MUST-B & IRC data eploration

Matthias Spangenberg

## 0.1 Motivation

BEEHAVE would benefit from a set of default scenarios representing foraging conditions within Europe, think for example of 10-15 scenarios for locations along a north-south gradient. Needed for the development are scenario-specific 1. nectar and pollen availability for each scenario, with a daily resolution, 2. numbers of foraging hours per day and 3. a matching egg-laying rate. The number of foraging hours per day can be derived from temperatures and irradiation, both preferably with an one-hour resolution. Egg-laying can either be estimated from region-specific bee observations, or probably also from a combination of pollen availability and foraging conditions.  
We would like to discuss the suitability of the available MUST-B project colony performance records and how to derive location specific foraging hours from IRC data. In particular, **feedback** about the recorded bee colony performance measurements would be very appreciated. We have the impression that measurement uncertainty masks potential colony dynamics for many recorded parameters in the MUST-B data. For example, colony weight shows clear trends, but we do not see such trends for the number of eggs (Colony performance figures below). Are eggs expected to show less pronounced trends, or is it indeed measurement uncertainty?

Another question is whether you see more potential in the data than we see, or if you are aware of additional data which are available. And maybe you have an idea what else these data could be good for (e.g. refinement of some aspects of the BEEHAVE model).

## 0.2 Data availability

The MUST-B project data was downloaded from Zenodo <https://zenodo.org/record/4953762#.YYQmzJso_s2>. Floral resource and honeybee data were collected at six locations, of which 4 are in Denmark and 2 are in Portugal, in 2019 and 2020. Floral resource availability was recorded for polygons within 1.5 or 3 km distance to the locations repeatedly. Recorded colony performance measures were number of beebread cells, worker brood cells larvae, other cells, eggs, nectar cells, honey cells, adult bees, departing foragers, returning foragers and colony weight.  
The EFSA report about the MUST-B project indicates that also temperatures and irradiaten were recorded for the locations, but these data were not presented in the Zenodo data set.

Since GPS positions of MUST-B locations were given, we downloaded temperature and irradiation data with an hourly resolution from <https://re.jrc.ec.europa.eu/pvg_tools/en/#TMY>.

## 0.3 Small analysis

As a starter, we aimed to derive location-specific nectar and pollen availability for only one location, namely *Hinnerup* in Denmark. We hoped to get at least a qualitative impression about the resource availability over the course of the study years.

In the MUST-B data, resource availability was recorded by transect walks, and per-polygon records were classified. Furthermore, vegetation types per polygon were “Herb”, “Shrub” or “Tree”. For example, for Herbs, class 3 is defined as: “covers at least 75% of the area and flowers are abundant”, and class 1 is: “covers less than 25% and/or flowers are not abundant”. In my opinion, these classifications are too vague to derive quantitative measures from them, as required for the parameterization of BEEHAVE flower patches. Furthermore, even when within-class records are on an ordinal scale, comparisons between vegetation types are still impossible.  
As an example, we present the total area of all polygons, separated by resource availability classes, for the “Herb” vegetation type (Fig. 1.

![Figure 1: Resource availability of the “Herb” vegetation type of the Hinnerup location. We present the area sum of all polygons with the same floral abundance class (1=low, 2=medium, or 3=high). Interpretation: we do not see clear trends within each floral abundance class, nor do we see potential to tun these findings into quantitative estimates for BEEHAVE flower patch parameterization. Do you agree?](data:application/pdf;base64,)

Figure 1: Resource availability of the “Herb” vegetation type of the Hinnerup location. We present the area sum of all polygons with the same floral abundance class (1=low, 2=medium, or 3=high). Interpretation: we do not see clear trends within each floral abundance class, nor do we see potential to tun these findings into quantitative estimates for BEEHAVE flower patch parameterization. Do you agree?

Daily foraging hours for the “Hinnerup” locations were derived from IRC data, using “typical meteorological year” data, from 2013. We considered this as suitable for this small analysis, but would of course go for more recent data for the actual scenario development. Notably the last recent available record was from 2016, maybe it is possible to get newer IRC data from other sources? All considered all hours with a temperature below 12 °C and a global radiation below 200 W/m² as not suitable for bee foraging, and summed up the remaining hours to derive the daily foraging hours.

![Figure 2: Daily foraging hours for a typical meteorological year (tmy) for the Hinnerup location ,Denmark. Foraging hours are all hours with a temperature greater than or equal to 12 degree Celsius and a global irradiation greater than 200 W/m^2.](data:application/pdf;base64,)

Figure 2: Daily foraging hours for a typical meteorological year (tmy) for the Hinnerup location ,Denmark. Foraging hours are all hours with a temperature greater than or equal to 12 degree Celsius and a global irradiation greater than 200 .

## 0.4 Colony performance

We want to quickly produce some bee performance plots from the MUST-B observations. We choose hives from the Hinnerup region, Denmark, for simplicity here. To explore all data, please use the shiny app in the shine\_app folder (works only from within RStudio).

library(ggplot2)  
library(dplyr)  
  
perf <- read.csv("../data/raw-data/Table\_V\_Colony\_inspection.csv")  
  
perf$site\_name <- newbee::get\_site\_names(perf)  
perf <- perf[perf$site\_name == "Hinnerup", ]  
  
perf$day\_of\_year <- newbee::calculate\_day\_of\_year(perf$sampM, perf$sampD)  
  
# remove colony records from the year 1970   
perf <- perf[perf$sampY > 1970, ]   
  
outcomes <- unique(perf$paramText)  
  
for(OUTCOME in outcomes){  
   
 df <- perf[perf$paramText == OUTCOME, ]  
   
 df <- df %>%  
 group\_by(siteNo, hiveNo, sampY, day\_of\_year, paramText, resUnit\_desc) %>%  
 summarize(max = max(resVal))  
   
 print(ggplot(data=df, aes(x = day\_of\_year, y = max, col = factor(hiveNo))) +  
 facet\_grid(rows = vars(sampY), scales = "free") +  
 geom\_point()+  
 geom\_line(alpha = 0.2, lwd = 0.8)+  
 xlim(0, 365)+  
 xlab("Day of the year")+  
 ylab(paste0( OUTCOME, " [", unique(df$resUnit\_desc), "]"))+  
 ggtitle(OUTCOME)+  
 theme\_classic()+  
 theme(legend.position="bottom"))  
}

