Integrated Environmental Assessment and Management — Volume 18, Number 5—pp. 1399–1413
Received: 2 June 2021 | Revised: 28 September 2021 | Accepted: 29 November 2021

Health & Ecological Risk Assessment

Holistic evaluation of long-term earthworm field studies with a fungicide

Stephan Jänsch,[1] Sonja Braaker,[2] Jörg Römbke,[1] Frank Staab,[3] and Tobias Pamminger[3,4]
[1]*ECT Oekotoxikologie GmbH, Flörsheim, Germany*
[2]*BASF S.A.S., Ecully, France*
[3]*BASF SE, Ludwigshafen, Germany*
[4]*Current affiliation: Bayer CropScience, Monheim am Rhein, Germany*

Abstract

Plant protection products to be placed on the market in the European Union need to meet rigorous safety criteria including the testing of lumbricid earthworms, the functionally most important soil organism group in Central European agricultural ecosystems. To address uncertainties and investigate the potential long-term in-crop effects of the fungicide Cantus® containing 50% boscalid as an active substance, a series of standardized earthworm field studies with an overall duration of 5 years per study program was carried out in four German agricultural fields under realistic crop rotation conditions. A two-step approach was chosen to analyze the potential overall long-term effects on earthworms in agricultural fields: (i) an assessment of the earthworm abundance development in the course of the four study programs in relation to the determined actual content of boscalid in soil and (ii) an effect size meta-analysis of earthworm abundance 1 year after treatment for each consecutive year and study program. Measured boscalid concentrations in the soil after multiple applications were well above the maximum boscalid residues observed in agricultural soils across Central Europe. There were isolated statistically significant reductions of earthworm abundance for some species and groups at some time points during the studies, but no consistent relationship to the Cantus® treatments was observed. These results were supported by the meta-analysis, indicating no adverse effects on earthworm populations. Therefore, fluctuations of abundance reflect the natural variation of the populations rather than a concentration-related response. Based on this comprehensive analysis, we conclude that there is no application rate-related effect of the 5-year use of Cantus® on the development of the earthworm communities. The four study programs, paired with a comprehensive evaluation, directly address the concerns about the potential long-term effects of boscalid on earthworms in the field and suggest that multiyear applications do not adversely affect earthworm populations. *Integr Environ Assess Manag* 2022;18:1399–1413. © 2021 ECT Oekotoxikologie GmbH and BASF SE. *Integrated Environmental Assessment and Management* published by Wiley Periodicals LLC on behalf of Society of Environmental Toxicology & Chemistry (SETAC).

KEYWORDS: Boscalid, Environmental risk assessment, Higher-tier, Lumbricidae, Meta-analysis

INTRODUCTION

This article presents an evaluation of the potential long-term in-crop adverse effects of the fungicide Cantus® containing 50% boscalid as an active substance (a.s.) on earthworms in the field. Adverse ecological effects represent changes that are considered undesirable because they alter valued structural or functional characteristics of ecosystems or their components (US Environmental Protection Agency [US EPA], 1998), in this context, a reduction in the abundance of the earthworm population, single species, age stages, or ecological groups. Plant protection products (PPP) to be placed on the market in the European Union (EU) need to meet rigorous criteria that among others include the safety for the environment (European Commission [EC], 2009). To prove environmental safety, the regulation of the EU requires the comprehensive testing of both the a.s. (EC, 2013a) and the formulated products (EC, 2013b) following a tiered environmental risk assessment (ERA) scheme utilizing test guidelines standardized by the Organization for Economic Co-operation and Development (OECD) and the International Organization for Standardization (ISO). These requirements comprise the testing of soil invertebrates, including lumbricid earthworms, the functionally most important soil organism group in Central European agricultural

This article contains online-only Supporting Information.

Correspondence Stephan Jänsch, ECT Oekotoxikologie GmbH, Flörsheim, Germany.
Email: s-jaensch@ect.de

Published 2 December 2021 on wileyonlinelibrary.com/journal/ieam.

ecosystems (Edwards, 1983; Edwards & Bohlen, 1996; Lee, 1985; Sims & Gerard, 1999; Uvarov, 2009; Van Groenigen et al., 2014). They are commonly categorized according to their different ecological roles (Bottinelli et al., 2020; Bouché, 1977). If a potential risk is identified in laboratory earthworm reproduction testing with *Eisenia fetida/andrei*, an earthworm field study must be performed according to ISO standard 11268-3 (ISO, 2014).

Based on the laboratory earthworm reproduction study provided to the registration authorities, the long-term risk of the formulation Cantus® for earthworms could not be excluded as laid out in Anonymous (2002) and EC (2008). Consequently, two field tests with the formulation Cantus® and a duration of 2 years were conducted at grassland sites in Germany to assess the risk under realistic conditions. However, uncertainty remained, in particular regarding the relationship between the use of Cantus® and its potential long-term adverse effects on earthworms. To address this remaining uncertainty and to investigate the potential adverse effects in a realistic agricultural setting, a series of standardized earthworm field studies with an overall duration of 5 years per study program (i.e., a series of five sequential field studies on the same experimental field) was carried out with Cantus® in four agricultural fields in Germany under realistic crop rotation conditions and according to the OECD principles of Good Laboratory Practice (OECD, 1998) from 2005 to 2010. These studies were initially evaluated individually for each year in a conventional manner according to the ISO standard 11268-3 requirements (ISO, 2014) and current best scientific practice. In this initial analysis, statistically significant effects for individual species at certain time points were detected, but no long-term adverse effects at the population level were seen. Analyzed in isolation, individual studies can suffer from limitations like the low number of earthworm samples and high variation in the measured biological variables, which may result in elevated levels of statistical uncertainty regarding the conclusions. For this reason and given the available comprehensive data set, a novel two-step approach was chosen to quantify potential overall long-term adverse effects on earthworms in agricultural fields: (1) In the first step, an assessment of the earthworm abundance development in the course of the four study programs in relation to the actual content (measured or calculated) of boscalid in soil, expressed as the so-called expected application rate, was performed. (2) In the second step, a formal effect size meta-analysis of earthworm abundance 1 year after treatment for each consecutive year and study program was conducted. A meta-analysis is a statistical approach used to combine the results of a set of studies into a common analytical framework (Borenstein et al., 2011; DerSimonian & Laird, 1986). It aims at combining the available information (e.g., effect sizes and associated confidence intervals) and estimating an overall effect and the associated error. This approach enables the integration of the results of multiple studies, thereby facilitating a more comprehensive conclusion. In recent years, likely due to the increased availability of comparable data, meta-analytical thinking has gained traction in the ecotoxicological community (Bundschuh et al., 2011; Cresswell, 2011; Garcia-Reyero et al., 2011; Lavoie et al., 2013; Pelosi et al., 2013; Scholz et al., 2018; Veltman et al., 2007) and first steps in the context of ERA have been published combining multiple results of surface water residues in US surface waters (Wolfram et al., 2018). Such approaches are vital for addressing the critical issue of statistical uncertainty in ecotoxicological risk assessment (European Food Safety Authority [EFSA], 2018, 2019; Lofstedt & Bouder, 2021; Mair et al., 2020). The combination of the applied evaluation approaches promises to more accurately characterize the potential long-term risk of Cantus® application to earthworm populations under realistic agricultural conditions as opposed to the standard approach in regulatory ERA where usually only individual earthworm field studies with a duration of 1 year are being considered.

MATERIALS AND METHODS

Boscalid

The pyridinecarboxamide a.s. boscalid (CAS-number: 188425-85-6) contained in the formulation Cantus®, is a broad-spectrum fungicide developed by BASF and first placed on the European market in 2003. It is designed to control several plant pathogenic fungi (such as *Botrytis* spp., *Alternaria* spp., and *Sclerotinia* spp.), targets the succinate dehydrogenase, and thus inhibits spore germination, germ tube elongation, mycelial growth, and sporulation.

Study programs

Four study programs (Table 1) were performed in parallel from 2005 to 2010 in arable fields in Germany with Cantus® applied to a 2-year interval vegetable crop rotation at two different sites located within the continental biogeographical region of Europe (European Environment Agency, 2017). The study programs were designed as five sequential earthworm field studies according to ISO standard 11268-3 (ISO, 2014; 1999 edition, considering recommendations by Kula et al., 2006) with each earthworm sampling after 1 year of study duration simultaneously being the preapplication sampling for the subsequent study. The individual studies followed the standard limit test approach with two to six annual applications. Since the total annual application rates were above the registered use rate and applications were repeated over 5 consecutive years, these study programs are considered to represent a realistic worst case. The fungicides benomyl and carbendazim were applied as reference substances (positive control).

Earthworm sampling was performed following ISO standard 23611-1 (ISO, 2018) (i.e., a combination of hand-sorting and formalin extraction) at four sampling spots per plot. The earthworm community composition (first-year preapplication sampling data, means of all 12 plots) is given in Table 2, confirming a total earthworm abundance above 60 individuals/m² as required by the ISO standard

TABLE 1 Location, design, application scheme, crop rotation regime, and soil properties of the study programs

Study program code	A1	A2	B1	B2
Site (federal state) and field	Kraichtal (Baden-Württemberg), field Neuenbürg		Leipzig (Free State of Saxony), field Gerichshain	
pH ($CaCl_2$)	6.8–7.0	6.9–7.0	5.8–6.0	5.8–5.9
Texture	Loamy silt	Loamy silt	Sandy loam silt	Silty loam sand
Sand (%)	10.7–14.3	12.5–16.1	35.8	35.9
Silt (%)	76.3–78.5	75.5–78.4	51.1	47
Clay (%)	9.3–10.7	8.3–10.5	13	10.2
C_{org} (%)	1.56–2.28	1.83–2.24	1.20–1.25	1.03–1.30
Treatments	Negative control, test item, positive control (reference)			
No. of plots	Four per treatment (12 in total)			
Plot size (m)	12 × 12		10 × 10	
Test item application rate (g boscalid/ha) (crop)	1^{st}, 3^{rd}, and 5^{th} year: 2 × 400 (green lettuce); 1 × 500 (beans)	1^{st}, 3^{rd}, and 5^{th} year: 2 × 400 (green lettuce); 2 × 500 beans	1^{st}, 3^{rd}, and 5^{th} year: 2 × 250 (cabbage); 2 × 500 (peas)	1^{st} year: 250 + 500 (cabbage); 500 + 1000 (peas), 3^{rd} year: 3 × 250 (cabbage); 3 × 500 (peas), 5^{th} year: 3 × 250 (cabbage); 1000 + 500 (peas)
	2^{nd} and 4^{th} year: 2 × 350 (winter wheat)	2^{nd} and 4^{th} year: 2 × 350 (winter wheat)	2^{nd} and 4^{th} year: 2 × 350 (winter cereal mix)	2^{nd} and 4^{th} year: 2 × 350 (winter cereal mix)
	Total over 5 years: 5300	Total over 5 years: 6800	Total over 5 years: 5900	Total over 5 years: 8150
Reference application rate (kg a.s./ha)	1^{st} year: 6 + 8 (a.s. benomyl)		1^{st} year: 4 + 8 (a.s. benomyl)	
	2^{nd} and 3^{rd} year: 8 (a.s. benomyl)		2^{nd} year: 10 (a.s. benomyl)	
	4^{th} and 5^{th} year: 10 (a.s. carbendazim)		3^{rd} year: 5 (a.s. carbendazim)	
			4^{th} and 5^{th} year: 10 (a.s. carbendazim)	
Earthworm sampling method	Four sampling spots per treatment and sampling date (0.25 m², 20–30 cm depth), combined handsorting and formalin extraction			

Abbreviation: a.s., active substance.

11268-3 (ISO, 2014). Also, both endogeic and anecic species (Bouché, 1977) were present at a sufficiently high density in all arable fields. The earthworm communities differed between the study programs to varying extent. The overall abundance of adults was higher in study programs A1 and A2 than in B1 and B2. However, species number and composition are better suited for site comparisons than species abundance, mainly because species numbers are less variable than abundance. Earthworm abundance can strongly vary seasonally and from year to year depending on local weather conditions, in particular temperature and moisture (Edwards & Bohlen, 1996). In arable soils, soil cultivation additionally influences abundance (Edwards, 1983). Study programs A1 and A2 exhibited some minor differences among each other but were still comparable (Figure 1). This is also true for the two study programs B1 and B2 that held an almost identical earthworm community (i.e., the same species and dominance spectrum). Comparing these two pairs of study programs, the main difference lay in the total earthworm abundance rather than the species spectrum. The same four species made up most of the community at all four study programs: *Allolobophora chlorotica*, *Aporrectodea caliginosa*, *Aporrectodea rosea*, and *Lumbricus terrestris*. The similarity of the earthworm species composition was additionally confirmed by the Sørensen-Dice similarity coefficients of 0.86 (A1:A2), 0.67 (A1:B1, A1:B2, A2:B1, A2:B2), and 1.00 (B1:B2).

The location, design, application scheme, crop rotation regime, and soil properties of the study programs are listed in Table 1 (see also Table S1). The soils of the study programs were similar since the differences in some soil properties (mainly sand content and C_{org}) are not considered large enough to have a strong impact on earthworm community structure or function (Jänsch et al., 2013).

In the course of each of the 5-year long-term earthworm field study programs, a total of 16 postapplication earthworm samplings were conducted for study programs A1 and A2 as

 DOI: 10.1002/ieam.4562

TABLE 2 Earthworm community composition and dominance spectrum (adults) in the four study programs (first-year preapplication data, means of 12 plots)

Study program	A1		A2		B1		B2	
Species	Abundance (individuals/m^2)	Dominance (%)	Abundance (individuals/m^2)	Dominance (%)	Abundance (individuals/m^2)	Dominance (%)	Abundance (individuals/m^2)	Dominance (%)
Allolobophora chlorotica	11.6	12.2	21.4	16.7	2.0	6.6	5.5	20.1
Aporrectodea caliginosa	54.9	57.7	66.3	51.6	12.5	41.5	13.3	48.5
Aporrectodea longa	0.0	0.0	0.0	0.0	0.3	1.0	0.3	1.1
Aporrectodea rosea	13.7	14.4	24.0	18.7	2.0	6.6	1.1	4.0
Lumbricus castaneus	2.5	2.6	2.8	2.2	0.0	0.0	0.0	0.0
Lumbricus rubellus	0.0	0.0	1.5	1.2	0.0	0.0	0.0	0.0
Lumbricus terrestris	11.3	11.9	11.4	8.9	13.3	44.2	7.2	26.3
Octolasion cyaneum	0.9	0.9	1.0	0.8	0.0	0.0	0.0	0.0
Octolasion tyrtaeum	0.2	0.2	0.0	0.0	0.0	0.0	0.0	0.0
Total epigeic adults	2.5	2.6	4.3	3.3	0.0	0.0	0.0	0.0
Total endogeic adults	81.3	85.5	112.7	87.8	16.5	54.8	19.9	72.6
Total anecic adults	11.3	11.9	11.4	8.9	13.6	45.2	7.5	27.4
Total adults	95.1	-	128.4	-	30.1	-	27.4	-
Total juveniles	152.0	-	225.6	-	100.2	-	71.9	-
Undetermined	17.7	-	22.5	-	3.0	-	2.5	-
Total earthworms	264.8	-	376.5	-	133.3	-	101.8	-
No. of species	7	-	7	-	5	-	5	-

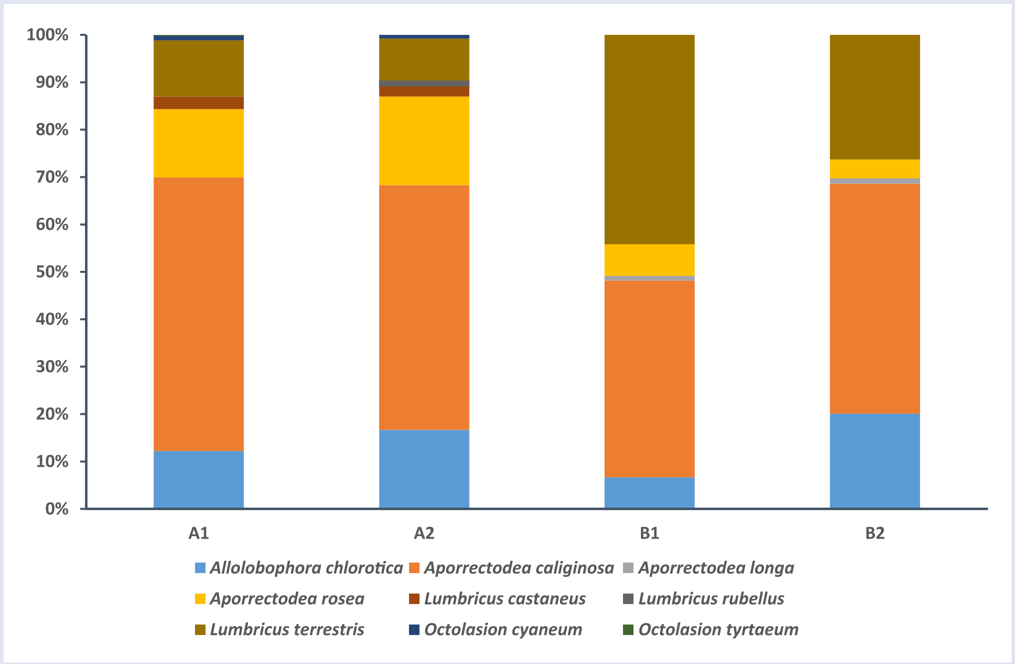

FIGURE 1 Earthworm community composition and dominance spectrum (adults) in the four study programs (preapplication data)

well as 14 samplings for study programs B1 and B2. These were performed 1 (not in year 2 of study programs B1 and B2), 2 (only in year 3 of study programs A1 and A2), 4–6 and 10–12 months after the first annual application of Cantus®.

Chemical analyses of boscalid residues in soil were regularly performed over the course of the study programs (sample depth up to 25 cm, five samples per plot, analysis via liquid chromatography-mass spectrometry/mass spectrometry):

- Year 1: 6–7 days after the second and 4–27 days after the last application;
- Years 2 and 4: 1–19 days before the first and 30–70 days after the second (= last) application;
- Year 3: 16–70 days before the first, 2–23 days after the second (B2: third), and 7–22 days after the last application;
- Year 5: 16–97 days before the first and immediately (A1 and A2) or 3 days (B1 and B2) after each application.

Representativeness of the arable fields regarding soil parameters and earthworm community

To gauge the meaningfulness of our results for crop sites across Europe in general, we analyzed whether the investigated arable fields are representative of a wide range of such sites in terms of both soil properties and earthworm community composition.

The physicochemical properties of the arable fields (Table 1) were compared with the typical ranges of ecologically relevant soil parameters for agricultural soils in Germany (Bussian et al., 2005) as well as maps of the European land cover (Copernicus Land Service, 2020) and relevant soil properties, such as pH (Ballabio et al., 2019), C_{org}, or sand content (EC, 2005). Additionally, they were compared with the ranges of soil parameters covered in the German "RefeSol" (Bussian et al., 2005) and the European "EUROSOILS" (Gawlik et al., 2001) reference systems that were among others developed to contain soils that are typical representatives for a large area of Germany and the EU, respectively.

To assess the ecological significance of the test results obtained from the four arable fields, we checked whether their earthworm communities were representative of German crop sites based on the so-called reference approach (Breure et al., 2005; Römbke et al., 2016). The principle of this approach is a comparison of site-specific earthworm data with a reference community, using abundance, species number, and community composition as assessment endpoints (Jänsch et al., 2013). The basic idea is that each site features specific soil and site parameters resulting in a site-specific earthworm community (Römbke & Breure, 2005). Sites can be classified into habitat types, based on soil and site parameters, such as the European Nature Information System (EUNIS) habitat type classification (Davies et al., 2004). For selected top-level habitat types including crop sites, the relative frequency of the 10 most common earthworm species in Germany, the mean number of earthworm species and their range of total adult abundance per site expected to occur in Germany has been established by Jänsch et al. (2013). Their habitat type classification was based on a German system (Riecken et al., 2009), which, at upper classification levels, is compatible with the EUNIS classification.

Expected application rates and their visualization in relation to earthworm abundance

To visualize the earthworm abundance in % of control in relation to the actual content of boscalid in soil, a so-called expected (pseudo) application rate (g boscalid/ha) was utilized. Pseudo application rates (kg/ha) were calculated based on the measured residue concentrations (mg/kg) (Table S2), the depth of the analyzed soil layer (cm), and an assumed soil bulk density of 1.5 g/cm³ (European and Mediterranean Plant Protection Organization, 2003). A worst-case field half-life (DT_{50}) for a temperature of 15 °C was derived from the results of three field dissipation trials conducted in Germany (Anonymous, 2002). The maximum DT_{50} at 20 °C determined for these trials (212 days) was recalculated to 15 °C (340.5 days) with the Arrhenius equation.

Arrhenius equation for calculating the worst-case field half-life (DT_{50})

$$DT_{50(act)} = DT_{50(ref)} \times Q_{10}^{\left(\frac{T_{ref} - T_{act}}{10}\right)}, \quad (1)$$

with

$DT_{50(act)}$ is the half-life at actual temperature (days), $DT_{50(ref)}$ the half-life at reference temperature (212 days), $Q_{10} = 2.58$, T_{act} the actual temperature (15 °C), T_{ref} the reference temperature (20 °C).

The expected application rates were calculated based on the worst-case DT_{50} at 15 °C, the pseudo application rate at the previous residue sampling (kg/ha) and the nominal application rate (kg/ha) for the following time points:

- day before an application;
- day of application;
- day before a residue sampling;
- day of earthworm sampling.

The following equations were used:

Calculation of pseudo application rates (kg/ha) based on residues measured in soil samplings (mg/kg)

$$MR = (RC_{(layer1)} \times d_{(layer1)} + RC_{(layer2)} \times d_{(layer2)}) \times bd_{soil} \times 100/1000, \quad (2)$$

with

MR is the measured residues as pseudo application rate (kg/ha), RC the measured residue concentration (mg/kg), *d* the depth of analyzed soil layer (cm), bd_{soil} the soil bulk density (1.5 g/cm³).

Calculation of expected residues (as pseudo application rate) at the day of application

$$ER_A = e^{-(\ln 2/DT_{50} \times T)} \times MR + AR. \quad (3)$$

Calculation of expected residues (as pseudo application rate) at the day before an application or residue sampling or the day of an earthworm sampling

$$ER_S = e^{-(\ln 2/DT_{50} \times t)} \times ER_A, \quad (4)$$

with

ER_A is the expected residues at the day of application (kg/ha), ER_S the expected residues at the day before application or residue sampling or day of earthworm sampling (kg/ha), DT_{50} the worst-case field half-life at reference temperature (15 °C) (days), *T* the time between previous sampling and application (days), *t* the time between application and day before following application or residue sampling or day of earthworm sampling (days), MR the measured residues (as pseudo application rate) at the previous sampling (kg/ha), AR the nominal application rate (kg/ha).

The data points for the calculated expected application rates and measured concentrations (transferred from mg boscalid/kg to g boscalid/ha) were connected to visualize the development of the expected application rate for each study program over time. Additionally, the rates calculated for each day of earthworm sampling were used to display the development of the earthworm abundance (% of control) in relation to the current expected application rate. This way, any apparent concentration-related response of the earthworm abundance or an overall decrease over the course of the study programs would become visible. Finally, linear regression analysis of the expected application rate (g boscalid/ha) and earthworm abundance (% of control) was performed. The slope of the regression and the coefficients of determination (r^2) were an indicator of the relationship between these two parameters.

All species and groups with comparable abundance data for all four study programs were evaluated (Table S3). Earthworm abundance and biomass data in principle follow the same trends. However, the variability of biomass data is even higher than for abundance and is often more biased by the presence of a few older individuals with high weight. Hence, earthworm abundance data are more robust and the evaluation was based on earthworm abundance only.

Effect size meta-analysis

The primary goal of this analysis was to provide a quantitative synthesis of the data gained in a cumulated 20 years of earthworm field studies, investigating the effects in a wide range of realistic and worst-case boscalid exposure scenarios on earthworm populations in arable soils. We have focused the analysis on:

1) The effect observed 12 months after the first application (MAA) for each study year and before the next crop cycle;
2) The effects of the toxic reference;
3) Nonstandardized effects sizes (changes in earthworm abundance) of the total earthworm community (adults and juveniles) and the three relevant functional (ecological) groups (Bottinelli et al., 2020; Bouché, 1977) represented by anecic, endogeic, and epigeic species (adults only). This was necessary because usually only adult individuals can be unambiguously identified to species level and consequently assigned to the corresponding functional group.

We excluded studies with no sufficient effect of the toxic reference (meaning exposure was not verified) and comparisons with an average of less than 5 earthworms/m² sampled in the control plots. The selection procedure and resulting sampling sizes are summarized in Figure S1, and the analyzed data are listed in Table S4.

Statistics meta-analysis. Study selection and all calculations were performed in R statistical environment version 3.5.1 using the "metafor" package (R Core Team, 2013; Viechtbauer, 2010). Absolute effect sizes (assuming a normal distribution of mean abundance values) and associated 95% confidence intervals (95% CI) were estimated using the "escalc" command in "metafor." To estimate overall effect sizes, we fitted a random mixed effects model with individual study years nested in the study program as a random factor for both overall abundance and the three functional earthworm groups. This model structure was chosen to account for the nonindependence of the individual trials (i.e., repeated measures of the same field) and model fitting was performed using restricted maximum-likelihood estimation (Viechtbauer, 2005). For all models we present: (1) Q values and associated *p* values as estimates for effect size heterogeneity (Higgins et al., 2003; Huedo-Medina et al., 2006) and (2) effect size estimates and associated CI.

RESULTS

Measured boscalid concentrations in soil and recovery rates

The mean measured boscalid concentrations (0–10 cm) ranged from 0.36 to 1.75 mg boscalid/kg (Table S2). Across all study programs, the calculated recovery rates of boscalid residues in soil (based on analytical measurements and the worst-case field half-life of 340.5 days at 15 °C; see above) for single applications ranged from 53.5% to 189.4%, with a mean overall recovery per study program of 98.6%–101.5%.

Study selection

Out of the 20 individual study years, two were excluded from the analyses because the toxic reference did not exhibit a statistically significant effect, casting doubt on the adequacy of the exposure regime at this time point (B1 2007 & B2 2007). On the functional group level, the epigeic group was excluded from the analyses because its abundance was less than 5 earthworms/m² in the majority of cases, rendering their population estimates unreliable (Figure S1).

Representativeness of the arable fields regarding soil parameters and earthworm community

Soil parameters. According to Bussian et al. (2005), the typical ranges of ecologically relevant soil parameters for agricultural soils in Germany are:

- Sand content: up to 90%;
- Clay content: up to about 50%;
- pH value ($CaCl_2$): 5 to >7;
- C_{org}: 0.5% to >8%.

Considering the physicochemical properties of the arable fields (Table 1), we deduce that they are typical for German crop sites. Also, comparing them to maps of the European land cover (Copernicus Land Service, 2020) and relevant soil properties, such as pH (Ballabio et al., 2019), C_{org}, or sand content (EC, 2005), this assessment may be extended to the European level. Additionally, they fit well in the ranges of soil parameters covered in both the German "RefeSol" (Bussian et al., 2005) and the European "EUROSOILS" (Gawlik et al., 2001) reference systems. Moreover, the arable fields did not feature conditions that might have masked potential adverse effects of boscalid such as exceptionally high organic matter or clay content to which boscalid ($\log P_{ow} = 2.96$) could have adsorbed, resulting in low bioavailability.

Earthworm communities. Based on the reference approach described above, the earthworm communities in the four study programs (presampling data; Table 3) were assessed regarding their correspondence with the expectation values for crop sites as described in Jänsch et al. (2013). In study programs A1 and A2, the number of species was higher than expected (7 instead of 3.3 ± 1.9) and the total adult abundance was 2–2.5 times the average but within the expected range (49.3 ± 86.2 individuals/m²). All three typical species (relative frequency >50%: *A. caliginosa*, *A. rosea*, *L. terrestris*) were present. Additionally, *A. chlorotica*, *Lumbricus castaneus*, *Lrubellus rubellus* (only A2), *Octolasion cyaneum* (not included in Jänsch et al., 2013), and *Octolasion tyrtaeum* (only A1) were identified. While the occurrence of epigeic species suggests that these arable fields represented relatively good conditions for earthworms, the presence of *A. chlorotica* and *O. cyaneum* indicates slightly more moist conditions than usually found at crop sites (Bouché, 1977; Lehmitz et al., 2014; Sims & Gerard, 1999). In study programs B1 and B2, both numbers of species and total adult abundance were within the expected range and all three typical species as well as *A. chlorotica* and *Aporrectodea longa* were present. Thus, the earthworm communities in both fields mostly corresponded to the expected values. Hence, the arable fields are regarded as representative for German and Central European crop sites and thus the results gained at these sites are meaningful in the context of the ERA for boscalid as currently required in the EU.

Holistic evaluation

In Figure 2, the development of the expected (pseudo) application rate (Panel A; in g boscalid/ha), total earthworm, and total juvenile earthworms abundance (Panels B and C; in % of control) over time in the four study programs are presented. Boscalid content in soil increased after multiple applications within the season. Between application seasons,

TABLE 3 Summary of the results of the random effect models analyzing the total, anecic, and endogeic communities for both the boscalid treatment and the respective toxic reference 12 months after application

			Change EW population (number)			Change EW population (%)				
Community	Treatment	*df*	Estimate	LB	UB	Estimate	LB	UB	Sig.	Con.
Total	Boscalid	17	−2.8	−19.3	13.6	−1.0	−6.7	4.7	NS	288
	Toxic reference	17	−89.4	−107.7	−71.1	−31.0	−37.4	−24.7	Sig	288
Anecic	Boscalid	17	−0.5	−5.1	3.9	−1.1	−11.6	8.9	NS	44
	Toxic reference	17	−29.26	−51.1	−7.43	−73.2	−127.8	−18.6	Sig	44
Endogeic	Boscalid	17	6.43	−1.4	14.3	5.5	−0.9	10.1	NS	116
	Toxic reference	17	−20.1	−34.9	−5.4	−17.3	−30.1	−4.7	Sig	116

Note: Here we present model degrees of freedom (*df*), estimates for earthworm (EW) changes, and associated 95% confidence interval (upper bound [UB] and lower bound [LB]) both in terms of total abundance and in percent changes compared to the control earthworm population. We indicate significant differences (Sig.) and the abundance of earthworms in the control plot (Con. [mean EW/m²]).

boscalid degraded to a soil concentration dependent on the application pattern.

A statistically significant reduction of total earthworm abundance compared to the control by 18.2% was observed at the second sampling date (4 MAA) in the third-year study of study program B2. However, since that study was excluded from the analyses due to insufficient effects in the toxic reference treatment, this data point is not contained in Figures 2 and 3. Overall, there was no linear relationship (slope of the regression line near zero) between expected application rate and total earthworm abundance (Figure 3A, $r^2 = 1.5 \times 10^{-3}$).

In addition to total earthworms, comparable data for all four study programs were available for adults of the dominant species *A. chlorotica*, *A. caliginosa*, *A. rosea,* and *L. terrestris* as well as the groups of endogeic adults, anecic adults, and total adults (Table S3). There were isolated statistically significant differences to the control, but for none of these was a consistent trend or an apparent relationship between expected application rate and abundance observed:

- *A. chlorotica*: Reductions compared to the control were more frequent during the first 3 years but never statistically significant; $r^2 = 3.2 \times 10^{-3}$;
- *A. caliginosa*: Reductions compared to the control were never statistically significant; $r^2 = 3.8 \times 10^{-4}$;
- *A. rosea*: Reductions compared to the control were more frequent during the first 3 years; a statistically significant reduction of *A. rosea* abundance compared to the control by 43.2% was observed at the third sampling date (6 MAA) in the third-year study of study program A1; $r^2 = 2.4 \times 10^{-2}$;
- *L. terrestris*: Reductions compared to the control were never statistically significant; $r^2 = 4.2 \times 10^{-2}$;
- Endogeic adults: Reductions compared to the control were never statistically significant; $r^2 = 3.6 \times 10^{-2}$;
- Anecic adults: Statistically significant reductions of anecic adult abundance compared to the control by 43.2% and 23.1% were observed in the third-year study at the fourth sampling date (12 MAA) of study program A2 and at the third sampling date (10 MAA) of study program B1, respectively; $r^2 = 3.6 \times 10^{-2}$;
- Adult earthworms: Reductions compared to the control were never statistically significant; $r^2 = 2.4 \times 10^{-3}$;
- Juvenile earthworms: Statistically significant reductions compared to the control were observed on five occasions:
  - o In the study program A2, at the first sampling date (1 MAA) in the third-year study by 28.7% and at the first sampling date in the fourth study year 6 weeks after the first annual application by 35.2%;
  - o In the first-year study of study program B1, at the first (1 MAA) and second (4 MAA) sampling date after the first annual application by 45.7% and 36.4%, respectively;
  - o In the third-year study of study program B2 at the second sampling date 4 MAA by 19.0%.

However, recovery of juvenile abundance was observed in subsequent samplings; Figure 3B, $r^2 = 5.4 \times 10^{-3}$.

Effect size meta-analysis

Total abundance. Combining all available data, there was no indication of lasting adverse effects on earthworm populations at 12 MAA (estimates −2.86; 95% CI: −19.3 to 13.6; see Figure 4 and Table 3) and no indication for effect size heterogeneity ($Q = 7.2$; $p = 0.9$) supporting the robustness of the obtained effect size estimates. In contrast, there were pronounced adverse effects of the toxic reference (estimates −89.4; 95% CI: −107.7 to −71.1; see Figure 4), but similarly to the treatment effect no indication for effect size heterogeneity ($Q = 18.6$; $p = 0.35$).

Anecic adults. Looking at the effects of treatment on anecic adults, in support of the results regarding total abundance, there was no indication of overall adverse effects at 12 MAA

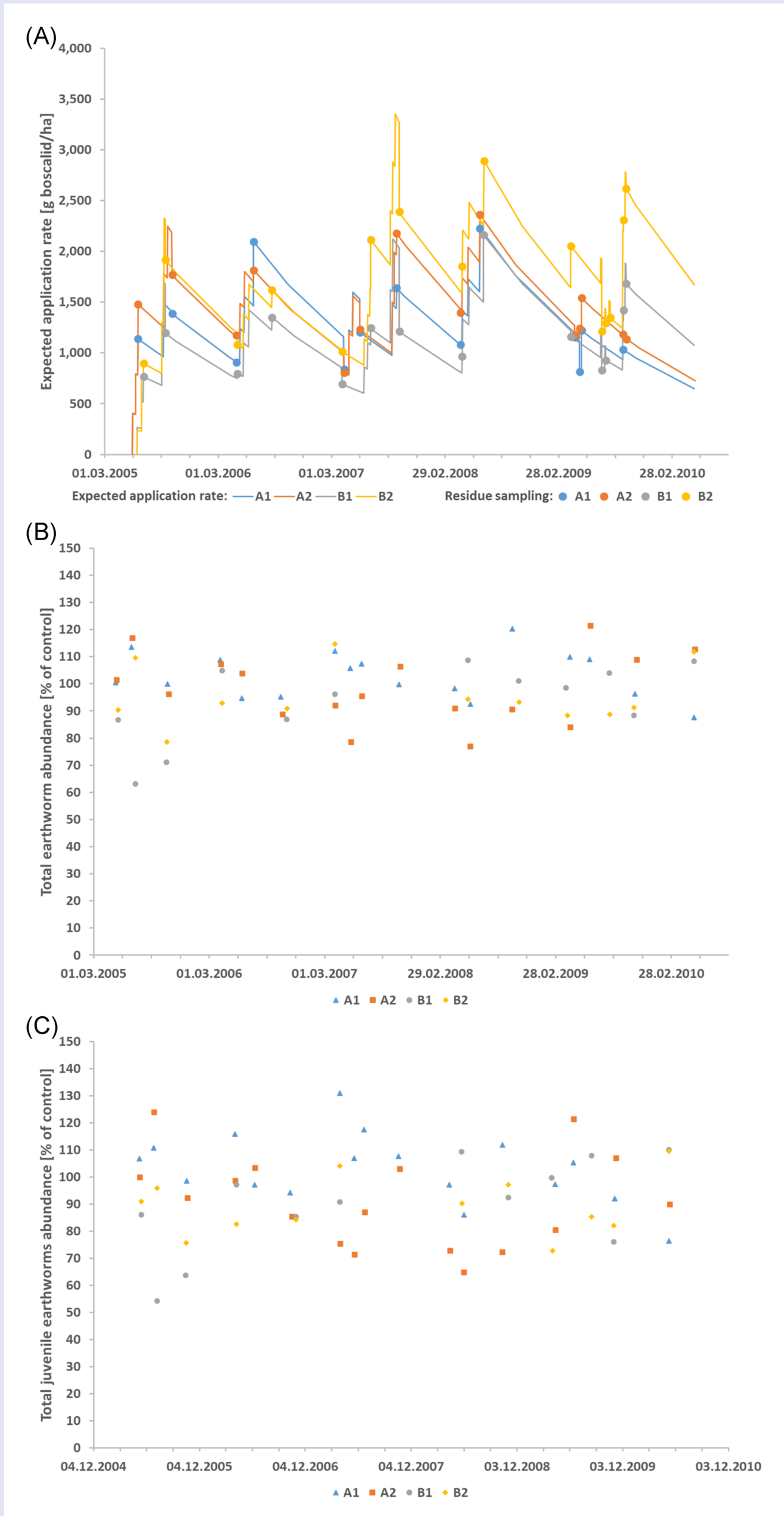

FIGURE 2 (A) Development of expected application rate (g boscalid/ha) over the 5 study years. Points for measured and calculated expected application rates are linearly interpolated. (B) Total earthworm and (C) total juvenile earthworms abundance (% of control) in the four study programs (study year 2007 for B1 and B2 excluded)

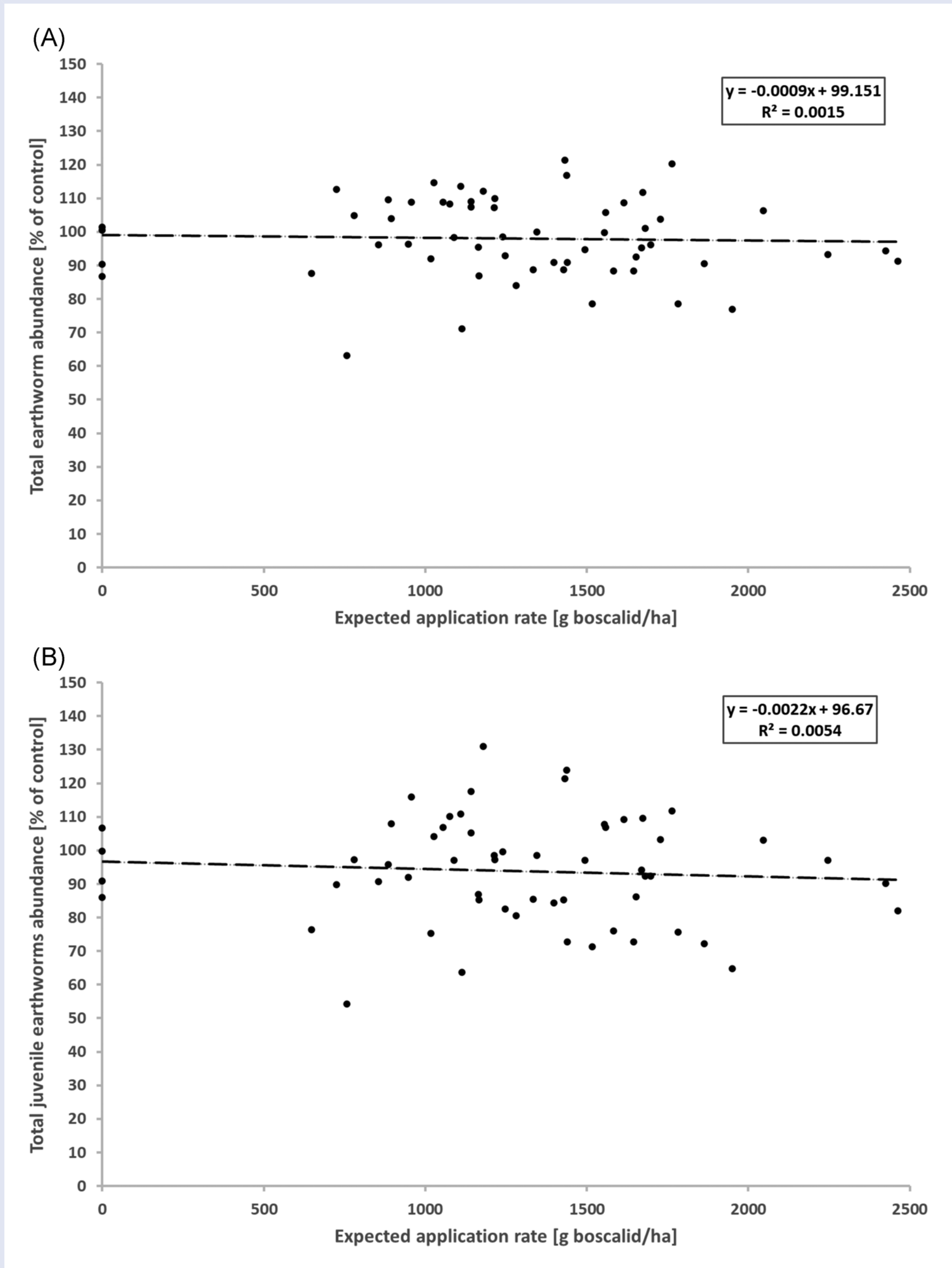

FIGURE 3 Relation between expected application rate (g boscalid/ha) and (A) total earthworm and (B) total juvenile earthworms abundance (% of control) (study year 2007 for B1 and B2 excluded)

(estimates −0.54; 95% CI: −5.03 to 3.95; see Figure S2 and Table 3). The observed effect size heterogeneity ($Q = 30.3$; $p = 0.02$; Figure S2) at the B site suggests higher variability in the populations' response compared to site A over the study program duration (see Figure S2). Regarding the effect of the toxic reference on anecic adults, we found a similar pattern compared to total abundance with pronounced adverse effects at 12 MAA (estimates −29.26; 95% CI: −51.1 to −7.43) and an indication of differences between programs ($Q = 261.02$; $p < 0.001$; see Figure S3).

Endogeic adults. The group of endogeic adults was by far the most abundant taxon (mean abundance 116 individuals/m²). At 12 MAA, there was no evidence of treatment-related effects (estimates 6.43; 95% CI: −1.4 to 14.3; see Figure S4 and Table 3) and no indication of effect size heterogeneity ($Q = 16.3$; $p = 0.5$). The toxic reference had clear negative effects on endogeic adults (estimates −20.1; 95% CI: −34.9 to −5.4; see Figure S5 and Table 3), and there was some indication that populations might react differently to the toxic reference ($Q = 26.6$; $p = 0.06$).

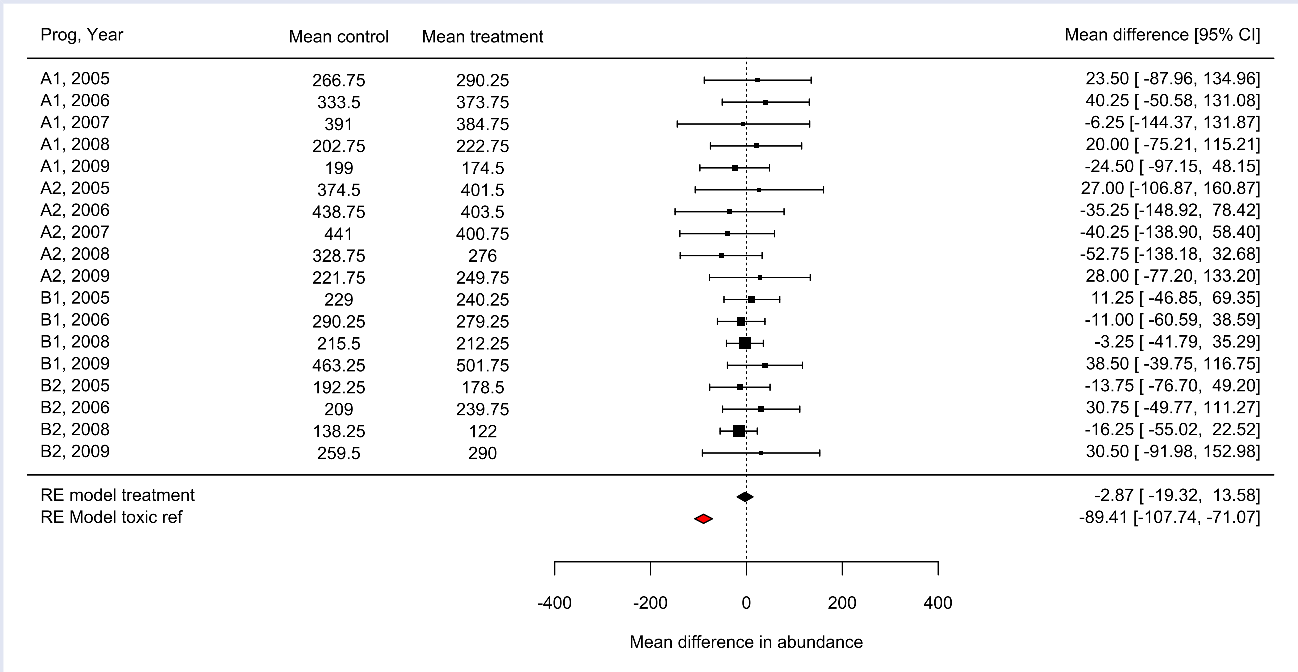

FIGURE 4 Forest plot summarizing the observed changes in earthworm populations at the end of each program year (individual years of the study programs are listed in chronological order) and the outcome of the nested random effects model for the treatment and toxic references across all years. We present the mean earthworm abundance (EW/m²) in the control plots (mean control) the boscalid treatment (mean treatment), the resulting effect sizes and associated 95% CI. In the case of the RE model treatment and RE Model toxic ref, the width of the rhombus indicates 95% CI. CI, confidence interval; toxic ref, toxic reference

DISCUSSION

Boscalid residues in Central European arable soils

Because of its physicochemical properties boscalid is expected to adsorb in the upper soil layers (Anonymous, 2002; Vallée et al., 2014) and is classified as having a moderate leaching potential with a calculated Groundwater Ubiquity Score leaching potential index of 2.68 (Pesticide Properties DataBase [PPDB], 2020). Several studies recently investigated the residues of currently used a.s., including boscalid in Central European arable soils, in particular Hvězdová et al. (2018) for the Czech Republic, Silva et al. (2019) for 11 EU member states, and Pelosi et al. (2021) for France. Boscalid was among the most frequently detected and highest concentrated a.s. Hvězdová et al. (2018) reported boscalid residues >0.01 mg/kg in the plough layer (0–25 cm) of 5 out of 75 soils with a mean concentration of 0.020 mg/kg and a maximum of 0.029 mg/kg. Silva et al. (2019) detected boscalid residues >0.01 mg/kg (= limit of quantification) in 87 out of 317 soils (27%) with a median content (0–15/20 cm soil depth) of 0.04 mg/kg and a maximum content of 0.41 mg/kg. Pelosi et al. (2021) detected boscalid residues in 155 out of 180 samples (86%) at concentrations (0–5 cm soil depth) of up to 1.212 mg/kg. The highest median concentration of 0.0047 mg/kg was measured in cereal crops. Because of differing sampling depth and unknown management practices (e.g., tillage, product, application rates, and times) these values are not easily comparable with each other and with the current study programs.

Across all four study programs and analytical samples, the measured residues in soil cores from 0 to 10 cm depth ranged from 0.36 to 1.75 mg boscalid/kg with a median of 0.80 mg/kg and a geometric mean of 0.83 mg/kg. The maximum calculated expected application rate was 3355 g boscalid/ha (Figure 2A), corresponding to 2.24 mg/kg assuming a bulk density of 1.5 g/m³ and a depth of 0–10 cm, 4.47 mg/kg when assuming worst case (i.e., 0–5 cm). These values are above both the measured residues reported by the above-mentioned authors as well as the laboratory chronic no observed effect concentration (NOEC) for Cantus® of 1.2 mg boscalid/kg (Anonymous, 2002). This emphasizes the relevance of the results from the four study programs for agricultural soils across Central Europe and confirms the worst-case character of the exposure situation achieved by using higher application rates than in actual agricultural practice.

Effects of the test substance boscalid

Because of the potential long-term risk to earthworms identified for the formulation Cantus® at the laboratory level, initially, two field tests with this formulation were conducted at grassland sites in Germany with an application of three times 0.3 kg a.s./ha and three times 0.6 kg a.s./ha, respectively. Based on the results of these field studies a maximum yearly application rate of 1 kg a.s./ha was defined by the European regulatory authorities (Anonymous, 2002; EC, 2008). In response, the four 5-year study programs described in this article were performed to address the

 DOI: 10.1002/ieam.4562

remaining uncertainties regarding the potential adverse effects in a realistic and long-term agricultural setting. The applied amounts in the four study programs were higher than the registered use rates (see Table 1) and the measured soil concentrations (Table S2) frequently exceeded the laboratory chronic NOEC for Cantus® of 1.2 mg boscalid/kg (Anonymous, 2002).

To our knowledge, no other data on the effects of boscalid on earthworm populations in the field have been published so far. The amount of data generated in the four earthworm field study programs is probably the highest ever collected for one a.s., at least when referring to studies performed according to the current field test guideline (ISO, 2014). The tests with Cantus® were conducted for up to 5 years, while published earthworm studies are usually stopped after 1 year, in rare cases after 2.

From the holistic evaluation of the four study programs we conclude there is no application rate-related adverse effect of the 5-year use of Cantus® on the development of the respective earthworm communities. The statistically significant reduction of total earthworm abundance by 18.2% during the third year of study program B2 is considered a random event likely to occur considering the high number of statistical comparisons in these study programs. Likewise, there were isolated statistically significant reductions of earthworm abundance for some species and groups at individual time points, but no consistent relationship to the Cantus® treatments was established. Hence, exposure to boscalid was apparently not the cause for differences in earthworm abundance between control and treatment plots. This is additionally confirmed by the slopes and coefficients of determination (r^2) from the linear regression analysis of the expected application rate (g boscalid/ha) and earthworm abundance (% of control) that are close to zero for the four dominant earthworm species (*A. chlorotica*, *A. caliginosa*, *A. rosea*, *L. terrestris*) as well as for the total number of earthworms, ecological groups, and age stages. Juveniles appear to react more sensitively (some statistically significant reductions) than adults but an indication of recovery in subsequent samplings has been observed in these cases. Juveniles are also the group whose abundance is subject to the strongest dynamics over the duration of the study programs, in particular due to the phenology of the earthworm population. Although earthworms are (semi-) continuous breeders, reproduction is among others dependent on soil temperature and moisture. In the northern hemisphere, most cocoons are produced in late spring and early summer and cocoon production may vary from year to year (Edwards & Bohlen, 1996). Therefore, random statistically significant differences to the control may be expected to occur more frequently than for other more stable age stages and aggregated groups. We hence conclude that the observed short-term fluctuations of earthworm abundance are rather a natural variation of the populations, for example, caused by the course of the weather influencing the micro-climate and thus local population dynamics throughout the study program duration (Edwards & Bohlen, 1996; Herwig et al., 2020), than a concentration-related response.

The results of the meta-analysis fully support the conclusions of the holistic evaluation. The overall estimated effect size suggests a mean and nonsignificant reduction of 3 earthworms/m² after 12-month exposure for each year of a program, which corresponds to about 1% of the total control population (see Figure 4 and Table 3) with an upper bound of the 95% CI of +4.7 and a lower bound of −6.7% change. Considering there is no indication of a time-dependent trend of the earthworm populations within a study program (see Figure 4), which would indicate additive effects of repeated boscalid application, it is unlikely that in-crop earthworm populations suffer adverse effects as a result of long-term boscalid exposure.

Our findings are in contrast to those observed in the field for other PPPs, in particular carbendazim and copper-based fungicides. The toxicity of carbendazim on earthworms in the field is well-documented (Römbke et al., 2004; van Gestel, 1992). Most recently, Römbke et al. (2020) performed an extensive pilot earthworm field study at a crop site in Germany with similar soil properties (pH: 7.2, C_{org}: 1.46%, texture: silt loam) as in the present study programs using six application rates of carbendazim (0.6–31.5 kg a.s./ha) resulting in median effective concentration (EC50) values for total earthworm abundance of 1.089 and 9.441 kg a.s./ha at 1 and 12 MAA, respectively. Jänsch and Römbke (2009) performed a literature review on the effects of copper on soil invertebrates. They found that strong adverse effects on earthworms in the field were frequently reported when the soil concentration of copper exceeded 50 mg/kg soil. The toxicity of these compounds on earthworms ultimately led to carbendazim being selected as a reference substance in both standard laboratory and field earthworm testing and a strong reduction of the permissible annual application rates of copper-based fungicides.

In the wider discussion of ERA, in particular in the EU, questions regarding the statistical power of standard test systems have gained prominence (EFSA, 2018; Römbke et al., 2020). The capability of test systems to detect an adverse effect is of fundamental importance to ensure the reliability of an ERA. While this question is certainly important, the currently discussed approaches such as post hoc power calculations, for example, the minimum detectable difference (Duquesne et al., 2020), suffer from clear but not immediately obvious shortcomings (Mair et al., 2020). While individual studies can suffer from low statistical power due to low replication or sample variability—effectively rendering species-based assessments difficult for most and impossible for rare species—the use of a meta-analytical framework could be applied to more abundant species as well and might provide a drastic increase in power to reliably detect adverse effects (Cohn & Becker, 2003; Ellis, 2010; Valentine et al., 2010). By combining such meta-analysis approaches with a CI-based characterization of statistical uncertainty as suggested by Mair et al. (2020), it becomes possible to directly and intuitively quantify the robustness of observed study results. Using a meta-analytical framework, it might be possible to achieve sufficient statistical power to ensure a conservative and robust risk assessment utilizing established and trusted test systems.

CONCLUSIONS

The soils of the four arable fields as well as their respective earthworm communities are regarded as typical and representative for German and Central European crop sites. Therefore, and given the worst-case character of the exposure situation, the results from the individual earthworm field studies as well as their holistic evaluation and meta-analyses are meaningful for the assessment of the field effects of boscalid on earthworms at all crop sites with comparable soil and site conditions. The outcome of every single study separately as well as the holistic approach and meta-analysis presented in this article are complementary to each other and offer the same overall result. We conclude there is no application rate-related effect of the 5-year use of Cantus® on the development of the earthworm communities in the four study programs. The performance of these study programs was therefore useful to address concerns about long-term in-crop adverse effects of boscalid on earthworms in the field and to refute the potential risk initially identified based on the data available during the ERA. Considering the immense efforts in conducting the numerous studies presented here, this approach can hardly serve as an example for a standard field-testing program in future risk assessment. However, this article illustrates possible ways to further refine risk assessment procedures, especially in cases where available data raised during the standard approach are not conclusive.

ACKNOWLEDGMENT

This work was funded by BASF SE, Ludwigshafen, Germany.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

DATA AVAILABILITY STATEMENT

Additional data and associated metadata are available through supplemental files.

SUPPORTING INFORMATION

FIGURE S1. Summary of study selection procedure and resulting sampling sizes.

FIGURE S2. Forest plot summarizing the observed changes in anecic adult abundance between control and treatment.

FIGURE S3. Forest plot summarizing the observed changes in anecic adult abundance between control and toxic reference.

FIGURE S4. Forest plot summarizing the observed changes in endogeic adult abundance between control and treatment.

FIGURE S5. Forest plot summarizing the observed changes in endogeic adult abundance between control and toxic reference.

TABLE S1. Application dates and rates.

TABLE S2. Sampling dates and measured residue concentrations.

TABLE S3. Sampling dates and abundance data for all species and groups with comparable data from all four study programs.

TABLE S4. Abundance data used for effect-size meta-analysis.

ORCID

Stephan Jänsch http://orcid.org/0000-0002-7281-8069
Jörg Römbke https://orcid.org/0000-0003-1341-634X
Tobias Pamminger http://orcid.org/0000-0003-1257-3829

REFERENCES

Anonymous. (2002, November 8). Nicobifen. Alternative common name proposed to ISO in August 2002: Boscalid. Volume 1. Report and Proposed Decision [Monograph]. Rapporteur Member State, Germany. Retrieved May 5, 2021, from: www.bvl.bund.de/SharedDocs/Downloads/04_Pflanzenschutzmittel/02_eu_berichte/Boscalid-DAR.html

Ballabio, C., Lugato, E., Fernández-Ugalde, O., Orgiazzi, A., Jones, A., Borrelli, P., Montanarella, L., & Panagos, A. (2019). Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression. *Geoderma*, *355*, 113912.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). Introduction to meta-analysis (450 pp.). John Wiley & Sons. https://onlinelibrary.wiley.com/doi/book/10.1002/9780470743386

Bottinelli, N., Hedde, M., Jouquet, P., & Capowiez, Y. (2020). An explicit definition of earthworm ecological categories—Marcel Bouché's triangle revisited. *Geoderma*, *372*, 114361. https://doi.org/10.1016/j.geoderma.2020.114361

Bouché, M. B. (1977). Stratéegies lombriciennes. In: Lohm U, Persson T, editors. Soil organisms as components of ecosystems. *Ecol Bull NFR* 25, 122–132.

Breure, A. M., Mulder, C., Römbke, J., & Ruf, A. (2005). Ecological classification and assessment concepts in soil protection. *Ecotoxicology and Environmental Safety*, *62*, 211–229.

Bundschuh, M., Zubrod, J. P., Seitz, F., Stang, C., & Schulz, R. (2011). Ecotoxicological evaluation of three tertiary wastewater treatment techniques via meta-analysis and feeding bioassays using *Gammarus fossarum*. *Journal of Hazardous Materials*, *192*(2), 772–778.

Bussian, B., Kördel, W., Kuhnt, G., Ohnesorge, S., & Weinfurtner, K. (2005). Das RefeSol-Projekt: Grundlagen eines deutschen Referenzbodensystems. *Wasser und Abfall*, *11*, 43–49.

Cohn, L. D., & Becker, B. J. (2003). How meta-analysis increases statistical power. *Psychological Methods*, *8*(3), 243–253. https://doi.apa.org/doiLanding?doi=10.1037%2F1082-989X.8.3.243

Copernicus Land Service. (2020). *Corine Land Cover (CLC) 2018, Version 2020_20u1*. Retrieved December 4, 2020, from: https://land.copernicus.eu/pan-european/corine-land-cover/clc2018

Cresswell, J. E. (2011). A meta-analysis of experiments testing the effects of a neonicotinoid insecticide (imidacloprid) on honey bees. *Ecotoxicology*, *20*(1), 149–157.

Davies, C. E., Moss, D., & Hill, M. O. (2004). *EUNIS Habitat Classification Revised 2004*. European Environment Agency. European Topic Centre on Nature Protection and Biodiversity. Retrieved July 21, 2020, from: https://www.eea.europa.eu/data-and-maps/data/eunis-habitat-classification/documentation/eunis-2004-report.pdf/at_download/file

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*(3), 177–188.

Duquesne, S., Alalouni, U., Gräff, T., Frische, T., Pieper, S., Egerer, S., Gergs, R., & Wogram, J. (2020). Better define beta-optimizing MDD (minimum detectable difference) when interpreting treatment-related effects of pesticides in semi-field and field studies. *Environmental Science and Pollution Research*, *27*(8), 8814–8821.

Edwards, C. A. (1983). Earthworm ecology in cultivated soils. In J. E. Satchell (Ed.), *Earthworm ecology—From Darwin to vermiculture* (pp. 123–137). Chapman & Hall.

Edwards, C. A., & Bohlen, P. R. (1996). *Biology and ecology of earthworms* (3rd ed., 426 pp.). Chapman & Hall.

Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results* (193 pp.). Cambridge University Press. https://doi.org/10.1017/CBO9780511761676

European Commission (EC). (2005). *Soil atlas of Europe. European Soil Bureau Network* (128 pp.). Office for Official Publications of the European Communities. https://esdac.jrc.ec.europa.eu/content/soil-atlas-europe

European Commission (EC). (2008). *Review report for the active substance boscalid*. Health & Consumer Protection Directorate-General. SANCO/3919/2007-rev. 5. Retrieved May 5, 2021, from: https://ec.europa.eu/food/plant/pesticides/eu-pesticides-database/active-substances/?event=as.details&as_id=472

European Commission (EC). (2009). Regulation (EC) No 1107/2009 of the European parliament and the council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. OJEU L, 309, 1–50.

European Commission (EC). (2013a). Commission Regulation No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation EC No. 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. 84 pp.

European Commission (EC). (2013b). Commission Regulation No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. 72 pp.

European Environment Agency (EEA). (2017). *Biogeographical regions in Europe*. Retrieved May 5, 2021, from: www.eea.europa.eu/data-and-maps/figures/biogeographical-regions-in-europe-2

European Food Safety Authority (EFSA). (2018). Guidance on uncertainty analysis in scientific assessments. *EFSA Journal*, *16*(1), 5123.

European Food Safety Authority (EFSA). (2019). Guidance on communication of uncertainty in scientific assessments. *EFSA Journal*, *17*(1), 5520.

European and Mediterranean Plant Protection Organization (EPPO). (2003). EPPO-Standards. Environmental risk assessment scheme for plant protection products. Chapter 8: Soil organisms and functions. *EPPO Bulletin*, *33*(2), 195–209. https://doi.org/10.1046/j.1365-2338.2003.00654.x

Garcia-Reyero, N., Habib, T., Pirooznia, M., Gust, K. A., Gong, P., Warner, C., Wilbanks, M., & Perkins, E. (2011). Conserved toxic responses across divergent phylogenetic lineages: A meta-analysis of the neurotoxic effects of RDX among multiple species using toxicogenomics. *Ecotoxicology*, *20* (3), 580–594.

Gawlik, B. M., Lamberty, A., Muntau, H., & Pauwels, J. (2001). EUROSOILS—A set of CRMs for comparability of soil-measurements. *Fresenius Journal of Analytical Chemistry*, *370*, 220–223. https://doi.org/10.1007/s002160100782

Herwig, N., Felgentreu, D., & Hommel, B. (2020). Auswirkungen von natürlichen Standortbedingungen und ackerbaulichen Maßnahmen auf Bodenorganismen—Erhebungen in den Langzeitversuchen des Julius Kühn-Instituts in Dahnsdorf (Hoher Fläming, Land Brandenburg). *Journal fur Kulturpflanzen*, *72*(7), 327–337.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*, 557–560. https://doi.org/10.1136/bmj.327.7414.557

Huedo-Medina, T. B., Sánchez-Meca, J., Botella, J., & Marín-Martínez, F. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I² index? *Psychological Methods*, *11*(2), 193–206.

Hvězdová, M., Kosubová, P., Košíková, M., Scherr, K. E., Šimek, Z., Brodský, L., Šudoma, M., Škulcová, L., Sáňka, M., Svobodová, M., Krkošková, L., Vašíčková, J., Neuwirthová, N., Bielská, L., & Hofman, J. (2018). Currently and recently used pesticides in Central European arable soils. *Science of the Total Environment*, *613–614*, 361–370.

International Organization for Standardization (ISO). (2014). *No. 11268-3. Soil quality—Effects of pollutants on earthworms, Part 3: Guidance on the determination of effects in field situations*.

International Organization for Standardization (ISO). (2018). *No. 23611-1. Soil quality—Sampling of soil invertebrates—Part 1: Hand-sorting and extraction of earthworms.*

Jänsch, S. & Römbke, J. (2009). Einsatz von Kupfer als Pflanzenschutzmittel-Wirkstoff: Ökologische Auswirkungen der Akkumulation von Kupfer im Boden. UBA-Texte 10/09. 70 pp. https://www.umweltbundesamt.de/publikationen/einsatz-von-kupfer-als-pflanzenschutzmittel

Jänsch, S., Steffens, L., Höfer, H., Horak, F., Roß-Nickoll, M., Russell, D., Toschki, A., & Römbke, J. (2013). State of knowledge of earthworm communities in German soils as a basis for biological soil quality assessment. *Soil Organisms*, *85*, 215–232.

Kula, C., Heimbach, F., Riepert, F., & Römbke, J. (2006). Technical recommendations for the update of the ISO earthworm field test guideline (ISO 11268-3). *Journal of Soils and Sediments*, *6*, 182–186.

Lavoie, R. A., Jardine, T. D., Chumchal, M. M., Kidd, K. A., & Campbell, L. M. (2013). Biomagnification of mercury in aquatic food webs: A worldwide meta-analysis. *Environmental Science and Technology*, *47*(23), 13385–13394.

Lee, K. E. (1985). *Earthworms—Their ecology and relationships with soils and land use* (411 pp.). Academic Press.

Lehmitz, R., Römbke, J., Jänsch, S., Krück, S., Beylich, A., & Graefe, U. (2014). Checklist of earthworms (Oligochaeta: Lumbricidae) from Germany. *Zootaxa*, *3866*, 221–245.

Lofstedt R., & Bouder F. (2021). Evidence-based uncertainty analysis: What should we now do in Europe? A view point. *Journal of Risk Research*, *24*(5), 521–540. http://doi.org/10.1080/13669877.2017.1316763

Mair, M. M., Kattwinkel, M., Jakoby, O., & Hartig, F. (2020). The MDD concept for establishing trust in non-significant results—A critical review. *Environmental Toxicology and Chemistry*, *39*(11), 2109–2123.

Organisation for Economic Co-operation and Development (OECD). (1998). OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 1. OECD Principles on Good Laboratory Practice (as revised in 1997). Paris (FR): Environment Directorate, OECD. OECD ENV/MC/CHEM(98)17.

Pelosi, C., Bertrand, C., Daniele, G., Coeurdassier, M., Benoit, P., N´elieu, S., Lafay, F., Bretagnolle, V., Gaba, S., Vulliet, E., & Fritsch, C. (2021). Residues of currently used pesticides in soils and earthworms: A silent threat? *Agriculture, Ecosystems & Environment*, *305*, 107167.

Pelosi, C., Joimel, S., & Makowski, D. (2013). Searching for a more sensitive earthworm species to be used in pesticide homologation tests—A meta-analysis. *Chemosphere*, *90*(3), 895–900.

Pesticide Properties DataBase (PPDP). (2020). Boscalid (Ref: BAS 510F). Retrieved December 4, 2020, from: http://sitem.herts.ac.uk/aeru/ppdb/en/Reports/86.htm#none

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved May 5, 2021, from: www.R-project.org

Riecken, U., Finck, P., Raths, U., Schröder, E., & Ssymank, A. (2009). *German Red Data Book on endangered habitats (short version July 2009)*. German Federal Agency for Nature Protection.

Römbke, J., & Breure, A. M. (2005). SPECIAL ISSUE: Ecological soil quality: Classification and assessment. *Ecotoxicology and Environmental Safety*, *62*(2), 185–308. https://www.sciencedirect.com/journal/ecotoxicology-and-environmental-safety/vol/62/issue/2

Römbke, J., Förster, B., Jänsch, S., Kaiser, F., Scheffczyk, A., Roß-Nickoll, M., Daniels, B., Ottermanns, R. & Scholz-Starke, B. (2020). Necessary adaptations for a harmonized field-testing procedure and risk assessment of earthworms (terrestrial) (UBA-Texte 193/2020). 118 pp. https://www.umweltbundesamt.de/en/publikationen/necessary-adaptations-for-a-harmonized-fieldtesting

Römbke, J., Gardi, C., Creamer, R., & Miko, L. (2016). Soil biodiversity data: Actual and potential use in European and national legislation. *Applied Soil Ecology*, *97*, 125–133.

Römbke, J., Van Gestel, C. A. M., Jones, S. E., Koolhaas, J. E., Rodrigues, J. M. L., & Moser, T. (2004). Ring-testing and field-validation of a Terrestrial Model Ecosystem (TME)—An instrument for testing potentially harmful substances: effects of carbendazim on earthworms. *Ecotoxicology*, *13*(1–2), 105–118. https://doi.org/10.1023/b:ectx.0000012408.58017.08

Scholz, S., Schreiber, R., Armitage, J., Mayer, P., Escher, B. I., Lidzba, A., Léonard, M., & Altenburger, R. (2018). Meta-analysis of fish early life stage tests—Association of toxic ratios and acute-to-chronic ratios with modes of action. *Environmental Toxicology and Chemistry*, *37*(4), 955–969.

Silva, V., Mol, H. G. J., Zomer, P., Tienstra, M., Ritsema, C. J., & Geissen, V. (2019). Pesticide residues in European agricultural soils—A hidden reality unfolded. *Science of the Total Environment*, *653*, 1532–1545.

Sims, R. W., & Gerard, B. M. (1999). *Earthworms. Synopses of the British Fauna (New Series) No. 31* (171 pp.). E. J. Brill/W. Backhuys.

US Environmental Protection Agency (US EPA). (1998). *Guidelines for ecological risk assessment* (US EPA/630/R-95/002F). US Environmental Protection Agency, Washington, DC, USA.

Uvarov, A. V. (2009). Inter- and intraspecific interactions in lumbricid earthworms: Their role for earthworm performance and ecosystem functioning. *Pedobiologia*, *53*, 1–27.

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, *35*(2), 215–247. https://doi.org/10.3102%2F1076998609346961

Vallée, R., Dousset, S., Billet, D., & Benoit, M. (2014). Sorption of selected pesticides on soils, sediment and straw from a constructed agricultural drainage ditch or pond. *Environmental Science and Pollution Research*, *21*, 4895–4905.

Van Gestel, C. A. M. (1992). Validation of earthworm toxicity tests by comparison with field studies: A review of benomyl, carbendazim, carbofuran, and carbaryl. *Ecotoxicology and Environmental Safety*, *23*, 221–236.

Van Groenigen, J. W., Lubbers, I. M., Vos, H. M. J., Brown, G. G., De Deyn, G. B., & Van Groenigen, K. J. (2014). Earthworms increase plant production: A meta-analysis. *Scientific Reports*, *4*, 6365.

Veltman, K., Huijbregts, M. A. J., Hamers, T., Wijnhoven, S., & Hendriks, A. J. (2007). Cadmium accumulation in herbivorous and carnivorous small mammals: Meta-analysis of field data and validation of the bioaccumulation model optimal modeling for ecotoxicological applications. *Environmental Toxicology and Chemistry*, *26*(7), 1488–1496.

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Wolfram, J., Stehle, S., Bub, S., Petschick, L. L., & Schulz, R. (2018). Meta-analysis of insecticides in United States surface waters: Status and future implications. *Environmental Science and Technology*, *52*(24), 14452–14460.

 DOI: 10.1002/ieam.4562