Freie Universität Berlin

Master thesis at the Department of Mathematics and Computer Science of Freie Universität Berlin

Dahlem Center for Machine Learning and Robotics

Decoding temporal and social factors driving entrained circadian rhythm learning in honeybees through pattern recognition

Julia Mellert
studentID: 5180332
julia.mellert@fu-berlin.de

First examiner:	Prof. Dr. Tim Landgraf
Second examiner:	Prof. Dr. Daniel Göhring

Berlin, November 20, 2023

In this thesis patterns of circadian rhythms in honey bee colonies are investigated, focusing on age-dependent variations, the effects of external influences, and the role of social interactions. Using comprehensive tracking data over two time periods (2016 and 2019), bee circadian rhythms are modeled by a cosinor model and statistically analyzed to examine their evolution over the life course. It is shown that circadian rhythms can be already observed in young bees and that they undergo a nonlinear maturation process with age. Synchronization of rhythms and earlier peaks of activity are more pronounced in older bees. The influence of external factors, especially weather conditions, on bee activity is investigated, showing that bees with robust circadian rhythms are more sensitive. In addition, the contribution of physical contact interactions to activity transmission within the colony is examined, which may form a component of social learning of circadian rhythms.

Contents

List of Figures

1 Introduction

1.1 Motivation

Circadian rhythms are rooted in the Earth's day-night cycles and shape behavior to adapt to the environment and gain a competitive advantage. Circadian rhythms are defined as biological rhythms that: (1) last around 24 hours without external stimuli, (2) are synchronized with light and or temperature stimuli, and (3) maintain a consistent duration across different temperatures[1]. These rhythms are ubiquitous and span all levels of biological organization, from unicellular cyanobacteria [2] to complex multicellular organisms [3]. Environmental factors and social influences jointly affect genetic and endocrine functions, and these factors together control circadian rhythms.

Social insects such as honeybees provide a unique perspective on the different layers of circadian rhythms at an organism and individual level because of their complex colony structures and colony level as a whole organism. This is analogous to the organism-level circadian rhythms seen in organisms such as humans, where a central biological clock determines the overarching rhythm, but secondary rhythms exist in various physiological processes.

Age-dependent division of labor is critical for honeybee colony efficiency and is consistent with their circadian rhythms [4, 5] and among various age groups bees exhibit different rhythmic behavior. Light-dark cycles, temperature, and feeding influence rhythmicity [6, 7, 8, 9, 10, 11, 12]. Individuals synchronize their rhythm by following the rhythm of the colony [13]. Bees communicate through multiple channels, including chemical, tactile, vibrational, visual, and auditory cues, often using multiple modalities simultaneously [14, 15, 16, 17, 18]. However, the full mechanisms of social entrainment of circadian rhythms in bees are still unknown. Surrogate signals such as vibrations on the substrate and volatiles in the hive synchronize circadian rhythms, suggesting the collaboration of multiple social cues [19].

The study of circadian rhythms in whole honeybee colonies in their natural habitat has been poorly explored due to the difficulty of tracking numerous individuals in a colony. Therefore, important questions about collective rhythms arise: Who exhibits rhythmicity, to what extent, and how do these rhythms synchronize or shift within the colony? Do external factors such as weather affect the rhythms of all bees, or do they affect foragers specifically? Given the importance of physical social interactions and the emerging role of vibrations on the ground, the question arises: does physical body contact contribute to the development of rhythms in young bees? To uncover the complexity of honeybee colony circadian rhythms, studying these questions in their natural environment is of utmost importance.

1.2 Contribution

Previous work [20, 21] tracking an entire colony of bees over a long period of time at individual level allows these questions to be addressed in a natural hive setting, which is the goal of this study: Decoding temporal and social factors driving entrained circadian rhythm learning in honeybees through pattern recognition. To achieve this goal, the following approach is undertaken, utilizing comprehensive tracking data spanning two time periods in 2016 and 2019:

- **Analysis and Modeling of Bee Circadian Rhythms**: Bee activity over time is calculated based on their positions within the hive, and it is compared across different age groups. Subsequently, a cosinor model is developed for each individual bee, enabling the extraction of key rhythm parameters. The existence of rhythmicity is assessed through a zero-amplitude test, and the prominence and temporal peak of the bee's rhythmicity are evaluated, with a primary focus on its development throughout their lifetimes.

- **Influence of External Factors**: Weather conditions and sun positions during the specified time periods are examined, and their impact on the circadian rhythms of honeybees is assessed through time-lagged cross-correlation among these time series.
- **Impact of Physical Contact**: Leveraging position data, an interaction model of neighboring bees is constructed. The role of body contact in the spread and synchronization of activity is investigated by analyzing the post-interaction change in velocity. Furthermore, characteristics such as bee age, the power of the circadian rhythm, and foraging behavior are used to identify groups of bees involved in speed transfer.

In the second chapter further inside among age variations in bees behavior and factors influencing circadian rhythms in bees is given and methods used in related work to model and evaluate circadian rhythms are reviewed. This work aims to improve the understanding of circadian rhythms in honeybee colonies and to investigate the interplay between age, environment, and social dynamics that shape these rhythms.

2 Related Work

2.1 Circadian Rhythm in Bees

Age-related division of labor is a fundamental aspect of honeybee behavior that ensures the efficient functioning of the colony. As worker bees age, they progress through distinct tasks that are closely tied to their physiological development and the needs of the hive. Young bees (1-2 days) are mostly cell cleaning and then emerging to nurses (age 4-12) caring for brood, feeding, and regulating temperature. Intermediate age (12-21 days) involves house bee tasks like building honeycombs, guarding entrances, and processing nectar and pollen. Beyond day 22, bees become foragers, gathering nectar, pollen, water, and propolis for hive sustenance and pollination. [22]

Circadian rhythm develop in bees after pupation [13]. Nurses, who are responsible for brood care, do not have defined rhythmic behaviors and operate irregularly throughout the day, with direct brood contact influencing their circadian rhythms [23]. Pre-foragers exhibit ontogeny of rhythmicity by having increased rest periods during the night [11]. Foragers exhibit robust circadian patterns [24], and their departures depend on time, indicating time consciousness [11].

Bees can adapt their rhythmic patterns: Isolated nurse bees rapidly develop robust circadian rhythms that coincide with day-night cycles [25, 23], and rhythmicity occurs earlier in young bees in colonies without foragers [13]. Even in 2-day-old bees, the rhythm can change due to placement at the entrance of the hive [1, 26].

Outside temperature cycles are identified playing a role as a zeitgeber for circadian locomotor rhythms [9] and also in-hive temperature has oscillational patterns [10]. External food sources are also influencing rhythmicity as any time of the day can be entrained with feeding as a zeitgeber [11] and restricted feeding schedules regulating foraging behavior [12] is shown.

Social signals significantly influence bee's circadian rhythms, development, and contextual expressions [13]. Chemical cues synchronize circadian phases [27], even in spatially separated bees controlled by volatile pheromones and colony environment [28, 13]. Temporal coordination occurs, such as when nectar receivers coordinate with foragers or when pre-foragers prepare for their first flight [13]. Strong social cues can override light-based entrainment [26].

2.2 Analyzing Circadian Rhythms

2.2.1 Parameters of Circadian Rhythm Models

In the context of circadian rhythm analysis, several parameters are used to model and describe rhythmicity. *Period (P)* refers to the total duration of a cycle. The *phase (ϕ)*, also called acrophase, is indicated by the onset of a coherent period of activity following a period of rest within each circadian cycle. The *amplitude (A)* quantifies half of the range of deflection within a given period. The *MESOR (M)* is the central value around which the oscillation occurs and corresponds to the mean of the data, unless the data points are non-equidistant and non-stationary, in which case the rhythm-adjusted mean MESOR is considered. Waveform is another parameter that may play a role. However, this is not discussed further below for simplicity. [29]

In the context of circadian rhythm analysis, phase comparisons and tests prove complicated [30, 31, 32] and require special statistical tests because of their cyclical nature (e.g., in the context of the circadian bee rhythm, when bees peak at 11 p.m. compared to 1 a.m. - who peaks earlier?).

Prominence captures the intensity and persistence of the rhythms. Its mathematical precision may be compromised by non-stationarity, but the model estimate remains biologically relevant. A commonly used metric is the R^2-metric [33], which is often combined with a sliding window approach [34]. The Q_p statistic from the X^2-periodogram is also used, but is susceptible to noise, for which Refinetti [35] proposes a noise filtering technique as a solution.

2.2.2 Numerical Methods for Evaluating Circadian Rhythms

ANOVA. ANOVA (Annotated Variance Analysis) is a statistical method for comparing group means and is therefore used to distinguish between noisy rhythms and random oscillations when the period is approximately known [29].

ANOVA calculates an F-statistic by comparing variability between and within groups (time periods). When the group means are equal (null hypothesis), the F-value is approximately 1. When rhythmicity is present, the F-value is significantly greater than 1 because of the greater variability between groups compared to the variability within groups[29].

The period is assumed to be approximately known, and it should be noted that only the variance is compared, not a specific rhythmic pattern, so periodicity cannot be directly demonstrated[29]. The method is sensitive to the choice of the number of time periods, especially for non-equidistant data [36].

Cosinor and cosine fit. The cosinor method [37] is used for the analysis of circadian rhythms, which are smooth patterns with additional noise. In this technique, a model of cosine curves with a known period (usually 24 hours) is fitted using least squares [38, 37, 39]. Each data point x_i is a function of the rhythm-adjusted mean (M), the amplitude (A) of the oscillation, the phase of the maximum value at a fixed reference time (ϕ), and the period (P), which is often known a priori.

Trigonometric angles (θ_i) corresponding to sampling times t_i are calculated as $\theta_i = \frac{(2\pi t_i)}{P}$. A cosinor model consisting of one component is represented by the following equation:

$$y_i = M + A * cos(\theta_i + \phi) + e_i \tag{1}$$

Where e_i is an assumed independent and normally distributed error term with a mean of zero and a fixed unknown variance σ^2 [40].

The fit can be conceived as a regression problem reformulated by a linear regression if P is known:

$$\begin{aligned} y_i &= M + A * cos(\theta_i + \phi) + e_i && |cos(a * b) = cos(a) * sin(b) - \\ & && |sin(a) * cos(b) \\ &= M + A * cos(\theta_i) * cos(\phi) - A * \sin(\theta) * sin(\phi) + e_i && |x_i = cos(\theta_i), \beta = A * cos * \phi, \\ & && |z_i = sin(\theta), y = -A * sin(\phi) \\ &= M + \beta * x_i + \gamma * z_i + e_i && \end{aligned}$$

and solved by ordinary least squares. Then the amplitude can be deduced by calculating $A = \sqrt{\beta^2 + \gamma}$ and $\phi = arctan(\frac{-\gamma}{\beta}) + K\pi$ where K is an integer.[40]

The model is valid under the following assumptions: (1) well data-fitting of the model, (2) normally distributed residuals, (3) homogeneous variance, (4) independent residuals, and (5) the parameters remain constant over time, indicating stationarity [40]. These assumptions can be tested by:

1. 1) a F-test comparing the pure error sum of squares ($SSPE$) and lack of fit sums of squares ($SSLOF$) [41] with
   $F = \frac{\left[\frac{SSLOF}{m-1\breve{}2p}\right]}{\left[\frac{SSPE}{N\breve{}m}\right]} > F_{1-\alpha}(m-1-2p, N-m)$
2. (2) the Shapiro-Wilk-test of normality for small sample sizes [42] or X^2-test for goodness of fit for lager sample sizes [43]
3. (3) a F-test [40] according to $F = \frac{(N-2p-2))r}{1-r} > F_{1-\alpha}(1, N-2p-2)$ (with N = total number ob observations and p = number of components in cosinor) for which the square of the estimated values are fitted to obtain the residuals r

The statistical significance of the model is evaluated with an F-test according to Equation 2. In this test, the total sum of squares (TSS) is divided into the regression model sum of squares (MSS) and the residual sum of squares (RSS), where 2 and $N-3$ are the number of degrees of freedom associated with the model ($k = 3$ parameters -1) and the error term ($N-k$).[40]

$$F = \frac{\frac{MSS}{2}}{\frac{RSS}{N\breve{}3}} > F_{1-\alpha}(2, N-3) \tag{2}$$

If the null hypothesis can be rejected, so that the amplitude is statistically greater than zero tested ("zero amplitude test"), this indicates 24-hour rhythmicity - circadianity. The confidence intervals of the model parameters can also be derived, as described in more detail in [40].

The empirical mean and amplitude of the time series can be also determined by creating various cosine curves for all potential phase angles and then identifying the curve that minimizes the sum of squared deviations between the generated cosine functions and the actual data [44, 45]. As an indicator of strength of rhythmicity the quotient between the sums of squares (or mean squares) of the best- and worst-fitting curves can be used [46]. Disadvantages of this is the lack of mathematical formalism, approximate nature, lack of straightforward generalization to multiple components [29].

The cosinor method is particularly suitable for cross-sectional studies. It can be used for non-equidistant, sparse, short-period and serially-independent data.[29]

Kalgomorov-Smironov-test. If non-numerical data, not on interval or ratio scales, are involved, it is often better to analyze them as counts within chosen intervals instead of using cosine curves, especially if step functions are used instead, avoiding quantization issues. For nominal data, an alternative is to assess uniformity in the frequency distribution using non-parametric tests such as the Kolmogorov-Smirnov-test [47, 48].

Like ANOVA, the Kolmogorov-Smirnov-test can confirm the absence of a distinct circadian rhythm if the null hypothesis is upheld, but it can't confirm a specific rhythmic pattern if the null hypothesis is rejected.

Enright- or X^2-periodogram. For serially-dependent time series data with equidistant measurements and a period of at least 10 days [49], the Enright-periodogram method [50] can be compared with a one-way ANOVA adjusted for rhythmicity tests.

In this approach, the data is segmented into distinct periods, and an index of variability is computed for each. This index is then used for comparing within-class and between-class means using statistics like the F-statistic [51] or X^2 [52]. This adapted method is often referred to as the X^2-periodogram [29].

For a dataset with N values (x_i for $i = 1...N$), divided into K blocks of period P, the Q_P-statistic in the X^2-periodogram is calculated using the formula:

$$Q_P = \frac{KN\sum_{h=1}^{P}(M_h M)^2}{\sum_{i=1}^{N}(x_i M)^2} \tag{3}$$

Here, M_h represents the mean of K values within each period segment $[M_h = \frac{\sum x_i}{K}$, for $i = 1$ to $(N - K)$ in P increments], and M signifies the mean of all N values $[M = \frac{\sum x_i}{N}$, for $i = 1$ to $N]$.[29]

In a time series containing a single rhythmic component, the highest peak in the periodogram corresponds to the true period within the series. Evaluating the significance of this peak involves comparing it to the X^2-distribution, as the Q_P-statistic follows a X^2-distribution with $P - 1$ degrees of freedom. However, it's essential to ensure that K is greater than or equal to 10 for the result to be considered significant.[29]

Fourier analysis. Fourier analysis, also known as spectral analysis, utilizes Fourier's concept [53] that any time series can be represented using sine and cosine waves of different frequencies. If spectral analysis of a time series detects a major component in the circadian frequency range, it suggests the series has a circadian rhythm [54, 55].

In Fourier analysis, spectral energy (R_j^2) for each frequency $\frac{j}{N}$ [for $j = 1$ to $\frac{(N-1)}{2}$] is obtained by:

$$R_j^2 = a_j^2 + b_j^2 \tag{4}$$

where:

$$a_j = \frac{2}{N}\sum_{i=1}^{N} x_i * cos(\frac{\frac{2}{p_{ij}t_i}}{N}) \tag{5}$$

$$b_j = \frac{2}{N}\sum_{i=1}^{N} x_i * sin(\frac{\frac{2}{p_{ij}t_i}}{N}) \tag{6}$$

Here, N is the data points count, x_i is the variable value at time t_i, and angular measures are in radians. By calculating R_j^2 for all j from $1 to \frac{N-1}{2}$, a periodogram is created. [29]

Additionally this can be combined with a significance test by [56, 57, 58] testing whether a component of the periodogram is significantly high to indicate a rhythm. With Fisher's [57] method the periodogram's significance level, $p(R^2)$, is computed as:

$$p(R^2) = N(1 - g)^{N-1} \tag{7}$$

where N is the data points count in the time series, and g is the ratio of the largest R_j^2 to the sum of all R_j^2 values in the periodogram.

Lomb–Scargle-periodogram. The Lomb-Scargle-periodogram developed by N.R. Lomb [59] uses a statistical method similar to the X^2-periodogram in Section 2.2.2. However, it is a method that accounts for unequal data spacing by modifying it using Fourier analysis as described in Section 2.2.2. The rhythm is assumed to correspond to a sinusoidal pattern. The method has been first used for circadian rhythms by [60].

For a dataset with N values (x_i for $i = 1$ to N) recorded at various time points t_i, the PN statistic in the Lomb-Scargle-periodogram process is computed using the following formula:

$$PN = \frac{1}{2 * \sigma^2}\{\frac{[\sum_{i=1}^{N}(x_i - M)cos(\frac{2\pi}{P}(t_i - \delta))]^2}{\sum_{i=1}^{N} cos^2(\frac{2*\pi}{P}(t_i - \delta))} + \frac{[\sum_{i=1}^{N}(x_i - M)sin(\frac{2\pi}{P}(t_i - \delta))]^2}{\sum_{i=1}^{N} sin^2(\frac{2\pi}{P}(t_i - \delta))}\} \tag{8}$$

Here, M represents the mean of all N values ($M = \frac{\sum x_i}{N}$), σ^2 is the variance of all N values ($\sigma^2 = \frac{[\sum(x_i - M)^2]}{N}$), P denotes the tested period, and δ adjusts the times for unequally spaced data. The calculation of δ is as follows:

$$\delta = \frac{1}{\frac{4\pi}{P}} tan^{-1}[\frac{\sum_{i=1}^{N} sin(\frac{4\pi}{P} t_i)}{\sum_{i=1}^{N} cos(\frac{4\pi}{P} t_i)}] \tag{9}$$

The significance level of the periodogram, $p(PN)$, is calculated using:

$$p(PN) = 1 - (1 - e^{-PNmax})^N \tag{10}$$

$PNmax$ signifies the highest PN value in the periodogram.[29]

Rayleigh-test. The Rayleigh-test [61] proves valuable when there is an absence of an amplitude estimate and the objective is to determine directionality of a distribution within a circle. In this context, a clock can be analogized as a circle, and the Rayleigh-test is employed to examine the distribution of times when events take place.

By assigning trigonometric angles θ_i (where $i = 1$ to N) corresponding to observed times t_i, the Rayleigh-statistic (nR^2) can be computed using the equation:

$$nR^2 = \frac{1}{N}[(\sum_{N}^{i=1} cos(\theta_i))^2 + (\sum_{N}^{i=1} sin(\theta_i))^2]. \tag{11}$$

To establish statistical significance, the X^2-distribution is utilized since 2 times nR^2 follows a X^2-distribution with two degrees of freedom.[29]

2.2.3 Noise Handling in Rhythm Data

Handling noise is a common problem dealing with data in circadian rhythm analysis. Several strategies handling this are existing.

Often encountered in noisy data, high-frequency oscillations are effectively addressed and discarded automatically through smoothing, mean gliding window methods, and cosine fitting techniques. Non-stationarity in the data can be circumvented using moving window approaches [62, 63] for parameter estimations. Data scarcity can be tackled by introducing None values, which minimizes bias. However, when equidistant data is required, interpolation may be employed, albeit at the expense of introducing more bias into the model.[29]

These noise reduction strategies should only be employed when dealing with isolated, randomly distributed noise; otherwise, if the noise reveals an underlying pattern, their application might not be suitable.

3 Methods

3.1 Data Collection and Preprocessing

3.1.1 Experimental Hive Setup

The data used in this study were collected in two separate experimental periods: 2016 [20] and 2019 [21].

A one-frame observation hive located at coordinates 52.457130, 13.296285 was created by establishing a queenright colony with young honey bees. Newly hatched bees were obtained by incubating the brood from both the host colony (2016) and a sister colony (2016, 2019) in an external incubator at a temperature of 34 °C. Within 24 hours of hatching on weekdays, the bees were marked by being removed from the brood comb, with hairs being taken off their thorax using a wet toothpick, shellac being applied, and a curved, round tag being attached. To avoid the emergence of unmarked bees inside the colony, the marked bees were introduced via a rear entrance, and the comb was renewed approximately every 21 days. A tube connected to the observation hive provided the bees with access to the outside world.

In 2016, hive recording spanned from July 24 to September 19, capturing a colony that began with around 2000 young bees, where a total of 3166 bees were marked, and throughout the observation period, we recorded 1920 individuals from 30 cohorts. Similarly, in 2019, hive was recorded from July 4 to October 15, featuring a starting number of 1500 young bees.

In 2019 alongside the hive-monitoring cameras, an additional camera was positioned at the end of the exit tube to capture the bee's ingress and egress from the hive at a frequency of 10 Hz.

3.1.2 Data Collection

The automatic recording system "BeesBook" [64, 65] was used for data acquisition, together with infrared flashes serving for illumination, synchronized with the exposure times of the cameras. In 2016, the honeycombs were captured in high resolution at a rate of 3 Hz, while in 2019, the frequency was increased to 6 Hz. To prevent backlighting and ensure optimal contrast, the imaging alternated between the sides of the observation hive.[20, 21]

Throughout the recording period, the BeesBook system autonomously detected and monitored all tagged bees, resulting in data containing distinctive bee IDs, timestamps, plane coordinates, orientations, and confidence scores for the decoded IDs. To enable easier analysis, the processed data were transferred to a PostgreSQL database.[20, 21]

Four reference markers were affixed to the edges of the bee hive, positioned directly outside the honeycombs, to enable the calculation of a homography transforming image coordinates into a common reference frame.[20, 21]

Weather data are obtained from the German Wetterdienst [66] provided via their Python API and are taken from the weather station in Berlin-Tempelhof. The weather data include the measurements of wind speed [m/s], temperature [°C], sunshine duration [h], precipitation duration [h], precipitation height [mm] and humidity [%]. Sun position and its reference times are determined using the Python API suncalc [67].

3.1.3 Data Preprocessing

The resulting tracking data are post-processed for tracking reliability, age, hive orientation, and speed before analysis. The detections of lifetime bee tracks with uncertain decoding (below 0.1 confidence) are excluded due to uncertainty in the machine learning vision pipeline.

Calculating the death date of an individual bee based on the first day of undetected hive presence is hindered in practice due to decoding errors leading to multiple daily detections for most IDs, as well as instances of non-detection on uncertain days despite the bee being alive. Therefore a Bayesian changepoint model is adopted for robustly estimating individual death dates, defining viability from emergence until a changepoint $(d = e + l)$ using a weakly informative $N(35, 50)$ prior for l, where e is emergence day and l is days lived. Detection probability is modelled with a Bernoulli distribution, considering detection above a threshold t during viability (modeled with a Beta(5, 1) prior) and below t after, using an informative Beta(25, 1) prior for t to account for deceased bee's lower detections. Normalizing detections to $[0, 1]$, pymc3 and NUTS sampler [68] are applied for model fitting, performing 2000 tuning samples and 1000 samples per bee; the death date is determined using the mean of the last 1000 Monte Carlo samples.[20]

Bees having an age of 0 days are discarded as it is their day of entering the colony and thus dependent on their entering time, they are leaking data at this day.

Addressing misalignment of certain bee tags relative to their bodily orientation, post-processing is applied. By leveraging the correlation between fast bee movement and forward trajectory, offsets are derived to correct decoded tag orientations. Calculating trajectory directions from tracks within a second, circular mean disparities between body orientation and movement direction are computed. Short or highly variable tracks are pruned, retaining the central 50%. Using a histogram with 10-degree bins, values incongruent with the histogram are discarded. The final orientation correction is determined as the median of the remaining values.[21]

All subsequent analyses are based on data collected between August 1 and 25, 2016, and between August 20 and September 14, 2019. The tracking data, therefore, contains gaps due to falsely filtering out correct detections, but also due to occlusions, such as when bees inspect cells or depart the nest on foraging trips. Due to that the tracking data is non-equidistantly distributed and consists of the information of the time series of an individual's bee's positions and the meta-information of their age [days] per day. The bee's age distributions for the two time periods are displayed in Figure 1.

3.2 Bee Activity over Time

To obtain an indicator of the activity level of an individual bee, the speed of bee movement [mm/s], synonymously with it velocity, is calculated by dividing the Euclidean distance between two consecutive observations by the elapsed time (which is a multiple of a fraction of a second, either $\frac{1}{3}$ for 2016 or $\frac{1}{6}$ for 2019):

$$V_t = \sqrt{(x_t - x_{t+1})^2 - (y_t - y_{t+1})^2} \tag{12}$$

Velocities being higher than a threshold of 15mm/s are discarded as being unauthentic and very likely measuring errors.

To obtain an overview of the bee activity, for each bee, the velocity is averaged over each 10-minute time window and then grouped by age and the mean velocity per 60min was calculated for each age. These are subdivided by age groups (1-4, 5-11, 12-19, 2016: 20-58, 2019: 20-48) and a 1-D Gaussian filter with $\sigma = 4$ per bin is applied to their mean.

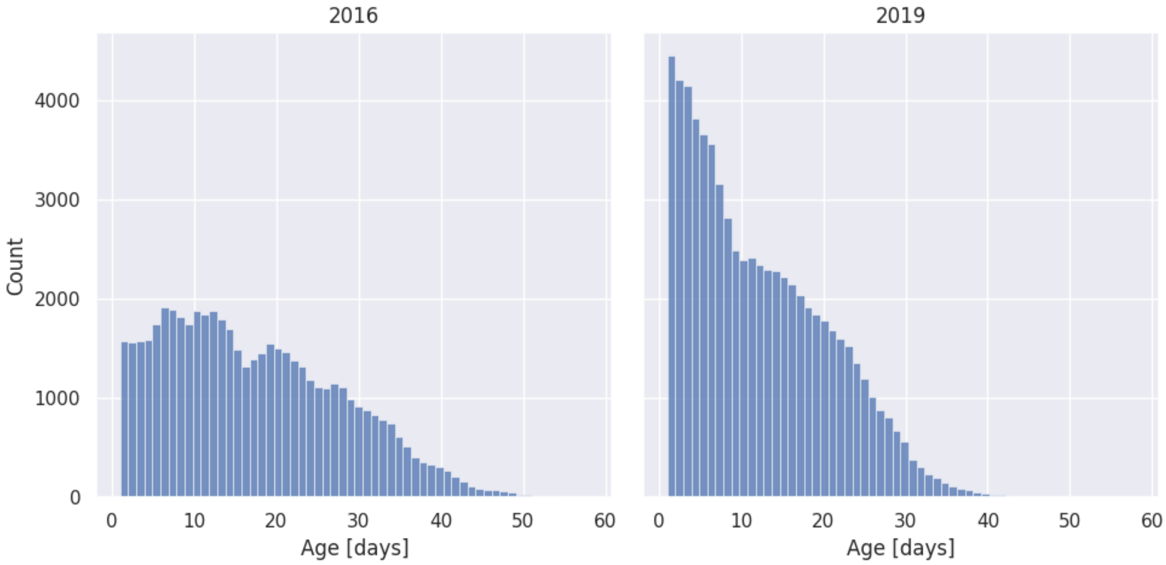


Figure 1: **Age distribution.** Age distribution for the two time periods 2016 and 2019. It is displayed on the left side the distribution of age among the colony for 2016 and on the right side the distribution for 2019.

3.3 Cosinor Model of Circadian Rhythmicity

3.3.1 Model

The expression of circadian rhythm is based on the level of activity of each bee which is measured by their velocities. As the velocities are noisy, non-equidistant and partially sparse data, due to tracking inconsistencies or foraging, the single cosinor model is used to analyze the rhythm of the bee. For that the velocities per bee are taken and a single cosinor reformulated as a linear regression problem of a fixed period of 24h is fitted with ordinary least squares according to the method described in 2.2.2.

The velocities are divided into an interval of three consecutive days and a new fit is calculated by shifting the time window by one more day to cancel out noise (due to more than one period) and avoid non-stationarity (due to subdividing the whole time period into days). Therefore, None values and velocities greater than 15mm/s considered are removed to obtain minimal bias. Velocities are low pass filtered by taking the median for a certain time period to reduce residual dependence. The model is calculated with the period of 60, 120, 300, 600, 3800, 10800 and 21600 seconds and the best fitting model is selected.

This results in this equation with $P = 24h$, and $i = 1...72$ being the time steps in hours:

$$v_i = M + A * (\frac{2 * \pi * t_i}{24} + \phi) + e_i \tag{13}$$

To test the underlaying assumptions of the model the following statistical tests were calculated:

1. F-test for model validity as described in 2.2.2 with $\alpha = 0.05$ and $p = 1$
2. Kolgomorov-Smirnoff-test to evaluate normality of residuals
3. F-test for homogenity of variance according to 2.2.2 with $\alpha = 0.05$ and $p = 1$

4. Durbin-Watson-statistic to evaluate independence of residuals
5. Augmented-Dickey-Fuller-test to detect non-stationarity

3.3.2 Existence of Rhythm

Existence of rhythm is then tested by the "zero-amplitude test" as described in 2 which is based on the assumption if the circadian model is significant, and the amplitude is greater than zero a 24h rhythm is indicated.

The statistical significance of the model is evaluated with an F-test. In this test, the total sum of squares (TSS) is divided into the regression model sum of squares (MSS) and the residual sum of squares (RSS), where 2 and $N - 3$ are the number of degrees of freedom associated with the model ($k = 3$ parameters - 1) and the error term ($N - k$).

$$F = \frac{\frac{MSS}{2}}{\frac{RSS}{N-3}} > F_{1-\alpha}(2, N-3) \tag{14}$$

The null hypothesis can be rejected with $p < 0.05$ indicating statistical significance of the model. This is in the following referred as a bee being circadian. As the test is based on the assumption of overall model validity described above in 3.3.1, this is also taken into account when analyzing whether a bee has rhythm.

3.3.3 Parameters of Cosinor Model

Another advantage of using cosinor, besides the detection of rhythm, is that the resulting parameters can be used as a model for the bee's circadian rhythm. Therefore, for each adaptation, the amplitude of the cosine wave (difference between its maximum and minimum), the phase (shift of the adaptation's activity maximum), and the prominence of the adaptation (R^2) are extracted as parameters.

Prominence is a commonly used indicator of the strength of expression of circadian rhythm, as it reflects how much of the velocity variation can be explained by the cosoidal oscillation or circadian rhythm. Hereafter, it will be referred to as circadian power and used as an indicator of how circadian a bee is. Amplitude reflects the strength of activity per day and phase reflects the temporal peak of a bee's daily activity.

To evaluate the validity of the R^2 as a good indicator for prominence, it is compared via a regression to the model's amplitude and the day-night differences of the velocities.

3.3.4 Comparison of Peak of Activity among Age Groups

To analyze the times of maximum activity the phase of the model is selected per bee and day and mapped on a 24h day. The bees are divided into three age groups: 00+, 10+ and 25+ days old, as these are approximate changepoints of behavior in life indicated by the literature as described in 2.1.

Since the rhythm is centered around midday and the peak time of the day is to be compared, not the general peak time, a one-sided unequal variance T-test (Welch-test) is used to compare the different peak times of the day. The bees have inhomogeneous phase variance between age groups and the alternative hypothesis the bees have an earlier activity peak than the comparison age group is reflected by the one-sidedness.

3.4 Influence Weather Parameters and Sun Position

For assessing weather influence a time-lagged sample cross-correlation analysis is performed as proposed by [69] to capture environmental influences. This method gauges the magnitude and direction of the link between bee activity, depicted by their velocity, and weather parameters. This process involves matching samples from different time series separated by a fixed time delay θ and then evaluating their correlation. This has the advantage that even time-lagged responses of bees to weather can be deciphered. The correlation falls within a range from -1 (indicating a perfect negative correlation) to 1 (indicating a perfect positive correlation), with 0 signifying no correlation.

The time-lagged cross-correlation analysis is conducted using the entire time periods of 2016 and 2019. A moving consecutive time window of three days is applied to the mean velocity of each bee measured at 10-minute intervals, along with the corresponding weather parameters also measured at 10-minute intervals. For each cross-correlation, both the maximum and minimum correlation values, as well as their associated time lags, are extracted. Subsequently, the time lags are converted to hours for further analysis.

The foundational presumption underlying this analysis is the assumption of stationarity, where both the mean and variance remain relatively constant. To validate the stationarity, the Augmented-Dickey-Fuller-test is employed. If the resulting p-value exceeds 0.05, the data is deemed non-stationary, necessitating the application of a detrending process.

To investigate the sun's role as a significant zeitgeber for circadian rhythms, a time-lagged cross-correlation analysis is conducted, like the approach taken for weather parameters. Specifically, the time series data of the sun's altitude is conducted. The altitude as a parameter for the sun's position is chosen, because the cross-correlation is consistently calculated within a three-day time window, and azimuth displays a sharp discontinuity at midnight when it shifts from maximum to minimum values. This abrupt change doesn't align with the gradual growth in bee activity throughout the day compared to the altitude showing a smoother transition.

To assess the impact of a bee's circadian power on its correlation with the sun's position, a regression analysis is performed. In this analysis, the degree of circadianity (measured by R^2 from the cosinor fit) is regressed against the correlation coefficient with the sun's altitude.

3.5 Social Factors and Physical Speed Transfer

To investigate the social factors contributing to circadian rhythm development, a model of physical interactions for the period in 2019 is developed. This model involves extracting the positions of individual bees within the hive and analyzing their post-interaction velocity changes. Additionally, various factors that could potentially influence the circadian rhythms of bees and serve as indicators of different behaviors are also extracted and patterns emerging from the combination of bees interacting are evaluated.

An interaction between two bees (one focal and one non-focal) is established under specific criteria: both bees must be detected simultaneously with a confidence threshold of 0.25, their marking's distance at the center of their bodies must not exceed 14mm, and their bodies must be classified as overlapping. To determine overlap, binary rectangular masks are employed, aligned with the focal bee's marker at (0,0) and an angle of 0°. These masks, sized at 13 x 5mm to approximate a bee's body, including a 1mm margin, are evaluated using a logical "AND" operation.

Multiple detections of the same interaction within a 1-second time interval are merged into a single interaction. In such cases, the initial body positions are used to determine the overlap. The interaction is counted twice because both bees involved are considered focal once.

The absolute post-interaction velocity change is calculated as the mean velocity from the end time of the interaction to 30 seconds after the interaction ("end velocity of the interaction"), subtracted from the mean velocity from 30 seconds before the interaction to the start of the interaction ("start velocity of the interaction").

As a reference an interaction null model is created by taking the distribution of the start and end times of the interactions and selecting two random bees at those times that are detected in the hive at that time. These pairs of bees are considered as "interacting" and their velocity change is calculated.

To gain further insight into social dynamics, the effect of an interaction on speed is compared between subgroups. The following parameters play a role in the selection of subgroups: the presence and power of rhythm, the age, the starting velocity at the beginning of the interaction and whether bees are forager, determined by whether the bee is registered in the exit camera of the hive.

For the subgroups of differentially expressed circadian bees and bees with different takeoff speeds, the speed changes to their counterparts of the null model are compared using a Welch-test to determine if there is a significant difference when bees have physical contact or not, in addition to visual impressions.

4 Analysis and Modeling of Bee Circadian Rhythms

4.1 Existence of Circadian Rhymicity of Bees

This chapter presents the results of the analysis of rhythmic patterns in honeybee activity, specifically examining variations across different age groups based on their observed hive velocities. The analysis comprises an initial exploration of velocity patterns, followed by an evaluation of a cosinor model to characterize these patterns. Subsequently, the presence of rhythmicity is assessed within distinct age groups, with particular attention given to bees aged 0-5 days.

4.1.1 Bee Activity over Time

In the study of honeybee activity over time, several characteristics in analyzing the bees mean movement speed as shown in Figure 2 and 3, reflecting the bees locomotion, are emerged and are described in the following.

The mean movement speed of honeybees falls within a range of approximately 1-5 mm/s, reflecting their typical locomotion. The highest average activity peak across all bees recorded at 5.34 mm/s occurred on September 3, 2019, at 12:10 am. Notably, individual bees, aged 33 days, displayed a peak of 13.46 mm/s on September 1, 2019, at 07:50 am. Conversely, the lowest recorded activity peak for all bees averaged at 0.98 mm/s, recorded on September 14, 2019, at 00:00 am. On an individual age level, the lowest peak of 0.07 mm/s is observed in a 38-day-old bee on September 9, 2019, at 11:10 pm.

These activity peaks tend to occur around noon, suggesting a diurnal preference in honeybee activity patterns. Overall, honeybee activity levels follow a cosinoidal pattern, with variations in activity levels between daytime and nighttime, indicating a consistent internal timekeeping mechanism. There are varying activity levels for each day but there is no discernible seasonal trend in honeybee activity levels during the specified time period. The rhythmic pattern is accompanied by bursty tendencies, characterized by non-daily-related oscillations of high and low activity.

Moreover, when observing bees by age, distinctions become apparent as shown in Figure 4. Younger bees exhibit less variation in their activity levels compared to their older counterparts. As bees age, their activity patterns become narrower and more prominently cosinoidal. Notably, at a certain age it varies which age group shows the maximum activity level and the highest daytime variation.

These observations suggest that honeybee activity follows a circadian pattern, with fluctuations observed between different age groups. The older the bees become, the higher the cosinoidal amplitude of their activity compared to younger bees. Regardless of age, all bees exhibit differences in activity levels between day and night, suggesting a circadian rhythm in their behavior.

4.1.2 Evaluation of Cosinor Model

In Section 3.3.1, the cosinor model is introduced, which relies on a linear regression framework. The model's validity hinges on certain assumptions, which are examined outlined in there. To ensure the model's reliability, different cosinor models are fitted to the dataset and a thorough evaluation of their goodness of fit is conducted. The distribution of statistical properties for each model is visually presented in Figure 5 for 2016 and in Figure 6 for 2019, while

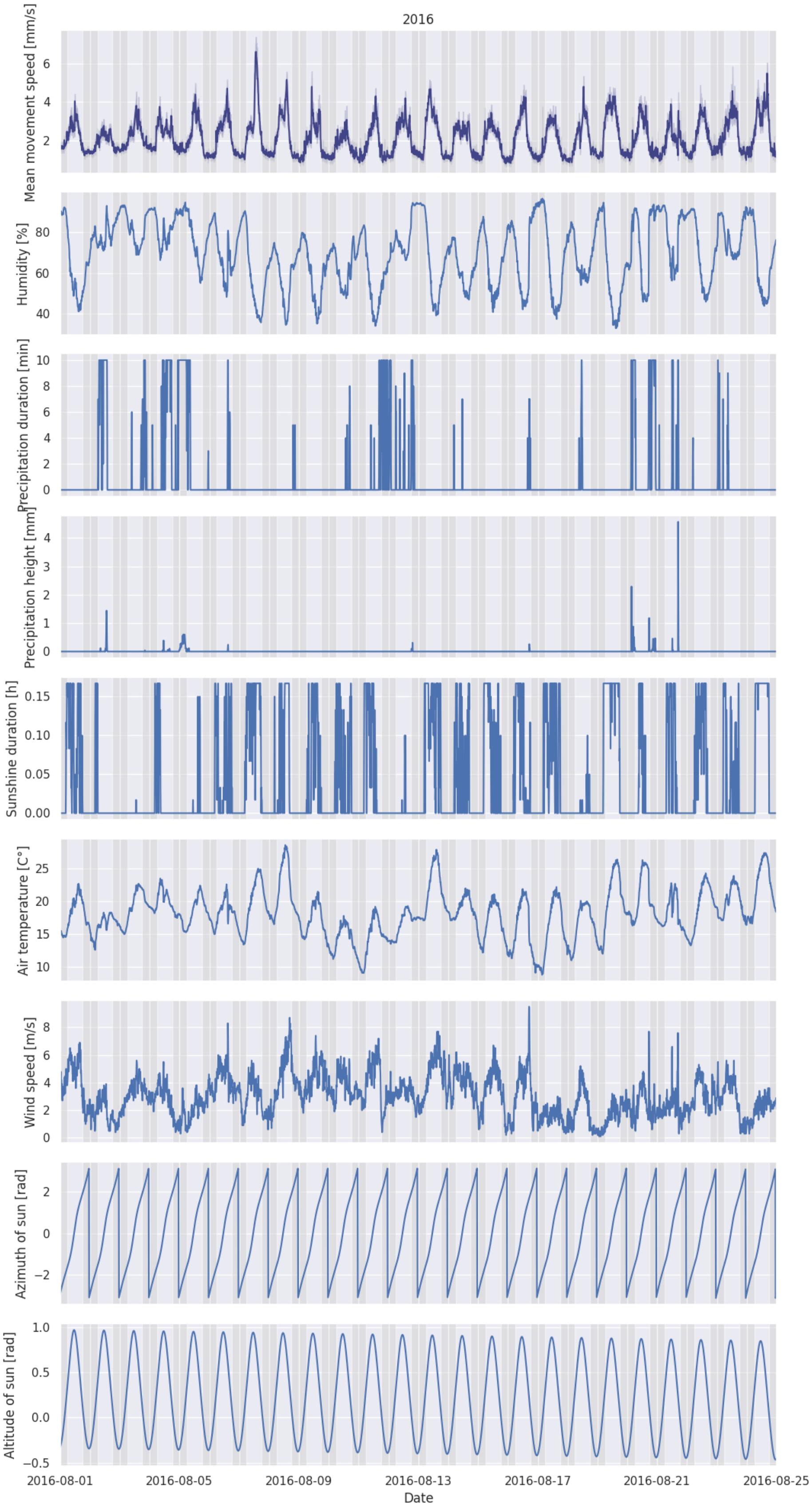


Figure 2: **Velocity and weather over period 2016.** It displays the mean velocity, humidity [%], precipitation duration [min], precipitation height [mm], sunshine duration [h], air temperature [C°], wind speed [m/s], sun azimuth [rad], sun altitude [rad] over the time period 2016.

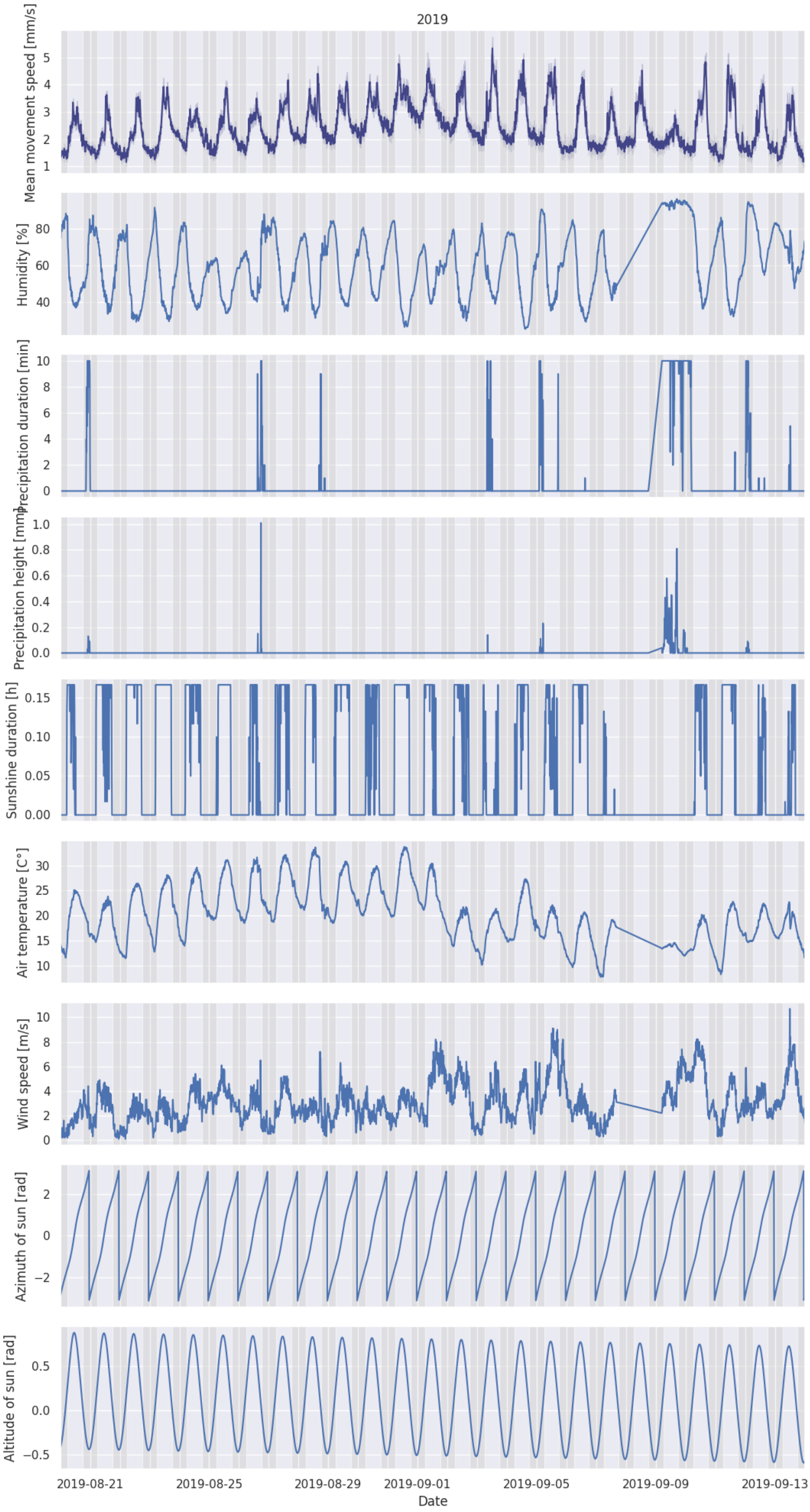


Figure 3: **Velocity and weather over period 2019.** It displays the mean velocity, humidity [%], precipitation duration [min], precipitation height [mm], sunshine duration [h], air temperature [C°], wind speed [m/s], sun azimuth [rad], sun altitude [rad] over the time period 2019.

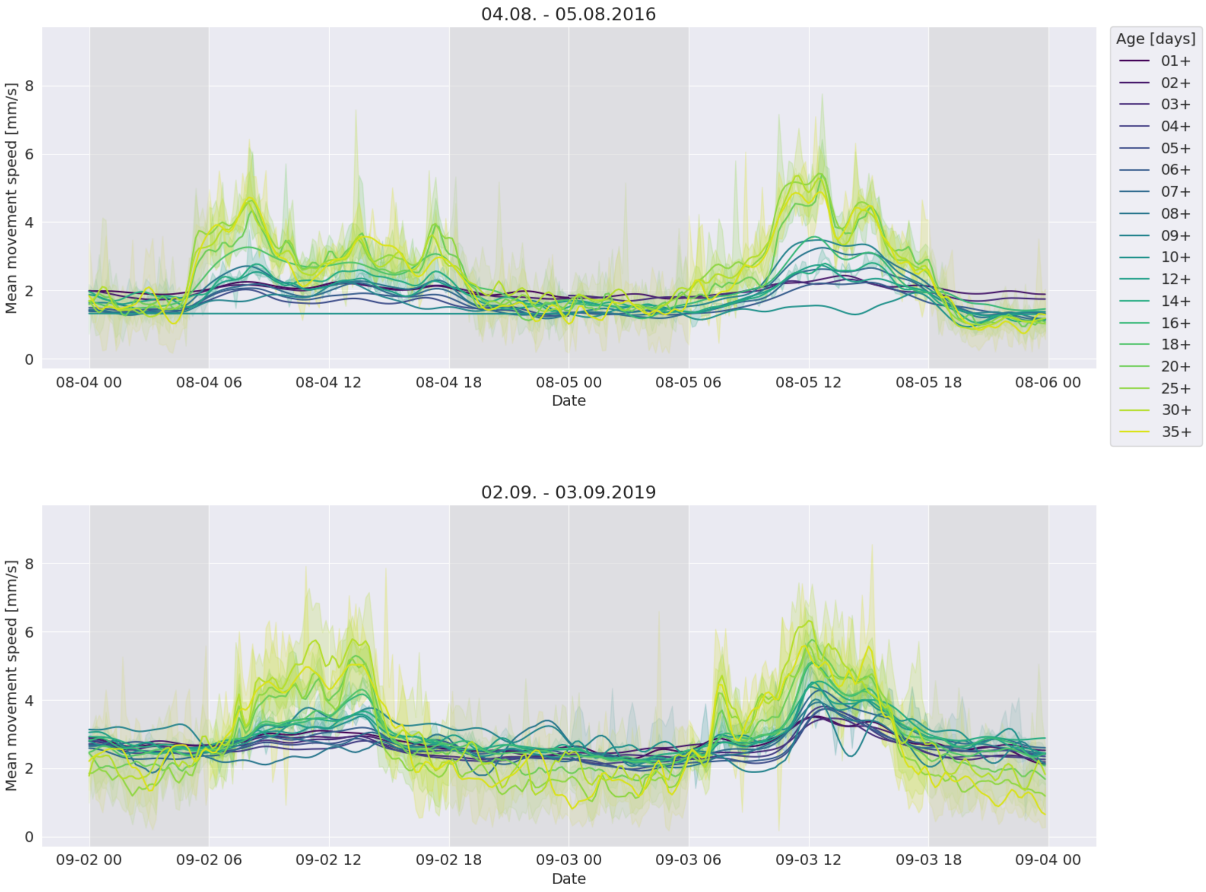


Figure 4: **Velocities among different age groups of bees.** The mean movement speed [mm/s] is displayed over time subdivided by different age groups. The darker areas reflect the nighttime.

the estimated model parameters are depicted in Figure 7.

The cosinor model involves estimating three primary parameters: MESOR, amplitude, and phase. These parameters exhibit slight variations across different models, but it's noteworthy that time periods ranging from 60 to 3600 seconds yield fairly consistent parameter estimates. This suggests that the choice of a specific model may not be critical. What's essential is that the model remain valuable to provide the best-fitting parameters, even when the model's assumptions aren't fully met[40].

The p-value of the F-test for model validity tends to decrease with longer time periods, but all time periods still meet the 95% quantile criterion for significant model validity. Similarly, while the p-value for homogeneity of variance decreases with longer time periods, a substantial proportion (>95%) remains statistically significant.

Conversely, the p-value from the Augmented-Dickey-Fuller-test increases with longer time periods, with the two longest time period models in 2016 failing to exhibit >95% significance for stationarity. The Durbin-Watson-statistic, indicating autocorrelation of residuals, shows a wider distribution with a higher mean for longer time periods. For the 3600- and 10800-second time periods, the 95% quantile falls between 1 and 3 in 2016, indicating acceptable autocorrelation. In 2019, the left boundary is 0.96, still within acceptable limits. Moreover, the p-values indicating a significant share tend to rise with longer time periods, potentially due to positive autocorrelation of residuals leading to an overestimation of p-values and hence more significant fits [70]. As the positive autocorrelation (DW-statistic) diminishes, so do the p-values.

The Kolmogorov-Smirnov-test, which assesses the residual distribution against a normal distribution, yields higher p-values with longer time periods, indicating that more samples exhibit significantly normally distributed residuals. However, in 2016, a significant share 95% is observed only for the longest period, while in 2019, it holds for the two longest models. For the 3600-second model in 2019, there is 90% compliance. This phenomenon may be attributed to the inability of velocities to fall below zero, causing the fit to lack a negative curve. Consequently, the distribution of residuals skews right-tailed, with more extreme positive values.

Among the various models analyzed, the model with a time period of 3600 seconds emerges as the best in terms of statistical properties, boasting a 41% in 2016 and 69% in 2019 compliance rate with model assumptions with the share among age groups visible in Figure 8. However, it's essential to acknowledge that over 95% of the intervals do not satisfy the criterion of normally distributed residuals, and this fact should be considered, especially when employing the p-value of the model. Furthermore, the parameter estimates obtained from this model closely align with those from higher-resolution data, reinforcing their validity and consistency.

4.1.3 Test of Rhythmicity

In the study, a rhythm test is conducted to investigate the presence of diurnal patterns in the activity of honeybees. The results of this test are shown in Figure 8. It illustrates the classification of bees into circadian and non-circadian groups, which are further divided by age groups. Additionally, the results are subdivided based on the reliability of the p-value, taking into consideration whether the assumptions required by the cosinor model are met.

The lowest proportion of rhythmic behavior is observed in the youngest bees of the +03 day age group, with 60% in 2019 and the 01+ age group with 28% in 2016. Conversely, the highest proportion of rhythmic behavior is seen in the oldest age group, +35 days, with 92% in 2019

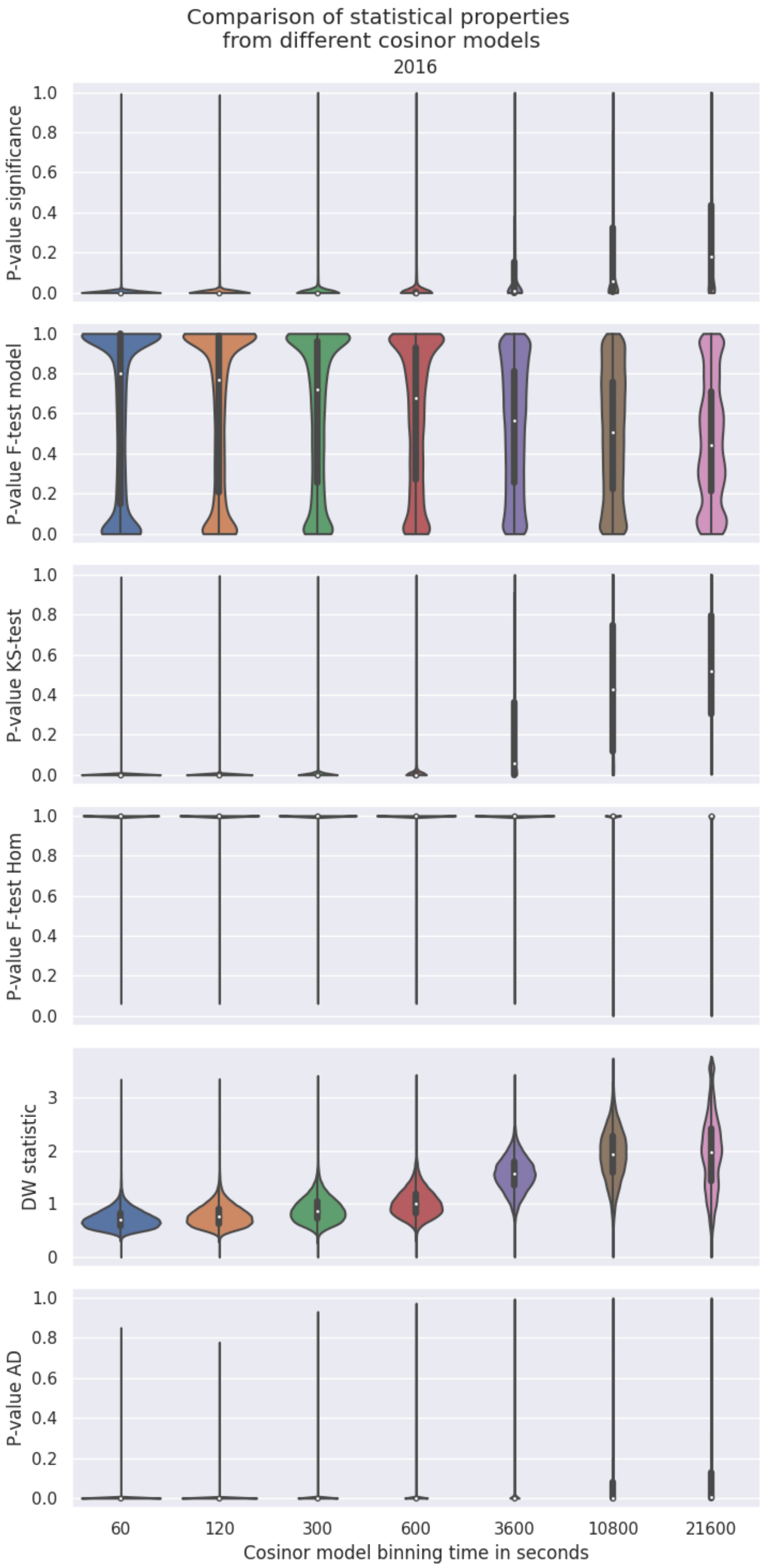


Figure 5: **Distribution of regression assumption tests 2016.** The distributions of the statistical indicators are displayed testing the validity of the assumptions made for the linear regression for different cosinor models. The rows display (1) p-value of the model, (2) p-value of a F-test whether the model fits to the data, (3) Kalgomorv-Smirnoff-test for normality of residuals, (4) F-test for variance homogeneity, (5) Durbin-Watson-statistic and (6) Augmented-Dickey-Fuller-test.

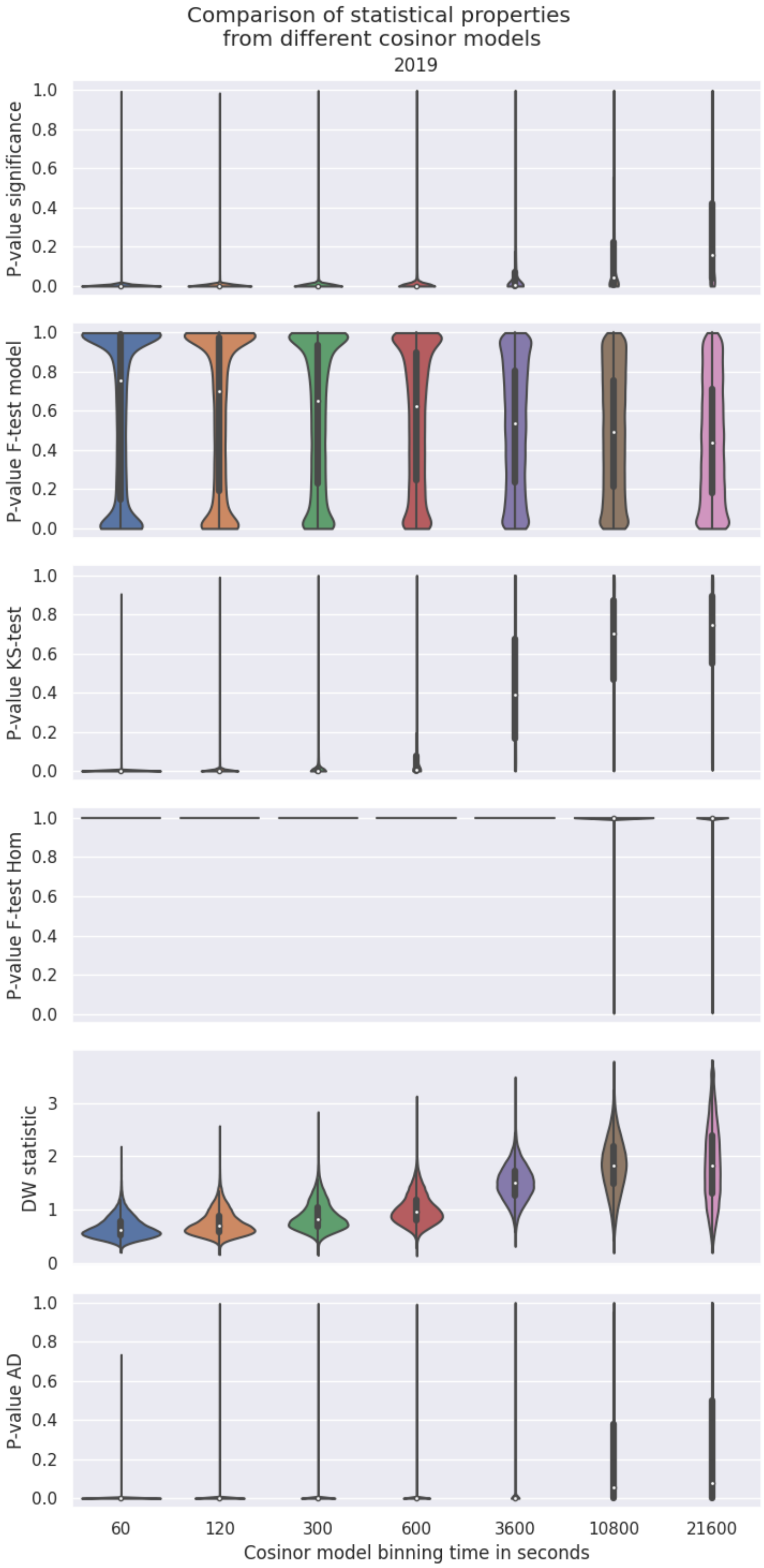


Figure 6: **Distribution of regression assumption tests 2019.** The distributions of the statistical indicators testing the validity of the assumptions made for the linear regression for different cosinor models are displayed. The rows display (1) p-value of the model, (2) p-value of a F-test whether the model fits to the data, (3) Kalgomorov-Smirnoff-test for normality of residuals, (4) F-test for variance homogeneity, (5) Durbin-Watson-statistic and (6) Augmented-Dickey-Fuller-test.

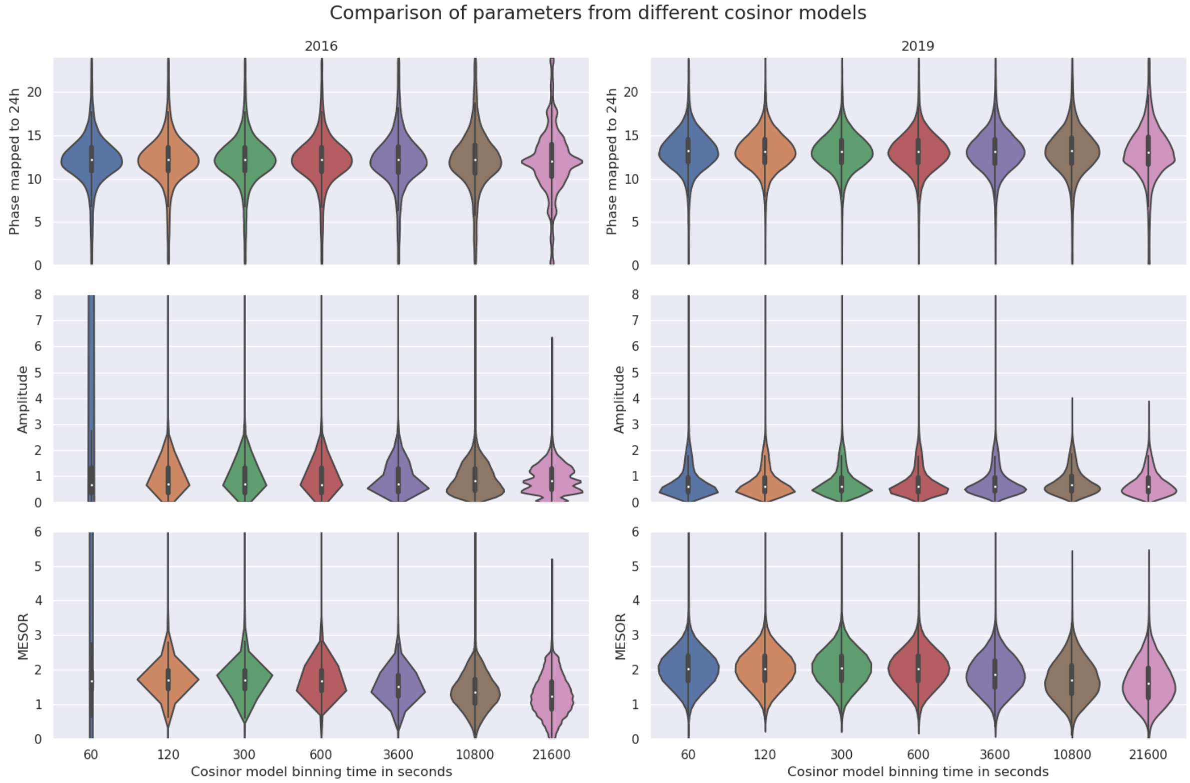


Figure 7: **Comparison estimated parameters cosinor models.** The figure displays the distribution of estimated parameters from the different cosinor models for 2016 (left) and for 2019 (right). In the first row the phase is compared, the second row the amplitude and the third the MESOR. The different cosinor models are named by the time period the median is taken from in seconds as described in Chapter 3.3.1. Some distributions are cut due to high outliers.

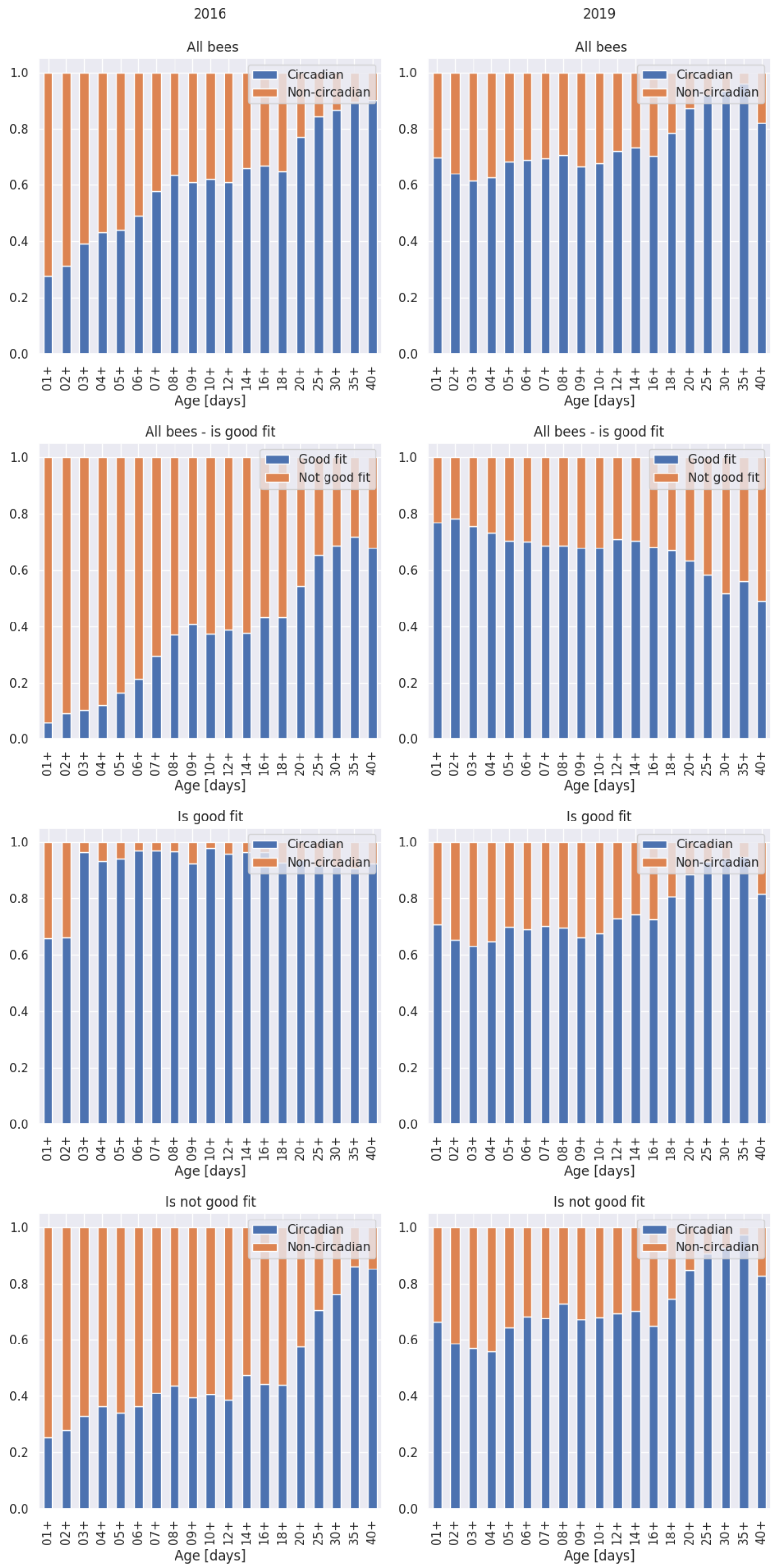


Figure 8: **Proportion [%] of rhythmic tested bees among age groups.** The proportion of rhythmic tested bees (blue) for 2016 (left) and for 2019 (right) are displayed. The first row displays the proportion among all bees, the second row the proportion of bees fulfilling the model assumptions, the third row shows the proportion of circadian tested bees among the bees fulfilling the assumptions and the fourth rows the proportion among the ones not fulfilling.

and 89% in 2016.

The majority of bees exhibit rhythmic activity patterns in each age group, consistently more than 50% for 2019 and 30% for 2016. The proportion of rhythmic bees increases trend wise with age, with the group of oldest bees again being less proportionally rhythmic. This observation might be explained by the fact that this age group includes the day of death. On the day of death, a bee's activity might be weaker, or in the case of a morning death, no clear diurnal rhythm can be recorded. This could lead to the absence of a discernible rhythm. Additionally, test values on the day before death can be influenced by the bee's activity on the day of death, as the rhythmicity test is applied within a 3-day window.

When considering reliable p-values, a higher proportion of rhythmic bees is found, ranging from 63% to 92% in 2019 and from 66% to 95% in 2016. Conversely, when the model assumptions are not met, the proportion of circadian bees is lower, ranging from 53% to 95% in 2019 and from 27% to 83% in 2016. This discrepancy suggests that failure to meet the assumptions of the cosinor model does not result in an overestimation of the proportion of circadian bees.

In summary, rhythmicity increases with age but decreases in the oldest group, possibly due to the inclusion of the day of death. Reliable p-values are associated with higher rhythmicity, while failure to meet model assumptions indicates lower circadian behavior, which could indicate a lack of cosine pattern in bee activity. However, it can be concluded that a proportion of bees in each age group already exhibits rhythmic behavior.

4.1.4 Rhythm in Young Bees

Since the test for circadianity shows the unexpected result that a significant proportion of young bees already have rhythmic patterns at 1-5 days of age, these are examined more closely for differences between those with circadian and non-circadian rhythms which is shown in Figure 9 for 2016 and for 2019 in Figure 10.

The circadian bees show little variation in mean velocity. They whiten slightly larger amplitudes, with a higher velocity during the day of difference of 0.3-0.5 mm/s, while the velocity at night is lower compared to the non-circadian bees. In terms of median velocity, circadian bees show greater day-night variation on certain days. Compared to non-circadian bees, circadian bees exhibit nights with no or shorter inactive periods (e.g., on August 2, 3, 7, and 8, 2016 (opposite on August 12) and on September 2, 2019 (opposite on September 16 and September 17)) and more intermediate or surge activity periods (observed on August 4, 6, 9, 10, 11, 14, 15, 17, and 23, 2016, and August 23, 26, 27, 28, 30, and September 6, 7, and 8, 2019).

Both circadian and non-circadian bees show variation in variance during the day and also at night, with higher variance during daytime hours. Circadian bees show slightly lower variability. With few exceptions in 2019, the minimum velocity is consistently 0. The maximum velocity is consistently higher during the day and lower at night for both circadian and non-circadian bees.

As for age-related differences, young circadian bees in 2016 were found to follow the pattern, the older the bees are, the higher the amplitudes, except for the first 5 days in the year 2016. In contrast, non-circadian bees show a less pronounced age-related pattern. In 2019, younger non-circadian bees show increased nocturnal activity, while circadian bees show less age-related differences.

Closer inspection supports that young bees do indeed exhibit a circadian pattern. The differences between young circadian and non-circadian bees are most apparent at night: Non-circadian bees exhibit shorter resting periods and are more active at night. This nocturnal

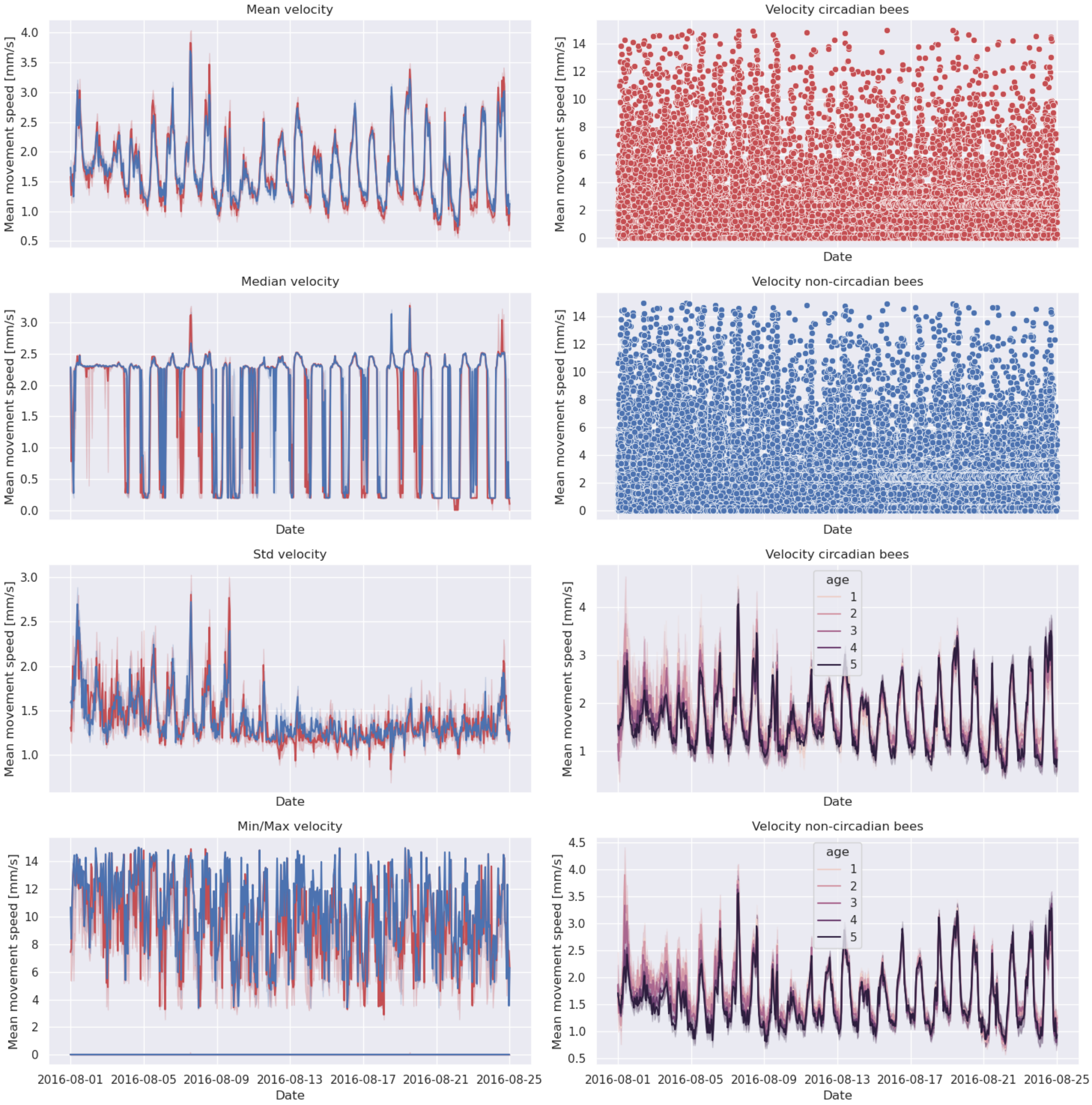


Figure 9: **Velocity of young bees (0-5 days) over period 2016.** The median velocity of each hour for each single bee is displayed over the time period 2016 with blue being the non- and red the circadian bees. The left side of the plot shows the mean of those velocities, the standard deviation, and the median the maximum and minimum. The left side displays all velocities and among the circadian tested bees the mean subdivided by age and among the non-circadian bees.

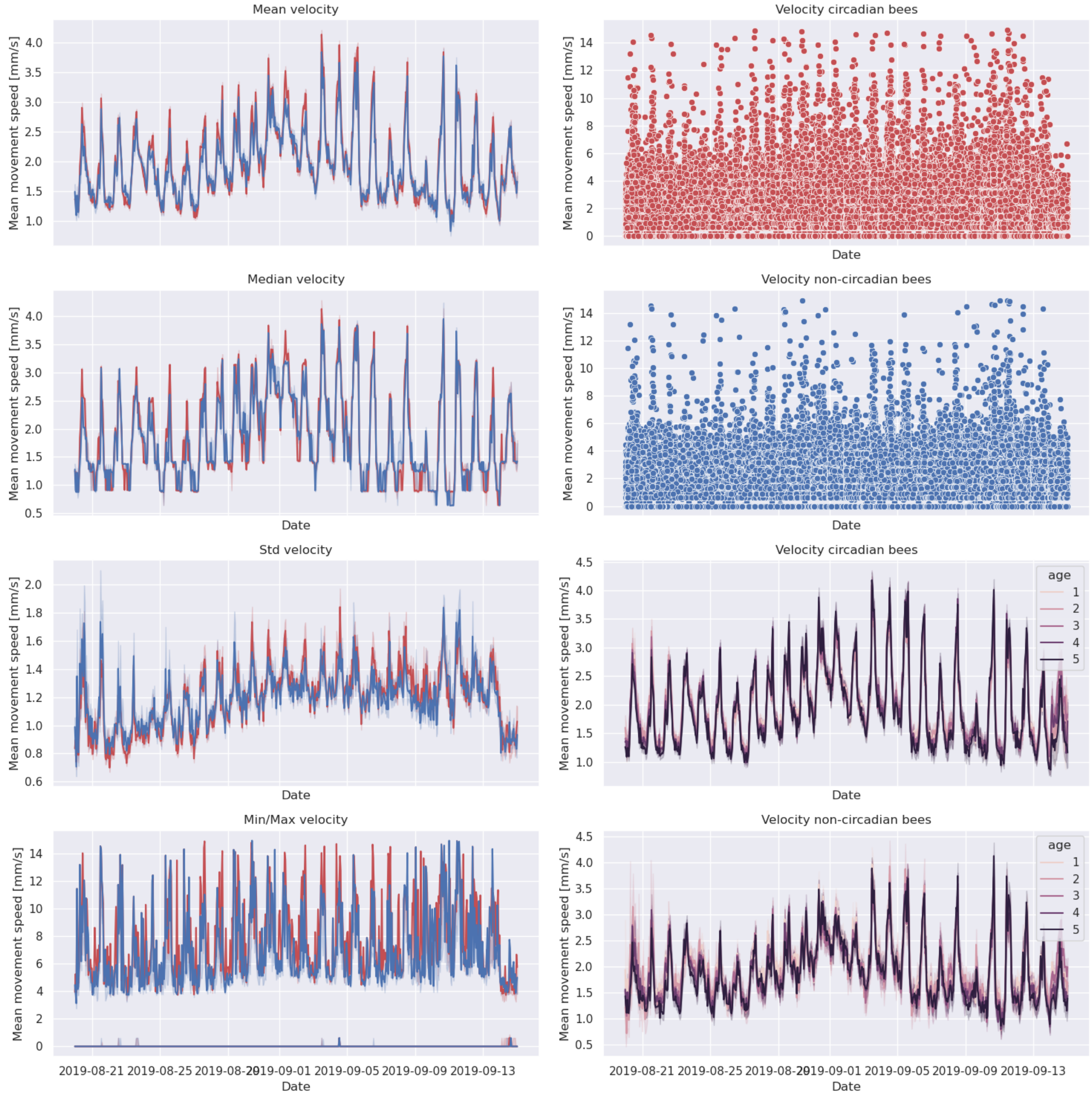


Figure 10: **Velocity of young bees (0-5 days) over period 2019.** The median velocity of each hour for each single bee is displayed over the time period 2019 with blue being the non- and red the circadian bees. The left side of the plot shows the mean of those velocities, the standard deviation, and the median the maximum and minimum. The left side displays all velocities and among the circadian tested bees the mean subdivided by age and among the non-circadian bees.

activity is particularly pronounced in very young non-circadian bees (1-3 days old).

4.1.5 Summary

In summary, the movement rate of honeybees has a cosinoid character, with marked variations between daytime and nighttime activity, with no seasonal trend in this time periods. The activity patterns of honeybees vary with age and become more cosinoid as the bees age. The presence of a rhythm is statistically confirmed by the zero-amplitude-test. Remarkable is the presence of rhythmic activity in a proportion of young bees, even at the age of only 1-5 days. These young bees with circadian rhythms showed mainly differences in nocturnal activity compared to non-circadian bees.

4.2 Circadianity over the Lifetime of a Honeybee

In this chapter, the maturation of rhythmic patterns is analyzed and differences among age groups are unveiled. The analysis commences by examining the prominence and how it evolves with the aging process of bees. Subsequently, the phase of their activity is explored, as determined by the cosinor model, and it is related to age and to the power of the bee's rhythmic behavior.

4.2.1 Power and Development of Rhythm

The following analysis delves into the development of circadian rhythms in bees throughout their lifetimes. The power, represented by R^2, is compared to the amplitude of the bee models and the day-night differences in bee velocities. Additionally, the prominence of the rhythm is examined among various age groups.

For assessing the validity of rhythm prominence, the R^2 of the cosinor fit is regressed against both the amplitude of the fit and the day-night differences in bee velocities. The regression with amplitude yielded an R^2 of 0.825 in 2016 and 0.855 in 2019, indicating a strong correlation. Conversely, the day-night velocity differences resulted in an R^2 of 0.716 for 2016 and 0.771 for 2019, showing slightly lower but still high correlation. These regressions confirm the reliability of R^2 as an indicator for rhythm prominence, both in terms of its alignment with the model (amplitude regression) and its representation of raw data (day-night differences), thereby making it the sole indicator for subsequent analysis.

The circadian power, as determined by the cosinor fit of the bees, is depicted in Figure 11. The median prominence ranges from 0.03 to 0.07 in 2016 and 0.04 to 0.09 in 2019 for the non-circadian bees. In contrast, for the circadian bees, the median prominence ranges from 0.14 to 0.39 in 2016 and 0.17 to 0.45 in 2019 across different age groups. Among younger bees, weaker rhythmicity is prevalent, but some individuals exhibit highly pronounced rhythmic patterns. For circadian bees, as they age, their rhythm becomes more prominent, as indicated by the median values, with the highest median observed in the 35+ age group in 2019 and the 40+ age group in 2016. Interestingly, in 2019, the oldest group has a lower median than the second oldest group, possibly due to the same effect observed in the proportion of rhythmic bees described in Section 4.1.3.

In 2019 a trend is observed where bees reach a plateau in rhythmicity from the age range of 06+ to 20+ days, followed by a resurgence of rhythmic behavior beyond 20+ days. This pattern is somewhat visible in 2016, but it becomes more pronounced from the age range of 12+ to 25+ days, albeit not as clear as in 2019.

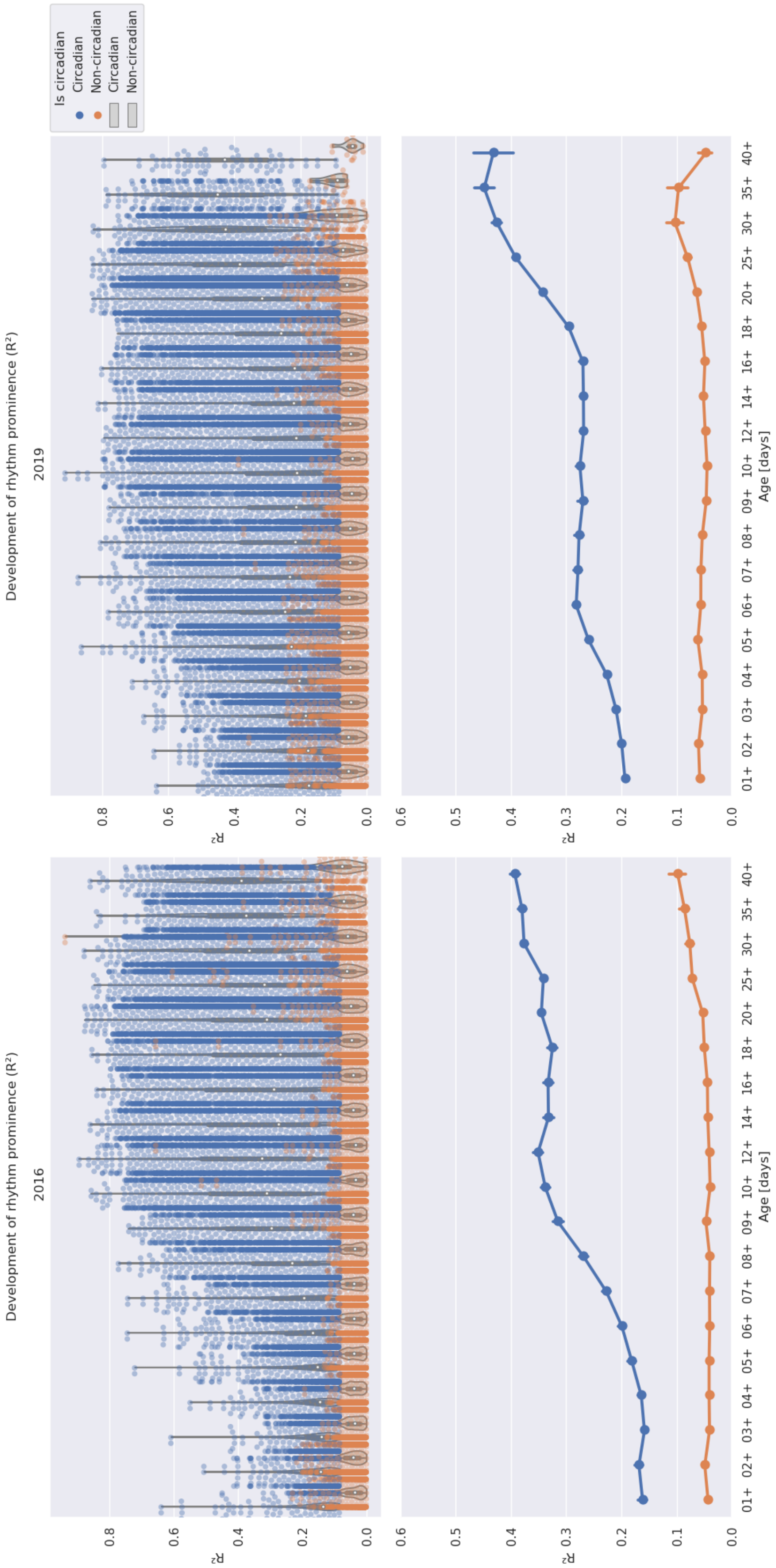


Figure 11: **Development of circadian power (R^2).** In the first row the distribution of circadian power among age groups is displayed and it is subdivided whether the bees are tested circadian or not for 2016 and for 2019. The second row displays the median among those groups.

The findings suggest that R^2 serves as a reliable estimator for indicating the power of the circadian rhythm. Both tested circadian bee groups exhibit a development towards a more prominent rhythm over their lifetime in both years. Uncircadian bees may also show a slight increase in prominence from the age of 25+ onward, but their prominence remains consistently low before that point. Interestingly, there is no linear trend in the development of prominence, with two distinct patterns emerging for the two time periods. In 2016, there is a strong growth in prominence at the beginning of a bee's lifetime, which flattens out at the age of 12+. In contrast, in 2019, there is growth in the early stages, followed by a plateau during the intermediate ages, and then a stronger growth rate re-emerges from the age of 20+ onward.

4.2.2 Temporal shift of Activity Peak

In this section, it is analyzed how the timing of peak activity, the phase of cosinor fit, shifts with age and circadian performance level. Here, peak activity is measured by the phase of cosinor fit per bee per day, see section 3.3.3, and mapped to a 24-hour cycle.

The distribution of phase shows clear age-dependent characteristics, as shown in Figure 12. Among the different age groups, the oldest bees show the earliest peak of activity, with 11.78/12.48 hours as the median phase. This is followed by the middle-aged bees in 2016, which peak around 12.28 hours and in 2019 13.27 hours being the last, while the youngest bees peak last in 2016, with a median phase of 12.35 hours in 2016 and second in 2019 with 13.18 hours. Furthermore, the variance of these distributions also varies between age groups. The oldest bees have the narrowest distribution with a low standard deviation 2.04/1.81 hours, while the middle-aged bees have a standard deviation of 2.97/2.32 hours and the youngest bees have the widest distribution with a standard deviation of 4.03/2.67 for 2016/2019. It is found that the variance of the distributions differed statistically significantly, and the oldest group is significantly earlier active than the younger groups. No significant difference is demonstrated between the peaks of the middle and youngest groups.

The distributions of circadian rhythm, divided into groups of circadian and non-circadian bees, shows a more differentiated distribution. Within the group of circadian bees, there are clear shifts in peak times, with a ranking from early to late: old, middle old, and young. These shifts are statistically significant in 2016, but for 2019 again just the oldest compared to the others.

Differences to this are observed in the non-circadian group. In 2019 there is no clear peak shift visible and in 2016, there is a reverse shift in peak times, with a ranking of young, middle age, and old. In addition, the variance varies within these distributions, with the oldest bees having the lowest variance, followed by the middle-aged bees, while the youngest bees have the highest variance for 2016 and 2019.

The pattern of the distribution of non-circadian bees persists when the phase distribution is subdivided by circadian performance rather than age by dividing the bees into three equally sized groups. Figure 12 illustrates this distribution. Bees with the highest circadian power peak at the earliest, with a median phase of 11.87/12.71 hours, followed by those with the lowest circadian power peaking with a median phase of 12.43/13.39 hours, and those with medium circadian power a shift of the peak toward the time of the oldest bees with 12.48/13.46 hours for 2016/2019. The variance of phase distribution in bees with low circadian power is higher than that of the group with the youngest bees.

Analysis per age rather than age groups, as displayed in Figure 13, reveals a non-linear trend in which younger bees initially peak earlier, then gradually transition to later peak

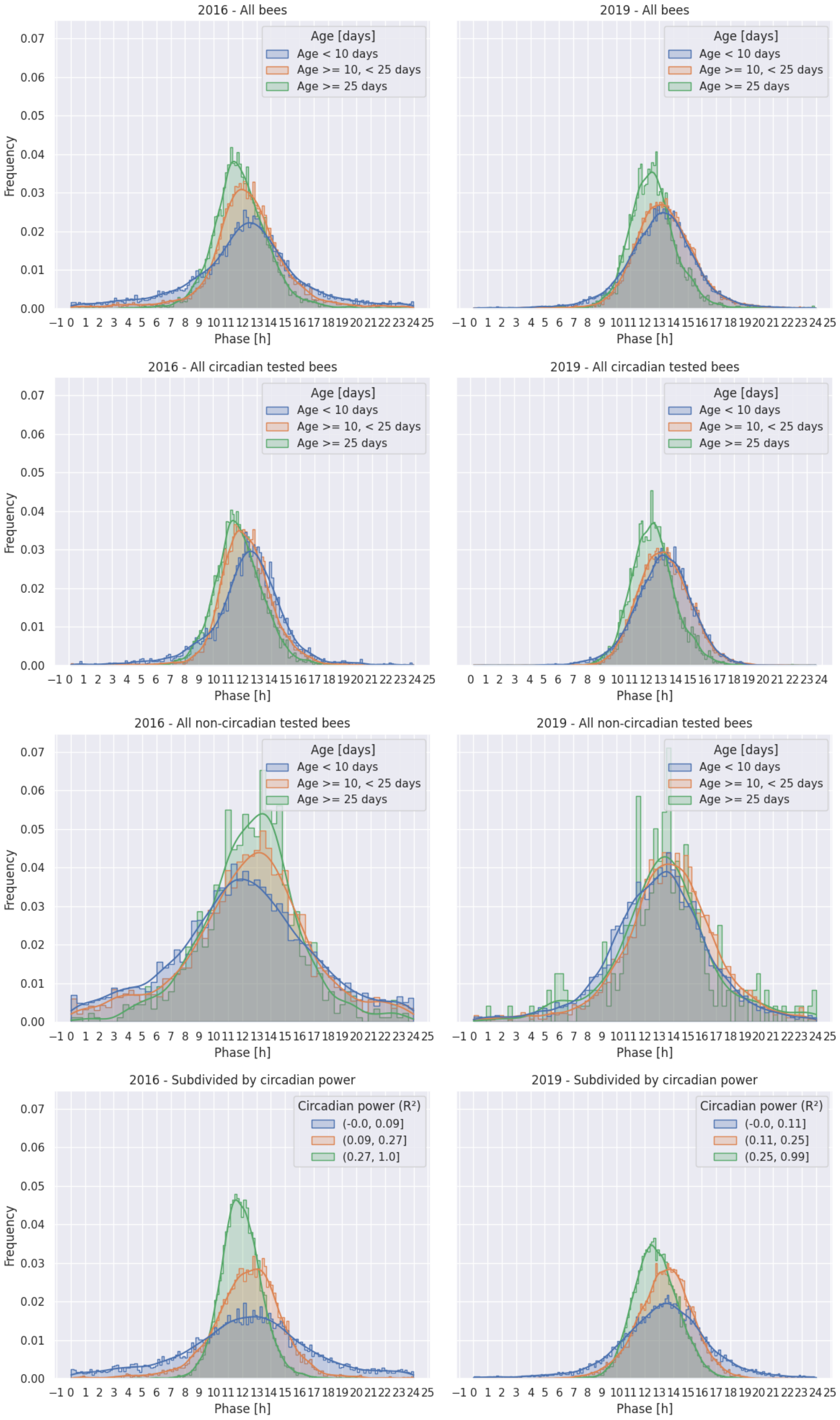


Figure 12: **Distribution of actvity peak - phase [h].** The distribution of phase is displayed mapped to 24h among three age groups for 2016 on the left side and on the right side for 2019. The first row displays the distribution among all bees, the second row the distribution of bees being tested circadian, the third row shows the distribution of not-circadian tested bees and the fourth row the distribution binned by circadian power (R^2).

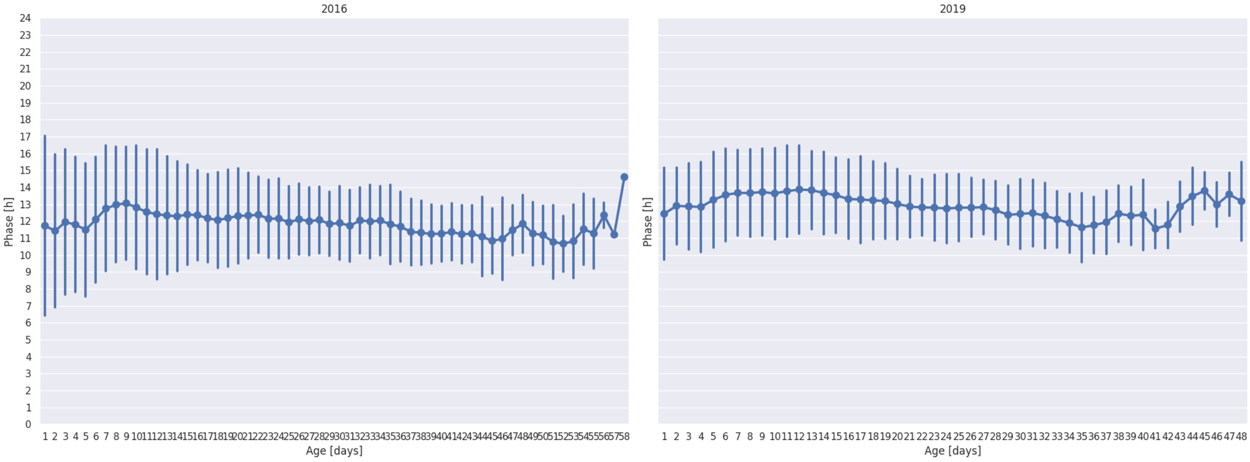


Figure 13: **Actvity peak - phase [h] - per age.** The median phase is displayed mapped to 24h in relation to their age with the error bars displaying the standard deviation for the age for 2016 on the left side and on the right side for 2019.

times until a reversal to earlier peak times occurs around age 13 in 2019 and age 10 in 2016. The second pattern, that younger bees tend to show more variability in their peak times, nevertheless remains apparent. This trend also explains the different peak sequence for 2016 and 2019. In 2019 more really young bees exist in the hive compared to 2016, which can be seen in the histogram in Figure 1, and those shift the young bee group to peak earlier than the intermediate ones.

The results of this study show clear patterns. High circadian performance and older age (25 years and older) are associated with an earlier and synchronized peak in activity. Very young bees are found to have an early activity peak first, get late and then return to an earlier one. Uncircadian bees have a wide distribution of their activity peaks, with no or reverse shift in age groups, and their peak is around midday. The differences in distribution depending on whether circadian performance or age is subdivided confirm the results from Section 4.2.1 that bees individually develop circadian rhythms at different ages.

From these patterns, several behavioral insights can be derived:

1. Age and Circadian Rhythm Strength: Depending on their age and the maturity of their circadian rhythm strength, bees exhibit a temporally shifted peak of activity.

2. Synchronization Among Circadian Bees: Circadian bees tend to synchronize their activity peaks with their peers, depending on their level of circadian rhythm development.

3. Young Bees: It is hypothesized that very young bees, compared to those with moderate circadian strength, initially develop a later peak in activity. However, as they age, around 10 to 12 years, they synchronize with older and highly circadian bees, resulting in an earlier peak. This may represent a critical age for the establishment of a robust circadian rhythm or starting development of forager behavior.

4. Non-Circadian and Very Young Bees: Both non-circadian and very young bees do not display randomly distributed phases; instead, their activity peaks also occur around midday.

4.3 Summary

Analysis of rhythmicity in bees across their lifespan reveals age-related patterns. Bees become more rhythmically mature as they age, but this development isn't linear.

Initially, in their early days, rhythmicity intensifies, then plateaus, with some bees highly matured and others less. Subsequently, they continue maturing until reaching their final circadian strength. These trends differ between two periods, with 2019 showing a more pronounced and earlier plateau. Additionally, bee's peak activity times vary with age and circadian maturity, with a trend of earlier, then later, and finally earlier peaks. The oldest bees peak earliest and are most synchronized, middle-aged bees show moderate variance, and the youngest bees have the highest variance. Overall, peaks tend to occur around noon, not uniformly throughout the day.

Due to the observations, it is hypothesized that the middle bees synchronize their activity phase after the older bees, by transferring activity from bees with a strong circadian rhythm to bees with a weaker circadian rhythm.

5 External Factors influencing Circadian Rhythm

Given the influence of environmental and external factors on bee rhythms outlined in 2.2.1, it is important to evaluate the extent of this influence, possible variations due to the degree of circadian strength, and the effects in a natural environment. Therefore, a correlation analysis is performed between the activity time series of each bee and the time series of weather parameters and sun positions.

5.1 Influence Weather

One of the primary external factors influencing bee activity is the weather. To understand this influence, a time-lagged cross-correlation analysis of bee activity and various weather parameters is conducted. The results, including the maximum and minimum correlation values for each parameter in relation to the circadian prominence of the bees, are presented in Figure 14.

In both time periods, the highest median positive correlation is observed with air temperature (0.22-0.8 in 2016, 0.22-0.8 in 2019), followed by sunshine duration (0.3-1.0 in 2016, 0.26-0.75 in 2019), wind speed (0.27-1.0 in 2016, 0.25-0.64 in 2019), and humidity (0.25-0.75 in 2016, 0.22-0.57 in 2019). The two lowest correlations are observed for precipitation duration (0.29-0.6 in 2016, 0.3-0.45 in 2019) and height (0.22-0.8 in 2016, 0.28-0.35 in 2019). The highest median negative correlation is observed with humidity (-0.22 to -1.0 in 2016, -0.22 to -0.8 in 2019), followed by air temperature (-0.22 to -0.075 in 2016, -0.22 to 0.62 in 2019), wind speed (-0.23 to -0.8 in 2016, -0.22 to -0.61 in 2019), sunshine duration (-0.21 to -1.0 in 2016, -0.25 to -0.78 in 2019), precipitation duration (-0.23 to -0.57 in 2016, -0.22 to -0.31 in 2019), and the least negative correlation is observed with precipitation height (-0.13 to 0.22 in 2016, -0.18 to -0.21 in 2019). Due to the limited amount of data for older bees, the correlation is particularly noisy in these age groups, where the highest but relatively unreliable correlation is mostly achieved. Apart from the precipitation parameters, there is a clear trend that the more circadian the bee, the higher the correlation. The non-existence of this trend may be explained by the unequal distribution of precipitation on different days, resulting in lower correlations.

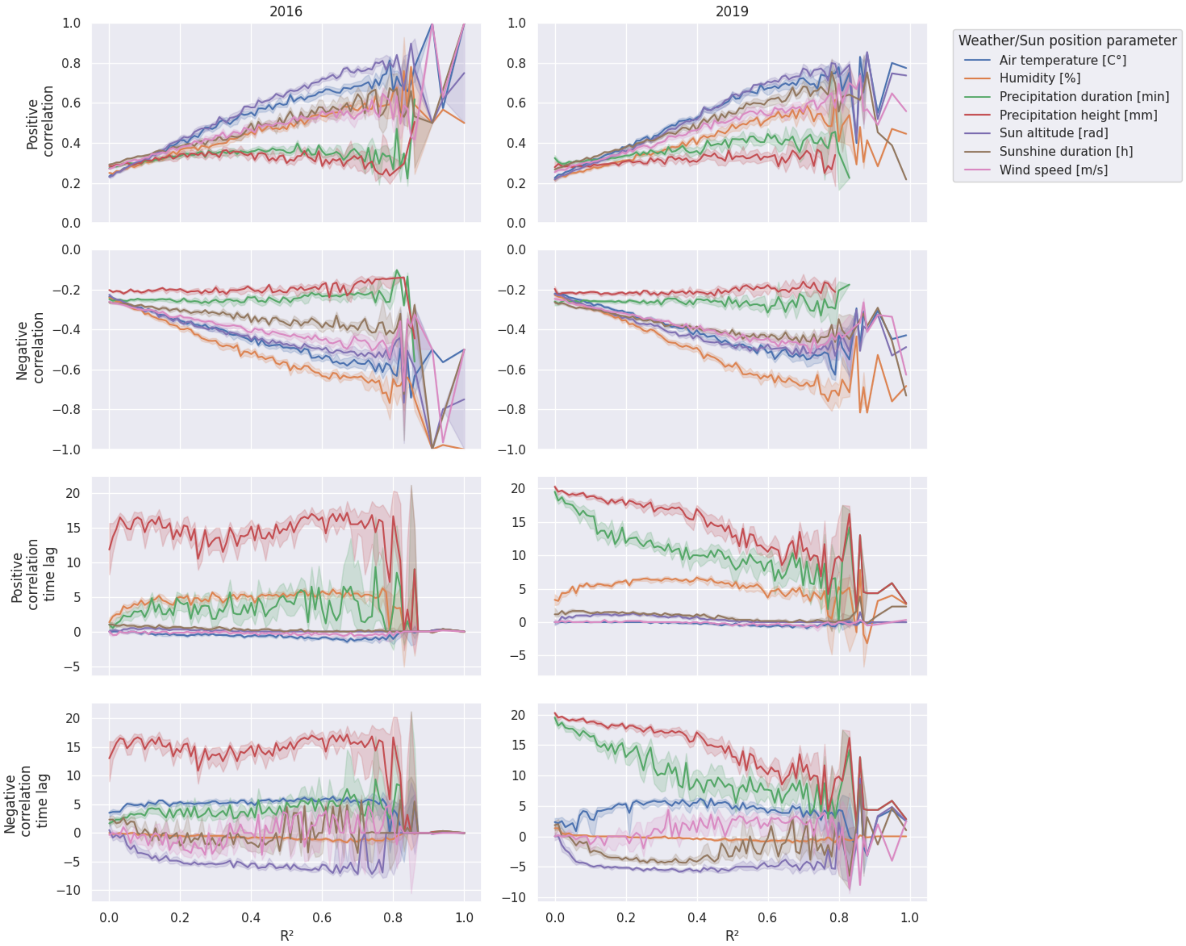


Figure 14: **Time-lagged cross-correlation for environmental factors and bee velocity.** The maximum (first row) and minimum (second row) cross-correlation is displayed in relation to the circadian power of the bee and the according time lags (third and fourth row).

The time lag patterns for these distributions exhibit different trends. In 2016, there is no clear pattern, while in 2019, for the precipitation parameters, both positive and negative correlations show a trend of a faster reaction with increased circadianity. However, since these parameters do not exhibit high correlations, this trend may not be meaningful.

In summary, the influence of weather parameters on bee activity shows that parameters related to sunny and warm weather have the highest positive correlation, indicating that more warmth and sunshine lead to higher bee activity. Conversely, the highest negative correlation is observed with humidity, indicating that higher humidity is associated with lower bee activity. In general, weather parameters with transient high variability, such as precipitation, have a lower impact on bee activity than stably changing parameters such as temperature. This suggests that the circadian rhythm of bees is not easily disrupted by external factors. The tendency for bees with stronger circadian rhythms to also exhibit increased correlation may be due to greater weather sensitivity or the fact that bees with stronger rhythms also exhibit greater variability consistent with the variable nature of weather parameters.

5.2 Influence Sun Position

The sun's position holds significance for bees, as it plays a vital role in their communication of directions to food sources through the waggle dance and serves as a natural daily zeitgeber. This study explores the potential impact of the sun's position on circadian rhythms, with results presented through time-lagged cross-correlation and regression analysis. The Figures 2 and 3 display the altitude of the sun over the time periods, while Figure 14 illustrates the maximum and minimum cross-correlation values and their corresponding time lags in relation to the prominence of the circadian rhythm.

The cross-correlation analysis with altitude reveals median maximum/minimum correlation values of 0.23-0.9/-0.23–1.0 for 2016 and 0.22-0.85/-0.23–0.5 for 2019. Interestingly, the results suggest that the sun's influence as a zeitgeber becomes more pronounced in circadian bees. However, in the case of very old bees, there is more variability, likely due to a smaller dataset for bees of that age group.

To further explore this relationship, regression analysis is conducted between altitude and maximum/minimum correlation values. The positive correlation regression shows a moderately high R^2 of 0.6 in 2016 and 0.55 in 2019, while the negative correlation regression is weaker with R^2 values of 0.39/0.32. This weaker negative correlation can be attributed to the formula of cross-correlation, which results in smaller correlations for larger time lags. In the case of negative correlations, a larger time lag of approximately 12 hours is needed for the negative correlation to become evident.

The time lag for positive correlation has median values ranging from 0-1h in 2016 and 0-1.5h in 2019, while for negative correlation, it ranges from 0.1–7h/0–6h. There's a slight tendency (difference of 1/1.5h) indicating that the time lag is smaller for more circadian bees with the exception of similar small-time lag for very low circadian bees. This fits to the shifted peak times described in 4.2.2. This supports that bees with a stronger circadian rhythm exhibit greater synchronization with the sun's position and their phase alignning with sun's noon position.

Conversely, there is a stronger time-lag tendency for negative correlation. This could be explained by the need for a time shift, resulting in a longer time lag for negative correlation. And accordingly, as bees become more rhythmic, their correlation with the sun's position becomes more significant, compensating for the loss of correlation as the time shift increases.

In summary the cross-correlation shows that bees follow the more circadian they are the more the suns position as a zeitgeber and synchronize their rhythms with the expected tendency of a positive correlation rather than a negative correlation. Still there is no perfect correlation indicating other parameters also play a role influencing the prominence of the bee's activity.

5.3 Summary

Cross-correlation analysis between bee velocity and weather parameters reveals several patterns. First, there is a stronger correlation between all parameters in bees with a more pronounced circadian rhythm. Second, sunny and warm weather has a positive effect on bee activity, while humid weather has more of a negative effect. In addition, bees with a more mature circadian rhythm show a smaller time delay in the peak of their activity compared to the position of the sun at noon.

6 Social Factors influencing Circadian Rhythm

Since a possible behavioral pattern of high-circadian bees transferring their activity to low-circadian bees is observed in 4.2.2, this hypothesis needs further investigation. To develop a deeper understanding of how bees learn to synchronize their circadian rhythms, the effects on the change in velocity and thus activity following a physical interaction with another bee is investigated. This analysis is done for the time period of 2019.

6.1 Transfer of Speed via Physical Contact

First, it is investigated whether speed, and thus activity, is transferred between interacting bees. For this purpose, bees are divided according to their speed at the beginning of the interaction, see Figure 15. When bees with high velocity meet bees with lower velocity, a velocity exchange takes place. In this dynamic, the faster bee loses speed while the slower bee gains speed and vice versa. It can also be observed that there is a threshold for the increase in speed above which the bees benefit from the increase in speed. The bees that benefit most from such interactions tend to be those whose speed is in the median range rather than the slowest. This tendency holds for both median and mean speed.

However, there is also the effect that the faster the bee is, the more likely it is to slow down. To see if the interaction really provides a speed transfer, the speed change per bin is compared to the null model. In the null model the effect is less strong and a Welch-test confirms that after an interaction the effect is statistically significant. The results of the test are summarized in Figure 16.

Interactions seem to be activating overall, as a positive change in speed after interaction is observed in the mean and median of all subcombinations of bees visible in Figure 17. This positive effect is confirmed by the results of comparing the by circadian power binned velocity changes to the complements in the null model with a Welch-test with all compared bins having a p-value lower than $0.1e^{-6}$.

The studies show that interaction with another bee has on average an activating character, which not only activates the other bee, but also causes a transfer of speed between bees. Thus, the transfer of speed through physical contact is a fundamental effect of bee interactions and could be used to distribute activity in the colony.

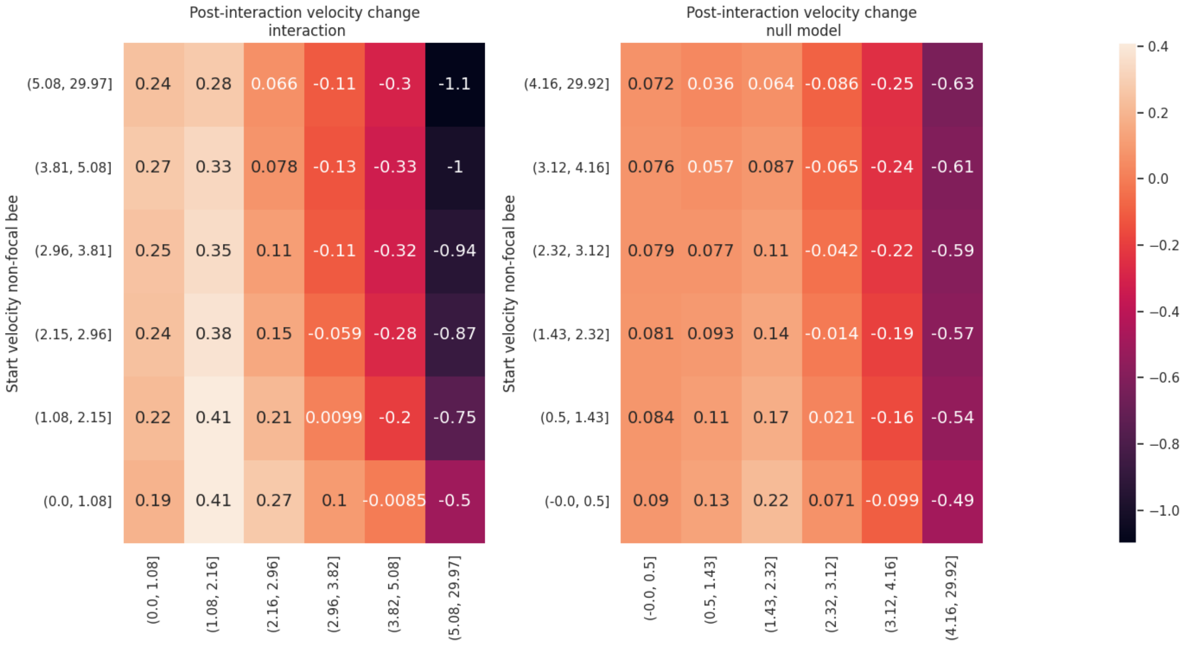


Figure 15: **Speed is transferred.** It displays the median of post-interaction velocity change of the focal bee binned by the start velocities on the y-axis of the non-focal bee and the x-axis the focal. The left side displays the interactions and the right side the null model.

6.2 Identification of Bee Groups Transferring and Receiving Speed

The investigation delves further into the diverse subgroups of bees, extending the analysis to identify intricate patterns of speed transmission. This multifaceted exploration encompasses the roles by bees with varying circadian power on speed transfer. Additionally, it is explored how age, phase, and foraging behavior impact speed dynamics, to further investigate on whether they serve as speed donors or recipients.

To determine whether these velocity transmissions play a role in the development of circadian rhythms, the interactions are subdivided according to their circadian performance and thus their activity level. Patterns in such a combination could reveal whether circadian bees activate uncircadian bees during the day, causing them to develop a rhythm. This is displayed in Figure 17a and 17b.

The analysis reveals a clearer pattern: for both the mean and median, it can be observed that the subgroup with the most circadian bees clearly transmits the most velocity when interacting. Moreover, the group with the highest circadianity is the one that achieves the least change in velocity after an interaction. The other subgroups, except they interact with the most circadian group, have a very similar effect on the change in velocity for the median. The median also shows a higher speed transition from intermediate circadian bees to intermediate circadian bees. The median has a wider range (negative and positive), suggesting that some interactions have a large effect on velocity transfer. This difference is most visible when the mid-circadian groups are velocity receivers, meaning that the positive-impact interactions occur with the participation of this group.

The data is also broken down by the duration of the interactions. The more extreme values

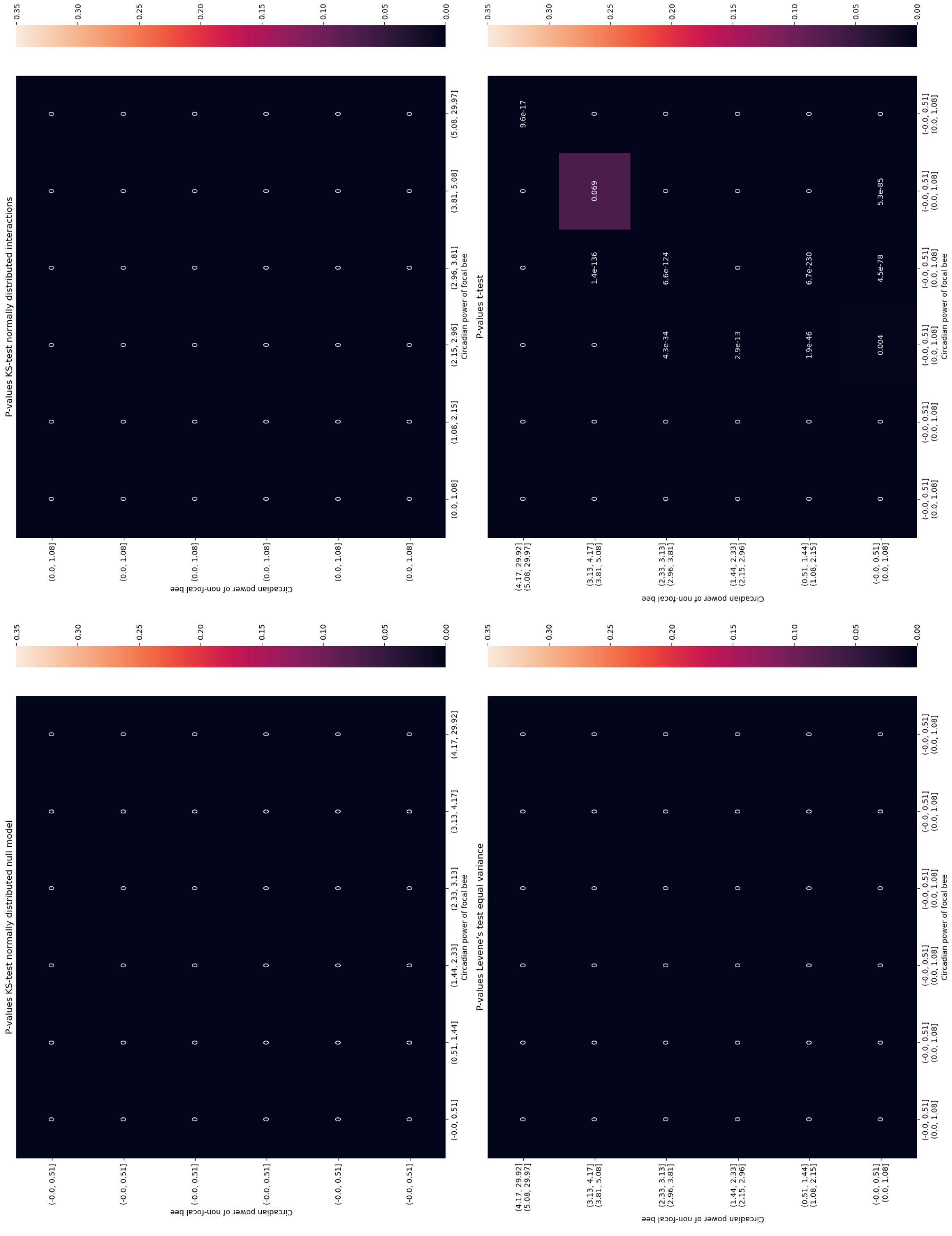
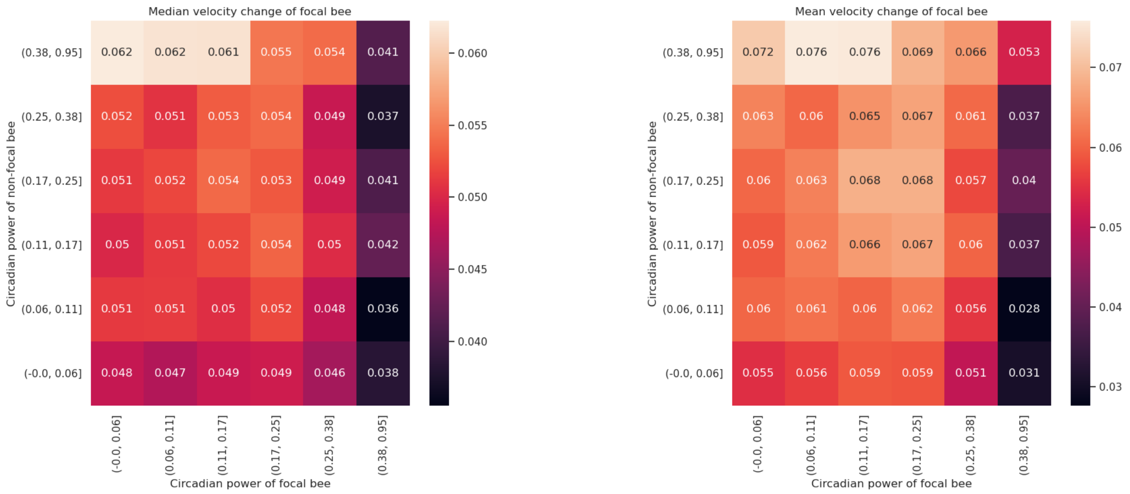
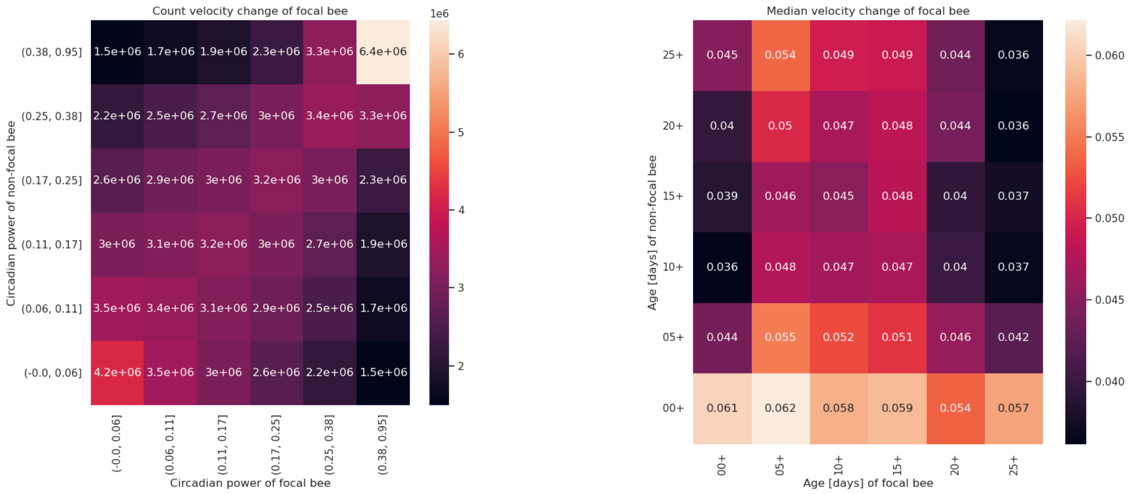


Figure 16: **Test result Welch-test (T-test with unequal variance) of binned velocity changes.** It displays on the first row the p-values for whether the velocity changes of the focal bee, binned by the start velocities of the non-focal and focal bee is normally distributed and accordingly the null model bins. All are not normally distributed, nevertheless a T-test can be applied due to large sample size. The second row displays the results whether the bins have equal variance according the Levenes-test (left side). All bins have unequal variance. The right side displays the result of the Welch-test with all bins being significantly different except one bin.



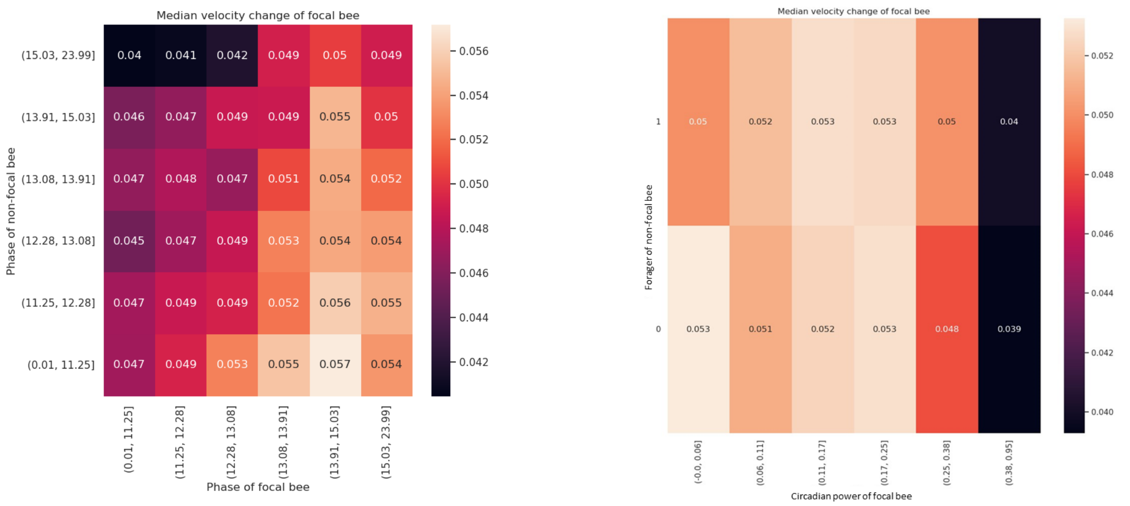
(a) Median - Circadian power (R^2)

(b) Mean - Circadian power (R^2)



(c) Count - Circadian power (R^2)

(d) Median - Age



(e) Median - Phase

(f) Median - Circadian power (R^2) - is forager

Figure 17: **Post-interaction velocity change focal bee.** Here the median, mean or count of post-interaction velocity change of the focal bee binned by different metrics is displayed with on the y-axis the non-focal bee and the x-axis the focal.

occur in longer interactions (30 - 120 seconds). On the other hand, there is not as much data in these ranges, so this could also be an effect of the smaller number of samples and therefore the patterns are noisier. There is also a difference between the different times of the day: The pattern is recognizable that the highest velocity transmission is reached between 12 and 16 o'clock. Bees with low circadian power constantly increase their speed. When interacting with the group with the highest circadian power, differences can be observed between bees with low and medium circadian power. In bees with low circadian power, the increase in speed is most pronounced at night. Conversely, there is a slight tendency for bees with medium circadian power to speed up mainly during the day than at night. These effects can be seen for duration in Figure 18 and for different day times in Figure 19.

Looking at the differentiation of interactions of non-circadian bees with circadian bees, displayed in Figure 20, shows that the circadian tested bees have a greater variation in transmitted velocity. On average, the non-circadian bees show the pattern that the more circadian a bee is, the more velocity is transmitted. This differs from the circadian bees, where it is observed that the less circadian the bee, the more velocity is transmitted when interacting with a bee with high circadianity. In terms of number of interactions, bees interact most with bees with similar circadianity. This pattern is also seen in the circadian tested bees but differs in the non-circadian tested bees. Here, the opposite pattern is seen: bees with low circadianism interact more frequently with bees with high circadianism and vice versa. It should be noted that groups 1 to 4 in the non-circadian subgroup have a comparable range of circadian power as the first group in the circadian subgroup, so that one field in the circadian subgroup corresponds to 16 bins in the non-circadian subgroup.

The bees are also classified according to their age, visible in 17d. This shows that really young and the old bees gain little speed from the interaction. There is also the pattern evident that the interaction with a very young bee transfers the most velocity.

In addition, bees are also divided into subgroups according to their peak activity time during the day, as shown in Figure 17e. Bees that reach their peak activity later in the day receive more velocity than bees that reach their peak activity earlier, and that the velocity senders are the bees that have their peak activity earlier.

Interactions between bees classified as foragers and bees with different circadian performances shows the following and is displayed in Figure 17f. Interaction with bees classified as foragers primarily activates mid-circadian groups. Interaction with non-foragers is found to activate medium-circadian group the most at median, and at mean the lowest circadian group. Low-circadian groups experience a smaller speed increase from interaction with foragers compared with interaction with non-foragers.

Based on the observations, the most prominent speed receivers among bees are those with low und intermediate circadian power. These bees tend to experience an increase in speed during interactions, especially when they engage with highly circadian bees. On the other hand, the most significant speed transmitters are the bees with the highest circadian power. These highly circadian bees consistently transmit the most velocity during interactions, and their own speed showing the least positive change after interaction. Among the subgroup of highly circadian bees, foragers exhibit a distinct propensity for transferring greater speed to intermediate circadian bees. An intriguing observation in terms of age reveals that very young bees emerge as speed transmitters too. Notably the observations reveal a nocturnal phenomenon, where low circadian bees receive the most speed. These nuances suggest that different mechanisms may underlie the speed transfer and do not undermine the hypothesis that very young and uncircadian bees acquire speed from rhythmic bees during the day.

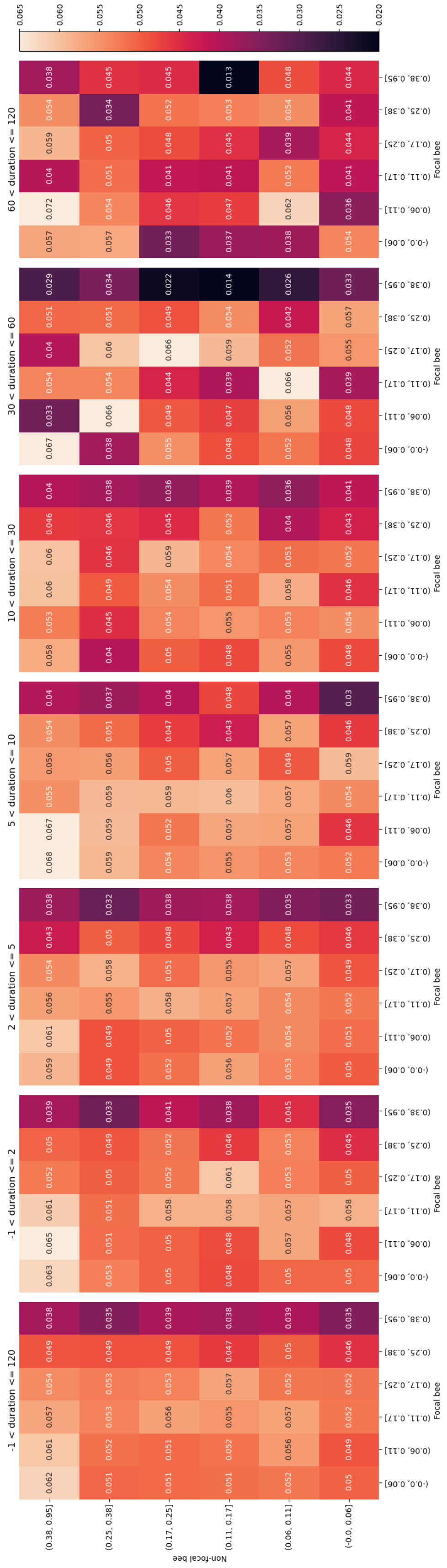


Figure 18: **Velocity change for different durations of interactions.** The velocity change of the focal bee binned the circadian power of non-focal and focal bee for different durations (all, 0-2s, 2-5s, 5-10s, 10-30s, 30-60s, 60-120s) is displayed.

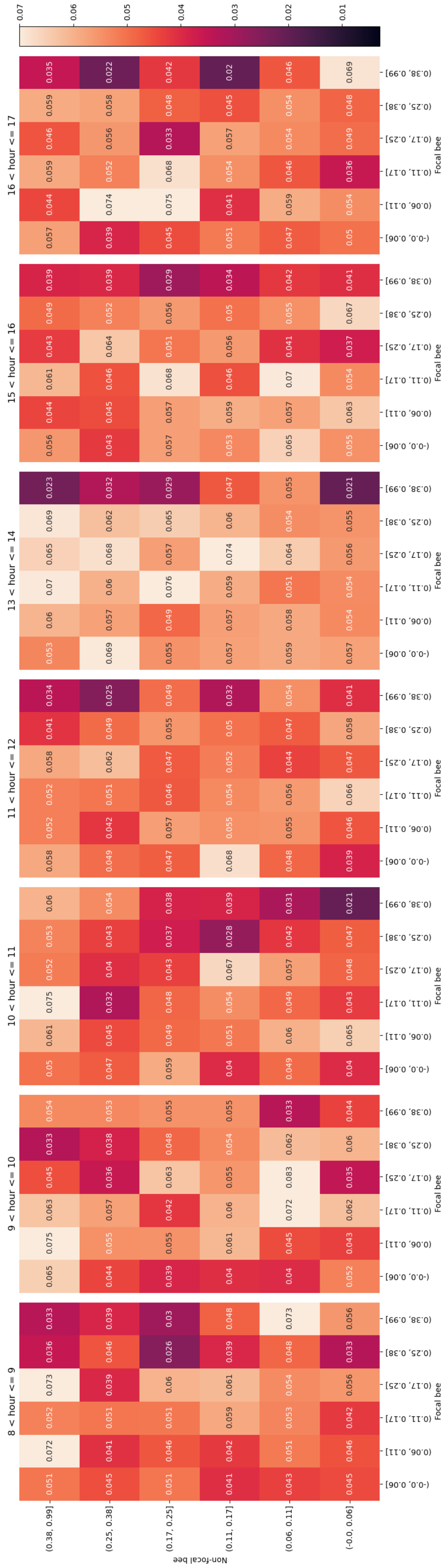
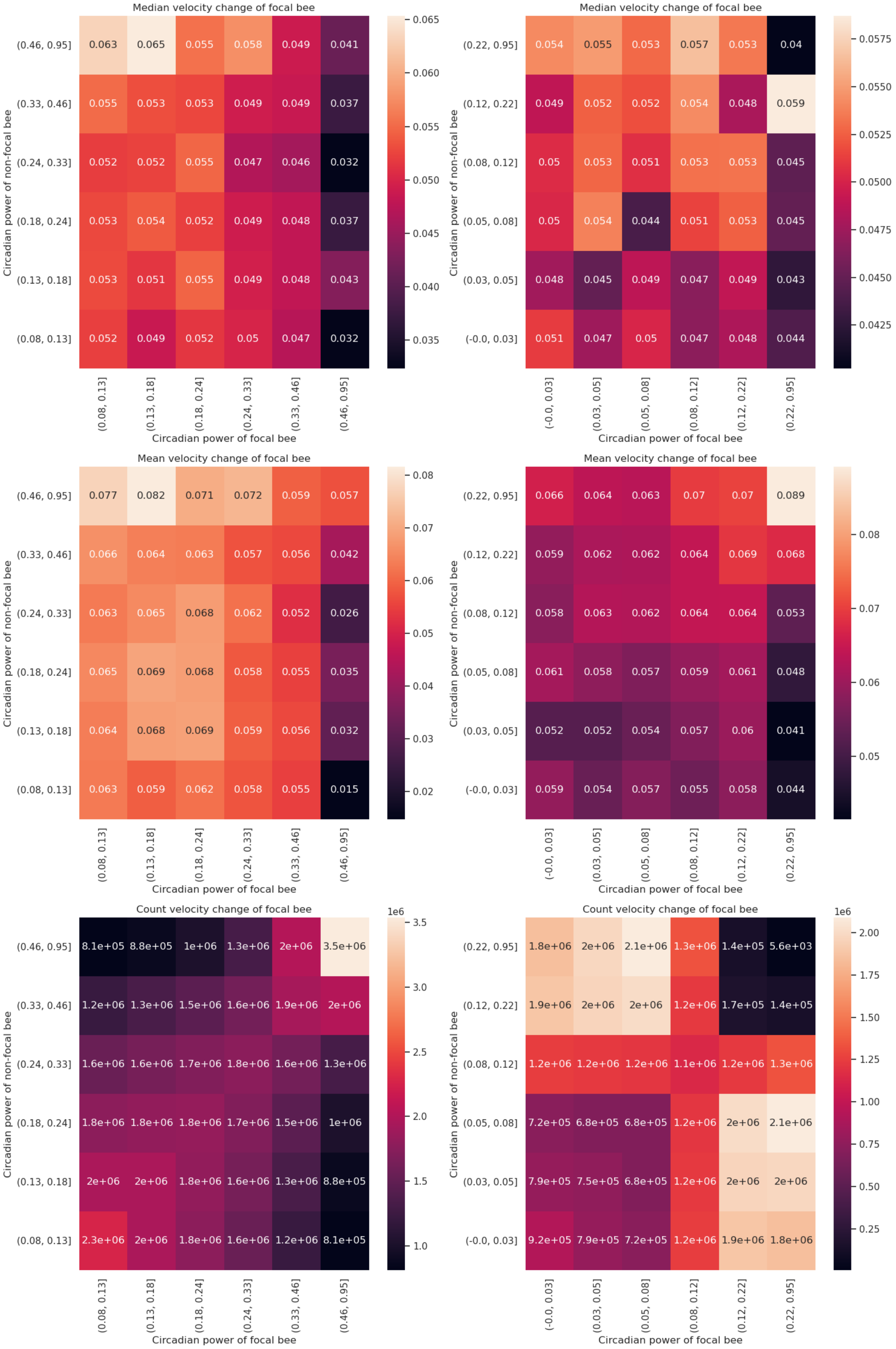


Figure 19: **Velocity change for different daytimes interactions take place.** The velocity change of the focal bee binned the circadian power of non-focal and focal bee for among the day (all, 0-5h, 5-9h, 9-12h, 12-16h, 16-20h, 20-24h) is displayed.



Median velocity change of focal bee

	(0.08, 0.13]	(0.13, 0.18]	(0.18, 0.24]	(0.24, 0.33]	(0.33, 0.46]	(0.46, 0.95]
(0.46, 0.95]	0.063	0.065	0.055	0.058	0.049	0.041
(0.33, 0.46]	0.055	0.053	0.053	0.049	0.049	0.037
(0.24, 0.33]	0.052	0.052	0.055	0.047	0.046	0.032
(0.18, 0.24]	0.053	0.054	0.052	0.049	0.048	0.037
(0.13, 0.18]	0.053	0.051	0.055	0.049	0.048	0.043
(0.08, 0.13]	0.052	0.049	0.052	0.05	0.047	0.032

Median velocity change of focal bee

	(-0.0, 0.03]	(0.03, 0.05]	(0.05, 0.08]	(0.08, 0.12]	(0.12, 0.22]	(0.22, 0.95]
(0.22, 0.95]	0.054	0.055	0.053	0.057	0.053	0.04
(0.12, 0.22]	0.049	0.052	0.052	0.054	0.048	0.059
(0.08, 0.12]	0.05	0.053	0.051	0.053	0.053	0.045
(0.05, 0.08]	0.05	0.054	0.044	0.051	0.053	0.045
(0.03, 0.05]	0.048	0.045	0.049	0.047	0.049	0.043
(-0.0, 0.03]	0.051	0.047	0.05	0.047	0.048	0.044

Mean velocity change of focal bee

	(0.08, 0.13]	(0.13, 0.18]	(0.18, 0.24]	(0.24, 0.33]	(0.33, 0.46]	(0.46, 0.95]
(0.46, 0.95]	0.077	0.082	0.071	0.072	0.059	0.057
(0.33, 0.46]	0.066	0.064	0.063	0.057	0.056	0.042
(0.24, 0.33]	0.063	0.065	0.068	0.062	0.052	0.026
(0.18, 0.24]	0.065	0.069	0.068	0.058	0.055	0.035
(0.13, 0.18]	0.064	0.068	0.069	0.059	0.056	0.032
(0.08, 0.13]	0.063	0.059	0.062	0.058	0.055	0.015

Mean velocity change of focal bee

	(-0.0, 0.03]	(0.03, 0.05]	(0.05, 0.08]	(0.08, 0.12]	(0.12, 0.22]	(0.22, 0.95]
(0.22, 0.95]	0.066	0.064	0.063	0.07	0.07	0.089
(0.12, 0.22]	0.059	0.062	0.062	0.064	0.069	0.068
(0.08, 0.12]	0.058	0.063	0.062	0.064	0.064	0.053
(0.05, 0.08]	0.061	0.058	0.057	0.059	0.061	0.048
(0.03, 0.05]	0.052	0.052	0.054	0.057	0.06	0.041
(-0.0, 0.03]	0.059	0.054	0.057	0.055	0.058	0.044

Count velocity change of focal bee

	(0.08, 0.13]	(0.13, 0.18]	(0.18, 0.24]	(0.24, 0.33]	(0.33, 0.46]	(0.46, 0.95]
(0.46, 0.95]	8.1e+05	8.8e+05	1e+06	1.3e+06	2e+06	3.5e+06
(0.33, 0.46]	1.2e+06	1.3e+06	1.5e+06	1.6e+06	1.9e+06	2e+06
(0.24, 0.33]	1.6e+06	1.6e+06	1.7e+06	1.8e+06	1.6e+06	1.3e+06
(0.18, 0.24]	1.8e+06	1.8e+06	1.8e+06	1.7e+06	1.5e+06	1e+06
(0.13, 0.18]	2e+06	2e+06	1.8e+06	1.6e+06	1.3e+06	8.8e+05
(0.08, 0.13]	2.3e+06	2e+06	1.8e+06	1.6e+06	1.2e+06	8.1e+05

Count velocity change of focal bee

	(-0.0, 0.03]	(0.03, 0.05]	(0.05, 0.08]	(0.08, 0.12]	(0.12, 0.22]	(0.22, 0.95]
(0.22, 0.95]	1.8e+06	2e+06	2.1e+06	1.3e+06	1.4e+05	5.6e+03
(0.12, 0.22]	1.9e+06	2e+06	2e+06	1.2e+06	1.7e+05	1.4e+05
(0.08, 0.12]	1.2e+06	1.2e+06	1.2e+06	1.1e+06	1.2e+06	1.3e+06
(0.05, 0.08]	7.2e+05	6.8e+05	6.8e+05	1.2e+06	2e+06	2.1e+06
(0.03, 0.05]	7.9e+05	7.5e+05	6.8e+05	1.2e+06	2e+06	2e+06
(-0.0, 0.03]	9.2e+05	7.9e+05	7.2e+05	1.2e+06	1.9e+06	1.8e+06

Figure 20: **Velocity change for non- and circadian bees.** It displays the mean, median and count of velocity change per focal bee binned by the circadian power of the non-focal and focal bee and subdivided by the left side the circadian bees and on the right side the non-circadian bees.

On the other hand, especially the effect that medium-circadian bees receive speed from very-circadian bees during the day is visible, which fits the hypothesis that activity of very-circadian bees is transferred to medium-circadian bees. This is confirmed by the observation that bees that peak late in the day receive speed from bees that peak earlier in the day.

6.3 Summary

Studying the effects of interactions on a bee's activity shows that bees transmit speed and thus activity via a physical interaction. The faster the bee is, the more speed is transmitted, and the slower the bee is, the more speed is received, while not being too slow. In summary, bees with low circadian power and late phase are the primary beneficiaries of speed during interactions, while bees with high circadian power are the main contributors to speed transmission within the bee colony. Furthermore, intermediate circadian bees receive the highest speed boosts during the daytime and more speed from forager bees compared to low circadian bees. These findings undermine the possible hypothesis of a spread of activity via speed transfer from highly circadian to intermediate circadian bees during daylight hours.

7 Conclusion and Outlook

This work involves a comprehensive analysis of circadian rhythms in two bee colonies in a natural environment which reveals several behavioral patterns.

The analysis is based on observational data of bee activity from two colonies, which are modeled by a cosinor model to effectively study the rhythmic behavior of bees. The proportion of bees exhibiting rhythmic behavior, the expression of these rhythms, the temporal organization over the course of a day, and over the course of a bee's life are examined, with a focus on the development and maturation of rhythm. In addition, the effects of external factors such as weather conditions and sun position on the activity patterns of bees are studied, with a particular emphasis on variations depending on the strength of circadian rhythms. An interesting aspect of the analysis concerns the temporal organization of these rhythms differing throughout the life of a bee, suggesting a potential transfer of activity from bees with high circadian rhythms to bees with lower circadian rhythms. Subsequently, the role of physical contact interactions between bees on their activity and potential activity transmission in social entrainment is investigated.

Several conclusions can be drawn from the questions posed in the study:

Rhythmicity in bees and its extent: the circadian rhythm of honey bee activity follows a cosine-like pattern characterized by marked fluctuations between day and night activity periods. This rhythmic behavior becomes more pronounced with age. Remarkably, circadian rhythms are evident even in young bees as young as 1-5 days old. These young bees mainly show differences in more pronounced nocturnal inactivity compared to their non-circadian conspecifics. This observation that potential nurse and cleaning bees are also proportionally rhythmic is a new observation compared to the reviewed literature.

The age-dependent maturation of circadian rhythms: As honey bees age, their circadian rhythms undergo a maturation process that results in increased rhythmicity. This maturation process is not linear nature. In the first days of a bee's life, its circadian rhythm increases, then reaches a plateau phase. Within this plateau phase, there are differences within the bee population, with some individuals showing advanced circadian maturation while others show a less pronounced circadian rhythm. Subsequently, bees continue to mature in terms of circadian rhythm strength until they reach a final level of rhythmicity. These trends differ between the two periods, with the plateau pattern being more pronounced in 2019 and occurring at earlier ages.

Synchronization and temporal shifts of rhythms within the colony: Bees show differences in the timing of their main activity, which are influenced by both their age and the degree of circadian maturity. In particular, older bees, which have stronger circadian rhythms, show earlier activity peaks and the highest degree of synchronization. In contrast, younger bees and those with lower circadian rhythms show greater variability in the timing of their peak activity the younger or more uncircadian they are. A common feature of all bees is that a majority of activity peaks around midday, unlike the assumption that they are evenly distributed throughout the day in non-circadian and very young bees. In addition, there is a trend where bees initially peak at an earlier hour, then shift to a later time, and finally return to an earlier peak. An interesting aspect of the analysis concerns the temporal shift from strong circadian/old bees to mid-circadian/medium-aged bees, suggesting a possible transfer of activity from bees with strong circadian rhythms to bees with lower circadian rhythms.

Influence external factors to bees rhythms: Sunny and warm weather positively affects bee activity, while humid weather has a negative impact. The study demonstrates that bees with stronger circadian rhythms are more sensitive to weather conditions, indicating that external factors can influence the prominence of bee activity. However, the circadian rhythm of bees remains relatively resilient to external disruptions like precipitation.

Contribution of Physical Body Contact to Rhythm Development: Speed, and thus activity, can be transferred between bees through physical contact. Bees with low circadian power benefit most from speed transfer during interactions, especially when interacting with bees with high circadianism. Conversely, high circadian bees contribute most to speed transmission within the colony. This could be one of the mechanisms to socially entrain circadian rhythmicity in the colony. However, very young bees have also been identified as velocity transmitters, suggesting further investigation into what type of interaction plays a role in activity transmission.

Concluded when comparing the two time periods, several similar patterns emerge in the analysis. At every age, a share exhibits rhythmic behavior, while young uncircadian bees tend to express non-rhythmic behavior with increased activity during nighttime. The older the bees become, the more prominent their rhythmicity, following a non-linear maturation process with high variance observed in middle-aged bees. This suggests that some bees mature earlier than others. As bees age, their temporal organization in peak activity becomes more synchronized, showing an unlinear trend of early-peaking, followed by late-peaking, and then early-peaking again. The oldest bees and most circadian bees tend to peak the earliest among all age groups. All bees exhibit activity distributions centered around noon, but circadian bees tend to have narrower and more synchronized distributions.

However, there are also differing patterns between the two years. In 2019, a higher proportion of bees display rhythmic behavior compared to 2016. Furthermore in 2019, there is a longer, earlier, and stronger plateau pattern of rhythm prominence and phase observed in intermediate-aged bees. In 2019, the intermediate age group does not peak earlier than the youngest bees, while in 2016, the opposite pattern is observed. Even thought both distributions show a very close peak time of a difference of 4 minutes in 2016 and 2019. These differences could be attributed to the presence of more very young bees in 2019 and the longer aging of bees in 2016, serving as an explanation for the different distributions in phase shifts, as 2016 had fewer very young bees and thus the group of bees 10 days old do not shift as a group that early. The observed plateau in 2019 bees and the higher proportion of very young circadian bees, along with the steeper maturation process in 2016, warrants further investigation and opens room for interpretation of biological significance. It might be beneficial to conduct further research with a third colony to better evaluate these differences.

The bees' activity and, consequently, their rhythmic behavior are currently modeled using a single-component cosinor, which serves as a suitable model for describing their rhythms. However, there is room for improvement in terms of reliability in rhythmicity tests, particularly with regard to the normality of the residuals. To enhance the model, potential approaches include further transformations of the velocity data, extending beyond low-pass filtering and median calculations. Alternatively, employing a multi-component cosinor model could address the issues associated with the extended, inactive periods at night and the bursty nature of bee behavior. Possible future research directions include also a spatial analysis of hive dynamics to better understand who peaks when and their roles in the colony which can be better understand locating them in the hive. Additionally, confirming the influence of body contact on activity patterns could be explored in a second colony and the kind of interactions transferring speed more detailed described.

Acknowledgements

A special thanks goes to Professor Tim Landgraf for supervising and encouraging my work. I also thank David Dormagen for introducing me to the topic and providing data, knowledge, and ideas for deeper analysis. I also want to thank Professor Giovanni Galizia and Professor Michael Smith for helping me with biological interpretations and deriving more advanced biologically important analysis topics. A big thanks and appreciation also goes to Weronika Klos, with whom I worked intensively and who is writing a supplementary paper on this topic. I thank the HPC service of ZEDAT, Freie Universität Berlin, for computing time and cluster infrastructure [71].

I would also like to mention the use of ChatGTP [72] as a formulation aid for the text of the thesis. It was used in a form in which self-defined notes and text modules are generated into a complete text with the help of ChatGTP and this text is then manually corrected, improved and modified for final use.

Abbreviations

ANOVA	Annotated Variance Analysis
A	Amplitude
DW	Durbin Watson
M	MESOR
MSS	Regression model sum of squares
P	Period
RSS	Residual sum of squares
TSS	Total sum of squares

References

[1] Guy Bloch, Esther Hazan, and Ada Rafaeli. “Circadian rhythms and endocrine functions in adult insects”. In: *Journal of Insect Physiology* 59.1 (2013), pp. 56–69. ISSN: 0022-1910. DOI: `10.1016/j.jinsphys.2012.10.012`. URL: `https://www.sciencedirect.com/science/article/pii/S0022191012002648`.

[2] T. Kondo and M. Ishiura. “Circadian rhythms of cyanobacteria: monitoring the biological clocks of individual colonies by bioluminescence”. In: *Journal of bacteriology* 176.7 (1994), pp. 1881–1885. ISSN: 0021-9193. DOI: `10.1128/jb.176.7.1881-1885.1994`.

[3] Deborah Bell-Pedersen et al. “Circadian rhythms from multiple oscillators: lessons from diverse organisms”. In: *Nature Reviews Genetics* 6.7 (2005), pp. 544–556. ISSN: 1471-0064. DOI: `10.1038/nrg1633`. URL: `https://www.nature.com/articles/nrg1633`.

[4] G. E. Robinson. “Regulation of division of labor in insect societies”. In: *Annual Review of Entomology* 37.1 (1992), pp. 637–665. ISSN: 0066-4170. DOI: `10.1146/annurev.en.37.010192.003225`.

[5] Thomas D. Seeley. *The Wisdom of the Hive: The Social Physiology of Honey Bee Colonies*. Cambridge, MA: Harvard University Press, 2022. ISBN: 9780674043404. DOI: `10.4159/9780674043404`. URL: `https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=3300714`.

[6] Wilhelm BEIER and Martin LINDAUER. “DER SONNENSTAND ALS ZEITGEBER FÜR DIE BIENE”. In: *Apidologie* 1.1 (1970), pp. 5–28. ISSN: 0044-8435. DOI: `10.1051/apido:19700101`.

[7] Hayward G. Spangler. “Daily Activity Rhythms of Individual Worker and Drone Honey Bees1, 2”. In: *Annals of the Entomological Society of America* 65.5 (1972), pp. 1073–1076. ISSN: 0013-8746. DOI: `10.1093/aesa/65.5.1073`.

[8] DARRELL MOORE and MARY ANN RANKIN. “Circadian locomotor rhythms in individual honeybees”. In: *Physiological Entomology* 10.2 (1985), pp. 191–197. ISSN: 0307-6962. DOI: `10.1111/j.1365-3032.1985.tb00034.x`.

[9] DARRELL MOORE and MARY ANN RANKIN. “Light and temperature entrainment of a locomotor rhythm in honeybees”. In: *Physiological Entomology* 18.3 (1993), pp. 271–278. ISSN: 0307-6962. DOI: `10.1111/j.1365-3032.1993.tb00599.x`.

[10] Manuel A. Giannoni-Guzmán et al. “The Role of Colony Temperature in the Entrainment of Circadian Rhythms of Honey Bee Foragers”. In: *Annals of the Entomological Society of America* 114.5 (2021), pp. 596–605. ISSN: 0013-8746. DOI: `10.1093/aesa/saab021`.

[11] D. Moore. “Honey bee circadian clocks: behavioral control from individual workers to whole-colony rhythms”. In: *Journal of Insect Physiology* 47.8 (2001), pp. 843–857. ISSN: 0022-1910. DOI: `10.1016/S0022-1910(01)00057-9`. URL: `https://www.sciencedirect.com/science/article/pii/s0022191001000579`.

[12] BRIGITTE FRISCH and JÜRGEN ASCHOFF. “Circadian rhythms in honeybees: entrainment by feeding cycles”. In: *Physiological Entomology* 12.1 (1987), pp. 41–49. ISSN: 0307-6962. DOI: `10.1111/j.1365-3032.1987.tb00722.x`.

[13] Ada Eban-Rothschild, Yair Shemesh, and Guy Bloch. "The colony environment, but not direct contact with conspecifics, influences the development of circadian rhythms in honey bees". In: *Journal of biological rhythms* 27.3 (2012), pp. 217–225. DOI: 10.1177/0748730412440851.

[14] G. Bloch, D. P. Toma, and G. E. Robinson. "Behavioral rhythmicity, age, division of labor and period expression in the honey bee brain". In: *Journal of biological rhythms* 16.5 (2001), pp. 444–456. DOI: 10.1177/074873001129002123.

[15] J. H. Hunt and F.-J. Richard. "Intracolony vibroacoustic communication in social insects". In: *Insectes Sociaux* 60.4 (2013), pp. 403–417. ISSN: 1420-9098. DOI: 10.1007/s00040-013-0311-9. URL: https://link.springer.com/article/10.1007/s00040-013-0311-9.

[16] Edward O. Wilson. *The insect societies*. 5. print. Cambridge, Mass.: Belknap Pr. of Harvard Univ. Pr, 1979. ISBN: 0674454901.

[17] Yves Le Conte and Abraham Hefetz. "Primer pheromones in social hymenoptera". In: *Annual Review of Entomology* 53 (2008), pp. 523–542. ISSN: 0066-4170. DOI: 10.1146/annurev.ento.52.110405.091434.

[18] Sara Diana Leonhardt et al. "Ecology and Evolution of Communication in Social Insects". In: *Cell* 164.6 (2016), pp. 1277–1287. ISSN: 0092-8674. DOI: 10.1016/j.cell.2016.01.035.

[19] Oliver Siehler and Guy Bloch. "Colony Volatiles and Substrate-borne Vibrations Entrain Circadian Rhythms and Are Potential Cues Mediating Social Synchronization in Honey Bee Colonies". In: *Journal of biological rhythms* 35.3 (2020), pp. 246–256. DOI: 10.1177/0748730420913362.

[20] Benjamin Wild et al. "Social networks predict the life and death of honey bees". In: *Nature communications* 12.1 (2021), p. 1110. DOI: 10.1038/s41467-021-21212-5.

[21] David M. Dormagen et al. "Machine learning reveals the waggle drift's role in the honey bee dance communication system". In: *PNAS Nexus* 2.9 (2023), pgad275. DOI: 10.1093/pnasnexus/pgad275. URL: https://academic.oup.com/pnasnexus/article/doi/10.1093/pnasnexus/pgad275/7251052.

[22] Brian R. Johnson. "Division of labor in honeybees: form, function, and proximate mechanisms". In: *Behavioral Ecology and Sociobiology* 64.3 (2010), pp. 305–316. ISSN: 1432-0762. DOI: 10.1007/s00265-009-0874-7. URL: https://link.springer.com/article/10.1007/s00265-009-0874-7.

[23] Yair Shemesh et al. "Molecular dynamics and social regulation of context-dependent plasticity in the circadian clockwork of the honey bee". In: *Journal of Neuroscience* 30.37 (2010), pp. 12517–12525. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.1490-10.2010. URL: https://www.jneurosci.org/content/30/37/12517.short.

[24] DARRELL MOORE et al. "Timekeeping in the honey bee colony: integration of circadian rhythms and division of labor". In: *Behavioral Ecology and Sociobiology* 43.3 (1998), pp. 147–160. ISSN: 1432-0762. DOI: 10.1007/s002650050476.

[25] Yair Shemesh, Mira Cohen, and Guy Bloch. "Natural plasticity in circadian rhythms is mediated by reorganization in the molecular clockwork in honeybees". In: *The FASEB Journal* 21.10 (2007), pp. 2304–2311. ISSN: 1530-6860. DOI: 10.1096/fj.06-8032com.

[26] Taro Fuchikawa et al. "Potent social synchronization can override photic entrainment of circadian rhythms". In: *Nature communications* 7 (2016), p. 11662. DOI: 10.1038/ncomms11662.

[27] Robin F. A. Moritz and Per Kryger. "Self-organization of circadian rhythms in groups of honeybees (Apis mellifera L.)" In: *Behavioral Ecology and Sociobiology* 34.3 (1994), pp. 211–215. ISSN: 1432-0762. DOI: 10.1007/BF00167746.

[28] Guy Bloch. "The social clock of the honeybee". In: *Journal of biological rhythms* 25.5 (2010), pp. 307–317. DOI: 10.1177/0748730410380149.

[29] Roberto Refinetti, Germaine Corné Lissen, and Franz Halberg. "Procedures for numerical analysis of circadian rhythms". In: *Biological rhythm research* 38.4 (2007), pp. 275–325. ISSN: 0929-1016. DOI: 10.1080/09291010600903692.

[30] *Inferential statistical methods for estimating and comparing cosinor parameters*. 1982. URL: https://psycnet.apa.org/record/1983-29398-001.

[31] E Batschelet. "Circular Statistics in Biology". In: *(No Title)* (1981). URL: https://cir.nii.ac.jp/crid/1370285712588263044.

[32] Mardia K V. *Statistics of Directional Data*. 1972.

[33] Halberg F. In: *Int J Chronobiol* 1 (1973), p. 31.

[34] Franz Halberg et al. "Chronobiology's progress. Part I, season's appreciations 2004-2005: time-, frequency-, phase-, variable-, individual-, age- and site-specific chronomics". In: *Journal of Applied Biomedicine* 4.1 (2006), pp. 1–38. ISSN: 1214-0287. DOI: 10.32725/jab.2006.001.

[35] Roberto Refinetti. "Non-stationary time series and the robustness of circadian rhythms". In: *Journal of Theoretical Biology* 227.4 (2004), pp. 571–581. ISSN: 0022-5193. DOI: 10.1016/j.jtbi.2003.11.032.

[36] Hillman D. In: *Prog Clin Biol Res* 341 (1990), p. 417.

[37] *The cellular aspects of biorhythms*. 1967.

[38] Bingham C. In: *Chronobiologia* 9 (1982), p. 397.

[39] Nelson W. In: *Chronobiologia* 6 (1979), p. 305.

[40] Germaine Cornelissen. "Cosinor-based rhythmometry". In: *Theoretical biology & medical modelling* 11 (2014), p. 16. DOI: 10.1186/1742-4682-11-16.

[41] Sanford Weisberg. *Applied linear regression*. 3rd ed. Wiley Series in Probability and Statistics. Hoboken, N.J: Wiley-Interscience, 2005. ISBN: 9780471704089. DOI: 10.1002/0471704091.

[42] S. S. SHAPIRO and M. B. WILK. "An analysis of variance test for normality (complete samples)". In: *Biometrika* 52.3-4 (1965), pp. 591–611. ISSN: 0006-3444. DOI: 10.1093/biomet/52.3-4.591.

[43] Rohlf, F. James, Sokal, Robert R. "Biometry : the principles and practice of statistics in biological research". In: *(No Title)* (). URL: https://cir.nii.ac.jp/crid/1130000795537706496.

[44] M. Fernanda Ceriani et al. "Genome-wide expression analysis in Drosophila reveals genes controlling circadian behavior". In: *Journal of Neuroscience* 22.21 (2002), pp. 9305–9319. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.22-21-09305.2002.

[45] G. Piccione, G. Caola, and R. Refinetti. “Temporal relationships of 21 physiological variables in horse and sheep”. In: *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* 142.4 (2005), pp. 389–396. ISSN: 1095-6433. DOI: `10.1016/j.cbpa.2005.07.019`.

[46] R. Refinetti. “Variability of diurnality in laboratory rodents”. In: *Journal of Comparative Physiology ? A* 192.7 (2006), pp. 701–714. ISSN: 0340-7594. DOI: `10.1007/s00359-006-0093-x`.

[47] Kanji G K. *100 Statistical Tests*. 1993.

[48] Siegel S. *Nonparametric Statistics for the Behavioral Sciences*. 1956.

[49] Roberto Refinetti. “Analysis of the circadian rhythm of body temperature”. In: *Behavior Research Methods, Instruments, & Computers* 24.1 (1992), pp. 28–36. ISSN: 1532-5970. DOI: `10.3758/BF03203466`.

[50] J. T. Enright. “The search for rhythmicity in biological time-series”. In: *Journal of Theoretical Biology* 8.3 (1965), pp. 426–468. ISSN: 0022-5193. DOI: `10.1016/0022-5193(65)90021-4`.

[51] G. J. Dörrscheidt and L. Beck. “Advanced methods for evaluating characteristic parameters (t, a,) of circadian rhythms”. In: *Journal of Mathematical Biology* 2.2 (1975), pp. 107–121. ISSN: 0303-6812. DOI: `10.1007/BF00275923`.

[52] P. G. Sokolove and W. N. Bushell. “The chi square periodogram: its utility for analysis of circadian rhythms”. In: *Journal of Theoretical Biology* 72.1 (1978), pp. 131–160. ISSN: 0022-5193. DOI: `10.1016/0022-5193(78)90022-X`.

[53] Fourier J. *Théorie Analytique de la Chaleur*. 1822.

[54] Halberg F. In: *Lancet* 73 (1953), p. 20.

[55] Koehler F. In: *Environ Med Surg* 14 (1956), p. 5.

[56] Arthur Schuster. “On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena”. In: *Terrestrial Magnetism* 3.1 (1898), pp. 13–41. ISSN: 0272-7528. DOI: `10.1029/TM003i001p00013`.

[57] Fisher R A. *Statistical Methods for Research Workers*. 1925.

[58] Andrew F. Siegel. “Testing for Periodicity in a Time Series”. In: *Journal of the American Statistical Association* 75.370 (1980), pp. 345–348. ISSN: 1537-274X. DOI: `10.1080/01621459.1980.10477474`.

[59] N. R. Lomb. “Least-squares frequency analysis of unequally spaced data”. In: *Astrophysics and Space Science* 39.2 (1976), pp. 447–462. ISSN: 0004-640X. DOI: `10.1007/BF00648343`.

[60] T. Ruf. “The Lomb-Scargle Periodogram in Biological Rhythm Research: Analysis of Incomplete and Unequally Spaced Time-Series”. In: *Biological rhythm research* 30.2 (1999), pp. 178–201. ISSN: 0929-1016. DOI: `10.1076/brhm.30.2.178.1422`.

[61] *On the problem of random vibrations, and of random flights in one, two, or three dimensions*. 1919.

[62] Box G EP. *Time-Series Analysis, Forecasting and Control*. 1970.

[63] Granger C WJ. *Forecasting Economic Time Series*. 1986.

[64] Franziska Boenisch et al. “Tracking All Members of a Honey Bee Colony Over Their Lifetime Using Learned Models of Correspondence”. In: *Frontiers in robotics and AI* 5 (2018), p. 35. DOI: `10.3389/frobt.2018.00035`.

[65] Fernando Wario et al. “Automatic methods for long-term tracking and the detection and decoding of communication dances in honeybees”. In: *Frontiers in Ecology and Evolution* 3 (2015). DOI: `10.3389/fevo.2015.00103`.

[66] Benjamin Gutzmann and Andreas Motl. *wetterdienst*. 2023. DOI: `10.5281/ZENODO.8352276`.

[67] Kyle Barron. *Suncalc*. URL: `https://github.com/kylebarron/suncalc-py` (visited on 2023).

[68] John Salvatier, Thomas V. Wiecki, and Christopher Fonnesbeck. *Probabilistic programming in Python using PyMC3*. PeerJ, 2016. DOI: `10.7287/peerj.preprints.1686v1`.

[69] William Menke. “Chapter 9 - Detecting and understanding correlations among data”. In: *Environmental Data Analysis with MatLab or Python*. Ed. by William Menke. San Diego: Elsevier Science & Technology, 2022, pp. 277–317. ISBN: 978-0-323-95576-8. DOI: `10.1016/B978-0-323-95576-8.00006-4`. URL: `https://www.sciencedirect.com/science/article/pii/B9780323955768000064`.

[70] Lothar Sachs. *Applied statistics: a handbook of techniques*. Springer Science & Business Media, 2012.

[71] Loris Bennett, Bernd Melchers, and Boris Proppe. *Curta: A General-purpose High-Performance Computer at ZEDAT, Freie Universität Berlin*. 2020. DOI: `10.17169/REFUBIUM-26754`.

[72] OpenAI. *ChatGTP Version 3.5*. 2023.