General answer and precisions

Following the comments of the reviewer 1, we no longer use temperatures within the range of 32-36°C, but selected temperature data from sensors that were well surrounded by brood (on both adjacent sides of frames) so that we specifically analysed brood thermoregulation. The selection of these temperatures was made possible thanks to the colony evaluation, which gave information on the presence or not of brood on frames adjacent to the sensor. The results were similar excepted for the link between temperature variation and the number of bees in the colony, which was no longer significant.

For the third part of the paper (prediction based on data to help beekeepers estimating colony size), we kept the 32-36°C range selection method (temperature range for brood thermoregulation). Indeed, i) our goal was to help beekeepers to estimate the size of their colony without having to open the hive, an ii) they generally do not have access to the information of brood presence around the sensor.

Response to reviewers

Review by anonymous reviewer, 23 Sep 2022 19:22

The factors influencing honeybee colony thermoregulation are a topic of growing importance, because these need to be known if we are to model effects of changing climates on colonies and define strategies to breed climate-resilient stock. The present study is generally well-written (except maybe for the abstract), and the authors deploy an impressive set of statistical methods to model the efficiency of colony thermoregulation on size indicators and meteorological data. The main weakness of the study, apart from the fact that it is based on a very limited number of locations (1), colonies (39) and time points (9), is the fact that brood temperature was measured only at one – fixed - location within the colony.

One of the goals of the study was to assess whether automated temperature monitoring systems can be used by beekeepers to estimate colony strength, mainly to reduce the frequency of visits which might be stressful for colonies and saving time. However, it is unrealistic to expect beekeepers to place ten sensors in their hives. We, therefore, decided to study the data based on only one sensor, which is generally placed between central frames by beekeepers. In lines 110-111 we indicated that we tested whether the link between colony size and thermoregulatory capacities could be used to assess colony size in surveillance networks and beekeeping operations.

Our main hypothesis was that colony size (number of adult bees and amount of brood) positively influenced the efficiency of brood thermoregulation. To test this hypothesis, we did not need to have multiple environmental locations. We had 236 observations for 29 colonies: colony size estimations associated with temperature data, which, we believe, provided a strong dataset to test our hypothesis.

As the authors state themselves (L365 ff), the results they obtain (greater stability of temperature and a trend towards higher temp in stronger colonies) would also be expected under the hypothesis that the thermosensor in small colonies was frequently outside of the brood area.

We agree. We modified the data selection and specifically used temperatures from sensors that were well surrounded by brood (i.e. brood on both adjacent frames): see lines 146-148.

We no longer used temperature within the range of 32-36°C, but selected temperature data from sensors that were well surrounded by brood (on both adjacent sides of frames) so that we specifically analysed brood thermoregulation. The selection of these temperatures was made possible thanks to the colony evaluation, which gave information on the presence or not of brood on frames adjacent to the sensor. The results were similar excepted for the link between temperature variation and the number of bees in the colony, which was no longer significant.

A second major shortfall is that mean brood nest temperature is used as a proxy of "precision" of thermoregulation. In metrics, this term usually refers to the degree of dispersal of measurements around their mean (as opposed to accuracy, which refers to the average distance from a "true" value). In the present study, the proxy for what is termed precision is the mean temperature around the brood (line 311). This means that thermoregulation in colonies with a permanently elevated brood nest temperature would have been rated as "precise", and, according to the terminology of the study, this would have contributed to the "efficiency" of thermoregulation, even if the outside temperature would have been 40 °C, so that an efficient thermoregulatory system would be one that is able to keep the brood near its physiological optimum. In my opinion, it would have been more interesting to measure accuracy of thermoregulation, defined as the average difference from a value seen as the optimum (based on a literature review of physiological/developmental studies).

Thanks for the suggestion. We replaced the term 'precision of the temperature close to its brood optimum' by 'temperature close to the optimum value for honeybee brood rearing', whenever it was possible or 'brood temperature value'.

We understand the rationale for using the accuracy metrics, however, the optimal brood temperatures range between 32-36°C; 35±0.5°C being the optimal for pupae. Within the goal to estimate colony size remotely, this is not possible to know whether the brood is at the larval or pupae stage and therefore to use a specific optimal value for data analysis. We thus focused on average brood temperatures and hypothesize that the more there will be brood around the sensor the higher the temperature would be (within this range) due to higher brood stimulation. We added the hypothesis at the end of the introduction (lines 84-87).

Finally, the main conclusion – colony size matters for thermoregulation and weather also plays a role – is of limited novelty and value if the numerical boundaries of brood and bees that limit the capacity of the colony to maintain developmental stability of the brood are not given, and their dependency on outside temperature, irradiation and humidity is not determined in numbers. A statement like "colonies with fewer than XY workers were unable to maintain broodnest temp within the physiologically optimum range of X – Y °C if outside temp surpassed a value of XY °C" would be more useful.

We found that whatever their size, colonies are able to perform an efficient brood thermoregulation (even small colonies between 1512 – 13419 bees). And so, it was not possible to correctly assess colony size based on the brood thermoregulation efficiency.

Furthermore, we did not have failing colonies during our experiment (foraging season) and this was not our goal to study temperature outside the brood thermoregulation range. The relationship between colony temperature (all in-hive temperatures, i.e. no pre-selection of brood temperatures) and brood amount has been studied by Meikle et al (2015). They found that a strict control of temperatures (low temperature variation) was indicative of colonies with brood and large temperature amplitudes were indicative of colonies with little or no brood (see lines 425-427). Our goal was therefore to move further on this relationship and focus specifically on brood temperature.

Detailed comments can be found below.

Title: a little confusing, because it suggests that numbers are stable – instead, what you mean to dsay I believe is that something (temperature) stays stable because of a number (of bees) – maybe "stability through numbers"?

Based on the new analysis and results, we changed the title to: Brood thermoregulation efficiency is positively linked to the amount of brood but not to the number of bees in honeybee colonies

Keywords: are lacking. None required for this publication platform?

We did not find keywords requirement in the "Guide for the authors" page. However, this can be added if necessary. Keywords: *Apis mellifera*, homeostasis, colony monitoring, colony size, beekeeping.

L14: strictly speaking, thermoregulation means that temperature is regulated, not that it is kept stable – in the winter cluster, temp drops to values below brood rearing conditions, and still is regulated. We replaced 'maintain' by 'regulate'.

L19: increased efficiency of homeostasis This modification was included

L19: Therefore, we determined … This modification was included

L23: temperature has no brood optimum (being able to breed temperatures would be useful in the current energy crisis, though) – replace by "temperature close to the optimum value for honeybee brood rearing" This modification was included

L29: brood temperature – if you mean the temperature of brood, then this formulation is correct. If you mean the temperature at which brood rearing is possible, then you should write this. We indeed meant the temperature of brood.

L36-38: I do not understand why the fact that brood temperature constancy was insufficient for approximating colony size should signify that thermoregulation is flexible and efficient.

The relation between brood temperature regulation and colony size was not strong enough to make predictions, meaning that even small size colonies are capable of strict temperature regulation around the brood. Thus, colonies seem to be resistant to this shortcoming (small size), at least in terms of temperature regulation around the brood, showing a good efficiency even in less good conditions. But we removed the term flexibility to avoid any confusion.

L46: citation lacking Two citations were added : (Schmickl & Crailsheim, 2004; Stabentheiner et al., 2021).

L52: is, therefore, This modification was included

L55: optimum, This modification was included

L79: see above (L23); maybe "average deviation from the optimum brood rearing temp"?

We changed the term by "mean brood temperature, supposedly close to the optimum value for brood rearing" (line 23)

L87: is, therefore, This modification was included

L117: please indicate the subspecies/breed of bees used, as well as the size and material of hive boxes used – both may be important for thermal homeostasis

We added it in line 119.

L126: for the setup t properly determine brood temperature, it is clearly essential that the sensor is at the centre of the brood nest at all times. How was this assured, if the sensor was stably installed in a given position?

The ultimate goal was to assess whether beekeepers could monitor remotely their colonies based on the brood temperature. For this purpose, we changed our rational for the data selection and used brood presence around the sensor to ensure that the data represents the brood temperature: see lines 146-147.

L137: as this method is based on the "Liebefeld" method for measurement of colony strength, the original citation should be given here

The method we used was actually the ColEval method. We cited the corresponding paper describing the method (Hernandez et al 2020), which is based on the visual estimation in percentages of the area covered by the different colony parameters, and the posteriori correction for under- or over-evaluation made by the observer. We added the name of the method in line 133.

L150: this clearly means that the sensors were outside of brood nests in some cases. So how can you guarantee that they were firmly within brood nests in those cases where the temperature amplitude was smaller? It might be that the sensor was just at the border of the brood nest, so that fluctuations observed were stronger than they might have been if measured at the centre.

We agree and therefore changed our data selections. See our comments above and the method section (lines 146-147).

L165: what was the distance between the weather station and the hives? Was the microclimate the same? Were all the hives exposed to sunlight to the same degree?

The weather station was around 10 m away from the apiary (added line 170). All colonies were in the same apiary and disposed into a single row, so the microclimate was the same, as well as the sunlight exposure.

L190: please explain what an "identity link" function is. Identity link function is the link function where the model directly predicts the outcome. This is the canonical link function for the linear model.

L226: you state that you had no a priori-reason for considering predictor interactions – what about the well-known fact that both mean temp and stability of homeostasis drop in the absence of brood? There were also cases with 0 brood cells in the dataset…

These cases were not accounted for because the temperature data was now selected for presence of brood around.

L242: the latter This modification was included

L245: were individual measurements or colonies used as replicates here (n=29 or n=236)? Please indicate in the main text and not in the appendix. Individual measurements were used as replicates. We specified in lines 155-157.

Fig. 1, caption: please explain abbreviations used for predictors in the figure caption. What is meanS – shouldn't this be meanTS? This modification was included in Figure 3.

Fig.2: Seems like the model systematically overestimated the effects – doesn't this mean that something went wrong with model adaptation? The model first underestimates the mean temperature then overestimate the mean temperature. The sense of the bias is not always in the same direction.

L329: apparently there is a problem with a text mark here. This was corrected in the text.

Methods: it would be good to mention the boundaries of the classes of bees and brood in the main text – this information is important to interpret the results. This modification was included

L340: regression analysis is adequate if a cause-and-effect – relationship is to be studied… but is it realistic to use temperature as the explanatory variable here? I would expect that the efficiency of thermoregulation would be increased by the colony as a consequence of the presence of brood, not vice versa.

Although the sense of the cause-effect relationship would indeed be that the amount of brood influences the thermoregulatory behaviour of bees and therefore the temperature (and not the reverse), our study aimed to use the temperature as an indication of the size of the brood. One being influenced by the other, a relationship could allow us to make predictions.

L360 ff: this sentence does not seem to make sense to me – does the likelihood of the individual to respond to a stimulus not in itself depend on the intensity of the stimulus?

We agree; the sentence is confusing, and we revised it to (lines 368-370): “Social homeostasis, and thus the ability to thermoregulate, depends on the likelihood of individuals to react to a stimulus, which should increase with the intensity of the stimulus (exceeding the response threshold of more individuals)”.

L385-87: stated like this, the fact that weak colonies have a reduced capacity to thermoregulate is a banality. For your analysis to make sense, it would be necessary to discuss the boundary of the number of bees underneath which thermoregulation reaches a level of instability that could become detrimental to brood development. As this boundary likely depends on the level *of irradiation and outside temp, the critical threshold values of irradiation, outside temp and outside humidity (limiting the capacity to cool through evaporation) should be assessed.*

We understand, but our goal here was not to study colony failure in thermoregulation but to evaluate the link between colony size and thermoregulation efficiency, and potentially use this relationship to assess colony size. In addition, we did not have data on colony failure to make such predictions.

L424: an alternative explanation for this is would be that temperatures usually drop when there is precipitation, and heat may be the bigger challenge to temp stability than cold during the summer

We removed this part, which is now not necessary to the discussion of the results.

Review by anonymous reviewer, 12 Sep 2022 20:13

The article "Stability in numbers: a positive link between honeybee colony size and thermoregulatory efficiency around the brood" was found to be well written and well cited. The analyses were thorough that included appropriate modeling techniques. Moreover, this article addresses an interesting question of whether automated temperature monitoring systems can be used to predict colony strength in terms of brood and adult bee numbers.

The authors would like to thank the reviewer for the enthusiastic review of our work.

Introduction

The introduction is well written and well cited.

Methods

The experimental design matches the analysis. The analyses are performed correctly. Extra steps were taken to account for random effects and multicollinearity. The data was scaled to make it comparable for model analysis. The methods are well cited.

Results

Line 330: There is an error message with one of the references to the appendix.

The automatic reference system had an issue, this was repaired.

In figure 5, I would remove the p-values above the bars and would stick with the asterixis, it is too busy and too much information with the p-values added in.

This was modified for ease of reading. Non-significant results were also removed.

Discussion

The interpreted results match the analysis and study design. The discussion is written concisely and is to the point. An appropriate amount of speculation is used based on the results obtained.