OPEN ACCESS

CrossMark

Forum Paper

Prototype Biodiversity Digital Twin: honey bees in agricultural landscapes

Jürgen Groeneveld[‡], Tomas Martinovic[§], Tuomas Rossi[|], Ondrej Salamon[§], Kata Sara-aho[|], Volker Grimm[‡]

‡ Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany
§ IT4Innovations, VSB – Technical University of Ostrava, Ostrava - Poruba, Czech Republic
| CSC – IT Center for Science Ltd., Espoo, Finland

Reviewed   v 1

Abstract

Honey bees are vital to human well-being and are under multiple stresses. We need to be able to assess the viability and productivity of honey bee colonies in different landscapes and under different management and climate-change scenarios. We have developed a prototype digital twin, HONEYBEE-pDT, based on the BEEHAVE model, which simulates foraging, population dynamics and *Varroa* mite infestation of a single honey bee colony. The main input data are land-cover maps and daily weather data. We have developed the pDT for simulating large areas and have tested it for the whole of Germany. We have also developed a web-based GUI that users can use to run the pDT for specific sites. Hive weight data from hundreds of hives will be used for calibration and validation.

Keywords

pollination, biodiversity, honey bees, multiple stressors, agricultural landscapes, resilience

Introduction

Pollinators are ubiquitous in ecosystems and play a critical role in our food supply, although the risks of their decline, including to biodiversity, are not fully understood (Goulson et al. 2015). Of particular importance for crop pollination (Garibaldi et al. 2013) and wild plant biodiversity are honey bees (*Apis mellifera*; (Potts et al. (2016))). Despite being a managed species, they are severely affected by climate change, emerging parasites and diseases, modern agricultural land use and possibly inappropriate beekeeping practices. In Europe, winter colony losses have increased to nearly 20% in recent decades (Gray et al. 2022) and, in the USA, annual losses can reach 50% (Steinhauer et al. 2021).

While single stressors, such as modern pesticides, may play an important role, the current general consensus is that the combination of multiple stressors impairs the resilience of honey bee colonies. Even if each stressor has no detectable effect at the colony level, their combination can lead to colony mortality (Henry et al. 2017). However, empirically quantifying the effects of stressors and their combination on honey bees is challenging. Bee colonies, even from the same apiary, show large variation in behaviour, which would require a large number of replications. In addition, most stressors, such as extreme weather, gaps in forage availability or parasites and pathogens, are virtually impossible to control.

Numerous simulation models have, therefore, been developed to support and extrapolate empirical research (Becher et al. 2013, Chen et al. 2021, European Food Safety Authority et al. 2021), but so far, only one of these, BEEHAVE (Becher et al. 2014), appears to be both available and able to link within-hive dynamics with foraging in a dynamic agricultural landscape (European Food Safety Authority et al. 2021).

BEEHAVE is a typical high-resolution ecological model: it has a relatively small spatial extent. It represents only the landscape around a single hive, i.e. 5 x 5 km². As such, it cannot be used to assess the status of honey bees and their habitat across regions, countries or beyond. Existing workflows for BEEHAVE rely on maps of fields and crops in the surrounding landscape, which are rarely available, as are data to test model predictions of colony performance. BEEHAVE has been used in more than 25 studies (Suppl. material 1), but its use to support policy development at national or European level has been limited. Such policies include important aspects of the Common Agricultural Policy (CAP) of the European Communities. To support the development of such policies, but also to assist farmers and beekeepers and their associations in developing sustainable and biodiversity-friendly practices, it would be necessary to extend the scope and predictive power of BEEHAVE towards a Digital Twin (DT), taking into account the specific challenges of developing a DT for biodiversity conservation (de Koning et al. 2023). The Digital Twin allows us to apply BEEHAVE in a consistent way from a local site-specific application, to regional up to national extent.

Objectives

As a first step, a prototype DT, HONEYBEE-pDT, was developed to enable the automated application of the BEEHAVE model for the whole of Germany. This includes two types of applications. First, to produce maps of Germany that visualise, for example, the number of adult bees before winter or the amount of honey that has been produced during a year. For such maps, we have run the HONEYBEE-pDT on a raster with a resolution of 5 km on the EuroHPC supercomputer LUMI (see Performance section). Second, to run BEEHAVE for specific hive locations. Users only need to specify the coordinates of the hive, but they can also modify the model parameters and the parameters of the floral resources. This user execution of HONEYBEE-pDT is possible via a web interface on a cloud environment (http s://app.biodt.eu, see Interface and Outputs section). The pDT can also be used for education and training in sustainable practices.

Workflow

Fig. 1 provides an overview of the HONEYBEE-pDT. The user GUI is implemented as an R Shiny application and the workflows are being prepared for execution using LEXIS (Large-scale Execution for Industry & Society, Golasowski et al. (2022)). Scripts have been developed to specify the input data (drivers: land-cover data and weather data) and to transform the input data into input files that can be read into the BEEHAVE simulation model. The input of weather data is done using the R package *rdwd*. The simulation experiments are also specified and executed by an R script using the *nlrx* package (Salecker et al. 2019). The execution of the HONEYBEE-pDT has been parallelized to take advantage of high-performance computing capabilities as described in the Performance section. The HONEYBEE-pDT can be applied to other countries where data on land cover, conversion of land-use type to nectar and pollen resources and weather data are available.
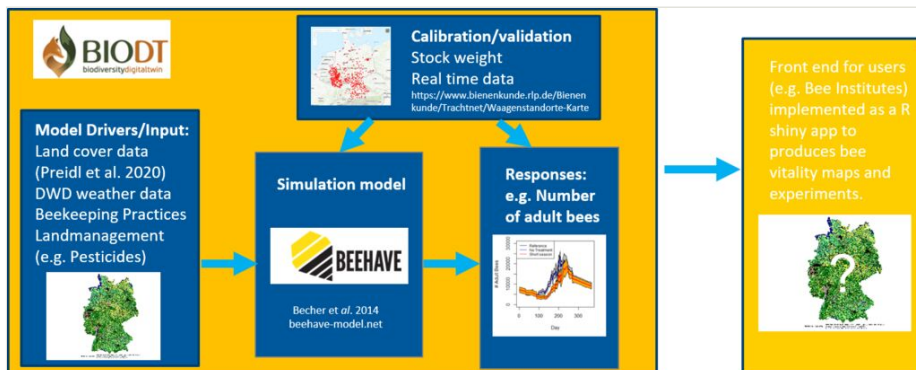

Figure 1. doi

Overview of the prototype HONEYBEE-pDT (see text for details).

Data

The pDT requires land-cover data, weather data and the specification of model parameters and flower resource parameters. In the pDT, the land-cover data are based on a map by Preidl and colleagues (Preidl et al. 2020), which provides information on 19 different land-cover classes, for example, crops such as oilseed rape or grassland. The data are freely available on the Pangea server (https://doi.org/10.1594/PANGAEA.910837). The data come as raster data and need to be converted into polygons for our application. We use the R package terra to manipulate the land-use data (https://cran.r-project.org/web/packages/terra/index.html). The conversion of land-cover types into floral resources is done by a look-up table that can be specified by the user; default values will be provided, based on previous BEEHAVE applications. We request weather data using an API provided by the R package rdwd (https://cran.r-project.org/web/packages/rdwd/index.html). Daily sunshine hours and daily maximum temperatures from the nearest DWD weather station are requested and converted into daily foraging hours. The weather data are freely available. There are data gaps in the DWD data, so we plan to replace the DWD data input with another product using the building block to download data from the Copernicus platform (https://cds.climate.copernicus.eu/). Other input options, such as beekeeping practices, can be customised by the user. In addition to the input data, it is planned to use monitoring data from the TrachNet project (Otten and Berg 2018, Johannesen et al. 2022), where weight changes of more than 500 hives in Germany are recorded These data will be used for calibration and validation. Currently, the data can be accessed by anyone via a web interface (https://www.bienenkunde.rlp.de/Bienenkunde/Trachtnet/Waagenstandorte-Karte). Access through this web interface is not feasible within this project, as it would require a manual download. The host of the data has provided us with the full dataset. We plan to develop a workflow to request subsets of these data. The automatic download procedure will be used internally in the beginning, but it is intended to make the data and data requests available to everyone.

So far, HONEYBEE-pDT is limited to Germany, but the workflows can be applied to other countries if the relevant data, such as land-cover maps, are available.

Model

BEEHAVE (Becher et al. 2014) is a simulation model implemented in NetLogo (Wilensky 1999) and is freely available (https://beehave-model.net). BEEHAVE consists of three modules: colony, foraging and mite module. The colony module runs with daily time steps. It describes age cohorts of larvae, worker bees and drones. These dynamics are driven by the daily egg laying rate of the queen, which is imposed by a hump-shaped distribution with a maximum in early summer.

The foraging module is agent-based, with one agent representing 100 bees. It simulates the foraging behaviour of bees, including scouting for new rewarding floral resources in the landscape and recruiting foragers via a waggle dance that communicates the foraging efficiency of particular flower fields. Foragers collect nectar and pollen in the given

landscape, but only when the weather permits. The temporal resolution of the foraging module is implicit and takes into account flight and handling time in seconds.

The mite model represents the dynamics of the *Varroa* mite population in the hive. Mites can be either inside the brood cells or phoretic, i.e. attached to an adult bee. Mites transmit viruses that increase the mortality of infected larvae or adult bees. The mite module includes optional control measures, such as treatment with acaricides. Other optional beekeeping practices include honey harvesting and swarm control.

BEEHAVE can be run with stylised settings for theoretical studies, i.e. all floral resources in the landscape are represented by two resource patches not representing a real landscape. Resource patches are the model entities describing areas with floral resources (e.g. fields or meadows) that are characterised by their size, distance to the honey bee colony and amount of nectar and pollen. However, it is also possible to import land-cover and weather data for specific locations and years. The landscape is represented as a list of fields, or patches, that provide nectar and/or pollen sooner or later in the year. Each patch is characterised by its distance from the hive, the likelihood of detection by foragers, the flowering period, the nectar and pollen supply and the handling time for the bees. The latter increases with increasing use of the patch, i.e. the foraging efficiency, for example, the resources of a patch can change over the course of a day. Weather data on temperature and rainfall are converted into the number of foraging hours per day, as bees do not forage in rain and low temperatures. BEEHAVE comes with example datasets for a landscape in England. The input file for BEEHAVE is a text file that can be created manually or by using the software tool BEESCOUT (Becher et al. 2016). The BEEHAVE implementation Beehave_BeeMapp2015 (https://beehave-model.net) includes additional features for setting up the model; this is the version used for the digital twin prototype presented here.

BEEHAVE was implemented in NetLogo (Wilensky 1999), a software platform and programming language based on Java and Scala. NetLogo is specifically designed for implementing agent-based models and provides tools for assembling a graphical user interface (GUI). Both BEEHAVE and NetLogo are freely available on the Internet and run on all major operating systems. BEEHAVE comes with detailed documentation of the model in ODD (Overview, Design concepts, Details) format (Grimm et al. 2020) and its code, as well as a tutorial and user manual. It has been used in more than 25 studies (Suppl. material 1).

Fig. 2 provides an overview of the main model components of BEEHAVE: foraging, demographics of honey bees and *Varroa* mites. Please note that the user of the pDT will not interact with BEEHAVE directly, but through the developed GUI.

FAIRness

The BEEHAVE model is well documented and freely accessible. The BEEHAVE version used and all developed scripts are published as open source on the BioDT GitHub repository (https://github.com/BioDT). All input data are freely available (see Data for

details). Currently, the model outcomes of the HONEYBEE-pDT produced in the web GUI will not be stored long term and the GUI user is responsible for the data.
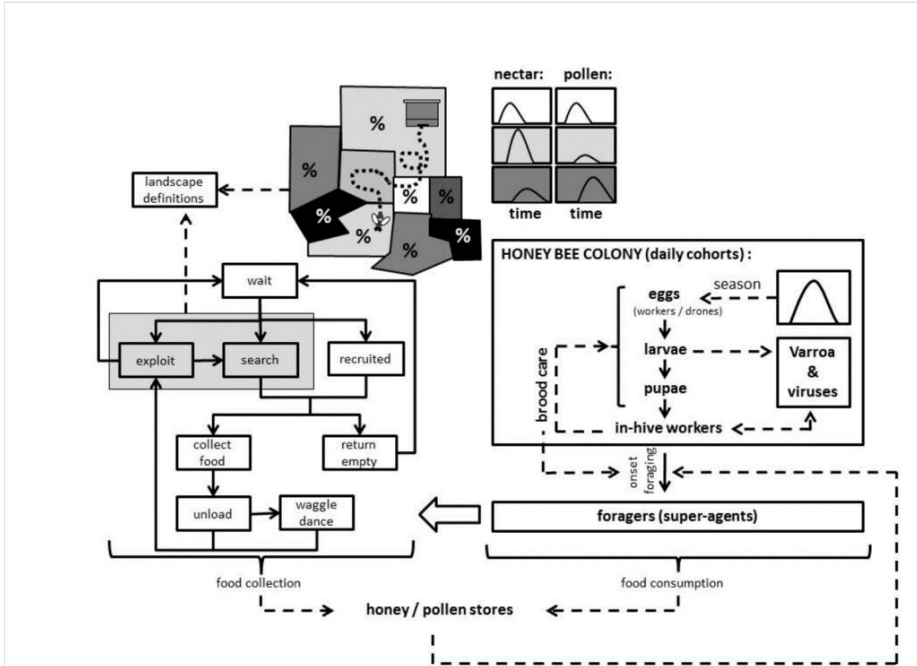

Figure 2. doi

Overview of the BEEHAVE model from the model description (ODD protocol available at https://beehave-model.net).

Performance

The simulation experiments can be specified and executed by R scripts using the *nlrx* package (Salecker et al. 2019). The software required for executing the model (NetLogo, Java, R with required packages) have been bundled in a Docker container image that can be pulled and executed on the CPU partition of the LUMI supercomputer through Apptainer/Singularity and on a cloud through Docker. The execution of the containerised BEEHAVE model has been parallelized on LUMI over individual inputs by using HyperQueue task scheduler. As an exploratory study, we used the pDT to predict the number of surviving bees and honey storage using a regular grid spanning around 3500 locations in Germany, based on the surrounding land-cover types and weather data. We ran the model for three years at each location. By utilising the developed parallelisation scheme, this calculation took about an hour on eight LUMI-C nodes. As a rough estimate, the same calculation would have taken more than a week on a regular laptop. While the run configuration on LUMI still requires optimisation for maximum efficiency, it is clear that the capability to execute the pDT in parallel over hundreds or thousands of cores and to

leverage the large computing capacity of LUMI-C is highly advantageous. The containerised solution here provides also a cleanly deployable environment for the pDT and directly enables also execution on a cloud environment for the workloads that do not need extensive computing resources, such as the current implementation in the web GUI.

Interface and Outputs

The communication between the user and the pDT is done by a R shiny web application hosted at https://app.biodt.eu. The user can vary parameters of the model and the floral resources. In addition, a location within Germany can be chosen. As outputs, the number of adult bees, honey production and flight time are visualised. A screenshot of the GUI for the site specific application is shown in Fig. 3.
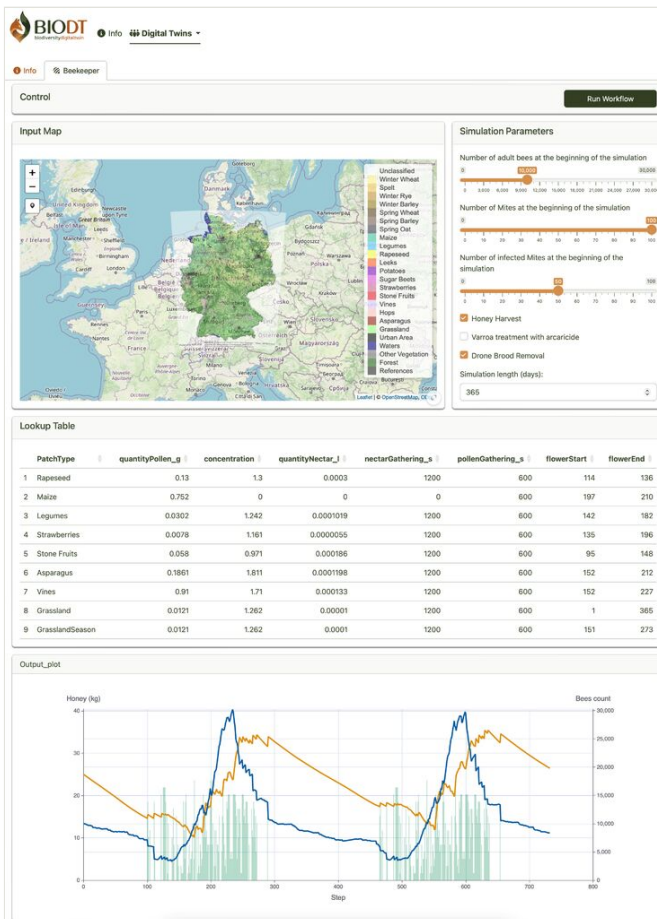

Figure 3. doi
Screenshot of a simplified GUI of the HONEYBEE-pDT.

Integration and sustainability

During the project lifetime, we already have run the pDT on different HPCs (LUMI and Karolina). Thus, in principle, the pDT can be easily migrated between computational infrastructure. One option after the end of the project is to host the HONEYBEE-pDT using resources from the Helmholtz Association to which the lead authors of this paper belong.

The HONEYBEE-pDT would benefit from links with other DT initiatives such as DestineE and EOSC, as information on extreme events, droughts and other environmental information is crucial for reliable prediction of honey bee flight and foraging behaviour. It would also be beneficial to attempt to link the HONEYBEE-pDT with DTs of vegetation DTs such as the GRASSMIND-pDT.

Application and impact

The prototype presented here, HONEYBEE-pDT, demonstrates the concept of a digital twin for supporting two important aspects of biodiversity conservation, pollination and agricultural land use. DTs are intended to support decisions in a more robust and relevant way than traditional models. Two characteristics of DTs are that

1. their inputs are regularly updated and their outputs are regularly compared to new monitoring data for calibration and validation and
2. they cover spatial scales that are relevant to stakeholders, including farmers and policy-makers.

Turning a simulation model, such as BEEHAVE, into a DT requires infrastructure and expertise far beyond what is normally available for modellers. Expertise is required to create data structures and workflows for key relevant input data, to create workflows for running BEEHAVE in parallel on a supercomputer, to containerise these workflows and the many complex software tools required and to create a professional GUI. The infrastructure required to run BEEHAVE at all relevant spatial scales was a supercomputer such as LUMI. The pDT development has been a team effort; while the modellers involved would not have had the time and expertise to create HONEYBEE-pDT on their own, the data and computer scientists involved would not have been able to take a model like BEEHAVE off the shelf and plug it into the workflows and infrastructure, as this would have required expertise in modelling and honey bee ecology. Certainly, frequent meetings and updates were needed to develop a mutual understanding of all the elements of the pDT, but the effort was well worth it, as the results and the prospect of the final, fully implemented DT far exceeded our expectations. Biodiversity modellers have always struggled with the choice between large-scale models that are too unrealistic at the local scale and small-scale models that are realistic, but too small in scale to be useful for supporting management and policy development. HONEYBEE-pDT was an important milestone in the adoption of the concept of DTs for biodiversity research, management and conservation (de Koning et al. 2023). This will enable a wide range of applications with highly relevant impacts.

HONEYBEE-pDT is aimed at different end-users. Firstly, we encourage beekeepers to simulate a virtual honey bee colony at a location of interest to them and compare the simulation results with their own experience and give us feedback. As it is difficult for academic researchers to reach the practitioners, we work closely with the German bee institutes and present at their annual meetings. We have also organised workshops and training on the BEEHAVE model to disseminate our tools. As a second target group, we have identified other researchers. At our user workshop in Leipzig in November, we realised that we need to allow them to upload customised versions of the BEEHAVE simulation model so that they can use the pDT for their work. The same goes for the third target group, industry. Companies, such as Bayer, also use BEEHAVE and may be interested in using a service such as the HONEYBEE-pDT, but they would want to use their own version of BEEHAVE, which includes a pesticide exposure and effects module (Preuss et al. 2022). In theory, pDT can also be used by national and European policy-makers to optimise CAP greening scenarios, by farmers and their associations to develop biodiversity-friendly cropping systems and pesticide use and by beekeepers and their associations to optimise beekeeping practices, in particular *Varroa* mite control.

Acknowledgements

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Becher MA, Osborne JL, Thorbek P, Kennedy PJ, Grimm V (2013) REVIEW: Towards a systems approach for understanding honeybee decline: a stocktaking and synthesis of

existing models. Journal of Applied Ecology 50 (4): 868-880. https://doi.org/10.1111/1365-2664.12112

- Becher MA, Grimm V, Thorbek P, Horn J, Kennedy PJ, Osborne JL (2014) BEEHAVE: a systems model of honeybee colony dynamics and foraging to explore multifactorial causes of colony failure. Journal of Applied Ecology 51 (2): 470-482. https://doi.org/10.1111/1365-2664.12222

- Becher MA, Grimm V, Knapp J, Horn J, Twiston-Davies G, Osborne JL (2016) BEESCOUT: A model of bee scouting behaviour and a software tool for characterizing nectar/pollen landscapes for BEEHAVE. Ecological Modelling 340: 126-133. https://doi.org/10.1016/j.ecolmodel.2016.09.013

- Chen J, De Grandi-Hoffman G, Ratti V, Kang Y (2021) Review on mathematical modeling of honeybee population dynamics. Mathematical Biosciences and Engineering 18 (6): 9606-9650. https://doi.org/10.3934/mbe.2021471

- de Koning K, Broekhuijsen J, Kühn I, Ovaskainen O, Taubert F, Endresen D, Schigel D, Grimm V (2023) Digital twins: dynamic model-data fusion for ecology. Trends in ecology & evolution 38 (10): 916-926. https://doi.org/10.1016/j.tree.2023.04.010

- European Food Safety Authority, Ippolito A, Focks A, Rundlöf M, Arce A, Marchesi M, Neri FM, Rortais A, Szentes C, Auteri D (2021) Analysis of background variability of honey bee colony size. EFSA Supporting Publications. https://doi.org/10.2903/sp.efsa.2021.EN-6518

- Garibaldi LA, Steffan-Dewenter I, Winfree R, Aizen MA, Bommarco R, Cunningham SA, Kremen C, Carvalheiro LG, Harder LD, Afik O, Bartomeus I, Benjamin F, Boreux V, Cariveau D, Chacoff NP, Dudenhoffer JH, Freitas BM, Ghazoul J, Greenleaf S, Hipolito J, Holzschuh A, Howlett B, Isaacs R, Javorek SK, Kennedy CM, Krewenka KM, Krishnan S, Mandelik Y, Mayfield MM, Motzke I, Munyuli T, Nault BA, Otieno M, Petersen J, Pisanty G, Potts SG, Rader R, Ricketts TH, Rundlof M, Seymour CL, Schuepp C, Szentgyorgyi H, Taki H, Tscharntke T, Vergara CH, Viana BF, Wanger TC, Westphal C, Williams N, Klein AM (2013) Wild Pollinators Enhance Fruit Set of Crops Regardless of Honey Bee Abundance. Science 339 (6127). https://doi.org/10.1126/science.1230200

- Golasowski M, Martinovič J, Křenek J, Slaninová K, Levrier M, Harsh P, Derquennes M, Donnat F, Terzo O (2022) The LEXIS platform for distributed workflow execution and data management. In: Terzo O, Martinovic J (Eds) HPC, Big Data, and AI Convergence Towards Exascale. Taylor & Francis

- Goulson D, Nicholls E, Botías C, Rotheray EL (2015) Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. Science (New York, N.Y.) 347 (6229): 1255957. https://doi.org/10.1126/science.1255957

- Gray A, Noureddine A, Arab A, Ballis A, Brusbardis V, Douglas AB, Cadahia L, Charriere JD, Chlebo R, Coffey MF, Cornelissen B, da Costa CA, Danneels E, Danihlik J, Dobrescu C, Evans G, Fedoriak M, Forsythe I, Gregorc A, Arakelyan I, Johannesen J, Kauko L, Kristiansen P, Martikkala M, Martin-Hernandez R, Mazur E, Medina-Flores CA, Mutinelli F, Omar EM, Patalano S, Raudmets A, San Martin G, Soroker V, Stahlmann-Brown P, Stevanovic J, Uzunov A, Vejsnaes F, Williams A, Brodschneider R (2022) Honey bee colony loss rates in 37 countries using the COLOSS survey for winter 2019-2020: the combined effects of operation size, migration and queen replacement. Journal of Apicultural Research https://doi.org/10.1080/00218839.2022.2113329

- Grimm V, Railsback SF, Vincenot CE, Berger U, Gallagher C, DeAngelis DL, Edmonds B, Ge JQ, Giske J, Groeneveld J, Johnston ASA, Milles A, Nabe-Nielsen J, Polhill JG, Radchuk V, Rohwader MS, Stillman RA, Thiele JC, Ayllon D (2020) The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism. Jasss-the Journal of Artificial Societies and Social Simulation 23 (2). https://doi.org/10.18564/jasss.4259
- Henry M, Becher MA, Osborne JL, Kennedy PJ, Aupinel P, Bretagnolle V, Brun F, Grimm V, Horn J, Requier F (2017) Predictive systems models can help elucidate bee declines driven by multiple combined stressors. Apidologie 48 (3): 328-339. https://doi.org/10.1007/s13592-016-0476-0
- Johannesen J, Wöhl S, Berg S, Otten C (2022) Annual Fluctuations in Winter Colony Losses of *Apis mellifera* L. Are Predicted by Honey Flow Dynamics of the Preceding Year. Insects 13 (9). URL: https://www.mdpi.com/2075-4450/13/9/829
- Otten C, Berg S (2018) TrachtNet 5.0—Die Bundesweite elektronische Trachtbeobachtung ist für jeden nutzbar. Bienen Nat..
- Potts SG, Imperatriz-Fonseca V, Ngo HT, Aizen MA, Biesmeijer JC, Breeze TD, Dicks LV, Garibaldi LA, Hill R, Settele J, Vanbergen AJ (2016) Safeguarding pollinators and their values to human well-being. Nature 540 (7632): 220-229. https://doi.org/10.1038/nature20588
- Preidl S, Lange M, Doktor D (2020) Introducing APiC for regionalised land cover mapping on the national scale using Sentinel-2A imagery. Remote Sensing of Environment 240 https://doi.org/10.1016/j.rse.2020.111673
- Preuss TG, Agatz A, Goussen B, Roeben V, Rumkee J, Zakharova L, Thorbek P (2022) The BEEHAVEecotox Model—Integrating a Mechanistic Effect Module into the Honeybee Colony Model. Environmental Toxicology and Chemistry 41 (11): 2870-2882. https://doi.org/10.1002/etc.5467
- Salecker J, Sciaini M, Meyer KM, Wiegand K (2019) The nlrx r package: A next-generation framework for reproducible NetLogo model analyses. Methods in Ecology and Evolution 10 (11): 1854-1863. https://doi.org/10.1111/2041-210X.13286
- Steinhauer N, VanEngelsdorp D, Saegerman C (2021) Prioritizing changes in management practices associated with reduced winter honey bee colony losses for US beekeepers. Science of the Total Environment 753 https://doi.org/10.1016/j.scitotenv.2020.141629
- Wilensky U (1999) NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-based Modeling. Northwestern University. Evanston, IL.

Supplementary material

Suppl. material 1: Supplementary Table 1 doi