Load Indicator in Denmark (Kudsk et al. 2018), an environmental and human risk assessment in Greece (Tsaboula et al. 2016) and the SYNOPS indicator in Germany 149 150 (Gutsche and Rossberg 1997, Strassemeyer et al. 2017). A detailed spatial mapping of pesticide use 151 and associated risks, using pesticide reports, is for example already possible in California based on 152 the CalPip database (http://calpip.cdpr.ca.gov/) or in Denmark (Kudsk et al. 2018). An extensive 153 discussion of existing pesticide risk indicators does not lie within the scope of this article, but we 154 refer to Labite et al. (2011), OECD (2016) and Yang et al. (2017) for an overview and a discussion. The 155 analysis in this article is restricted to three indicators. More specifically, the widely used, quantity-156 based indicators QA and TFI are compared to the Danish, risk-based LI indicator. QA, TFI and LI are 157 chosen because they are i) currently implemented in the framework of the NAP; ii) allow pesticide 158 use to be scaled to different aggregation levels (product, application, farm, landscape); and iii) can be 159 computed based on pesticide administration data. They are therefore policy-relevant, can be used as 160 a foundation for market-based policy instruments (such as taxes) and can be easily implemented in 161 other countries. Note that easy scalability and low data requirements are accompanied by the 162 drawback that the indicators utilized do not consider the application context (e.g. timing, weather 163 conditions, soil, and distance to bodies of water) and different risk components are aggregated (in 164 the case of the LI indicator).

#### 165 3. Methodology and Data

#### 166 **3.1 Indicator calculation**

The pesticide use records are based on observations of single applications from various farmers.
Farmers may grow the same crop on several fields of varying size. The observations include a wide
range of different pesticides and span across different years (see Data, Section 3.5 below for a
detailed overview). Thus, the data exhibit a high level of spatial and temporal heterogeneity of
pesticide use. This also implies that pesticide use indicators cannot be simply transformed into each
other without knowledge of field level information. Figure 1 gives an overview of the procedure used
for computation of the QA, TFI and LI indicators.



#### 175 Fig. 1. Intermediate steps in the computation of the QA, TFI and LI indicators

176 The pesticide products applied (i.e. A, B and C) may contain one or several active ingredients in

177 different concentrations. It is possible that some products contain similar, or identical active

178 ingredients. Products are applied in different quantities and on different fields (or parts of fields). The

179 computation procedure for the QA, TFI and LI indicators is described in detail below. All indicators

180 are computed on different aggregation levels.

181 Firstly, all applications were standardized to obtain per ha measures of pesticide use per product.

182 (1) Pesticide Use<sub>ijk</sub> per ha =  $\sum_{l=1}^{n}$  Pesticide Use<sub>ijkl</sub> in kg \*  $\frac{application surface in ha}{surface field in ha}$ 

- where, for a given crop and year (indicators omitted for clarity), i denotes the farm, j the field, k
  the pesticide product and *l* a single application of the product.
- The Quantity of Active Ingredient (QA) indicator is a simple indicator of the quantity of active
  substances used per ha (Eurostat 2017; USDA 2017). It was first computed on a product level:

187 (2)  $QA_{ijk} = Pesticide Use_{ijk} * concentration of active ingredient_k$ 

188 where, for a given crop and year, i denotes the farm, j the field and k the pesticide product.

Field-level pesticide use was computed in a second step. Finally, farm-level pesticide use wascomputed as the area weighed mean over all fields per farm, crop and year.

191 (3) 
$$QA_{ij} = \sum_{k=1}^{m} QA_{ijk}$$

192 (4)  $QA_i = \sum_{j=1}^n QA_{ij} * \frac{surface_j}{surface_i}$ 

where, for a given crop and year, i denotes the farm, j the field, k the pesticide product and *n* thenumber of fields per farmer, crop and year.

The Treatment Frequency Index (TFI) indicator is an indicator of pesticide use intensity (Coll andWajnberg 2017). It was first computed on a product level:

197 (5) 
$$TFI_{ijk} = \frac{Pesticide Use_{ijk}}{Recommended (standard) dosage for product k per ha in the given culture}$$

where, for a given crop and year, i denotes the farm, j the field and k the pesticide product. If a
range of possible dosages was indicated the maximum dosage was taken, respectively.

Field-level pesticide use was computed in a second step. Finally, farm-level pesticide use was computed as the weighed mean over all fields per farmer, crop and year.

202 (6) 
$$TFI_{ij} = \sum_{k=1}^{m} TFI_{ijk}$$

203 (7)  $TFI_i = \sum_{j=1}^n TFI_{ij} * \frac{surface_j}{surface_i}$ 

where, for a given crop and year, i denotes the farm, j the field, k the pesticide product and *n* the number of fields per farmer, crop and year.

The Danish Load Index (LI) indicator (Miljøministeriet 2012; Kudsk et al. 2018) is an indicator which takes into account health, fate and toxicity properties of pesticides. It was first computed on a product level:

209 (8) 
$$LI_{ijk} = \frac{Pesticide Use_{ijk}}{Recommended (standard) dosage of product k per ha in the given culture} * Load_k$$

where, for a given crop and year, i denotes the farm, j the field and k the pesticide product. *Load<sub>k</sub>*denotes the potential, product specific environmental and health impacts of product *k*. More
precisely, Load is the sum of the Human Health Load, Environmental Fate Load and

- 213 Environmental Toxicity Load. The Human Health Load is based on Hazard and Precautionary
- 214 Statements; the Environmental Fate Load on half-time in soil, half-time in water, bio-
- 215 concentration factors and the SCI-GROW (Screening Concentration in Ground Water) index; the
- 216 Environmental Toxicity Load on long- term toxicity for fish, daphnia and earthworms and short-

217 term toxicity for birds, mammals, daphnia, algae, aquatic plants, earthworms and bees (see e.g.

- 218 Kudks et al., 2018). Human Health Load, Environmental Fate Load and Environmental Toxicity
- 219 Load was first computed separately. More specifically, each product was compared to all other
- 220 products for each sub-indicator. The relative scores obtained for all sub-indicators were then
- summed up in the respective categories Human Health, Environmental Fate and Environmental
- Toxicity. Thus, a single value per product was obtained for each of the three Load categories.
   Load values in the three categories were finally added up to a total Load, using equal weights for
- 224 each category (sub-indicator scores were weighted according to the scheme in Miljøministeriet
- 225 2012).

Field-level pesticide use was computed in a second step. Finally, farm-level LI of pesticide use wascomputed as the weighed mean over all fields per farmer, crop and year:

228 (9) 
$$LI_{ij} = \sum_{k=1}^{m} LI_{ijk}$$

229 (10)
$$LI_i = \sum_{j=1}^n LI_{ij} * \frac{surface_j}{surface_i}$$

where, for a given crop and year, i denotes the farm, j the field, k the pesticide product and *n* thenumber of fields per farmer, crop and year.

#### 232 **3.2 Confidence intervals and correlation coefficients**

233 Confidence intervals for the respective moments were computed to assess the statistical significance 234 of differences in the distribution moments. It was then checked to see if confidence intervals overlap. 235 If confidence intervals do not overlap, the hypotheses of significantly different parameters can be 236 accepted (given the chosen error rate). In this way it can, for example, be shown that two 237 distributions have significantly different means, variances or skewness. Payton et al. (2003) show 238 that 85% of the confidence intervals should be used to check for overlaps, given a maximum error 239 rate of 5% to falsely reject the null hypothesis. To compute confidence intervals, 10 000 non-240 parametric bootstrap replications were drawn from the empirical distribution of the data, from which a bootstrap cumulative distribution function and confidence intervals were then calculated 241 242 (e.g. Davison and Hinkley 1997).

- 243 Correlation coefficients were computed first to assess the dependence between indicators. As
- 244 dependence between indicators was highly non-linear, a measure of ordinal dependence was
- chosen. Kendall's tau was selected due to its direct connection to the concept of copulas (see Section
- 246 3.3) and its better efficiency properties and robustness compared to Spearman's rho (Croux and
- 247 Dehon 2010). Given two sets of observations of the joint random variables *X* and *Y*, Kendall's tau is
- 248 composed of the difference between concordant and disconcordant pairs of values  $x_i$  and  $y_i$ , i.e.
- 249 pairs where rankings "agree" or "disagree". See Croux and Dehon (2010) for a formal definition.

### 250 3.3 Copulas theoretical background

- 251 (Bivariate) copulas are functions that allow (two) distribution functions to be "coupled" to a
- 252 multivariate distribution function. Formally, consider the two (continuous) distribution functions
- 253  $F(x) = P[X \le x]$  and  $G(y) = P[Y \le y]$  of the random variables X and Y. Then the joint distribution
- function H(x, y) can be expressed by the copula function C(.) as follows:
- 255 (11) H(x, y) = C[F(x), G(y)]

- where the copula function C(.) captures the dependence structure between the two distribution
- 257 functions. Empirical applications of the copula concept can be found in various disciplines such as
- finance and insurance, as well as in applied agriculture. Copulas can also be used to model joint
- 259 occurrence of unfavorable weather conditions for agricultural production. These models then serve
- to improve insurances against weather shocks, where the tail of the loss distribution is particularly
- relevant (Xu et al. 2010, Okhrin et al. 2013). Copulas are linked directly to concordant measures of
- 262 dependence, such as Kendall's tau (Nelsen 2006). Indeed, such measures of dependence are
- functions of the copula value (see Nelsen (2006) for an overview). This can be illustrated for Kendall's

tau as follows (where F(x) = u and G(y) = v):

265 (12) 
$$\tau = 4 * E[C(u, v)] - 1$$

The upper and lower tail dependence coefficients describe dependence in the upper and lower tails of bivariate distributions. Given the copula function defined in equation (11), the upper and lower tail dependence coefficients (TDCs),  $\lambda_U$  and  $\lambda_L$ , are defined as:

269 (13) 
$$\lambda_U = 2 - \lim_{t \to 1} \frac{1 - C(t,t)}{1 - t}$$
 and

270 (14) 
$$\lambda_L = \lim_{t \to 0} \frac{C(t,t)}{t}$$

where  $\lambda_U$  and  $\lambda_L$  lie in the range of [0,1] and t denotes the 100t-th percentile of F and G,

272 respectively. Therefore, TDCs depend solely on the copula of the (continuous) random variables and

"theoretical" TDCs can be computed directly from copula parameters (for an overview see Joe (1997)
and Nelsen (2006)). The values obtained for TDCs can be interpreted in such a way that larger values

275 indicate a higher probability of joint extremes in the respective upper or lower tail than lower values

276 of the respective TDC. A TDC of zero implies asymptotic (tail) independence. In the context of this

article, a larger "upper TDC" implied a higher probability of jointly capturing major application events

with the two indicators compared, while a smaller "upper TDC" signified that indicator choice had a

279 large influence on the analysis of pesticide applications, as major application events were not

280 captured jointly. Differences in upper and lower tail dependence further indicated an asymmetric

281 dependence structure.

#### 282 **3.4 Copula selection and estimation procedure**

The following estimation procedure and robustness checks were applied in this article for the
 estimation of copulas and TDCs<sup>1</sup>:

285 (i) *Test independence of indicator pairs analytically with independence tests*. Before the dependence

- 286 structure of indicator pairs was analyzed, independence of all indicator pairs was tested analytically
- with a bivariate, asymptotic independence test based on Kendall's tau (Genest and Favre 2007).
- 288 Independence was rejected for all indicator pairs, which led to the next step, the estimation of
- 289 copulas and TDCs as indicated above.

<sup>&</sup>lt;sup>1</sup> The described estimation and testing procedure was implemented using the R statistical software (R Core Team 2013); especially the VineCopula (Schepsmeier et al. 2017), CopBasic (Asquith 2017) and gofCopula (Okhrin et al. 2016) packages. All R-codes used for the implementation are documented in the online Appendix to provide full transparency of the procedure.

- ii) Estimate copula parameters and TDCs for a wide range of copula families with a semi-parametric
   estimation procedure with empirical margins. Given the sample size, Frahm et al. (2005) find semi parametric estimation procedures perform best for the estimation of TDCs. The semi-parametric
   estimation consisted of two steps: firstly, parametric copula families with empirical margins were
   fitted to the data with a maximum likelihood estimation (log 1007). Bivariate Caussian, t
- fitted to the data with a maximum likelihood estimation (Joe 1997). Bivariate Gaussian-, t-,
- 295 Archimedean- and two parameter Archimedean copulas were considered, therefore covering a wide
- range of possible dependence structures between indicator pairs and all of the most important and
- 297 most used copula classes (compare Joe (1997) and Nelsen (2006)). More precisely, 39 different
- parametric copula types were fitted to the data with the VineCopula package in R. See Schnepsmeier
- et al. (2017) for a full list of these copula types. Secondly, theoretical TDCs were computed on the
- 300 basis of the copula parameters estimated in step one.

301 iii) Choose the three copula families which fit best overall indicator pairs. It was then analytically and 302 graphically evaluated which of the copula families computed in step ii) provided the best fit for the 303 data at hand. Firstly, all the copula families tested were ranked for each indicator pair according to 304 Akaike Information criteria (AIC), Bayesian Information Criteria (BIC) and maximum likelihood values. 305 Copula fit was likewise checked graphically with density plots in which the observed data were 306 plotted against 5000 simulated observations (based on the copula parameters estimated in step 307 ii)).The three copula families with the best fit for each crop were selected to provide a broad and 308 robust overview.

- iv) *Identify copula family with best goodness-of-fit*. The copula family with best goodness-of-fit
- 310 compared to all other copula families considered in step ii) was identified out of the three copula
- families selected in step iii). Goodness-of-fit was tested with Vuong and Clarke tests (Vuong 1989,
- Clarke 2007), and the copula family with the highest score among all copula families was chosen. In
- addition, Akaike and Schwarz corrections (Vuong 1989, Clarke 2007) were applied for the number of
- estimated parameters and then compared to the original results as a robustness check.
- v) Check robustness of obtained copula results with a parametric bootstrap test and the fully non-
- 316 *parametric stable tail dependence function approach*. Firstly, to support the copula selection and the
- 317 findings on tail dependence obtained with the semi-parametric approach described in steps ii) iv), a
- 318 parametric bootstrap test for copula goodness-of-fit (Genest et al. 2006) was performed, based on
- 319 Kendall's process (using Cramer-von-Mises test statistics (Anderson, 1962)). Finally, a second, fully
- 320 non-parametric robustness check was carried out using the stable tail dependence function (Kiriliouk
- 321 et al. 2016), which is an approach for inference on tail dependence.

#### 322 **3.5 Data description and validity checks**

323 The dataset used here was provided by the Swiss Central Evaluation of Agri-Environmental Indicators 324 (CE-AEI). The data is obtained from Agroscope, the Swiss Federal center of excellence for agricultural 325 research, that supervises the collection of the here used data. Farmers participate voluntarily in the 326 data collection scheme and receive compensations for participation (for detailed information on 327 payment schemes and confidentiality agreements see www.agrarmonitoring.ch). Data collection is 328 handled by farmers with the AGRO-TECH software, which is also used to document compliance with 329 regulations and direct payment schemes, and therefore integrated in daily working routines. Primary 330 data were collected and anonymized by intermediaries to guarantee anonymity.

331 The dataset used for the analysis is unique in its level of detailed information on farmers' 332 management practices. It consisted of an unbalanced panel (i.e. not all entities of farmers appear in 333 the sample in every year) of non-organically producing Swiss farmers reporting on their input use 334 (e.g. fertilizer, pesticides, machinery) from 2009 to 2013. While the absolute number of observations 335 on a farm-level was comparatively small (in total 462 farm-level observations), the dataset consisted 336 of around 7500 and 6000 field-level management measures reported with daily resolution for winter 337 wheat and potatoes, respectively. Detailed information on the properties of the pesticides used was 338 available on a product level for each pesticide application. This allowed the robustness of results 339 from the analysis to be tested over different years, during single years, over different aggregation 340 levels and for the most important subgroups of pesticides (herbicides and fungicides). Detailed tables 341 indicating the number of farmers and fields per crop and year and the types of pesticides used per 342 field are documented in the Appendix (Tables A1 and A2). The analysis considered a wide range of 343 application regimes, including 118 different pesticides for potatoes and 131 different pesticides for 344 winter wheat.

- 345 As data were self-reported by farmers, double entries and typos might appear in the data set. These
- 346 entries had to be identified meticulously and removed before the dataset could be used. Therefore, a
- 347 transparent procedure was set up to remove such observations. The procedure is summarized in the
- 348 Appendix, Section B1. To calculate pesticide indicators, CE-AEI data were merged on a product level
- 349 with information on fate, toxicity and formulation of products from the Pesticide Properties
- 350 Database (Lewis et al. 2016), which collects publically available information on pesticides (e.g. from
- 351 pesticide admission). Information on recommended standard application dosages was obtained from
- 352 the Swiss pesticide register (e.g. BLW 2013). A detailed description of the databases used and data
- 353 sources, as well as their linkages, can be found in Figure A1 in the Appendix.
- 354 Sample farms differed with regard to farm types, climate conditions, topographic and geographic
- 355 regions: thirty-six percent of sample observations consisted of purely crop producing farms, whereas
- 356 sixty-four percent of the farms kept at least one animal unit per ha. Seventy-five percent of
- 357 observations were located in valleys or hilly regions and twenty-five percent in mountain regions.
- 358 The altitude of farm locations extended from 360m to 950m above sea level. Farm locations were
- distributed over the most important crop production regions of Switzerland (see Figure 2).



#### 361 Figure 2. Location of sample farms in Switzerland

Source: Own depiction. Shapefiles from Bundesamt für Statistik (BFS), GEOSTAT (2017). Points show farm
 locations per municipality.

364 Spycher et al. (2013) find that average pesticide use in the here used database is well in line with 365 national pesticide use statistics for Switzerland, which also addresses general concerns regarding 366 self-reporting and self-selection of farmers. They state that the dataset covers field crops, i.e. winter 367 wheat and potatoes and the most common pesticide types, i.e. herbicides and fungicides quite well. 368 However, Spycher et al. (2013) find a slightly lower average pesticide use than in national sales statistics, as commonly reported for similar datasets (e.g. Krujine et al. 2012). Furthermore, a good fit 369 370 of average sample statistics and population averages was found (see Appendix, Section B2). Any 371 concerns regarding biased self-reporting of extreme pesticide use were also addressed by analyzing 372 the compliance of Swiss farmers with pesticide regulations, i.e. pesticide overdosing (see Appendix, 373 Section B2). Results suggest that farmers have reported actual management decisions (i.e. extreme 374 applications), and were not concerned about anonymity issues.

- 375 Possible concerns about the unbalanced structure of the sample and pooling of years were
- addressed by testing for significant differences in skewness of pesticide use distribution across single
- 377 years (Tables A3 and A4 in the Appendix). In general, none of the five years was found to differ
- 378 significantly in skewness from all other years (at the 5% significance level), except for LI for winter
- 379 wheat in 2010. As 2010 was the year with the highest number of participants with winter wheat, this
- 380 could indicate that skewness of the LI indicator in winter wheat was biased downwards by pooling.

The opposite was observed for the QA indicator in winter wheat, where the skewness was found to be significantly smaller in 2011 and 2012, compared to 2009 and 2010.

#### 383 **4. Results**

#### 384 **4.1** Pesticide use trends shift by indicator

385 Yearly averages of QA, TFI and LI indicators for winter wheat and potato production in Switzerland

386 (Figure A2, Appendix) showed that pesticide use trends reversed when using different indicators.

387 More specifically, the TFI and LI indicators pointed towards a positive trend in pesticide use, while

- the QA indicator pointed towards a negative trend for the period (2009-2013) under consideration.
- 389 Data for all years were then pooled and focus was placed on observations at farm and crop levels, as
- the crucial pest management decisions are generally taken at the farm-level (Waterfield and
- Zilberman 2012). Histograms of pesticide use by indicator and crop are shown in Figures A3 and A4 in
- the Appendix. Results showed that QA and LI indicators had a significantly larger positive skewness
- than the TFI indicator (Tables A5 and A6, Appendix), suggesting that QA and LI were more likely to
- indicate extreme applications than TFI. Median, 90% and 99% quantiles of indicators were likewise
- 395 compared, leading to the same conclusions (Tables A7 and A8, Appendix).

#### 396 **4.2** Analysis of the dependence structure

- 397 Significant, positive correlations were found between the QA and LI (in potatoes n=192, Kendall's
- 398 tau= 0.27 *P*< 0.01; in winter wheat n=270, Kendall's tau= 0.48, *P*<0.01) and between TFI and LI (in
- potatoes n=192, Kendall's tau= 0.41, P< 0.01 and in winter wheat n=270, Kendall's tau= 0.59, P<
- 400 0.01). Correlation for TFI and LI indicators was higher than for QA and LI indicators, and higher for
- 401 winter wheat than for potatoes (Table 1). Copulas and tail dependence coefficients (TDCs) were then
- 402 used to reveal possible asymmetric dependence structures and provide information on the degree of
- 403 dependence in the tails of the distributions, i.e. between extreme observations. Firstly, the best
- fitting copulas for the indicator pairs (see Methods, Section 3.4) were identified. Then copula density
- was plotted. Figure 3 visualizes how the dependence structure shown in the scatterplot (Fig. 3a) is
   captured by the estimated copula function (Fig. 3b). The figure focuses on the comparison of QA and
- 406 captured by the estimated copula function (Fig. 3b). The figure focuses on the comparison of QA and407 LI indicators for farm-level pesticide use in potatoes and shows a relatively high dependence
- 408 between the lower tails of QA and LI and almost no dependence in the upper tails. Respective figures
- for the LI and TFI indicators, as well as farm-level pesticide use in winter wheat can be found in the
- 410 Appendix (Figures A5, A6 and A7).



## Fig. 3. Copulas reveal differences in the dependence structure of indicators between upper and lower tails

411

414 The scatterplot in Figure 3a) shows the relationship between QA and LI values for farm-level pesticide use in 415 potatoes. Pesticide use in terms of the QA and LI indicators is plotted on the X- and Y- axis, respectively. 416 Extremely high values of the QA and LI indicator, which do not correspond with extremely high values of the 417 other indicator, are thus located in the bottom right and top left of the plot respectively. Figure 3b) shows a 418 contour plot of the fitted copula model for the same observations and indicators (Survival Gumbel copula, see 419 Methods, Section 2.4 for selection and estimation procedure). X- and Y- axis in the contour plot show 420 distribution functions for QA and LI values, respectively (see Methods, Section 3.3). The Z-Axis ('Copula 421 Density') shows the probability density of the fitted Survival Gumbel copula. The copula captures the 422 dependence structure between the distributions of the two indicators, not only on average, but over the whole 423 distributions. The plot shows a high copula density for joint low values of QA and LI indicator (bottom left 424 corner) and a low density of the copula for jointly high values of the QA and LI indicator (top right corner).

425 Upper and lower TDCs of the estimated copula models were then computed to quantify the degree 426 of dependence in the tails of the distributions. TDCs are a measure for dependence in the upper and 427 lower tails of distributions and lie in the range from 0 to 1, where larger values indicate a higher 428 probability of joint extremes. Table 1 shows that all copulas reveal a clearly asymmetric dependence 429 structure between all indicator pairs. While they show a moderate to strong dependence in the 430 lower tails, they exhibit very weak to no dependence in the upper tails. Thus, the analysis suggested 431 that high quantities do not necessarily correspond to high risks. Results were found to be robust over 432 the three best fitting copula types (see Methods, Section 3.4 for selection procedure) and copula 433 goodness-of-fit was confirmed (see Appendix, Section B3 on robustness checks for copulas). TDCs were larger between TFI and LI than between QA and LI, in line with results from the correlation 434 435 coefficients. Similarly, a stronger dependence was observed for winter wheat than for potatoes. In 436 addition, the robustness of the results was checked further by performing the same analysis again 437 separately for herbicides and fungicides. These are the most relevant pesticide types in the sample 438 and in European agriculture in general (e.g. Eurostat, 2017). Results were robust for both indicator 439 pairs over all analyzed crops, pesticide types and copula families, leading to a total of 36 separately 440 estimated copulas and TDCs (see Appendix Table A9).

## Table 1. Significant correlation of quantitative indicators and farm-level pesticide risks – but very low explanatory power for extreme risks

#### Potatoes

Comparing LI	Kendall's τ	Survival Gumbel		Survival E	BB1	Survival BB7	
Indicator to:	correlation	Copula		Copula		Copula	
		$L_{TDC}$	Utdc	L <sub>TDC</sub>	UTDC	L <sub>TDC</sub>	Utdc
QA Indicator	0.27	0.32	0.00	0.32	0.00	0.36	0.02
TFI Indicator	0.41	0.49	0.00	0.45	0.06	0.50	0.28

#### Winter Wheat

Comparing LI	Kendall's τ	Survival Gumbel		Survival BB1		BB7	
Indicator to:	correlation	Copula		Copula		Copula	
		L <sub>TDC</sub>	U <sub>TDC</sub>	L <sub>TDC</sub>	$U_{\text{TDC}}$	L <sub>TDC</sub>	$U_{\text{TDC}}$
QA Indicator	0.48	0.56	0.00	0.56	0.00	0.62	0.16
TFI Indicator	0.59	0.63	0.00	0.61	0.01	0.65	0.36

#### 443

444 Level of dependence between farm-level indicators of pesticide use in potato (n=190) and winter wheat 445 (n=270) production. QA, TFI and LI denote the Quantity of Active Ingredient, Treatment Frequency Index and 446 Load Index Indicators, respectively. LTDC and UTDC are in the range of [0,1] and indicate lower and upper tail 447 dependence coefficients, i.e. the dependence for extremely low and extremely high observations between two 448 indicator distributions, respectively. The three copula families analyzed were chosen per crop for the farm-level 449 observations, according to the six-step selection and testing procedure described in Section 3.4. Survival 450 Gumbel, BB1 and BB7 copulas refer to the 180 degrees rotated Gumbel, Clayton-Gumbel and Joe-Clayton 451 copulas, which are one parametric and two parametric Archimedean copulas, respectively. The LTDC of all three 452 copula families lies in a positive range. The  $U_{TDC}$  of the Survival Gumbel copula is zero by definition and in a 453 positive range for Survival BB1 and BB7 copulas. See Nelsen (1997) for exact definitions of the selected copula 454 families. Bold numbers indicate the best fitting copula type per indicator pair, according to Vuong and Clarke 455 goodness-of-fit tests (Vuong 1989, Clarke 2007). Kendall correlation coefficients measured the (average) 456 ordinal dependence between two indicators and all were significant at the 1% level. Detailed robustness and 457 goodness-of-fit tests for copulas and TDCs can be found in Appendix, Section B3.

458 Results continued to remain robust when pesticide use was aggregated on a field level instead of a 459 farm level (Appendix, Table A10).

#### 460 4.3 Temporal hotspots shift by indicator

So far, extreme pesticide use was analyzed as "one-year application regimes", located in the upper 461 462 tail of observed pesticide use distributions, without considering the spatial and time dimension of 463 applications. It was then analyzed if results remained consistent when defining extreme pesticide use 464 as temporal "hotspots", i.e. single, daily applications with high risks for the environment and human 465 health compared to the other applications. Analysis with copulas and TDCs was again applied on a 466 level of non-aggregated, single pesticide applications. In line with field- and farm-level results, these 467 results showed that the purely quantitative QA and TFI indicators cannot capture hotspots in terms of environmental and health risks, as indicated by the LI indicator. Although goodness-of-fit of the 468 469 copula models is weaker at the application level than at the farm-level, results remained qualitatively 470 consistent for both crops, all copula families and different pesticide types tested (Appendix, Table 471 A11). Single applications, pooled for all farmers and years, were plotted on a daily scale (Figure 4) to 472 illustrate differences between QA, TFI and LI. Temporal hotspots as indicated by QA, TFI and LI,

- 473 respectively (the highest 1% of all applications), were then highlighted. The graph for potatoes
- 474 illustrates how indicator choice may shift policy targets. While hotspots for the QA occurred between
- 475 May and July, the TFI identifies June to August as the relevant timeframe. Most importantly, the risk-
- adjusted LI identified another timeframe (July-August) and other applications as relevant hotspots.
- 477 This was similarly observed for winter wheat (see Appendix, Figure A8).



#### 480 Fig. 4. Temporal hotspots of pesticide use in potatoes shift with indicator choice

481 Bars represent single pesticide applications in potatoes, plotted on a daily scale. Plotted observations include

482 application data for the years 2009-2013 and all sample farmers. Pesticide use is expressed in Quantity of

483 Active ingredient (QA), Treatment Frequency Index (TFI) and Load Index (LI) indicators, respectively. "Temporal

484 Hotspots" indicate the Top 1% of all applications per indicator, respectively.

#### 485 5. Discussion

Throughout the world, current pesticide policies have been implemented using diverse indicators for
 the quantification of pesticide use. Inconsistencies between different pesticide indicators were

488 analyzed in this paper to aid the identification of targets for pesticide policies, i.e. use patterns and 489 profiles associated with the greatest risks for the environment and human health. Results show that 490 the most important, commonly utilized, quantitative pesticide indicators fail to identify the use 491 patterns with the greatest risks for the environment and human health from pesticide use, as 492 indicated by the LI indicator. Therefore, reliance on these indicators for the implementation of high-493 level pesticide reduction targets and market-based pesticide policies designed to incentivize a 494 reduction in environmental and health impacts from pesticide use, can lead to a misspecification of 495 policy targets and, consequently, to biased policy incentives and outcomes. The goal of such policies 496 is a reduction of risks from pesticide use relative to current levels. It is therefore important to identify 497 the use patterns associated with the greatest risks, even if they are deemed acceptable in terms of 498 potential risks by regulatory authorities (e.g. the United States Environmental Protection Agency and 499 the European Food Safety Authority).

500 More specifically, the analysis showed that the quantitative indicators QA and TFI have a good 501 average fit to proxy potential environmental and human health risks of pesticides indicated by the LI 502 indicator but are unsuitable for identifying the use patterns associated with the greatest risks, i.e. 503 extreme risks. Differences between indicators may occur due to variations in standard dosages (i.e. 504 between QA and both TFI and LI), inherent pesticide properties (i.e. between TFI and LI) or 505 combinations of both (i.e. between QA or TFI and LI). Positive, significant correlations between both 506 QA and TFI with LI were found. This is in line with Kudsk et al. (2018), who find a good average fit of 507 TFI and LI indicators. The analysis of the dependence structure was then extended with copulas and 508 tail dependence coefficients. Very small coefficients of upper tail dependence were found between 509 both the QA and TFI with the LI. The finding of small "upper TDCs" suggested that indicator choice 510 has a great influence on the analysis of pesticide applications, as extremely high values are not 511 captured jointly. These findings show that the two quantitative indicators (QA, TFI) have almost no 512 explanatory power for the pesticide use observations with the greatest risks for the environment and

513 health as indicated by the LI indicator.

514 The findings relating to the very low explanatory power of QA and TFI for extreme pesticide risks 515 remained robust over different crops (winter wheat and potatoes), when separately testing the main 516 pesticides types in potato and winter wheat (herbicides and fungicides) and for all tested aggregation 517 levels of pesticide use (single applications, fields, farms). The existence of higher correlation 518 coefficients and TDCs for herbicides compared to fungicides as well as for potatoes compared to 519 winter wheat highlights that the degree of dependence is determined by registered (available) 520 products and farmers' application patterns in the analyzed cropping system. Differences in results 521 between pesticide groups may be an indication of different levels of heterogeneity with regard to 522 pesticide properties (e.g. toxicity) in the respective group of products. Differences between crops 523 (winter wheat < potatoes) may be an indication that the more input intensive a crop is, the greater 524 the inconsistencies between indicators, as one would expect. The comparison of different 525 aggregation levels is especially important, as the definition of "extreme pesticide use" (i.e. the 526 greatest observations, located in the upper tail of the distribution) is central for the analysis in this article. Two policy-relevant definitions of "extreme pesticide use" were covered in the analysis. 527 528 Extremes might be relevant in a context of a high risk of aggregated, farm-level pesticide use over the 529 cropping year ("extreme application regimes"), but also in a context of high temporal and spatial 530 concentrations of pesticide risks from single applications ("extreme application events"). Peaks, or 531 "temporal hotspots", of pesticide use are especially relevant for their effects on ecosystems, like

bodies of water. Results revealed that in both cases, quantitative indicators are unsuitable for

- 533 pinpointing the pesticide applications and application regimes associated with the greatest risks (as
- shown by the LI indicator), highlighting the broad relevance of the findings.

535 In addition, results show that trends in national pesticide use may shift with the choice of pesticide 536 indicators, as different indicators capture different properties of the pesticides used. To this end, 537 trends of average, yearly pesticide use were computed and compared for the QA, TFI and LI 538 indicators. The findings confirmed results of Kniss (2017) for the US, who has observed a reversal in 539 pesticide use with different indicators. National pesticide use trends serve to evaluate the efficiency 540 of public policies to regulate pesticide utilization. It is therefore important, from a regulatory 541 perspective, that the representation of pesticide trends reflects potential pesticide impacts in a 542 realistic manner. The scope of our study was not to infer on long-term pesticide use trends in 543 Switzerland, but to identify the differences between indicators. Inference on long-term pesticide use 544 trends demands a longer time series, which would be more robust to yearly shocks, e.g. caused by

545 adverse weather events.

The above findings are transferable to other countries and regions with a similar or even greater
 heterogeneity of pesticide products in use. The pesticides used by farmers in the sample are similar

to those used by producers in the EU. 97.5% and 95% of the 40 most used herbicides and fungicides
in the non-EU country Switzerland were also approved in Germany (BLV 2017) and France (Anses
2017), and pesticide use patterns are similar between these countries (CH: Baan et al. 2015, GER: JKI
2017, FR: Agreste 2017). Structural differences in the computation of the three indicators were
compared to facilitate the transfer of results to other countries in a more general way. It was found
that in comparison to the QA indicator, the lower skewness and shorter right tail of the TFI indicator

- 554 (Tables A5 and A6 and Figures A3 and A4 in the Appendix) are based on the normalization of applied
- dosages used to compute it. The greater skewness and longer tail of the LI indicator when compared
  to the TFI indicator (see Methods, Section on "Indicator calculation"; Tables A5 and A6 and Figures

557 A3 and A4 in the Appendix) are based on the heterogeneity of qualitative pesticide properties (i.e.

- toxicity) that are also considered. It can therefore be argued that the identified inconsistencies
- 559 between indicators translate to countries with similar application patterns and registered products.
- 560 Along these lines, inconsistencies should be even more pronounced for countries with a potentially
- 561 greater range of qualitative properties (e.g. toxicity) across the pesticides used.

#### 562 6. Conclusions

563 In conclusion, results suggest that when implementing pesticide policies, policymakers should 564 consider choice and design of underlying pesticide indicators with great care, as they have significant 565 impacts on the incentives of policies. This is highly relevant in the context of the present 566 implementation and revision of NAPs in Europe, which aim to reduce "risks and impacts of pesticide 567 use on human health and the environment" (Directive 2009/128/EC). More specifically, given the 568 results, market-based policies with underlying quantitative indicators currently in force might create 569 mis-incentives for farmers. Policy measures, such as pesticide taxes (Finger et al. 2017) based on 570 quantitative indicators, could even encourage the use of more risky products if these were to 571 become cheaper than products with low inherent risks, but high standard dosages or a large number 572 of applications. Similarly, the adoption of high-level pesticide reduction targets using quantitative 573 indicators might lead to the implementation of policies which incentivize the use of more risky 574 pesticides. Countries which plan to introduce new or revise existing market-based pesticide policies

- or reduction targets should effectively target the "risk components" of pesticides (e.g. risks for the
- 576 environment and human health, etc.), to prevent possible mis-incentives. Such "risk components"
- 577 might vary in their importance depending on the country and should be adapted to country specific
- 578 circumstances. The results of the study are especially relevant in the light of recent evidence on the 579 effects of extreme pesticide use on the health of residents in agricultural areas (Larsen et al. 2017)
- and on the environment (Releya and Hovermann 2006; Gordon et al. 2012, Bundschuh et al. 2013
- 581 and Topping and Elmeros 2016). Larsen et al. (2017) conclude that there is a need for the
- 582 implementation of pesticide policies that tackle extreme applications. Focusing on incentives of
- 583 policies, this study shows that the choice of a non-target-specific indicator can render such policies
- 584 inefficient or even lead to detrimental effects. Given the large number of policies currently
- 585 introduced to reduce adverse effects from pesticide use and achieve greater sustainability in
- agriculture, the results potentially have large-scale implications for farmers' pesticide use worldwide.
- 587 The results of this study also have implications for the analysis of the economic productivity of
- 588 pesticides (e.g. as in Hossard et al. 2014 and Lechenet et al. 2017). More specifically, they emphasize
- the need to investigate the lack of precision in the assessment of negative external effects with
- 590 purely quantitative indicators. Further research is needed in this field to compare the commonly
- 591 used quantitative indicators to actual indicators of pesticide efficacy.
- 592 The study can be extended in various directions. Firstly, other crops should be considered. This is 593 particularly important for special crops, vegetables and fruits that are characterized by very high 594 absolute levels of pesticide use. Secondly, an analysis of determinants of extreme pesticide use is 595 necessary to design efficient and effective policies for the reduction of this practice. Thirdly, other 596 indicators might be included to quantify environmental and health risks (e.g. Labite et al. 2011; OECD 597 2016; Kudsk et al. 2018). These indicators could explicitly include information on the application context, thus improving the quality of risk assessment (Pierlot et al. 2017). In addition, indicators 598 599 based on more complex, site-specific risk assessment models might become more widely available in 600 the future, given the increasing collection of timely, detailed and spatially explicit data in agriculture. 601 Finally, the creation of a harmonized risk indicator for pesticides in Europe could further improve the 602 coherence of policies in the NAP and help to monitor pesticide risks on a detailed level across time 603 and space in Europe.
- 604

#### Acknowledgements

We thank Agroscope for the provision of the used dataset. We further thank the five anonymous reviewers for their constructive feedback on earlier versions of this article. 

#### 611 7. References

- 612 Agreste. Pratiques Culturales sur les grandes cultures et prairies.
  613 http://agreste.agriculture.gouv.fr/enquetes/pratiques-culturales/grandes-cultures-prairies.
  614 Accessed November 2017 (2017).
- Anderson, T.W. On the distribution of the two-sample Cramer-von Mises criterion. *The Annals of Mathematical Statistics*, pp.1148-1159 (1962).
- Anses. Le catalogue des produits phytopharmaceutiques et de leurs usages, des matières fertilisantes
  et des supports de culture autorisés en France. https://ephy.anses.fr. Accessed December 2017
  (2017).
- 620 Asquith, W. Package 'copBasic'. R package version 2.0.5 (2017).
- Baan, de L., Spycher, S., & Daniel, O. Einsatz von Pflanzenschutzmitteln in der Schweiz von 2009 bis
  2012. Agrarforschung Schweiz 6, 48-55 (2015).
- Baumol, W.J., Oates, W.E., Bawa, V.S., Bawa, W.S. and Bradford, D.F. *The theory of environmental policy*. Cambridge university press (1988).
- Beketov, M.A., Kefford, B.J., Schäfer, R.B. & Liess, M. Pesticides reduce regional biodiversity of stream
  invertebrates. *Proceedings of the National Academy of Sciences*, *110*(27), pp.11039-11043 (2013).
- Böcker, T. & Finger, R. European Pesticide Tax Schemes in Comparison: An Analysis of Experiences and
  Developments. *Sustainability 8*(4), 378 (2016).
- Böcker, T., Britz, W., Finger, R. (2018). Modelling the effects of a glyphosate ban on weed management
  in silage maize production. *Ecological Economics* 145: 182–193.
- Bundesrat der Schweizerischen Eidgenossenschaft. Aktionsplan zur Risikoreduktion und nachhaltigen
   Anwendung von Pflanzenschutzmitteln. Bern, Schweiz (2017).
- Bundesamt für Statistik (BFS), GEOSTAT 2017. BFS GEOSTAT Gemeindegrenzen.
   https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/geostat/geodaten-
- bundesstatistik/administrative-grenzen/generalisierte-gemeindegrenzen.html. Accessed March2017 (2017).
- 637 Bundesamt für Statistik (BFS). BFS Landwirtschaftliche Strukturerhebung 2016.
  638 https://www.bfs.admin.ch/bfs/de/home/statistiken/land-
- 639 forstwirtschaft/landwirtschaft/strukturen.html. Accessed October 2017 (2017).
- Bundesamt für Lebensmittel und Verbraucherschutz (BLV). Zugelassene Pflanzenschutzmittel.
   https://www.bvl.bund.de/DE/04\_Pflanzenschutzmittel/psm\_node.html. Accessed November
   2017 (2017).
- Bundschuh, M., Zubrod, J.P., Klemm, P., Elsaesser, D., Stang, C. & Schulz, R. Effects of peak exposure
  scenarios on *Gammarus fossarum* using field relevant pesticide mixtures. *Ecotoxicology and environmental safety*, *95*, pp.137-143 (2013).
- 646BundesamtfürLandwirtschaft(BLW).BLWPflanzenschutzmittelverzeichnis.647http://www.psm.admin.ch/psm/produkte.Accessed December 2013 (2013).
- 648 Clarke, K. A Simple Distribution-Free Test for Nonnested Model Selection. *Political Analysis*, 15, 347 649 363 (2007).

- Claeys, S., Vagenende, B., De Smet, B., Lelieur, L. and Steurbaut, W. The POCER indicator: a decision
   tool for non-agricultural pesticide use. *Pest management science*, *61*(8), pp.779-786 (2005).
- Coll, M., & Wajnberg, E (eds.). Environmental Pest Management: Challenges for Agronomists,
   Ecologists, Economists and Policymakers. Hoboken, NJ, USA: Wiley (2017).
- 654 Croux, C., & Dehon, C. Influence functions of the Spearman and Kendall correlation measures.
  655 Statistical methods & applications, 19(4), 497-515 (2010).
- Damalas, C. A. Understanding benefits and risks of pesticide use. *Scientific Research and Essays 4*(10),
  945-949 (2009).
- Davison, A. C., & Hinkley, D. V. Bootstrap methods and their application (Vol. 1). Cambridge University
   press (1997).
- European Commission (EC) (ed.).*Annex II: Refit Initiatives to the Commission Work Programme 2016*.
  European Commission, Strasbourg, France (2015).
- European Commission Scientific Advice Mechanism (EC SAM) (ed.). *EU authorization processes of Plant Protection Products from a scientific point of view*. Group of Chief Scientific Advisors, Brussels,
   Belgium (2018).
- 665 European Union (EU). EU Pesticides database. http://ec.europa.eu/food/plant/pesticides/eu-666 pesticides-database/public. Accessed December 2017 (2017).
- European Union, DG Health and Food Safety (ed.). "Overview Report Sustainable Use of Pesticides".
  Luxembourg: Publications Office of the European Union (2017).
- 669Eurostat.Agri-environmentalindicator-consumptionofpesticides.670http://ec.europa.eu/eurostat/statistics-explained/index.php/Agri-environmental\_indicator\_-
- 671 \_\_\_\_\_consumption\_of\_pesticides#Data\_used\_and\_methodology. Accessed April 2017 (2017).
- Finger, R., Möhring, N., Dalhaus, T. & Böcker, T. Revisiting pesticide taxation schemes. Ecological
  Economics, 134, pp.263-266 (2017).
- 674 Finger, R. Take a holistic view when making pesticide policies stricter. Nature 556, 174 (2018).
- Frahm, G., Junker, M. & Schmidt, R. Estimating the tail-dependence coefficient: properties and pitfalls.
  Insurance: mathematics and Economics, 37(1), pp.80-100 (2005).
- Gaba, S., Gabriel, E., Chadœuf, J., Bonneu, F. & Bretagnolle, V. Herbicides do not ensure for higher
  wheat yield, but eliminate rare plant species. Scientific Reports, 6 (2016).
- Genest, C., Quessy, J. F., & Rémillard, B. Goodness-of-fit procedures for copula models based on the
   probability integral transformation. Scandinavian Journal of Statistics, 33(2), 337-366 (2006).
- Genest, C., & Favre, A. C. Everything you always wanted to know about copula modeling but were
  afraid to ask. Journal of hydrologic engineering, 12(4), 347-368 (2007).
- Genest, C., & Rémillard, B. Validity of the parametric bootstrap for goodness-of-fit testing in
  semiparametric models. In Annales de l'Institut Henri Poincaré, Probabilités et Statistiques (Vol.
  44, No. 6, pp. 1096-1127). Institut Henri Poincaré (2008).
- Ghimire, N., & Woodward, R. T. Under-and over-use of pesticides: An international analysis. Ecological
   Economics, 89, 73-81 (2013).

- Gordon, A.K., Mantel, S.K. & Muller, N.W. Review of toxicological effects caused by episodic stressor
   exposure. Environmental toxicology and chemistry, 31(5), pp.1169-1174 (2012).
- Gutsche, V. and Rossberg, D. SYNOPS 1.1: a model to assess and to compare the environmental risk
  potential of active ingredients in plant protection products. *Agriculture, ecosystems & environment, 64*(2), pp.181-188 (1997).
- Hoop, D. & Schmid, D. Grundlagenbericht 2013: Zentrale Auswertung von Buchhaltungsdaten.
  Ettenhausen: Agroscope INH (2014).
- Hossard, L., Philibert, A., Bertrand, M., Colnenne-David, C., Debaeke, P., Munier-Jolain, N., Jeuffroy,
  M.H., Richard, G. & Makowski, D. Effects of halving pesticide use on wheat production. *Scientific reports*, 4 (2017).
- Hossard, L., Guichard, L., Pelosi, C. and Makowski, D., 2017. Lack of evidence for a decrease in synthetic
  pesticide use on the main arable crops in France. *Science of the Total Environment*, 575, pp.152161 (2014).
- Houbraken, M., Bauweraerts, I., Fevery, D., Van Labeke, M.C. and Spanoghe, P. Pesticide knowledge
   and practice among horticultural workers in the Lâm Đồng region, Vietnam: A case study of
   chrysanthemum and strawberries. *Science of the Total Environment*, *550*, pp.1001-1009 (2016).
- Julius Kühn-Institut (JKI). Statistische Erhebungen zur Anwendung von Pflanzenschutzmitteln in der
   Praxis. Ergebnisse. http://papa.jki.bund.de. Accessed November 2017 (2017).
- Joe, H. Multivariate models and multivariate dependence concepts. CRC Press (1997).
- Kiriliouk, A., Segers, J. & Warchoł, M. Nonparameteric estimation of extremal dependence: in Extreme
   Value Modeling and Risk Analysis, D.K. Dey and Jun Yan (eds.), Boca Raton, FL, CRC Press (2016).
- Kniss, A. R. Long-term trends in the intensity and relative toxicity of herbicide use. Naturecommunications, 8 (2017).
- Kruijne R., van der Linden A.M.A., Deneer J.W., Groenwold J. G. & Wipfler E.L. *Dutch Environmental Risk Indicator for Plant Protection Products*. Wageningen, Netherlands, Alterra-Report 2250.1.
  (2012).
- Kudsk, P., Jørgensen, L.N. & Ørum, J.E. Pesticide Load—A new Danish pesticide risk indicator with
   multiple applications. Land Use Policy, 70, pp.384-393 (2018).
- Labite, H., Butler, F. and Cummins, E. A review and evaluation of plant protection product ranking tools
   used in agriculture. Human and Ecological Risk Assessment, 17(2), pp.300-327 (2011).
- Larsen, A. E., Gaines, S. D., & Deschênes, O. Agricultural pesticide use and adverse birth outcomes in
   the San Joaquin Valley of California. Nature Communications, 8(1), 302 (2017).
- Lechenet, M., Dessaint, F., Py, G., Makowski, D. & Munier-Jolain, N. Reducing pesticide use while
   preserving crop productivity and profitability on arable farms. Nature Plants, 3, p.17008 (2017).
- Lefebvre, M.; Langrell, S.R.H. & Gomez-y-Paloma, S. Incentives and policies for integrated pest
   management in Europe: A review. Agron. Sustain. Dev., 35, 27–45 (2015).
- Lewis, K. A., Tzilivakis, J., Warner, D. J. & Green, A. An international database for pesticide risk
   assessments and management. Human and Ecological Risk Assessment: An International Journal
   22(4), 1050-1064 (2016).

- Malaj, E., Peter, C., Grote, M., Kühne, R., Mondy, C.P., Usseglio-Polatera, P., Brack, W. & Schäfer, R.B..
   Organic chemicals jeopardize the health of freshwater ecosystems on the continental scale.
   Proceedings of the National Academy of Sciences, 111(26), pp.9549-9554 (2014).
- Miljøministeriet. The Agricultural Pesticide Load in Denmark 2007-2010. Environmental review no. 2,
   Miljøstyrelsen, Kopenhagen (2012).
- Ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt (MAAF). Plan Ecophyto II.
  http://agriculture.gouv.fr/sites/minagri/files/151022\_ecophyto.pdf. Accessed January 2018
  (2015).
- Ministère de l'Agriculture et de l'Alimentation (MAA). Indicateur de Fréquence de Traitements
  phytosanitaires (IFT). http://agriculture.gouv.fr/indicateur-de-frequence-de-traitementsphytosanitaires-ift. Accessed November 2017 (2017).
- 738 Nelsen, R. B. An introduction to copulas, 2nd. New York: SpringerScience Business Media (2006).
- Organisation for Economic Co-operation and Development (OECD). Report on the online pesticide risk
   indicators database (PRIER). Environment Directorate, OECD, Paris (2016).
- Okhrin, O., Odening, M. and Xu, W. Systemic weather risk and crop insurance: the case of China.
   *Journal of Risk and Insurance*, *80*(2), pp.351-372 (2013).
- Okhrin, O., Trimborn, S., Zhang, S. & Zhou, Q. M. gof-Copula: Goodness-of-Fit Tests for Copulae. R
   package version 0.2.3 (2016).
- Osteen, C. D., & Fernandez-Cornejo, J. Economic and policy issues of US agricultural pesticide use
   trends. Pest management science, 69(9), 1001-1025 (2013).
- Payton, M. E., Greenstone, M. H., & Schenker, N. Overlapping confidence intervals or standard error
  intervals: what do they mean in terms of statistical significance? Journal of Insect Science, 3(1),
  34 (2003).
- Perry, E. D., Ciliberto, F., Hennessy, D. A., & Moschini, G. Genetically engineered crops and pesticide
  use in US maize and soybeans. Science advances, 2(8), e1600850 (2016).
- Pierlot, F., Marks-Perreau, J., Real, B., Carluer, N., Constant, T., Lioeddine, A., Van Dijk, P., Villerd, J.,
   Keichinger, O., Cherrier, R. and Bockstaller, C. Predictive quality of 26 pesticide risk indicators and
   one flow model: A multisite assessment for water contamination. *Science of the Total Environment*, 605, pp.655-665 (2017).
- Pimentel, D. and Burgess, M. "Environmental and economic costs of the application of pesticides
   primarily in the United States." In Pimentel, D, Peshin, R., ed. Integrated pest management
   Pesticide Problems Vol. 3. Dordrecht, Netherlands: Springer, pp. 47-71 (2014).
- 759 R Core Team. R: A language and environment for statistical computing (2013).
- Relyea, R. & Hoverman, J. Assessing the ecology in ecotoxicology: a review and synthesis in freshwater
   systems. Ecology Letters, 9(10), pp.1157-1171 (2006).
- Schepsmeier, U., Stoeber, J., Brechmann, E. C., & Graeler, B. VineCopula: statistical inference of vine
   copulas. R package version 2.1.3 (2017).
- Skevas, T., Lansink, A.O. and Stefanou, S.E. Designing the emerging EU pesticide policy: A literature
   review. *NJAS-Wageningen Journal of Life Sciences*, *64*, pp.95-103 (2013).

- Spycher, S., Badertscher, R., & Daniel, O. Indikatoren für den Einsatz von Pflanzenschutzmitteln in der
   Schweiz. Agrarforschung Schweiz 4(4), 192–199 (2013).
- Stehle, S. & Schulz, R. Agricultural insecticides threaten surface waters at the global scale. Proceedings
   of the National Academy of Sciences, 112(18), pp.5750-5755 (2015).
- Storck, V., Karpouzas, D.G. and Martin-Laurent, F. Towards a better pesticide policy for the European
   Union. *Science of the Total Environment*, *575*, pp.1027-1033 (2017).
- Strange, R.N. & Scott, P.R. Plant disease: a threat to global food security. Annu. Rev. Phytopathol., 43,
   pp.83-116 (2005).
- Strassemeyer, J., Daehmlow, D., Dominic, A.R., Lorenz, S. and Golla, B. SYNOPS-WEB, an online tool for
   environmental risk assessment to evaluate pesticide strategies on field level. *Crop protection*, *97*,
   pp.28-44. (2017)
- Sun, B., Zhang, L., Yang, L., Zhang, F., Norse, D., & Zhu, Z. Agricultural non-point source pollution in
   China: causes and mitigation measures. Ambio, 41(4), 370-379 (2012).
- Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., & Polasky, S. Agricultural sustainability and
   intensive production practices. Nature, 418(6898), 671 (2002).
- Topping, C.J. and Elmeros, M. Modeling Exposure of Mammalian Predators to Anticoagulant
   Rodenticides. *Frontiers in Environmental Science*, *4*, p.80 (2016).
- Tsaboula, A., Papadakis, E.N., Vryzas, Z., Kotopoulou, A., Kintzikoglou, K. and Papadopoulou Mourkidou, E. Environmental and human risk hierarchy of pesticides: A prioritization method,
   based on monitoring, hazard assessment and environmental fate. *Environment international*, *91*,
   pp.78-93 (2016).
- 787 United States Department of Agriculture (USDA). USDA/NASS QuickStats Ad-hoc Query Tool.
   788 https://quickstats.nass.usda.gov. Accessed July 2017 (2017).
- Vercruysse, F. and Steurbaut, W. POCER, the pesticide occupational and environmental risk indicator.
   *Crop Protection*, *21*(4), pp.307-315 (2002).
- Vuong, Q. H. Ratio tests for model selection and non-nested hypotheses. Econometrica 57 (2), 307-333
  (1989).
- Waterfield, G. & Zilberman, D. "Pest management in food systems: an economic perspective." Annual
   review of environment and resources 37: 223-245 (2012).
- Xu, W., Filler, G., Odening, M. and Okhrin, O. On the systemic nature of weather risk. *Agricultural Finance Review*, *70*(2), pp.267-284 (2010).
- Yang, Y., Lian, X.Y., Jiang, Y.H., Xi, B.D. and He, X.S. Risk-based prioritization method for the
   classification of groundwater pesticide pollution from agricultural regions. *Integrated environmental assessment and management*, *13*(6), pp.1052-1059 (2017).
- Zhang, K. M., & Wen, Z. G. Review and challenges of policies of environmental protection and
  sustainable development in China. Journal of environmental management, 88(4), 1249-1261
  (2008).
- Zhang, S., Okhrin, O., Zhou, Q.M. and Song, P.X.K. Goodness-of-fit test for specification of
   semiparametric copula dependence models. Journal of Econometrics, 193(1), pp.215-233 (2016).

## 805 Appendices

#### 806 Table A1. Sample size

Year	Winter Wheat	Potatoes
2009	61 (195)	44 (101)
2010	66 (193)	38 (103)
2011	52 (158)	40 (108)
2012	48 (137)	32 (82)
2013	43 (134)	38 (88)
Total	270 (817)	192 (482)

807 Number of farms (fields) per year and crop

#### 809 Table A2. Share of treated fields per pesticide type

	V	Vinter Whe	at	Potatoes			
Year	Herbicide	Fungicide	Insecticide	Herbicide	Fungicide	Insecticide	
2009	0.92	0.76	0.07	0.88	0.97	0.25	
2010	0.94	0.84	0.08	0.93	0.99	0.27	
2011	0.96	0.76	0.25	0.92	0.97	0.31	
2012	0.96	0.85	0.12	0.96	1.00	0.35	
2013	0.97	0.80	0.04	0.93	0.98	0.39	
Total	0.95	0.80	0.11	0.92	0.98	0.31	

810 Share of fields treated with at least one pesticide, per pesticide type and year

Pesticide Type	Year	(	QA		TFI		LI	
		Lower	Upper	Lower	Upper	Lower	Upper	
	2009	1.9	4.5	-0.0	0.6	0.3	1.7	
	2010	0.9	2.1	-0.1	0.8	0.9	2.9	
All Pesticides	2011	1.8	4.4	-0.4	0.3	0.8	1.9	
	2012	1.4	3.6	0.1	1.1	0.3	1.3	
	2013	2.2	5.2	-0.5	0.2	0.1	1.0	

812 Table A3. 85% confidence intervals for skewness of yearly pesticide use distributions in potatoes

813 Skewness is calculated for each year (2009-2013) from farm and culture level pesticide use distributions. Lower 814 and Upper indicate the lower and upper bound of the 85% confidence interval for each statistic, respectively. 815

Payton et al. (2003) show that, given a maximum error rate of 5%, 85% confidence intervals should be used to

816 check for overlaps (see Methods, Section 2.2). Confidence intervals are computed with non-parametric 817 bootstrapping techniques with 10000 repetitions, respectively. QA, TFI and LI denote the Quantity of Active

818 Ingredient, Treatment Frequency Index and Load Index Indicators, respectively.

819

## Table A4. 85% confidence intervals for skewness of yearly pesticide use distributions in winter wheat

Pesticide Type	Year	(	QA		TFI		LI
		Lower	Upper	Lower	Upper	Lower	Upper
	2009	1.1	3.0	-0.3	0.3	0.9	1.7
	2010	1.3	3.1	-0.1	0.5	-0.0	0.7
All Pesticides	2011	0.2	0.8	0.0	0.6	2.0	3.7
	2012	-0.1	0.6	-0.3	0.3	1.2	2.9
	2013	0.5	1.3	0.3	1.8	1.6	2.9

823 Skewness is calculated for each year (2009-2013) from farm and culture level pesticide use distributions. Lower

and Upper indicate the lower and upper bound of the 85% confidence interval for each statistic, respectively.

Payton et al. (2003) show that, given a maximum error rate of 5%, 85% confidence intervals should be used to

826 check for overlaps (see Methods, Section 2.2). Confidence intervals are computed with non-parametric

827 bootstrapping techniques with 10000 repetitions, respectively. QA, TFI and LI denote the Quantity of Active

828 Ingredient, Treatment Frequency Index and Load Index Indicators, respectively.

829

#### 832 Table A5. 85% confidence intervals of summary statistics for farm-level pesticide use in potatoes

Dosticido Tupo	Confidence	(	QA	T	FI		LI
Pesticide Type	Interval of	Lower	Upper	Lower	Upper	Lower	Upper
	Mean	2.46	2.82	1.99	2.23	3.92	4.65
Herbicides	Sd	1.63	1.82	1.05	1.24	3.09	4.00
	Skewness	0.12	0.50	0.33	0.83	1.23	2.20
	Mean	5.98	6.66	5.54	6.19	4.41	5.08
Fungicides	Sd	3.10	3.47	2.87	3.31	2.99	3.67
	Skewness	-0.01	0.30	0.27	0.66	0.90	1.58
	Mean	10.43	11.87	8.48	9.22	8.86	9.84
All Pesticides	Sd	5.87	7.94	3.31	3.82	4.27	5.23
	Skewness	1.80	2.60	0.05	0.53	0.79	1.47

833 Summary statistics are calculated for the pooled data (2009-2013) on a farm and culture level. Lower and

834 Upper indicate the lower and upper bound of the 85% confidence interval for each statistic, respectively.

Payton et al. (2003) show that, given a maximum error rate of 5%, 85% confidence intervals should be used to

836 check for overlaps (see Methods, Section 2.2). Confidence intervals are computed with non-parametric

837 bootstrapping techniques with 10000 repetitions, respectively. QA, TFI and LI denote the Quantity of Active

838 Ingredient, Treatment Frequency Index and Load Index Indicators, respectively.

839

## Table A6. 85% confidence intervals of summary statistics for farm-level pesticide use in winter wheat

Posticido Tupo	Confidence	(	QA	Т	FI	LI	
- resticide Type	Interval of	Lower	Upper	Lower	Upper	Lower	Upper
	Mean	0.76	0.93	1.23	1.34	0.79	1.03
Herbicides	Sd	0.89	1.05	0.57	0.65	1.08	1.62
	Skewness	1.25	1.65	0.35	0.82	3.39	4.47
	Mean	0.71	0.84	1.25	1.42	1.32	1.55
Fungicides	Sd	0.67	0.86	0.92	1.01	1.22	1.4
	Skewness	1.29	2.71	-0.11	0.2	0.76	1.07
All Pesticides	Mean	1.9	2.15	3.13	3.37	2.32	2.73
	Sd	1.28	1.55	1.33	1.5	1.99	2.6
	Skewness	0.84	1.84	0.08	0.56	1.75	2.74

842 Summary statistics are calculated for the pooled data (2009-2013) on a farm and culture level. Lower and

843 Upper indicate the lower and upper bound of the 85% confidence interval for each statistic, respectively.

Payton et al. (2003) show that, given a maximum error rate of 5%, 85% confidence intervals should be used to

845 check for overlaps (see Methods, Section 2.2). Confidence intervals are computed with non-parametric

846 bootstrapping techniques with 10000 repetitions, respectively. QA, TFI and LI denote the Quantity of Active

847 Ingredient, Treatment Frequency Index and Load Index Indicators, respectively.

848

#### Table A7. Median, 90% decile and 99% percentiles of farm-level pesticide use in potatoes

Indicator	cator Median		99% percentile		
QA	9.85	17.32	41.92		
TFI	8.65	14.00	17.55		
LI	8.75	16.44	23.66		

Parameters are calculated for all years pooled (2009-2013) on a farm and culture level. QA, TFI and LI denote

the Quantity of Active Ingredient, Treatment Frequency Index and Load Index Indicators, respectively. The ratio

of 50% to 90% and 99% percentiles give information about the skewness of the distribution and the weight ofthe distribution on upper tails.

855

#### 857 Table A8. Median, 90% decile and 99% percentiles of farm-level pesticide use in winter wheat

Indicator	Median	90% decile	99% percentile
QA	1.78	3.9	6.1
TFI	3.17	5.14	6.53
LI	2.13	4.71	12.63

Parameters are calculated for all years pooled (2009-2013) on a farm and culture level. QA, TFI and LI denote

859 the Quantity of Active Ingredient, Treatment Frequency Index and Load Index Indicators, respectively. The ratio

of 50% to 90% and 99% percentiles give information about the skewness of the distribution and the weight ofthe distribution on upper tails.

## Table A9. Significant correlation of quantitative indicators and farm-level herbicide and fungicide risks – but very low explanatory power for extreme risks

Pesticide Type	Comparing LI Indicator to:	Kendall's τ correlation	Survival Gumbel Copula		Survival BB1 Copula	Survival BB7 Copula		
			L <sub>TDC</sub>	$U_{\text{TDC}}$	L <sub>TDC</sub>	$U_{\text{TDC}}$	L <sub>TDC</sub>	U <sub>TDC</sub>
Herbicides	QA Indicator	0.17	0.32	0.00	0.32	0.00	0.46	0.00
	TFI Indicator	0.46	0.58	0.00	0.58	0.00	0.67	0.20
Fungicides	QA Indicator	0.36	0.45	0.00	0.42	0.01	0.48	0.18
	TFI Indicator	0.62	0.70	0.00	0.68	0.10	0.74	0.52

#### Potatoes

#### Winter Wheat

Pesticide	Comparing LI	Kendall's τ	Survival Gumbel		Survival BB1		BB7	
Type	Indicator to:	correlation	Copula		Copula		Copula	
			L <sub>TDC</sub>	$U_{\text{TDC}}$	L <sub>TDC</sub>	$U_{\text{TDC}}$	L <sub>TDC</sub>	U <sub>TDC</sub>
Herbicides	QA Indicator	0.54	0.55	0.00	0.64	0.00	0.71	0.25
	TFI Indicator	0.48	0.57	0.00	0.50	0.24	0.57	0.47
Fungicides	QA Indicator	0.58	0.71	0.00	0.71	0.00	0.85	0.06
-	TFI Indicator	0.66	0.77	0.00	0.77	0.00	0.89	0.28

Level of dependence between farm-level indicators of pesticide use in potato and winter wheat production.

866 QA, TFI and LI denote the Quantity of Active Ingredient, Treatment Frequency Index and Load Index Indicators,

respectively. LTDC and UTDC are in the range of [0,1] and indicate lower and upper tail dependence

868 coefficients, i.e. the dependence for extremely low and extremely high observations, respectively. The three

analyzed copula families are chosen per crop for farm-level observations, according to the six-step selection

and testing procedure described in the Methodology, Section 2.4. Bold numbers indicate the best fitting copula

type per indicator pair, according to Vuong and Clarke goodness-of-fit tests (Vuong 1989, Clarke 2007). The

UTDC of the Survival Gumbel copula is zero by definition. Kendall correlation coefficients measure the
 (average) ordinal dependence between two indicators and are all significant at the 1% level. Detailed

robustness and goodness-of-fit tests for copulas and TDCs can be found in the Appendix, Section B3.

# Table A10. Significant correlation of quantitative indicators and field-level pesticide risks – but very low explanatory power for extreme risks

Comparing LI	Kendall's τ	Survival G	umbel	Survival BB1		Survival BB6	5
Indicator to:	correlation	Copula		Copula		Copula	
		L <sub>TDC</sub>	$U_{\text{TDC}}$	L <sub>TDC</sub>	UTDC	L <sub>TDC</sub>	$U_{\text{TDC}}$
QA Indicator	0.47	0.28*	0.00*	0.28	0.00	0.28*	0.00*
TFI Indicator	0.58	0.42	0.00	0.44	0.00	0.44	0.00

#### Winter Wheat

**Potatoes** 

Comparing LI	Kendall's τ	Survival Gum	bel	Survival BB1		Survival BB8	
Indicator to:	correlation	Copula		Copula		Copula	
		L <sub>TDC</sub>	UTDC	L <sub>TDC</sub>	Utdc	L <sub>TDC</sub>	Utdc
QA Indicator	0.22	0.57	0.00	0.57	0.00	0.00*	0.00*
TFI Indicator	0.37	0.63	0.00	0.62	0.00	0.00	0.00

878 Level of dependence between field-level indicators of pesticide use in potato and winter wheat production. QA, 879 TFI and LI denote the Quantity of Active Ingredient, Treatment Frequency Index and Load Index Indicators, 880 respectively. LTDC and UTDC are in the range of [0,1] and indicate lower and upper tail dependence 881 coefficients, i.e. the dependence for extremely low and extremely high observations, respectively. LTDC and 882 UTDC are in the range of [0,1] and indicate lower and upper tail dependence coefficients, respectively. The 883 three analyzed copula families are chosen per crop for field-level pesticide use observations, according to the 884 six-step selection and testing procedure described in the Methodology, Section 2.4. Bold numbers indicate the 885 best fitting copula type per indicator pair, according to Vuong and Clarke goodness-of-fit tests (Vuong 1989, 886 Clarke 2007). The UTDC of the Survival Gumbel, Survival BB6 and Survival BB8 copulas is zero by definition. 887 Kendall correlation coefficients measure the (average) ordinal dependence between two indicators and are all 888 significant at the 1% level. \* denotes acceptance of Kendall copula-goodness-of-fit tests (Genest et al. 2006). 889 Cramer-von Mises test statistics were used to accept/reject the null-hypothesis of a matching copula family at 890 the 5% level. P-values were computed according to the parametric bootstrap procedure described in Genest et 891 al. (2006).

# Table A11. Significant correlation of quantitative indicators and application-level pesticide risks – but very low explanatory power for extreme risks

Comparing LI	Kendall's τ	Studer	nt t	Survival	BB6	Survival	BB8
Indicator to:	correlation	Copula		Copula		Copula	
		L <sub>TDC</sub>	UTDC	$L_{TDC}$	UTDC	L <sub>TDC</sub>	Utdc
QA Indicator	0.10	0.06	0.06	0.28	0.00	0.00	0.00
TFI Indicator	0.23	0.07	0.07	0.25	0.00	0.00	0.00

## Potatoes

#### Winter Wheat

Comparing LI	Kendall's τ	Student t		Frank		Survival BB8	
Indicator to:	correlation	Copula		Copula		Copula	
		$L_{TDC}$	UTDC	L <sub>TDC</sub>	UTDC	L <sub>TDC</sub>	Utdc
QA Indicator	0.33	0.13	0.13	0.00	0.00	0.00	0.00
TFI Indicator	0.36	0.07	0.07	0.00	0.00	0.00	0.00

Level of dependence between application-level indicators of pesticide use in potato and winter wheat

896 production. QA, TFI and LI denote the Quantity of Active Ingredient, Treatment Frequency Index and Load

897 Index Indicators, respectively. LTDC and UTDC are in the range of [0,1] and indicate lower and upper tail

898 dependence coefficients, i.e. the dependence for extremely low and extremely high observations, respectively.

LTDC and UTDC are in the range of [0,1] and indicate lower and upper tail dependence coefficients,

900 respectively. The three analyzed copula families are chosen per crop for application-level pesticide use

901 observations. They represent the best fitting copulas according to the six-step selection and testing procedure

902 described in the Methodology, Section 2.4. Bold numbers indicate the best fitting copula type per indicator

pair, according to Vuong and Clarke goodness-of-fit tests (Vuong 1989, Clarke 2007). The UTDC of the Frank,

904 Survival BB6 and Survival BB8 copulas is zero by definition. All estimates of Kendall correlation coefficients are

significant at the 1% level. Kendall (Genest et al. 2006) goodness-of-fit tests with Cramer-von Mises test
 statistics were used to test the null-hypothesis of matching Survival BB6, Survival BB8 and Frank copulas. P-

907 values were computed according to the parametric bootstrap procedure described in Genest et al. (2006). PIOS

goodness-of-fit tests were used for the Student t copula (Zhang et al. 2016). Copula goodness-of fit could not
be confirmed at the 5% level with the above tests.

910



#### 913 Figure A1. Linkages of used databases

- 914 Source: Own depiction. 1 : Central Evaluation of Agri-Environmental Indicators, Switzerland (2009-2013,
- 915 compare: Spycher et al. 2013); 2 : Pesticides Properties Database, March, 2016 (compare: Lewis et al. 2016); 3:
- 916 Pesticide Load Indicator (compare Miljøministeriet 2012); 4: BLW Pflanzenschutzmittelverzeichniss (2009-2013,
- 917 compare BLW 2013).



919

920 921 Fig. A2. Pesticide use trends shift by indicator

222 Linear pesticide use trends in winter wheat and potatoes from 2009-2013 with the QA, TFI and LI indicators.

923 Linear trends are significant for TFI and LI indicators in winter wheat at the 10% and 5% level, respectively. No

924 significant trend was found for potatoes. Shaded areas indicate 95% confidence intervals of linear trends.



#### 927 Fig. A3. Distributions of LI and QA indicators show more skewness than the TFI

928 Histograms of farm-level pesticide use in potato production (pooled observations from 2009-2013). QA, TFI and

229 LI denote the Quantity of Active Ingredient, Treatment Frequency Index and Load Index Indicators,

930 respectively.



#### 932 Fig. A4. Distributions of LI and QA indicators show more skewness than the TFI

933 Histograms of farm-level pesticide use in winter wheat production (pooled observations from 2009-2013). QA,

TFI and LI denote the Quantity of Active Ingredient, Treatment Frequency Index and Load Index Indicators,respectively.





Fig. A5. Copulas reveal differences in the dependence structure of indicators between upper and
 lower tails (TFI and LI for potatoes).

939 Comparing farm-level pesticide use in potatoes measured with the TFI and LI indicators. Figure a) shows a

940 scatterplot of TFI and LI values. Figure b) shows a contour plot of the fitted copula model (Survival BB1 copula,

941 see Methods, Section 2.4 for selection and estimation procedure). TFI and LI denote the distribution functions

942 for TFI and LI values, respectively (see Methods, Section 2.3). The copula captures the dependence structure

943 between the distributions of the two indicators, not only on average, but over the whole distribution.





## Fig. A6. Copulas reveal differences in the dependence structure of indicators between upper and lower tails (QA and LI for winter wheat).

948 Comparing farm-level pesticide use in winter wheat measured with the QA and LI indicators. Figure a) shows a

scatterplot of QA and LI values. Figure b) shows a contour plot of the fitted copula model (Survival Gumbel

950 copula, see Methods, Section 2.4 for selection and estimation procedure). QA and LI denote the distribution

951 functions for QA and LI values, respectively (see Methods, Section 2.3). The copula captures the dependence

952 structure between the distributions of the two indicators, not only on average, but over the whole distribution.

#### b) Copula Contour Plot





# Fig. A7. Copulas reveal differences in the dependence structure of indicators between upper and lower tails (TFI and LI for winter wheat).

957 Comparing farm-level pesticide use in winter wheat measured with the TFI and LI indicators. Figure a) shows a

958 scatterplot of TFI and LI values. Figure b) shows a contour plot of the fitted copula model (Survival Gumbel

959 copula, see Methods, Section 2.4 for selection and estimation procedure). TFI and LI denote the distribution

960 functions for TFI and LI values, respectively (see Methods, Section 2.3). The copula captures the dependence

961 structure between the distributions of the two indicators, not only on average, but over the whole distribution.



#### 964 Fig. A8. Temporal hotspots of pesticide use shift with indicator choice (winter wheat).

965 Single pesticide applications in winter wheat, plotted on a daily scale. Observations include application data for

the years 2009-2013 and all sample farmers. Pesticide use is expressed in Quantity of Active ingredient (QA),
 Treatment Frequency Index (TFI) and Load Index (LI) indicators respectively. "Temporal Hotspots" indicate the

968 Top 1% of all applications per indicator, respectively.

970		<b>B1. Documentation of data cleaning procedure</b>
971	The fol	lowing detailed procedure was used to clean data entry mistakes from the raw data:
972	1)	Selection. Only farmers with complete observations (field calendar and socio-
973		economic/bookkeeping data), only relevant cultures, only pesticide applications where all
974		indicators are available, only farmers with a focus on arable farming (exclude those which
975		have an area share of more than 5% wine or fruit).
976	2)	<i>Remove outlier fields</i> . Only keep those fields where seeding <u>and</u> harvesting activity (in this
977		order) is reported. Eliminate fields where a yield of zero is indicated or no seeding and or
978		harvesting takes place.
979	3)	Remove outlier applications. Remove applications with double accounting (same product,
980		same day, same amount).
981	4)	Correct farmers' typos. According to the following procedure: Identify applications where the
982		product is applied with an amount of (mean +- $1$ *sd). For applications (mean + $1$ *sd) also
983		check if the amount exceeds the recommended standard dosage. For those outliers, check
984		every individual application in the context of the whole field calendar for the given field/year
985		and correct if necessary. In addition, compare those applications to applications by other
986		farmers and applications of the same farmer in other years. In particular, consider the
987		possibility of split applications/packages/mixtures for those outliers.
988	Each a	oplication was tested according to the above procedure with the software package R (R Core
989	Team 2	2013) to avoid removing or changing actual documented behavior and only remove individual
990	mistak	es and thus guarantee a high data quality. The procedure was designed to test the applications
991	in the o	context of the given regulations, of applications of the similar products on other Swiss farms,

- other applications on the same farm and the data quality of information on other management
- 993 measures provided by the farmer. Given the richness of the data, applications could therefore be
- 994 subject to context-specific evaluation thus guaranteeing a higher reliability than simple outlier
- 995 identification algorithms.

#### B2. Validity checks for self-reporting and self-selection

Swiss pesticide regulations explicitly forbid pesticide overdosing (application dosages higher than 998 999 recommended dosages). Therefore, to address any concerns regarding self-reporting by farmers 1000 checks were carried out to verify that they complied with these regulations according to the reported 1001 data, or reported any violations of these regulations. The latter would be a sign for a high degree of 1002 trust in anonymity assurances given to farmers. Applications were rated as overdosed when the 1003 maximum allowed amount in the respective culture was exceeded by more than 5%. 5-year average 1004 values of the share of overdosed applications were then computed per culture and pesticide type. 1005 The share of overdosed applications in potato and winter wheat production is 2% and 4.8% for 1006 herbicides and 5.6% and 3.4% for fungicides, respectively. This suggests that farmers were not 1007 concerned about reporting and anonymity issues during data collection. We also compared 1008 important socio-economic characteristics of our sample to the population averages, as reported by 1009 Agroscope for 2013 (Hoop and Schmid 2014) and the Swiss Institute for Statistics for 2016 (BFS 1010 2017). The average farm size in the sample is 26.8 ha, which is representative of Swiss crop 1011 producers with an average farm size of 26.5 ha. We also find a representative age structure in our 1012 sample. The percentage of farmers in a given age group (population averages in parenthesis) is as 1013 follows: <39 years: 28(21)%; 40-49 years: 30(30)%; 50-59 years: 35(37)% and >60 years 7(13)%. 1014 Finally, we found the ratio of farmers with no education lies at 15% compared to 2% as the 1015 population average. The higher rate of educated farmers, compared to population averages, might 1016 be explained by the focus on crop producing farms in our sample, compared to a high share of milk 1017 producing farms in the population. Moreover, our dataset does not cover very remote regions, as 1018 crop farms are located in more accessible regions (compare Figure 2).

1019

#### **B3.** Assessing robustness of farm-level results for copulas and tail dependence 1020

1021 As described in the Methodology Section, step v) of the article, robustness was also checked 1022 regarding the results obtained for copula estimation and tail dependence with a parametric 1023 bootstrap test and the stable tail dependence function, respectively. The first robustness check was a 1024 goodness-of-fit test based on Kendall's process as described in Genest et al. (2006). Cramer-von-1025 Mises test statistics were used to assess if the null-hypothesis, that the chosen copula family matches 1026 the empirical copula, is rejected or accepted at the 5% level (Genest et al. (2006) report a higher test 1027 power for Cramer-von-Mises test statistics than for Kolmogorov-Smirnov test statistics). Genest et al. 1028 (2006) and Genest and Rémillard (2008) reported that test power for some of the copula families is 1029 low. Furthermore, it was not possible to identify a power study for the two-parameter copulas used 1030 to assess tail dependence in this article. However, the results of the test still give some indication of 1031 whether the distributional assumptions of the selected copulas are met, or if further analysis is 1032 needed.

Potatoes Survival Gumbel Survival BB1 Survival BB7 Herbicides QA-LI Reject Reject Reject TFI-LI Accept Accept Accept

Accept

Accept

Accept

QA-LI

TFI-LI

QA-LI

TFI-LI Reject

#### 1033 Table B1. Results of copula goodness-of-fit tests

Fungicides

**All Pesticides** 

#### Winter Wheat

		Survival Gumbel	Survival BB1	BB7
Herbicides	QA-LI	Accept	Accept	Reject
	TFI-LI	Reject	Accept	Accept
Fungicides	QA-LI	Reject	Reject	Reject
	TFI-LI	Reject	Reject	Reject
All Pesticides	QA-LI	Accept	Accept	Reject
	TFI-LI	Reject	Reject	Reject

Accept

Accept

Accept

Accept

Accept

Accept

Accept

Reject

1034 Note: The goodness-of-fit test described in Genest et al. (2006) was used. Cramer-von Mises test statistics were 1035 used to accept/reject the null-hypothesis of a matching copula family at the 5% level. P-values were computed 1036 according to the parametric bootstrap procedure described in Genest et al. (2006).

1037 Table B1 reports outcomes of the goodness-of-fit test described by Genest et al. (2006) for all

1038 indicator pairs and copula families for farm-level pesticide use. With the exception of four copula

1039 pairs (QA-LI for herbicides in potatoes, fungicides in winter wheat, TFI-LI for all pesticides in winter

1040 wheat) goodness-of-fit for the best fitting copula is confirmed. The semi-parametric approach used

1041 to assess tail dependence in this article is expected to show a bad performance if distributional

1042 assumptions are not met (Frahm et al. 2005). In addition, the fully non-parametric stable tail

1043 dependence function (Kiriliouk et al. 2016) was used to assess tail dependence and test if the results

1044 on tail dependence remain valid for those four indicator pairs. 1045 Computation of the stable tail dependence function was implemented in accordance with Kiriliouk et

- al. (2016) and using the R package copBasic (Asquith 2017). The function was estimated for thefollowing indicator pairs:
- 1048 i) QA and LI indicator for herbicides in potato production
- 1049 ii) QA and LI indicator for fungicide use in winter wheat production
- 1050 iii) TFI and LI indicator for fungicide use in winter wheat production
- 1051 iv) TFI and LI indicator for overall pesticide use in winter wheat production

1052 More specifically, it was checked to verify that the statement of a low degree of upper tail 1053 dependence for i) to iv) respectively, can be supported by the results of the stable tail dependence 1054 function method. To this end, the stable tail dependence function was plotted for all indicator pairs, 1055 as suggested by Kiriliouk et al. (2016). This allows qualitative confirmation or rejection of findings on 1056 tail dependence. The approach consists of plotting several levels of the stable tail dependence 1057 function, where levels are ordered according to their relative distance to the upper endpoint of the 1058 joint distribution of the indicator pairs. The stable tail dependence function l(x) is defined as 1059 follows.

1060 
$$l(x) = \lim_{t \downarrow 0} t^{-1} \mathbb{P}[F_1(X_1) > 1 - tx_1, \dots, F_d(X_d) > 1 - tx_d]$$

1061 where  $X = (X_1, ..., X_d)$  is a random vector with continuous marginal distribution functions and a 1062 joint distribution function.  $\mathbb{P}(X_1 \le x_1, \dots, X_d \le x_d), t > 0$  is small and the numbers  $x_1, \dots, x_d \in X_d$  $[0,\infty)$  parametrize the relative distances to the upper endpoints of the *d* variables (in this case two, 1063 1064 namely the respective indicator pair). Tail dependence is therefore investigated close to the upper 1065 endpoints (1,1) of the joint (bivariate) distribution. An example of how to interpret plots of the stable 1066 tail dependence function is given in Figure B1 and Figure B2. These illustrate plots of the stable tail dependence function for several levels of two Gumbel-Hougaard copulas with theoretical upper TDCs 1067 of 0.78 and 0.01 indicating strong upper tail dependence and almost no upper tail dependence, 1068 1069 respectively.



Figure B1. Example plot of the stable tail dependence function for a Gumbel-Hougaard copula witha TDC of 0.78





- 1078 Level sets would show 90° bends and straight lines, respectively for the two extreme cases of
- 1079 complete asymptotic dependence and asymptotic independence. Given this indication, the upper
- 1080 figure clearly depicts the high degree of tail dependence present, whereas the lower figure points

towards (asymptotic) independence, as expected. Figures B3, B4, B5 and B6 show plots of the stable
 tail dependence function for the indicator pairs i) - iv), respectively.













1090 Figure B5. Plot of the stable tail dependence function for indicator pair iii)





1093 The plots clearly indicate a low level of tail dependence for indicator pairs i) to iv) as the level sets are 1094 close to straight lines. However, level sets of indicator pair iv) get closer to the curves shown in 1095 Figure A10 with a higher relative distance to the end point which indicates a higher degree of 1096 asymptotic dependence than indicator pairs i) to iii). These findings are in line with the values 1097 reported for the upper TDCs in Tables 1 and A5, where the estimated upper TDCs for indicator pairs i) 1098 to iii) were {0.00; 0.00;0.00}, {0.00;0.00;0.06} and {0.00;0.00;0.28} and {0.00;0.01;0.32} for indicator 1099 pair iv), respectively. The findings for the stable tail dependence function method are therefore in 1100 line with the findings reported in Tables 1 and A9 and qualitatively support the result indicating only 1101 a weak upper tail dependence.

1102

```
1103 Online Appenix: R-Code
```

```
1104
      # Niklas Möhring, 13.06.2018
1105
      # Appendix to the paper "Quantity based indicators fail to identify extreme
      pesticide risks" by Möhring, Gaba and Finger
1106
       # Documentation of the computation procedure of tail dependence
1107
1108
      coefficients and correlation coefficients in R
1109
       # A detailed description of the procedure and sources can be found in
1110
      section 3.4 of the paper
1111
1112
1113
      library(dplyr)
1114
      library(VineCopula)
1115
      library(copBasic)
1116
1117
1118
      # i) Test independence of indicator pairs analytically with independence
1119
      tests ####
1120
1121
1122
      # Independence tests
1123
      # Repeat for all indicator combinations, crops and aggregation levels
1124
1125
      BiCopIndTest(dat[1]][,1], dat[[1]][,2]) # Convert observations to pseudo
1126
      observations with pobs() before testing
1127
1128
1129
1130
1131
       # ii) Estimate copula parameters and TDCs for a wide range of copula
1132
      families with a semi-parametric estimation procedure with empirical margins
1133
      ####
1134
1135
1136
       # Repeat for all indicator combinations, crops, aggregation levels and
1137
      copula families
1138
      summary(BiCopSelect(dat[[1]],dat[[2]],familyset = xy,rotations = FALSE)) #
1139
      Convert observations to pseudo observations with pobs() before testing,
1140
      insert specific copula in familyset
1141
1142
1143
1144
1145
      # iii) Choose the three best fitting copula families over all indicator
1146
      pairs ####
1147
1148
1149
       # Choose the three best fitting copula families according to AIC and BIC
1150
      values, as well as graphical fit with simulated copula distributions
1151
       # Repeat for all indicator combinations, crops and aggregation levels
1152
1153
      BiCopCompare(pobs(dat$ind1), pobs(dat$ind2))
1154
1155
1156
1157
1158
1159
       # iv) Identify copula family with best goodness-of-fit ####
1160
1161
       # Rank copula families according to Vuong and Clarke tests (with and
1162
      without corrections)
1163
       # Repeat for all indicator combinations, crops and aggregation levels
```

```
1164
1165
      BiCopVuongClarke (pobs (dat$ind1), pobs (dat$ind2)) # 20 - 16 -19
1166
      BiCopVuongClarke(pobs(dat$ind1), pobs(dat$ind2), correction = "Akaike")
1167
      BiCopVuongClarke(pobs(dat$ind1), pobs(dat$ind2), correction = "Schwarz")
1168
      Based on the results of all tests identify the copula with best goodness of
1169
      fit out of the three families in step iii)
1170
1171
1172
1173
1174
1175
      # v) Check robustness of obtained copula results with a parametric
1176
      bootstrap test and the fully non-parametric stable tail dependence function
1177
      approach ####
1178
1179
1180
1181
      # Semi-parametric bootstrap test for copula goodness of fit based on
1182
      Kendalls process (and CvM statistics)
1183
      # Repeat for all indicator combinations, crops and aggregation levels
1184
1185
      BiCopGofTest(dat[[1]], dat[[2]], family = XY,method = "Kendall",B=100) #
      Test if the above chosen copula families resemble the empirical copula well
1186
1187
      based on CvM results
1188
1189
1190
1191
      # Fully non-parametric robustness checks for copulas using the stable tail
1192
      dependence function
1193
      # Repeat for all indicator combinations, crops and aggregation levels
1194
1195
1196
      stabtaildepf(dat[[1]]) # the data must contain two transformed (pobs())
1197
      columns with the indicators to compare
1198
      stabtaildepf(dat[[1]], smooth=TRUE, ploton=F, col=2)
1199
1200
1201
1202
1203
      # Estimate Kendalls tau correlation coefficients and p-values ####
1204
      # Repeat for all indicator combinations, crops and aggregation levels
1205
1206
      cor.test( dat[[1]], dat[[2]], method = "kendall", exact = FALSE)
1207
```