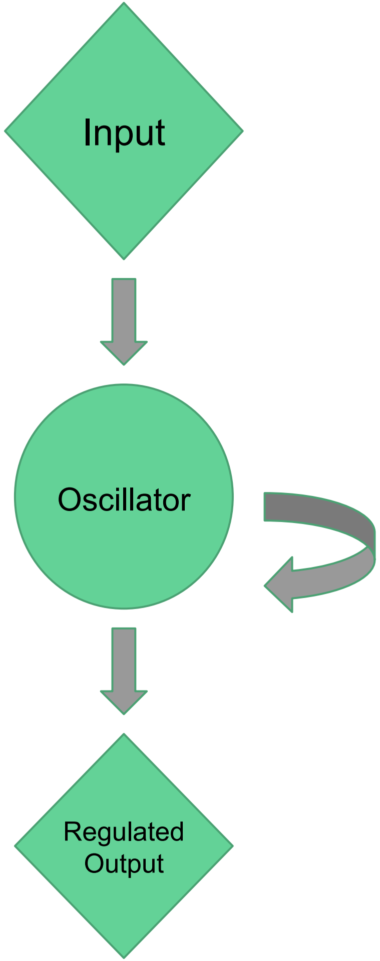
Who turned out the lights?

Boolean Network Models and Their Shortcomings

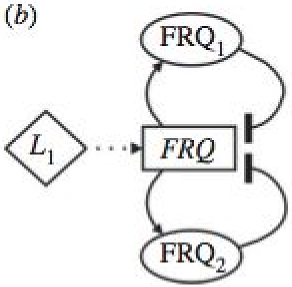
Circadian Rhythms

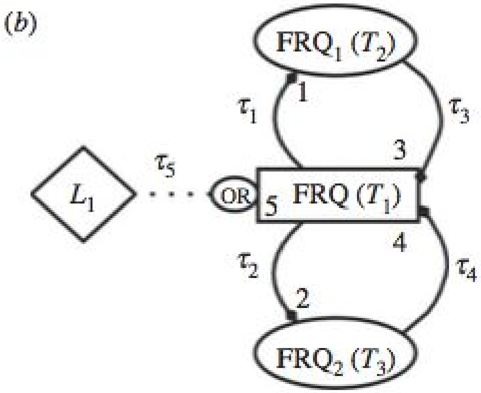
- Daily Oscillations occurring in most organisms
  - Affect metabolism, behavior etc.
- Produced by positive and negative gene regulatory feedback networks
- Sync to the environment using light and temperature



Boolean Models

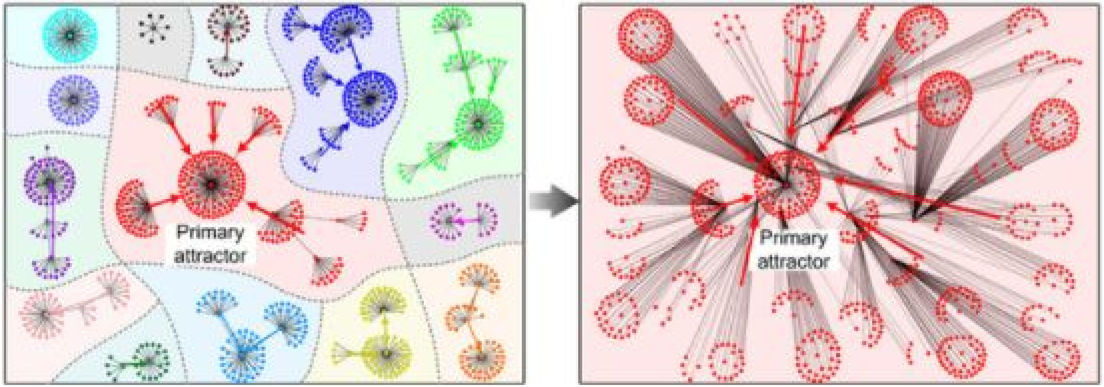
- Vertices
  - On or Off
- Edge
  - Where edges meet Vertices
  - Each a boolean operator
- Types of Boolean Operators
  - And
  - Or
  - Not
  - Identity
  - Plus some more





The Original Research Question

- Draw on a 2013 paper by Junil Kim et. al.
  - Control kernel for boolean networks
  - By controlling a subset of nodes, we can control end behavior of the whole model
- Can we isolate the control kernel for a circadian system?
  - Is it just the light nodes?
  - Does it include all the light nodes?



No

This didn't work

- Turns out, the models are too small for interesting behavior concerning multiple attractors
    - Uninteresting
- The added complexity introduced by light nodes and signalling delays made applying the concepts from the paper difficult
    - Unrealistic given length of project

New Research Question

- As we were working, we modeled turning off all the lights
  - This broke the model almost immediately
  - Surprising to us, expected some altered continuation of behavior at least for some timesteps
- These models, like all models, have limitations in their applications to biology
  - Especially because these models were selected by a computer
- Our new goal for the project
  - Find the limitations of these circadian models
  - Understand how these problems reflect composition of the model
  - Interpret biological implications
  - Attempt to fix

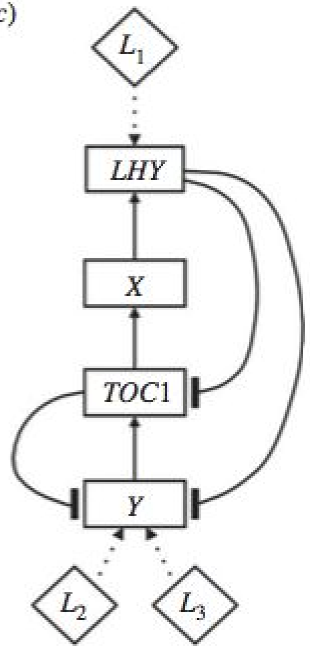
The Model

- We decided to focus our efforts on the 2-loop *Arabidopsis* model.
    - Large enough to see some interesting behaviors
    - Small enough to implement quickly

2-loop *Arabidopsis*

$$
\begin{aligned}
LHY(t) &= X(t-\tau_3)\cdot L_1(t-\tau_7)\\
TOC1(t) &= \overline{LHY}(t-\tau_1)\cdot Y(t-\tau_6)\\
X(t) &= TOC1(t-\tau_2)\\
Y(t) &= \left(\overline{LHY}(t-\tau_4)\cdot \overline{TOC1}(t-\tau_5)\right)\cdot\left(L_2(t-\tau_8)+L_3(t-\tau_9)\right)
\end{aligned}
$$

(c)



The Plan

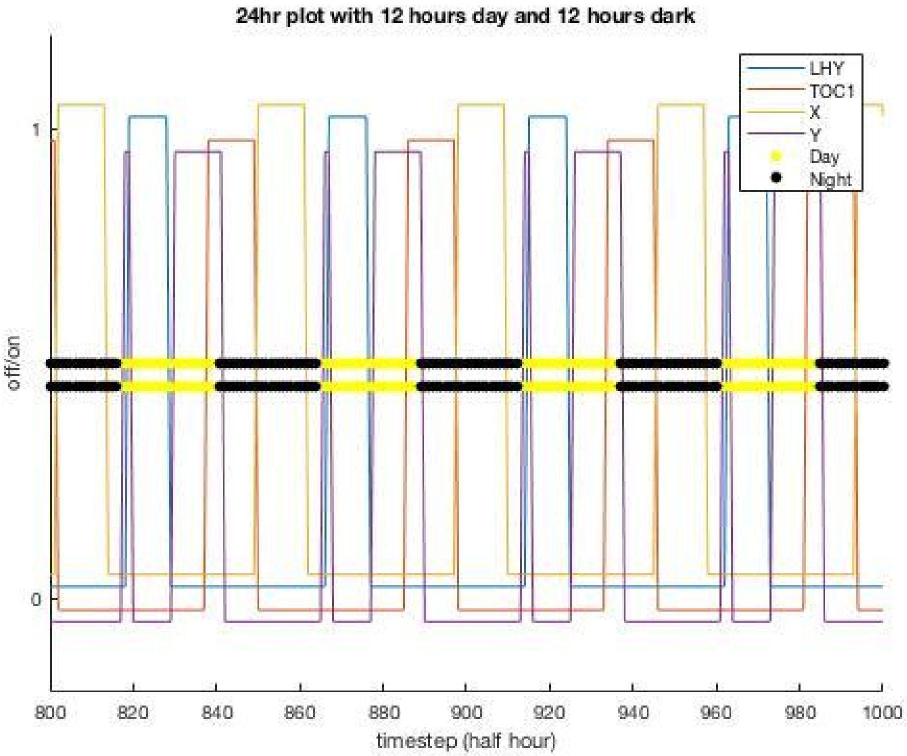
- Every possible 24 hr photoperiod (24 dark -> 24 light)
  - Modeled over unnecessarily long amounts of time
- Every possible 48 hr photoperiod (48 dark -> 48 light)
  - Modeled over unnecessarily long amounts of time
- Starting the model with subsets of lights
- Running lights on asynchronous photoperiods
- Turning off lights midday
  - Turning off subsets of lights midday
- Turning on genes before simulating

github.com/gnorthrup/BooleanControl

The Method

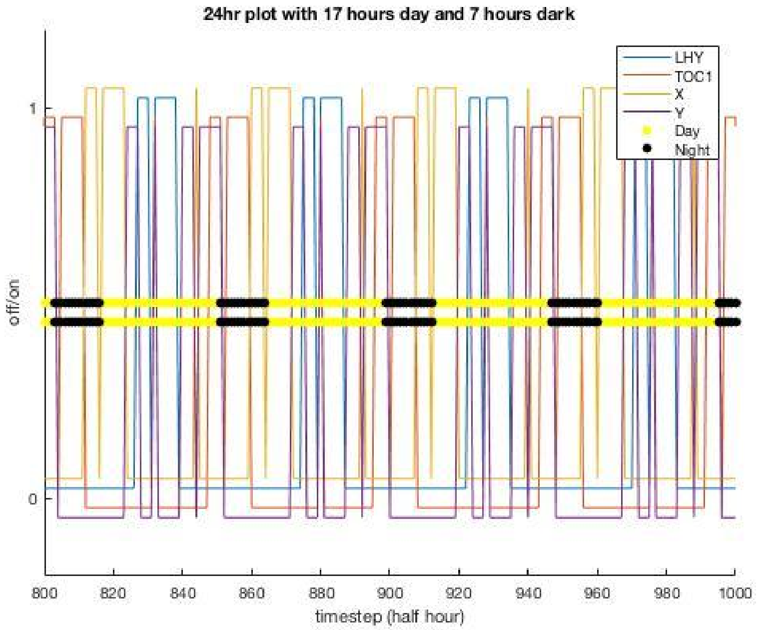
- 3 manually defined classes
  - EvalTree
  - FetchNode
  - LightNode
- Boolean networks are a big recursive mess to implement
  - EvalTree used recursively for every logic gate
  - FetchNode & LightNode used as recursion base to return actual values of the nodes
- Matlab is not typesafe
  - This ended up helping us
  - We can pass multiple node types through EvalTree and it was fine
- With all this setup, we created a big array of zeros and iterated over every entry

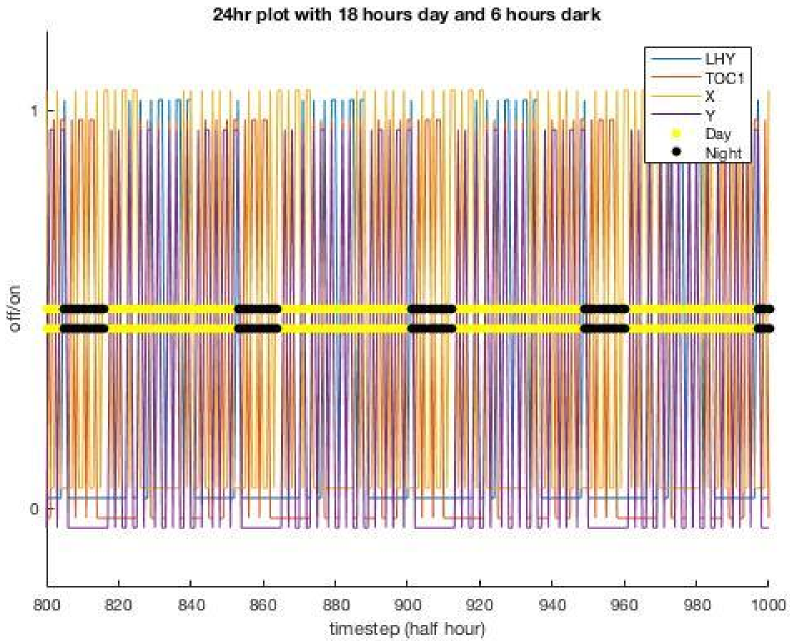
24hr photoperiods - Results



- 12 hr day and 12 hr night produces the results seen in the original paper
    - We did something right

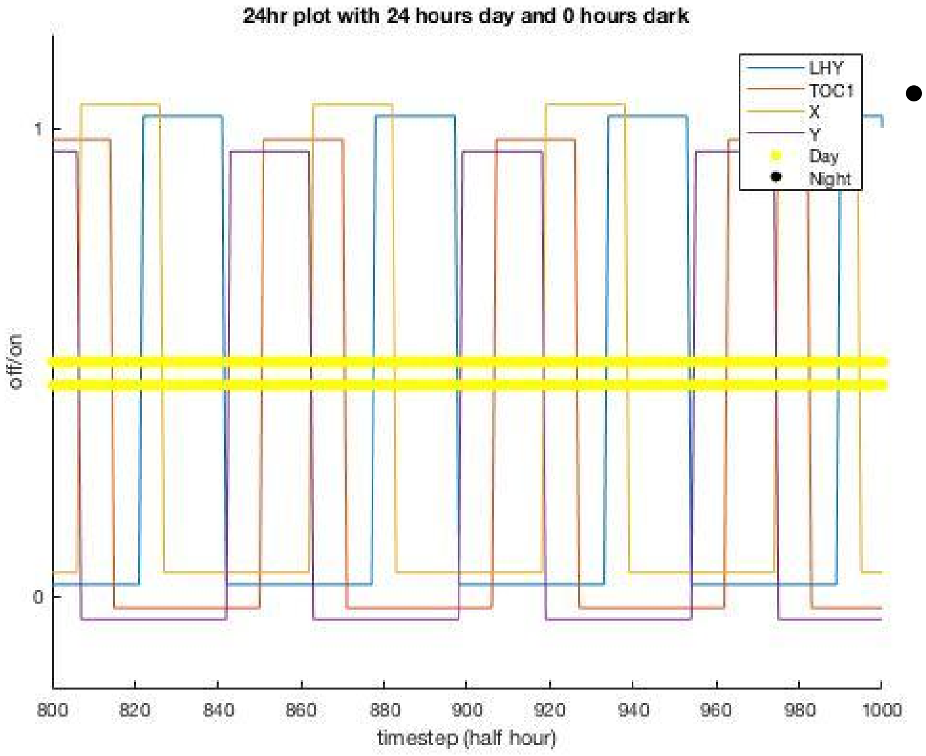
24hr photoperiods - Results



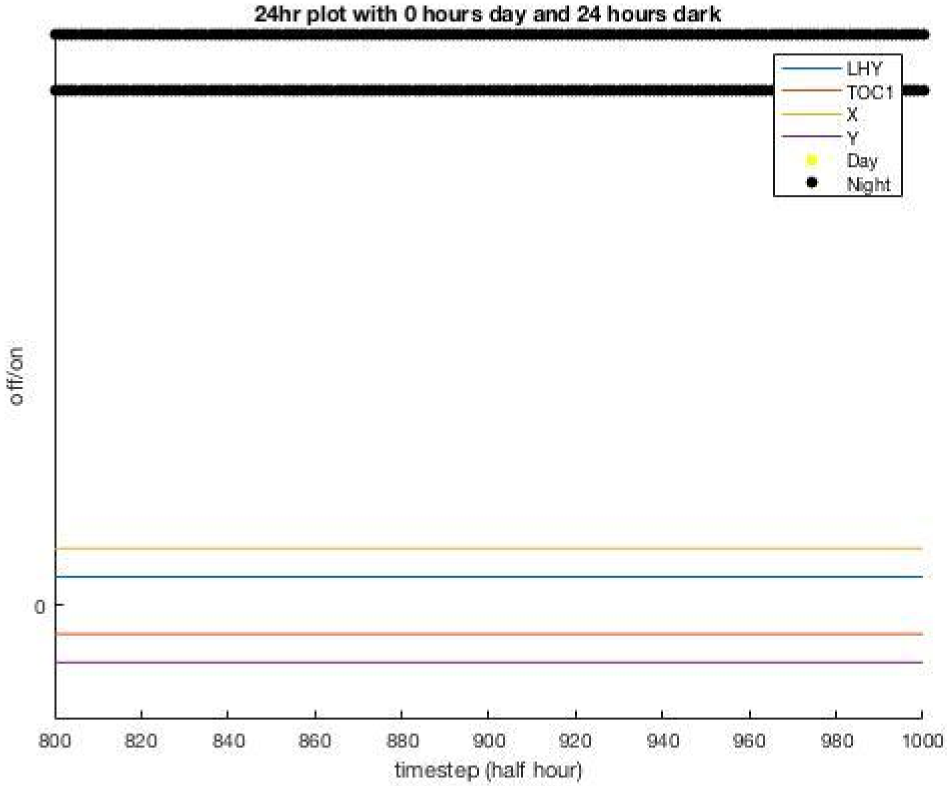


- At 17hr day: unusual results, multiple spikings. Behavioral origin clear.
- At 18hr+ day: entirely rapid switchings. Behavioral origin utterly unclear.

24hr photoperiods - Results



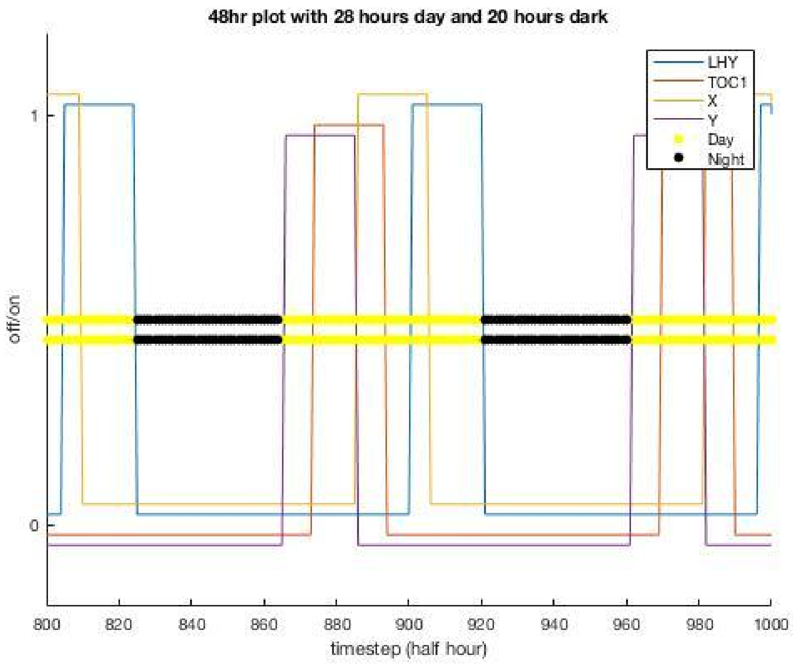
- The theoretical ‘innate attractor’ of the model
  - 28 hr period
  - Agrees with biological understandings
  - Model was designed with this in mind



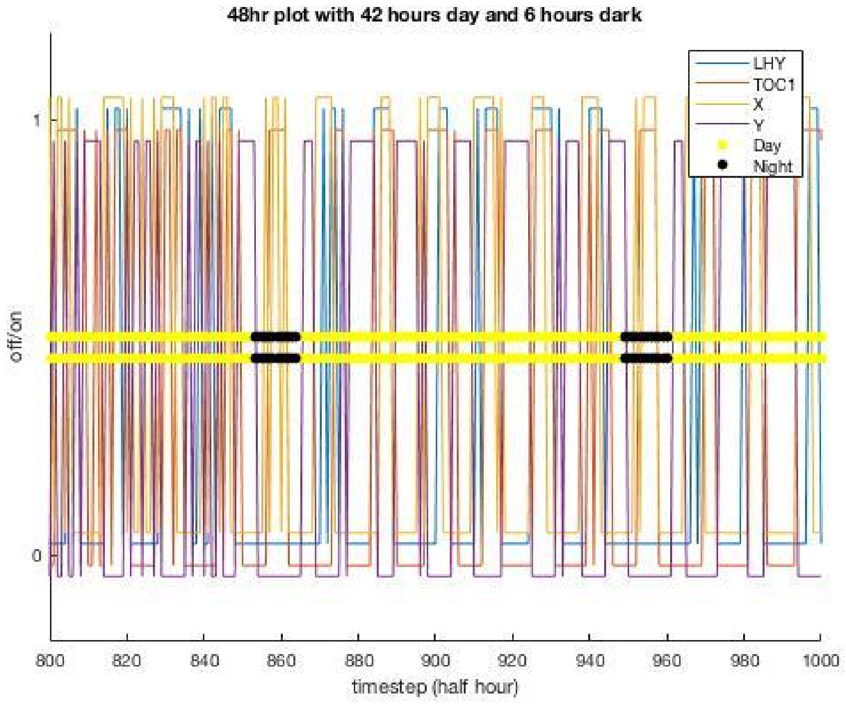
- Also, rather unsurprisingly if you never turn on the lights the model never does its thing

48-hr photoperiod - Results

- When the night gets too long (20hrs) the model shuts off behavior at night
  - Similar to just turning off the lights when night is longer than the longest signal delay + time to turn off
- Also 28 hr day is the exact innate period of the model
  - For models with longer days (but not so long they break) we see the same 28hr pattern then a secondary pattern



48-hr Photoperiod - Results

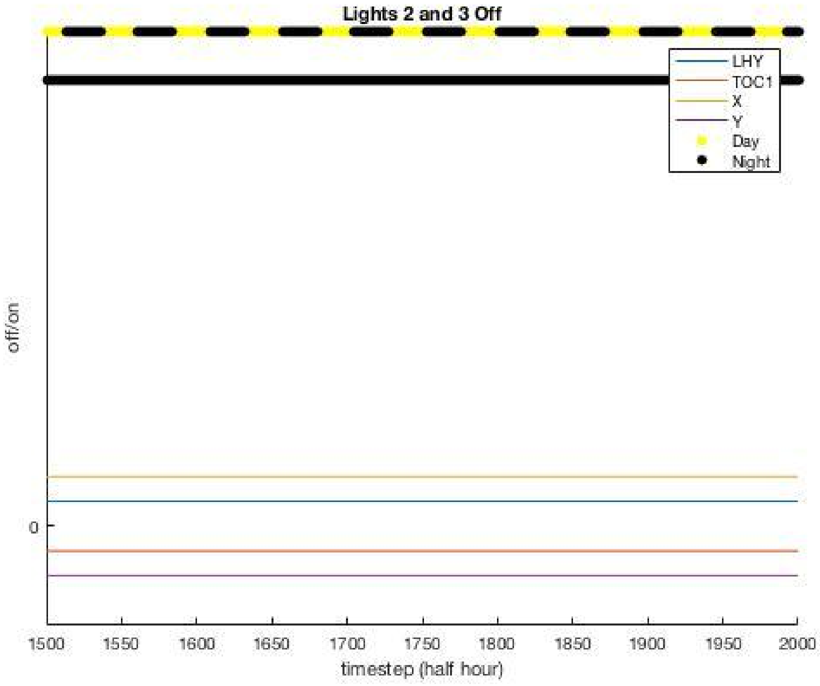


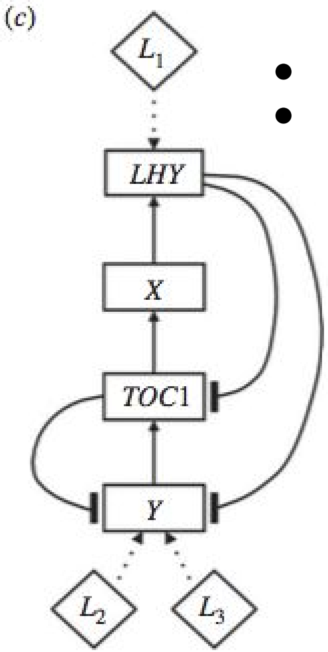
- Full disclosure we have no clue what is happening here
- The model goes from crazy to less crazy, but is never periodic
  - Even after many many timesteps
- Previous whacky or broken plots still reflected periodic behavior
  - Boolean Chaos?

Photoperiod - Conclusions

- The model does poorly (i.e. does not reflect biological behavior) for photoperiods outside the range of usual photoperiods
  - This model was built to fit data on usual photoperiods, without changing the parameter values we can't improve the behavior
- From a theory standpoint, this is caused by the size of the model
  - Small networks propagate changes exceptionally quickly
  - A network with more nodes is more robust to these perturbations

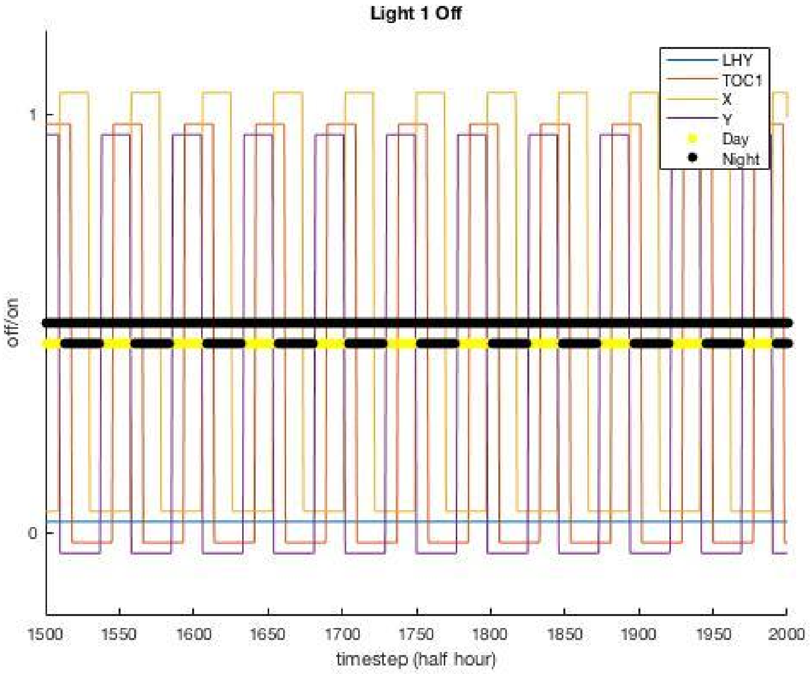
Subsets of Lights - Results

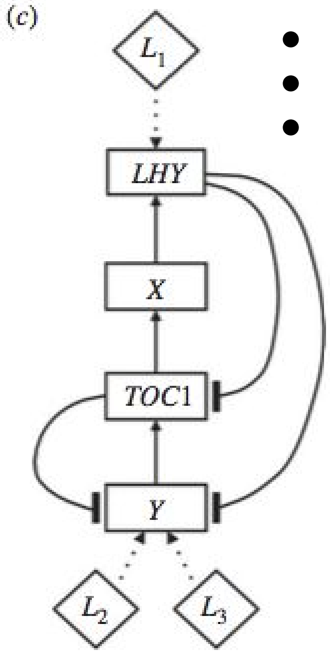




- Without L_2, L_3 Y shuts off
- Y propagates positively through system

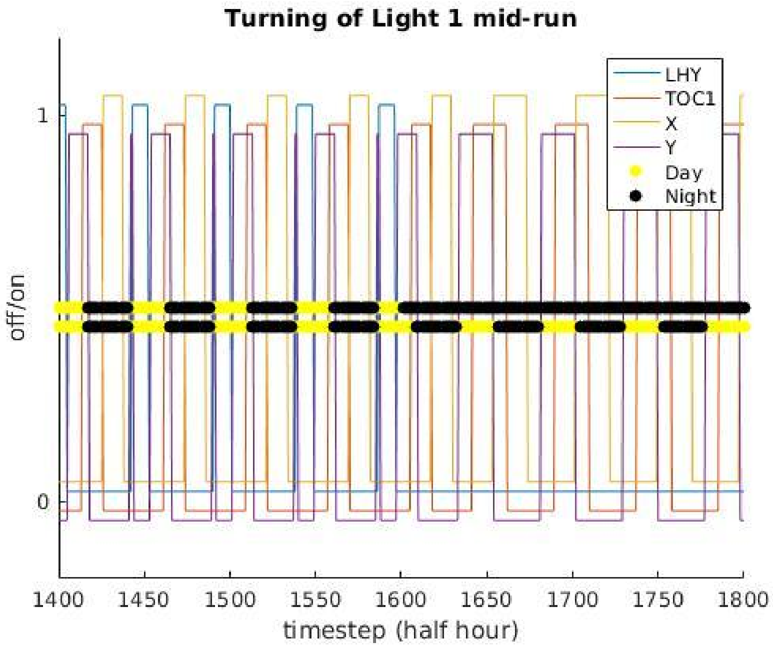
Subsets of Lights - Results

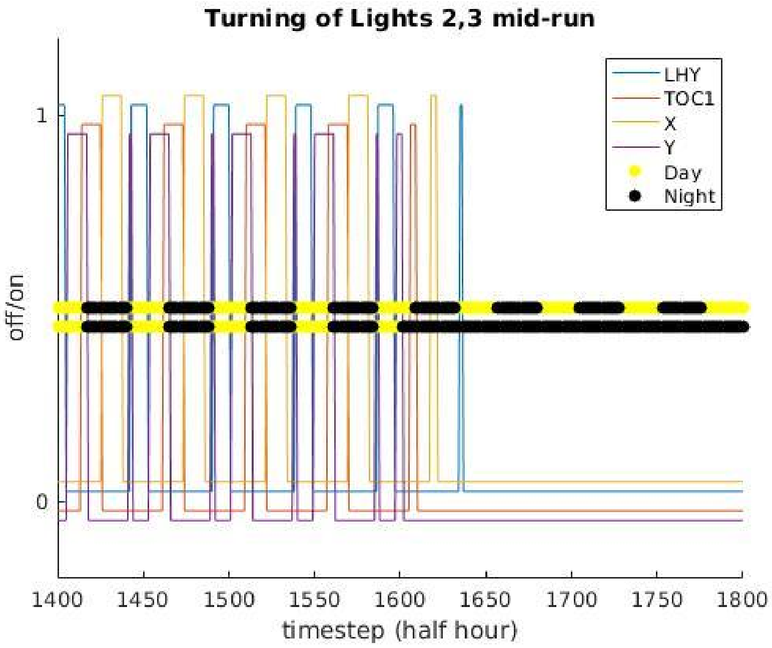




- LHY is primarily inhibitory
- System functions without it
- Y can keep everything else going

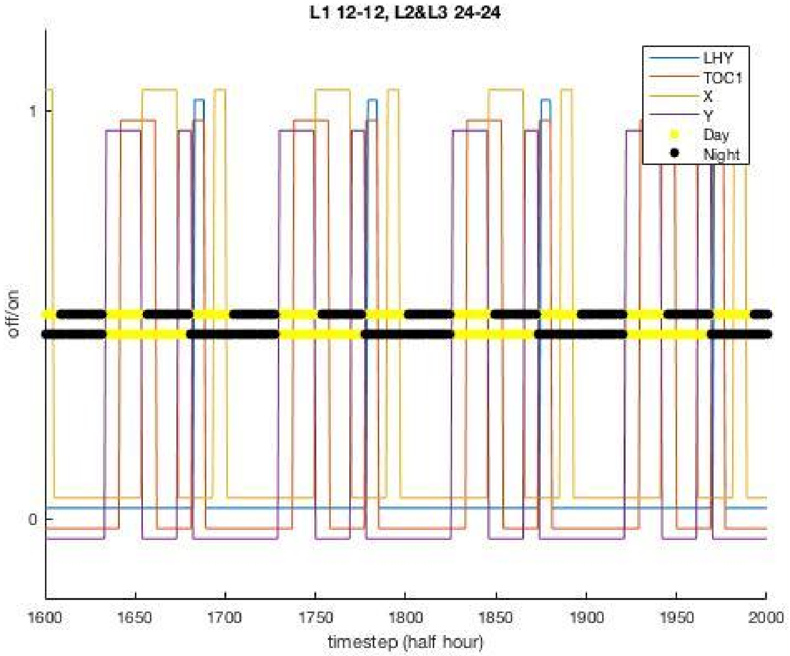
Turning Lights off Mid-Simulation - Results



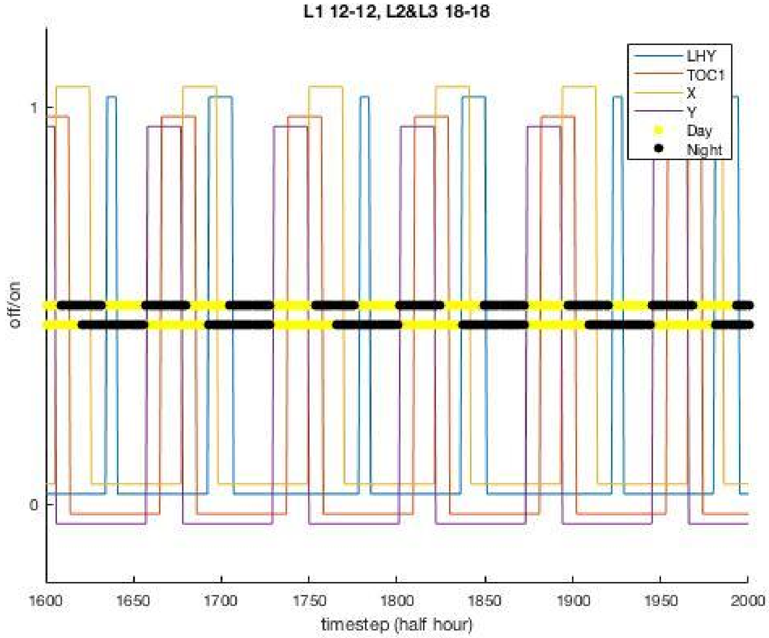


- Changes propagate within a day or so

Asynchronous Photoperiods - Results

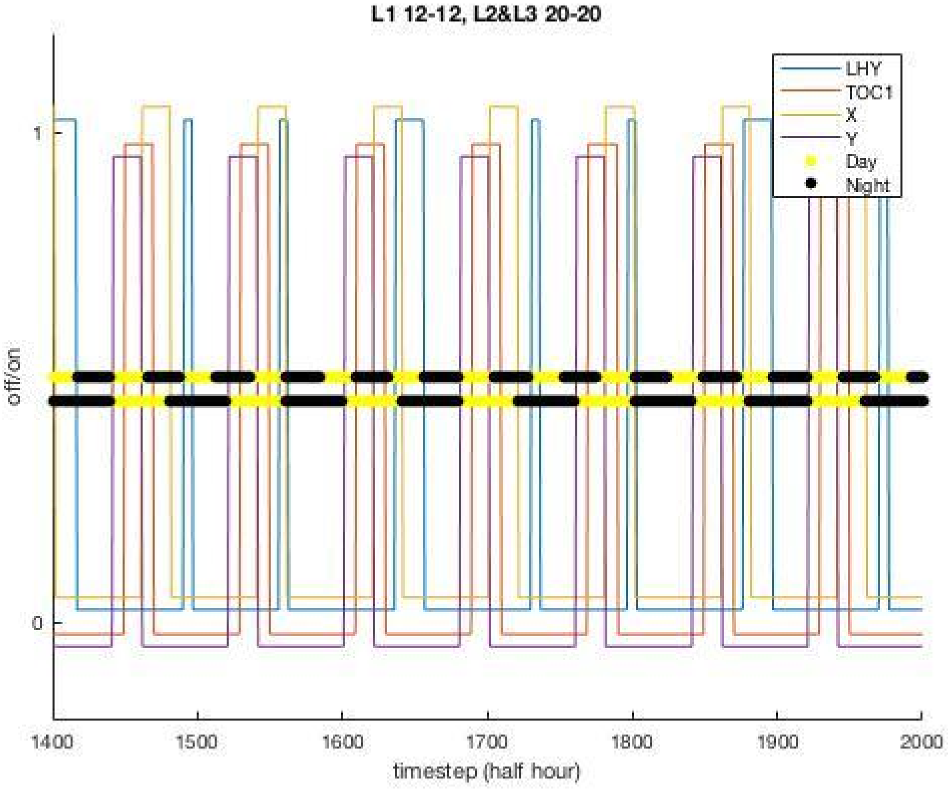


24 hr + 48 hr lights = 48 hr period



24 hr + 36 hr lights = 72 hr period

Asynchronous Photoperiods - Results



- Occam's Razor
- The resulting period of the model is the LCM of the two photoperiods
- Here 24 and 40 hour periods produce 120 hour periodic behavior

Asynchronous Lights - Conclusions

- Clearly the sun doesn't run asynchronously
    - Are there any biological implications here?
- Different colors of lights
    - Different photoreceptors may result in differing signaling cascades
    - Trigger different genes, impact different parts of circadian network
- Genetic knockout
    - For lights off, corresponds to something wrong with the gene/signaling pathway for that 'light'

Asynchronous Lights - Conclusions

- The model was very robust to changes in light behavior, even rather ridiculous changes
    - ‘Robust’ = The model produces expected AND biologically plausible results
- As opposed to photoperiod changes which produced unwanted and non-biological results even for photoperiods which weren’t that dramatic
- We expected the model to crumble quickly under these more aggressive changes, especially after the photoperiod experiments
    - After examining the behavior of the model in many cases, the behavior of gene Y has the most to do with the behavior of the model. As long as Y is turning on and off it can exert control over the network

Wrap-Up Thoughts

- Why did we not fix these issues?
  - In most cases, the problematic behavior of the model was way more entwined with actual intended behaviors of the model than we actually thought when we started
    - Mostly due to the fact that the model size is such that changing an edge or logic gate changes an appreciable amount of the model's total parameters
  - In short, producing a model that could handle both the standard cases where the current model functions, as well as the edge cases where it doesn't, would require a significant investment in understanding and potentially overhauling the model
  - The intended behavior of the model under duress in some cases is unresearched, so we have no frame of reference for what a 'fixed' model would act like

Wrap-Up Thoughts

- Where do we go from here?
  - The paper on circadian rhythms also proposes a 3-loop model for *Arabidopsis*
    - Perhaps this model is more robust to the changes we saw in the 2-loop model
  - More biological research
    - A lot of the scenarios tested here are not found in the literature
    - Without comparison to biological data, it is difficult to decide if our model is producing biologically relevant results
- Isn't the point that you don't need the data?
  - Yes and no. If the goal is to make a model for already researched processes, then having the data is a huge boon for confirming success
    - A certain number of 'confirmed' models are necessary before we begin to trust our unconfirmed models
  - Even if you are using data, we've seen that building boolean models can offer novel insight into regulatory pathways

Papers

Akman, Ozgur E., et al. “Digital Clocks: simple Boolean models can quantitatively describe circadian systems.” *Journal of the Royal Society Interface*. No. 9 (2012).

Kim, Junil, et al. “Discovery of a kernel for controlling biomolecular regulatory networks.” *Scientific Reports*. No 3. (2013).

Questions?