UCL CoMPLEX

Behavioural Influence On Circadian Patterns in *Drosophila Melanogaster*

A Modelling Approach

# Case Presentation 1

*Author:*
Mark Ransley

*Supervisors:*
Dr. Joerg Albert
Prof. Peter Dayan

January 17, 2013

**Abstract**

Experimental evidence through the literature suggests that the locomotive activity of a creature plays a role in entraining the central circadian clock. However, such activity also serves as the main observable output variable through which information about diurnal oscillations may be inferred. By examining data from entrainment experiments on Drosophila Melanogaster we identify several behavioural patterns that occur during photic entrainment and their relationship with later free-running activity. We discuss the relevance of this with respect to a new model encompassing environmentally and endogenously driven locomotion. We then attempt to construct a generative mathematical model involving said effects, and investigate whether such a model could explain the variety of profiles recorded in the lab.

# Contents

# 1 Background

## 1.1 The internal timekeeper

The circadian clock is a well studied system enabling the occurrence of time-dependent biological functions even in the absence of external stimulus. The underlying genetic mechanisms differ little from mammals to insects [Glossop], and essentially involve a transcriptional feedback loop with period $\approx$ 24hours. Whilst certain individual cells and organs contain their own temporal circuitry for localised procedures [Cahill], most biological oscillations are controlled by a neuronal "master clock" that is calibrated through sensory information. Such environmental cues for setting the clock are termed *zeitgeber*. As a consequence of mankind's manipulation of the environment, exposure to non-natural zeitgeber profiles has resulted in a number of clock related diseases such as insomnia, narcolepsy and possibly bipolar disorder [Milhiet]. Using animal and *in silico* models to further probe the nuances of the circadian clock may ultimately allow for more effective and less invasive treatments of such conditions.

## 1.2 Synchronisation

The most extensively studied zeitgeber is light, and there have been shown to exist photoreceptors in the eye specialised for relaying external brightness levels directly to the neuronal clock [Cahill]. Experiments have shown wild type Drosophila can be robustly entrained through only a few days of controlled light/dark cycles, and afterwards will continue to display diurnal activity in complete darkness. Free running circadian rhythm will often drift gradually away from that of entrainment, suggesting an endogenous period slightly off 24h, and illustrating the need for zeitgeber. Similar results have been obtained for temperature, and environmental elements such as social interaction can be seen to configure the clock through more indirect pathways [Golombek].

## 1.3 A novel entrainment pathway

In a paper recently submitted to Science [Albert *et al.*], it was shown that the circadian rhythms of Drosophila can be entrained by vibrating the insect, which

may at first seem surprising since the environment does not naturally exhibit diurnal vibrational cycles as with light and temperature. Comparing with knockout mutants lacking chordotonal organs (ChOs) that could not be entrained by vibration, but adhered to their initial photically trained patterns, it was shown that ChOs front a mechanosensory pathway to the neuronal clock, entirely separate from the one used by light. ChOs are stretch receptors and have previously been linked to temperature entrainment, however they are also responsible for detecting sounds and other airborne vibrations, and providing feedback to the brain of the organism's own muscle movement.

In the same paper, $Per^{01}$ mutants lacking a neuronal clock (but with ChOs) also responded to the vibration, but were not entrained by it, indicating that certain stimuli can interact with the clock, but also may bypass it and affect motion directly. The simple linear model of the circadian system involves stimuli setting the clock, which in turn governs activity. However, it is clear that the relationship is more complicated than that.

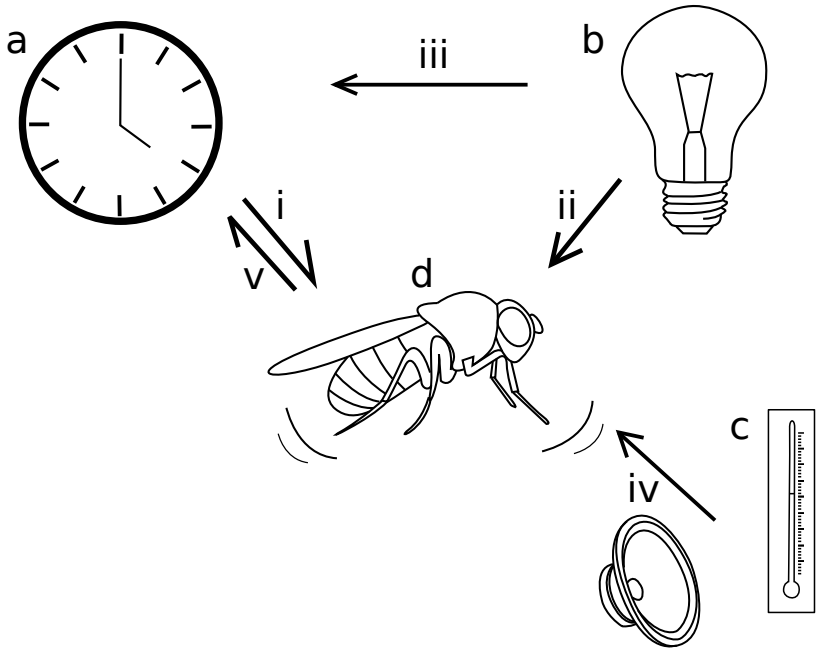



Figure 1: A multidirectional model of the circadian system. See Section 2.1

# 2 Integration of circadian elements

## 2.1 A new model

Figure 1 illustrates the relationships known to exist between (a) the neuronal clock, (b) photic stimuli, known to have a direct pathway to influence (a); (c) non-photic stimuli, which communicate with the (a) through indirect means such as (d), locomotive activity.

Evidence of pathway (i) is seen in experiments during free-run, where the neuronal clock is the only driver of activity. Pathway (ii) is manifest in a phenomena known as *masking* (observable in Appendix), where the sudden presence of light renders the insect paralysed for some time, whilst simultaneously setting the phase of the neuronal clock via pathway (iii). In [Mrosovsky] *et al.* the group administered benzodiazepine triazolam to hamsters and found no effect on circadian rhythm unless specimens were granted access to an exercise wheel, in which case activity phase shifted from that of the wheel-accessing control arm. This implies that the drug interacts with the neuronal clock through affecting the creature's own movement, i.e. pathway (v). We say the drug is a type of non-photic zeitgeber along with temperature and vibration. Qualititative differences between the two types of entrainment can be elucidated through plotting the phase response curve, discussed in [Golombek], though we will not use this in our analysis. Pathway (iv) is distinct from (ii) in that category (c) effectively hijacks (i) in order to communicate with the neuronal clock.

## 2.2 Key objective

During free run, paths (ii), (iii) and (iv) are redundant, leaving only (i), which we know to exist during free run, and (v), for which we know little about in the absence of non-photic stimuli. A mathematical model describing free run behaviour through necessary use of (v) would add weight to arguments supporting the pathway's existence in these conditions, and provide possible quantification of its influence and role in the circadian system.

# 3 Data

We obtained wild-type Canton S Drosophila activity observations from the Stanewsky Lab at Queen Mary, University of London. Following three days of 12-hour photic entrainment cycles (LD), specimens were monitored in complete darkness with external zeitgeber such as temperature held constant (DD) for a further 21 days, $7\frac{1}{2}$ hours. Under these conditions, flies were placed in individual tubes, each equipped with an infrared beam, of a Drosophila Activity Monitor (DAM), where movement was defined as the number of times the fly crossed its beam in a 30 minute interval.

Data was given in the form of a time series vector for each fly, containing the output from the DAM. Attached to each time series was the experiment start time (e.g. 9:30am) and a binary-valued vector indicating whether the light was on during the corresponding interval.

A large number of tools for handling DAM data have been consolidated into Matlab's FlyToolBox. The most useful three for our analysis are pictured in Figure 2.

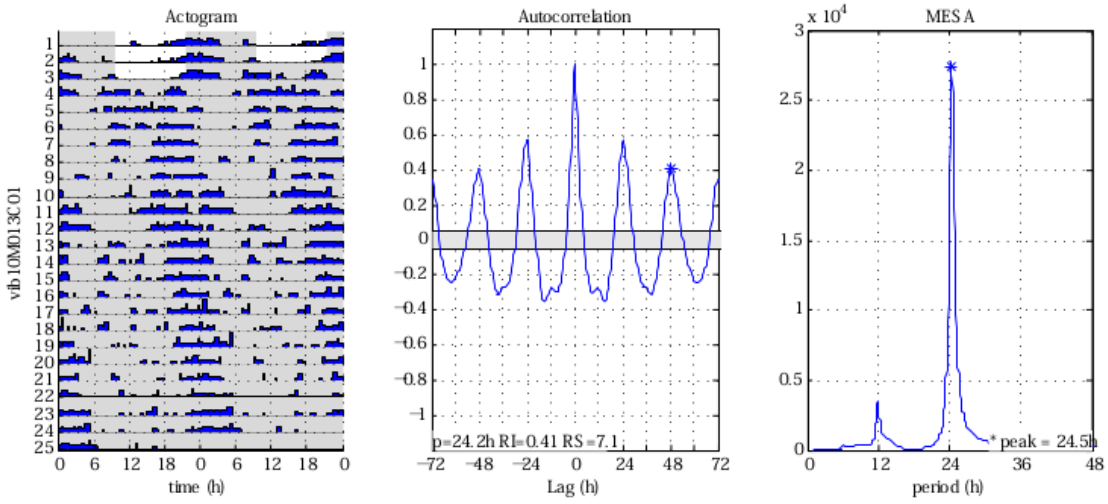



Figure 2: The Actogram, Autocorrelogram and MESA diagrams for one fly.

## 3.1 Analytical tools

### 3.1.1 Actogram

The Actogram displays activity in consecutive 30-minute intervals, i.e. it plots bar charts of data directly from the DAM. Each line represents a day, but runs for 48-hours so that there is overlap with the lines around it. This allows for any complete 24 hour interval to be examined (without having to move to the next line) and also permits vertical day-by-day comparison, rendering qualities such as period more noticeable.

### 3.1.2 Autocorrelogram

The autocorrelogram is a key temporal analysis tool for validating apparent periodicities. The data is smoothed and detrended, and then the correlation coefficient between the treated data and itself at various lags is plotted, so that peaks occur when the data is in phase with itself at a particular lag, i.e. the period. The grey bars represent a correlation of 0.05, weak enough to assume there is no periodicity in data that remains in this region. The star at 48h is termed the Rhythms Index (RI) and is defined as the value of the third peak. It provides a numerical evaluation of long-term rhythmicity. The third peak is chosen as a compromise between long-term assessment and diminution of dataset (data points have to be discarded at the tail ends with each successive lag).

### 3.1.3 Fourier Techniques & Maximum Entropy Spectral Analysis

Once the autocorrelogram has shown there is indeed periodicity to the data, the nature of its oscillations can be studied in more detail using Fourier Analysis. This involves decomposing the detrended function into a sum of $n$ *sine* and *cosine* components with coefficients $\{a_n\}$ and $\{b_n\}$ respectively. This form can be used to approximate any periodic function to arbitrary accuracy, though since the data is finite and discrete there are lower and upper bounds on the frequencies of components. For a time series with $N$ points, frequencies lower then $\frac{1}{N}$ cannot be studied as there is not enough data to observe repeats. Also, oscillations with period $\leq$ 15mins are indistinguishable from higher harmonics due to the sampling rate of 30mins, and clearly oscillations with period 30mins will cancel themselves

out and not appear in the time series. The 15mins lower bound is referred to as the Nyquist frequency.

Fourier methods can be used to transform the time series into a frequency series like the one seen in Figure 2c. The height of each spike provides an indication of that period's prominence in the data, so circadian output typically has significant peaks at around 24h, with shorter wavelengths indicating ultradian oscillations, or possibly just the Fourier decomposition attempting to describe noise. It is even possible that such spectra reveal patterns within the data that are hidden to the human eye.

Since techniques for discrete Fourier transforms are computationally slow, and the so-called Fast Fourier Transform works optimally for a dataset of length $2^k$ for some integer $k$, the technique of Maximum Entropy Spectral Analysis is used instead. MESA calculates the auto covariance of the dataset and uses ideas from information theory to calculate the frequency spectrum of the most random possible time series with equivalent autocovarience. The principles underpinning this method are advanced but the algorithm is described in [Numerical Recipes].

## 3.2 A new visualisation tool

To aid our analysis and render trends more obvious we constructed a function to plot the total activity (TAct) of an individual in a given interval $h$ (Figure 3), where the interval is set at both fixed and moving times. When h is equal to the period of the insects diurnal rhythm, circadian oscillations cancel themselves out, with the moving window plot showing little more than noise about the fixed one (Fig. 3a). For some flies the oscillations persist, due to a period not consistent with the 30-minute intervals or possibly even variable. Diurnal oscillations become pronounced when h is set to half of the circadian period (Fig. 3b).

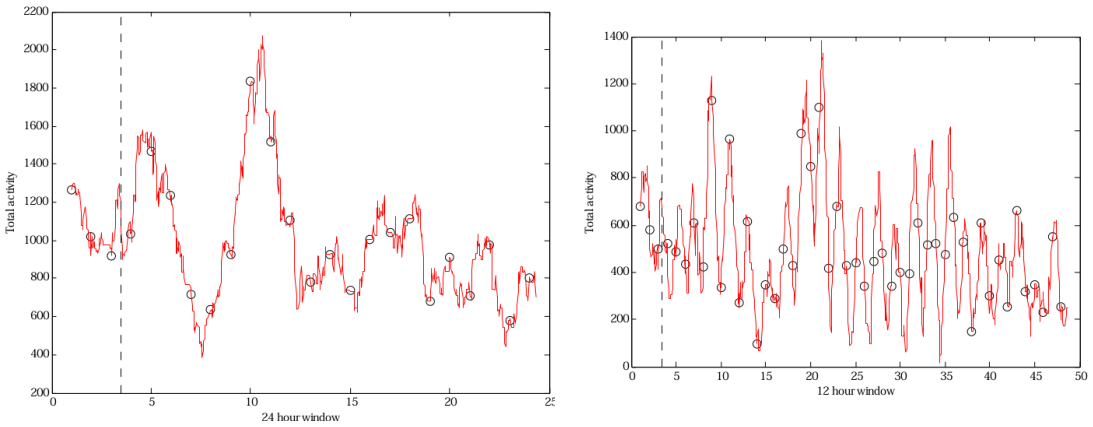



Figure 3: Total activity plots per h-hours (black) and in a moving h-hour window (red) for h = 24 (left) and h = 12 (right).

# 4 Analysis

In Appendix 1 we perform a qualitative data analysis on 16 individual flies taken from Stanewsky's data (Experiment vib10M13[1:16]) in an attempt to identify recurrent themes, anomalous profiles, possible causes or indicators of activity patterns and elements that should be considered when modelling circadian output. Tools used for this investigation were individual actograms, autocorrelograms (for free run), MESA plots (for free run) and Tact plots with dashed black lines indicating start of free run, and lines of best fit for free run. Since Specimen 6 died in the second day of the experiment, it has been excluded from further population analyses. A few hypotheses were formed from observations:

1. Higher total activity during entrainment appears to indicate higher activity in the free run.

2. Specimens with more noise during entrainment night appear to have more noise during the free run "night".

3. The masking effect described in Section 2.1 is apparent in many cycles, with very little activity in the first 6 hours of each day of entrainment, but a surge in activity during the second half of each entrainment day that tends to bleed over into the night. However, following entrainment the insects immediately

display motion synchronous with the previous daytime, illustrating entrainment despite the effects of masking. This will not be modelled as we are only interested in the free run. However, there appears to be a link between masking and free run activity - those with a higher resistivity to masking appear to age faster.

Three other features of the data should be remarked upon, and will be considered when constructing the model:

1. A decrease in overall activity is apparent in the actograms and Tact plots of the majority of flies (very apparent in $\frac{9}{15}$ specimens).

2. For some insects (very clear for Specimens 2 and 5, possible also for Specimens 4, 9 and 14) free running period begins $> 24$h, but shifts following a day of decreased activity to $< 24$h. We speculate that such behaviour could be the result of propriosensitive feedback.

3. There is much variation throughout the population in the length of the free running "day", and for some specimens this length appears to change gradually.

Since we are only concerned with modelling free runs, we can use data from entrainment to infer parameter values for the free run based on Hypotheses 1, 2 & 3 - provided they are valid hypotheses.

## 4.1 Entrainment activity as indicator of free run activity

To test Hypothesis 1 we plotted the average total activity per day during entrainment against the average total activity per day during free run (Figure 4) and observed a strong positive correlation (correlation coefficient 0.753). Consequently, the total activity during entrainment will be used to guide a magnitude parameter in the model.

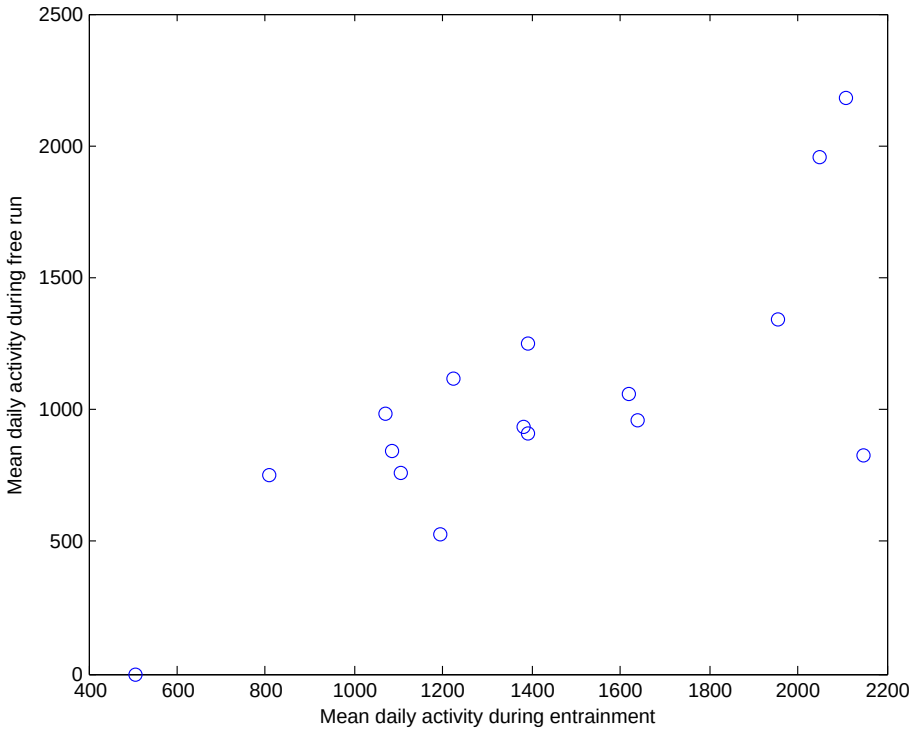



Figure 4: Comparison of mean activity per day during entrainment against mean activity per day in free run, correlation coefficient = 0.753

## 4.2 Noise

To test Hypothesis 2 we defined the "noise window" between the hours of 5-9am, due to this being an interval when all specimens (except for Specimen 1) could be considered asleep on every day of the experiment. We then compared the average daily activities in this window during entrainment with those in free run (Figure 5) and found the correlation coefficient to be 0.791, confirming that the level of noise during entrainment does indeed indicate the level of noise within free run. Due to the high period of Specimen 1 causing phase-shift type effects, this specimen was discarded from the noise-window analysis.

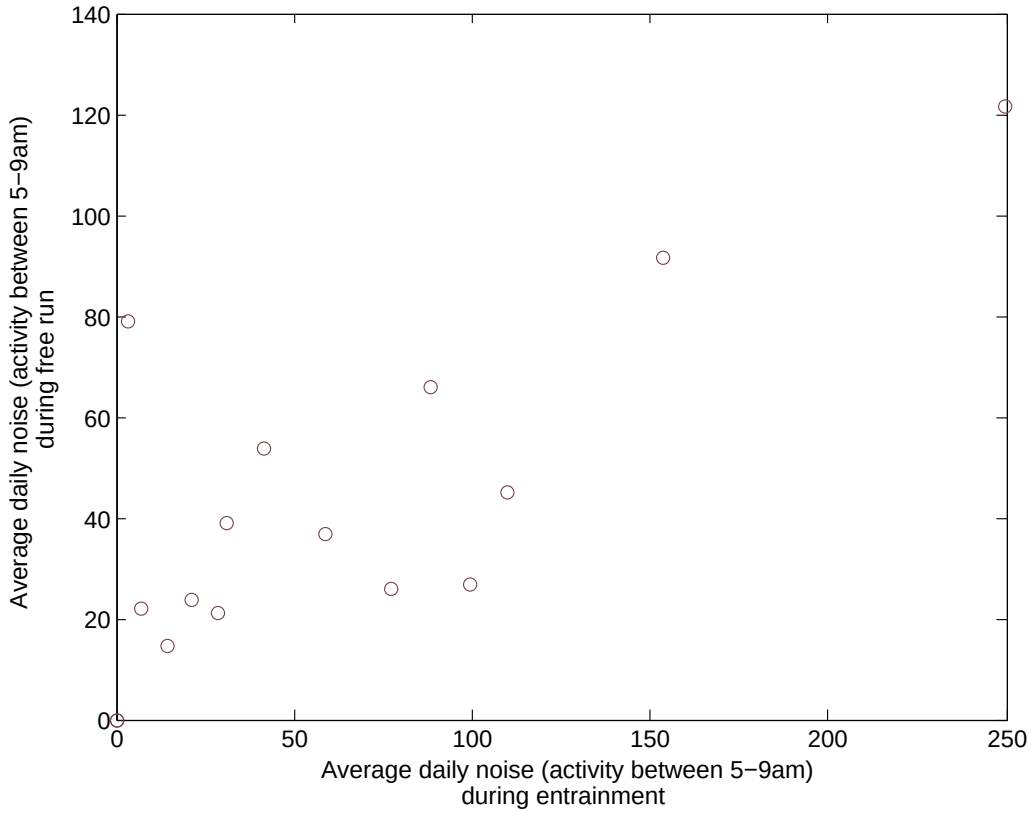



Figure 5: Comparison of average daily activity in the noise window during entrainment and free run. Correlation coefficient = 0.791

## 4.3 Ageing

To test Hypothesis 3, the proportion of activity during light entrainment that occurred during the first six hours of daily light was plotted against the gradient of the line of best fit for free running TAct values (Figure 6). The result shows effectively no correlation, with the invalidated hypothesis most likely being formed due to an observation bias caused by the outlier at $Proportion = 0.55$ (Specimen 8).

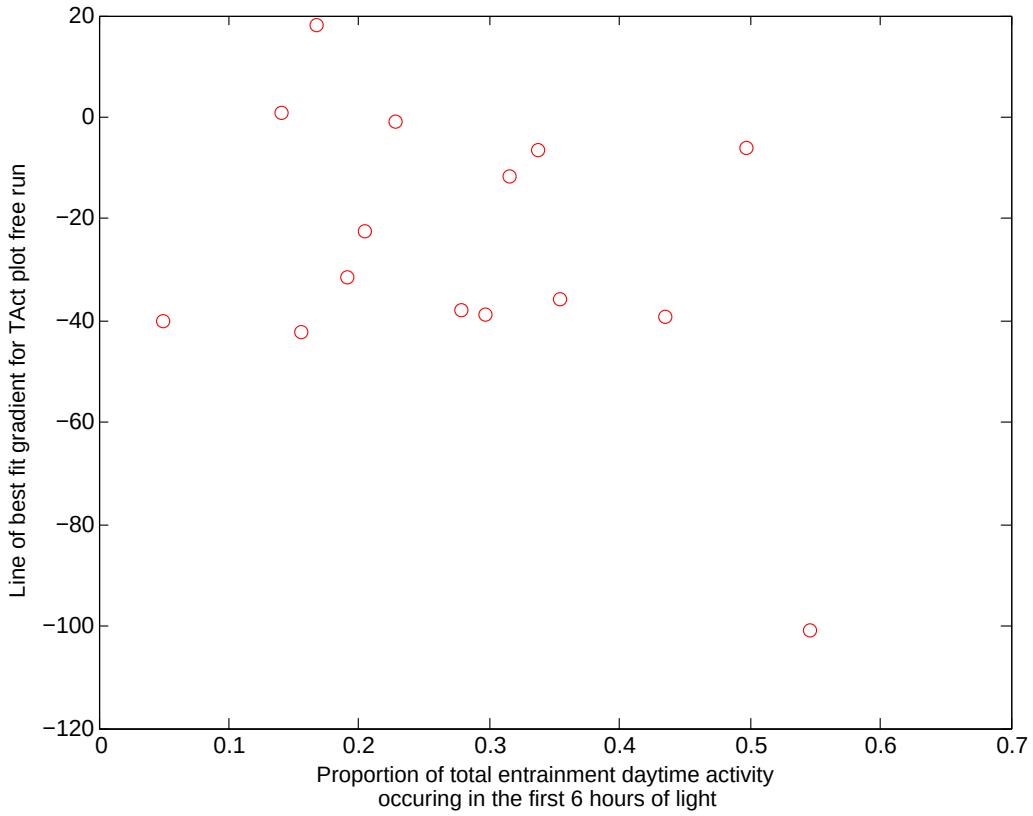



Figure 6: Comparison of resistance to masking and rate of ageing. Correlation coefficient of -0.004 after removal of outlier at 0.55 invalidates Hypothesis 3. Note that with one exception all gradients are $\leq 0$, corroborating Remark 1.

# 5 Modelling

To generate data consistent with experimental output, activity during interval t ($Act_t$) was modelled by a Poisson distribution with variable mean $\lambda_t$. To represent the cyclic nature of the activity, $\lambda_t$ was defined by a sinusoidal function with period $\approx$ 24hours, and phase $\omega_t$. To map $\lambda_t$ out of the range $[0, 1]$, the sin term was given coefficient $\tau$ to represent an average activity level per interval. Negative values of the oscillation were set to zero, and in line with the observation (Section 4 Remark 3) regarding the variability in length of endogenous "daytime", parameter $s$ was added to shift more or less of the sin wave above zero. To permit changes in period

(Remark 2), $\dot{\omega}_t$ was modulated by parameter $\gamma$. To investigate the idea that the insect's activity influences $\omega_t$, we made $\gamma$ a function of a weighted average $\phi$ of the previous twelve values of $Act_t$ with respect to a constant $\mu$, and gave this term coefficient $\alpha$. The decrease in activity due to ageing was modelled by coupling $Act_t$ to an exponential decay term, with exponent $A$. Subsequently the answer had to be rounded to produce discrete output. A further stochastic term $N$ was added to account for noise during the "night" and add further noise during the "day" (without this double stochasticity, activity generated by the model was still very sinusoidal in shape, whereas the real data appears more random). The noise term will be discussed in more detail below. Thus we created the following model:

$$
\begin{aligned}
Act_t &= e^{-At}\left(Pois(\lambda_t) + N\right) && (1)\\
\lambda_t &= max(\tau \sin(\omega_t) + s, 0) && (2)\\
w_t &= t\gamma_t \frac{2\pi}{48} && (3)\\
\gamma_t &= 1 - \alpha(\phi(Act_{t-12:t-1}) - \mu) && (4)
\end{aligned}
$$

The noise term $N$ is sampled from a Gaussian distribution $\mathcal{N}(0, 1)$ which is multiplied by parameter $\delta$, representing the mean activity per interval and rounded to the nearest integer, since $Act_t \in \mathbb{Z}$ . Any values $< -0.5$ are output as zero. Such a distribution was chosen (over another Poisson, for example) because of its tendency towards zero, no matter how high the remaining values are scaled to by $\delta$.

The effect of $\alpha$ and thus the activity-period coupling mechanism is shown in Figure 7

where

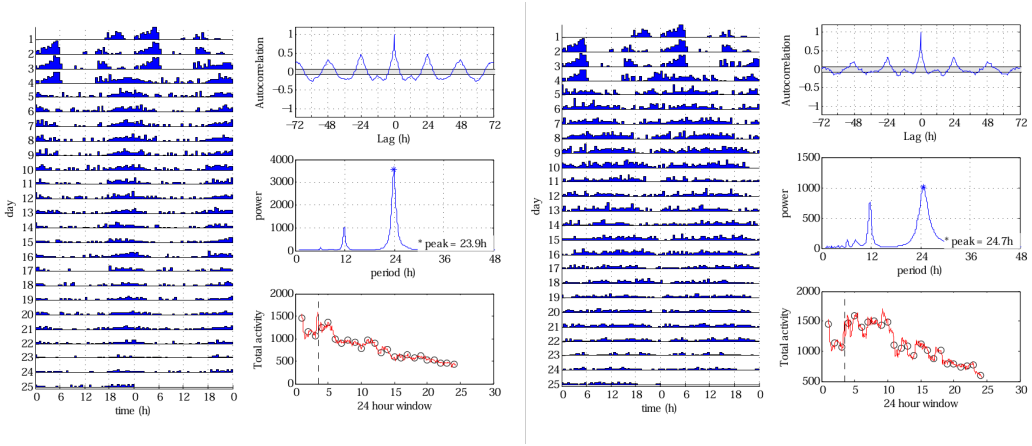



Figure 7: Two flies generated by the model, both with parameters $\tau = 50$, $A = 0.001$, $\mu = 1$, $\delta = 30$, $s = 0$ and starting values borrowed from Specimen 5. The fly on the left has $\alpha = 0$ whilst the one the right has $\alpha = 0.0015$

When fitting the model to a data profile, we start at 24 hours after commencement of the final entrainment ($t = 145$) so as to avoid any of the masking effects that the model does not account for. The first 144 values are taken from the data being fitted to, which provides the model with values for $Act_{t-12:t-1}$ in the first 12 generations.

# 6 Estimation of parameters

We attempted to use Maximum Likelihood Estimation to obtain parameter values for each specimen. Let $\mathbf{x}$ be the time series we are trying to fit, and $\mathbf{X}$ be the corresponding simulated model output.

To reduce the number of parameters used in the optimisation algorithm, we used Sections 4.1-4.2 to find values for $\tau$ and $\delta$. Fitting a line-of-best-fit to Figure 4 we were able to measure each fly's mean daily activity during entrainment and obtain an approximate value for mean daily activity during free-run, and subsequently $\tau$, the average activity per free-running interval. An analogous procedure was applied to obtain a value for $\delta$. The remaining parameters to be fitted are $\theta = (\alpha, \mu, s, A)$.

There are n=1023 values for which the model has to fit, since the first 144 are considered entrainment. Working backwards from the final value, each simulated term in the time series is dependent on the one that preceded it, until we arrive at the first one, created from boundary conditions (the final 12 entrainment values). Hence the *likelihood* of generating $\mathbf{x}$ under $\theta$ is

$$p_\theta(\mathbf{X} = \mathbf{x}) = \left( \prod_{n=1023}^{2} p_\theta(X_n = x_n | X_{n-1} = x_{n-1}) \right) p_\theta(X_1 = x_1) \tag{5}$$

With such a large time series, this can generate some extremely small values, so instead we work with the *log likelihood*:

$$\log p_\theta(\mathbf{X} = \mathbf{x}) = \sum_{n=1023}^{2} \log \left( p_\theta(X_n = x_n | X_{n-1} = x_{n-1}) \right) + \log p_\theta(X_1 = x_1) \tag{6}$$

which instead produces large negative values. By maximising (6) for $\theta$ we find the parameters most likely to produce the desired time series.

To calculate the likelihood function requires working with the probability mass function for the Poisson component:

$$p_\lambda(Z_1 = z_1) = \frac{\lambda^{z_1} e^{-\lambda}}{z_1!} \tag{7}$$

and the cumulative density function of the half-Gaussian distribution described in the previous section. Since the output from the half-Gaussian is rounded, the CDF is

$$p(Z_2 = z_2) = \frac{1}{\sqrt{2\pi}} \int_{\frac{z_2 - \frac{1}{2}}{\delta}}^{\frac{z_2 + \frac{1}{2}}{\delta}} e^{-\frac{t^2}{2}} dt \tag{8}$$

From (1) we get

$$e^{At} Act_t = Pois(\lambda_t) + N \tag{9}$$

Clearly here there is a problem, since the term on the left is not an integer, yet the term on the right is, so the probability of generating the LHS is zero unless $A = 0$.

Due to time constraints, a complete re-examination of the model is not possible. The decay term is essential for getting a fit to the data, since ageing is one of its most noticeable features. Perhaps if this work is resumed in the future the ageing term could be incorporated into (2).

# 7 Discussion

In spite of the parameters not being fitted for each individual, this project has still uncovered some new insights into the nature of circadian rhythms in Drosophila, and how to model (or not model) it. The data itself was examined in ways previously unseen in the literature, and the model constructed was able to display many of the features observed.

On reflection, the sine function was probably not the best one to use for the periodic component; a periodic square function may have been a better choice since the data rarely appears sinusoidal. Also the distribution of data should be further investigated. We took data from one of the vibrational experiments and examined data during the longer periods of vibrational entrainment, where behaviour followed a more robust 24hr pattern. Estimating the mean and variance of the data in each time interval of daily entrainment across all the days when it occurred revealed that whilst sometimes the mean and variance were close, indicating a Poisson distribution, this was not true in the majority of cases. An experiment where robust light:dark cycles were used for the entire 24 days would allow for the distributions, and the rate at which they change with every time-step, to be better understood.

Models exploring different ways of coupling activity to the other parameters, for example $s$, should be constructed in future work, since we have only considered one possible motion-feedback relationship, and there are almost certainly more.

# References


[Albert *et al.*] Albert, Joerg. Stanewski, Ralf. Wolfgang, Werner, Simoni, Alekos. A mechanosensory input pathway to the Drosophila circadian clock - Science: submitted

[Mrosovsky] Mrosovsky, N. Salmon, PA. Triazolam and phase-shifting acceleration re-evaluated - Chronobiology International [1990, 7(1):35-41]

[Milhiet] Milhiet V, Etain B, Boudebesse C, Bellivier F. Circadian biomarkers, circadian genes and bipolar disorders - J Physiol Paris. 2011 Dec;105(4-6):183-9. doi: 10.1016

[Golombek] Golombek, Diego. Rosenstein, Ruth. Physiology of circadian entrainment - Physiol Rev 90: 10631102, 2010; doi:10.1152/physrev.00009.2009.

[Cahill] Cahill, Gregory. Hasegawa, Minoru. Circadian oscillators in vertebrate retinal photoreceptor cells - Biol. Signals 1997

[Glossop] Glossop, Nicholas. Circadian timekeeping in Drosophila Melanogaster and Mus Musculus - Biochemical Society Essays Biochem. 2011 49, 19-35; doi:10.1042/BSE0490049

[Hall] Hall, Jeffrey. Dowse, Harold. Funes, Pablo. Levine, Joe. Signal analysis of behavioral and molecular cycles - BMC Neuroscience: 2002

[Ashworth] The role of behaviour in temperature entrainment of the Drosophila circadian clock - UCL CoMPLEX Summer Project: 2012

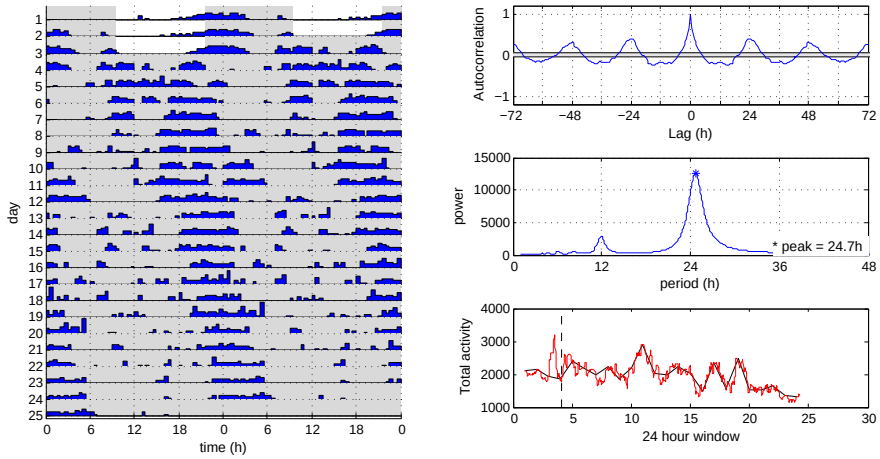
[Bushey] Bushey, Daniel. Hughes, Kimberly. Tononi, Giulio. Cirelli, Chiara. Sleep, aging, and lifespan in Drosophila - BMC Neuroscience 2010, 11:56 doi:10.1186/1471-2202-11-56

[Dowse] Dowse, HB. Ringo, JM Do Ultradian Oscillators Underlie the Circadian Clock in Drosophila? - Ultradian Rhythms in Life Processes 1992, pp 105-117

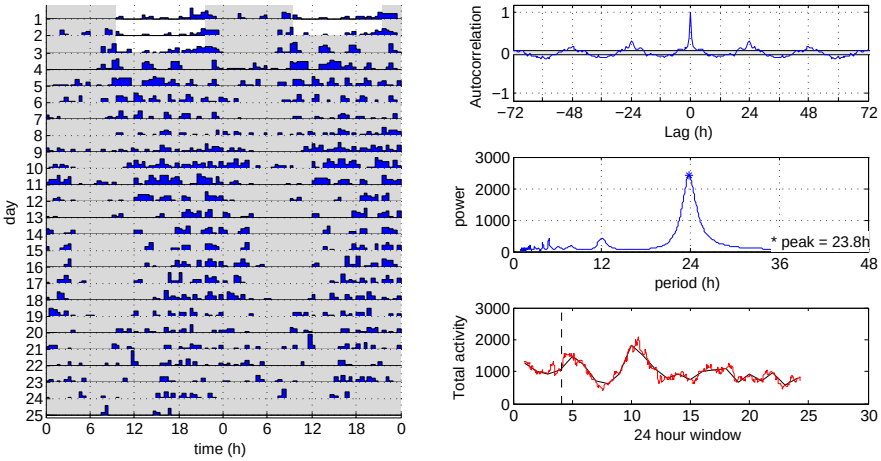
[Numerical Recipes] Numerical Recipes: The Art of Scientific Computing Third Edition (2007), 1256 pp. Cambridge University Press ISBN-10: 0521880688

# Appendices

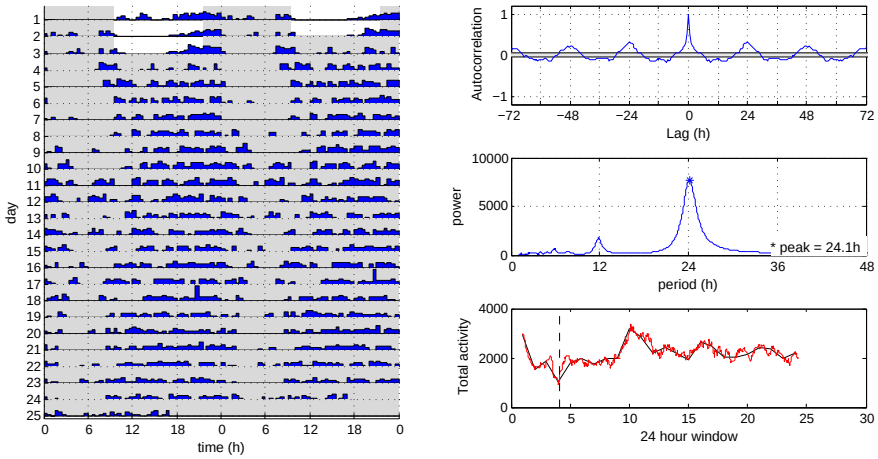
## A Qualitative analysis of individual activity profiles



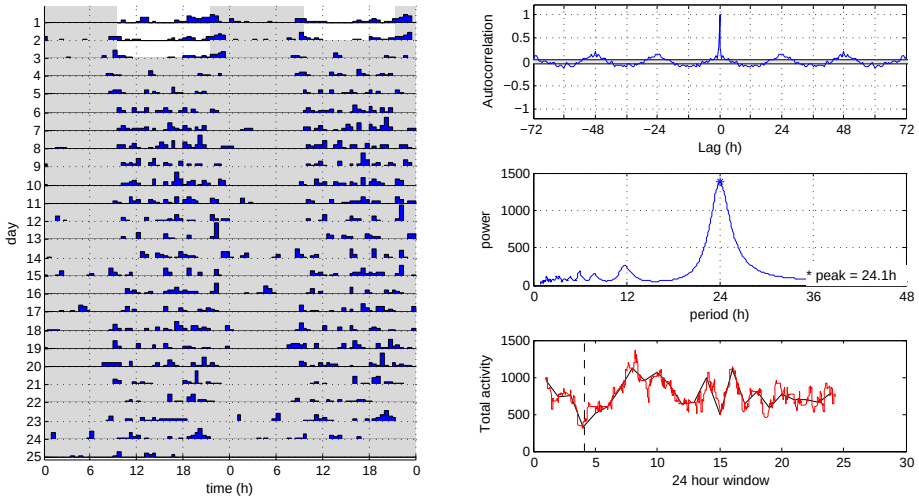

Specimen 1: Well defined activity rest intervals with period > 24h, high total activity levels with a clear decrease over time. Both duration and magnitude of activity decrease. Some noise during rest.



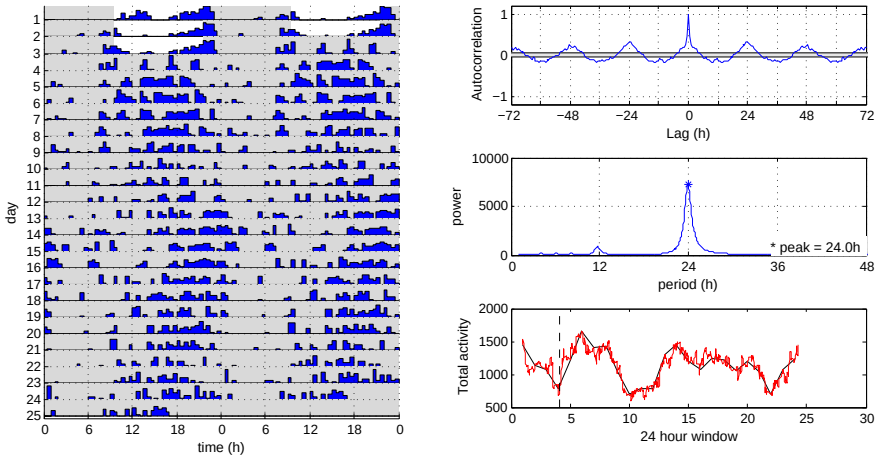

Specimen 2: Sparse activity during daytime contained in subsets of activity. Period seems to change from >24h to <24h following a day of low but consistent activity, as evidenced by strong MESA peak and weak autocorrelogram.



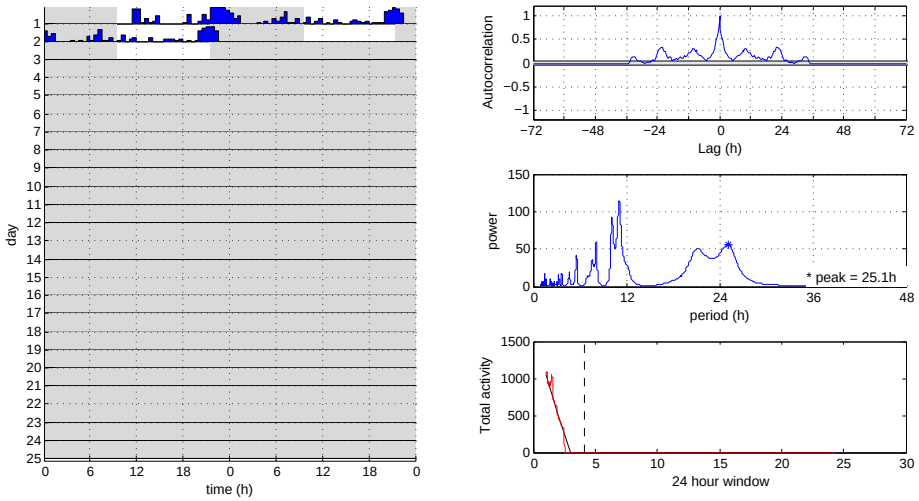

Specimen 3: Settles to a 24h pattern of relatively high activity, with longer days than nights.



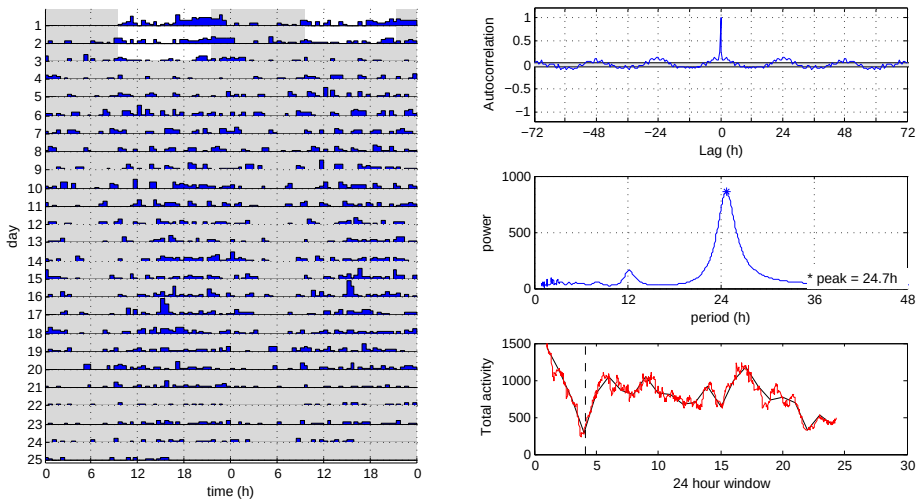

Specimen 4: Low and sparse activity, especially following entrainment, though 24h cycles continue throughout with little activity during the nights. Noise is comparable in magnitude to daytime movement.



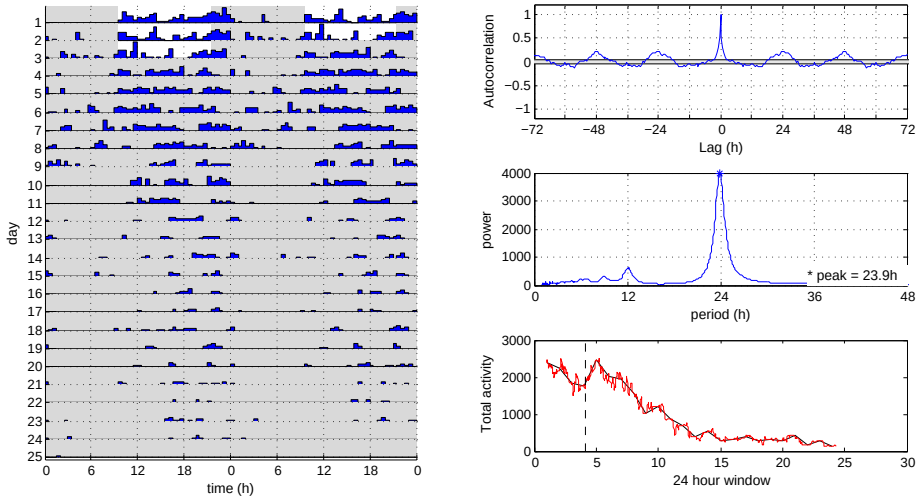

Specimen 5: Again period seems to change from >24h to <24h, with a few days of lower activity in between, and a hint of a second period reversal or phase shift at day 24.



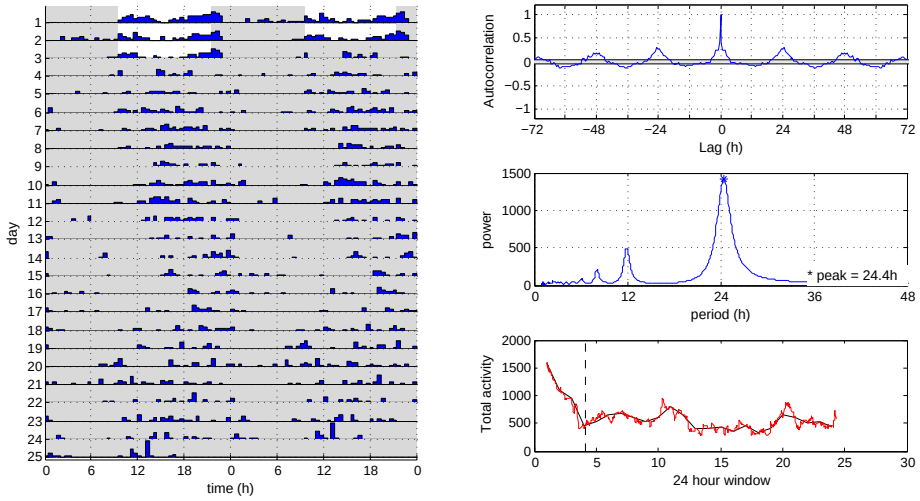

Specimen 6: Did not survive entrainment.



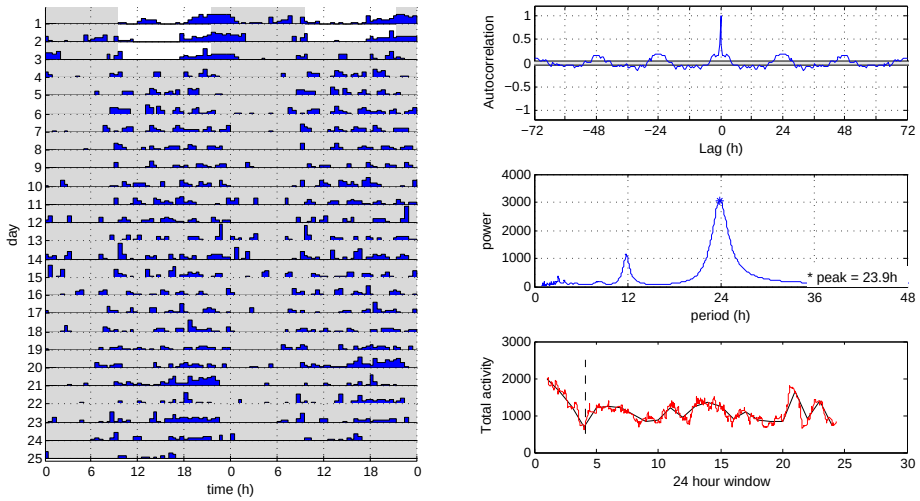

Specimen 7: Activity more evenly distributed over 24h, though there are comparatively short periods of lower activity corresponding to 24h cycles.



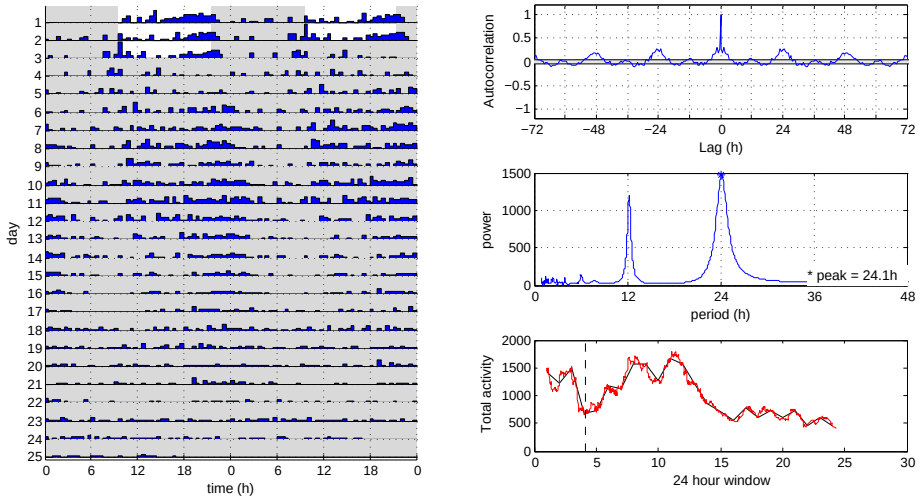

Specimen 8: Pronounced decrease in activity throughout free-run. Some noise but still clear 24h patterns, as evidenced by large MESA peak.



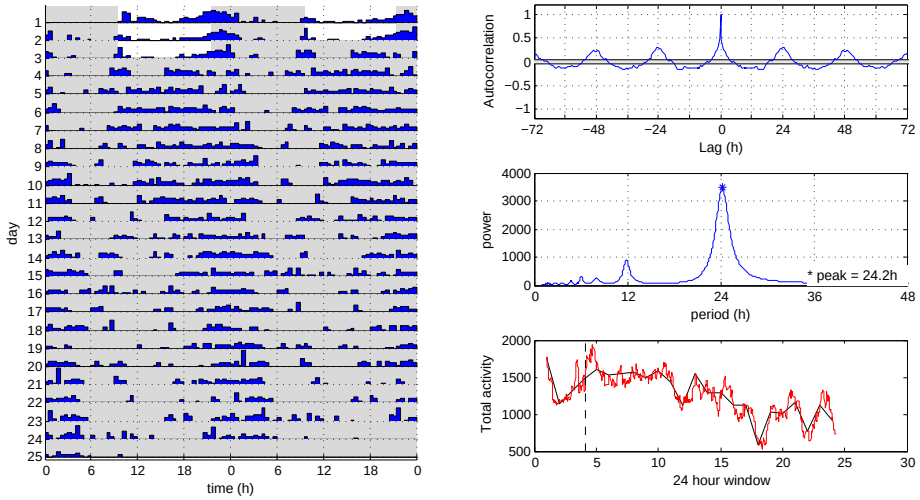

Specimen 9: Low activity following entrainment, and a possible period reversal or continued phase shift around day 15. Day and night remain approximately equal in duration.



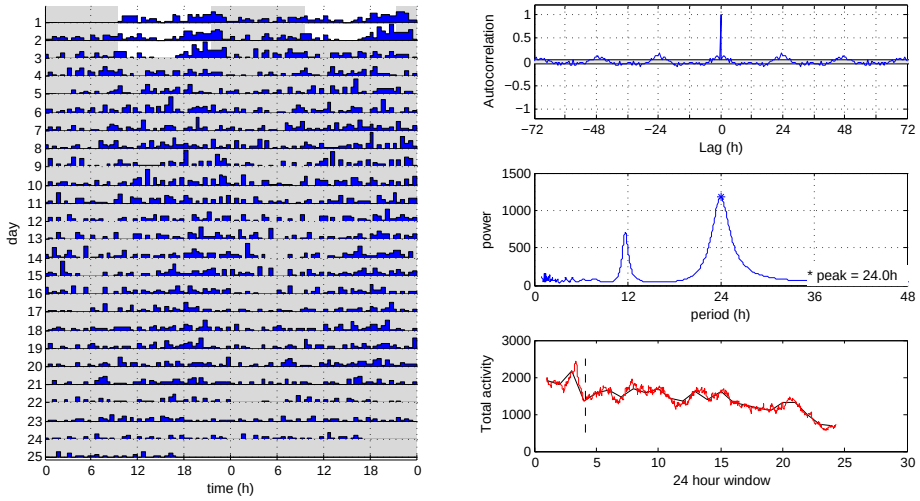

Specimen 10: Strong 24h cycle, with lots of rests within each day.



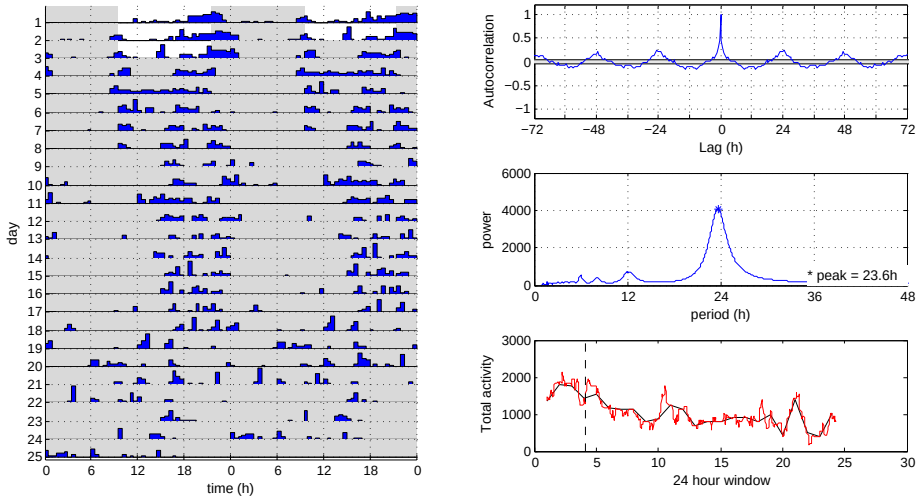

Specimen 11: Low activity following entrainment, followed by a rise and then a gradual loss of circadian rhythm from day 12. Secondary MESA peak at 12h.



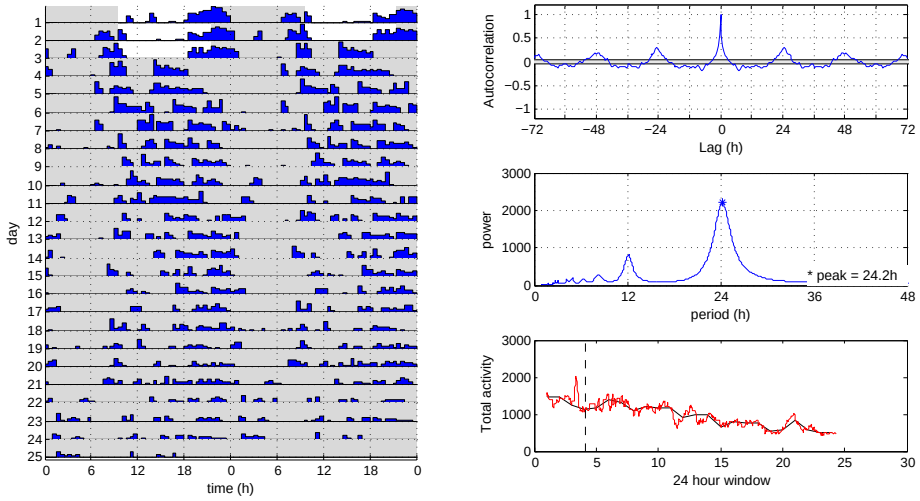

Specimen 12: Decrease in activity as a result of reduced day duration whilst adhering to 24h cycles. Days seem to start later but end around the same time and activity is fairly consistent throughout the day.



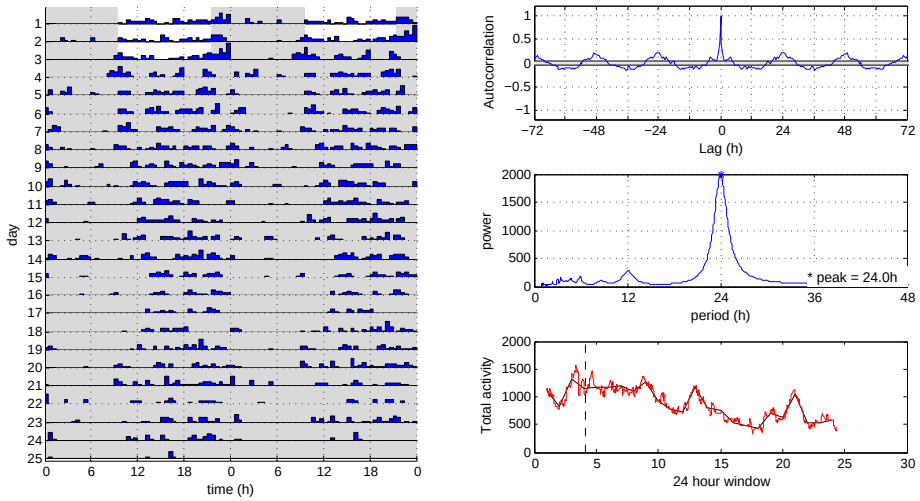

Specimen 13: Very noisy signal with decrease in activity. Increased activity during daytime but weak autocorrelogram and MESA readouts.




Specimen 14: Unraveling of circadian rhythm from day 19. Previously clear day and night cycles with little noise.




Specimen 15: Consistent decrease in activity. Clear 24h diurnal activity though days appear to get longer as activity decreases. Periods of wake and rest throughout the day.




Specimen 16: Sparse activity with general decrease, but clear 24h cycles of roughly equal day and night intervals.