



**HAL**  
open science

## Early warnings of bee colony failure

Théotime Colin, Vasilis Dakos, Andrew Barron, William Meikle, Eduardo Altmann, Tanya Latty

► **To cite this version:**

Théotime Colin, Vasilis Dakos, Andrew Barron, William Meikle, Eduardo Altmann, et al.. Early warnings of bee colony failure. 2024. hal-04794232

**HAL Id: hal-04794232**

**<https://hal.science/hal-04794232v1>**

Preprint submitted on 22 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Early warnings of bee colony failure

**Théotime Colin** (✉ [theotime.colin@gmail.com](mailto:theotime.colin@gmail.com))

Macquarie University <https://orcid.org/0000-0003-0223-4479>

**Vasilis Dakos**

ISEM, CNRS, Université de Montpellier, IRD, EPHE

**Andrew Barron**

Macquarie University <https://orcid.org/0000-0002-8135-6628>

**William Meikle**

USDA-Agricultural Research Service <https://orcid.org/0000-0002-1759-6950>

**Eduardo Altmann**

School of Mathematics and Statistics, The University of Sydney

**Tanya Latty**

University of Sydney

---

## Article

### Keywords:

**Posted Date:** January 12th, 2024

**DOI:** <https://doi.org/10.21203/rs.3.rs-3712390/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** There is **NO** Competing Interest.

---

# Abstract

How to identify ecological systems at risk of failure is a central question of modern biology and agriculture. Due to human impacts and global change, there is a growing need for early warning signals that identify when an ecological system is at risk of a state change before those changes become irreversible or extremely complex and costly to remediate. Honeybees (*Apis mellifera*) are an urgent case because our food crops heavily rely on them for pollination and annual bee colony losses reported by beekeepers across the globe are unsustainable. Given enough warning, beekeepers can rescue dying colonies, but early warning signals of death for individual bee colonies are lacking. Here we used early warning signals to investigate whether fluctuations and dynamical patterns in internal hive temperature can be used as an early indicator of impending colony failure. Across three distinct datasets we found that temperature regulation of failing colonies was different enough to distinguish them from healthy colonies weeks before they died. This signal comes early enough to intervene and assist colonies with standard beekeeping practices. Our study shows that early warning theory can help to identify practical signals of risk of state change even in systems that change state relatively rapidly, such as a dying bee colony, early enough to intervene and prevent losses.

## Main

Global change is exposing ecological systems to new stressors and new magnitudes of stress. These stressors include landscape degradation, the spread of introduced pests and climate change. In response to chronic stress, ecological systems can often manifest what appears to be a “tipping point” response: ecological systems are resilient to stress up to a point, beyond which they show rapid and dramatic changes of state. These dramatic changes could be population collapses of whale populations<sup>1</sup>, bee colonies<sup>2</sup>, eutrophication of lakes<sup>3</sup>, tree diebacks during droughts<sup>4</sup>, tree deaths in forests<sup>5</sup> and transitions from forests to savanna<sup>6</sup>. What we need are practical methods to detect ecological systems at risk of passing a tipping point while we still have time to employ protective management to avoid such risk.

Early warning theory predicts that changes of states may often be detectable before a system passes a tipping point that leads to a new state<sup>7</sup>. As a system approaches a tipping point, external perturbations of the system should result in variability and higher correlation in the system dynamics (Figure 1)<sup>3,8</sup>. Early warning signals are receiving increasing attention in ecological and agricultural systems but practical examples outside of the labs are still scarce<sup>7,9</sup>. The few examples of early warning signals applied to field environmental or biological systems are based on multi-year studies and strongly rely on modelling, which has perhaps slowed down practical applications in ecology. Non-experimental examples include the transition between savannah and forests<sup>10</sup>, the eutrophication of lakes<sup>3</sup>, and the death of trees<sup>11</sup>. But if the theoretical framework behind early warning signals is correct, indicators of vulnerability to state change should be common and detailed knowledge of a system should not be needed to detect them.

Here we examined whether we could identify a practical and measurable early indicator of a Western honeybee (*Apis mellifera*) colony at risk failure. Honeybee colony losses are an urgent case where early

warning signals may help effective intervention. The death of bee colonies happens on a much shorter time-scale<sup>2</sup> than the transition of trophic states in lakes, the death of trees or plant community changes at the landscape levels. Bee colony failure is a relatively rapid process: an apparently healthy and well-functioning bee colony can decline in weeks, which has caused several authors to label it as a “collapse”. Globally, honey bee colony failure rates have increased in industrialised nations in the last decade<sup>12</sup>. At the same time, many industrialised nations have increased their reliance on honey bees for pollination service for food crops<sup>13</sup>. Hence there is concern that high honey bee colony failure rates could pose a risk to food chain security<sup>14</sup>. Traditionally, an assessment of colony condition requires opening the hive every few weeks for a manual inspection of the frames. This has become too labour intensive to be sustainable at the scale at which colony failure now occurs. At the same time, electronic sensors specifically designed for hive monitoring are increasingly available. If there were a practical early indicator of a colony at risk of failure, in time, colony failures could be reduced or managed using the wide range of already existing beekeeping techniques.

Temperature regulation is considered an indicator of bee colony health in ecotoxicology<sup>15,16</sup> because it directly reflects the capacity of the colony to maintain optimal brood development conditions. In warmer months, honeybee colonies keep the temperature of the brood between 33°C and 36°C. In Winter, the amount of brood is reduced<sup>17</sup> and colonies maintain the temperature of the hive at a lower average, with higher fluctuations. The presence of equilibria in this easily measurable and actively controlled variable of the honeybee colony makes temperature an ideal candidate to explore early warning signals of colony failure (Figure 2). When bee colonies become stressed, we expect their ability to regulate the temperature to decrease, leading to higher temperature fluctuations and higher minimum and maximum (Figure 1). We also expect the average temperature to decrease in stressed colonies which may have less brood and thus require less heat or may be unable to warm up the hive.

Here, we use three independent datasets containing temperature time-series, brood population estimates and census data for a total of 97 bee colonies to determine how dying, weak, and healthy bee colonies controlled temperature. We compare the performance of simple summary statistics used as early warning signals as classifiers of the survival and health status of colonies. Next, we simplified our approach by setting a classification threshold relevant to real-world beekeeping practices, to test the practicality of our method. Our results confirm that early warning signals based on hive temperature are useful predictors of hive collapse.

## Results And Discussion

We measured the two most used early warning signals<sup>8,11</sup>, changes in the autocorrelation and standard deviation, along records of daily beehive temperature. Standard deviation and autocorrelation are expected to increase when the system is close to a tipping point due to an increase in the sensitivity of the system to stochastic fluctuations or to a lower capacity of the system to follow rapid environmental changes<sup>8,11</sup>. We also calculated minimum and maximum temperature within hives, which show the

colonies' ability to control temperature during the coldest times of night and the hottest times of the day. We also used average temperature as a proxy for the overall capacity of colonies to maintain a warm, stable temperature. Temperature values were calculated over a full daily biological cycle, from midnight to midnight, to simplify analyses and help with interpretation (Figure 2a).

To compare the utility of temperature measures at detecting failure, we used three datasets ("California 2016", "Sydney 2018", "Sydney 2020") where the survival outcome and brood population were known for each colony. In our analyses, we first ignored this knowledge to simulate a real-time scenario similar to what beekeepers would experience. We then checked our results against the known outcomes.

Simple thresholds could not be used to predict the death of colonies as these parameters are expected to vary in all bee colonies due to changes as colonies approach winter and transition to a natural overwintering state<sup>18</sup> and as a result of extreme weather changes. This is visible on Figure 2a, where all colonies, regardless of their survival outcome, experience a drop in average temperature from Autumn to Spring. Instead, we compared the daily values of temperature parameters (daily autocorrelation, daily standard deviation, daily minimum, daily maximum and daily average) for each single colony compared to all other colonies from the same experiment and regardless of their survival outcome to obtain an individual z-score for each individual hive on each specific day. The z-score represents the difference between a single colony and the mean value of all other colonies in the same environmental conditions, for any given variable.

This early warning method requires selecting a Z-score threshold above which colonies are considered at risk of dying. To show how different Z-score thresholds changed the rates of false positives and true positives, we varied a threshold from 0 to the maximum Z-score obtained. We used the rates of true and false positive hive death predictions to build a Receiver Operating Characteristic (ROC, Figure 2b) and calculate the associated Area Under the Curve (AUC) which quantifies the performance of our early warning signals at different Z score thresholds. For our largest dataset, the California 2016 dataset, the AUC was the highest for the daily average (0.65), followed by the daily minimum (0.64), the daily maximum (0.61), the daily autocorrelation (0.59) and the daily standard deviation (0.58). Similar results were obtained for the Sydney 2018 and Sydney 2020 datasets (Supplementary material 1). Although these numbers may seem rather low, they correspond to the false assumption that a hive recorded dead or with low brood at the end of the experiment was always in this state, from the beginning of the experiment. Given that the nature of the colonies at the beginning of the experiment could not be differentiated by the beekeeper or the observer (only colonies visually assessed to be of similar "strength" were used), the information of when colonies started to fail cannot independently inform our ROC. This lowers the AUC for all variables measured.

While we think that determining a z-score threshold should be a choice left to end-users to optimise their own costs and benefits in using our method, we have determined, for each dataset, a z-score to threshold simply to illustrate our findings. We measured, from our ROC curves, the threshold that corresponded to the first decrease in the growth rate of the true positive to false positive ratios. We ran a piecewise

regression on the ROCs of the daily average and determined the Z-score threshold corresponding to the main breakpoint for each dataset. To limit the effects of unusual weather events or beekeepers opening hives, we considered that a colony should exceed the Z-score threshold for three consecutive days to be considered at risk of dying, regardless of future measurements. Using these values on the largest of our datasets (the California 2016 dataset), where colonies were cared for by a commercial beekeeper, we were able to detect 5/6 colonies that died later (true positives), 10/14 low brood colonies (true positives), and got 9/40 false positives (false positives; note that in this dataset, these might have initially been failing colonies that were rescued by a beekeeper) (Figure 2c). Among the five dead colonies, the colonies were on average first detected at risk of dying 98 days before death was confirmed (29, 35, 61, 155 and 210 days). Colonies in the California 2016 dataset were checked monthly outside of November and December, with most of them dying in late Autumn or Spring when they were subject to regular inspections (Figure 2a). Such warnings would give beekeepers a practical time period in which to act to rescue their colonies. When expressed as the percentage of days with warnings for each group, we found that warnings occurred in 35.4% of the days for the dead colonies group, 20.3% of the days for the low brood group, and that 1.0% of the days had false positive warnings in the healthy colonies group.

Similar results were obtained in Sydney in 2018 and 2020 (Supplementary Material 1). In the Sydney 2018 dataset, we detected 8/9 colonies that died later, some of the low brood colonies (2/4), but false positives occurred in the healthy colonies group (5/8). The probability that most colonies would have at least one false-detection increases with time, and it is not surprising that this dataset, the longest of all, has the largest number of false-positives. These false positives were however short-lived. False positives in the healthy colonies group represented only 2.0% of all days overall, while warnings occurred for 25.1% of days in the dead colony group and 6.3% of days in the low brood group. The eight colonies that failed were detected on average 56 days before they died (9, 9, 30, 33, 59, 59, 119 and 133 days before they were observed dead or were considered too small to recover and were removed to prevent the spread of disease or pests).

Bee colonies in the Sydney 2020 dataset experienced higher than expected mortality, likely due to massive forest fires, which exposed our colonies to significant amounts of smoke. Even though deaths occurred during the summer, which was comparatively unusual, and with a smaller sample size, our method was able to detect 3/3 that later died. We detected one low brood colony (1/2), and a few false positives occurred in the healthy colonies group (3/10). Among the three dead colonies detected as failing, hives were detected on average 20 days before death was confirmed (8, 21 and 31 days). Estimating the overall percentage of daily warnings for each group, we found that dead colonies were detected as dying 58.6% of the days, low brood colonies 20% of the days, and false positives represented 3.8% of the total number of days in the healthy group. The efficacy of temperature as an early warning signal during summer was surprising, given the expectation that honeybees would not struggle to warm their hives during warmer months. However, in Sydney it is rare that ambient temperature reaches 35°C at night, so bee colonies that were failing to warm their hives still had different temperature profiles from other, presumably healthy, bee colonies.

The cause underlying the lack of temperature regulation in failing bee colonies is unclear. Here, we purposefully ignored known failure mechanisms and additional information about beekeeping practices (hive migration) and environmental stress (food stress and pesticide exposure) as we were interested in developing an early warning indicator that functioned irrespective of the specific stressor(s). Typically, the development of early warning systems has relied on understanding the existence and mechanisms behind tipping points to model the system. Obtaining and integrating such knowledge in early warning systems likely improves their accuracy and is an inviting scientific prospect but is not practical for most of the pressing environmental issues we face. A requirement to include extensive information about stressors and their impacts would most likely hamper the development of real-world applications, where extensive data collection is often unfeasible. Instead, we showed that useful early warning signals can be obtained for practical applications even without refining their interpretation with the knowledge needed to build mechanistic models, qualify tipping points and demonstrate critical transitions.

Here, we determined a threshold based on the change in the ratio of true to false positive simply to illustrate our findings. We do not wish to recommend a particular threshold, however, because it influences the number of false-positives (leading to unnecessary efforts from beekeepers) or false-negatives (decreasing the number of colonies rescued). This trade-off would be better solved by stakeholders who are likely to have a variety of cost-benefit ratios depending on place and time.

Early warning signals can guide the development of markers of biological systems at risk of failing or transitioning to sub-optimal states. Our application of this method to detect failing bee colonies is the first example of the use of early warning signals in a biological or agricultural system where it can realistically be used to minimise losses. We hope this will encourage the development of similar methods in related fields. We suggest that the development of early warning tools for critical transitions in biological and eco-systems should focus on measuring and analysing processes that are highly dynamic, self-regulated by the system and very variable. In our changing world, there are an abundance of biological systems similar to the bee colony, where stable equilibriums occur, and failures have costly consequences to society. The insights from this early warning study demonstrate that early warning signals can be applied to rapid processes in a meaningful way to solve urgent environmental and agricultural issues, such as the unsustainable losses of bee colonies worldwide.

## Methods

### Sites

We used temperature measurements from the centre of 97 bee colonies from three experiments at two locations. Full details on the methodologies used to collect these data are available in the original publications for the California dataset (2016-2017 experiment)<sup>19</sup> and the Sydney 2018 dataset<sup>20</sup>.

Two of these experiments were located in Sydney, Australia, and received an experimental treatment (pesticide or control, food stores removal or control) to reproduce environmental stress, while in the third

experiment hives were located in California, USA and not subject to any treatments but were dispatched in different crops and were managed independently by a beekeeper<sup>19</sup>.

In the first experimental dataset, in 2018, temperature was measured from 21 hives during a year<sup>20</sup>. 7 colonies were treated with a commercial formulation of tau-fluvalinate, 7 with a commercial formulation of thymol and 7 were kept as control. One control hive was removed from the dataset following storm damage. Of the remaining 20 hives, seven died including 2 in the tau-fluvalinate group, 3 in the thymol group, and 2 in the control group. This study concluded there were no important effects of the treatments at the colony-level<sup>20</sup>.

In the second experimental dataset, in 2020, all the frames containing only food stores were removed every fortnight from 8 hives, while these stores were left in another 8 colonies. One colony was removed from the dataset because it absconded on a hot day when a nearby bushfire emitted a large amount of smoke around the hives. Three colonies were considered dead including one in the food removal group and two in the control group. This second experiment only lasted two months due to delay caused by bushfires and abruptly ending the experiments because of a COVID-19 lockdown. In these two experimental datasets, the distribution of colony mortality between treatment groups suggests no or weak effects of the treatments on colony survival.

In the California dataset, 60 hives were managed by a commercial beekeeper in California, USA over a year, in a migratory pollination operation described in an associated publication<sup>19</sup>. Hives were assessed once in October, January, February and March. We acknowledge this creates a long gap in our records (November and December) but cold winter prevented regular access to the bee colonies. This study found differences in weight gains in different crops and argued temperature monitoring provided useful information to compare hives from different groups.

## Measurements

Measurements were taken every 15 min using sensors with a 0.0625°C precision (iButton Thermochron model DS1922), which corresponds to 96 data points per day for each hive. Every two months, iButtons were removed, data saved, and iButtons were reinitialised and placed back into the hives as soon as weather and practical considerations permitted. This occasionally created small gaps of about 48h in the temperature records. When supplies permitted, a new set of iButtons was installed in the hives after the previous set was collected, reducing the gap in the temperature record. The first few records of each hive button were removed to avoid capturing data corresponding to the colony warming up the hive immediately after opening or to the sensor's casing warming up.

Regular inspections were conducted during which brood frames were brushed free of bees, photographed and capped brood numbers were measured using combcount<sup>21</sup> and used as a proxy for the overall brood amount in the hives. The honeybee's immature stage typically requires a 21 days development window including 9 days where larvae are uncapped and 12 where brood larvae are capped, during and after



pupation<sup>22</sup>. Brood amount data from the last evaluations were used to classify colonies into “low brood amounts” colonies and “healthy colonies”. This was done based on expert opinion separately for each dataset based on a visual comparison of brood patches between hives of the same datasets and corresponded to gaps in the distribution of colonies’ brood amounts, which were found to resemble a left-skewed normal distribution (SM1).

Colonies in the Sydney experiments were censused at least every two weeks in the Sydney 2018 dataset during evaluations or by assessing flight activity on Sunday days during winter, and weekly in the Sydney 2020 dataset during food store manipulations. Colonies were considered dead and removed when they contained less than 300 workers to prevent the spread of pests and diseases between hives in Sydney. Hives in the California 2016 dataset were checked at least once a month in August, October, January, February and March. In November and December<sup>19</sup> temperatures were too cold to census colonies.

## Analyses

All analyses were conducting in R v4.2.3, data handling was done using ‘tidyverse’ v2.0.0 and ‘lubridate’ v1.9.2, figure panels were built using ‘cowplot’ v1.1.1, breakpoints for the ROC curves were obtained using ‘segmented’ v1.6.4.

Average, min, max, standard deviation and autocorrelation were calculated for each colony and day using base R function (‘mean’, ‘min’, ‘max’, ‘sd’ and ‘autocor’ with ‘lag=1’). For each colony and each day within each of the three datasets, a daily Z-score was calculated by calculating the average value and standard deviation of the focal variable for all colonies on the day at the exception of the focal one on this day. The z-score was calculated daily as the absolute difference between the average of the variable of one focal hive and that of all the other hives, divided by the standard deviation of the variable for all the other hives on a specific day. The z-score is a standard score which corresponds to the difference between  $x$ , the observed value and  $\mu$ , the mean of the sample divided by the  $\sigma$ , the standard deviation of the sample:

$$Z = \frac{x - \mu}{\sigma}$$

We included all colonies in the calculation of the z-score, regardless of their outcome in terms of survival or brood population, to place ourselves in a real situation where a beekeeper does not know the health status of bee colonies, as it is the end measurement they are trying to obtain.

Since the z-score in itself is an absolute value, a z-score threshold is needed to classify observations in groups. To quantify the effect of a threshold on the ratio of true positive to that of false negative rate we used a receiver operating characteristic (ROC) curve<sup>23</sup>. Receiver operating characteristic curves are used to show the performances of a classifier at every possible classification value. For each focal variable, we calculated the true positive rate (number of true positives divided by the sum of the number of true positives and false negatives) and the false negative rate (number of false positive divided by the sum of

the number of true negatives and false positives) for each threshold by varying thresholds from 0 to the maximum z-score measured for the focal variable, with a step of 0.1.

The shapes of our curves usually showed a clear breakpoint for low values of true positive rates and false positive rates after which curve growth slowed down. Typically, this breakpoint is used as a threshold because it corresponds to the point after which the true positive rate grows slower than the false positive rate. While we preferred to avoid providing one such threshold as a recommended threshold value, we chose to provide one example with a breakpoint-based example to calculate an example of the number of days that bee colonies were detected as dying before they died. Using this breakpoint is imperfect, and in our case the false positive rate is inflated because we have no knowledge of when colonies that died during experiments stopped being “healthy”, but it allows us to provide a neutrally selected threshold value. The ROC breakpoint was calculated for the daily average measurements of each dataset (the value with the highest overall AUC in all three datasets) using the ‘segmented’ function of the ‘segmented’ library and we retained the threshold value the closest to this breakpoint.

We then applied this threshold to the z-scores of the average daily temperature. Colonies with Z-scores above the thresholds were considered at risk of death. This led to flickering warnings during hot and cold weeks in all colonies, so we removed warnings that occurred for less than three days in a row. We then calculated the number of days from the first warning to the observed death of the colony or when it was considered too weak to survive and had to be terminated to prevent the spread of pests.

## References

1. Clements, C. F., Blanchard, J. L., Nash, K. L., Hindell, M. A. & Ozgul, A. Body size shifts and early warning signals precede the historic collapse of whale stocks. *Nat. Ecol. Evol.* **1**, 1–6 (2017).
2. Perry, C. J., Søvik, E., Myerscough, M. R. & Barron, A. B. Rapid behavioral maturation accelerates failure of stressed honey bee colonies. *Proc. Natl. Acad. Sci.* **112**, 3427–3432 (2015).
3. Wang, R. *et al.* Flickering gives early warning signals of a critical transition to a eutrophic lake state. *Nature* **492**, 419–422 (2012).
4. Camarero, J. J., Gazol, A., Sangüesa-Barreda, G., Oliva, J. & Vicente-Serrano, S. M. To die or not to die: Early warnings of tree dieback in response to a severe drought. *J. Ecol.* **103**, 44–57 (2015).
5. Cailleret, M. *et al.* Early-warning signals of individual tree mortality based on annual radial growth. *Front. Plant Sci.* **9**, 1–14 (2019).
6. Hirota, M., Holmgren, M., van Nes, E. H. & Scheffer, M. Global Resilience of Tropical Forest. *Science (80- )*. **334**, 232–235 (2011).
7. Scheffer, M. *et al.* Early-warning signals for critical transitions. *Nature* **461**, 53–59 (2009).
8. Dakos, V., van Nes, E. H., D’Odorico, P. & Scheffer, M. Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology* **93**, 264–271 (2012).

9. Dakos, V. *et al.* Tipping Point Detection and Early-Warnings in climate, ecological, and human systems. 1–35 (2023).
10. Hoffmann, W. A. *et al.* Ecological thresholds at the savanna-forest boundary: How plant traits, resources and fire govern the distribution of tropical biomes. *Ecol. Lett.* **15**, 759–768 (2012).
11. Dai, L., Vorselen, D., Korolev, K. S. & Gore, J. Generic Indicators for Loss of Resilience Before a Tipping Point Leading to Population Collapse. *Science (80-. )*. **336**, 1175–1177 (2012).
12. Goulson, D., Nicholls, E., Botias, C. & Rotheray, E. L. Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. *Science (80-. )*. **347**, 1255957–1255957 (2015).
13. Aizen, M. A. & Harder, L. D. The Global Stock of Domesticated Honey Bees Is Growing Slower Than Agricultural Demand for Pollination. *Curr. Biol.* **19**, 915–918 (2009).
14. Klein, A. M. *et al.* Importance of pollinators in changing landscapes for world crops. *Proc. R. Soc. B Biol. Sci.* **274**, 303–313 (2007).
15. Colin, T., Meikle, W. G., Paten, A. M. & Barron, A. B. Long-term dynamics of honey bee colonies following exposure to chemical stress. *Sci. Total Environ.* **677**, 660–670 (2019).
16. Meikle, W. G., Colin, T., Adamczyk, J. J., Weiss, M. & Barron, A. B. Traces of a neonicotinoid pesticide stimulate different honey bee colony activities, but do not increase colony size or longevity. *Ecotoxicol. Environ. Saf.* **231**, 113202 (2022).
17. Meikle, W. G. & Holst, N. Application of continuous monitoring of honeybee colonies. *Apidologie* **46**, 10–22 (2015).
18. Döke, M. A., Frazier, M. & Grozinger, C. M. Overwintering honey bees: biology and management. *Curr. Opin. Insect Sci.* **10**, 185–193 (2015).
19. Meikle, W. G., Weiss, M. & Beren, E. Landscape factors influencing honey bee colony behavior in Southern California commercial apiaries. *Sci. Rep.* **10**, 1–16 (2020).
20. Colin, T. *et al.* Effects of late miticide treatments on foraging and colony productivity of European honey bees (*Apis mellifera*). *Apidologie* **52**, 474–492 (2021).
21. Colin, T., Bruce, J., Meikle, W. G. & Barron, A. B. The development of honey bee colonies assessed using a new semi-automated brood counting method: Combcount. *PLoS One* **13**, 1–14 (2018).
22. Winston, M. L. *The biology of the honey bee*. (Harvard University press, 1987).
23. Boettiger, C. & Hastings, A. Quantifying limits to detection of early warning for critical transitions. *J. R. Soc. Interface* **9**, 2527–2539 (2012).

## Figures

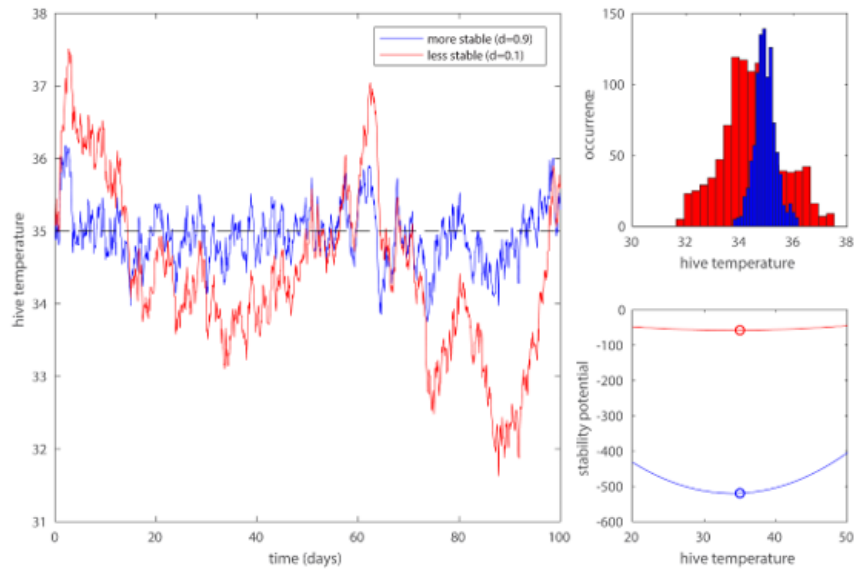
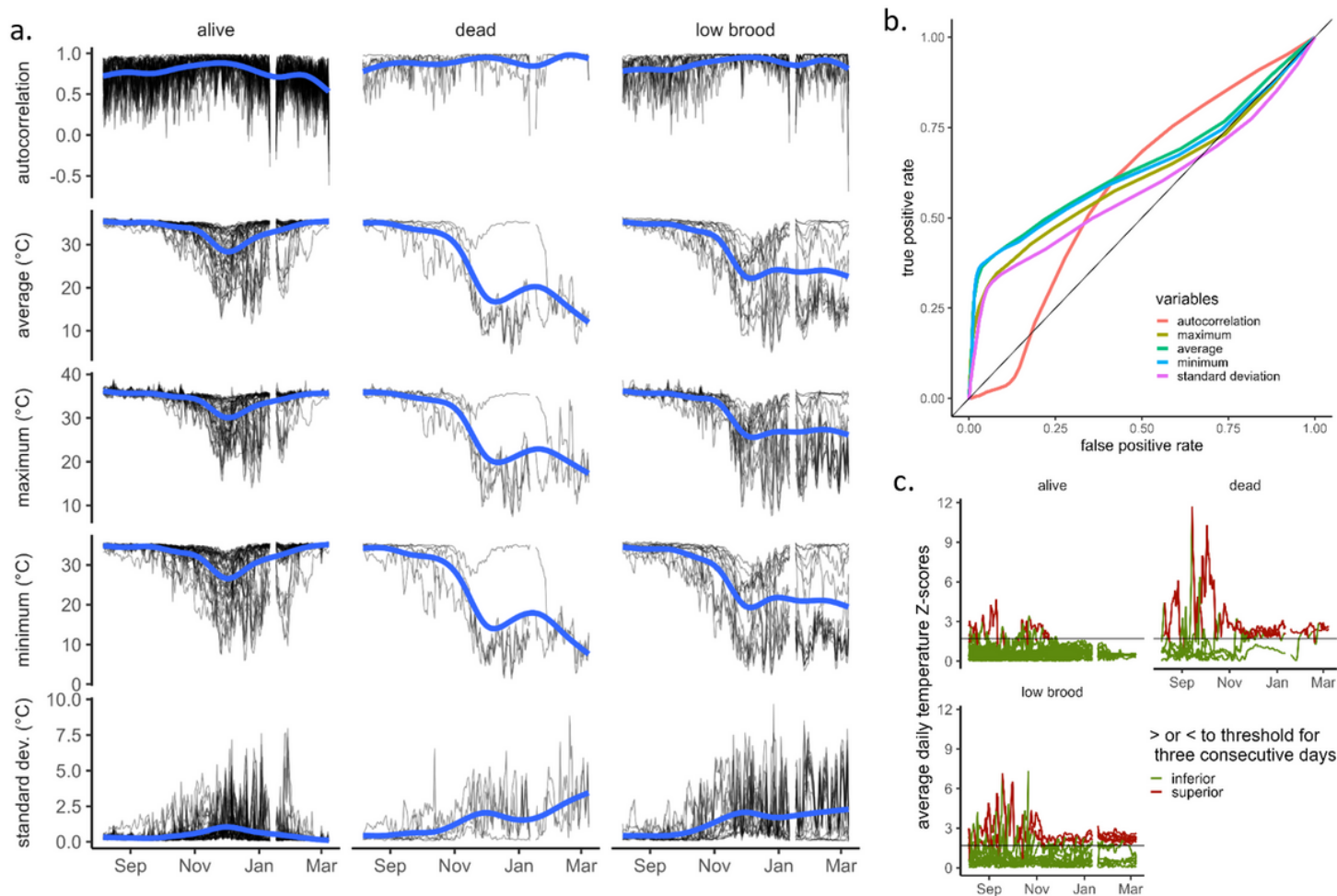


Figure 1. Toy model illustrating the effect of the weakening of the temperature regulation capacity in a bee colony. We simply assume that temperature in the hive  $T_h$  is regulated with a rate  $d$  to reach the desired temperature  $T_d = 35^\circ\text{C}$  according to the model  $dT_h = -\alpha(T_h - T_d)d_t + \sigma d_w$ . We added a stochastic term of normally distributed noise  $d_t$  with strength  $\sigma$ . In this model the equilibrium hive temperature always converges to the desired temperature  $T_d$  as long as  $\alpha > 0$ . For low values of rate  $\alpha$  close to 0 the potential around  $\alpha(T_h - T_d)^2$  around the equilibrium  $T_h = T_d$  becomes less steep, which means that disturbances (for instance due to external temperature fluctuations) would take longer to die away and fluctuations in the hive temperature will become more variable. We show this in the figure where we simulated hive temperature fluctuations for high ( $\alpha = 0.9$ ) and low ( $\alpha = 0.1$ ) regulation rate that leads to high and low variability in hive temperature fluctuations. Low regulation rates also cause higher maximal and minimal temperatures. This can also be seen in the histogram of the temperature values. The increase in variability is direct consequence of the flattening of the stability potential as can be seen in the low right panel. Simulations were based on the following values:  $T_d = 35$ ,  $\alpha = 0.1$  and  $0.9$ ,  $\sigma = 0.5$  and solved in Matlab R2017b (Mathworks) using Ito-calculus for iterating the stochastic equation.

## Figure 1

See image above for figure legend.



**Figure 2**

a. Daily hive temperature parameters for the California dataset. Bee colonies were either considered alive, dead, or had abnormally low brood levels at the end of the experiment (SM1). Daily autocorrelations, daily standard deviations, daily averages, daily minimums and daily maximums suggest differences between alive, dead and low brood bee colonies that increase during winter. Black lines show the daily data, blue lines generalised additive models (default parameters from *geom\_smooth* in R v4.0.0). b. Receiver operating characteristic (ROC) curves showing the relationship between the true positive rate and the false positive rate for the five variables measured. Although the Area under the curve is higher for the autocorrelation, the positive rate is mainly high when the false positive rate is also high. The other measurements perform similarly. c. Detection of failing hives with a threshold of 1.7 and >3 consecutive warnings (false-positive conservative) as an example of potential software application. Top panel shows the Z-Scores for the daily average temperature, with colonies shown in red when they cross the threshold (thin black line) for over three consecutive days, or green otherwise. Bottom panel shows the same classification applied to the daily temperature average curves.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [CodeforColinetal2023Earlywarningsofbeecolonyfailure20231206.rmd](#)
- [DataforColinetal2023Earlywarningsofbeecolonyfailure20231206.zip](#)
- [SupplementaryMaterials.docx](#)