Stability in numbers: a positive link between honeybee colony size and thermoregulatory brood thermoregulation efficiency around is positively linked to the amount of brood but not to the number of bees in honeybee colonies

Ugoline Godeau¹, Maryline Pioz¹, Olivier Martin², Charlotte Rüger³, Didier Crauser¹, Yves Le Conte¹, Mickael Henry¹, Cedric Alaux¹

¹ INRAE, Abeilles et Environnement, 84914 Avignon, France

² INRAE, Biostatistique et processus SPatiaux (BioSP), 84914 Avignon, France

³ ANSES, Epidémiologie et appui à la surveillance (EAS), 69364 Lyon Cedex 07, France

Corresponding author:

Ugoline Godeau, godeau.ugoline@gmail.com
Abstract

To ensure the optimal development of brood, a honeybee colony needs to maintain regulate its temperature within a certain range of values (thermoregulation), regardless of environmental changes in biotic and abiotic factors. While the set of behavioural and physiological responses implemented by honeybees to regulate the brood temperature has been well studied, less is known about the factors that may influence the efficiency of this thermoregulation. Based on the response threshold model of task allocation, increased efficiency of colony homeostasis should be driven by increases in group size. Therefore, we determined whether colony size (number of adult bees and amount of brood) positively influenced the efficiency of brood thermoregulation that we measured via two criteria: (i) the precision of the mean brood temperature, supposedly close to the optimum value for brood optimum rearing, and (ii) the stability of the temperature around this optimum. Finally, within the applied perspective of honeybee colony monitoring, we assessed whether the efficiency of thermoregulation could be used as a proxy of colony size.

For that purpose, we followed 29 honeybee colonies over two years, measured both brood and adult population size regularly over the beekeeping season, and monitored the in-hive brood temperature over the 24 hours preceding the inspections of these colonies. We then studied the effect of the size of the colony (number of adult bees and number of brood cells), as well as meteorological variables, on the efficiency of thermoregulation (mean and stability of brood temperature, i.e. between 32 and 36°C).

In addition to, we found a clear link meteorological conditions, we found that the mean and brood thermoregulation (mean temperature and the stability of this temperature were
both positively linked to the size of colonies. The mean brood temperature was also positively linked to the amount quantity of the brood, while its stability did not seem influenced by the size of the colony (number of bees). However, these relationships or brood size. The relationship between colony brood size and thermoregulation were: mean temperature was however too weak for clearly discriminating colony population size based solely on the brood thermoregulatory efficiency. This demonstrates an extremely high flexibility and efficiency of honeybee colonies to thermoregulate the brood regardless of the amount of brood and the group colony size.

Introduction

Homeostasis denotes the ability of living organisms to actively maintain steady internal conditions necessary for survival. A classic example of organismal homeostasis is the regulation of body temperature within certain boundaries, even when environmental temperatures change. Such a phenomenon is also found in social insects, like the honeybee *Apis mellifera*, for which the maintenance of nest conditions within a certain range of values (homeostasis), regardless of environmental changes in biotic and abiotic factors, is crucial for their colony development and survival. (Schmickl & Crailsheim, 2004; Stabentheiner et al., 2021).

The maintenance of inner hive conditions is one of the most crucial functions of honeybee colonies. While adult bees are rather eurytherms (*i.e.* can live under a wide range of temperatures), with a minimum of 18°C for normal muscle function (Esch & Bastian, 1968) and a maximum for survival above 50°C (Coelho, 1991; Kovac et al., 2014), the brood is stenothermic (*i.e.* only able to survive and develop within a narrow temperature range) (Seeley, 1985). Accurate temperature regulation
is, therefore, essential for proper development, with brood temperature strictly controlled within a temperature range of 32 to 36°C (Seeley, 1985) with regulation even more precise during the pupal period (35±0.5°C, Jones et al. 2004; Kronenberg & Heller 1982; Stabentheiner et al. 2010, 2021). Maintaining this optimal temperature window is crucial for the colony. Indeed, extended deviations are known to increase mortality (Koeniger, 1978; Wang et al., 2016), cause morphological defects (Fukuda & Sakagami, 1968; Himmer, 1932; Winston, 1987), disrupt synaptic organization in the brain of adult bees (Groh et al., 2004) and affect behavioural performances (Becher et al., 2009; Jones et al., 2005; Tautz et al., 2003). The colony, through the cooperation and coordination between individuals, therefore implements a set of behavioural and physiological responses to ensure proper temperature regulation of the hive (Jones & Oldroyd, 2006). When temperature is perceived as being too high, workers regulate it by fanning hot air out of the nest with their wings and may simultaneously spread water to induce evaporative cooling (Prange, 1996). At a finer scale, young workers can passively absorb heat by placing themselves between the heat source and the brood cells. This behaviour is called heat shielding, and it is usually carried out by placing the ventral side against the hot surface (Bonoan et al., 2014; Siegel et al., 2005; Starks et al., 2005; Starks & Gilley, 1999). When temperature is perceived as being too low, workers can contract their thoracic muscles to produce heat (Esch et al., 1991; Heinrich, 1980, 1985, 1993; Heinrich & Esch, 1994). Another efficient heating strategy consists of entering an empty cell to warm the adjacent cells containing brood (Bujok et al., 2002; Kleinhenz et al., 2003). Finally, during longer periods of cold, workers can cluster together and generate metabolic heat (Kronenberg & Heller, 1982; Meikle et al., 2016; Seeley & Heinrich, 1981; Stabentheiner et al., 2010).
The thermoregulatory mechanisms within the hive are therefore numerous, of different natures (behavioural, physiological or passive), flexible and interlaced, resulting in an effective brood temperature homeostasis (Kronenberg & Heller, 1982; Stabentheiner et al., 2021), even in extreme ambient conditions (e.g. Himmer, 1932 and Lindauer, 1955). However, temperature can fluctuate around its optimal value (Stabentheiner et al., 2021). Within the goal of maintaining temperature homeostasis, the efficiency of thermoregulation can be gauged through two criteria: (i) the precision of the mean brood temperature, supposedly close to its optimum value for brood rearing, and (ii) the stability of the temperature around this optimum value. Many studies have investigated how bees perform thermoregulation (see above), but little is known about the factors that can influence the efficiency of this thermoregulation. Based on the response threshold model of task allocation (Beshers & Fewell, 2001), the probability that an individual bee will engage in thermoregulation will depend on the level of the task stimulus and her threshold for that stimulus, i.e. the likelihood of reacting to the task–associated stimuli. A greater between-individual variability and within-individual consistency (specialisation) in task performance is therefore expected to increase behavioural homeostasis within the colony (Ulrich et al., 2018). This was confirmed by an increased stability in temperature changes within colonies composed of genetically diverse worker bees as compared to colonies with a low level of genetic diversity (Jones et al., 2004). A more recent study showed in ants that increased colony behavioral homeostasis is also driven by increases in group size (number of adult ants), likely via a stabilization of task performance frequency and a decrease in task neglect (Ulrich et al., 2018). We could therefore expect a similar influence of colony size on thermoregulatory efficiency in honeybees. A link between temperature regulation and the number of honeybees has been suggested (Seeley & Heinrich 1981; e.g. Southwick, 1985), as well as the
role of the stimulus intensity (here brood amount), but this remains to be investigated and characterized.

We therefore investigated in this paper whether the thermoregulatory efficiency around the brood was related to colony size (number of adult bees and brood amount). For that purpose, we monitored outside meteorological conditions, inner hive temperature and bee population level of several colonies over two years. We then investigated the influence of colony size on the thermoregulatory level (mean brood temperature) and stability (fluctuations around the mean brood temperature). Finally, by using the relationship between the thermoregulatory efficiency and colony size, we investigated whether the ability to regulate temperature around the brood could be used to estimate the colony size (for instance, whether high variability in thermoregulatory capacities could be an indicator of a relatively weak colony, and vice versa), without needing more data such as climate data. Indeed, within the context of severe colony losses observed around the world over the past years (Ellis et al., 2010; Neumann & Carreck, 2010; Potts et al., 2010), there is a clear need for surveillance networks and beekeeping operations to identify simple and non-intrusive proxies of colony state for monitoring and assessing their development and potential decline (López-Uribe et al., 2020). Such proxies could be extremely useful given that connected hives now allow us to monitor real-time data on physical variables such as weight, temperature, humidity and respiratory gases (Marchal et al., 2020; Meikle & Holst, 2015; Zacepins et al., 2011, 2012).

Methods
I. **Experimental setup and colony monitoring**

Experiments were performed at INRAE (Avignon, France, 43°540N-4°-520E) with hybrid honeybee colonies (a mix of *Apis mellifera*; *ligustica* and *Apis mellifera mellifera*). A total of 28 colonies, established in ten-frame colonies, wooden Dadant hives (L452 x L483 x H310 mm), were randomly selected from our local apiary in 2018 plus a new colony in 2019, for a total of 29 different colonies over the two years. Each colony was equipped with a temperature sensor (SHT35-DIS-B2.5KS, Sensirion AG) measuring in-hive temperature every 5 min. with a precision of 0.1°C within a temperature range of 20 to 60°C. The sensor precision was verified and validated beforehand using a climatic chamber. Colony strength was found to be more related to temperature data from sensors nearest to the geometric centre of the hive (Cook et al., 2022). Therefore, the sensor was placed in between the two central frames 5 and 6 and at mid-height, in order to be as close as possible to the brood, which generally occupy the central place in the hive.

The sensor was wired to a STM32 microcontroller (STMicroelectronics) and data were stored on a memory card (SanDisk Ultra SDHC 16 Go).

Colonies were inspected six times in 2018 (i.e. every two to three weeks between July and October) and five times in 2019 (i.e. approximatively every three weeks between April and July) by using the ColEval method (Hernandez et al., 2020) and to estimate three parameters: the number of open and closed brood cells and the number of adult bees. During colony visits, each side of each frame was visually inspected and the area covered by each of these parameters was reported as a percentage (one full side = 100%). Considering that a full side of a Dadant Hoffmann frame has a surface of 9.0311.34 dm² and contains in theory 1,400,400 bees and 3,100,400 brood cells, percentages were ultimately converted into number of open brood cells, number of closed brood
cells, and number of adult bees inside the hive (Alaux et al., 2018; Hernandez et al., 2020). Initial population size was different for each colony, and ranged from 4,536 to 40,131 adult bees, 15,200 to 44,250 open and closed brood cells in July 2018, and from 5,292 to 40,950 adult bees and 0 to 37,600 open and closed brood cells in April 2019 (Appendix S2 Figure S1, Appendix S3 Table S3).

II. Data analysis

1. Link between thermoregulatory efficiency, meteorological conditions and colony size

Thermoregulatory variables

Using temperature sensor data, we calculated the mean of in-hive temperatures reflecting brood temperature, we only kept temperature data from sensors that were surrounded by brood (containing brood on both adjacent sides of frames). Such information was retrieved from colony inspection (see above the ColEval method). We then calculated the mean of brood temperatures over the 24 hours preceding the day of colony evaluation (hereafter MeanT – Appendix S2 Figure S2). We chose a 24-hour time period since brood population can rapidly evolve over days (e.g. adult emergence). We therefore minimized the risk of having brood population changes between the temperature and population monitoring. We then discarded observations for which this mean (MeanT) was outside the range 32-36°C, since it is unlikely that brood develop near the sensor at such temperatures (Seeley, 1985), resulting in a total number of 236 observations for the 29 colonies. The mean brood temperature in the dataset was 34.3°C (min = 32.15, Q1 = 33.9, Median = 34.4, Q3 = 34.81, max = 35.85), and rarely exceeded 35°C (n = 29 events for 236 observations). We then calculated the coefficient of variation (CV; i.e. standard deviation expressed as a percentage of MeanT) within this same 24-hour period, to obtain...
a dimensionless variable representing the variation of temperature as a percentage of the mean
temperature. In the end, we had a total of 236 observations (replicates) of each variable (colony
population + temperature data) for the 29 colonies. The mean brood temperature in the dataset was
34.77°C (min = 33.00, Q1 = 34.62, Median = 34.82, Q3 = 34.96, max = 38.75), and exceeded
35.5°C in one case only (38.75°C). The mean temperature coefficient of variation in the dataset
was 0.74 (min = 0.21, Q1 = 0.50, Median = 0.68, Q3 = 0.87, max = 3.04). The final response
variables are therefore (i) the in-hive mean temperature (MeanT), representing the
thermoregulatory precision efficiency to attain the optimum brood rearing temperature and (ii) the
coefficient of variation of the in-hive temperature (CV), representing the thermoregulatory
stability.

**Predictor variables**

Both colony size and environmental meteorological conditions can potentially influence the
regulation of the in-hive temperature (Stabentheiner et al., 2021). Regarding colony size, we studied two predictor variables: (i) the number of adult bees (Nbees), and (ii) the total number of
brood cells (Nbrood - as the sum of the number of capped and uncapped brood cells). Regarding
meteorological conditions, we retrieved data from a local INRAE weather station, located ten
meters away from the apiary, and investigated the three following variables as relevant indicators
of environmental conditions: (i) the daily mean external temperature over the 24 hours preceding
the day of colony evaluation (temperature mean TM, in degree Celsius), (ii) the daily global
radiation (GR, in joule/cm², Burrill & Dietz 1981) and (iii) the daily precipitation (rainfall rate RR, in mm). In addition, to take into account a possible effect of
phenological advancement of the colony, we have adapted the cumulative growing degree-day,
usually used to estimate the growth and development of plants, to the foraging activity of bees,
which largely contributes to colony development. The cumulative growing degree-day (GDDcum) was calculated as the sum of mean daily temperatures (TM) above 12.5°C, from the beginning of each year, i.e. 2018 and 2019 (Appendix S2 Figure S3). A temperature of 12.5°C corresponds to the minimum temperature at which honeybee foraging activity starts (Vicens & Bosch, 2000). For a given year the GDDcum was calculated as follows (where t=1 corresponds to the 1st of January):

\[ GDDcum_t = \sum_{i=1}^{t} (\max(TM_i - 12.5, 0)) \]

Finally, colony replicate was included as a random effect to take into account potential variation in thermoregulatory capacity inherent to the colony (such as colony genetics, Jones et al. 2004).

The various meteorological predictors (TM, GR, RR and GDDcum) were tested in addition to the strength of the colony (via Nbrood and Nbees variables) to explain the thermoregulatory efficiency criteria: the mean temperature around the brood (MeanT) and the variability of temperatures around the brood (CV). For the latter, we also integrated MeanT as a predictor to assess its potential influence on thermoregulatory stability.

**Statistical analysis and model selection**

We conducted a model-averaging analysis in order to study in detail the predictors explaining the variations of MeanT and CV and their contributions. For this purpose, we first used generalized linear mixed models to model the relationship between the two response temperature variables (MeanT and CV) and the predictors. For MeanT, we specified a Gaussian distribution and the Identity link function. For CV, since this variable was continuous and severely skewed, we specified an inverse-Gaussian distribution with a Loga Inverse link function. Because variables are measured in different units, we centred and scaled (by dividing by the standard deviation) the
numerical variables when used as predictors (hereafter with an “S” at the end of their names). The two models are written as follows:

\[ g(\mu_{ij}) = \beta_0 + \sum_{h=1}^{p} \beta_h x_{hij} + \alpha_i \]

Where

- \( \mu_{ij} \) is the expectation of the variable \( Y_{ij} \)
- \( y_{ij} \) is the \( j \)th observation of the \( i \)th colony (either MeanT or CV),
- \( \beta_0 \) is the intercept,
- \( \beta_h \) is the regression coefficient for the \( h \)th predictor,
- \( x_{hij} \) is the \( j \)th value of the \( i \)th colony for the \( h \)th of \( p \) fixed-effect predictors,
- \( \alpha_i \) is the colony-specific effect and \( \alpha_i \sim Gaussian(0, \sigma^2) \),
- \( g \) is the link function (identity for the normal distribution and log for the inverse gaussian distribution).

We fitted the generalized linear models using the "glm" function (from the "stats" package - R Core Team 2021) for fixed effects models (without colony random effect), or the "glmer" function (from the "lme4" packages – Bates et al. 2015) for mixed models (with colony random effect).

In order to avoid multicollinearity, which is highly problematic in the case of model averaging (Banner & Higgs, 2017; Cade, 2015), we have excluded the possibility for the models to integrate simultaneously variables that introduce multicollinearity (for multicollinearity detection method...
see Appendix S1). In the end, the following pairs were not integrated into the same model: NBeeS and NBroodS, GRS and 
and GDDcumS and MeanTS.

A model selection procedure was applied by using the corrected Akaike Information Criterion (AICc). This procedure was done with the dredge function in the “MuMIn” package (Bartoń, 2020). Finally, we conducted model averaging based on AICc, a multimodel inference approach that allows one to derive inference from a subset of closely related best models, and not just from a single best model. Regarding the choice of the subset, we included models with a ΔAICc of less than seven points from the best model, grouping models that are likely to be the best models and that should all be used when making further inferences (ΔAICc<2) and models that are unlikely to be best models but that should not be discounted (ΔAICc ∈ [4,7], Burnham & Anderson, 2002).

Having previously forced some predictors not to be included in certain models, in order to avoid multicollinearities and a strong bias of underestimation of coefficients, we preferred the results of the conditional average. This model only averages over the models where the parameter appears, ignoring the cases where the model does not include the predictors when calculating the coefficients (unlike the full average for which the coefficients are set to 0.0 if the predictors are not included in the model). We did not consider any interaction between predictors because we had no a priori biological reason for doing so, and integrating these interactions, in particular with the random effects, was too ambitious in relation to the quantity of data available.

Based on the average models of MeanT and CV, for each predictor, we extracted its regression coefficient (to study effect size) and its P-value (to evaluate if its relationship with the studied variable is statistically significant). We also assessed the relative importance of each predictor by summing the Akaike weight across all the models in the set in which the predictor appeared. The
closer the sum is to 1.0, the more important the predictor is in the set of fitted models (Burnham & Anderson, 2004).

2. Thermoregulatory efficiency across colony size categories

Within the applied perspective of honeybee colony monitoring and by using the same dataset, we tested whether data on thermoregulatory capacities on their own (i.e. without the climatic information), could provide information on colony size. For that purpose, we transformed the two quantitative variables representing the strength of the colony (Nbees and Nbrood), into ordinal variables of four categories (respectively; catBees and catBrood) based on the quartiles (balanced in terms of number; Appendix S3 Table S3). Categories of catBees were distributed as follow: cat1 = [1512 ; 13419], cat2 = [13420 ; 19152], cat3 = [19153 ; 25326], cat4 = [25327 ; 47880]. Categories of catBrood were distributed as follow: cat1 = [0 ; 13800], cat2 = [13801 ; 20400], cat3 = [20401 ; 25326], cat4 = [27226 ; 46250]. We then assessed whether colony size categories were associated with specific in-hive temperature variables (MeanT and CV) by comparing the latter across the different categories representing the strength of the colony. Since the goal was to estimate colony size remotely and in that case it is not possible to know whether there is brood around the sensor, we selected MeanT (and corresponding CV) within the brood temperature range of 32-36°C. MeanT and CV data were not normally distributed (neither globally nor by categories), we applied the nonparametric test of Kruskal-Wallis to test whether the medians of the thermoregulatory variables (MeanT or CV) differ across colony-size categories. We also looked at the effect size of this Kruskal-Wallis test (as being the eta squared based on the H-statistic, with $\varepsilon^2 < 0.01$: very small effect, $0.01 < \varepsilon^2 < 0.08$: small effect, $0.08 < \varepsilon^2 < 0.26$: medium effect and $\varepsilon^2 \geq 0.26$: large effect, Cohen, 1988). In the case of significant results ($P < 0.05$), we
applied a Dunn post-hoc test to investigate multiple pairwise comparisons. These tests were carried out using the package “rstatix” (Kassambara, 2021).

Finally, we conducted ordinal logistic regression (with “clm” function from the “ordinal” package - Christensen 2019) in order to try to predict colony-size category (brood or number of bees) based on the in-hive mean brood temperature (with MeanT-scaled) data. We compared predictions of the models with the real observations and extracted accuracy (the proportion of all correctly classified validation points) and Cohen’s Kappa statistics \( \kappa = (\text{Observed Accuracy} - \text{Expected Accuracy}) / (1 - \text{Expected Accuracy}) \), which evaluate classification performance, taking into account the possibility of the agreement occurring by chance.

Statistical analyses were conducted on the filtered data (temperature of brood thermoregulation: MeanT ∈ [32, 36°C]), and therefore the sample size of each population category was no longer balanced (see Appendix S3 Table S8 & Appendix S3 Table S9 for the effective categories sizes).

Results

I. Link between the mean temperature around the brood, colony size and meteorological conditions

After setting aside models prone to multicollinearity, a total of 48 models remained (Appendix S3 Table S4S2) from which the average model was estimated (Appendix S3 Table S5S3). The combined outputs of the average model suggest that MeanT deviance was better explained by the scaled total number of brood cells (NbroodS), the scaled cumulative growing degree-day (GDDcumS), the scaled precipitation levels (RRS), the scaled global radiation (GRS),
the scaled mean external temperature (TMS), the scaled total number of brood cells (NbroodS), the scaled total number of adult bees (NBeesS), the scaled precipitation levels (RRS), and a colony random effect (sum of weights = 0.20). NbroodS, GRS and TMS had a significant positive effect on MeanT, contrary to GDDcumS and RRS that had a significant negative effect on MeanT (Figure 1). NbeesS, GRS and RRS had no significant effect on MeanT (Figure 1). When looking at the sum of Akaike weights, GRS (Figure 1), NBroodS and particularly GDDcumS, and RRS had high relative importance. The colony random effect had a low sum of weights (sum of weights = 0.05). The model was not very efficient in predicting the observed data, with weak MeanT strongly overestimated and strong MeanT heavily underestimated (Figure 2).
**Figure 1**: Left: Predictor estimates ($\beta$), and $P$-values ($P$) of fixed-effect parameters for the in-hive mean temperature ($\text{MeanT}$) average model. Bars represent the 95% confidence intervals of predictor estimates. Red points and bars are for predictors with significant effect at a level of 5%.

Right: Sum of weights across all models in the set where the variable occurred. $\text{NbroodS} =$ Scaled the number of adult bees, $\text{TMS} =$ Scaled daily mean external temperature, $\text{GRS} =$ scaled daily global radiation, $\text{RRS} =$ scaled daily precipitation, $\text{GDDcumS} =$ scaled cumulative growing degree-day.
II. **Link between the variability of the temperature around the brood, colony size and meteorological conditions**

After setting aside models prone to multicollinearity, a total of 318 models remained (Appendix S3 Table S6 S4) from which the average model was estimated (Appendix S3 Table S7 S5). The combined outputs of the average model suggest that CV deviance was better explained by the scaled mean in-hive temperature (MeanTS), the scaled precipitation levels (RRS), the scaled external mean temperature (TMS), the scaled number of adult bees (NBeesS), the scaled total...
number of brood cells (NbroodS), the scaled cumulative growing degree-day (GDDcumS), and a colony random effect. All models composing the average model included the colony random effect, which had a significant effect on CV. TMS had a significant positive effect on CV, contrary to MeanTS, RRS, and NBeesS which had a significant negative effect (Figure 3). All other predictors (NBroodS, RRS, NBees, GDDcumS and MeanTS) had no significant effect on CV (Figure 3). When looking at the sum of Akaike weights, (Figure 2), TMS, MeanTS, RRS and, to a lesser extent, NbeeS, RRS and NbroodS had high relative importance. The colony random effect also had a very high relative importance based on sum of weights (sum of weights = 1.0). The model was again not very efficient in predicting the observed data, with weak CVs slightly overestimated and the few strong CVs heavily underestimated, meaning that the model better explains weak CVs than strong CVs (Figure 4).
Figure 3: Left: Predictor estimates ($\beta$), and P-values ($P$) of fixed-effect parameters for the in-hive temperature CV average model. Bars represent the 95% confidence intervals of predictor estimates. Red points and bars are for predictors with significant effect at a level of 5%. Right: Sum of weights across all models in the set where the variable occurred. $NbroidS =$ Scaled the number of adult bees, $TMS =$ Scaled daily mean external temperature, $GRS =$ scaled daily global radiation, $RRS =$ scaled daily precipitation, $GDDcumS =$ scaled cumulative growing degree-day. $MeanTS =$ scaled mean brood temperature.
Figure 4: Coefficient of variation of in-hive temperature (CV) predicted by the average model as a function of the observed CV, with first bisector in blue (Predicted CV=Observed CV), deviations from this line in grey and regression line of the point cloud as a dotted red line.

III. Link between thermoregulatory efficiency and categorized colony sizes

We then assessed whether thermoregulatory efficiency (precision MeanT and its stability CV) differed between colony size categories (based on quartiles of brood amount: catBrood and of bee number: catBees) (Figure 5). We found significant variations across colony size categories with moderate magnitude for both MeanT (catBrood: P-value < 0.001, \(\epsilon^2 = 0.112239\) and catBees: P-value \(\leq 0.00100223\), \(\epsilon^2 = 0.07960499\), Kruskall-Wallis tests) and CV (catBrood only: P-value \(\leq 0.0010329\), \(\epsilon^2 = 0.4010\) and catBees: P-value < 0.001, \(\epsilon^2 = 0.0834\)).
efficiency of middle-0248). The MeanT was significantly higher in the largest size colonies (cat2 and cat3 of catBrood and catBees) did not consistently differ from thermoregulatory efficiency of (and cat4) as compared to the smallest and largest colonies and from each other (Figure 5). However, the MeanT and the CV were always significantly different between colonies of the smallest and largest size category (cat1) (for both catBees and catBrood). Significant differences in the MeanT were also observed between catBrood1 and catBrood2, as well as between catBrood2 and catBrood 4. However, CV values did not differ across colony size categories, both in terms of number of bees and brood amount (excepted between catBees2 and catBees4).

The effect of Nbrood or Nbees on the precision or stability of thermoregulation was not distinct enough to clearly discriminate different potential colony sizes. However, we can only notice that a low temperature (MeanT $\approx 33.8^\circ C$) combined with poor stability of this temperature (CV $\approx 1.67$) was somewhat associated with a weak colony at least in terms of number of bees (catBees cat1 $\leq 13,419$) bees, cf. Appendix S3 Table S8). In the same way, a temperature near the optimal brood temperature (MeanT $\approx 34.7^\circ C$) combined with a good stability of this temperature (CV $\approx 0.962$), seemed to indicate a strong colony, at least in terms of the amount quantity of brood (catBrood cat4 $\geq 27,225$ cat1 $\leq 13,800$) brood cells, cf. Appendix S3 Table S9S7). The analysis of the density plot (Appendix S2 Figure S1 Error! Reference source not found.) showed that, around below a MeanT of $35.34^\circ C$, colonies were unlikely to belong to the first category of NBrood. Conversely, below a MeanT of $34^\circ C$, colonies were unlikely to belong to the fourth category third and fourth categories of NBrood. Regarding the CV, we did not observe any clear discrimination of categories (NBrood and NBees, Figure 5 and Appendix S2 Figure S4).
Figure 5: Thermoregulatory efficiency (MeanT and CV) across categories of number of adult bees (catBees) and brood cells (catBrood). ns: not significant, stars indicate significant differences between colony categories (post-hoc Dunn test). **** P-value < 0.001, *** P-value < 0.005, ** P-value < 0.01, * P-value < 0.05, ns P-value ≥ 0.05. For category ranges see Methods and Appendix S3 Table S1.

Finally, the ordinal logistic regression showed that MeanT and CV were not significantly correlated to Nbrood categories (catBrood) was significantly correlated to categories of number of bees, catBees (MeanT: mean brood temperature (MeanTS P-value = 0.151; CV: P-value = 0.075), and the _ = 3.62e-11). However, the model predictions of the model based on the two predictors were predominantly incorrect in comparison with the observations (accuracy = 0.3136, _ = 0.0313; Table 1).
other hand, MeanT had a significant effect on Nbrood categories (catBrood) with ordinal logistic models (MeanT: P-value < 0.001; CV: P value = 0.601), but the model were not correct enough to make predictions were still not correct (accuracy = 0.34324492, κ = 0.0884246; Table 1).

Table 1: Confusion table for both ordinal logistic models predicting catBees or catBrood based on MeanT and CV. Grey highlights indicate true positives. For category ranges see Methods and Appendix S3 Table S2S1.

<table>
<thead>
<tr>
<th>Variable studied</th>
<th>Data observations catBrood observed</th>
<th>Model predictions catBrood predicted</th>
<th>catBrood 1 predicted</th>
<th>catBrood 2 predicted</th>
<th>catBrood 3 predicted</th>
<th>catBrood 4 predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>catBees</td>
<td>cat1</td>
<td>40</td>
<td>40</td>
<td>26</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cat2</td>
<td>5</td>
<td>5</td>
<td>50</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cat3</td>
<td>5</td>
<td>8</td>
<td>58</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>catBrood1</td>
<td>cat4</td>
<td>20</td>
<td>312</td>
<td>6</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>catBrood2</td>
<td>cat1</td>
<td>12</td>
<td>827</td>
<td>1529</td>
<td>216</td>
<td></td>
</tr>
<tr>
<td>catBrood3</td>
<td>cat2</td>
<td>22</td>
<td>4019</td>
<td>35</td>
<td>416</td>
<td></td>
</tr>
<tr>
<td>catBrood4</td>
<td>cat3</td>
<td>33</td>
<td>56</td>
<td>1326</td>
<td>4724</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cat4</td>
<td></td>
<td>1</td>
<td>5</td>
<td>38</td>
<td>11</td>
</tr>
</tbody>
</table>

Discussion
I. **Link between the mean temperature around the brood and colony size**

In order for the brood to develop normally, honeybee colonies need to regulate the brood temperature between 32°C and 36°C, and optimally at 35°C (Seeley, 1985). Brood temperature is therefore regulated within a narrow range of temperatures but according to the response threshold model of division of labour, we still expected that an increase in group size would generate a higher level of social homeostasis and therefore increased capacity of reaching optimal nest conditions, due to higher variability and task specialization between individuals (Ulrich et al., 2018). We did not find that mean brood temperature significantly increases with the number of adult bees within colonies. The link seems present but not strong enough in our data to be significant. However, the increase in thermoregulation mean temperature was significantly and positively related to the amount of brood. Social homeostasis, and thus the ability to thermoregulate, is not only due to the likelihood of individuals to react to a stimulus but also to, which should increase with the intensity of the stimulus and whether it exceeds (exceeding the individual response threshold of more individuals) (Theraulaz et al., 1998). Under this last scenario, it is possible that the stimulus intensity of the thermoregulatory tasks (brood amount) was high enough to surpass the threshold response of many individual bees, regardless of their respective thresholds. However, a non-mutually exclusive hypothesis is that the greater the quantity of brood, the greater the chance there is that the sensor is well surrounded by brood, and therefore to record optimal temperatures. By selecting temperatures within the thermoregulatory range, we expected to always have (containing brood in the vicinity on both adjacent sides of the sensor, which was confirmed during colony inspection, but a large amount of brood would exclude frames), and we therefore minimized situations in which the
temperature sensor would be located at the periphery of the brood patch where the brood temperature would be lower.

II. **Link between the variability of the temperature around the brood and colony size**

The variability of brood temperature was mainly linked to the mean temperature around which this variability was calculated: as the value of the brood temperature increased, approaching 35°C, the temperature variability decreased. This confirms that a finer-tuned temperature regulation occurs around the temperature of 35°C, which is the optimal temperature for pupal development (Jones et al., 2004; Kronenberg & Heller, 1982).

Variation in brood temperature was also correlated to the colony size (number of adult bees and, to a lesser extent, amount of brood), with larger colonies exhibiting reduced changes in brood temperature. This could be expected given that we first found that brood temperature increased with colony size (amount of brood, see above), and brood temperature increase was associated with lower temperature variability. However, when analysing the contribution of different environmental and colony variables, the ability to keep brood temperature stable seemed to be more linked to the number of adult bees (significant) than to the amount of brood (selected by models but not significant). This suggests that an increase in group size allows honeybee colony responses to better buffer against environmental fluctuation and therefore regulate hive temperature. This phenomenon could be attributed to differences among individuals, which generally increases with colony size in social insects. However, the variability of brood temperature was neither linked to the mean temperature around which this variability was calculated nor to the colony size (number of adult bees or amount of brood). This suggests that a decrease in group size does not particularly constitute an element
of loss of efficiency in brood thermoregulation. Bee colonies, regardless of their size, seem to have high abilities to keep a stable brood temperature. This phenomenon could be attributed in social insects to differences among individuals (Ulrich et al., 2018). Inter-individual behavioural variation was notably found to favour the collective control of nest climate in bumblebees (*Bombus terrestris*, Weidenmüller 2004). However, the contribution of group size to the stability of brood temperature was of low magnitude and needs to be confirmed in future studies with perhaps a greater range of colony sizes. More generally, we also found important inter-colony variation (with colony random effect of weight 1.0) in the ability of keeping brood temperature stable. These differences among colonies could be linked to various underlying reasons, for example, the exact location of the hive (more or less shaded) or the genetics of the bees (Graham et al., 2006; Jones et al., 2004).

### III. Link between thermoregulatory efficiency and environmental conditions

This efficiency of thermoregulation also depends on environmental conditions. The effect of the environmental temperature on the hive temperature has been highlighted previously. Stabentheiner et al. (2010) notably showed that the environmental temperatures have a non-negligible impact on the temperature regulation capacity inside the hive, in particular at the level of the brood. We also highlighted an influence of these environmental conditions on the regulation of brood temperature.

The external temperature had a non-negligible influence on both the mean brood temperature and the cumulative growing degree-day, which represents the quantity of foraging days, and so gives an indication of phenological advancement of the colony, was negatively associated with the mean brood temperature: the further we advanced in the beekeeping season, the lower the brood
temperature was. Such an association could be explained by the link between brood temperature and brood size, which declined between the spring and fall.

More interestingly, the external temperature had a high non-negligible influence on colony’s ability to maintain a stable brood temperature. Indeed, the hotter the environmental temperature, the hotter mean brood temperature was, and the more difficult it was for the colony to stabilise the brood temperature. This might be explained by the mechanisms used by colonies to compensate for high environmental temperatures, which consist of the collection and evaporation of water above the brood. Notably, in response to a simulated heat wave at 37°C (2°C above the optimal temperature), a 70% increase in forager traffic to sustain water needs was previously observed (Bordier et al., 2017). However, the efficacy of water collection not only depends on the foraging capacity but also on water availability in the environment (distance from the hive, water amount), which might lead to some degree of fluctuation in the regulation of brood temperature as compared to the more “passive” response to cold (changes in bee density and endothermy), especially in our experimental site characterized by high summer temperatures (between 35 and 40°C). In the context of climate change, where a general increase in temperatures is observed, this lower efficiency of colonies in maintaining optimal conditions for brood rearing may add another level of stress for honeybees and represent a cause of concerns for beekeeping activities.

The cumulative growing degree-day, which gives an indication of phenological advancement of the colony, was negatively associated with the mean brood temperature: the further we advanced in the beekeeping season, the lower the brood temperature was. Such an association could be easily explained by the link between brood temperature and brood size, which declined between the spring and fall. More interestingly, even though precipitation was not significantly associated with
an increase in the brood temperature, it seemed to substantially influence the maintenance of stable brood temperature, as the occurrence of precipitation was associated with a decrease in temperature variability. This observation confirms the influence of group size (number of bees) on the stability of brood temperature since higher numbers of bees are expected in the hive due to reduced or no foraging activity on rainy days. Finally, solar radiation seemed to influence the mean brood temperature, with an increase in radiation associated with an increase in mean temperature, likely due to the direct heating of the hive by the sun's rays on sunnier days.

IV. From thermoregulatory data to colony size evaluation

Measurements of in-hive temperatures have already been suggested and used for monitoring honeybee colony populations. It was notably found that the adult and brood mass of colonies were positively correlated with the in-hive temperatures (Cook et al., 2022; Meikle et al., 2016, 2017). Similarly, brood mass (but not adult mass) was inversely related to the amplitude of in-hive temperatures (Meikle et al., 2017). These results were obtained by including all in-hive temperatures (no pre-selection of brood temperatures ranging between 32 and 36°C). As a consequence, 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 (Meikle et al., 2017). The fact that we obtained similar results but on brood thermoregulation (i.e., link between mean brood temperature and brood amount) is promising within the goal of estimating colony population size. Indeed, in-hive temperatures could be used as a first filter to discriminate colonies with brood from colonies with little or no brood. Then, analysis of brood thermoregulatory efficiency could be used in a second step to evaluate in more detail the state of colonies with a relatively high amount of brood. The greater the brood temperature homeostasis,
the larger the colony would behave brood (as indicated by our results). By splitting colonies into size categories, we effectively found that large colonies with a large amount of brood had significantly better brood thermoregulation temperature than small colonies. However, when looking at the boxplot and density plot of colony size categories according to their thermoregulatory levels (Figure 5, Appendix S2 Figure S4), it was only possible to state that a temperature below 33.8°C was not indicative of large colonies in our dataset (≥27,225 adult bees or ≥27,225 brood cells were unlikely). While in-hive temperatures can easily discriminate colonies with brood from colonies with almost no brood, a higher level of colony-size discrimination was not possible when focusing on brood temperatures.

V. Conclusion

Brood thermoregulatory efficiency, more specifically increasing brood temperature, was associated with colony brood size, and the propensity to reach 35°C was more related to other hand, the amount size of brood near the sensor and brood did not seem to influence the temperature stability more related to of this temperature. Similarly, the number of adult bees. This highlights in the importance of increased group size notably for maintaining stable temperature conditions within the hive. However, when considering the size effect (magnitude), the influence of colony size on colony was not associated with an increase in thermoregulation was relatively marginal, which indicates efficiency (mean temperature or its stability). These results certainly highlight a role of the amount of brood on its thermoregulation, but above all they highlight the very high efficiency of the honeybee colony colonies to thermoregulate whatever the amount regardless of brood and the group their size.
As a consequence, the discrimination of colony population level based on brood thermoregulatory data was rather difficult. Nevertheless, within the applied perspective of honeybee colony monitoring, it would be useful to analyse a larger range of colony size variation, including depopulated or collapsing colonies, to fully conclude on the potential of brood temperature as a proxy for colony size estimation.

**References**


Acknowledgements

This project received funding from France Agrimer (ToolBee, 17-03-R) (MP, CR, DC, MH, YLC, CA). U. Godeau was supported by the European Horizon 2020 research and innovation program under grant agreement no. 817622 (B-GOOD). We are extremely grateful to Hervé Mugnier, Jean-Camille Lebreton, Aurélien Neveux and Sully Vitry for their technical assistance with the hive sensors.