Beehive Population Dynamics

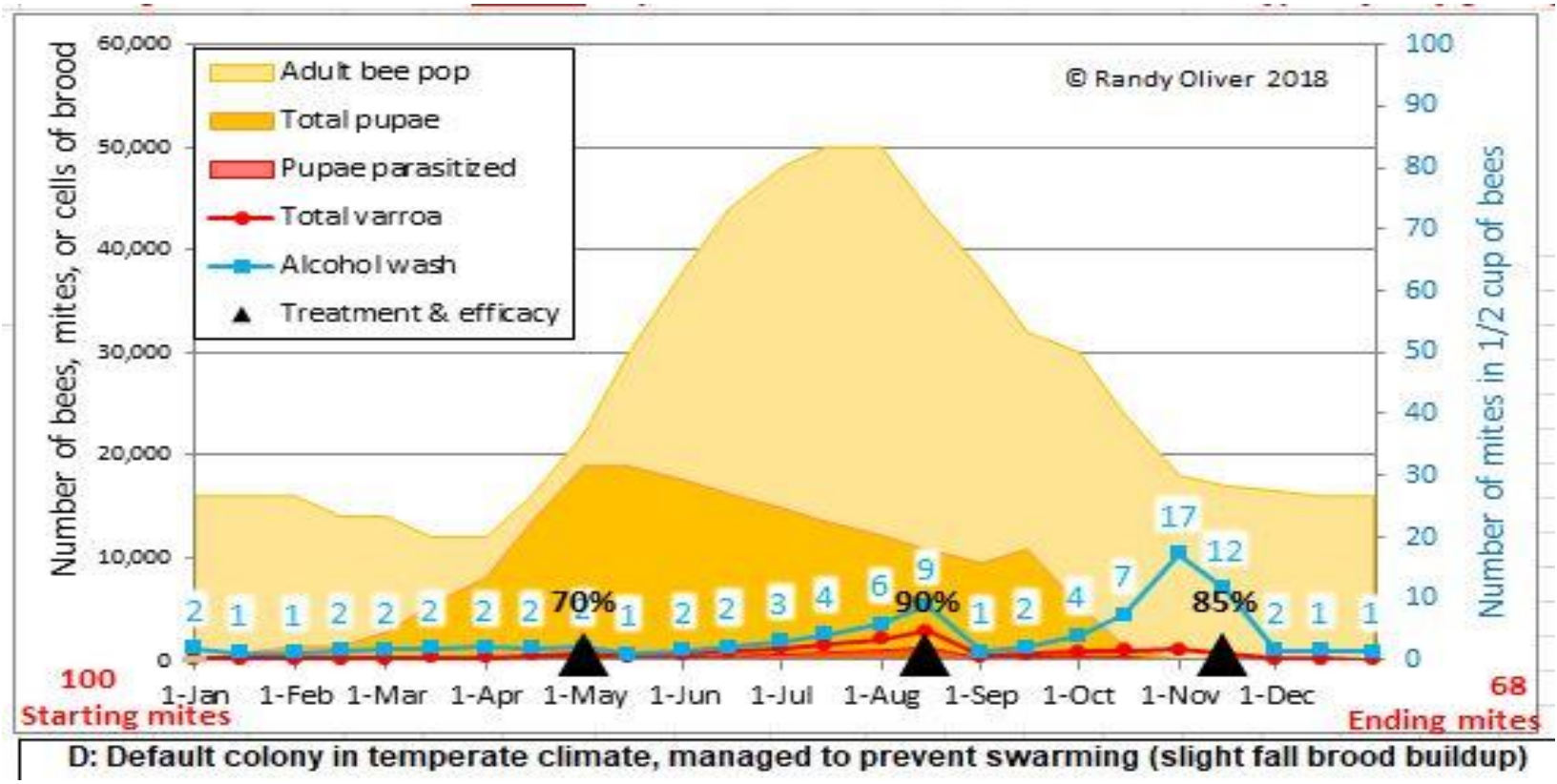
By Tyler Trotter

Part 1: My Model

I recently got into beekeeping. In researching how to take care of bees, the different roles of each bee, etc., I ran across some very good resources about the scientific data that some beekeepers have uncovered. These resources include everything from how much to feed a colony to the size and shape of their home. I even read an entire book about the basics of beekeeping. Intrigued, I continued studying the honey accumulation in the hive, the dependency of the hive on external feeding for some time after their initial installation, and the rate at which a freshly installed 'bee package' (roughly 3lbs of bees with a queen) will grow in population to fill the boxes I had built as their home.

For the purposes of this project, this is where my focus lies. Beehive population dynamics. I had spent several hours thinking about the population dynamics internally within my hive. Each week I had analyzed their progress, noting when to add the next brood box (where the queen will also lay eggs—different than an excess honey 'super'). One of the most in depth researchers of bees is author and biologist Randy Oliver. On his website, Randy has made some fascinating discoveries regarding bees. The following graphic was taken from that website:

(Graphic 1, www.ScientificBeekeeping.com)



D: Default colony in temperate climate, managed to prevent swarming (slight fall brood buildup)

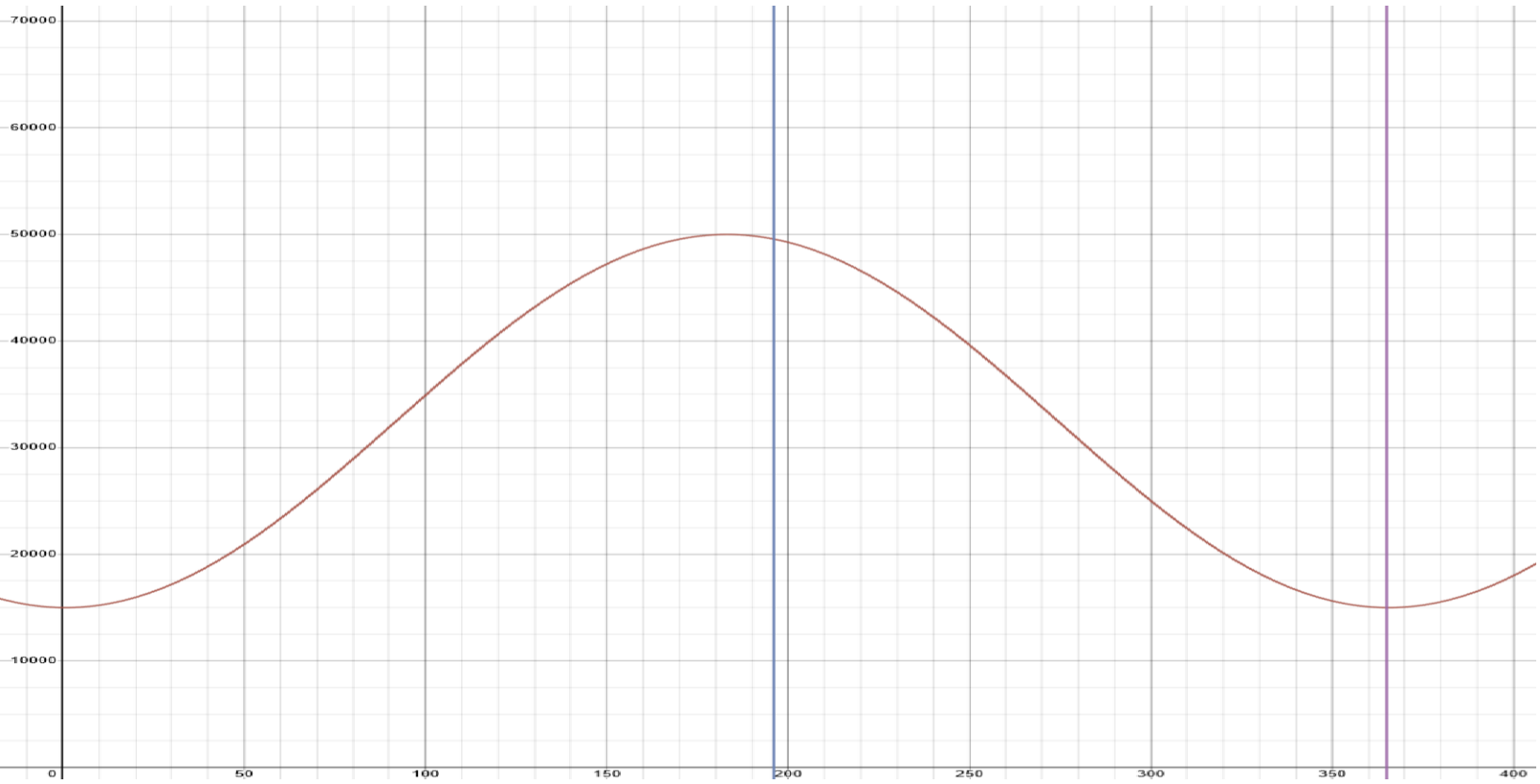
Here, Randy is trying to understand how the population of a predator to the bee, known as the Varroa Mite, changes as the population of the hive increases. My focus here is to analyze both the adult bee population and the total pupae. First, it is interesting to note that the population decreases during the winter months and the bounces back up to a maximum during the summer in apparent sinusoidal fashion. The simplest model I experimented with to represent what is witnessed in the graphic was:

$$P(t) = 17{,}500 * \text{Sin}(t) + 32{,}500$$

In the graph above, the population of bees reaches a minimum of ~15,000 during the coldest times of the year and achieves a maximum of 50,000 in mid-July. Here, recognize that (50,000 + 15,000) / 2 translates to a vertical shift of 32,500 and the amplitude is (50,000 – 15,000) / 2. Likewise, the period $0 \leq t \leq 2\pi$ correlates to a 1-year cycle of the beehive's population. With some experimentation, I settled in on (where t is in days and t = 0 translates to Jan. 1st):

$$P(t) = 17{,}500 * \text{Sin}(39.5 * t / (365*2\pi) + 4.7) + 32{,}500$$

(Graphic 2, used www.Desmos.com)



So now that I had a decent framework whereby to understand the population at a given day of the year, it was then that I started to think logically about the differential equation that undergirded the change in population.

A healthy beehive has exactly one egg layer known as the queen. She is the single most important bee in the hive and egg production is her primary role. This means the sole factor in hive growth is dependent on how much the queen lays at a given time. The average queen can lay an approximate maximum of 2000 eggs in a day. These usually sit in some form of incubation for 7-10 days until they spawn as a juvenile bee. Thus, at any time there can be as many as 20,000 brood in a hive. This is logical and coincides with Randy's graphic meaning that his queen is average in this way. Now, to highlight bee death it is important to start with the bee's lifecycle. On average, a bee lives for about 29 days. Likewise, bees require a certain amount of food and if there is a food deficit in a region, some number of bees will die due to starvation. For the following model, I have disregarded larvae casualty and bee predation (which can be highly significant if left unchecked—hence why Randy is studying the mite):

$\dot{P}$ = Births (Queen Laying) – Deaths (Starvation, Predation, Defects, Pesticides, etc)

$$\dot{P} = [1700*\text{Sin}^2(t) + 300] - [P/29 + (1 - \alpha)*P]$$

$$= [1700*\text{Sin}^2(t) + 300] - [P*(30 - 29*\alpha)/29]$$

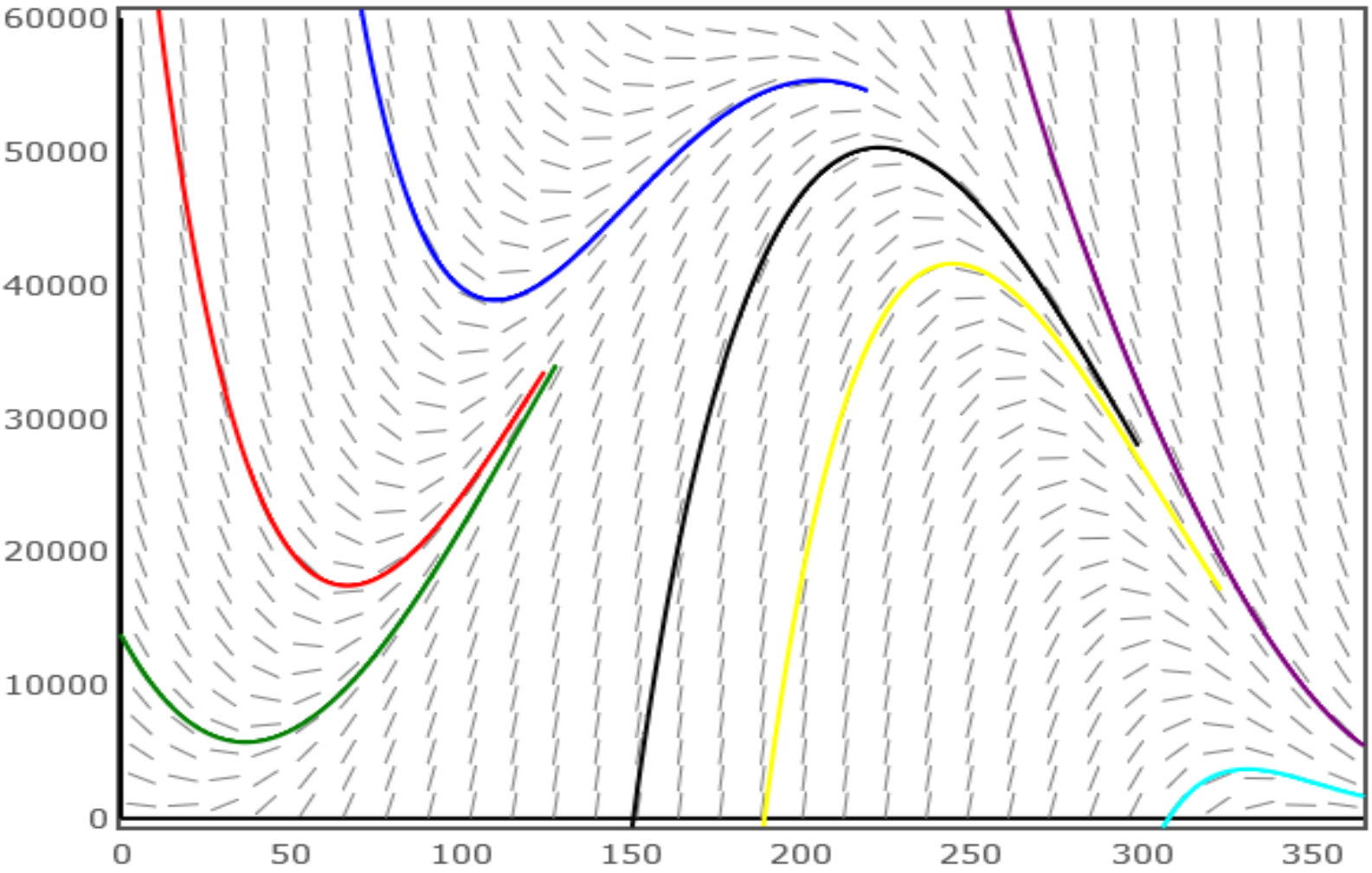
In the above model, which is a logical model I came up with, births solely come from the queen. Since fewer than zero eggs does not make physical sense, Sin was squared. Likewise, since she never lays 0 eggs by assumption in this model, there is a 300-egg baseline as the fewest eggs that could be lain per day. Thus, at her peak she lays 2000 and at the minimum she lays 300. For deaths, this model includes P/29 which represents the average number of the population which dies per day. $(1 - \alpha)*P$, $0 \leq \alpha \leq 1$ corresponds to a percentage of the population which would die due to food scarcity. In my current hive's situation, I am feeding my bees and therefore $\alpha = 1$. This is a parameter which is dependent of the prevalence of food in a region.

Since this is merely a logical model that I came up with by thinking about the nature of bees, it has some immediately noticeable shortcomings. This is not entirely rigorous since it over accounts for death during its peak and underestimates at its minimum for average daily deaths. However, since the population is changing gradually over the course of an entire year, I consider this part acceptable. Furthermore, for $\alpha = 0$, the population would have immediate die off, but in this model the queen continues to produce eggs and there would be a subsequent resurgence of the population that happened after this supposed mass death until it immediately died the next day (repeat).

As mentioned before, due to my feeding the bees I set my parameter α = 1 and thus there is no death due to food scarcity. With some further experimentation, I achieved the following function:

$$\dot{P} = 2000*\text{Sin}^2(20*t/(365*2\pi)) - P/29$$

(Graphic 3, used https://homepages.bluffton.edu/~nesterd/apps/slopefields.html)



This is amazing because it demonstrates the following carrying capacity:

$$\dot{P} = 2000*\text{Sin}^2(10*t/(365*\pi)) - P/29 = 0$$

$$\Rightarrow \quad 2000*\text{Sin}^2(10*t/(365*\pi)) = P/29$$

$$P = 29*2000*\text{Sin}^2(10*t/(365*\pi))$$

$$= 58{,}000\text{Sin}^2(10*t/(365*\pi))$$

And since $0 \leq \text{Sin}^2(10*t/(365*\pi)) \leq 1$, we get that the maximum population of the hive is 58,000 bees. Thus, the stable cycle in my model is $58{,}000\text{Sin}^2(10*t/(365*\pi))$. It is globally stable since all initial conditions tend towards it as $t \rightarrow \infty$. I interpret this as so, for a population greater than 58,000 the queen simply cannot keep up and thus more bees die than eggs were laid 10 days prior so there is a net population decrease. This model strikes me as being pretty accurate according to the population graph shown in the beginning, the experimental data from Randy. For this equation, I removed the 300 minimum births and bundled it into the amplitude of the Sin component. Re-implementing the 300 daily minimal births produces the following:

$$P = 49{,}300*\text{Sin}^2(10*t/(365*\pi)) + 8{,}700$$

Which implies a theoretical minimal population of 8700 and the same carrying capacity. Graphically, compared to the previous population function, appears as follows (where blue is the new and red is the old):

(Graphic 4, used www.Desmos.com)

Recall that the blue graph represents the 1-year periodic cycle for which the deaths and births are equalized. This graph is amazing because it shows that the stable cycle of the differential equation has predictive power for a hive's population.

Finally, consider a different queen who lays more eggs or another who lays less. It seems logical that a reformed equation would set this as a parameter dependent on the queen's laying capacity, like this:

$$\dot{P} = (B-M)*Sin^2(20*t/(365*2\pi)) - P/29 + M$$

Where B represents her maximal laying and M represents her minimal laying. Then the carrying capacity simply becomes 29*(B-M) + 29*M = 29*B. As a rule of thumb, this seems in line with the established data of hives reaching maximums somewhere in the ballpark of 50,000 – 80,000 bees. As a result of my model, I suspect that Randy didn't measure, or it became difficult to precisely measure the population of his bees in mid-July where they reached their exact peak.

This leads me to another interesting consideration. Bees are known to do something called swarming. This is how the bee population grows—not within a hive, but at the level of the hive. Bees will produce another queen and about half of the population will follow the new queen to a new location where they will establish a new colony. The other half stays with the old queen and regrows. This aspect of bees is very fascinating as it allows an entire hive to reproduce and duplicate itself, ultimately producing more bees than just the carrying capacity of one hive with one queen. The bees will not swarm their first year, but the second year after their hive has established itself it is prepared to split. However, the population of bees within the hive determines whether the hive will swarm or not. Let me introduce the following.

Firstly, this iterated map is an approximation for a neighborhood around the carrying capacity in the direction of increasing time:

$$P_{n+1} = \begin{cases} P_n + (B - M), & P_n < 29 * B \\ P_n/2, & P_n \geq 29 * B \end{cases}$$

And it represents the swarming effect of a hive's population as you exceed this critical value (which in this case is just the carrying capacity). After which, the aforementioned population dynamics get used. This represents an interesting fact about nature which in many things is probabilistic. For example, sometimes a bee lives a bit longer than exactly 29 days, or the queen started laying a few days early or stays in her peak laying for

longer than a day. It is during these times that the above model loses its predictive power. When the hive reaches a population larger than predicted by the model, they swarm (of course, only if they have enough time to fully establish a hive and survive the coming winter).

Secondly, suppose that every year this phenomenon happens to every single hive. Then every year any given hive splits and you end up with exponential growth. You start with 1 hive on year 1, that hive splits the next year and then there are 2 hives. After a year both of those hives split, and you have 4 hives. The model becomes:

$$\text{Number of Hives} = 2^{\text{years}-1}$$

This translates to:

$$\text{Total Bees} = \text{Number of Hives} * \text{Population} = 2^{\text{years}-1} * P_{n+1}$$ (or P if before swarm / ~50 days after swarm by my model)

This is something I am still exploring since there is a hive carrying capacity based upon the 'allowable' density of hives in a given region depending on the prevalence of food, and also the fact that not every hive swarms every year, hives die off, etc.. I would also like to introduce other variables, such as food availability, predator population dynamics, and I've been reading that bees live only 29 days during the summer due to their high workload ultimately killing them, but actually live for much longer during the winter months where their activity decreases significantly. So, even the average bee lifetime is periodic. This one-year periodicity is starting to manifest as a theme the more I think about bee population—which makes sense. Also, this could explain why Randy's bee population is about 15,000 compared to my predicted 8,700 since bees aren't working themselves to death during this time.

Part 2: Literature Models

For the next couple of pages, I comment on models I have found regarding beehive population dynamics from academic literature and how their considerations could improve my model. There are several papers which have delved into quantitatively understanding the dynamics of bee populations. In the paper, 'A Quantitative Model of Honey Bee Colony Population Dynamics', by David Khoury, Mary Myerscough, and Andrew Barron, they included the following models which differ from mine in that they refer to populations of 'hive bees' and 'forager bees', with unique lifecycles and roles:

Rate of change of hive bee numbers:

$$\frac{dH}{dt} = E(H,F) - HR(H,F)$$

eclosion recruitment to forager class

Rate of change of forager numbers:

$$\frac{dF}{dt} = HR(H,F) - mF.$$

recruitment death

In this model, m = death rate of forager bees, H = number of hive bees, E = eclosion and R = recruitment rate of a specific type of bee (they indicate that all bees start as hive bees but get recruited to foraging). This is interesting to consider hive bees and foraging bees as unique sub-populations since the number of bees which can be reared depends on how many hive bees there are in the hive to care for them. This means that if there are too few hive bees, even if the queen is laying her maximal number of eggs, most will not make it to maturity since they can't be tended to. Likewise, the number of foraging bees is very crucial since they're responsible for gathering food. Furthermore, they discuss that the death rate of hive bees is negligible. It's possible that the 29-day average lifespan of bees only applies to foraging bees and therefore adds an interesting layer to consider.

The authors of this paper included details for the functions R(H, F), E(H, F), and F_0, where L = queen laying rate:

$$E(H,F) = L\left(\frac{N}{w+N}\right) = L\left(\frac{H+F}{w+H+F}\right).$$

$$R(H,F) = \alpha - \sigma\left(\frac{F}{H+F}\right).$$

$$F_0 = \frac{L}{m} - w\frac{J}{J+1}, \quad H_0 = \frac{1}{J}F_0 \quad \text{where}$$

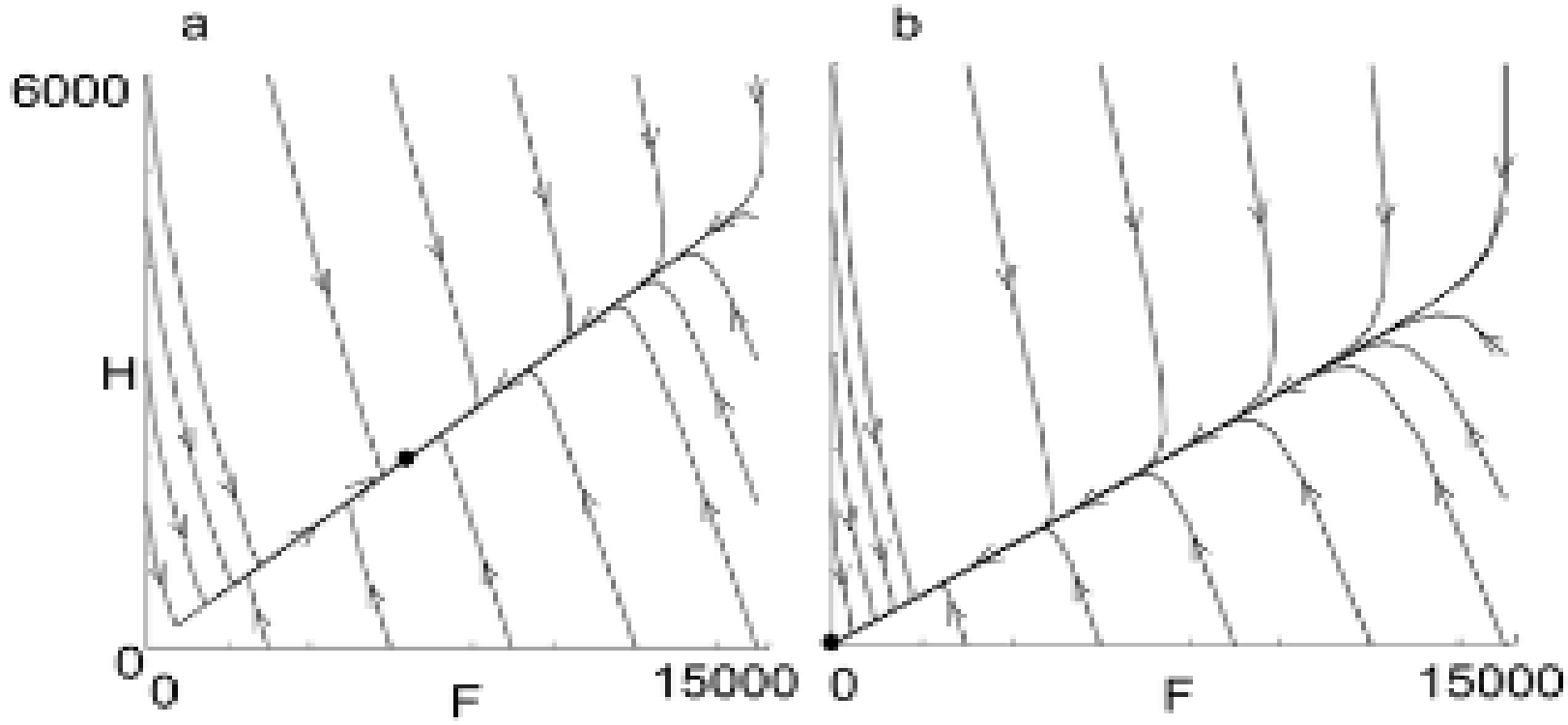
$$J = \frac{1}{2}\left[\left(\frac{\alpha}{m} - \frac{\sigma}{m} - 1\right) + \sqrt{\left(\frac{\alpha}{m} - \frac{\sigma}{m} - 1\right)^2 + 4\frac{\alpha}{m}}\right]$$

When

$$m < \frac{L}{2w}\left(\frac{\alpha+\sigma+\sqrt{(\alpha-\sigma)^2+4\frac{L\sigma}{w}}}{\alpha-\frac{L}{w}}\right) \quad \text{and} \quad \alpha-\frac{L}{w}>0.$$

They were able to use linear stability analysis to come up with the following phase plane diagram which represents the populations of hive bees by forager bees:

(Graphic 5, https://doi.org/10.1371/journal.pone.0018491.g003)



This shows that for parameter values L=2000, α=0.25, σ=0.75 and w=27 000, trajectories approach the dots. In the first picture, m = 0.24 and the population approaches a stable fixed point. For m = 0.40, all trajectories collapse into 0. Thus, they have narrowed the parameter m (forager death rate) to an amount which yields stable non-zero equilibrium.

The next model, 'Modeling Honey Bee Populations', by David Torres, Ulises Ricoy, and Shanae Roybal, includes a very interesting food scarcity component to their model which could prove useful. They begin by modeling the food supply with the following differential equation:

$$\frac{df}{dt} = s(t)(f_Q)pF + \gamma_L \sum_{i=4}^{i=8} w_i d_i - f_d, \;\; f \geq 0$$

They have essentially multiplied the consumption rate of each bee type (hive bee, forager bee, and larvae) by the number of them that exist—then sum up the total. This represents the change in their food supply. This is a very fair and logical model. I would like to use it and adapt it to also include the change in existing pollinate-able flora nearby, on which the hive survives, as a function of the number of forager bees collecting it as well as this overall consumption/demand.

The final model I investigated was called 'A Mathematical Model of Honey Bee Colony Dynamics to Predict the Effect of Pollen on Colony Failure', by Shahin Bagheiri and Medhi Mirzaie. Their approach was likewise split into different equations for each of the bee types. The following are those respective equations:

1. $$\frac{dB_o}{dt} = LS(.) - \phi_o B_o$$
2. $$\frac{dB_c}{dt} = \phi_o B_o - \varphi_c B_c - m_c B_c$$
3. $$\frac{dH}{dt} = \varphi_c B_c - HR_p(.) - HR_n(.)$$
4. $$\frac{dF_p}{dt} = HR_p(.) - m_p F_p$$
5. $$\frac{dF_n}{dt} = HR_n(.) - m_n F_n$$

In order, the equations represent the change in population of (1) uncapped brood; L = eggs lain by the queen and S(.) is the number of those eggs which survive, and the other term represents brood which either dies or gets capped over, (2) capped brood; in which $\phi_c B_o$ incoming capped brood, and "where $\phi_c B_c$ is the rate that young bees emerge from pupation per day, and the last term is the rate that capped broods die", (3) hive bees; "where $R_p(.)$ and $R_n(.)$are recruitment function, representing the proportional rate of hive bees that become pollen and nectar foragers, respectively. The death rate of hive bees is ignored because they are much safer than the external environment bees", (4) pollen foragers; "where the first term represents the hive bees to pollen foragers transition rate and the last term is the rate that pollen foragers die", and (5) nectar foragers; "where c is the maximum pollen brought in daily to the colony by each pollen forager.

The consumption of pollen by brood and hive bees is given by γ_{B_0} and γ_H, respectively." They have decided to focus on different kinds of foragers, both pollen (protein) and nectar (carbohydrates) which each serve function to satisfy the needs of the hive.

In conclusion, the models in the literature contained some very valuable insights into what could be made better about my model, namely getting more granular about the different bee types based off their role in or out of the hive. Likewise, the above models test many different control groups (such as the theoretical maximum if there were no predators, ample food and space, etc.) which should be what the model predicts if given these conditions are present in the bee's environment. However, my model proved to do a good job in predicting the stable, annual population cycle when compared to the data obtained from Randy Oliver. I am very pleased with the predicted population maximum during the peak season in my model and I will continue to explore deeper into what can be done to mathematically model the population dynamics of a beehive as it is something which is very interesting to me. The implications of this study could prove useful for those trying to optimize honey production, profitability, etc.. My bees are well on their way to the mid-Summer peak population, and I look forward to refining my model as I can. Here is a picture of one of my hive's frames full of capped-over brood:

References

1. Bagheri S, Mirzaie M (2019) A mathematical model of honey bee colony dynamics to predict the effect of pollen on colony failure. PLoS ONE 14(11): e0225632. https://doi.org/10.1371/journal.pone.0225632
2. Khoury DS, Myerscough MR, Barron AB (2011) A Quantitative Model of Honey Bee Colony Population Dynamics. PLoS ONE 6(4): e18491. https://doi.org/10.1371/journal.pone.0018491
3. Oliver, Randy. "Randy's Varroa Model." *Scientific Beekeeping*, 24 Sept. 2021, https://scientificbeekeeping.com/.
4. Torres DJ, Ricoy UM, Roybal S (2015) Modeling Honey Bee Populations. PLoS ONE 10(7): e0130966. https://doi.org/10.1371/journal.pone.0130966