Science of the Total Environment 912 (2024) 169096

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Reproductive toxicity in birds predicted by physiologically-based kinetics and bioenergetics modelling

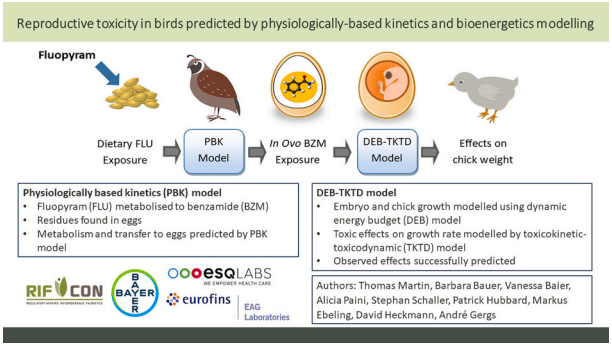
Thomas Martin [a,*], Barbara Bauer [a], Vanessa Baier [b], Alicia Paini [b], Stephan Schaller [b], Patrick Hubbard [c], Markus Ebeling [d], David Heckmann [d], André Gergs [d]

[a] Rifcon GmbH, Goldbeckstraße 13, 69493 Hirschberg an der Bergstraße, Germany
[b] esqLABS GmbH, Hambierich 34, 26683 Saterland, Germany
[c] Eurofins EAG Agroscience, LLC, Easton, MD, USA
[d] Bayer AG, Crop Science Division, Monheim, Germany

HIGHLIGHTS

- Egg injection studies allow further investigation of *in-ovo* toxic effects in birds.
- Resulting data are suitable to calibrate TKTD models based on DEB theory.
- Physiology based kinetics models can predict *in ovo* exposure from hens' intake.
- Carry-over effects observed after hatching despite no further exposure.
- Effects at hatching and 14 days post-hatch successfully predicted by DEB-TKTD model.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Rafael Mateo Soria

Keywords:
Dynamic energy budget
TKTD
Pesticide
Environmental risk assessment
Physiologically based kinetics
Avian

ABSTRACT

Effects on the growth and reproduction of birds are important endpoints in the environmental risk assessment (ERA) of pesticides. Toxicokinetic-toxicodynamic models based on dynamic energy budget theory (DEB) are promising tools to predict these effects mechanistically and make extrapolations relevant to ERA. However, before DEB-TKTD models are accepted as part of ERA for birds, ecotoxicological case studies are required so that stakeholders can assess their capabilities.

We present such a case-study, modelling the effects of the fluopyram metabolite benzamide on the northern bobwhite quail (*Colinus virginianus*). We parametrised a DEB-TKTD model for the embryo stage on the basis of an egg injection study, designed to provide data for model development. We found that information on various endpoints, such as survival, growth, and yolk utilisation were needed to clearly distinguish between the performance of model variants with different TKTD assumptions. The calibration data were best explained when it was assumed that chemical uptake occurs via the yolk and that benzamide places stress on energy assimilation and mobilisation.

To be able to bridge from the in vitro tests to real-life exposure, we developed a physiologically-based toxicokinetic (PBK) model for the quail and used it to predict benzamide exposure inside the eggs based on dietary

* Corresponding author.
E-mail address: Thomas.Martin@rifcon.de (T. Martin).

https://doi.org/10.1016/j.scitotenv.2023.169096
Received 28 June 2023; Received in revised form 25 November 2023; Accepted 2 December 2023
Available online 12 December 2023

exposure in a standard reproductive toxicity study. We then combined the standard DEB model with the TKTD module calibrated to the egg injection studies and used it to predict effects on hatchling and 14-day chick weight based on the exposure predicted by the PBK model. Observed weight reductions, relative to controls, were accurately predicted. Thus, we demonstrate that DEB-TKTD models, in combination with suitable experimental data and, if necessary, with an exposure model, can be used in bird ERA to predict chemical effects on reproduction.

1. Introduction

The use of pesticides is an important part of modern agriculture, combatting large losses to pest species and so contributing to food security (Popp et al., 2013). However, like all man-made chemicals that are released into the environment, pesticides have the potential to cause environmental damage and must be used safely. Environmental risk assessment (ERA) for pesticides is therefore vital to ensure that pesticides can be used effectively without harm to the environment.

Conventionally, environmental risk is extrapolated from dose-response measures such as the No Observed Effect Level (NOEL), the highest dose level at which no effects on any endpoints have been observed in laboratory studies, or the benchmark dose (BMD), the dose at which an X% adverse effect would be expected. However, summary statistics like the NOEL and BMD do not consider the temporal dynamics of exposure and so they only apply to a specific exposure duration. Additionally, their accuracy is limited by the number of treatments included in a study, especially the NOEL. Realistic exposure is dynamic and often much shorter than standard toxicity studies, in which exposure may be sustained at high levels for long periods. Therefore, measures like the NOEL and even BMDs are of limited suitability to the prediction of environmental risk under realistic, spatiotemporally variable conditions (Jager, 2012). To account for the mismatch between lab and field, European ERA uses safety factors to ensure that decisions are protective. However, greater realism and environmental relevance in the methods used to extrapolate from lab to field are clearly desirable (EFSA PPR et al., 2009; Schneeweiss et al., 2022).

One way to achieve this is through the use of mechanistic effect models in pesticide ERA. Mechanistic effect models simulate the mechanisms by which chemicals affect biological systems whether at the level of the individual, population or community (Grimm and Martin, 2013). The use of such models in ERA has been suggested for some time, for example to predict effects at the population level based on individual level data (Kramer et al., 2011). In recent years, one methodology that has gained prominence is the use of toxicokinetic-toxicodynamic models (TKTD) based on dynamic energy budget (DEB) theory. DEB simulates the growth, maturation, and reproduction of individual organisms based on energy acquisition and allocation over time (Kooijman, 2009). Once the physiological DEB model is calibrated to control data for a species, the effects of pesticide exposure can then be simulated through the addition of a TKTD module. TKTD modelling simulates the internal damage resulting from chemical exposure over time and the resulting changes to biological processes (Ashauer et al., 2011). These combined DEB-TKTD (also termed DEBtox) models were originally developed to gain greater insight into toxicological data but also have the ability to extrapolate beyond tested scenarios and so their potential for use in ERA is clear (Sherborne and Galic, 2020).

In contrast to conventional ERA methods, DEB-TKTD models do account for exposure dynamics and so are particularly well suited to making extrapolations to multiple exposure scenarios. For example, a DEB-TKTD model developed with laboratory data may be used to predict the effects of realistic exposure in the field, or predict the multiplication factor that would need to be applied to a realistic exposure profile to result in an X% effect (the exposure profile specific EPx (EFSA PPR et al., 2018)). This potential use in ERA has been recognised by EFSA in the recent ERA guidance for birds and mammals (EFSA, 2023). DEB-TKTD modelling can be especially beneficial to ERA for birds and mammals, as the modelled sublethal endpoints are of particular relevance for these taxa. Protection goals state that no visible mortality should occur as a result of pesticide application, so high application rates that could lead to any acute toxicity are not acceptable. This means that the risk of sublethal effects which may result from longer-term, low-level exposure is key to decision making (EFSA, 2023). DEB-TKTD models simulate effects on growth, reproduction, and the timing of events such as reaching sexual maturity, all of which may be considered the critical endpoint for reproductive toxicity in ERA. Despite its clear relevance, DEB-TKTD modelling has not yet been adopted into ERA protocols and is currently limited to research purposes. A key reason is the lack of case studies in which the effects of pesticides have been modelled for bird species (EFSA, 2023).

In this study, we provide such a case study by modelling the reproductive effects of the fungicide fluopyram (hereafter abbreviated to FLU) on the northern bobwhite quail (*Colinus virginianus*). Regulatory reproduction studies (Bayer, 2008a, 2008b) on this species showed that dietary intake of FLU was associated with reductions in offspring weight at hatching and 14 days post-hatch. Observed effects on 14-day old chicks were greater than those on hatchlings, despite chicks having no dietary exposure to the test compound themselves. This observation suggests that embryo growth was impacted by chemical exposure inside the egg, with some carry-over effects after hatching. As part of the regulatory framework, a metabolism study was also carried out, in which pesticide residues in chicken (*Gallus gallus domesticus*) eggs were measured to assess the potential dietary exposure of humans. Results showed that FLU is largely metabolised to 2-(trifluoromethyl)benzamide (henceforth abbreviated to BZM) and that this metabolite accounts for almost all chemical recovery from eggs (Bayer, 2008c; FAO and WHO, 2011). Based on this finding, we conducted an experiment based on Farhat et al. (2020) in which fertilised eggs were injected with benzamide, and several developmental endpoints were observed during the incubation period. We then used the resulting data to calibrate a DEB-TKTD model for embryo growth under BZM exposure.

To be able to use that model to predict effects in a reproduction study where adults were exposed to FLU via their diet, it was necessary to estimate the resulting embryonal exposure by calculating BZM concentrations in eggs arising from maternal transfer. For this purpose, a physiologically based kinetic (PBK) model for birds by Baier et al. (2022) was extended to predict FLU concentrations in quail. PBK models simulate the absorption, distribution, metabolism and excretion of a substance within the body, represented as a network of interconnected compartments. The quail PBK model developed here was used to simulate the metabolism of FLU from the diet to BZM and the transfer of the latter into eggs. This allowed predictions of BZM concentrations in eggs associated with each dietary dose level in the regulatory reproduction study. For validation, the DEB-TKTD model, calibrated to embryo data, was then used to predict embryo growth and - with its TKTD module coupled to an existing DEB model for bobwhite quail (Marn et al., 2022) - the growth of chicks for 14 days after hatching, i.e. the most sensitive endpoint derived from the quail reproduction studies.

The study had two aims. First, to test whether DEB-TKTD modelling could accurately capture the effects of BZM concentration on embryo growth. Second, to test whether the resulting TKTD parameters, together with the quail PBK model predicting embryonal exposure based on maternal dietary intake, could be used to predict effects in a different life stage, i.e. 14 days after hatching. Successfully modelling reproductive

effects will provide a valuable DEB-TKTD modelling case study for a bird species. Moreover, it will demonstrate the utility of embryo data for modelling toxic effects at multiple life stages in cases where chemical residues in the egg drive the observed effects.

2. Methods

2.1. Study species

The focal species of this study was the northern bobwhite quail, *Colinus virginianus*. The species belongs to the order *Galliformes* and is native to North America, with introduced populations in Europe and Asia. Bobwhite quail are recommended for testing the toxicity of substances on birds (e.g. avian acute oral toxicity test OECD 223 (2016) or avian reproduction test OECD 206 (1984)), so they are an important test species providing input for ERA.

2.2. Egg injection study

To generate calibration data, we conducted a study in which developing embryos were exposed to known concentrations of the FLU metabolite BZM. In this study, BZM (provided by Bayer, 99.4 % purity) was injected into the air cells of eggs prior to incubation, with dose levels quantified as mg BZM $\times$ kg^{-1} egg weight (as per Farhat et al., 2020). The injection consisted of BZM dissolved in dimethyl sulfoxide (DMSO) and was serially diluted such that the injection rate of the whole solution was constant across treatments (1 μL $\times$ g^{-1} egg weight). Three control treatments were included: 1) a negative control in which eggs were injected with 1 μL $\times$ g^{-1} egg weight nanopure water; 2) a positive control in which eggs were injected with 41.1 mg chlorpyrifos $\times$ kg^{-1} egg weight (this chemical was shown to affect embryo mass in a previous egg injection study); 3) a solvent control in which eggs were injected with 1 μL $\times$ g^{-1} egg weight DMSO. The negative control was included to determine whether DMSO impacted the measured endpoints and the positive control to confirm effectiveness of the dosing method. The solvent control was used for direct comparison to treatment data and for modelling.

Eggs were collected from a breeder flock and stored at an average temperature of 10–16 °C and relative humidity of 50–90 % for no more than eight days before incubation. Prior to injection, eggs were candled to exclude any with shell cracks or internal abnormalities from the study. Eggs that were not excluded were then weighed to the nearest 0.01 g and fumigated with formaldehyde gas to reduce the chance of pathogen contamination. Eggs were then injected and incubated at 37.4 °C and 55 % humidity for up to 22 days. Ten eggs per treatment were assigned for sampling at each time point. At day seven, all eggs were candled and any 'clear' (no shadow of embryo visible) eggs were removed and examined to determine fertility. At each sampling time point, the assigned eggs were candled to determine viability and then euthanised with carbon dioxide. The embryos were then extracted and their wet weight was recorded. From day 14 onward, the length of the tarsus was also measured (this could not be measured earlier in development). On day 22, embryos were weighed with and without the yolk sac.

The study was conducted in two stages. The first stage comprised fewer treatment groups and sampling timepoints but covered a broader range of dose levels. The results of the stage one study were then used to inform the design of stage two. Stage one of the study comprised controls and three treatments: 100, 225, and 500 mg $\times$ kg^{-1}. For egg viability and embryo weight, there were three observation points, on days 7, 14, and 22. Tarsus length was measured only on days 14 and 22. Positive and negative control groups were sampled only on days 7 and 22 for viability and day 22 only for embryo weight and tarsus length. Due to low survival in the 500 mg $\times$ kg^{-1} dose group, stage two focused on lower doses, comprising treatments of 10, 50, 110, and 225 mg $\times$ kg^{-1}, as well as positive controls and solvent controls. Stage one data showed no significant effects of DMSO injection vs negative controls on any endpoint, so the negative control treatment was not included in stage two. Viability was measured on day 7 for all treatment groups and on days 14, 16, 18,19, 20, 21, and 22 for all treatment groups except the positive controls. Embryo weight was measured at day 7 for the solvent control treatment only, embryo weight and tarsus length were recorded on days 14, 16, 18, 19, 20, 21, and 22 in all other treatments except the positive control treatment, which followed the same observation schedule as in stage one.

2.3. Data for modelling

2.3.1. Calibration data

Calibration data came from the egg injection study, described above. Mean data per treatment from all measured endpoints (survival, tarsus length, embryo weight and yolk weight) were used for model calibration. Individual data were unsuitable for model fitting as each individual was only sampled once. Data for the positive and negative control treatments were not used as they were only observed at the end of the study. Moreover, no significant effects on any endpoint (relative to negative controls) were detected in the solvent control treatment. The solvent control treatment (for which data were collected at all time points), was included, and is hereafter referred to simply as the control treatment. Egg viability data were converted to survival percentage by expressing the number of viable eggs and a percentage of the number of fertile eggs assigned for observation at each time point. The survival data were then censored so that survival probability could not increase over time (censored data included only the lowest recorded survival proportion up to each timepoint point in each treatment). On day 22, embryo weight was measured with and without the yolk sac. The final yolk weight was calculated as the difference between the two measurements. At all other time points embryo weight was measured without the yolk sac.

2.3.2. Validation data

The validation data came from two chronic reproduction studies (Bayer, 2008b, 2008a) with *C. virginianus*, conducted according to OECD protocol 206 (OECD, 1984), in which the parental birds were exposed to FLU via the diet, and their reproductive performance was assessed over ~10 weeks of egg laying. Chick weight was measured at hatch and at 14 days after hatching but concentrations of BZM in the eggs at laying were not measured. Thus, exposure concentrations inside the eggs were predicted by the PBK model described below. Chicks received untreated food after hatching so effects observed on day 14 post-hatch were assumed to result from damage during the embryonal phase. While survival was measured, these data were not used in model validation as the aim of this study was to develop a model for representing sublethal effects. Additionally, the survival data were not analogous between the calibration and validation datasets, since the egg injection study was terminated before hatching. In the reproduction studies used for validation, eggs which failed to hatch were counted towards mortality although the embryos may have been alive up to day 22.

2.4. Model implementation

The PBK analyses were performed using PBK software MoBi (version 9.1). R (distribution 4.0.5) with RStudio (Version 2022.02.0) were used in the analysis for pre- and post-processing of data and model output.

All DEB-TKTD model calculations were performed using the MATLAB software vR2021a and a set of MATLAB scripts included in the Bring-Your-Own-Model (BYOM, www.debtox.info/byom.html) framework v6 beta 8, customized for the model and data presented here.

2.4.1. PBK model

The generic physiologically-based kinetic (PBK) model developed by Baier and colleagues (Baier et al., 2022) was used to assess

concentrations of FLU and its metabolite BZM in eggs. The generic avian PBK model was developed for male and female birds using PK-Sim and MoBi from the Open Systems Pharmacology Suite (OSPS, https://www.open-systems-pharmacology.org/). The PBK model includes an ovulation model (egg development) to predict concentrations of chemicals in eggs from dietary exposure. The generic model was previously parametrised for chicken (*Gallus gallus*), bobwhite quail (*Colinus virginianus*), and mallard duck (*Anas platyrhynchos*), see Scanes et al. (2022a, 2022b), and was tested with nine chemicals. Time-concentration profiles of chemicals reaching tissues and egg compartments were simulated and compared to in vivo data for model evaluation (Baier et al., 2022).

The generic avian PBK model was adapted to predict concentrations of FLU and its metabolite BZM in chicken tissues and eggs. The adaptation consisted of structural changes and re-parameterisation based on in vitro (physico-chemical properties of FLU and BZM) and in vivo (residual measurements following oral intake) chicken data (Bayer, 2008c; "Fluopyram: Feeding Study Laying Hens (Gallus Gallus Domesticus)", 2008). This model was then used to create the FLU quail PBK model by replacing species-specific physiological parameters of the chicken for the quail, as well as adapting the relative rate of FLU to BZM metabolism, derived from in vitro data (Bayer, 2022, 2022b). The nominal concentration in the in vitro experiments was converted using the 'virtual cell based assay' model (Comenges et al., 2017; Proença et al., 2019).

The PBK model was then used to predict egg concentrations in the two quail reproduction studies used for validation (see below). More information on the PBK model and the approach can be found in the Supporting Information.

2.4.2. The physiological DEB model

The growth of the embryo during incubation as well as the first 14 days of growth post-hatch were both modelled according to the Dynamic Energy Budget (DEB) framework, standard notation for model parameters is used throughout (Kooijman, 2009; van der Meer, 2006). Different model parameters were used for the embryo and chick stages. DEB describes the acquisition and allocation of energy by organisms as a set of differential equations and divides an organism's mass into 'reserve' and 'structure'. When energy is assimilated from food, it enters the reserve ('assimilation flux'). Energy from the reserve is then utilised by the organism to perform basic functions and to grow and/or reproduce ('mobilisation flux'). Fixed proportions of the mobilisation flux are allocated to maintenance & growth and to maturation & reproduction. The maintenance flux takes priority over growth and increases as the animal grows such that, at a given feeding rate, the animal eventually reaches a maximum size, at which no energy is available for further growth (Fig. 1, equations in Supporting Information). The model framework is well established and has been subject to a systematic sensitivity analysis, demonstrating the relationships between parameter values and model outputs, as well as the influence on changes to data on the values arrived at by model fitting (Matyja, 2023).

Because embryos do not feed, modelling growth inside the egg represents a special case of the standard DEB model. Rather than describing the uptake of energy from food, the model included the conversion of an already large reserve (mostly yolk) into structure (the embryo). This followed the approach used by Kooijman (2015) to model the development of the Australian Freshwater Crocodile (*Crocodylus johnstoni*) embryos in the egg. The model of chick growth used the standard DEB model and parameters for *C. virginianus* (Marn et al., 2022) from the 'Add my Pet' (AmP) database (Marques et al., 2018).

2.4.3. DEB-TKTD model

The physiological DEB model was combined with toxicokinetic-toxicodynamic modules for lethal and sublethal effects (Fig. 1, equations in Supporting Information). A 'damage module' combined toxicokinetics and the first stage of toxicodynamics. This module simulated the accumulation and repair of 'damage' resulting from BZM exposure and was the basis for modelling lethal and sublethal effects. The sublethal effects module then simulated stress placed on biological processes by the level of damage over time. Meanwhile, the effects of damage on survival probability were simulated using the reduced general unified

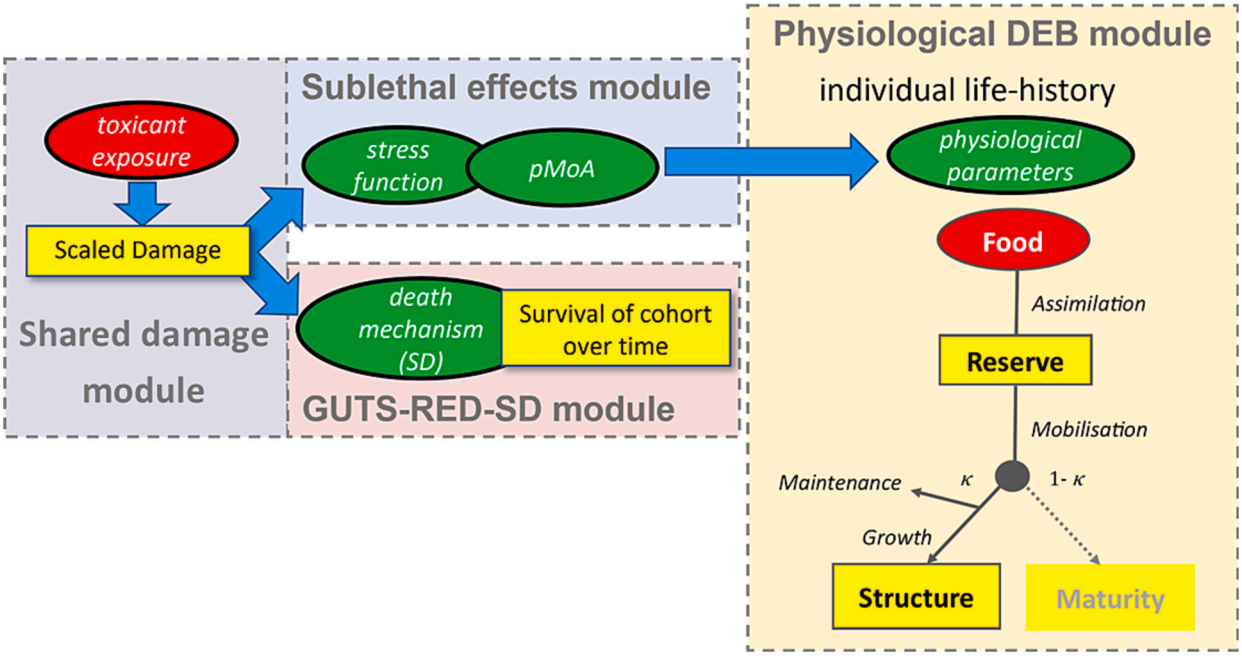


Fig. 1. Schematic of the conceptual model used here, based on (Jager, 2020). It illustrates the 3 main modules of our model: physiological DEB, GUTS-RED-SD and sublethal effects. Yellow boxes represent state variables, of which maturity is not presented, as no data are available for comparison. The red oval represents the forcing variable exposure, which is not affected by the modelled processes. Green ovals and thick blue arrows represent model parameters, functions and their linkages. Within the physiological DEB module, thin black arrows represent energy fluxes and the text in italics the 4 main processes that could be adversely affected by the toxicant, depending on the assumed physiological mode of action (pMoA). The processes of assimilation and reserve mobilisation are reduced by stress while the energetic costs of maintenance and growth increase with stress.

threshold model of survival, assuming a stochastic death mechanism (GUTS-RED-SD) (Ashauer et al., 2011; Jager and Ashauer, 2018).

2.4.4. Damage model

Both the sublethal and lethal effects modules share a reduced TKTD module that combines toxicokinetics and the first stage of toxicodynamics (see eq. S7 in Supporting Information). This module directly links exposure concentration (in this case, BZM concentration in the egg) over time *C(t)* [mg a.s. × kg^{-1}] to internal 'scaled damage', *D(t)* [mg a.s. × kg^{-1}], assuming first-order kinetics. The speed at which scaled damage approaches the exposure concentration (both have the same units) is determined by a single parameter, the 'dominant rate constant', k_d [d^{-1}], so called because it describes the rates of both damage accrual and repair (EFSA PPR et al., 2018). Such one-compartment models are widely used in ecotoxicology, where detailed toxicokinetics data are often unavailable (Jager and Ashauer, 2018).

We assumed that the BZM concentration inside the egg was constant during embryo development both in the calibration and validation studies. This was because the weight of the egg remains relatively constant and it was assumed that BZM cannot be further metabolised (flubenzoic acid, the metabolite of BZM was not found in eggs in a metabolism study (Bayer, 2008c)). We tested two model variants with different assumptions regarding exposure: 1) the 'immediate exposure model', in which the embryo is exposed to BZM from the point of injection; 2) the 'delayed exposure model', in which the embryo is only exposed once it starts to use yolk as an energy source (day 8). The delayed exposure variant assumes that this occurs because BZM is slightly lipophilic ($\log(K_{ow}) = 0.64$, NCBI (2022)) and so reaches a higher concentration in the yolk, due to its high lipid content (Ricklefs, 1977; Skadhauge, 1981).

2.5. Effects

In both the lethal and sublethal effects modules it is assumed that, up to a threshold level, damage can be tolerated by the organism and has no negative impact. Beyond the threshold level, damage leads to a linear increase in survival hazard, or physiological stress leading to sublethal effects. TD parameters, determining the thresholds and effect strength, are free to take different values for lethal and sublethal effects.

Lethal effects were modelled according to the stochastic death (SD) mechanism common for DEB-TKTD models (Jager, 2020). SD assumes that all animals in a population have the same tolerance for the chemical and so have the same probability of survival given the same exposure to a toxicant. The increased hazard resulting from damage is added to the background hazard and used to model survival probability over time, *S(t)*. The original SD and DEB models were developed alongside one another and are highly compatible (Bedaux and Kooijman, 1994; Kooijman and Bedaux, 1996). The reason for including a survival module was to better inform estimation of the dominant rate constant, k_d, and selection of TK assumptions (immediate vs delayed exposure) in the damage module.

Sublethal effects were modelled by altering one or more physiological model processes as a function of stress over time, *s(t)*. For an overview of standard processes see Jager et al. (2023). The physiological mode(s) of action (pMoA) determines which process(es) are subjected to stress. Here, only pMoAs that affect growth were tested.

1. Maintenance

Stress applied to maintenance increases the maintenance costs per unit of structure (the parameter $\left[\dot{p}_M\right]$, J × cm^{-3} × day^{-1}) by a factor of $(1 + s(t))$. This leads to the same proportional increase in overall maintenance flux. With more energy used on maintenance, less energy is available for growth and so growth rate is decreased, as is the maximum size that can be achieved. In the first instance (i.e. before any effects on body size and reserve density), the increased energy demands lead to higher mobilisation of reserve (Eq. S6 in Supporting information).

2. Growth

Stress applied to growth increases the cost of synthesising a unit of structure (the parameter $[E_G]$, J $_{\text{Reserve}}$ × cm $^{-3}_{\text{Structure}}$) by a factor of $(1 + s(t))$. This does not affect maximum size but reduces the growth rate. Like stress on maintenance, this pMoA leads to increased reserve mobilisation at a given body size and reserve density.

3. Reserve mobilisation and assimilation

With this pMoA, energy conductance, $\dot{v}$ [cm × day^{-1}], is multiplied by $(1 - s(t))$, leading to a reduction in mobilisation flux. The result is that reserve is utilised at a lower rate so there is decreased energy available for growth and maturation. Stress is also applied to assimilation. Specifically, the maximum area specific assimilation rate, $\left\{\dot{p}_{Am}\right\}$ [J × cm^{-2} × day^{-1}] is also multiplied by $(1 - s(t))$. The corresponding decrease in assimilation flux fixes the maximum reserve density, $[E_m] = \left\{\dot{p}_{Am}\right\} \Big/ \dot{v}$, stopping biologically unplausible reserve accumulation. This assumes that the maximum reserve density that can be reached at a given feeding level is fixed. Maintaining body composition is particularly important for flying birds (Lima, 1986) and energy excretion (analogous to reduced assimilation) has been suggested as a mechanism of reducing deleterious fat accumulation in humans (Lund et al., 2020; Lund and Clemmensen, 2023) and birds (Mathot et al., 2020). While this phenomenon has not been measured in birds, but it has been detected in overfeeding studies on mice (Klaus et al., 2005), and it is known that birds are able to adjust gut transit time in response to predators (Mathot et al., 2020).

2.6. Parameterisation

Most physiological parameters in the embryo model were fixed to values derived from the literature (Table 1), primarily the AmP entry for *C. virginianus* (Marn et al., 2022). Because no assimilation takes place in the egg, the value of the area specific assimilation rate, $\left\{\dot{p}_{Am}\right\}$, had no impact on the model results and so this was fixed as a multiple of the volume specific maintenance rate, $\left[\dot{p}_M\right]$, so that the maximum volume, $V_m = \left(\kappa\left\{\dot{p}_{Am}\right\} \Big/ \left[\dot{p}_M\right]\right)^3$, matched the value from the AmP entry. As is common practice in DEB modelling, it was assumed that the density of wet structure and reserve was equal to that of water (Lika et al., 2011; Martin et al., 2022).

To model chick growth (as part of model validation), parameters for *C. virginianus* from the AmP database (Marn et al., 2022) were used without recalibration of any physiological parameters. Initial wet weight and damage levels were the values predicted at hatching by the embryo model. Initial reserve density [J $_{\text{Reserve}}$ × cm^{-3} $_{\text{Structure}}$] was calculated as a function of predicted reserve use during incubation. The maximum reserve density, $[E_m]$, was multiplied by predicted reserve use as proportion of that predicted in the control treatment.

2.7. Calibration

Three physiological parameters and the background survival hazard were fitted to the control data (fitting routine described below). Their values were restricted to realistic ranges, based on literature data. The volume-specific maintenance rate, $\left[\dot{p}_M\right]$ was limited to the 2.5th and

Table 1
Fixed A) and estimated B) parameters of the DEB-TKTD model for embryo growth. Sources for the fixed values are indicated in the right-hand column.

A) Parameters taken from literature

Parameters of the embryo model

Symbol	Value	Unit	Definition	Source
κ	0.555	–	Allocation to soma, (0–1)	(Marn et al., 2022)
κ_G	0.800	–	fraction of growth energy fixed in structure	
μ_E	550,000	J × $mol^{-1}_{(Reserve)}$	Chemical potential of reserve	Calculated based on chemical indices for organics in Marn et al. (2022)
w_E	23.9	g × mol^{-1}	Molecular weight of reserve	
μ_V	500,000	J × $mol^{-1}_{(Reserve)}$	Chemical potential of structure	
w_V	23.9	g × mol^{-1}	Molecular weight of structure	
d_E	0.500	$g_{(dry)}$ × $cm^{-3}_{(wet)}$	Density of reserve	d_E = 1- fraction water content of yolk in Galliformes (Ricklefs, 1977)
d_V	0.150	$g_{(dry)}$ × $cm^{-3}_{(wet)}$	Density of structure	d_V = 1- water content of hatchlings (Skadhauge, 1981) Both rounded to the nearest 0.05.
$[E_G]$	3921	J× cm^{-3}	Energetic cost per cm^3 structural growth	$[E_G] = \frac{\mu_V d_V}{w_V \kappa_G}$ as per Kooijman (2009)
$\left\{\dot{p}_{Am}\right\}/\left[\dot{p}_M\right]$	10.6	1× cm^{-1}	Ratio of max. Area specific assimilation rate and $\left[\dot{p}_M\right]$	$\left\{\dot{p}_{Am}\right\}/\left[\dot{p}_M\right] = L_m/\kappa = \frac{5.85}{0.55} = 10.6$ Where L_m is the cube root of the observed ultimate volume (200cm^3) as per Marn et al. (2022)

B) Parameters estimated from data

Symbol	Unit	Definition
$\dot{v}$	cm × d^{-1}	energy conductance (velocity)
$\left[\dot{p}_M\right]$	J× cm^{-3}× d^{-1}	specific volume-linked somatic maintenance rate
Ww_{E0}	g	initial wet weight of reserve
k_d	1× d^{-1}	dominant rate constant
z_s	mg × kg^{-1}	threshold survival
b_s	kg × mg^{-1}× d^{-1}	effect strength survival
z_b	mg × kg^{-1}	threshold energy budget
b_b	kg × mg^{-1}	effect strength energy budget

97.5th percentiles of estimated values in the whole AmP library. The energy conductance, v, was estimated within the range [0,1] which covers virtually every entry in the AmP library. The initial values (at the start of the fitting routine) of both parameters were those of the 'generalised animal' from AmP. The final estimated parameter was the initial wet weight of reserve, Ww_{E0}. Literature data show that the yolk weight of the northern bobwhite, and related *Galliformes*, is typically around 30 % of total egg weight (Ricklefs, 1977; Smith et al., 1996; Woodard and Wilson, 1963) and up to 44 % (Okon et al., 2020). Allowing for the influx of water and nutrients from the albumen (which also contribute to reserve) and based on total egg weight of 10-11 g (Skewes et al., 1988; Smith et al., 1996), a range of 4.5–6.5 g was specified for estimation of this parameter, with a starting value of 5.5 g. Background survival hazard was fitted in the range 0 (no mortality) to 1 (no survival after one day). The DEB-TKTD model does not explicitly include tarsus growth. Thus, to be able to utilize tarsus length data for model fitting, an empirical relationship between embryo weight and tarsus length was derived (Supplementary Information).

The five TKTD parameters were estimated simultaneously for each pMoA. The growth data from the 500 mg × kg^{-1} bw treatment were omitted due to low sample size (only two individuals) at the final time point. Initial values for the survival parameters were derived using the dedicated function start_vals_guts in BYOM. The same initial threshold and effect strength values were used for the sublethal effects model as GUTS-SD and DEB-TKTD are based on similar equations.

The value of k_d, was constrained so scaled damage would be at least 5 % of external concentration by the end of the observation period. If k_d is allowed to take values very close to zero this can lead to high uncertainty around other parameters while making no meaningful improvement to the model fit. In essence, when data suggest that the value of k_d is low, even a very narrow confidence interval (in absolute terms) may cover infinite orders of magnitude without including zero. The lower values in the interval will result in reduced scaled damage over time, requiring the effect strengths, b_b and b_s, to be increased accordingly, towards infinity, to match observed effects. The result is an extremely wide confidence interval for b_b and b_s, making model predictions uninformative. This is a known issue with the reduced damage model used here (and indeed any model when calibration data are insufficient to precisely identify parameter values). As such, it is recommended practice to impose constraints on k_d to avoid issues with its value approaching zero or infinity (Jager and Ashauer, 2018).

Initial fitting was performed using a Nelder-Mead simplex algorithm, this method has been shown to be a robust method for DEB modelling and is widely used (AmP, 2023). However, to avoid convergence on local optima, likelihood profiling was also performed, also providing 95 % confidence intervals around parameter values. Final fitting was performed using the parameter space explorer algorithm (Jager, 2021), also generating 95 % confidence intervals around the modelled mean body size and survival probability over time. To make this process faster, parameter values were constrained based on the results of the likelihood profiling. The lower boundary for each parameter was set at 0.5 × lower 95 % confidence limit from likelihood profiling while the upper boundary was 2 × upper 95 % confidence limit.

2.8. Validation

Prior to validation only the scaled feeding rate, f, was fitted to validation control data. This value was expected to be greater than the default assumption of 1 (as is the case for hatchlings in the AmP entry for the species) to account for high feeding and rapid growth during the first weeks after hatching. The fitted value was 1.94, which was in line with the values estimated for juvenile datasets (1.47–1.91) (Marn et al.,

2022). All other DEB parameters were fixed at the values suggested by the AmP entry while the TKTD parameters were fixed at the values derived through model calibration.

Predicted BZM concentration inside the eggs varied depending on when they were laid, due to temporal variation in the feeding rate of hens. Dose levels in the validation dataset were therefore defined as the temporal mean predicted BZM concentration in eggs at each dietary concentration. The model was then used to predict mean hatchling weight and mean14-day chick weight at each dose level. The results, both observed and predicted, were converted to predicted mean effect sizes (%, calculated as $100 \times (W_w(t)_{Treatment} - W_w(t)_{Control}) / W_w(t)_{Control}$). For visualisation of the dose-response relationship, effect sizes at intermediate BZM concentrations were interpolated and plotted against BZM concentration in the egg to produce a modelled dose-response curve. The same procedure was followed to interpolate confidence intervals on effect size across the modelled range of BZM concentrations over the whole egg laying period. These modelled dose response curves were then compared to observed effect sizes in each treatment.

2.9. Model performance analysis

The quality of calibrated model fits was evaluated through a combination of qualitative and quantitative criteria. Qualitatively, the visual match between model predictions and data and the likelihood regions of the estimated parameters were assessed. Quantitatively, four criteria were considered. The normalised root mean square error (NRMSE) and the coefficient of determination (R^2) both consider the match over time while the survival probability prediction error (SPPE) considers the match between the model and survival data at the final timepoint (Focks et al., 2018). The SPPE is calculated per treatment and quantifies the error at the last timepoint as a percentage of the observed value. Both the NRMSE and SPPE were suggested by the recent EFSA scientific opinion on the use of TKTD modelling for pesticide risk assessment (EFSA PPR et al., 2018). Finally, the Akaike information criterion (AIC) was considered as a measure of prediction error to all data.

Quantitative performance criteria for validation were, similarly to calibrations, NRMSE and R^2. However, in contrast to calibrations, these metrics were not calculated with respect to simulated weight over time, but on the prediction of the relative effect data by the modelled dose response curve. As effects were observed at only one time point for each model (hatching for the embryo model and day 14 for the chick model), predictions are plotted against concentration rather than time. To quantify how well the model captures absolute weight and not only relative effect sizes, a metric analogous to the SPPE was also calculated, the growth prediction error (GPE). The GPE quantified the error in the prediction of hatchling or 14-day chick weight as a percentage of the observed values on a per treatment basis.

3. Results

3.1. DEB-TKTD calibration – egg injection studies

3.1.1. Fit to control treatment

When fitted to the control data the resulting parameter values (and 95 % CIs) were $\dot{v}$ =0.0542 cm × d^{-1} (0.0525–0.0561), $\left[\dot{p}_M\right]$ = 7.60 J × cm^{-3} (<7.60–21.3), and Ww_{E0} = 5.07 g (4.85–5.49). The volume specific maintenance rate, $\left[\dot{p}_M\right]$, was fitted to its lower boundary but the boundaries were not extended, as lower values were deemed unrealistic for an endotherm, even an embryo. Visually and quantitatively, these values produced a close fit to all endpoints (see supporting Fig. S4), with NRMSE = 12.8 % and 7.1 % for weight and tarsus length respectively ($R^2 > 0.93$ for both) and NRMSE = 4.8 % for final yolk weight (R^2 not valid with only one observation point and treatment). Separately, the background survival hazard rate was fitted to 0.00973 day^{-1}, the SPPE −0.73 %.

3.1.2. Fitting to the treatment groups

The TKTD model parameters were first fitted for all physiological modes of action, assuming immediate exposure. This showed a clear qualitative difference between the stress on reserve mobilisation and the other two pMoAs. When maintenance or growth was stressed, the effects on body weight and tarsus length over time were simulated relatively accurately ($R^2 \geq 0.89$, NRMSE <17 % for maintenance and $R^2 \geq 0.93$, NRMSE <14 % for growth) but both led to increased reserve use and therefore lower final yolk weight in the higher exposure levels - the opposite of the observed pattern (see supporting Fig. S6). When maintenance was stressed, modelled embryo weight began to decrease towards the end of the incubation period in the highest dose groups. For either pMoA, modelled reserve use increased with dose level earlier in the incubation period and decreased in the later stages of incubation due to smaller body size, this was more noticeable for stress on growth costs. However, modelled reserve use was still higher overall in the stressed groups. On the other hand, stress on reserve mobilisation allowed the model to accurately simulate the observed body weight and tarsus length over time, while also matching the observed pattern in final yolk weight (Fig. 2). This mode of action was therefore selected for further fitting and analysis.

As described in the Methods, final fits were generated for the pMoA reserve mobilisation using the parameter space explorer algorithm to generate confidence intervals on predictions (Table 2 and Fig. 2, supporting Figs. S9 & S10). This was also carried out assuming exposure began at day 8 when the yolk starts to be absorbed. For both lethal and sublethal effects, the delayed exposure model had lower thresholds (z_s and z_b) and higher effect strengths (b_s and b_b). This pattern was predictable, as the assumption of delayed exposure meant that there was less time for damage to accumulate and cause effects. The best fitting value of z_b in the delayed exposure model was 0. However, the parameter was not removed from the model, so that uncertainty around its value could still be included in simulations. For both exposure scenarios, the model simulations closely matched the data, with little difference in the goodness of fit measures (Table 3). R^2 was >0.90 for embryo weight and tarsus length for both model variants while NRMSE was ≤15 % for weight and < 7 % for tarsus length. For all treatments and both model variants, |SPPE| <18 % while mean |SPPE| was <8.5 % for both model variants. There was little difference between the performance of the model variants but the immediate exposure variant fit the sublethal endpoints slightly better. However, survival data exhibited a delayed response to exposure which was best captured by the delayed exposure variant. This was reflected by a lower Akaike information criterion, indicating that the delayed exposure variant was the most accurate model overall. As survival data were included primarily to inform the fitting of the toxicokinetic model, the delayed exposure model variant was chosen for validation.

3.2. DEB-TKTD validation – reproductive toxicity studies

3.2.1. Predicted BMZ concentrations in eggs

As part of the validation process, treatment groups in the validation dataset were adjusted from dietary concentration (in mg BZM × kg^{-1} diet) to BZM concentration in the egg (mg BZM × kg^{-1} egg) based on PBK model predictions. The variability of feeding rates of adult hens was reflected in the variability of predicted BMZ concentrations in the eggs. While predicted concentrations inside eggs varied depending on when they were laid, concentrations inside individual eggs were assumed to remain constant (see Methods, Damage Model). The results were used to derive the average BZM concentration inside eggs from each treatment group (Fig. 3). Feeding rates increased during the egg-laying phase. Predicted BMZ concentrations in the yolk were consistently higher than in the albumen.

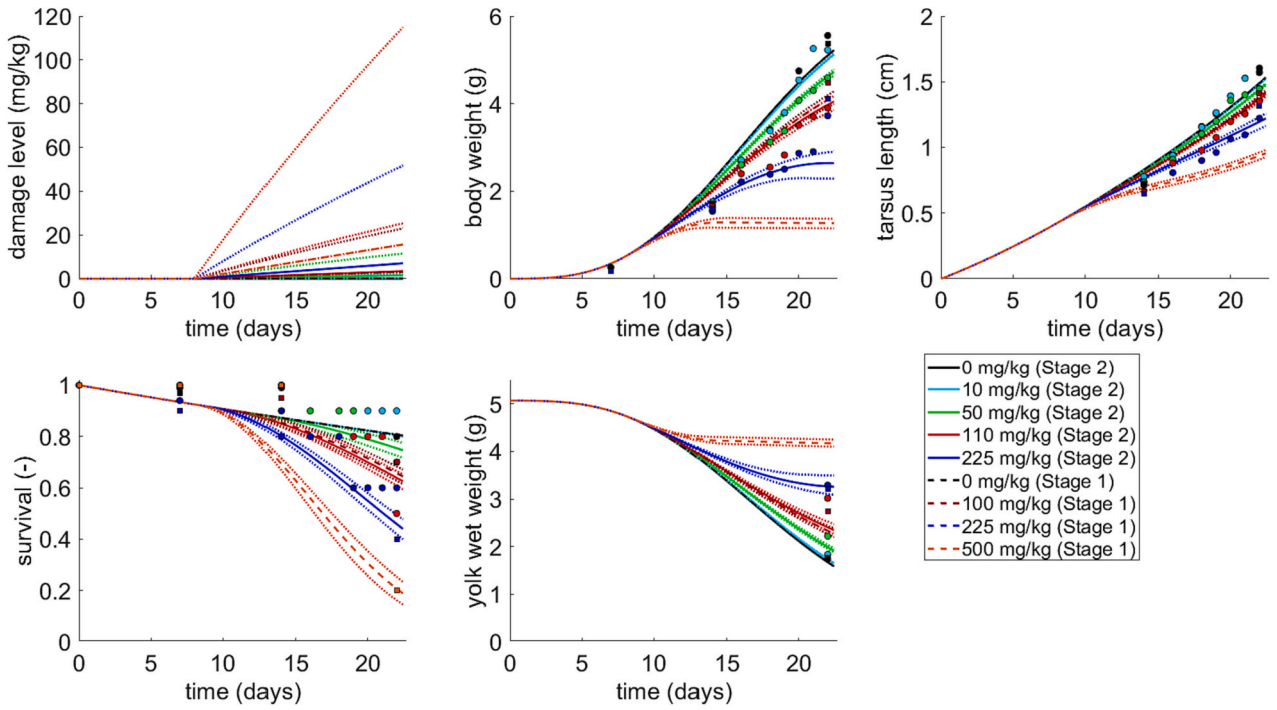


Fig. 2. Fit to data: model assuming the pMoA stress on energy conductance and delayed exposure. The figure shows simulation results by the calibrated model as lines with 95 % confidence intervals (dotted lines) and data from corresponding treatments as circles of the same colour. This mode of action and the assumption of exposure starting on day 8, after the embryo starts to use energy from yolk, gave a qualitatively good match to the observed patterns of growth and survival. Weight and yolk data from the 500 mg × kg^{-1} treatment was omitted when fitting the model.

Table 2
Fitted parameter values for the pMoA stress on reserve mobilisation, assuming delayed exposure in the egg.

Parameter	Best fit	95 % CI lower limit	95 % CI upper limit
k_d [1× d^{-1}]	0.0022[a]	0.0022	0.0148
z_s [mg × kg^{-1}]	0.522	0.251	3.77
b_s [−]	0.0142	0.00223	0.0177
z_b [mg × kg^{-1}]	0[a]	0	0.181
b_b [−]	0.145	0.0227	0.163

k_d = dominant rate constant.
z_b = threshold energy budget effects.
z_s = threshold survival effects.
b_b = effect strength energy budget.
b_s = effect strength survival.
[a] Equal to lower boundary. In the case of k_d, lower limit corresponds to 5 % saturation within 22 days.

3.2.2. *Effects predicted by DEB-TKTD*

Effect on mean observed weight was calculated for each treatment over the whole egg-laying period and was plotted against the mean predicted BZM concentration inside the egg over the same period. The effect on 14-day chick weight at the highest BZM concentration could not be predicted to 1 s.d. because that data point was based on only one surviving chick and so no standard deviation could be calculated. However, it was within the confidence intervals around model predictions. Due to its low sample size, the highest concentration was omitted from quantitative Goodness-of-Fit analyses for effects on 14-day chick weight. The NRMSE was below 50 % for 14-day chick weight and slightly above for hatchling weight while the R^2 was in the range 0.5–0.6 for both endpoints. Absolute predictions were also assessed by calculating the GPE, or the error in each treatment as a percentage of the observed value. In all cases, |GPE| was <19 % and the mean |GPE| was <8 % for both endpoints (Table 4).

Table 3
Performance criteria for DEB-TKTD models assuming immediate and delayed exposure– for calibration treatment data of *C. virginianus*. AIC is an estimator of prediction error, with lower values indicating the higher quality model. NMRSE ≤50 % and high R^2 values indicate good model performance. SPPE values of 0 indicate an exact prediction, negative values underestimation and positive values overestimation of effects.

	Immediate exposure	Delayed exposure
Akaike Information Criterion (AIC) [−]	2366	2260
Normalised Root Mean Square Error (NRMSE) [%] – embryo weight	13.3	15.0
Normalised Root Mean Square Error (NRMSE) [%] – tarsus length	5.77	6.61
R^2 [−] – embryo weight	0.933	0.916
R^2 [−] – tarsus length	0.950	0.934
Survival-probability prediction error (SPPE) [%]		
Stage 1 study- 100 mg × kg^{-1} benzamide	0.44	4.42
Stage 1 study- 225 mg × kg^{-1} benzamide	−9.13	−5.12
Stage 1 study- 500 mg × kg^{-1} benzamide	−2.36	0.35
Stage 2 study- 10 mg × kg^{-1} benzamide	9.27	9.27
Stage 2 study- 50 mg × kg^{-1} benzamide	−8.54	−5.63
Stage 2 study- 110 mg × kg^{-1} benzamide	−17.7	−13.7
Stage 2 study- 225 mg × kg^{-1} benzamide	10.9	14.9

4. Discussion

This study set out to address the lack of case studies of DEB-TKTD models predicting pesticide toxicity to birds (EFSA, 2023). Our model successfully predicted effects on the weight of *C.virginianus* hatchlings and 14-day old chicks resulting from dietary exposure of laying hens to

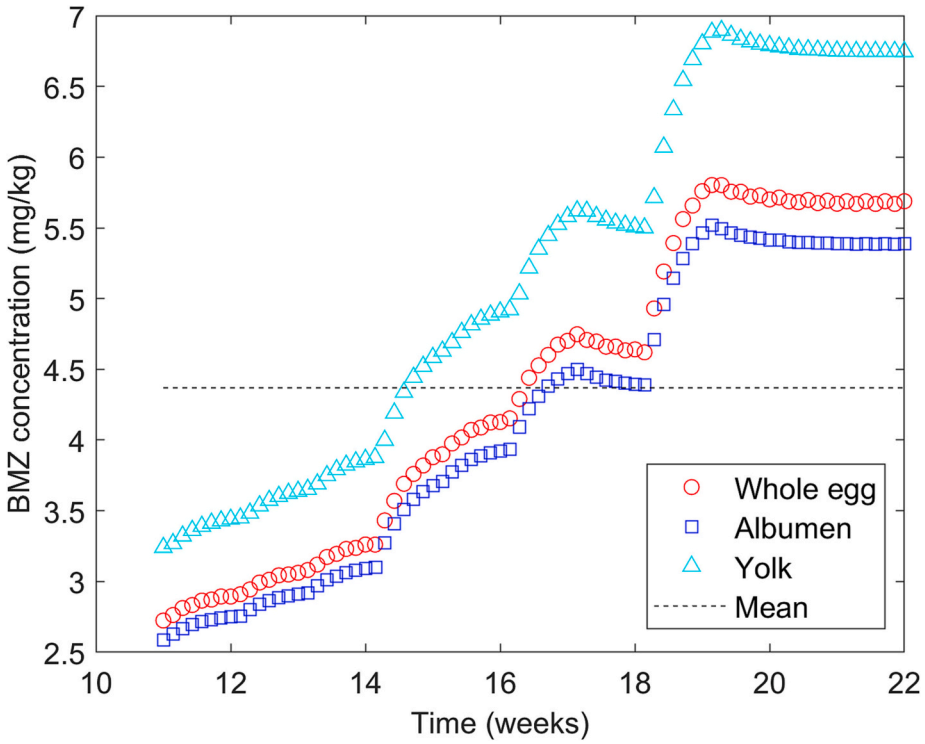


Fig. 3. BMZ concentrations in the whole egg (red circles), albumen (blue squares) and yolk (light blue triangles) during the egg-laying period, predicted by the PBK model for the 50 ppm dose group in a reproduction study with the *C. virginianus* (Bayer, 2008a). The temporal mean concentration in the whole egg (dashed horizontal line) was used as exposure for the embryo to make DEB-TKTD predictions on hatchling and 14-day chick weight, while the standard deviation of whole egg concentrations over time is used to derive standard deviations of effect sizes.

Table 4
Performance criteria for DEB-TKTD predictions of hatchling weight and 14-day chick weight.

Hatchling				Chick			
NRMSE	R^2	Max | GPE|	Mean GPE	NRMSE	R^2	Max | GPE|	Mean GPE
52.4	0.568	18.7	7.63	41.7	0.540	15.6	−0.29

FLU. Crucial to this success were the calibration dataset – generated by a specially designed experiment – and PBK modelling – which allowed the quantification of BZM concentration in the eggs following dietary dosing. However, these steps were only necessary to investigate *in ovo* toxicity and are not prerequisites for all DEB-TKTD applications.

The basis for DEB-TKTD models is a physiological DEB model which can accurately and realistically simulate growth alongside maturation and/or reproduction for a species under control conditions (Sherborne and Galic, 2020). *In ovo* embryonic growth represents a special case in DEB theory as the organism initially consists almost entirely of reserve which is utilised over time to grow structure, without any external source of energy or mass. An advantage, from a modelling perspective, is that 'reserve', which is normally abstract, can be approximately quantified as the energy content or mass of yolk and compared to data. The physiological model fit very well to control data for all endpoints, including final yolk weight. Parameter values were fixed based on literature data or fitted within a realistic range. The close fit to data lends support to the DEB framework as a physiological basis from which to model toxicity.

In contrast to many DEB model applications, calibration data were generated specifically for this study. The results provided vital insights into the reproductive effects observed in the reproduction study, suggesting these effects could be explained by toxic action inside the egg. As the egg is essentially a closed system, it appears that exposure to toxicants transferred from the mother can be accurately mimicked through egg injection. This dosing method is very precise, with little to no intra-treatment variability in dose, and the choice of measured endpoints produce a strong dataset for DEB-TKTD model calibration. Additionally, this study design requires relatively little time and, due to its precision, few animals, so it is preferable, in several respects, to further data generation through reproduction studies.

While the endpoints of interest for predictions were the weights of hatchling and 14-day old chicks, the inclusion of an endpoint other than growth, the amount of yolk used, was vital for model selection. With growth data alone, pMoAs must generally be differentiated based on the fit of the calibrated model to data, as each pMoA affects the modelled growth curve differently over time (Sherborne et al., 2020). Even with knowledge of how a chemical acts at a cellular level, there is usually no way to select mode of action a priori since DEB processes (assimilation, maintenance, growth, and reserve mobilisation) are broad and somewhat abstract (Perkins et al., 2019). However, in our case all the modes of action tested provided similar fits to growth data. Therefore, growth data alone would have been insufficient for a decisive choice of pMoA. For adult animals, reproduction data can provide the additional information needed to inform pMoA selection. Since embryos do not reproduce, and maturity is not directly measurable, yolk weight is a highly informative additional endpoint for sublethal effects, as it is directly related to reserve dynamics. Alternatively, respiration of the eggs could be measured in future studies (Hamidu et al., 2010), any increase relative to controls would be consistent with stress on maintenance or growth, while a decrease would be indicate stress on mobilisation.

In our study, yolk weight was measured only at the final timepoint and this information was critical in identifying the pMoA. Of the pMoAs tested, only stress on reserve mobilisation and assimilation could explain the pattern of decreased yolk utilisation alongside reduced embryo growth with increasing exposure concentrations. Stress on kappa could potentially result in reduced growth and reserve utilisation (eq. S6 in

Supporting Material). However, this would also result in more energy being allocated to maturation, so would imply that embryos mature faster in response to BZM exposure. Thus, the pMoA of BZM toxicity could be selected on decisively and not only based on small quantitative differences in fit quality to growth data. Yolk weight data collected at other timepoints, particularly near the beginning of the study, would be highly valuable for any similar studies in future. Firstly, this could better inform the initial reserve mass, which in this study was fitted within a range based on literature data. Secondly, time-course data could prove crucial in distinguishing between the pMoAs of stress on growth or maintenance. An accurate choice of pMoA is crucial for risk assessments, as although differences in fit to growth calibration data under different pMoAs may be small, the effect of pMoA choice on forward predictions can be large, especially at the population level (Martin et al., 2014).

The inclusion of embryo survival data also played a key role in understanding the *in ovo* kinetics of benzamide, where we tested two different assumptions. First, that exposure to embryo begins immediately after BZM injection into the egg or, second, that BZM accumulates in the yolk and so exposure begins alongside yolk absorption. The model fit well to growth data under both assumptions. However, survival data only showed effects later in the observation period, supporting the assumption of delayed exposure and resulting in a lower AIC value of the fitted model.

A limitation of the egg injection data was that there was no way to include a recovery period or to implement pulsed exposure. This meant that the information regarding damage dynamics was limited. The data suggested slow damage dynamics resulting in k_d being fitted to its (imposed) lower limit, although there was some uncertainty around its value. Where information is limited, a very similar fit to effects data can be achieved for a range of 'low' k_d values, by adjusting the values of TD parameters accordingly. This is reflected in the correlations between parameter values visible on the parameter space plots (supporting Figs. S10 & S11). Such correlations may be taken to indicate that the model is overfitted and that correlated parameters could be combined to reduce the number of free parameters. However, the modular and mechanistic nature of DEB-TKTD models means that this is not advisable. Such changes to the model would compromise its biological interpretability and would bring into question whether it still fits within the DEB-TKD framework. Additionally, several of the visible correlations are coincidental rather than causal relationships. For example, lethal and sublethal effect strength (b_s and b_b respectively) both depend on k_d and are therefore correlated despite having no mechanistic link to one another.

The model was then used to predict effects (as percentage reduction relative to controls) on hatchling weight and 14-day old chick weight observed in the reproduction study (OECD, 1984). Model validation was successful for both endpoints. This is particularly noteworthy, as predictions for the 14-day old chicks represented an extrapolation of toxic effects from one life-stage to another. This was achieved by applying the same TKTD parameters (fitted to embryo data) to a different DEB model, comprising the standard model parameters for *C. virginianus* from the AmP database (Marn et al., 2022). Chicks were not exposed to BZM post-hatching, yet effects were larger than at hatching, suggesting carry over toxicity from exposure in the egg. Due to the low dominant rate constant - k_d was fitted to its lower limit – the model predicted that damage accrued during incubation was repaired slowly, leading to further effects during the first 14 days of growth. The reduction in body weights 14 days after hatching were even predicted slightly more accurately than at hatching.

This result is highly encouraging. Firstly, it demonstrates the potential value of the egg injection study design. This experimental setup appropriately replicates the maternal transfer of the test compound to the eggs and elucidates the extent and mechanism of toxic action during development. Such experiments provide highly valuable data which can inform TKTD model calibration and contribute to the 3Rs (reduction, refinement and replacement) of animal testing. Non-mammalian embryos are not currently regarded as a sentient life-stage under EU law (European Council, 2010) and the study design requires far fewer individuals than standard reproduction studies (Farhat et al., 2020).

Secondly, model validation shows that TKTD parameters can be successfully transferred between models for different life-stages using parameter values from the AmP library. Standard DEB parameters are available for many bird species, including those commonly used in regulatory studies for risk assessment (Kooijman, 2022; Schuckink Kool and Kooijman, 2016; Teixeira et al., 2018), but despite this -and the fact that DEB was originally developed as a means of better understanding toxic effects (Kooijman and Bedaux, 1996) – applications of DEB models to predict toxic effects in birds are currently lacking. One obstacle to the use of standard AmP parameters has been the variability of control data, meaning that models must be recalibrated to new datasets (Jager et al., 2023). However, considering predictions in terms of relative effect sizes compared to control rather than absolute endpoints as done here, circumvents this issue to an extent. Applying TKTD model parameters together with an existing set of DEB parameters is a far more streamlined process than calibrating and validating the physiological model for each application. Moreover, the published parameter sets in AmP have been reviewed by experts and the DEB framework itself is highly standardised (Marques et al., 2018), so these model components do not need to be re-evaluated for each application. Such a database of parameter values for different species is not available for simplified DEB-TKTD models, so simply transferring TKTD parameters between models without recalibration is not an option. However, such models are simpler to recalibrate to control data (Jager, 2020).

While our model predictions were well within the range of the validation data, it must be noted that the validation data showed some variability. A potential explanation for this variability is the dosing method. While dosing with BZM was exact in the calibration experiments, the validation data came from reproduction studies on FLU. Laying bobwhite hens were fed a diet containing a fixed concentration of FLU in each treatment but feeding rate – and therefore ingested dose – could vary between individuals and over time. Since hatchlings are a life-stage of interest, the assessment of *in ovo* exposure in reproduction studies would certainly be beneficial for understanding and modelling observed effects. However, this study showed how PBK modelling could bridge this data gap, by predicting toxicant transfer from exposed birds into their eggs based on metabolism studies on chickens (OECD, 2007). Intra-treatment variability in exposure was accounted for by the model (the horizontal error bars on Fig. 4 reflect the temporal variability in predicted BZM concentration in eggs). This demonstrates how TKTD modelling can account for dynamic exposure, a key advantage over conventional analysis.

The toxicity exposure ratio (TER), used in chronic risk assessment for birds and mammals, considers constant exposure (average daily dietary dose, DDD) in lab and field scenarios which may differ greatly in duration (EFSA, 2023). By accounting for dynamic exposure, the TKTD model developed in this study could offer greater realism. For example, the model could be used to find the EC_{10} (the lowest dietary concentration where an 10 % effect on hatchling or 14-day chick weight would be predicted) by making predictions for eggs laid at various timepoints following dietary exposure. The predicted EC_{10} could then be compared to realistic residue levels in food items. Alternatively, it could be used to find the DP_{10}, the multiplication factor that would need to be applied to a realistic exposure profile for a 10 % effect to be predicted. Such applications in risk assessment have already been suggested as part of ERA by EFSA's guidance for birds and mammals (EFSA, 2023). At the population level, effects may also be considered through the implementation of DEB-TKTD modelling with individual-based population models (Martin et al., 2012; Vaugeois et al., 2021).

5. Conclusions

The transfer of toxicants into eggs can be an important driver of

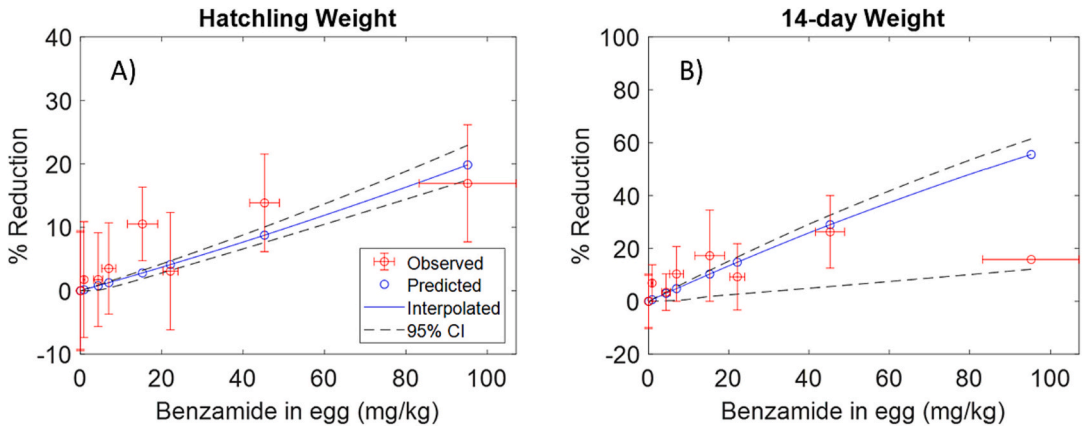


Fig. 4. Prediction of validation data expressed as % reduction compared to control for hatchling weight (A) and 14-day chick weight (B) from the delayed exposure model. Predicted % reduction as a function of BZM concentration in the whole egg is shown as a blue line (best fit, dashed grey line: 95 %-confidence interval), created as interpolation between predictions made at concentrations indicated by the blue circles. Independent empirical validation data is displayed as red circles along with their error bars. Vertical error bars are calculated as standard deviation of weight data for each treatment divided by the control mean. Horizontal error bars express the standard deviation of predicted whole-egg concentrations among the eggs laid at different times during the validation experiments by the PBK model (Fig. 3). The observed effect on 14-day chick weight at the highest treatment level has no standard deviation on the effect size (i.e., no vertical error bar) as it comprises weight data only from a sole survivor chick.

reproductive toxicity in birds and this process can be well replicated by egg injection studies (Farhat et al., 2020). With the appropriate choice of measured endpoints and sufficient temporal resolution, egg injection studies give greater insight into the mechanisms of *in ovo* toxic effects than standard datasets. Additionally, the study design requires fewer animals and less time than standard reproduction studies. The resulting data are very well suited for calibrating mechanistic models. In particular, the data improve confidence when selecting pMoA for DEB-TKTD models, a choice which is not always clear and can impact predictions substantially. Time-course data for yolk mass should be considered in future studies, as these data are particularly important for pMoA selection.

For birds exposed via diet, exposure (as ingested dose) can be calculated from concentration in the diet and food consumption. However, for developing embryos, characterising exposure to toxicants requires knowledge of the chemical's kinetics in laying birds, particularly its transfer to the eggs. In this regard, metabolism studies on chickens (OECD, 2007) provide a valuable data source. In the first instance such data can determine whether parent compound or metabolites are most relevant for embryo exposure, which is vital for the design of egg injection studies. Secondly, these data can be used to calibrate PBK models (Baier et al., 2022), simulating the transfer of the compound into the eggs. PBK model predictions are essential. Without them, observed effects in reproduction studies cannot be modelled, since exposure inside the egg is not measured.

Effects observed in the reproduction study were accurately predicted by the DEB-TKTD model, giving support to the model itself but also the whole approach, including the experimental design and the PBK model. We also showed that TKTD parameters can make accurate predictions when transferred between models for different life-stages of a species. Considering model predictions in terms of relative effects obviates the need for recalibration of DEB parameters from AmP to control data. This approach can greatly increase the utility of AmP models for species considered in ERA.

The potential applications for DEB-TKTD modelling in pesticide ERA are clear, having already been outlined in EFSA guidance (EFSA, 2023). This study provides a valuable case-study to support the methodology, demonstrating that pesticide effects on avian reproduction can be modelled accurately with the framework.

CRediT authorship contribution statement

Thomas Martin: Formal analysis, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Barbara Bauer:** Formal analysis, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Vanessa Baier:** Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Alicia Paini:** Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Stephan Schaller:** Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Patrick Hubbard:** Investigation, Writing – review & editing. **Markus Ebeling:** Conceptualization, Data curation, Investigation, Writing – review & editing. **David Heckmann:** Conceptualization, Investigation, Methodology, Project administration, Writing – review & editing. **André Gergs:** Conceptualization, Investigation, Methodology, Project administration, Supervision, Writing – review & editing.

Funding

The study was funded by Bayer AG. Already mentioned in Statement of interests.

Declaration of competing interest

A. Gergs, M. Ebeling and D. Heckmann are employees of Bayer AG. The work of the other co-authors on the project was funded by Bayer AG. Bayer AG is a manufacturer of the active substance fluopyram investigated in this paper.

Data availability

Empirical data is available for non-commercial use on request via email: cropscience-transparency@bayer.com. Please indicate in your request the test report numbers as listed in Supporting Information.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2023.169096.

References

AmP, 2023. AmP Parameter Estmation [WWW Document]. URL. https://debportal.debtheory.org/docs/AmPestimation.html (accessed 11.23.23).

Ashauer, R., Agatz, A., Albert, C., Ducrot, V., Galic, N., Hendriks, J., Jager, T., Kretschmann, A., O'Connor, I., Rubach, M.N., Nyman, A.-M., Schmitt, W., Stadnicka, J., van den Brink, P.J., Preuss, T.G., 2011. Toxicokinetic-toxicodynamic modeling of quantal and graded sublethal endpoints: a brief discussion of concepts. Environ. Toxicol. Chem. 30, 2519–2524. https://doi.org/10.1002/etc.639.

Baier, V., Paini, A., Schaller, S., Scanes, C.G., Bone, A.J., Ebeling, M., Preuss, T.G., Witt, J., Heckmann, D., 2022. A generic avian physiologically-based kinetic (PBK) model and its application in three bird species. Environ. Int. 169, 107547 https://doi.org/10.1016/j.envint.2022.107547.

Bayer, 2008a. AE C656948: A Reproduction Study with the Northern Bobwhite.

Bayer, 2008b. Effect of AE C656948 Technical on Reproduction of the Northern Bobwhite Quail.

Bayer, 2008c. Metabolism of [Phenyl-UL-14C]AE C656948 in the Laying Hen.

Bayer, 2022. Comparative In Vitro Metabolism of [Phenyl-UL-14C]Fluopyram in Hepatocytes from Chicken. Duck and Bobwhite Quail.

Bayer, 2022b. Estimation of Apparent Michaelis-Menten Constants for the In Vitro Degradation of Fluopyram in Hepatocytes of Chicken and Bobwhite Quail.

Bedaux, J.J.M., Kooijman, S.A.L.M., 1994. Statistical analysis of bioassays, based on hazard modelling. Environ. Ecol. Stat. 1, 303–314. https://doi.org/10.1007/BF00469427.

Comenges, J.M.Z., Joossens, E., Benito, J.V.S., Worth, A., Paini, A., 2017. Theoretical and mathematical foundation of the virtual cell based assay – a review. Toxicol. In Vitro, SI: The Virtual Cell Assay 45, 209–221. https://doi.org/10.1016/j.tiv.2016.07.013.

EFSA, 2023. Risk assessment for Birds and Mammals | EFSA [WWW Document]. URL. https://www.efsa.europa.eu/en/efsajournal/pub/7790 (accessed 2.10.23).

EFSA PPR, et al., 2009. Risk assessment for birds and mammals. EFSA J. 7, 1438. https://doi.org/10.2903/j.efsa.2009.1438.

EFSA PPR, Ockleford, C., Adriaanse, P., Berny, P., Brock, T., Duquesne, S., Grilli, S., Hernandez-Jerez, A.F., Bennekou, S.H., Klein, M., Kuhl, T., Laskowski, R., Machera, K., Pelkonen, O., Pieper, S., Smith, R.H., Stemmer, M., Sundh, I., Tiktak, A., Topping, C.J., Wolterink, G., Cedergreen, N., Charles, S., Focks, A., Reed, M., Arena, M., Ippolito, A., Byers, H., Teodorovic, I., 2018. Scientific opinion on the state of the art of Toxicokinetic/Toxicodynamic (TKTD) effect models for regulatory risk assessment of pesticides for aquatic organisms. EFSA J. 16, 5377, 188pp. doi:https://doi.org/10.2903/j.efsa.2018.5377.

European Council, 2010. Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. Off. J. Eur. Union L 276, 33.

FAO, WHO, 2011. Fluopyram, in: Pesticide Residues in Food 2010: Joint FAO/WHO Meeting on Pesticide Residues, Rome, Italy, 21–30 September 2010. FAO Plant Production and Protection Paper. FAO, Rome, pp. 263–282.

Farhat, A., Crump, D., Bidinosti, L., Boulanger, E., Basu, N., Hecker, M., Head, J.A., 2020. An early–life stage alternative testing strategy for assessing the impacts of environmental chemicals in birds. Environ. Toxicol. Chem. 39, 141–154. https://doi.org/10.1002/etc.4582.

Fluopyram: Feeding Study Laying Hens (Gallus Gallus Domesticus), 2008.

Focks, A., Belgers, D., Boerwinkel, M.-C., Buijse, L., Roessink, I., Van den Brink, P.J., 2018. Calibration and validation of toxicokinetic-toxicodynamic models for three neonicotinoids and some aquatic macroinvertebrates. Ecotoxicology 27, 992–1007. https://doi.org/10.1007/s10646-018-1940-6.

Grimm, V., Martin, B.T., 2013. Mechanistic effect modeling for ecological risk assessment: where to go from here? Integr. Environ. Assess. Manag. 9, e58–e63. https://doi.org/10.1002/ieam.1423.

Hamidu, J., Feddes, J.J.R., Ouellette, C., Fasenko, G.M., 2010. Measurement of avian embryo respiration by an indirect calorimetric system. Can. Biosyst. Eng. Genie Biosyst. Au Can. 52, 9–24.

Jager, T., 2012. Bad habits die hard: the NOEC's persistence reflects poorly on ecotoxicology. Environ. Toxicol. Chem. 31, 228–229. https://doi.org/10.1002/etc.746.

Jager, T., 2020. Revisiting simplified DEBtox models for analysing ecotoxicity data. Ecol. Model. 416, 108904 https://doi.org/10.1016/j.ecolmodel.2019.108904.

Jager, T., 2021. Robust likelihood-based approach for automated optimization and uncertainty analysis of Toxicokinetic-Toxicodynamic models. Integr. Environ. Assess. Manag. 17, 388–397. https://doi.org/10.1002/ieam.4333.

Jager, T., Ashauer, R., 2018. Modelling Survival under Chemical Stress - a Comprehensive Guide to the GUTS Framework, Version 2.0. ed. leanpub, York, UK.

Jager, T., Goussen, B., Gergs, A., 2023. Using the standard DEB animal model for toxicokinetic-toxicodynamic analysis. Ecol. Model. 475, 110187 https://doi.org/10.1016/j.ecolmodel.2022.110187.

Klaus, S., Pültz, S., Thöne-Reineke, C., Wolfram, S., 2005. Epigallocatechin gallate attenuates diet-induced obesity in mice by decreasing energy absorption and increasing fat oxidation. Int. J. Obes. 29, 615–623. https://doi.org/10.1038/sj.ijo.0802926.

Kooijman, B., 2009. Dynamic Energy Budget Theory for Metabolic Organisation, 3rd ed. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511805400.

Kooijman, B., 2015. AmP Crocodylus johnsoni. version (2015/09/21).

Kooijman, B., 2022. AmP Coturnix japonica. version (2022/11/04).

Kooijman, S.A.L.M., Bedaux, J.J.M., 1996. The Analysis of Aquatic Toxicity Data. VU University Press, Amsterdam.

Kramer, V.J., Etterson, M.A., Hecker, M., Murphy, C.A., Roesijadi, G., Spade, D.J., Spromberg, J.A., Wang, M., Ankley, G.T., 2011. Adverse outcome pathways and ecological risk assessment: bridging to population-level effects. Environ. Toxicol. Chem. 30, 64–76. https://doi.org/10.1002/etc.375.

Lika, K., Kearney, M.R., Freitas, V., van der Veer, H.W., van der Meer, J., Wijsman, J.W.M., Pecquerie, L., Kooijman, S.A.L.M., 2011. The "covariation method" for estimating the parameters of the standard dynamic energy budget model I: philosophy and approach. J. Sea Res. 66, 270–277. https://doi.org/10.1016/j.seares.2011.07.010. The AquaDEB project (phase II): what we've learned from applying the Dynamic Energy Budget theory on aquatic organisms.

Lima, S.L., 1986. Predation risk and unpredictable feeding conditions: determinants of body mass in birds. Ecology 67, 377–385. https://doi.org/10.2307/1938580.

Lund, J., Clemmensen, C., 2023. Physiological protection against weight gain: evidence from overfeeding studies and future directions. Philos. Trans. R. Soc. B Biol. Sci. 378, 20220229. https://doi.org/10.1098/rstb.2022.0229.

Lund, J., Gerhart-Hines, Z., Clemmensen, C., 2020. Role of energy excretion in human body weight regulation. Trends Endocrinol. Metab. 31, 705–708. https://doi.org/10.1016/j.tem.2020.06.002.

Marn, N., Lika, K., Augustine, S., Goussen, B., Ebeling, M., Heckmann, D., Gergs, A., 2022. Energetic basis for bird ontogeny and egg-laying applied to the bobwhite quail. Conserv. Physiol. 10, coac063. https://doi.org/10.1093/conphys/coac063.

Marques, G.M., Augustine, S., Lika, K., Pecquerie, L., Domingos, T., Kooijman, S.A.L.M., 2018. The AmP project: comparing species on the basis of dynamic energy budget parameters. PLoS Comput. Biol. 14, e1006100 https://doi.org/10.1371/journal.pcbi.1006100.

Martin, B., Jager, T., Nisbet, R.M., Preuss, T.G., Grimm, V., 2014. Limitations of extrapolating toxic effects on reproduction to the population level. Ecol. Appl. 24, 1972–1983. https://doi.org/10.1890/14-0656.1.

Martin, B.T., Zimmer, E.I., Grimm, V., Jager, T., 2012. Dynamic energy budget theory meets individual-based modelling: a generic and accessible implementation. Methods Ecol. Evol. 3, 445–449. https://doi.org/10.1111/j.2041-210X.2011.00168.x.

Martin, T., Hodson, M.E., Ashauer, R., 2022. Modelling the effects of variability in feeding rate on growth – a vital step for DEB-TKTD modelling. Ecotoxicol. Environ. Saf. 232, 113231 https://doi.org/10.1016/j.ecoenv.2022.113231.

Mathot, K.J., Kok, E.M.A., van den Hout, P., Dekinga, A., Piersma, T., 2020. Red knots (Calidris canutus islandica) manage body mass with dieting and activity. J. Exp. Biol. 223, jeb231993. https://doi.org/10.1242/jeb.231993.

Matyja, K., 2023. Standard dynamic energy budget model parameter sensitivity. Ecol. Model. 478, 110304 https://doi.org/10.1016/j.ecolmodel.2023.110304.

NCBI, 2022. PubChem Compound Summary for CID 2331, Benzamide.

OECD, 1984. Test No. 206: Avian Reproduction Test.

OECD, 2007. Test No. 503: Metabolism in Livestock. Organisation for Economic co-Operation and Development, Paris.

OECD, 2016. Test No. 223: Avian Acute Oral Toxicity Test, OECD Guidelines for the Testing of Chemicals. Section 2. OECD Publishing, Paris. https://doi.org/10.1787/9789264264519-en.

Okon, B., Ibom, L., Dauda, A., Ebegbulem, V.N., 2020. Egg Quality Traits, Phenotypic Correlations, Egg and Yolk Weights Prediction Using External and Internal Egg Quality Traits of Japanese Quails Reared in Calabar, Nigeria 6.

Perkins, E.J., Ashauer, R., Burgoon, L., Conolly, R., Landesmann, B., Mackay, C., Murphy, C.A., Pollesch, N., Wheeler, J.R., Zupanic, A., Scholz, S., 2019. Building and applying quantitative adverse outcome pathway models for chemical Hazard and risk assessment. Environ. Toxicol. Chem. 38, 1850–1865. https://doi.org/10.1002/etc.4505.

Popp, J., Pető, K., Nagy, J., 2013. Pesticide productivity and food security. A review. Agron. Sustain. Dev. 33, 243–255. https://doi.org/10.1007/s13593-012-0105-x.

Proença, S., Paini, A., Joossens, E., Benito, J.V.S., Berggren, E., Worth, A., Whelan, M., Prieto, P., 2019. Insights into in vitro biokinetics using virtual cell based assay simulations. ALTEX - Altern. Anim. Exp. 36, 447–461. https://doi.org/10.14573/altex.1812101.

Ricklefs, R.E., 1977. Composition of eggs of several bird species. Auk 94, 350–356.

Scanes, C.G., Witt, J., Ebeling, M., Schaller, S., Baier, V., Bone, A.J., Preuss, T.G., Heckmann, D., 2022a. Quantitative morphometric, physiological, and metabolic characteristics of chickens and mallards for physiologically based kinetic model development. Front, Physiol, p. 13.

Scanes, C.G., Witt, J., Ebeling, M., Schaller, S., Baier, V., Bone, A.J., Preuss, T.G., Heckmann, D., 2022b. Quantitative comparison of avian and mammalian physiologies for parameterization of physiologically based kinetic models. Front, Physiol, p. 13.

Schneeweiss, A., Juvigny-Khenafou, N.P.D., Osakpolor, S., Scharmüller, A., Scheu, S., Schreiner, V.C., Ashauer, R., Escher, B.I., Leese, F., Schäfer, R.B., 2022. Three perspectives on the prediction of chemical effects in ecosystems. Glob. Chang. Biol. gcb.16438 https://doi.org/10.1111/gcb.16438.

Schuckink Kool, P., Kooijman, B., 2016. AmP Gallus gallus WL. version (2016/11/24).

Sherborne, N., Galic, N., 2020. Modeling sublethal effects of chemicals: application of a simplified dynamic energy budget model to standard Ecotoxicity data. Environ. Sci. Technol. 54, 7420–7429. https://doi.org/10.1021/acs.est.0c00140.

Sherborne, N., Galic, N., Ashauer, R., 2020. Sublethal effect modelling for environmental risk assessment of chemicals: problem definition, model variants, application and challenges. Sci. Total Environ. 745, 141027 https://doi.org/10.1016/j.scitotenv.2020.141027.

Skadhauge, E., 1981. Introduction. In: Skadhauge, E. (Ed.), Osmoregulation in Birds. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 1–12. https://doi.org/10.1007/978-3-642-81585-0_1.

Skewes, P.A., Wilson, H.R., Mather, F.B., 1988. Correlations among egg weight, Chick weight, and yolk sac weight in bobwhite quail (Colinus virginianus). Fla. Sci. 51, 159–162.

Smith, J.L., Graczyk, T.K., Cranfield, M.R., 1996. The effect of egg yolk sampling on performance parameters and reproductive indices of northern bobwhite quail (Colinus Virginianus) eggs. J. Wildl. Dis. 32, 517–520. https://doi.org/10.7589/0090-3558-32.3.517.

Teixeira, C., Kooijman, B., Zimmer, E., 2018. AmP Anas platyrhynchos. version (2018/04/19).

van der Meer, J., 2006. An introduction to dynamic energy budget (DEB) models with special emphasis on parameter estimation. J. Sea Res. Dynamic Energy Budgets in Bivalves 56, 85–102. https://doi.org/10.1016/j.seares.2006.03.001.

Vaugeois, M., Venturelli, P.A., Hummel, S.L., Forbes, V.E., 2021. A simulation-based evaluation of management actions to reduce the risk of contaminants of emerging concern (CECs) to walleye in the Great Lakes Basin. Sci. Total Environ. 768, 144326 https://doi.org/10.1016/j.scitotenv.2020.144326.

Woodard, A.E., Wilson, W.O., 1963. Egg and yolk weight of Coturnix quail (Coturnix Coturnix Japonica) in relation to position in egg sequences. Poult. Sci. 42, 544–545. https://doi.org/10.3382/ps.0420544.