RESEARCH ARTICLE

Modelling daily weight variation in honey bee hives

Karina Arias-Calluari[1,2☯]*, **Theotime Colin**[2,3☯], **Tanya Latty**[2], **Mary Myerscough**[1], **Eduardo G. Altmann**[1]

1 School of Mathematics and Statistics, The University of Sydney, Sydney, New South Wales, Australia, **2** School of Life and Environmental Sciences, The University of Sydney, Sydney, New South Wales, Australia, **3** School of Natural Sciences, Macquarie University, Sydney, New South Wales, Australia

☯ These authors contributed equally to this work.
* karina.ariascalluari@sydney.edu.au

Abstract

A quantitative understanding of the dynamics of bee colonies is important to support global efforts to improve bee health and enhance pollination services. Traditional approaches focus either on theoretical models or data-centred statistical analyses. Here we argue that the combination of these two approaches is essential to obtain interpretable information on the state of bee colonies and show how this can be achieved in the case of time series of intra-day weight variation. We model how the foraging and food processing activities of bees affect global hive weight through a set of ordinary differential equations and show how to estimate the parameters of this model from measurements on a single day. Our analysis of 10 hives at different times shows that the estimation of crucial indicators of the health of honey bee colonies are statistically reliable and fall in ranges compatible with previously reported results. The crucial indicators, which include the amount of food collected (foraging success) and the number of active foragers, may be used to develop early warning indicators of colony failure.

Author summary

Honey bees are under threat and dying at an alarming rate due to pesticides, parasites, and other stressors. Obtaining information about the health of bee colonies is essential to understand how this happens and to identify measures that can prevent this from happening. Herein, we built a mathematical model of the daily dynamics of hives that allows such information to be extracted without detailed and expensive measurements. Based only on measurements of how the weight of the hive changes during the day, our model can be used to estimate how many bees are collecting food, how successful they are, and how much time they spend outside the hive. Due to its simplicity, the model presented here can be applied to a wide range of hive scale systems and help beekeepers track how healthy and productive their bees are.

1 Introduction

There is growing concern that crop pollination by honey bees might no longer be a sustainable operation in the near future [1]. Single and combined stress from pesticides, parasites, and diseases have led to an increase in bee colony mortality worldwide [2–7]. Despite beekeepers' efforts to control pests and diseases, annual colony losses of about 30% are frequently reported in countries across Europe [8] and North America [9–11]. Automated methods to track the status of hives can play a key role in helping beekeepers prevent the death of colonies [12–16].

The social organization of a bee colony is complex and relies on positive and negative feedback loops to maintain homeostasis. Feedback loops can help buffer the effect of stressors or cause snowballing effects which in some cases precipitate colony death. A well-known feedback loop that has been hypothesized to accelerate the failure of the colony is the one that regulates the ontogeny of foraging [17]. In a bee colony, workers typically spend the first few days of their adult life performing various maintenance, defence, and care tasks before transitioning into foragers [18]. The presence of a sufficiently large number of foragers inhibits the transition of younger bees into foragers [17, 19]. Under ideal conditions, this ensures enough bees are performing in-hive tasks, and that the colony is able to process all the food brought back by the foragers and to raise the next generation of workers. If an environmental stress reduces the number of foragers, then this same mechanism ensures the rapid replacement of the foraging force. With fewer foragers, the transition of younger bees into foragers is increased and the foragers' population is rapidly replenished. Under sustained stress, it has been hypothesized that this buffering mechanism can lead to generations of workers starting to forage at a younger age, which reduces their life expectancy and the foragers' population increasingly fast until colony death [20]. Thus, information about the number of forager bees and their activities can provide important guidance about the health of the hive.

Directly measuring the number of forager bees is difficult, but the use of new technologies provides opportunities to estimate them from more accessible data [12, 21]. Recently developed affordable and precise balances can easily be fitted to hives in the field to study the long-term weight changes [22] and shorter within-day weight variations of hives [12, 23]. Within-day hive weight variations are clearly visible and exhibit similar patterns that are consistent across multiple day as shown in Fig 1. In the morning, rapid weight loss occurs because foragers leave the hive [24]. For this reason, previous methods have focused on measuring the slope and low point of the morning departure segment of daily hive weight's curves [24], and have used this measure as likely representing forager activity [25]. However, these phenomenological methods provide limited information about the daily internal dynamics of the hive. Other methods perform piece-wise linear regression analysis where up to five breaking points are suggested after an abrupt change of slope [24, 26]. In particular, these methods do not provide quantitative estimations of the relevant bee activities such as the number of forager bees or the amount of food brought to the hive because the different bee activities are not explicitly modeled. In this paper, we develop a mechanistic mathematical model of the daily dynamics of hives which, in combination with hive weight time series, provides quantitative estimations of key indicators of the state of the hive such as the total number of active forager bees and the average mass of food they bring to the hive in each trip. We start in Sec. 2.1 with our model presentation, its qualitative description, and its main properties. In Sec. 2.2 we present how the parameters can be inferred from data. In Sec. 2.3 we apply our methodology to time series of ten different hives measured across a year and show how in most cases the results lead to biologically meaningful estimations. Finally, in Sec. 3 we summarize our conclusions. A detailed description of our data and computational methods is presented in the Materials and Method section. The data and codes used in our analysis are available in a repository [27].
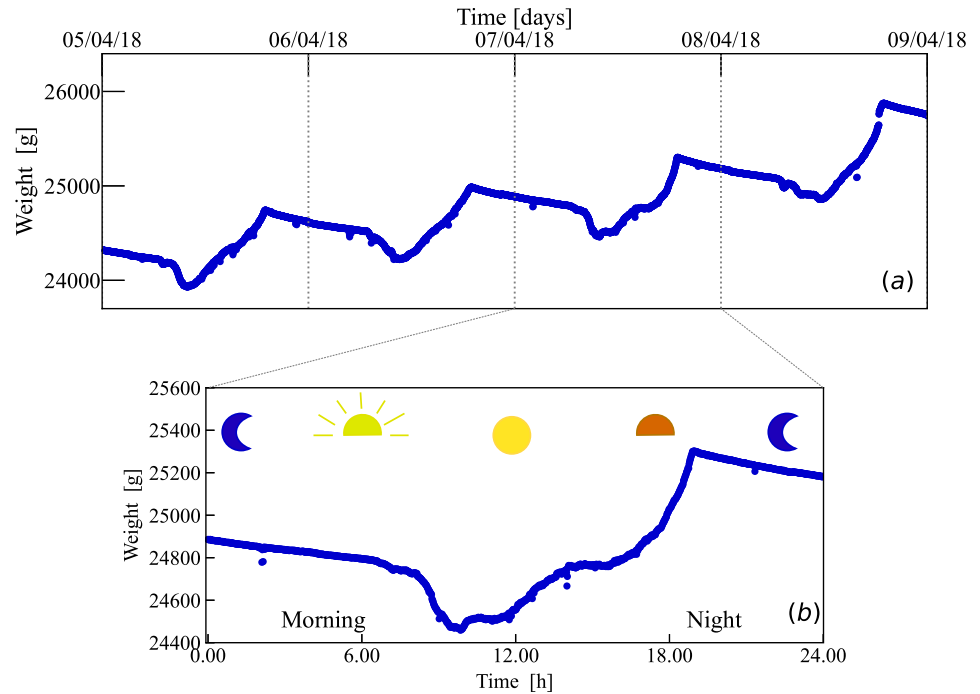
Fig 1. (a) The weight of a bee hive shows systematic variations throughout a day. The blue dots show the weight of a hive measured every minute for four consecutive days. (b) Each daily hive weight's curve present consistent patterns at specific time ranges of morning, afternoon, evening and night time.

https://doi.org/10.1371/journal.pcbi.1010880.g001

2 Results

2.1 Model

Our aim is to obtain a mechanistic model that contains the essential processes necessary to describe how the weight W of a bee hive changes throughout time t over a daily cycle.

Qualitative description. Our model will incorporate the following well-known processes and honey-bee behaviours that are also schematically illustrated in Fig 2:

1. In the morning, the weight of the hive rapidly falls as foragers leave the hive [24]. This initial weight drop usually lasts for a few hours, indicating that the weight loss is only partly counterbalanced by forager bees returning to the hive. A successful forager will return with pollen, nectar, water or propolis. The weight of a forager bee is expected to be higher when they return than when they leave, if the resources they seek are available in their environment.

2. Around midday, departures and arrivals reach roughly an equilibrium. Hives in environments with food resources start gaining weight, possibly indicating that the weight of departing foragers is less than the weight of effective foragers returning with food to the colony [24]. The weight increases until all foragers have returned at dusk.

3. In the late afternoon, the number of departures reduces until they stop completely around dusk. Once all the foragers have returned, the hive stops gaining weight.

4. At night, the weight of a hive declines slowly until dawn due to respiration, food consumption, and evaporation. The metabolic activity of a bee colony remains high at night because
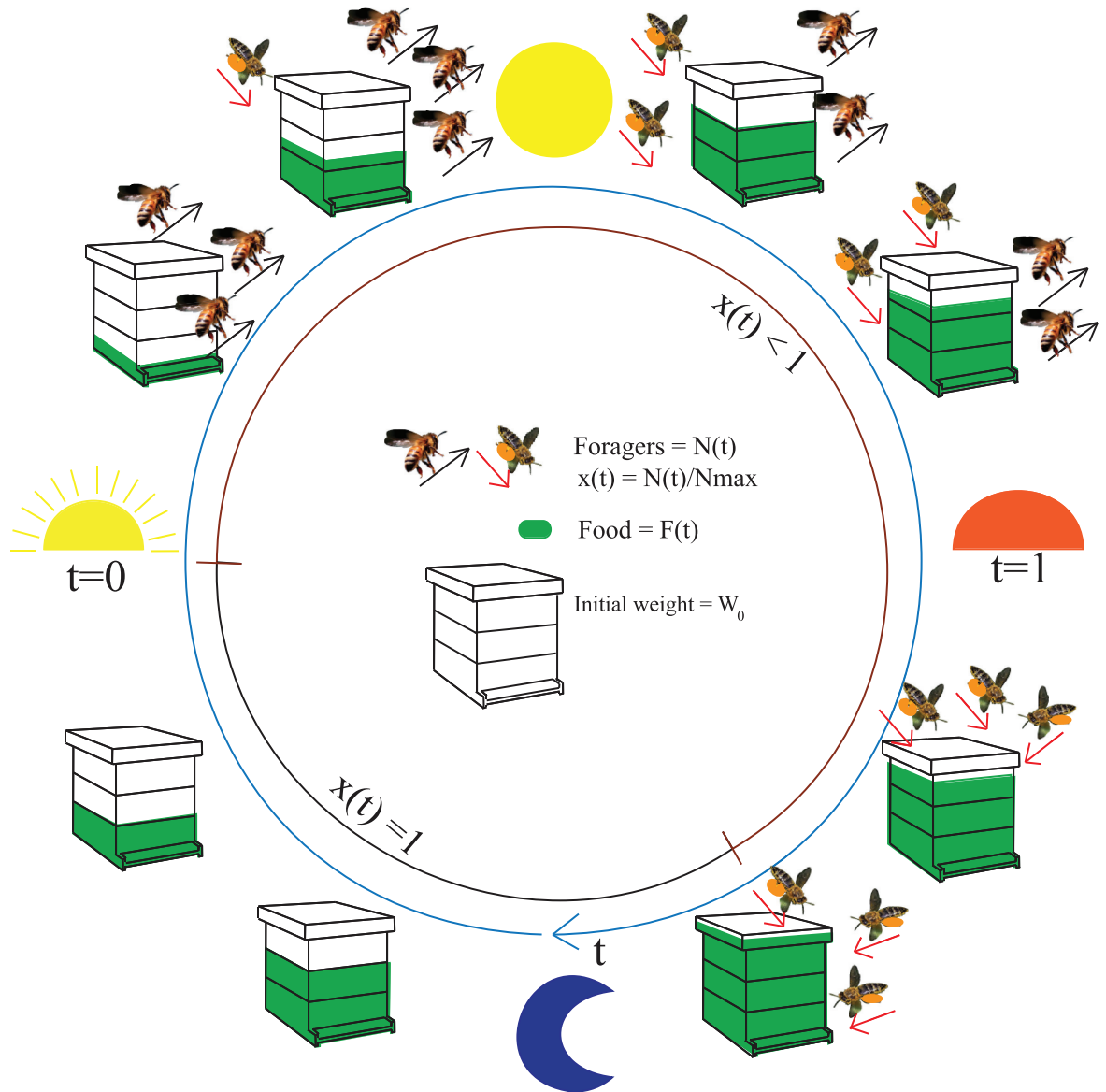
Fig 2. Schematic representation of the daily cycle of bee activities relevant to modeling colony weight, where time flows in a clockwise direction (blue line). The main time-dependent elements we model are the number of forager bees $N(t)$ and the food $F(t)$ (in green) inside the hive. The foragers are represented as active bees that depart the hive (black arrow) and return to it (red arrow) with food (in orange). The ratio of forager bees inside the hive at a particular time is represented by $x(t)$, which is constant before and after the dynamics of forager bees because the N_{max} foragers are all inside the hive. The static weight W_0 is composed of the hive structure and bees that do not leave the hive.

bees actively maintain the temperature of the brood at about 35 ˚C, as well as humidity and CO_2 concentration to optimize development. Bees actively warm the colony by shivering, or cool down and ventilate the hive by evapotranspiration and fanning. These activities require a great deal of movement and, thus, cause honey bees to consume significant amounts of honey at night. Sugars in the honey are broken down to produce energy and exhaled as CO_2. A colony also evaporates a large amount of water, either to facilitate

biological processes, concentrate sugars in food stores, or cool down the hive. Moreover, it is likely that a small amount of the weight loss at night results from cleaning activities, with bees likely getting rid of pests, dead bees and chewed brood cell caps throughout the night. The cycle resumes the next day again.

The net outcome of the activities described above depends on the hive, the environment, and the weather conditions. In favourable environments and at its peak population in summer, a single bee colony can harvest a few kilograms of food in a single day. Most of this weight comes in the form of nectar and honeydew, which generally contains more than 50% water and diluted sugars [28]. Hives in environments with little or no resources may lose more weight than they gain in a day as a consequence of honey consumption inducing respiration and evaporation (i.e., foragers expend more resources than they are able to bring back to the colony).

Mathematical model. We consider the time dependent weight $W \in \mathbb{R}^+$ (measured in g) to depend on the number of forager bees $N \in \mathbb{Z}, 0 \leq N \leq N_{max}$ (measured in bees) and on the food $F \in \mathbb{R}^+$ (measured in g) as

$$W(t) = F(t) + wN(t) + W_0, \tag{1}$$

where W_0 is the constant weight of the hive and w is the (average) weight of a bee. Next we discuss the variation in $N(t)$ and $F(t)$.

For mathematical convenience, we model the variation in the number of foragers bees $N(t)$ by focusing on the fraction $x \in \mathbb{R}$ (nondimensional) of all N_{max} foragers bees that are inside the hive:

$$x \equiv \frac{N}{N_{max}} \Rightarrow N(t) = N_{max}x(t). \tag{2}$$

We write the rate of change of x as

$$\frac{dx}{dt} = a(t)(1 - x) - d(t)x, \tag{3}$$

where $a(t)$, measured in $1/s$, is the *a*rrival rate of the fraction of forager bees outside the hive and $d(t)$, measured in $1/s$, is the *d*eparture rate of the fraction of forager bees inside the hive. The inverse of these rates can be thought of as the typical period a bee spends outside and inside the hive, respectively. These terms multiply the fraction of bees outside $(1 - x)$ and inside (x), respectively. The rate of bees arriving is

$$A(t) = a(t)(N_{max} - N(t)) = a(t)N_{max}(1 - x(t))$$

and the rate of bees departing is

$$D(t) = d(t)N(t) = d(t)N_{max}x(t),$$

where A and D are measured in *bees/s*.

The food inside the hive $F(t)$ increases due to the mass m (measured in g) of food brought in by arriving bees and decreases due to evaporation, respiration, and food consumption, which we model as a constant loss ℓ (measured in g/s) so that:

$$\frac{dF}{dt} = -\ell + mA(t) = N_{max}(ma(t)(1 - x)) - \ell. \tag{4}$$

Notice that $F(t)$ is linear in x and can be computed directly from $x(t)$ and $a(t)$ by integrating the above equation. Substituting $F(t)$ and $x(t)$ in Eq (1) we obtain the time-dependent weight $W(t)$ of the hive, our measurable quantity. This implies that the equations can be solved sequentially, starting from $x(t)$. The final step of our model is to specify the departure $d(t)$ and arrival $a(t)$ rates defined in Eq (3) and then solve for $x(t)$. We will consider a very simple choice for $a(t)$ and $d(t)$—piecewise constant in t—which allows for an explicit solutions for $x(t)$ and, thus, $W(t)$. Note that more sophisticated models of food inside the hive $F(t)$ could use a time-dependent loss $\ell(t)$ dependent on temperature, weather variations, or a term proportional to N (or x) to indicate a consumption rate proportional to the number of (forager) bees. Furthermore, more elaborate functions $a(t)$ and $d(t)$, such as delay equations, could be incorporated.

Day and night cycles. Without loss of generality, in our model, we consider:

$t = 0$ to be the time in the morning (around sunrise) at which the bees start exiting the hives.

$t = 1$ to be the time in the late afternoon (around sunset) when bees stop leaving the hive.

Therefore, in all cases below we assume $d(t) = 0$ for $t < 0$ and $t > 1$. The matching of the model time t to the clock time requires the estimation from data of two clock times: t_0 (corresponding to $t = 0$) and t_1 (corresponding to $t = 1$).

Night dynamics. Let us start with the (easiest) night case $t > 1$, when bees do not depart $d(t) = 0$ and arrive at a constant rate $a(t) = a$. Assuming that at the beginning of this regime at $t = 1$ there are x_1 bees inside, the particular solution of Eq (3) is

$$x(t) = 1 - (1 - x_1)e^{-a(t-1)}, \tag{5}$$

which corresponds to an exponential decay towards all the bees being inside the hive ($x = 1$) with a rate a. Therefore, we can consider $1/a$ proportional to the duration for which bees remain outside the hive.

Day dynamics. Let us consider the simple case where $d(t) = d$ and $a(t) = a$, both constant in time. The initial condition in this case is $x = 1$ at $t = 0$ because we assume all forager bees spent the night in the hive (i.e., $1/a \ll$ night period and all previous-day foragers return). In this case, Eq (3) leads to

$$x(t) = \frac{a}{d+a} + \frac{d}{d+a}e^{-(d+a)t}. \tag{6}$$

Explicit solutions. We are now able to combine the results above to write explicit solutions for the fraction of foragers inside the hive $x(t)$, the weight of food inside the hive $F(t)$, and the total weight of the hive $W(t)$. We consider the arrival rates during the day a_1 and at dusk a_2 to be different:

$$a(t) = \begin{cases} a_1 & \text{for } t \leq 1, \\ a_2 & \text{for } t > 1, \end{cases} \tag{7}$$

where we expect to retrieve $a_2 > a_1$ (bees spend less time out when it is getting dark).

Introducing these arrival rates and a constant departure rate d in Eqs (5) and (6) we find an explicit solution divided in three time intervals for $x(t)$:

$$x(t) = \begin{cases} 1 & \text{for } t < 0, \\ \dfrac{a_1}{d+a_1} + \dfrac{d}{d+a_1}e^{-(d+a_1)t} & \text{for } 0 \leq t \leq 1, \\ 1 - (1-x_1)e^{-a_2(t-1)} & \text{for } t > 1 \end{cases} \tag{8}$$

where $x_1 = \frac{a_1}{d+a_1} + \frac{d}{d+a_1}e^{-(d+a_1)}$ is the value of $x(t=1)$ and ensures the continuity of $x(t)$.

We can now turn to the dynamics of food inside the hive $F(t)$. From Eq (4) we obtain

$$F(t) = C + (\tilde{m} - \ell)t - \tilde{m}\int x(t)dt,$$

where $\tilde{m} = mN_{max}a$ and C is an integration constant to be fixed using the initial condition $F(0) = F_0$ and $F(1) = F_1$. The solution for $F(t)$ is then solved by integrating Eq (8), which can be done analytically in each of the three branches leading us to

$$F(t) = \begin{cases} F_0 - \ell t & \text{for } t < 0, \\ F_0 - \dfrac{\tilde{m}d}{(d+a_1)^2} + \left(\tilde{m}\left(1 - \dfrac{a_1}{d+a_1}\right) - \ell\right)t + \tilde{m}\dfrac{d}{(d+a_1)^2}e^{-(d+a_1)t} & \text{for } 0 \leq t \leq 1, \\ F_1 + \ell + \tilde{m}\dfrac{1-x_1}{a_2} - \ell t - \tilde{m}\dfrac{1-x_1}{a_2}e^{-a_2(t-1)} & \text{for } t > 1, \end{cases} \tag{9}$$

where $F_0 \equiv F(t=0)$—for this model an arbitrary constant that is incorporated into the static weight of the hive W_0 in Eq (1), consequently, without loss of generality, we take $F_0 = 0$—and $F_1 \equiv F(t=1)$ is computed setting $t=1$ in the second line of Eq (9) so that

$$F_1 = F_0 - \frac{\tilde{m}d}{(d+a_1)^2} + \left(\tilde{m}\left(1 - \frac{a_1}{a_1+d}\right) - \ell\right) + \frac{\tilde{m}d}{(d+a_1)^2}e^{-(d+a_1)}.$$

Finally, the weight of the hive $W(t)$ is obtained from Eq (1) as the sum of Eq (8) (times wN_{max}), Eq (9), and the constant weight of the hive W_0, leading to an explicit piece-wise continuous function $W(t)$ that will be compared to the measured data. A summary of the parameters used in the model is given in Table 1.

Effective description of the model solution W(t).   The mathematical model described above is divided in three regimes—$t < 0$, $0 \leq t \leq 1$, $t > 1$—with similar characteristics as shown in Fig 3. The three regimes contain segments of linear dependence of $W(t)$—a decay in the first and last regime, and typically an increase in the intermediate regime—with a transition period (shaded region) between the regimes:

$t < 0$ The first regime corresponds to all bees inside the hive ($x = 1$) and a linear decay of W on time t with rate ℓ—evaporation and respiration rate defined in Eq (4)—and with an intercept $A \equiv W(t=0)$.

$0 \leq t \leq 1$ The second regime starts with an exponential relaxation of $x(t)$ from 0 to the fixed point value $x^* \equiv \frac{a}{d+a}$, as described in Eq (6), which leads to a sharp decay of $W(t)$ until an inflection time t_c. Shortly after t_c, $W(t)$ shows a linear dependency with rate α and constant B.

Table 1. The parameters θ of our model. For simplicity, the average weight of bees w is fixed based on previously reported results. The remaining parameters are inferred from data within a prior range of admissible parameters (right column). The boundaries of the range of flat priors were chosen based on the cited references or on values smaller/ larger than reasonably possible values (e.g., the maximum weight max(W) is chosen to be the maximum total weight of the hive at all times). In all cases reported below, the inferred parameters remained away from the boundaries, indicating that our choice of boundaries has no influence on our results.

Parameter	Description	Prior Range
w	Average weight of a single bee, introduced in Eq (1).	0.113 g/bee [29, 30]
W_0	Static weight of the hive that includes the structure, food stores, non-foraging bees, etc., see Eq (1).	[5000, $max(W)$]g
N_{max}	Total number of foragers active on that day, introduced in Eq (2).	(0, 80000] bee
ℓ	Continuous loss, introduced in Eq (4).	(0, $max(W)/h$] g/h
m	Average mass of food brought by foragers, introduced in Eq (4).	(0, 0.78w] g/bee [31–33]
a_1, a_2	Rates of arrivals, see Eq (7), inversely proportional to the time bees spend out.	[0.10, 4.80] 1/h [34–36]
d	Rates of departures, see after Eq (7), inversely proportional to the time bees spend in-side the hive.	[0.81, 3.01] 1/h [34–36]
t_0, t_1	Clock time corresponding to $t = 0$ and $t = 1$, respectively.	[6.50, 9.00] h [14.5, 18.5] h

$t > 1$ The third regime shows an exponential relaxation of $x(t)$ from x^* to $x = 1$ as described in Eq (5), which leads to a sharp growth in $W(t)$ as all the bees return and before the linear decay of the first regime starts again completing the cycle.

The critical regime describing bee activities is the second regime that takes place during the day $0 \leq t \leq 1$. It can be effectively described by the 4 parameters A, α, B, and t_c (see Fig 3a), which can be rewritten in terms of the parameters of the model (see Table 1) as:

A: The total weight of the hive at $W(t = 0)$ defined as: $A = N_{max}w + W_0$

α: The rate of weight gain due to foraging, from Eq (9): $\alpha = \tilde{m}\left(1 - \frac{a_1}{d+a_1}\right) - \ell$

B: The intersection point of weight gain at $t = 0$: $B = -(\alpha + \ell)\left(\frac{w}{ma_1} + \frac{1}{d+a_1}\right) + A$

t_c: The time of the minimum weight: $t_c = \frac{-1}{(d+a_1)}\ln\left(\frac{\alpha}{dN_{max}w+\alpha+\ell}\right)$

Fig 3 shows an important mathematical property of our model: different choices of parameters lead to nearly identical curves $W(t)$ making our model effectively *underdetermined* by $W(t)$. Mathematically, this *underdetermination* happens because the *four* effective parameters introduced above (A, B, α, t_c), which provide a simplified quantitative description of the solution $W(t)$, depend on *five* model parameters (N_{max}, W_0, m, a_1, d, note that $\tilde{m} \equiv mN_{max}a$ and the mass of a bee w is taken as fixed). This means that there are different choices of the parameters θ of our model that lead to the same four effective parameters describing the second regime $W(t)$ (even if the curves $W(t)$ are mathematically distinct, their difference is indistinguishable in Fig 3 and much smaller than the experimental precision of the data). Biologically, this underdetermination appears due to our definition of which bees in the hive count as forager bees. Our model parameters are fixed and can be thought as an average over a spectrum of bees that ranges from very active to those that do only very few forager excursions in a day. Depending on whether we include the least active bees as foragers or not, we obtain different model parameters but the same effective parameters defined above. For instance, the same
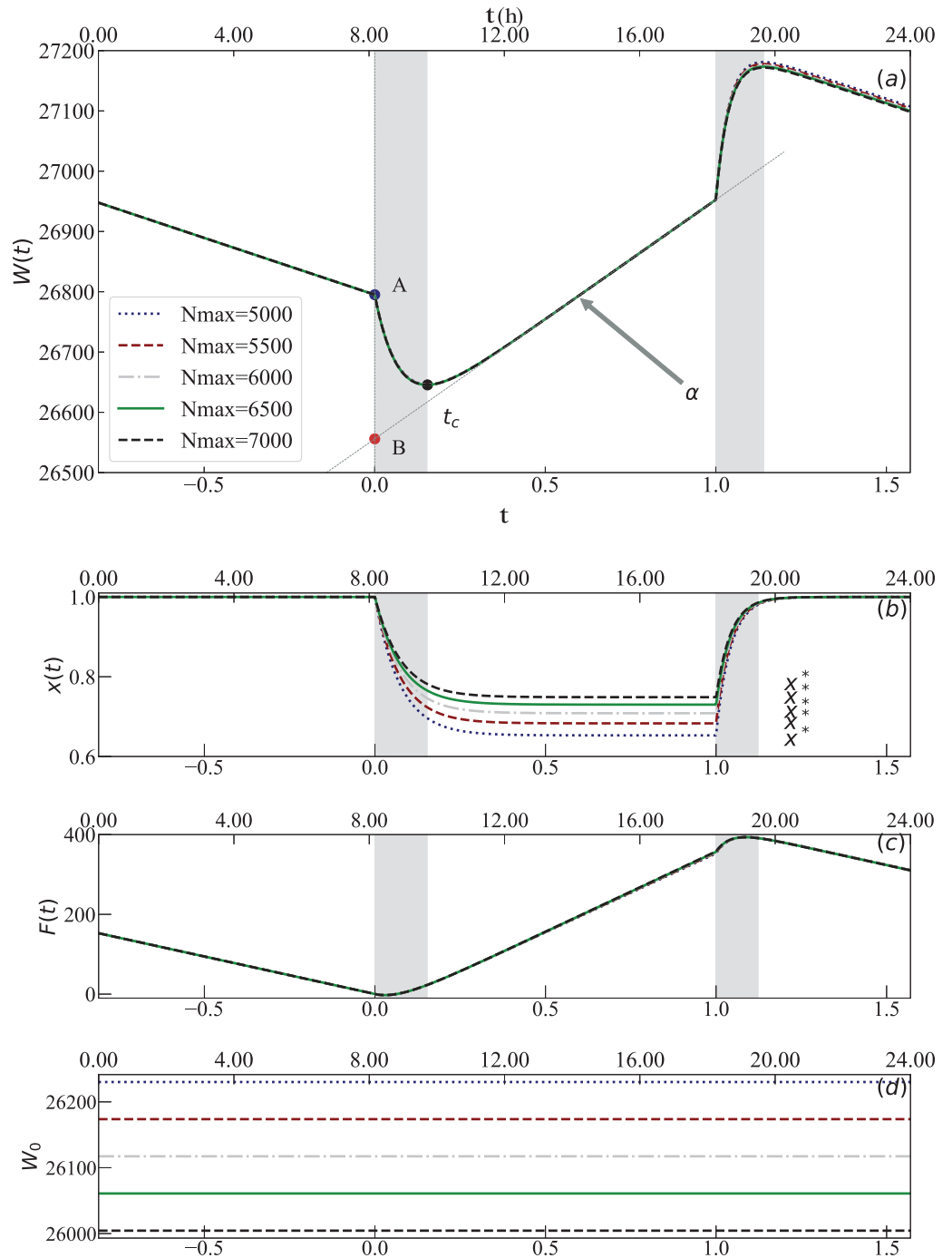
Fig 3. Graphical representation of the solution of our model. (a) The weight as a function of time, $W(t)$, in which the three linear regimes and the two transition regimes (shaded areas) are visible. (b-d) The three key elements that compose $W(t)$: (b) the number of foragers inside the hive $x(t)$ that converges to the constant value x^*; (c) the amount of food $F(t)$ inside the hive; and (d) the constant weight of the hive W_0. In all four plots, five curves are shown for different model parameters $\theta = \{N_{max}, W_0, a_1, a_2, d, m, \ell = 18.74 \ g/h, w = 0.113 \ g/bee\}$ that lead to the same effective parameters A, B, t_c, α shown in panel (a). The values of the remaining parameters in numerical order for N_{max} are $W_0[g] = 26230$, 26173, 26117, 26060, 26004; $a_1[1/h] = 0.882$, 0.914, 0.941, 0.964, 0.984; $a_2[1/h] = 2.329$, 2.328, 2.328, 2.327, 2.327; $d[1/h] = 0.469$, 0.424, 0.387, 0.356, 0.330; and $m[g/bee] = 0.038$, 0.037, 0.036, 0.035, 0.034. The two grey spans highlight the time intervals of exponential decay and growth in $x(t)$, respectively.

https://doi.org/10.1371/journal.pcbi.1010880.g003

initial weight $A \equiv W(t = 0) = N_{max}w + W_0$ can be obtained by including the less active bees as foragers—thus increasing N_{max}—or as part of the non-foragers bees that spend most of the time in the hive—thus contributing to the constant W_0. Similarly, counting the least active bees as foragers reduces the average mass m of food brought to the hive, increases the average time they spend in the hive (i.e., decreasing d), decreases the time spent outside the hive (i.e., increase a_1, a_2), and leads to a larger fraction of forager bees inside the hive (x^* increases). As shown in Fig 3, solutions of $W(t)$ with very different N_{max}, W_0 in Fig 3b and 3d, respectively—which imply also different d, a_1, a_2, and m—lead to essentially the same curves for the food F (t) in panel Fig 3c and weight $W(t)$ in Fig 3a.

2.2 Connecting model and data

Here we show how to combine our model with data to extract information about the state of hives. We describe the theoretical and computational methodology used to infer the model parameters θ from time series of hive weights $W'(t)$.

Theory.   We assume that the measured weight W' deviates from the actual weight W as

$$W' = W + \varepsilon,$$

where ε is an independent Gaussian distributed random variable with zero mean and standard deviation σ, i.e., $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, that accounts for measurement errors and random effects not accounted in our model.

If θ are the model parameters generating $W \equiv W_\theta$, at an arbitrary time t, as described in the model of the previous section, the probability likelihood of an observation W' is a Gaussian distribution centred at W_θ given by

$$P(W'|\theta) = \mathcal{N}(W_\theta; \sigma^2).$$

Assuming that observations are performed at times $q_1, q_2, \ldots q_n$ with $q_i = q_0 + n\delta q$ and are (conditionally) independent from each other, the log-likelihood of the observations is given by:

$$\log P(W'|\theta) = \sum_{i=1}^{n} \log \mathcal{N}(W_\theta(q_i); \sigma) = \sum_{i=1}^{n} -\frac{(W'(q_i) - W_\theta(q_i))^2}{2\sigma^2} - \log(\sqrt{2\pi}\sigma) \quad (10)$$

We can include further information on the parameters θ by considering a prior distribution $P(\theta)$. Here we consider flat priors for all parameters, i.e., $P(\theta) = cst.$ for $\theta \in [\theta_1, \theta_2]$ as specified in Table 1 above. Effectively, this restricts the range of admissible parameters without changing the probability of parameters within the range. The estimation of the parameters θ can be done maximizing the posterior distribution $P(\theta|D)$ which in our case is equivalent to the maximization of the log-likelihood (10) or to the minimisation of the squares of $W' - W$ as

$$L = \sum_{i=1}^{n} (W'(\tau_i) - W(\tau_i))^2 \quad (11)$$

in the range of admissible parameters set by the priors.

Computation.   The problem of obtaining the ten parameters introduced in our model— see Table 1—has been thus reduced to the optimization problem of finding the minimun of L in a ten-dimensional space. The computational method we use for this optimization is centred around the Levenberg–Marquardt algorithm as implemented in the library SciPy (See Materials and methods Sec-4.1). However, a reliable estimation of the parameters cannot be obtained directly through the naïve application of this method both because of the high-dimensionality of the problem and the effective underdetermination of our model reported at the end of last

section (which leads to essentially flat regions of L in the parameter space with multiple local minima). Here we use properties of bee hives and our model to simplify this problem and obtain a reliable estimation of parameters.

We start by considering four parameters:

w For simplicity, we fix the average weight of bees to $w = 0.113$ g/bee [29, 30].

t_0, t_1 We infer the times t_0 and t_1 in a grid of plausible values (separated by five minutes). During this minimization we do not distinguish between model parameters that lead to the same effective parameters A, B, α, and t_c because t_0 and t_1 are not involved in them. The estimation of t_0 and t_1 allow us to re-scale the data to the model time scale, facilitating the numerical estimation of the remaining parameters.

ℓ The continuous loss can be calculated directly from the negative linear decrease when $t < 0$.

These parameters, and the effective parameters (A, B, α, t_c), are not strongly affected by the underdetermination problem discussed above. The remaining six parameters—N_{max}, W_0, m, a_1, a_2, d—are directly connected to the underdetermination problem and need to be estimated with care. We start by obtaining an estimation of A, B, α, and t_c by direct minimization of L. The mapping of the effective parameters to the model parameters is done in two different ways:

(i) We seek a range of plausible model parameters determined as the values of W_0 and N_{max} that keep the four effective parameters constant and that maintain the parameters ℓ, m, a_1, a_2 and d within their prior range (set in Table 1). In practice, this is done by defining different linear combinations between N_{max} and W_0 that are introduced in the system of six equations that connect the parameters. For each possible N_{max}, we obtain a particular a_1, d, and m by solving the systems of equations given by A, B and t_c. The second rate of arrival, a_2, is obtained by the conservation of departed and returned bees. This method leads to the estimation of an *interval* of possible parameters.

(ii) Alternatively, by pre-defining one parameter, we calculate the N_{max} value directly because the system of equations presents a unique solution. We focus on the departure rate d because of the consistency among different reported studies on the time interval between trips that a forager bee spends inside the hive [37] and because d can be directly connected to the simplifying modeling decision of dividing bees into foragers and those that do not leave the hive (i.e., we consider as foragers the most active bees so that collectively their average departure rate is d, the other bees are considered to effectively stay inside the hive, and their weight is counted as part of W_0). In practice, we choose d so that the average time spent by forager bees inside the hive between trips is $\tau_d = 0.816$ h (the half-life τ is derived from the arrival or departure rate, and defined in Eq (12) below) [36].

Even if the first approach does not lead to a point estimation of the parameters, in practice most solutions can be restricted to relatively narrow intervals by considering physically-based constraints. Some of these constraints are the average mass of food carried by a single bee m, the interval of time between trips τ, and the number of trips needed to satisfy food requirements. Details on the computational methods appear in Materials and Methods Sec. 4.2 and the codes are available in the repository [27].

2.3 Results from data analysis

In this section we report and discuss the results obtained by applying the model defined in Sec. 2.1 to data using the inference methods explained in Sec. 2.2.
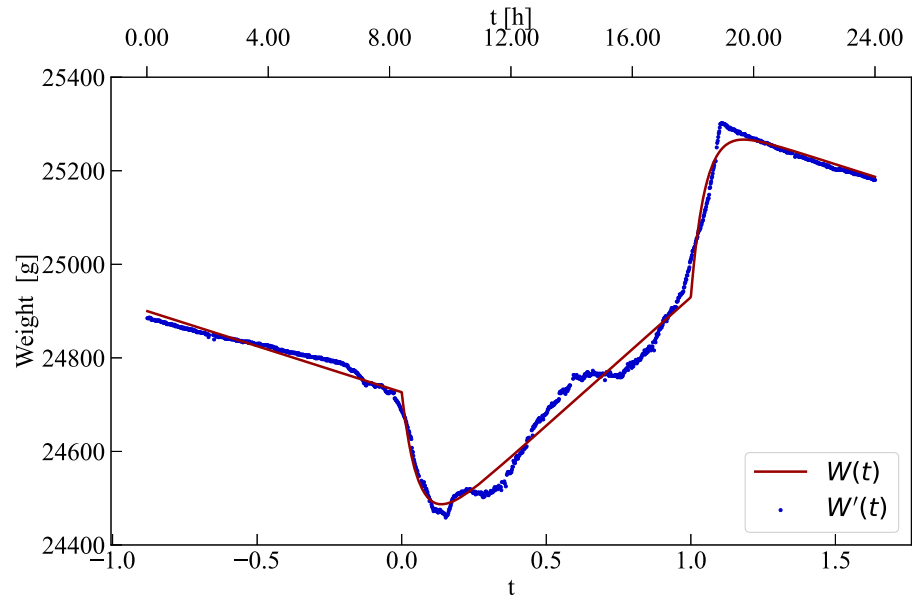
Fig 4. Comparison between the model and data. Data corresponds to measurements on 2018–4-7 for 'Hive 6'. By fixing τ_d = 49 min, the estimated parameters are N_{max} = 5830 bees, W_0 = 24068 g, m = 0.03 g/bee, a_1 = 0.99 h^{-1}, a_2 = 2.1 h^{-1}, d = 0.85 h^{-1}, l = 20.66 g/h, t_0 = 8.40 h and t_1 = 17.93 h.

Data. We used 24h time series of weights $W'(t)$ sampled every δt = 1 minute for ten different hives on different days. The data was collected between 27th of November 2017 and 2nd of March 2019, from hives located at Macquarie Park, NSW, Australia (33° 46′ 06.6″ S 151° 06′ 43.8″ E, see [34]). We show data from the 7th of April in Figs 4, 5a and 5b; additional dates appear in Fig 5c. In all cases, the daily maximum temperature was above 16°C (which is
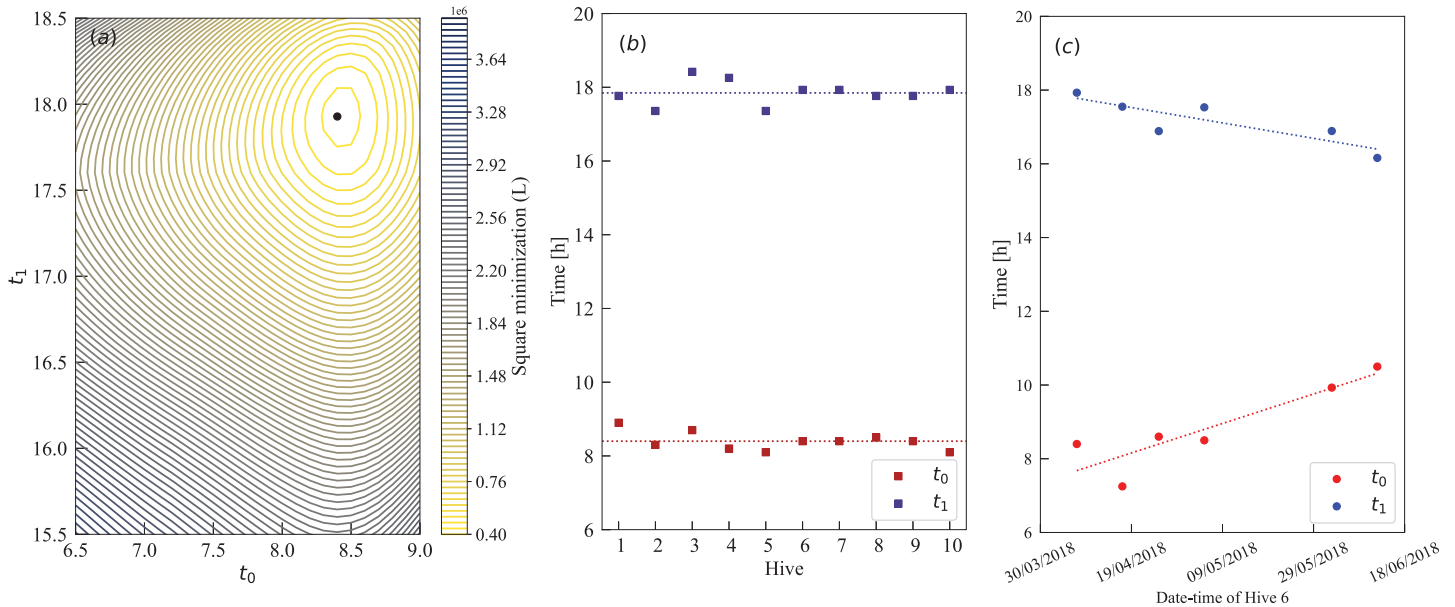
Fig 5. Inference of activity times t_0 and t_1. (a) Contour plot of L obtained performing multiple minimizations at fixed t_0 and t_1 ('Hive 6' at 2018–4-7). The minimum L is at t_0 = 8.40 h and t_1 = 17.93 h. (b) Estimated t_0 and t_1 fluctuate around similar values for ten different hives located in close proximity when analysing data sets during the same day. E.g. at 2018–4-7, the different activity times oscillate around t_0 = 8.4 h and t_1 = 17.85 h, respectively (dashed lines). (c) Time evolution for the optimum t_0 and t_1 for 'Hive 6' only. Transitioning from Autumn to Winter, we observe a linear increase for t_0 and a linear decrease for t_1 (linear regression).

sufficient for honeybees to forage [38]), and no rain was recorded (ensuring that the recorded weights are not affected by precipitation). This dataset also contains an independent estimation of the total number of bees in each of the hives, obtained from measures of the total weight of the hive with and without bees (performed two days after the day we use to infer the parameters of our model) [34]. As a pre-processing step to the analysis of weight time series, we removed outliers that correspond to exogenous influences as described in Materials and Methods Sec. 4.1.

Representative case. Fig 4 shows the comparison between our model with inferred parameters and data for one specific hive. We see how the combination of the three piece-wise continuous curves in our model (for $t < 0$, $0 \leq t \leq 1$, and $t > 1$, respectively) provides a reasonable account of the characteristic pattern of within-day hive weight variations. The main advantage of our methodology is the mechanistic generative model associated to this curve which ensures the parameters describing the curve have a relevant meaning. For instance, we estimate that on this day there were N_{max} = 5, 830 active forager bees, that they brought on average 0.03 g of food on each trip, and that trips were on average $\tau_{a_1} = 0.70$ h long. The difference between our model and the data is particularly pronounced at $t \gtrsim 1$, when a sharper than expected transition is observed. This suggests that modifications of our model are needed to better capture the return of the bees during the initial night period.

Activity time. As a further test of our inference method, and the meaningfulness of its outcome, we look in detail at the inferred times in which foragers start (t_0) and stop (t_1) departing the hive on different days and different hives. The results were summarized in panels of Fig 5. The Fig 5a shows a single well-defined optimal parameter value exists, confirming the first and crucial step of our inference procedure; Fig 5b displays similar values across multiple hives in the same day and similar location, in agreement with the expectation that the initial times are similar for hives at similar locations; and in Fig 5c the estimated values for the same hive varies throughout the year, with the period of activity becoming shorter as the days become shorter (in the southern-hemisphere winter). These results confirm the success of our approach which is to treat these times as additional parameters to be inferred from the data simultaneously with the other parameters.

Systematic analysis. Motivated by the success of our methodology in the cases above, we applied it systematically to 10 different hives on the same day (2018–4-7). The results obtained from our two inference methods are reported in Table 2. The comparison of the inferred parameters of these hives with expectations about the behaviour of bees and hives provides further evidence that our mathematical model is capturing meaningful information from the data. In particular:

- The time forager bees start (t_0) and stop (t_1) departing the hives are aligned with sunlight.

- The time of minimum weight was inferred to be around t_c = 9: 45 a.m. shortly after t_0 (for all hives except 'Hive 8', which presents a later minimum). This result is compatible with the minimum daily weights reported in experimental data [39], which reveal minimum weights around 9: 00 a.m. Recently, results from segmented linear regression showed minimum losses can take place even 4 hours after sunrise [40].

- The estimation of the number of forager bees N_{max} is around a few thousand bees, in line with alternative methods of estimation. In particular, Fig 6a shows that the total number of bees per hive correlates positively with our estimation of N_{max} (R^2 = 0.596, p = 0.036). The fraction of forager bees (assuming $\tau_d = 0.816h$) is estimated to be around 18%, comparable to previous estimations (around $25\% - 30\%$ in Refs. [41, 42]). Additionally, we obtained a positive correlation between N_{max} and rate of weight gained during the day α, (R^2=0.778,

Table 2. Inferred parameters for 10 different hives on the same day (2018–4-7). Values $\emptyset$ in the interval estimation indicates that there was no interval of parameters satisfying all imposed constraints for this hive. The values of τ next to the parameters a_1, a_2, and d correspond to the time spent outside (respectively, inside) the hive and were computed using Eq (12). The uncertainties in the values of the estimation assuming τ_d = 0.816 h minutes, reported under the curly brackets, were computed using bootstrapping, see Materials and methods Sec. 4.3.

	Hives									
θ	1	2	3	4	5	6	7	8	9	10
	Effective parameters									
t_0 (h)	8.90	8.30	8.70	8.20	8.10	8.40	8.40	8.50	8.40	8.10
t_1 (h)	17.77	17.36	18.42	18.26	17.36	17.93	17.93	17.77	17.77	17.93
l (g/h)	9.762	40.805	5.600	16.628	33.183	20.663	15.016	3.974	7.687	4.399
A(g)	17,503	23,384	17,878	26,809	23,412	24,731	21,244	18,057	19,485	19,364
B(g)	17,409	22,.849	17,808	26,562	22,905	24,369	20,940	17,925	19,342	19,262
α (g/h)	16.594	95.081	4.595	38.529	72.577	59.560	24.780	4.721	23.754	4.476
t_c (h)	9.521	9.598	9.852	9.633	9.565	9.706	10.525	13.485	9.953	11.423
	Interval estimation									
N_{max} (bees)	[1950, 5400]	[5050, 8200]	[940, 2500]	[2310, 3450]	[7180, 7650]	[4150, 6030]	$\emptyset$	$\emptyset$	$\emptyset$	[840, 940]
W_0(g)	[16892, 17282]	[22451, 22794]	[17596, 17772]	[26413, 26537]	[22542, 22594]	[24045, 24254]	$\emptyset$	$\emptyset$	$\emptyset$	[19257, 19267]
a_1(1/h)	[3.59, 4.79]	[0.61, 0.92]	[1.44, 2.63]	[0.54, 0.79]	[0.81, 0.85]	[0.76, 1.01]	$\emptyset$	$\emptyset$	$\emptyset$	[0.20, 0.25]
τ_{a1}(h)	[0.14, 0.19]	[0.75, 1.13]	[0.26, 0.48]	[0.88, 1.28]	[0.82, 0.86]	[0.69, 0.92]	$\emptyset$	$\emptyset$	$\emptyset$	[2.75, 3.40]
a_2(1/h)	[1.34, 1.34]	[1.01, 1.02]	[3.20, 3.21]	[2.43, 2.47]	[1.07, 1.07]	[2.10, 2.11]	$\emptyset$	$\emptyset$	$\emptyset$	[2.36, 2.35]
τ_{a2}(h)	[0.52, 0.52]	[0.68, 0.69]	[0.22, 0.22]	[0.28, 0.29]	[0.65, 0.65]	[0.33, 0.33]	$\emptyset$	$\emptyset$	$\emptyset$	[0.29, 0.30]
d(1/h)	[0.82, 2.38]	[0.82, 1.41]	[0.82, 2.28]	[0.82, 1.26]	[0.82, 0.87]	[0.82, 1.24]	$\emptyset$	$\emptyset$	$\emptyset$	[0.82, 0.94]
τ_d(h)	[0.29, 0.85]	[0.49, 0.85]	[0.30, 0.85]	[0.55, 0.85]	[0.79, 0.85]	[0.56, 0.85]	$\emptyset$	$\emptyset$	$\emptyset$	[0.74, 0.85]
$m\left(\frac{\times 10^{-3} g}{bee}\right)$	[7.00, 10.00]	[37.40, 60.00]	[7.00, 12.00]	[38.50, 59.80]	[32.40, 34.10]	[28.80, 39.60]	$\emptyset$	$\emptyset$	$\emptyset$	[39.00, 45.00]
	Estimation assuming τ_d = 0.816 h									
N_{max} (bees)	5220±0	7930±0	2410±0	3330±0	7400±0	5830±0	3540±0	970±0	1260±0	910±0
W_0(g)	16912±3	22481±7	17606±5	26426±3	22570±4	24068±5	20840±1	17946±4	19339±2	19260±1
a_1(1/h)	4.77±0.6	0.90±0.0	2.60±0.2	0.77±0.0	0.83±0.0	0.99±0.1	0.50±0.0	0.11±0.0	0.44±0.0	0.24±0.0
a_2(1/h)	1.34±0.1	1.02±0.1	3.21±0.3	2.47±0.2	1.07±0.1	2.11±0.1	2.00±0.3	1.56±0.1	1.84±0.1	2.35±0.2
d(1/h)	0.85±0.1	0.85±0.0	0.85±0.1	0.85±0.1	0.85±0.0	0.85±0.1	0.85±0.0	0.85±0.0	0.85±0.1	0.85±0.0
$m\left(\frac{\times 10^{-3} g}{bee}\right)$	10.0±1	38.0±2	10.0±1	40.0±3	33.0±1	30.0±3	33.0±1	53.0±2	77.0±3	41.0±2

https://doi.org/10.1371/journal.pcbi.1010880.t002

p = 0.0056) consistent with the previously reported link between the size of the population and the amplitude of weight variations during the active foraging period [40].

- The average amount of food brought back to the hive by a bee was inferred to be typically in the range $m \in [10, 40] \times 10^{-3}$ g/bee, which is comparable to reported crop loads carried by active foraging bees when returning to the hive with loads of pollen [31], water [32], and different sweet solutions [33]. The two exceptions (Hives 8 and 9) show values up to $m = 77 \times 10^{-3}$ g/bee and possible correspond to sugar concentrations with total dissolved solids (TDS) of 60% [33].

- In 90% of the cases, $a_1 < a_2$, which is in line with the expectation that bees return faster to the hive at dusk (after they stop leaving).

- The typical time spent by forager bees inside and outside the hive is computed through an analogy with the half-life, τ, which represents the time that half of the foragers' colony will need to go outside the hive or vice-versa. The half-life, τ depends on the departure or arrival
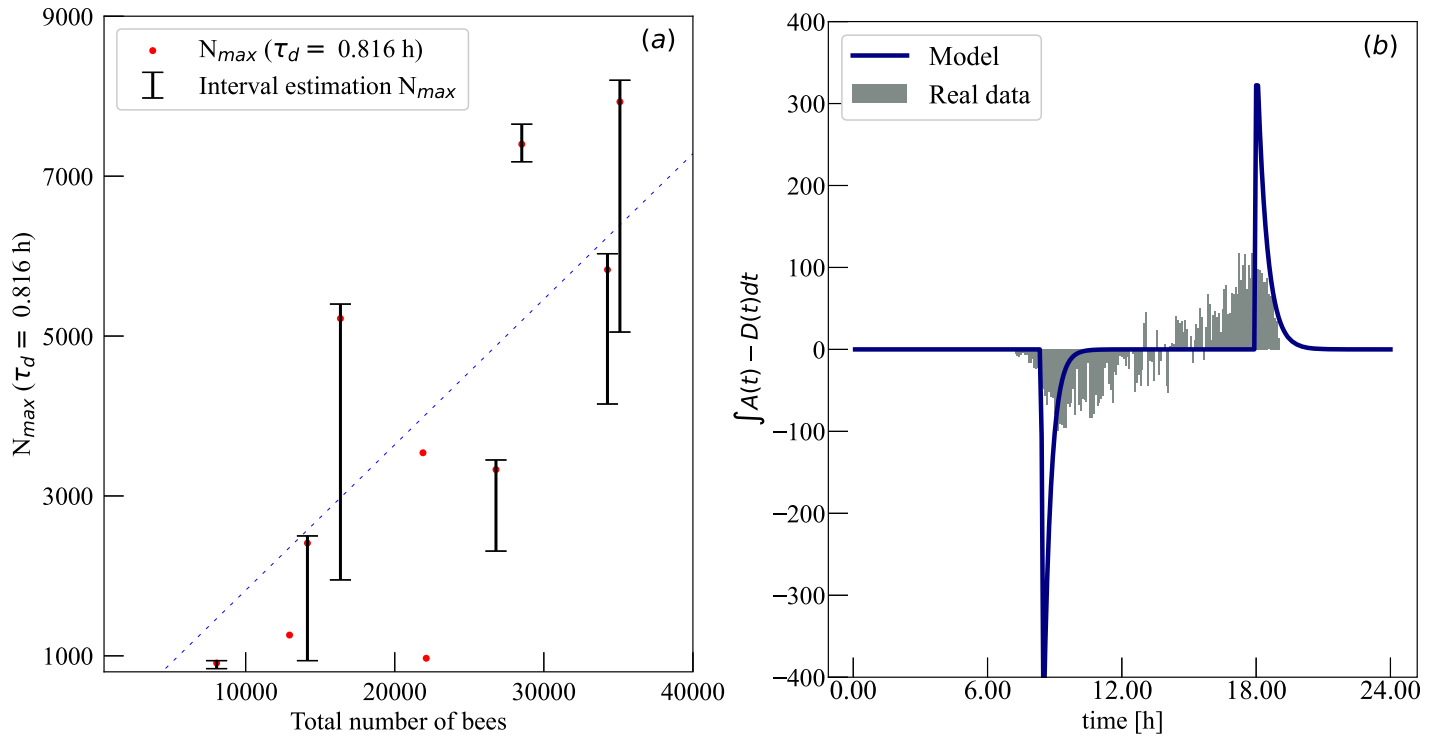
Fig 6. Comparing our model to independent measurements. (a) Our estimation of the number of active foragers N_{max} correlates positively with the total number of bees ($R^2 = 0.596$, $p = 0.036$). The dashed straight line corresponds to 18.2% of all bees being foragers, constant across all hives. The total number of bees was estimated through an independent measurement performed on 2018–04-9, two days after the data used in our inference [34]. The points without bars correspond to those that had no successful interval estimation, as discussed in Table 2. In all our analysis of correlation between variables, the reported p-value p is the probability of obtaining a positive coefficient of determination R^2 equal or larger than the reported R^2 under the null hypothesis that the variables are independent from each other. (b) The plot represents the difference between the number of bees arriving and bees departing a honey-bee hive during a 5 minutes interval. The model prediction (blue line) was computed integrating $A(t) - D(t)$ over each time interval. The real data was measured using bee-tracking methods.

rate $r \in \{d, a_1, a_2\}$ in hours as

$$\tau = (t_1 - t_0) \frac{\ln 2}{r}. \tag{12}$$

From the calculated values of d, $a_{1,2}$, and $t_1 - t_0$, we determine that the time in the hive between trips, τ_d, ranges from 0.29h to 0.85h and the typical time of a foraging trip, τ_a, is in between 0.14h to 3.40h. These estimations align with the results in Refs. [36, 37, 42]. The only hive that shows values up to 6.25 h for trip duration is Hive 8, this hive shows behaviour outside the ones predicted by our model, as evidenced by $\emptyset$ in the interval estimation. However, some reports points that bees are able to be outside the hive 6.25 h at the age of 36 days [35].

Arrivals and departures. Finally, we test the predictions of our model against independent experimental measures that tracked the number of bees that arrive and depart from a hive at a time interval [37]. Fig 6b compares these measurements to the assumption in our model and reveals the limitations of our simplifying choice of constant and piece-wise linear arrivals and departures in Eq (7). While our model captures the main characteristics of the data, unsurprisingly, bee activities do not start or end abruptly and the measured data shows smoother variations than the prediction of our model. A similar behaviour can be observed for $W(t)$—such as in Fig 3—where the measured data do not show the abrupt changes seen in the

model curve at $t = 0$ and $t = 1$ (discontinuous derivative). This limitation can be partially overcome within our modeling framework—Eqs (1) to (4)—by changing the assumptions of constant $a(t)$ and $d(t)$—e.g., in Eq (7). For instance, the next simplest alternative is to consider that the arrival remains constant $a(t) = a$ but that the departures varies continuously from 0 at time $t = 0$ to a maximum value d_{max} at time $t = 0.5$ and back to 0 at $t = 1$. The simplest polynomial with these characteristics is

$$d(t) = 4d_{max}t(1 - t). \tag{13}$$

In this case, Eq (3) is a linear first-order equation and the solution is:

$$x(t) = \frac{C}{e^{r(t)}} + a \int e^{r(t)dt}dt, \tag{14}$$

where $r(t) = at + 2d_{max}t^2 - \frac{4}{3}d_{max}t^3$ and C is the arbitrary constant to be fixed by the initial condition $x(0) = 1$. This alternative model has the same number of parameters (d_{max} instead of d), but requires additional computational resources to numerically find the solution for $x(t)$ through the integration of Eq (14). This example illustrate the flexibility of our model to consider more realistic settings and also the trade-off between model complexity and the need of computational approaches.

3 Discussion

We introduced a simple mathematical model of the foraging activities of bees that reproduces the observations of within-day weight variations of hives. Simplistic assumptions of the departure and arrival of bees allow us to obtain closed form solutions of the model that result in a continuous and piece-wise differentiable curve $W(t)$ with 10 biologically-relevant parameters. We inferred the parameters of $W(t)$ for 10 different hives and for different days across the year. Overall, the results reveal that the parameters we infer agree with existing knowledge of bee behavior and that the inferred number of forager bees is compatible with the expectation based on the independent measure of the total number of bees in the same hives. Further work is needed to quantitatively test the accuracy of our inferred parameters and the extent of validity of our model. For instance, future studies could simultaneously collect weight data and forager activities with tracking devices for individual bees, allowing for a direct comparison between our approach and estimations based on independent measurements.

The main advantage of our approach when compared to previous works is that it combines mechanistic mathematical models with data analysis to obtain interpretable quantitative information about the condition of the hive. Additional advantages of our model approach include: (i) indications from the inference of when results cannot be trusted (either because they are not capturing meaningful information or because optimal results lie outside of the allowed error range); (ii) that our framework allows for the proposal of more sophisticated models, e.g., by considering additional terms contributing to the food dynamics in Eq (4) or more detailed modes of arrival $a(t)$ and departure $d(t)$ such as in Eq (13); and (iii) that our model allows for a systematic exploration of the effect of foraging activities and food processing parameters on how hive weight changes over time. An important methodological expansion of our work would be to consider and compare different type of functional forms in our model. This will require considering cases in which solutions are not obtained in closed forms and the implementation of a Bayesian-model comparison that can compare models of different complexities, in line with modern simulation-based and Bayesian inference that are increasingly applied to dynamical systems of biological interest [43–45].

The success of our model in extracting meaningful information about bee hives opens the possibility of using it for a systematic analysis of the health of hives across different hives and across time. Our model can be used to obtain interval estimations of the following quantities, which are of particular interest:

- The number of active foragers in a bee colony N is a crucial parameter of the capacity of a colony at sustaining itself. Environmental stressors such as pesticides can cause a rapid depletion of the population of active foragers [46]. Estimations of N provide a rapid assessment of the daily foragers population and allow quick intervention and stress reduction to potentially prevent colony collapse [20].

- The time spent by foragers outside the hive τ_a can be connected to the efficiency of the colony because long times outside the hive lead to a reduction of the life expectancy of foragers [37].

- The amount of food collected by bees in each trip m allows for a better assessment of foraging difficulty and thus of environments that can best support beekeeping operations [22].

- The rate of weight growth of a hive α gives a reliable indicator of whether foragers are supplying their colony properly (leading to optimal colony development) or if instead the colony is endangered (suggesting the existence of stressors or environmental limitations).

Tracking the daily evolution of these parameters opens the perspective of developing weight-based early-warning indicators of bee colony failure that could lead to new strategies for risk control.

4 Materials and methods

4.1 Computational methods

Here we describe the computational methods used to estimate ranges of the parameters θ of our model presented in Table 1 from the within day time series of weights $W'(t)$. In Sec. 2.2 of this manuscript this issue has been formulated as the problem of finding the parameters θ that minimize the sum of squares L in Eq (11). The ten parameters proposed in this model were reduced to nine parameters after considering the weight of a bee w as 0.113 g/bee.

The minimisation was performed in a nine-dimensional space through a least-square method composed of three different approximation levels. First, the clock times when bees started departing t_0, and when they stopped leaving t_1. Next, a reliable estimation of the characteristic features from the time series of the weight fluctuations was performed. Finally, the remaining seven parameters were obtained from the connection with the effective parameters.

For the estimation of clock times for t_0 and t_1 we created a squared-grid for the times $t_0 \in$ [6:30am,9:00am] and $t_1 \in$ [2:30pm,6:30pm] spaced every 5 minutes. These times combinations provides 1440 different scaling options. Then, we applied to each ordered pair, (t_0, t_1), the Levenberg–Marquardt algorithm for square minimization. We used the python library SciPy (through "scipy.optimize.curve_fit" Pythons' function) with the inputs described in Table 3 to apply the algorithm.

We chose the combination of t_0 and t_1 that minimizes the error as the inferred scale time parameters. Then, after re-scaling our data set on the time axis, we proceeded with the estimation of effective parameters; A, B, t_c, and α from the daily-hive data. We infered the remaining seven model parameters from a system of six equations, which consist of: (i) four equations presented as the "effective description of the model"; (ii) one equation from the conservation of the number of departed and returned bees; and (iii) the last equation is the rate of evaporation ℓ, which can be calculated directly from the time series after the re-scaling process when

Table 3. The inputs required for the "scipy.optimize.curve_fit" Pythons' function.

Parameter	Description
y_{data}	Measured weight W' per minute over a full day.
x_{data}	Time per minute scaled for the corresponding measured weight W'.
f	Model function, introduced in Eq (1).
P_0	Initial guess for the θ parameters that are not bee facts W_0 = 15000 g, N_{max} = 15000 bees, l = 17 g/h m = 0.01 g/bee, a_1 = 0.21 h^{-1}, a_2 = 0.41 h^{-1}, d = 1 h^{-1}.
$Bounds$	Lower and upper bounds on parameters W_0 = [5000 $max(W')$] g, N_{max} = (0 80000) bees, l = (0 max (W')) g/s, m = (0 0.78w) g, a_1 = [0.10 4.80] h^{-1}, a_2 = [0.10 4.80] h^{-1}, d = [0.81 3.01] h^{-1}.
$maxfev$	Maximum number of evaluations allowed, for this function, 5000 iterations were used.
t_0, t_1	Clock times corresponding to t_0 and t_1, obtained previously.

$t < 0$. Then, we performed a simultaneous calculation with six equations and seven variables; consequently, the system is undetermined with no unique solution due to an extra degree of freedom. We proposed that our dependent system of equations must be solved for any possible N_{max} value. Note that the total weight of the hive at $W(t = 0)$ can be decomposed in two terms only: the weight given by the total number of forager bees N_{max}, and the initial constant weight W_0. Consequently, our solution will depend on picking a value for N_{max} and finding the corresponding W_0 value. Thus, an infinite number of solutions will be obtained and then reduced to a confidence interval as can be observed in S1 Fig. Alternatively, by pre-defining one value, such as departure rate d, we can reduce the variables of our equation's system and obtain a unique solution for the parameters $\hat{\theta}$. The solution has extended to 10 independent and different hives during the day 2019–4-7 and is presented in Sec 2.3.

4.2 Pre-processing of data

Before any computation, we removed outliers that do not reflect actual change in hive weight and that would otherwise affect our analysis. These outliers appeared due to equipment failure, animal or human interactions with the hives, or rainfall. In practice, we removed all measurement points that corresponded to a difference of 20g or more within two consecutive minutes, a conservative threshold which ensures that strong oscillations were excluded. From the ten different hives analysed, the hive that presented the highest number of outliers was hive '10', with 89 outliers, so at most only 6.1% of the data were excluded from this hive (S2 Fig).

4.3 Bootstrapping

Bootstrapping is used for assessing the effects of data uncertainty. We applied a bootstrapping method to estimate the uncertainty around the estimated parameters $\hat{\theta}$. For each daily data of 10 different hives, 200 re-samplings were generated, and the minimisation procedure described above was applied to them (for the fixed value of t_0, t_1 obtained in the full dataset). Suppose the inference fails or the number of iterations exceeds 5, 000 evaluation, and the optimal parameters are still not found. In that case, the re-sample is discarded (for the 10 analysed hives listed in Table 2 only 5% of the 200 re-samplings were discarded). The inferred parameters of the accepted re-samples are used to estimate the range of plausible parameters. The uncertainty reported in the main manuscript corresponds to the standard deviation over the parameters estimated in the re-sampled data. For 10 hives reported in Table 2 the uncertainty of the 7 inferred parameters corresponds to an area of 95.4% under the curve of the probability density function. The S3 Fig displays an example of the probability density to calculate the range of values for the seven parameters of 'Hive 6' on 2018–4-7.

Supporting information

S1 Fig. Multiple solutions are obtained from the Hive 6 on 2018–4-7, and then reduced to a confidence interval. (a) The time spent inside, τ_d, and outside the hive during the day, τ_{a_1} is compared with reported intervals of time inside the hive taken by the forager bees between trips, and an average of time outside the hive per flight. Then, a valid interval is selected considering previous reports. (b) The maximum reported food load collected per trip (crop load) is 0.077 g. The shaded area represents a region where an acceptable crop load is obtained. (c) Representation of two characteristic parameters, the rate of gained weight and the intersection with $t = 0$, as N_{max} decreases B and α become unstable. (d) Error calculated as the sum of the squared difference between the model and the real data. The shaded area represents a region where an acceptable error occurs based on bootstrapping simulation.
(PDF)

S2 Fig. Example of outlier detection for the Hive 10 on 2018–4-7. The outliers detected in red were removed before the data analysis.
(PDF)

S3 Fig. Probability Density Function for the parameters displayed at Table 1 after the bootstrapping process of the Hive 6 on 2018–4-7. These fitting parameters are: N_{max}(bees), W_0(kg), m(g), $a_1(h^{-1})$, $a_2(h^{-1})$, $d(h^{-1})$, l(g/h) with $t_0 = 8.4$ h and $t_1 = 17.93$ h after bootstrapping process of 'Hive 6' during the 2018–4-7.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Karina Arias-Calluari, Theotime Colin, Tanya Latty, Mary Myerscough, Eduardo G. Altmann.

Data curation: Karina Arias-Calluari, Theotime Colin.

Formal analysis: Karina Arias-Calluari, Theotime Colin, Eduardo G. Altmann.

Funding acquisition: Tanya Latty, Mary Myerscough, Eduardo G. Altmann.

Investigation: Karina Arias-Calluari, Theotime Colin.

Methodology: Karina Arias-Calluari, Theotime Colin, Eduardo G. Altmann.

Software: Karina Arias-Calluari.

Supervision: Eduardo G. Altmann.

Validation: Karina Arias-Calluari, Theotime Colin.

Writing – original draft: Karina Arias-Calluari, Theotime Colin, Eduardo G. Altmann.

Writing – review & editing: Karina Arias-Calluari, Theotime Colin, Tanya Latty, Mary Myerscough, Eduardo G. Altmann.

References

1. Goulson D, Nicholls E, Botías C, Rotheray EL. Combined stress from parasites, pesticides and lack of flowers drives bee declines. Science. 2015; 347(6229):1255957.

2. Goulson D, Nicholls E, Botías C, Rotheray EL. Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. Science. 2015; 347 (6229). https://doi.org/10.1126/science.1255957 PMID: 25721506

3. Hladik ML, Main AR, Goulson D. Environmental risks and challenges associated with neonicotinoid insecticides; 2018.

4. Siviter H, Bailes EJ, Martin CD, Oliver TR, Koricheva J, Leadbeater E, et al. Agrochemicals interact synergistically to increase bee mortality. Nature. 2021; 596(7872):389–392. https://doi.org/10.1038/s41586-021-03787-7 PMID: 34349259

5. Le Conte Y, Ellis M, Ritter W. Varroa mites and honey bee health: can Varroa explain part of the colony losses? Apidologie. 2010; 41(3):353–363. https://doi.org/10.1051/apido/2010017

6. Powney GD, Carvell C, Edwards M, Morris RK, Roy HE, Woodcock BA, et al. Widespread losses of pollinating insects in Britain. Nature communications. 2019; 10(1):1–6. https://doi.org/10.1038/s41467-019-08974-9 PMID: 30914632

7. Woodcock BA, Bullock J, Shore R, Heard M, Pereira M, Redhead J, et al. Country-specific effects of neonicotinoid pesticides on honey bees and wild bees. Science. 2017; 356(6345):1393–1395. https://doi.org/10.1126/science.aaa1190 PMID: 28663502

8. Corbet SA, Williams IH, Osborne JL. Bees and the pollination of crops and wild flowers in the European Community. Bee world. 1991; 72(2):47–59. https://doi.org/10.1080/0005772X.1991.11099079

9. Lee KV, Steinhauer N, Rennich K, Wilson ME, Tarpy DR, Caron DM, et al. A national survey of managed honey bee 2013–2014 annual colony losses in the USA. Apidologie. 2015; 46(3):292–305. https://doi.org/10.1007/s13592-015-0356-z

10. Currie RW, Pernal SF, Guzmán-Novoa E. Honey bee colony losses in Canada. Journal of Apicultural Research. 2010; 49(1):104–106. https://doi.org/10.3896/IBRA.1.49.1.18

11. Jacques A, Laurent M, Consortium E, Ribière-Chabert M, Saussac M, Bougeard S, et al. A pan-European epidemiological study reveals honey bee colony survival depends on beekeeper education and disease control. PLoS one. 2017; 12(3):e0172591. https://doi.org/10.1371/journal.pone.0172591 PMID: 28278255

12. Meikle W, Holst N. Application of continuous monitoring of honeybee colonies. Apidologie. 2015; 46 (1):10–22. https://doi.org/10.1007/s13592-014-0298-x

13. Francis RM, Nielsen SL, Kryger P. Varroa-virus interaction in collapsing honey bee colonies. PloS one. 2013; 8(3):e57540. https://doi.org/10.1371/journal.pone.0057540 PMID: 23526946

14. Morawetz L, Köglberger H, Griesbacher A, Derakhshifar I, Crailsheim K, Brodschneider R, et al. Health status of honey bee colonies (Apis mellifera) and disease-related risk factors for colony losses in Austria. PloS one. 2019; 14(7):e0219293. https://doi.org/10.1371/journal.pone.0219293 PMID: 31287830

15. Horn J, Becher MA, Johst K, Kennedy PJ, Osborne JL, Radchuk V, et al. Honey bee colony performance affected by crop diversity and farmland structure: a modeling framework. Ecological Applications. 2021; 31(1):e02216. https://doi.org/10.1002/eap.2216 PMID: 32810342

16. Charreton M, Decourtye A, Henry M, Rodet G, Sandoz JC, Charnet P, et al. A locomotor deficit induced by sublethal doses of pyrethroid and neonicotinoid insecticides in the honeybee Apis mellifera. PloS one. 2015; 10(12):e0144879. https://doi.org/10.1371/journal.pone.0144879 PMID: 26659095

17. Leoncini I, Crauser D, Robinson G, Le Conte Y. Worker-worker inhibition of honey bee behavioural development independent of queen and brood. Insectes Sociaux. 2004; 51(4):392–394. https://doi.org/10.1007/s00040-004-0757-x

18. Seeley TD, Kolmes SA. Age polyethism for hive duties in honey bees—illusion or reality? Ethology. 1991; 87(3-4):284–297. https://doi.org/10.1111/j.1439-0310.1991.tb00253.x

19. Robinson GE. Regulation of honey bee age polyethism by juvenile hormone. Behavioral ecology and sociobiology. 1987; 20(5):329–338. https://doi.org/10.1007/BF00300679

20. Perry CJ, Søvik E, Myerscough MR, Barron AB. Rapid behavioral maturation accelerates failure of stressed honey bee colonies. Proceedings of the National Academy of Sciences. 2015; 112(11):3427–3432. https://doi.org/10.1073/pnas.1422089112 PMID: 25675508

21. Henry M, Beguin M, Requier F, Rollin O, Odoux JF, Aupinel P, et al. A common pesticide decreases foraging success and survival in honey bees. Science. 2012; 336(6079):348–350. https://doi.org/10.1126/science.1215039 PMID: 22461498

22. Quinlan GM, Sponsler D, Gaines-Day HR, McMinn-Sauder HB, Otto CR, Smart AH, et al. Grassy–herbaceous land moderates regional climate effects on honey bee colonies in the Northcentral US. Environmental Research Letters. 2022;. https://doi.org/10.1088/1748-9326/ac7063

23. Meikle W, Weiss M, Stilwell A. Monitoring colony phenology using within-day variability in continuous weight and temperature of honey bee hives. Apidologie. 2016; 47(1):1–14. https://doi.org/10.1007/s13592-015-0370-1

24. Meikle WG, Holst N, Colin T, Weiss M, Carroll MJ, McFrederick QS, et al. Using within-day hive weight changes to measure environmental effects on honey bee colonies. PloS one. 2018; 13(5):e0197589. https://doi.org/10.1371/journal.pone.0197589 PMID: 29791462

25. Colin T, Meikle WG, Paten AM, Barron AB. Long-term dynamics of honey bee colonies following exposure to chemical stress. Science of the Total Environment. 2019; 677:660–670. https://doi.org/10.1016/j.scitotenv.2019.04.402 PMID: 31071668

26. Meikle WG, Rector BG, Mercadier G, Holst N. Within-day variation in continuous hive weight data as a measure of honey bee colony activity. Apidologie. 2008; 39(6):694–707. https://doi.org/10.1051/apido:2008055

27. Arias-Calluari K, Altmann EG. Codes for modelling weight of bee hive; 2022. Available from: https://doi.org/10.5281/zenodo.7272480.

28. Chalcoff VR, Aizen MA, Galetto L. Nectar concentration and composition of 26 species from the temperate forest of South America. Annals of botany. 2006; 97(3):413–421. https://doi.org/10.1093/aob/mcj043 PMID: 16373370

29. Thompson H. Extrapolation of acute toxicity across bee species. Integrated environmental assessment and management. 2016; 12(4):622–626. https://doi.org/10.1002/ieam.1737 PMID: 26595163

30. Thompson HM, Pamminger T. Are honeybees suitable surrogates for use in pesticide risk assessment for non-Apis bees? Pest management science. 2019; 75(10):2549–2557. https://doi.org/10.1002/ps.5494 PMID: 31124265

31. García-García MC, Ortiz PL, Dapena MJD. Variations in the weights of pollen loads collected by Apis mellifera L. Grana. 2004; 43(3):183–192. https://doi.org/10.1080/00173130410020350

32. Visscher PK, Crailsheim K, Sherman G. How do honey bees (Apis mellifera) fuel their water foraging flights? Journal of Insect Physiology. 1996; 42(11-12):1089–1094. https://doi.org/10.1016/S0022-1910(96)00058-3

33. Afik O, Dag A, Shafir S. The effect of avocado (Persea americana) nectar composition on its attractiveness to honey bees (Apis mellifera). Apidologie. 2006; 37(3):317–325.

34. Colin T, Forster CC, Westacott J, Wu X, Meikle WG, Barron AB. Effects of late miticide treatments on foraging and colony productivity of European honey bees (Apis mellifera). Apidologie. 2021; 52(2):474–492. https://doi.org/10.1007/s13592-020-00837-3

35. He X, Wang W, Qin Q, Zeng Z, Zhang S, Barron AB. Assessment of flight activity and homing ability in Asian and European honey bee species, Apis cerana and Apis mellifera, measured with radio frequency tags. Apidologie. 2013; 44(1):38–51. https://doi.org/10.1007/s13592-012-0156-7

36. Rodney S, Purdy J. Dietary requirements of individual nectar foragers, and colony-level pollen and nectar consumption: a review to support pesticide exposure assessment for honey bees. Apidologie. 2020; 51(2):163–179. https://doi.org/10.1007/s13592-019-00694-9

37. Colin T, Warren RJ, Quarrell SR, Allen GR, Barron AB. Evaluating the foraging performance of individual honey bees in different environments with automated field RFID systems. Ecosphere. 2022; 13(5):e4088. https://doi.org/10.1002/ecs2.4088

38. Clarke D, Robert D. Predictive modelling of honey bee foraging activity using local weather conditions. Apidologie. 2018; 49(3):386–396. https://doi.org/10.1007/s13592-018-0565-3

39. Hambleton JI. The effect of weather upon the change in weight of a colony of bees during the honey flow. 1339. US Department of Agriculture; 1925.

40. Holst N, Meikle WG. Breakfast canyon discovered in honeybee hive weight curves. Insects. 2018; 9(4):176. https://doi.org/10.3390/insects9040176 PMID: 30513723

41. Van Der Steen J. The foraging honey bee. BBKA News-The British Bee Journal. 2015; 2015(February):43–46.

42. Rodney S, Kramer VJ. Probabilistic assessment of nectar requirements for nectar-foraging honey bees. Apidologie. 2020; 51(2):180–200. https://doi.org/10.1007/s13592-019-00693-w

43. Cranmer K, Brehmer J, Louppe G. The frontier of simulation-based inference. PNAS. 2020. https://doi.org/10.1073/pnas.1912789117 PMID: 32471948

44. Roda WC. Bayesian inference for dynamical systems. Infectious Disease Modelling. 2020; 5:221–232. https://doi.org/10.1016/j.idm.2019.12.007 PMID: 32021948

45. Barnes CP, Silk D, Sheng X, Stumpf MP. Bayesian design of synthetic biological systems. Proceedings of the National Academy of Sciences. 2011; 108(37):15190–15195. https://doi.org/10.1073/pnas. 1017972108 PMID: 21876136

46. Colin T, Meikle WG, Wu X, Barron AB. Traces of a neonicotinoid induce precocious foraging and reduce foraging performance in honey bees. Environmental science & technology. 2019; 53(14):8252–8261. https://doi.org/10.1021/acs.est.9b02452 PMID: 31257879