

# Modelling with Matlab Assignment 5

Phil Ledwith

July 13, 2017

## 1 Question 1

I've used the same MATLAB files for questions 1 and 2. The files are Final\_Q1.m and Final\_Q1b.m and have been submitted with this pdf. Plots for all nine measured variables are given below, grouped together following the example of the authors in the paper. For this first question we were not specifically asked to simulate or plot the expression profile, so technically Final\_Q1.m should suffice. I've created the plots anyway, essentially reproducing the answer to Q2(i), because it was necessary in order to check that the equations had been entered faithfully and were performing as expected.

A few short notes on the supplementary material and how the code and graphs were constructed:

In eqn (3) the term  $[P]_p$  appears. This was assumed to be a typo: nowhere in the main paper or the supplementary material does it define or explain what  $[P]_p$  is supposed to represent. I have therefore assumed this was supposed to be written  $[P] = y(9)$  in my code. Similarly in eqn (9) we see a  $P$  which is not in square brackets, and for similar reasons I have assumed this was supposed to be written  $[P]$  and in my code it is represented by the variable  $y(9)$ . Anticipating question (2) of this assignment I checked both of these assumptions question 2 first, and reproducing not only figure 2(A) as required but figures 2(B), 2(C), and 2(D) as well. All were reproduced exactly, suggesting either that these choices were appropriate or that the model is not affected by two significant changes of variables, which would make it a poor model.

In both question(1) and question(2) I have avoided the problem of transients by first running the simulation for 20 days, long enough that the values of (t,y) at the end of each full period (day) were the same for several days in a row, and then used these end of day values as the initial conditions for the appropriate simulation. If you wanted to see the transients it is easy to replace these initial conditions with a zero vector or a random vector as has been done in several other assignments, but I tried both and little information is gained in this case with either approach.

for question 1 (and the figures below) only, the data has not been normalized; for question 2 we followed the example of the authors and normalized protein expression levels to their respective maximums. There are commented lines in the code to allow for either approach.

The first figure asked for is the expression profile for the "dawn genes" *CCA1* and *LHY*, which are combined together in the variable  $[CL]$  in the paper.

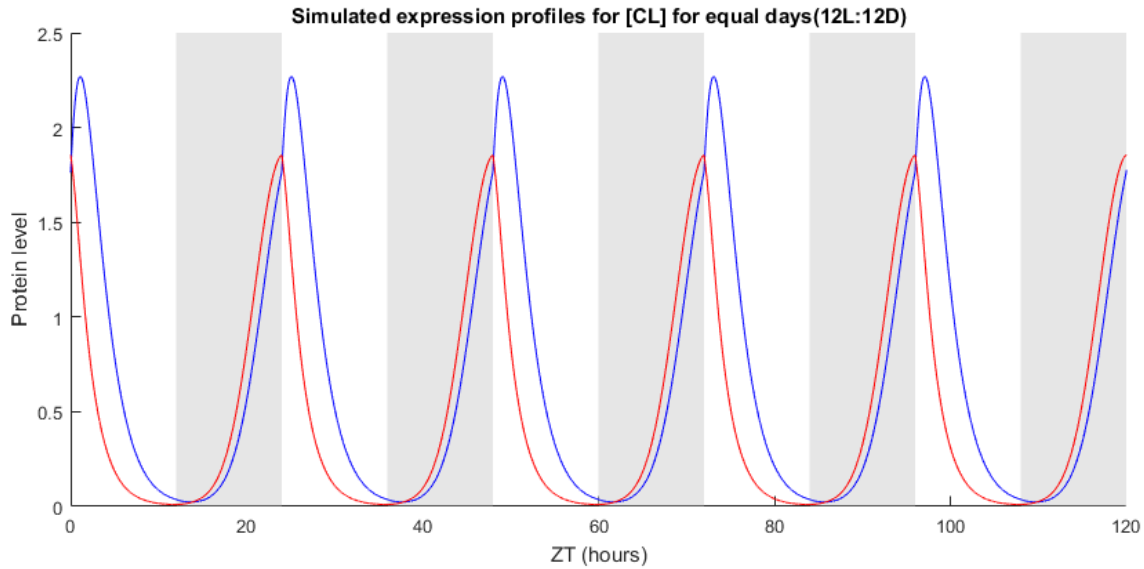


Figure 1: Expression levels for  $[CL]_m$  (red) and  $[CL]_p$  (blue). Grey bars represent dark periods.

The second figure is for the pair of clock genes *PRR9* and *PRR7*, which the authors combine into the variable  $[P97]$ . These are transcriptional repressors, and were also the most sensitive to initial conditions during setup.

In the original paper the figures are all normalized for maximum protein expression. I also did this when I reproduced figure 2 for Question 2 of the assignment, but I have not done so here so that we can more easily see the relative abundance of each protein expressed; e.g. the noticeable difference between  $[P97]_m$  and  $[P97]_p$  as seen in the plots. This becomes more useful later in Question3 when we consider absolute expression in the context of the robustness of the clock.

The third figure is for the equations for  $[P51]$ , which combines the actions of the transcriptional repressors *PRR5* and *TOC1*.

Figures 4 and 5 are for expression of  $[EL]$  and the dark accumulating protein  $P$ ; see section 2.2 of the paper. Again notice the contrast between  $[EL]_m$  and  $[EL]_p$  expression in the plot.

Turns out I need to put some filler of some kind here after all so that the text for question 2 is forced below the

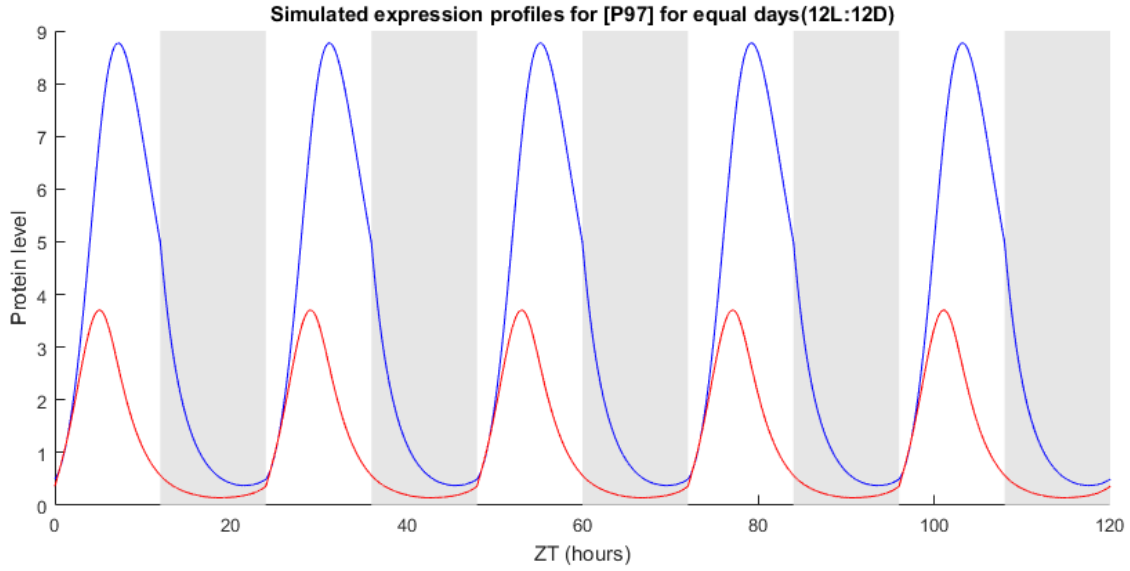


Figure 2: Expression levels for  $[P97]_m$  (red) and  $[P97]_p$  (blue). Grey bars represent dark periods.

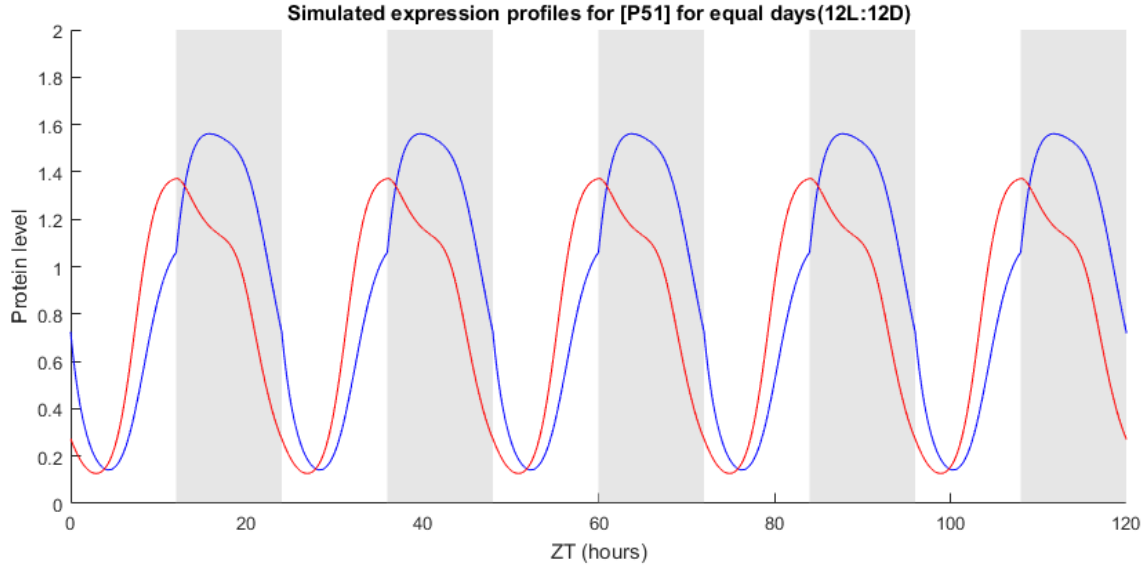


Figure 3: Expression levels for  $[P51]_m$  (red) and  $[P51]_p$  (blue). Grey bars represent dark periods.

images for question 1. blah blah blah blah blah blah.

## 2 Question 2

As I already mentioned, I essentially did question2(i) already in order to check that the code for question(1) was working as expected, given the presence of possible typos in the supplementary material. The same files (Final\_Q1.m and Final\_Q1b.m) were used. All plots were be of  $[x]_m$ , where  $x$  is the desired quantity, i.e. I used the red lines from the graphs in question(1).

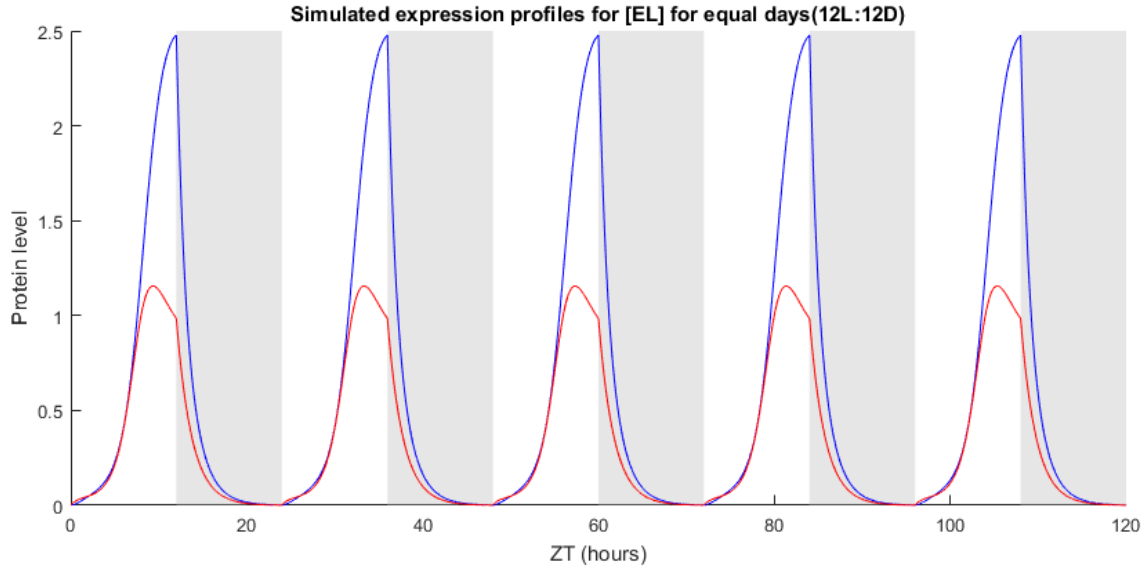


Figure 4: Expression levels for  $[EL]_m$  (red) and  $[EL]_p$  (blue). Grey bars represent dark periods.

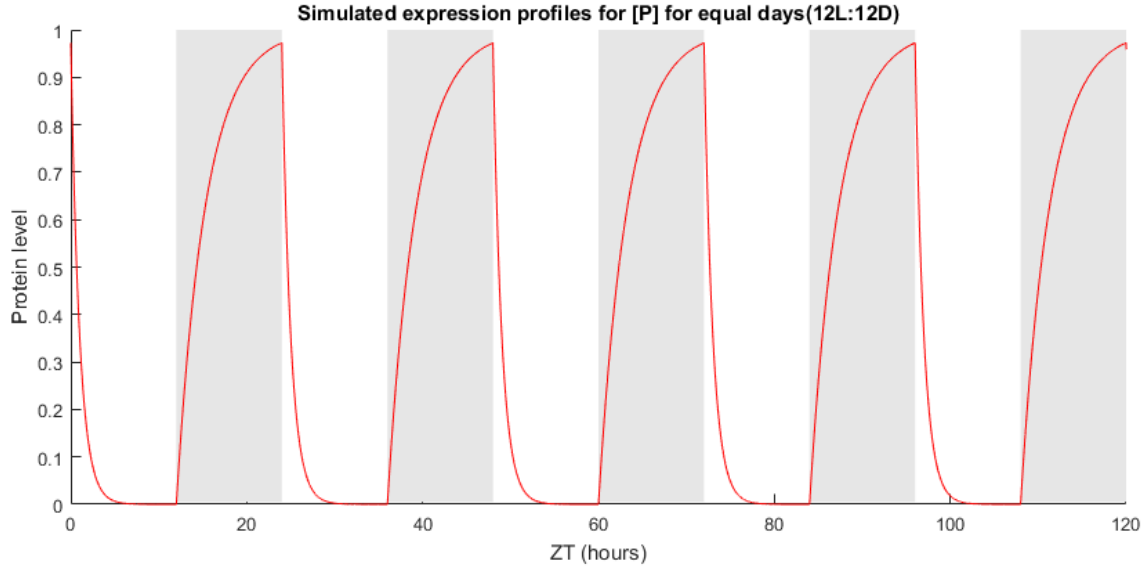


Figure 5: Expression levels for  $[P]$ . Grey bars represent dark periods.

In the legend for the figure the authors claim that "all values are normalized to their respective maximum". In order to reproduce this, as in question (1) I first ran the simulation for 20 days to obtain steady state maximum values. line 41 of the file Final\_Q1b.m

```
disp(max(y(:, :)));
```

displays in the command window a vector of the maximum values for each period of the simulation. 20 days was long enough that the values in the final six rows (representing one dark period and one light period for 3 consecutive days) were all repeats, and then for each column of **max(y(:, :))** the larger of the last two values was chosen as the entry for the variable **y<sub>max</sub>**. This variable was placed in the parameter data before the start of the simulation; since we are not using random start values, the runs are deterministic (this was confirmed just in

case by performing several repeated simulations). Along with the initial values for a simulation (determined in a similar manner as described in question(1)) This creates a parameter set for each scenario (see lines 24-27 and 29-32 of Final\_Q1b.m). We begin with the plots requested in questions 2(i) and 2(ii):

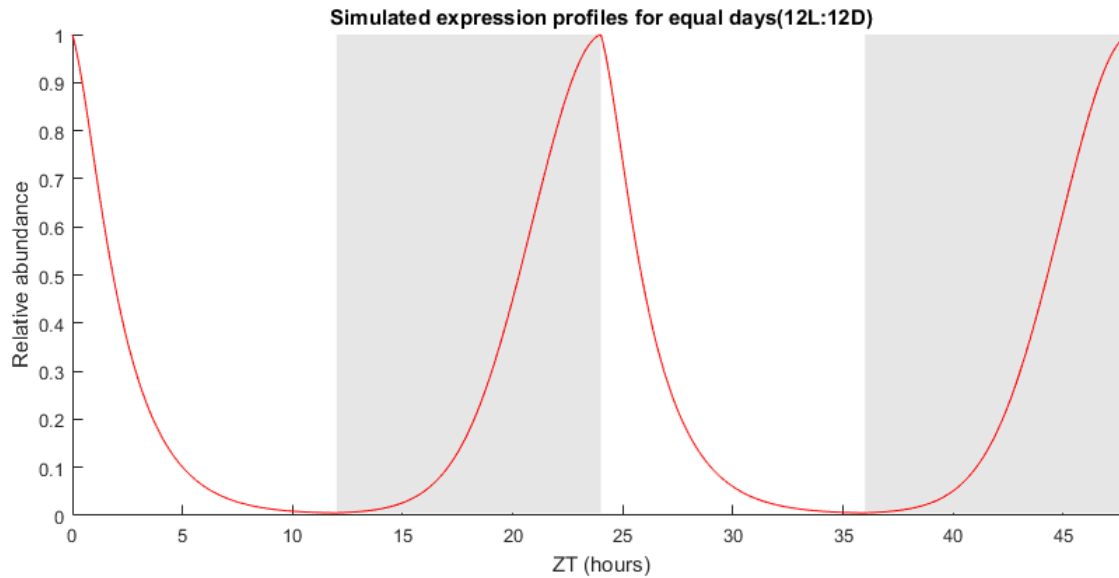


Figure 6: Normalised expression profile of *CCA1/LHY* in an equal dark/light cycle

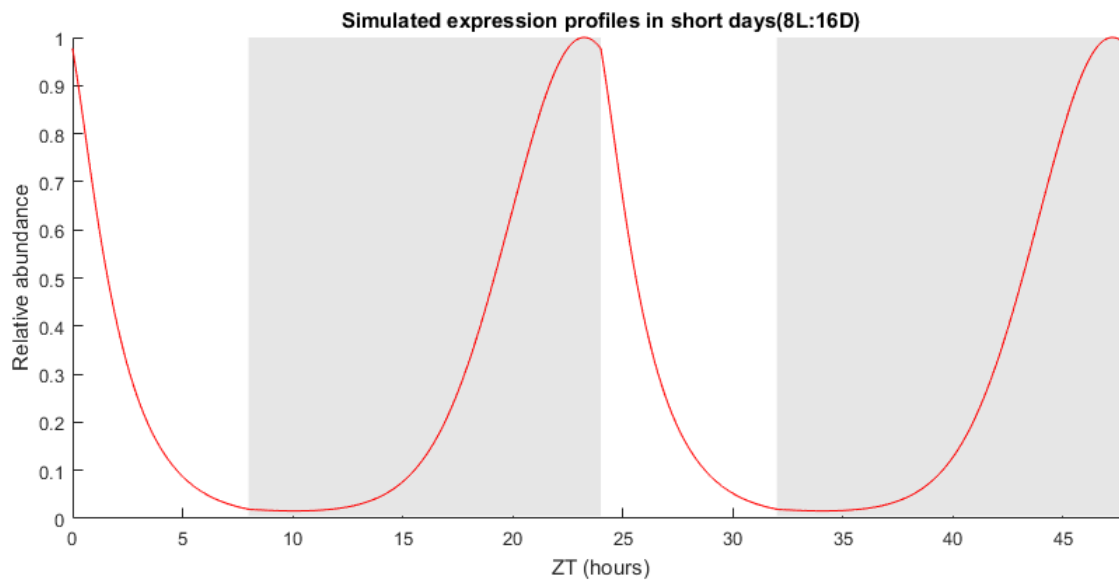


Figure 7: Normalised expression profile of *CCA1/LHY* in a short day cycle (8L:16D)

Although it was not asked for, the entirety of Figure(2) from the paper was reproduced anyway (once again, in order to confirm that the simulation was working as it was supposed to). It seems a waste therefore in the context of the question not to provide the full figure, which I have done in Figure 8.

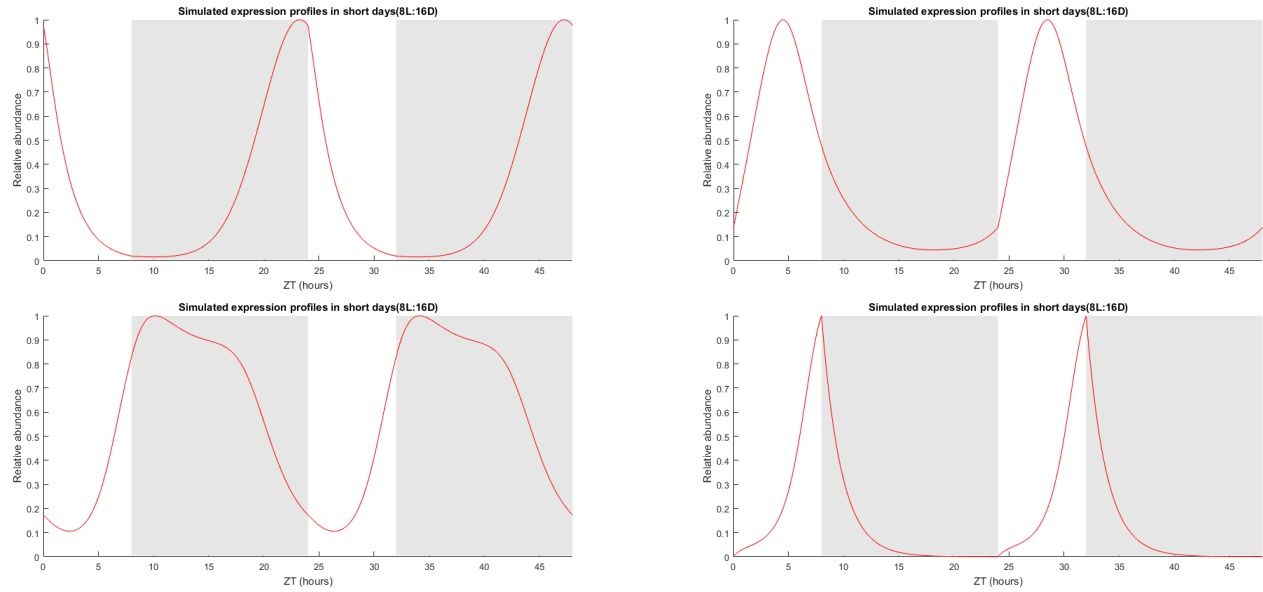


Figure 8: Normalised expression profiles in short days (8L:16D). (A) is CCA1/LHY. (B) (top right) is PRR/PRR7. (C) (bottom Left) is PRR5/TOC1. (D) (bottom right) is ELF4/LUX. Light bars are daylight hours, grey bars are dark periods.

### 3 Question 3

Please refer to file Final\_Q3b.m for the code, which was modified slightly. We'll begin with results from varying proportions of darkness and light in a fixed 24hour period, and further by restricting ourselves initially to CCA1/LHY expression.

Since the start data values determined for question 2 were seen to vary according to the amount of light in the day, and here we are investigating what happens when we change exactly that variable, there are no good values for the start data: As a compromise, I chose the start data for the 12L:12D day since that's the average. Simulations were run with a loop counter varying the amount of light in a day from 1h to 24h over a five day period. Refer to figure 9:

As can be seen, CCA1/LHY expression is fairly robust. Peak expression stays approximately the same as the number of daylight hours is varied, though it comes later and later as the number of daylight hours is increased. With fewer daylight hours expression seems not to diminish as much as with longer daylight periods. Figure 10 is the breakdown of figure 9 into four separate ranges of  $n$  = number of daylight hours.

Having looked at CCA1/LHY expression, and having found it to be mostly robust, the investigation was extended to the other variables. A similar analysis determined that PRR9/PRR7 expression is also fairly robust

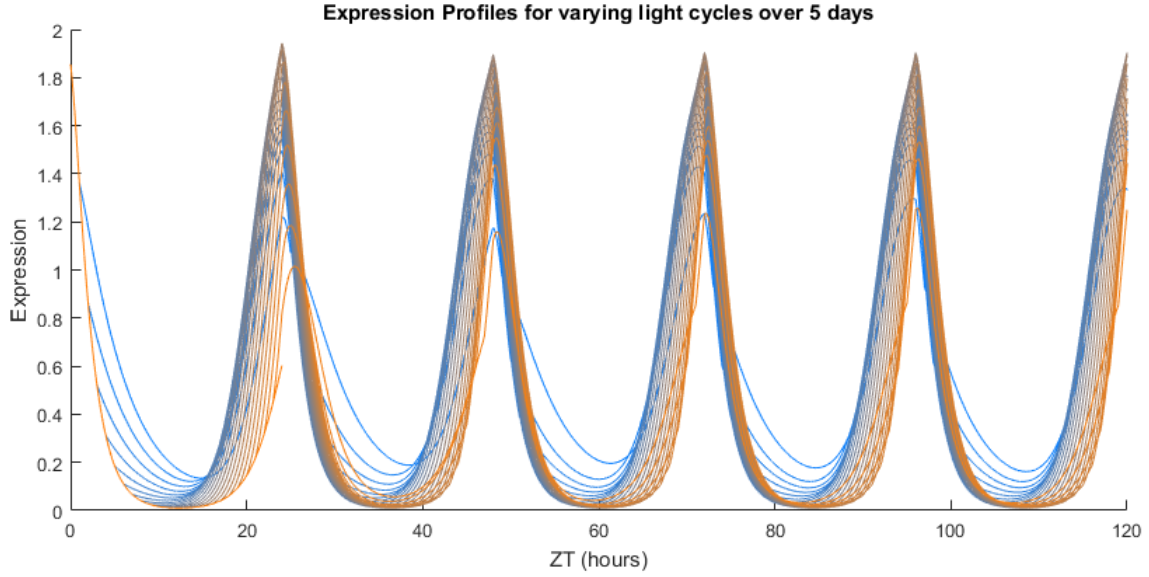


Figure 9: CCA1/LHY expression for 24h days with light period varying from 1 to 24h. Blue lines = less daylight, orange = more daylight

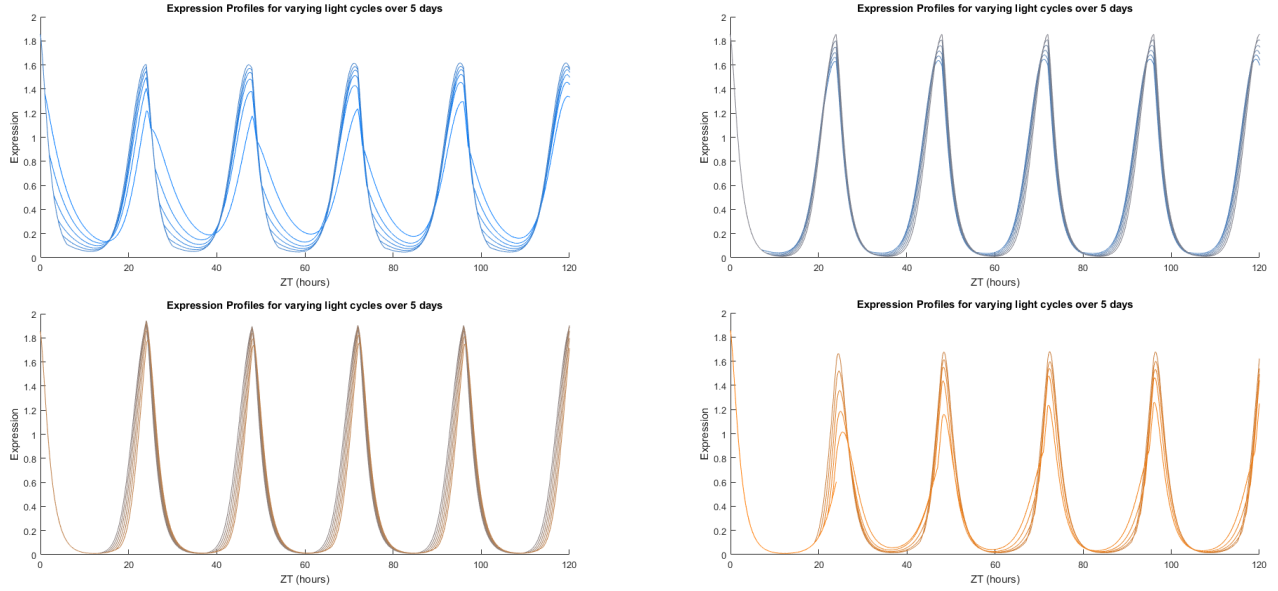


Figure 10: Breakdown of expression profiles for CCA1/LHY. top left: D = 1 to 6. top right: D=7 to 12. Bottom Left: D=13 to 18. Bottom right: D=19 to 24.

to changes in daylight hours, although absolute expression diminished as the number of daylight hours increased past 19 (these figures were omitted to save space and because they were uninteresting). The shape of the profile changes slightly but the same is also essentially true for PRR5/TOC1 expression, which is plotted in figure 11 (left column).

The situation was different for the  $[EL]$  complex, however. As can be seen in the right hand column of figure 11, ELF4/LUX expression is strongly affected by daylight, becoming significantly reduced as the number of daylight hours in a day decreases (blue lines). Section 2.1.2 of the paper notes that the evening complex(EC), which is made up of the the three genes ELF3, ELF4 and LUX, inhibits the expression of morning genes as well as

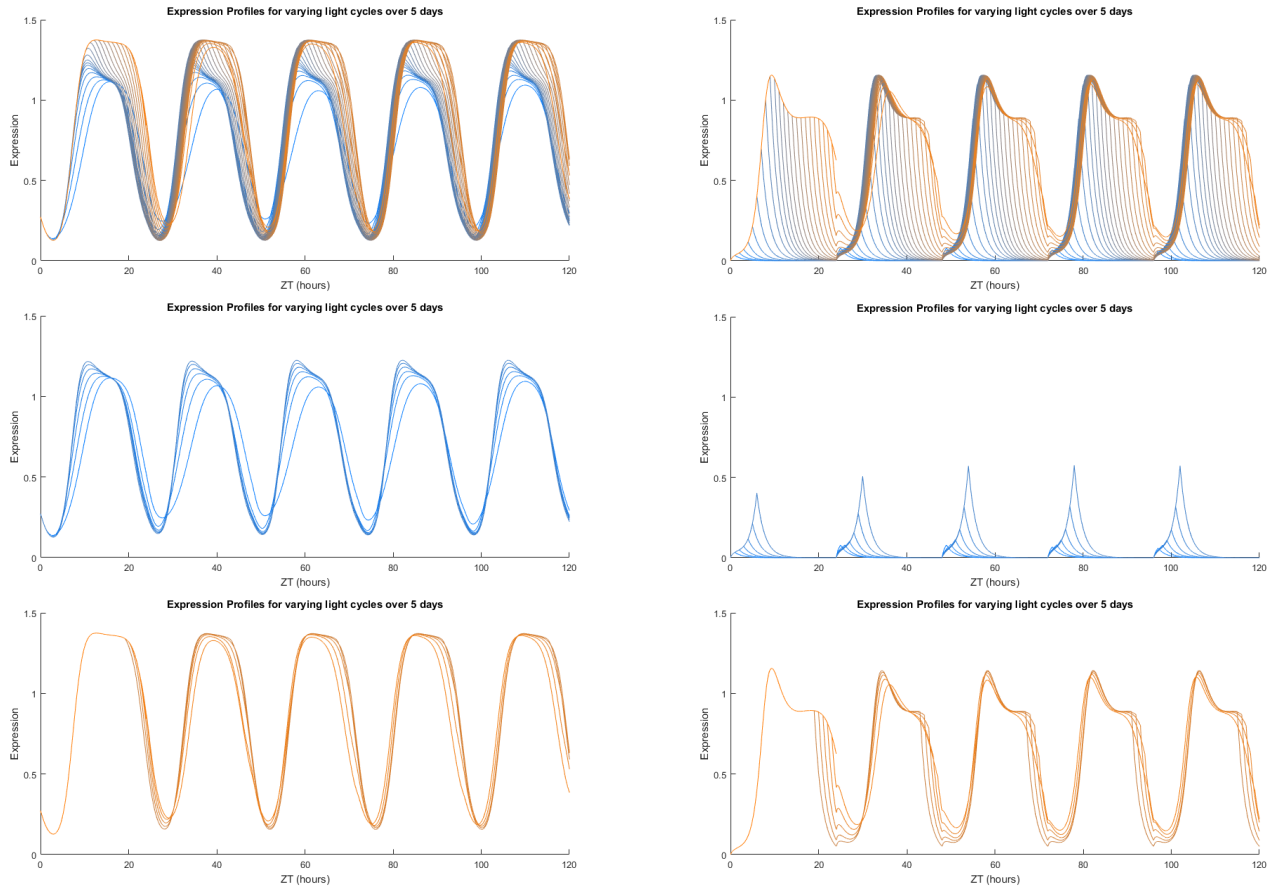


Figure 11: Breakdown of expression profiles for [P51] (left) and [EL] (right). top: entire profile for  $n=1$  to  $n=24$  hours of darkness. Middle: daylight hours ( $n$  in my code) =  $D = 1$  to 6. Bottom: daylight hours = 19 to 24.

growth and flowering related genes. The model then would seem to displaying the opposite of the desired behaviour; when there are very few daylight hours (i.e. winter), [EL] and by extension the EC repressor is least active, which would in theory allow the above mentioned growth and flowering genes greater expression. So the model essentially seems to be suggesting you should see more flowering plants in winter. This would be worth investigating further if time permitted.

The second part of the question asked for an investigation of equal periods of light and dark but with varying total day lengths. Refer again to Final.Q3b.m for the code used. A range of  $n = 10$  to 50 hours was initially used to test different lengths of day, keeping daylight hours = nighttime hours =  $n/2$ . In order to compare the effective periods each individual plot was rescaled by its own daily period so the results could all be compared on a single graph. As always, we begin with CCA1/LHY expression, summarized in Figure 12. As can be seen I've attempted to shade these graphs to highlight extreme values for total length of day: as the number of hours becomes very short, the red shading increases, and as the number of hours becomes very large, the amount of blue shading increases. Grey lines are mid-range values.

Comparing these two graphs it is apparent that for some middle range the circadian clock seems robust, but for short days (red) and for long days (blue) there is significant deviation from the standard expression profile for



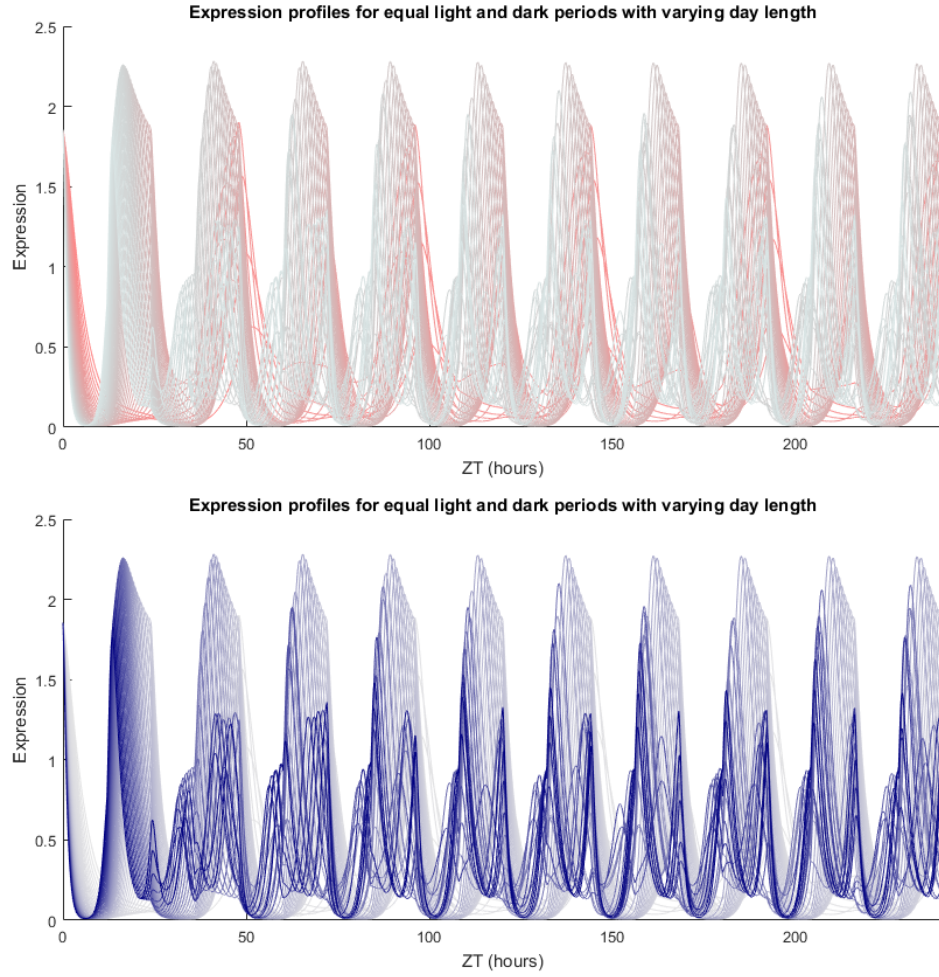


Figure 12: Expression profiles for CCA1/LHY for day lengths varying from 10 to 50 hours, normalized for period length on the x axis. Top: red lines indicate very short periods. Bottom: blue lines indicate very long periods.

CCA1/LHY. By splitting the data further into groupings of 10 days (figure 13) this can be made even more apparent.

Upon breaking the data down still further, it was found that the loss of periodicity of expression at the low end first occurs when  $n = (\text{total day length}) = 15$ . the expression profile then recovers briefly but completely breaks down at  $n=10$ . This can be seen in the top left plot in figure 13. At the upper end, full periodic expression of CCA1/LHY is lost at  $n=34$  hours, and this is shown in figure 14.

For completeness' sake I also ran simulations for the PRR family and for the dusk genes (ELF4 and LUX). I have included the plots obtained below in figure 15. Similar effects as with CCA1/LHY are seen again with PRR9/PRR7 and the periodicity seems to be lost at high values. With PRR5/TOC1 and ELF4/LUX the periodicity appears to be more robust than for CCA1/LHY expression and remains somewhat periodic even at extreme day lengths of 50 hours. Given these plots I think using more detailed analysis on CCA1/LHY and then using that as a benchmark for robustness as I have done is only partly justified; while it makes a good limiting factor (if the system fails anywhere it will likely be in the expression of CCA1/LHY or PRR9/PRR7 according

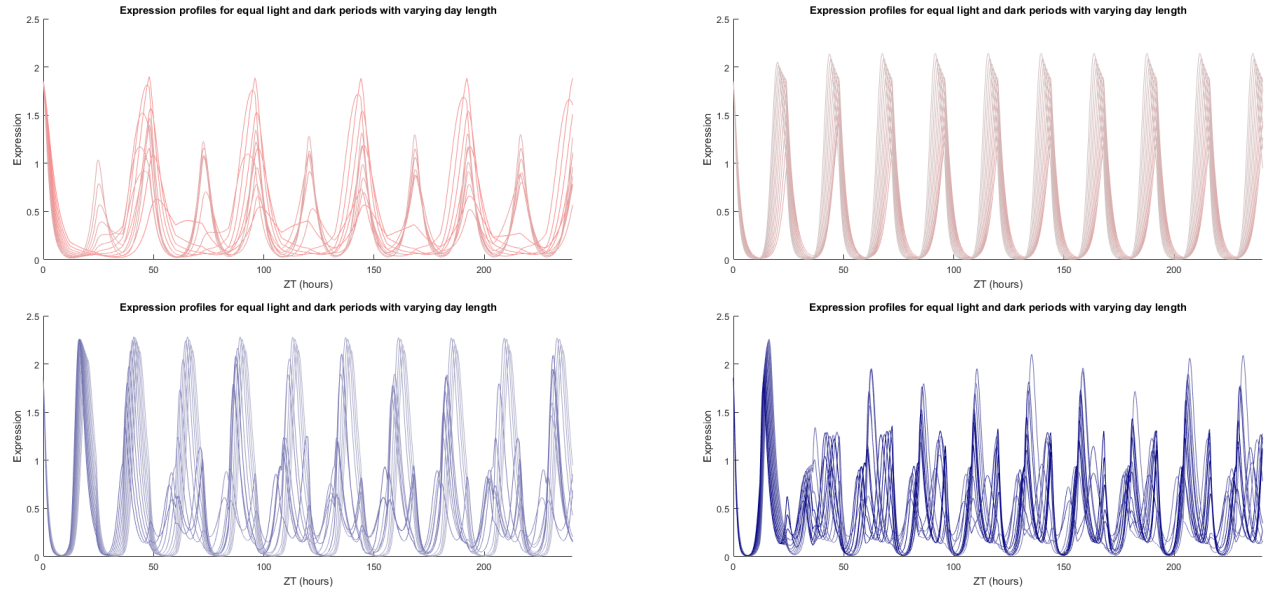


Figure 13: Top left, periods of 10-20 hours, appears to be devolving a staggered expression profile where full expression occurs only every other day. Top right: 20-30, stable. Bottom left: 30-40, approximately stable. Bottom right: 40-50 hours, periodicity of expression largely lost

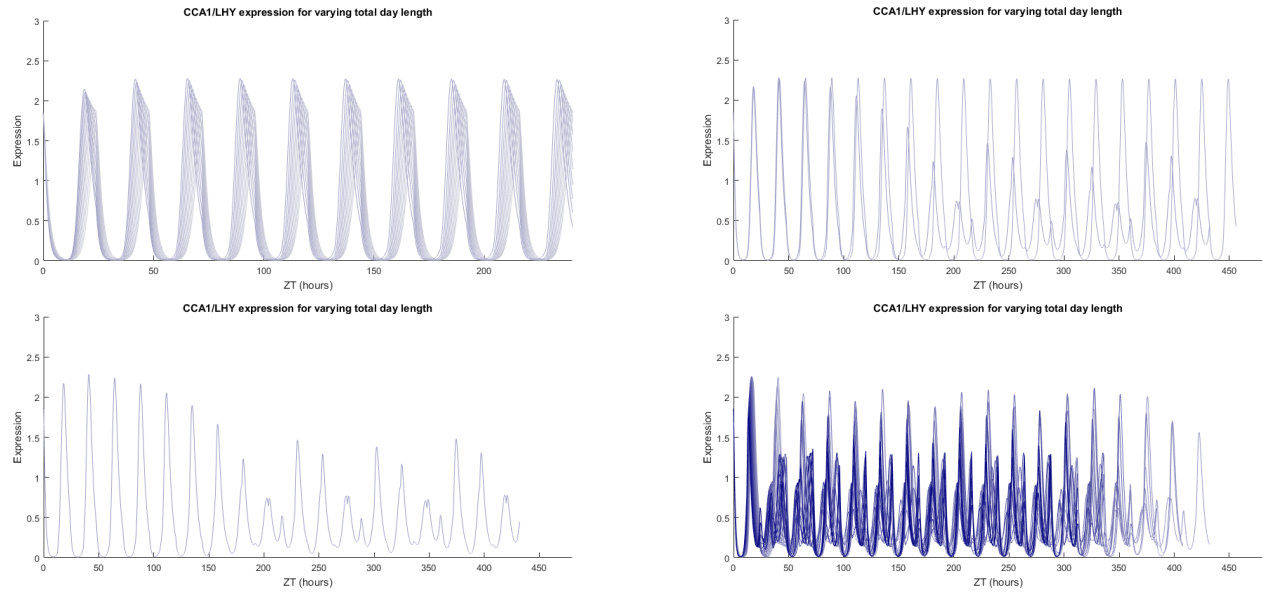


Figure 14: Breakdown of CCA1/LHY periodicity for variation in  $n$ =total day length (equal light/dark portions). Top left:  $n=23$  to  $33$ , periodicity is maintained. top right:  $n=33$  and  $34$ . Bottom Left:  $n=34$ . Bottom right:  $n=35$  to  $50$ . Note that after  $n=34$  total hours in a day periodic expression is never truly regained.

to the figures), saying that expression in CCA1/LHY becomes less than ideal ignores the fact that other parts of the circadian clock continue to hold up well even at high values.

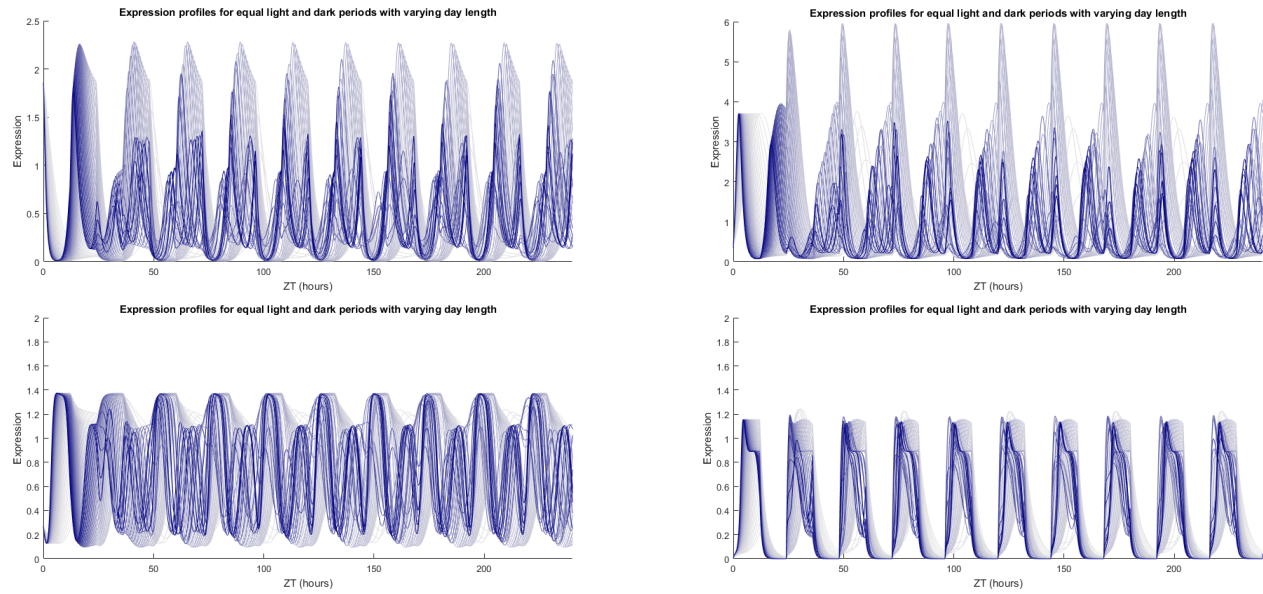


Figure 15: Comparisons of the effect of changing the period on expression. Top Left: CCA1/LHY. Top right: PRR9/PRR7. Bottom left: PRR5/TOC1. Bottom right: ELF4/LUX

## 4 Question 4

I've chosen to attempt (B). My source materials are as follows:

- (1) The document provided in moodle, DELAY "DIFFERENTIAL EQUATIONS IN SINGLE SPECIES DYNAMICS" by Shigui Ruan, O. Arino et al. (eds.), Delay Differential Equations and Applications, Springer, Berlin, 2006, pp.477-517.
- (2) "Numerical Treatment of Delay Differential Equations by Hermite Interpolation" by H.J. Oberle and H.J. Pesch, Numer. Math. 37, 235-255 (1981)
- (3) A matlab tutorial found at <https://uk.mathworks.com/matlabcentral/fileexchange/3899-tutorial-on-solving-ddes-with-dde23>  
This tutorial was a zip file containing example code and a PDF entitled "Solving Delay Differential Equations with dde23", by L.F. Shampine and S. Thompson, Appl. Numer. Math., 37 (2001), pp. 441-458 (and if I understand the text correctly they seem to have been responsible for creating the dde23 function for MATLAB)
- (4) My own experience and files from last year.

From what I understand of the source texts listed above and gather from my own files, the basic theory is as follows: to change supplementary equation(2) in De Caluwé et al. (2016), I need to define a history (for values of  $(t)$  less than the delay value) and a vector **lags** containing the delays themselves(in this case one delay, for  $y(1)$ ).

Since we need to keep the loop cycles for day/night parameter changes used in questions (1)-(3), and since I use a "steady state" to define the initial value anyway, I will use **ystart** as my history vector for the 3 daylight cycles I'm considering: a short day (8D:16L), a medium day (12D:12L) and a long day (16D:8L).

The hints provided for the question suggest: "...You may use either the linear chain trick with a weak delay kernel (see Ruan (2006), Figure 13) or..." Since DDE23 is awkward to implement over the matched conditions as we move from day to night in a cycle, I've chosen to try and implement the problem this way. The two equations that need updating from the model are:

$$\frac{d[CL]_m}{dt} = \frac{(v_1 + v_1 L * L * [P])}{1 + \left(\frac{[P97]_p}{K1}\right)^2 + \left(\frac{[P51]_p}{K2}\right)^2} - (k_{1L} * L + k_{1D} * D) * [CL]_m$$

And

$$\frac{d[CL]_p}{dt} = (p_1 + p_{1L} * L) * [CL]_m - d_1 * [CL]_p$$

In both cases, with a weak delay kernel, the time delayed form of  $[CL]_m$  becomes and is then replaced by

$$[CL]_m(t) \mapsto [CL]_m(t - \tau) = \int_{-\infty}^t \alpha e^{-\alpha(t-u)} [CL]_m(u) du = I(t)$$

In addition we now have the new constraint/equation

$$\frac{dI}{dt} = \alpha[CL]_m(t) - \alpha I(t)$$

So equations (1) and (2) in the model become replaced by the equations

$$\frac{d[CL]_m}{dt} = \frac{(v_1 + v_1 L * L * [P])}{1 + \left(\frac{[P97]_p}{K1}\right)^2 + \left(\frac{[P51]_p}{K2}\right)^2} - (k_{1L} * L + k_{1D} * D) * I$$

$$\frac{d[CL]_p}{dt} = (p_1 + p_{1L} * L) * I - d_1 * [CL]_p$$

$$\frac{dI}{dt} = \alpha[CL]_m(t) - \alpha I(t)$$

With these alterations the problem can still be solved numerically using ODE45. Refer to files Final\_Q4L.m and Final\_Q4runme.m. Ruan states that  $\alpha = 1/T$  for the weak delay kernel where  $T$  is the average delay; since the delay in mRNA expression is likely on the order of hours an average delay might be expected to be approximately 2, giving us a value of  $\alpha = 0.5$ . To begin with, and to check that very short delays converged to the original model, I ran the simulation for  $\alpha = 10$  which would correspond to an average delay of only 0.1 hours for our major four expression complexes ( $[\square], [\square], [\square], [\square]$ ), essentially reproducing a version of Figure 2 from the paper (but I ran my simulations for 10 days in case the delays produced transients). This is shown in figure 16.

Having established that the delayed system was implemented correctly and behaving as expected,  $\alpha$  was steadily

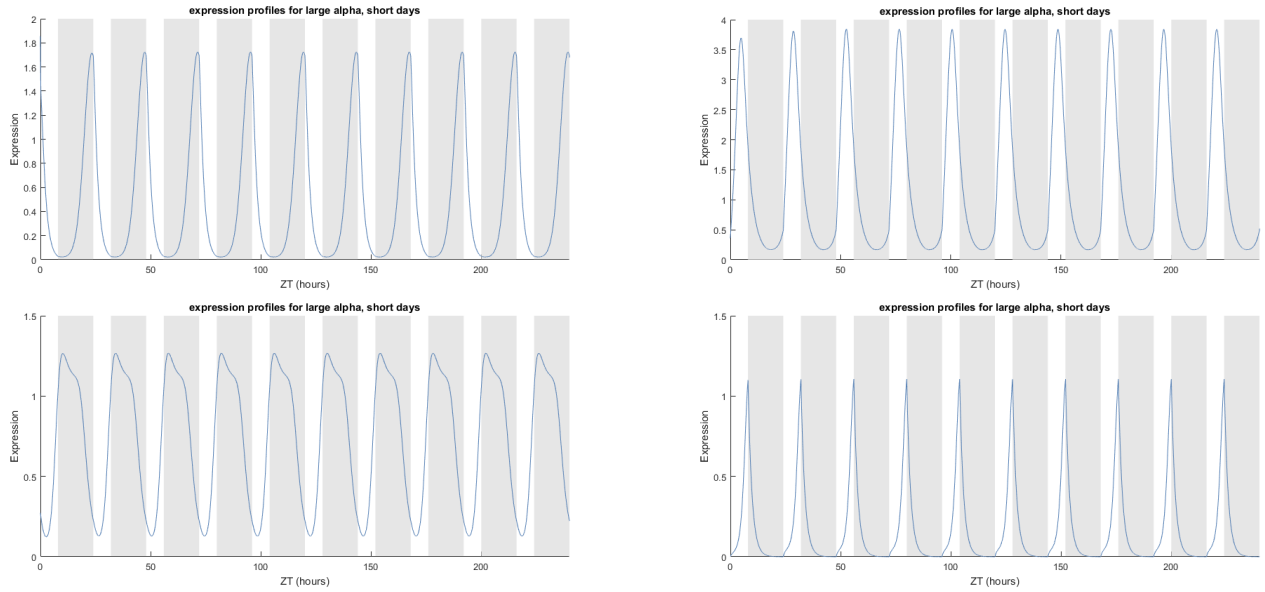


Figure 16: Expression profiles for a weak delay kernel and average delay 0.1h

decreased to simulate longer and longer average delays (still for short days only, i.e (8L:16D). Maximum expression levels of CCA1/LHY become slightly elevated at  $\alpha = 2$ , corresponding to an average delay of 0.5h or 30 mins (not shown); despite the fact that this is the delayed variable, however, the maximum expression peak is still at the same position and periodicity is still preserved.

As  $\alpha$  is decreased from 2 down to 0.5 (corresponding to an average delay on the order of hours), maximum expression of CCA1/LHY becomes further elevated and now expression is also zero for short periods. The peak is still in the same place relative to the cycle and periodicity is still maintained. The two repressors are still unaffected, though the shape of the Evening complex profile starts to be. All these results also continue to be true for a wide range of L:D ratios, so again no figures are included as they are mostly unremarkable.

As  $\alpha$  is decreased still further below the 0.5 threshold, effects become more interesting. At  $\alpha = 0.4$  the Evening Complex (ELF4/LUX) expression profile starts to change shape, and profiles for "short" days with little light start to diverge from "longer" days with more daylight (the total period of one day is still fixed to 24h). At  $\alpha = 0.35$  the maximum expression of CCA1/LHY starts to trend upwards each period, though in order to see this you have to run the simulation for 20 days. Figure 17 shows the expression complexes at this point, just before a breakdown of periodicity.

The next figure (figure 18) is at  $\alpha = 0.34$ . Note that a value of  $\alpha = 1/3$  corresponds to an average delay of 3 hours. This is the threshold value beyond which expression of CCA1/LHY behaviour becomes complex. For "short" days (i.e. 24h days with 8 or fewer hours of light), the situation becomes increasingly erratic. For 24h days with 12h periods of sunlight or longer, there appears to be some kind of bifurcation-like behaviour as the



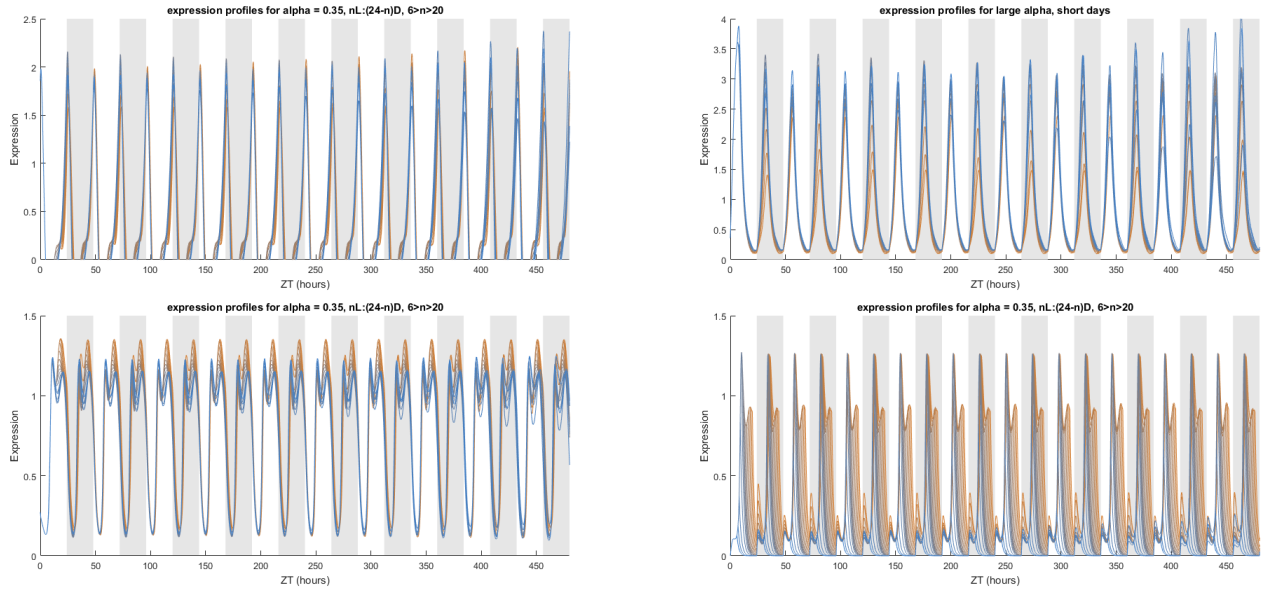


Figure 17: Expression profiles for  $\alpha = 0.35$ . Notice that expression profiles of CCA1/LHY and PRR9/PRR7 are increasing with time and therefore losing stability

expression of CCA1/LHY starts to alternate between two stable values, then this behaviour too collapses and behaviour becomes erratic for all ratios of light to darkness in the 24h cycle.

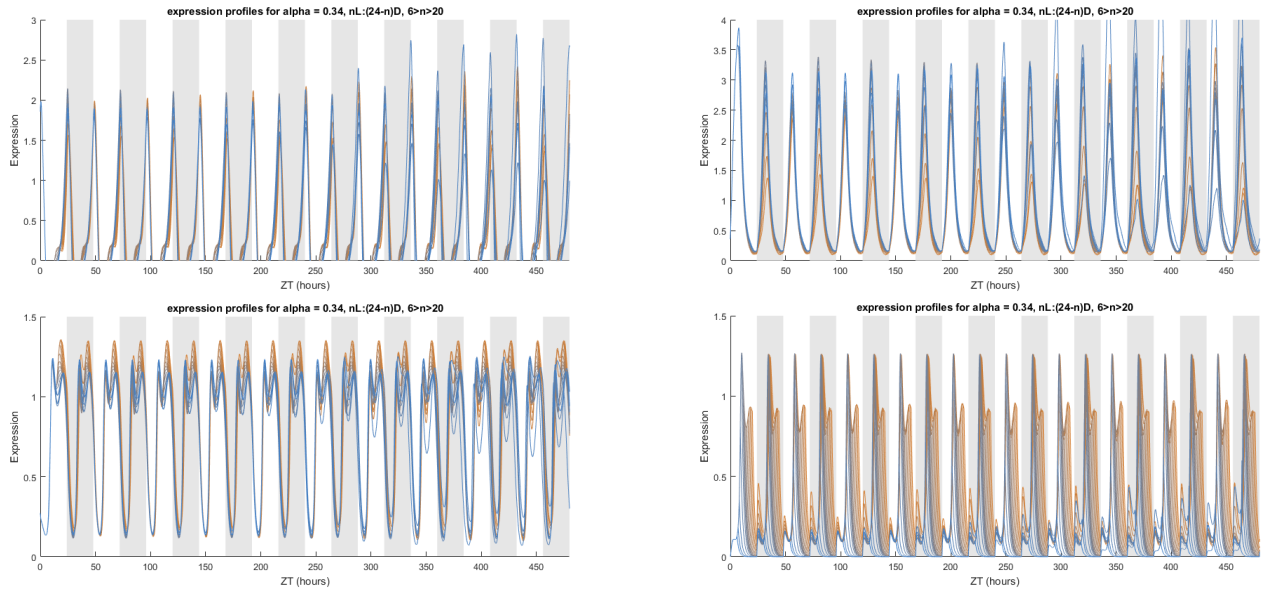


Figure 18: Expression profiles for  $\alpha = 0.35$ . Notice that expression profiles of CCA1/LHY and PRR9/PRR7 are increasing with time and therefore losing stability

A value of  $\alpha$  below 0.2 starts to sound unreasonable, as the corresponding average delay time is now 5hours; the last figure (figure 19) looks at the expression of CCA1/LHY for different ratios of light and dark at this value. It becomes difficult at some point to specify exactly what we mean by "robustness of the clock", or even just behaviour of the clock. It's extremely clear that expression level of CCA1/LHY is extremely erratic for all but the "longest" days, with random spiking and no identifiable period of any kind; and similar results obtain for PRR complexes and the ELF4/LUX complex. However, for 20L:4D days we still see a spike every 48 hours,

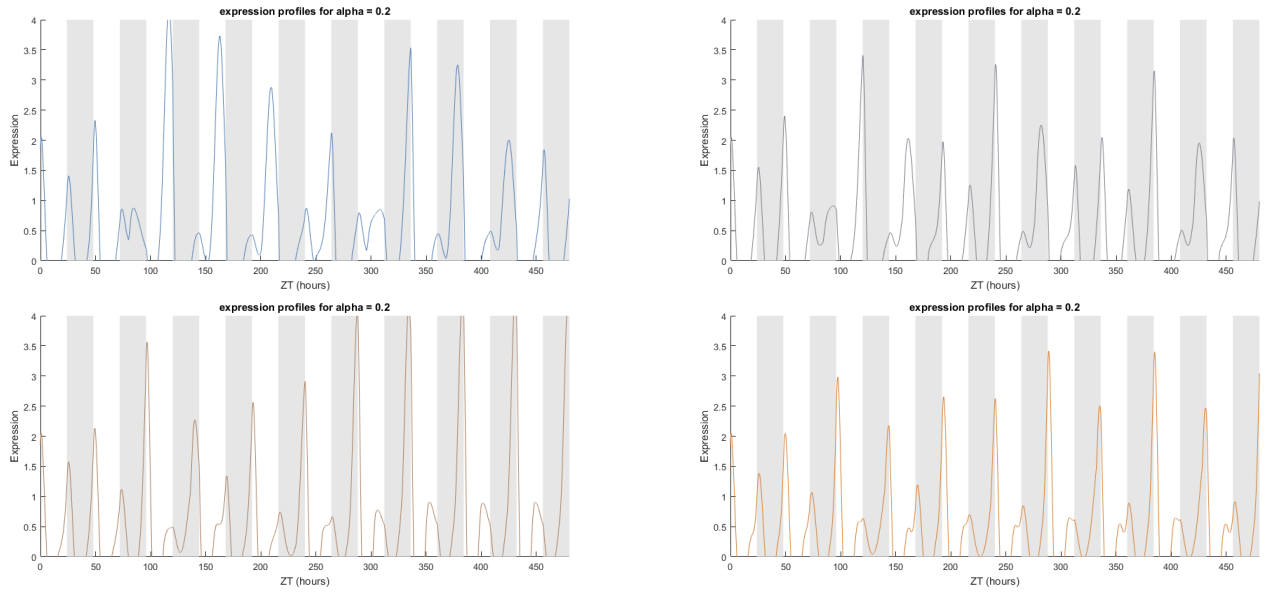


Figure 19: Expression profiles for  $\alpha = 0.2$ . at D:L = 8:16(top left), 12:12(top right), 16:8(bottom left), and 20:4(bottom right).

which is remarkable when we consider both the unusual length of daylight hours (summer in the arctic circle) and the average delay in  $[CY]_m$  expression.

On the other hand, plant proteomes are much larger than the limited number of proteins measured in the clock itself, and elevated expression levels of these proteins might already be enough of a problem to consider the clock badly broken; As mentioned earlier, for example, the evening complex (ELF4/LUX) also has an effect on plant growth and flowering, and spikes in that expression might have disastrous effects elsewhere in the plant. So on that basis even the behaviour observed at  $\alpha = 0.35$  is already cause for concern.