Honey bee performance in Germany - towards a Digital Twin prototype

Jürgen Groeneveld [1], Tuomas Rossi [2], Tomas Martinovic [3], Volker Grimm [1]
1. Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany
2. CSC – IT Center for Science Ltd., Espoo, Finland
3. IT4Innovations, VSB - Technical University of Ostrava

Introduction

Honey bees (*Apis mellifera*) provide important services by producing honey and pollinating commercial crops but also a wide range of wild flowers. Therefore, honey bees contribute to diversity but also rely to some extent on diverse landscapes, since bees depend on the continuous provision of nectar and pollen in particular in times when the mass flowering crops (e.g. oil seed rape) are not available. Honey bees face multiple stressors such as intensive agriculture, diseases and temporal limitation in nectar and pollen supply. Digital Twin applications can help to assess the quality of a given landscape for honey bee performance and honey production. Such a Digital Twin is presented in Figure 1. Freely accessible land use data, weather data and beekeeping regimes drive the open source simulation model BEEHAVE which can be calibrated by stock weight data.

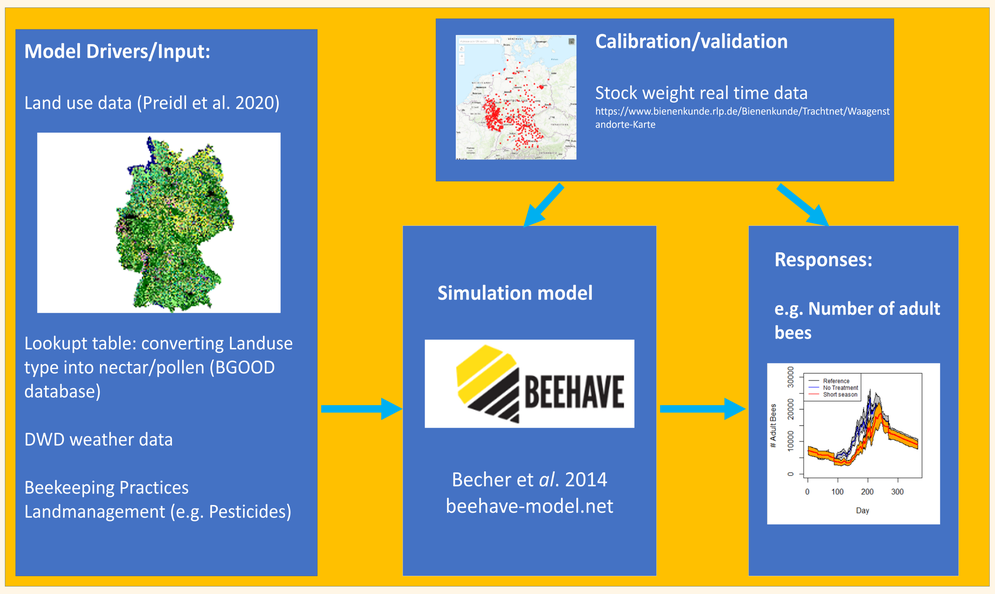

Figure 1: Summary of the envisioned Digital Twin Application

Technical description

R scripts have been developed to specify the input data (drivers) and to transform the input data into input files that can be read in to the simulation model. The simulation experiments will be specified and executed by a R script as well using the nlrx package. The software required for executing the model (NetLogo, Java, R with nlrx and other packages) have been bundled in a Docker container image that can be pulled and executed on LUMI through Apptainer / Singularity and on a cloud through Docker. The main performance bottleneck of the use case is the large number of input data that requires processing. To overcome this, the execution of the containerised BEEHAVE model has been parallelised on LUMI-C over individual inputs by using HyperQueue. This way the use case can be executed in parallel on hundreds or thousands of cores, leveraging the large computing capacity of LUMI-C.

The codes are available in BioDT GitHub (invitation needed for now):
https://github.com/BioDT/uc-beehave-singularity-for-lumi
https://github.com/BioDT/uc-beehave-execution-scripts

The Digital Twin can be applied to other countries, provided that data on land use, conversion of land use type into nectar and pollen resources, and weather data are available.

Outputs

Having the described workflow established it allows the user to run the Digital Twin application for each point in Germany (Fig. 2). The computational infrastructure of LUMI allows us to compute a map for Germany using roughly 3500 points (we have plotted the number of honey bees at the end of three years). This setting allows also computational demanding model calibration and analysis.

Questions: Where to put the digital twin and the result data? Any common design principles between projects? How to proceed with FAIR result data and software handling?

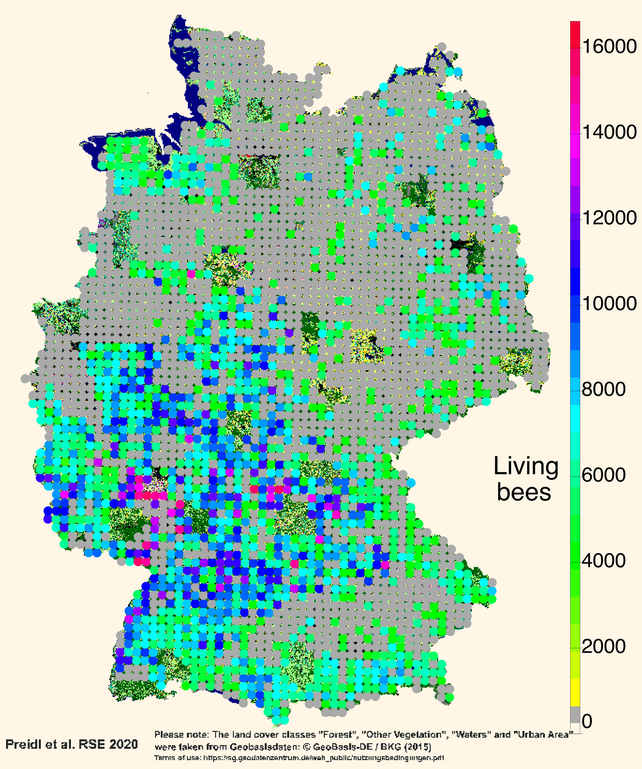

Figure 2: Example results to test the workflow. Calibration and validation are the next steps on the way to a Digital Twin Prototype that will be able to produce informative results for the end users.

Funded by the European Union

@BiodiversityDT | /company/biodt | BioDT | /communities/biodt
BIODT biodiversitydigitaltwin