Optimising time-series experimental design for modelling of circadian rhythms: the value of transients data

Laurent Mombaerts*, Alexandre Mauroy**, Jorge Goncalves*

** Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg (e-mail: laurent.mombaerts@uni.lu; jmg@uni.lu)*

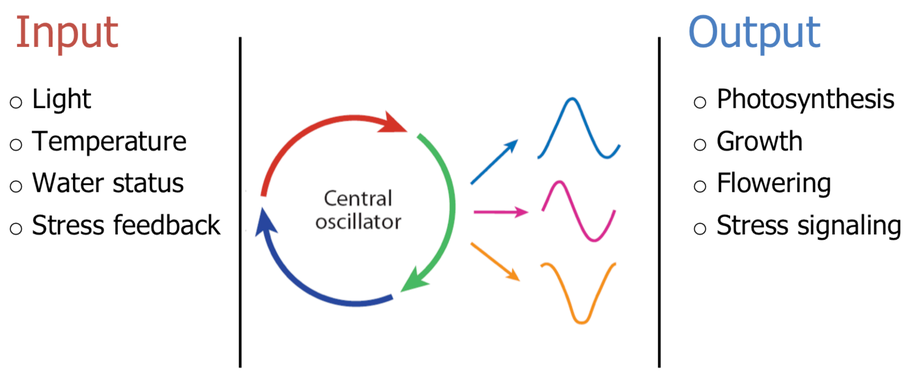
*** University of Namur (e-mail: alexandre.mauroy@unamur.be)*

Introduction and Background

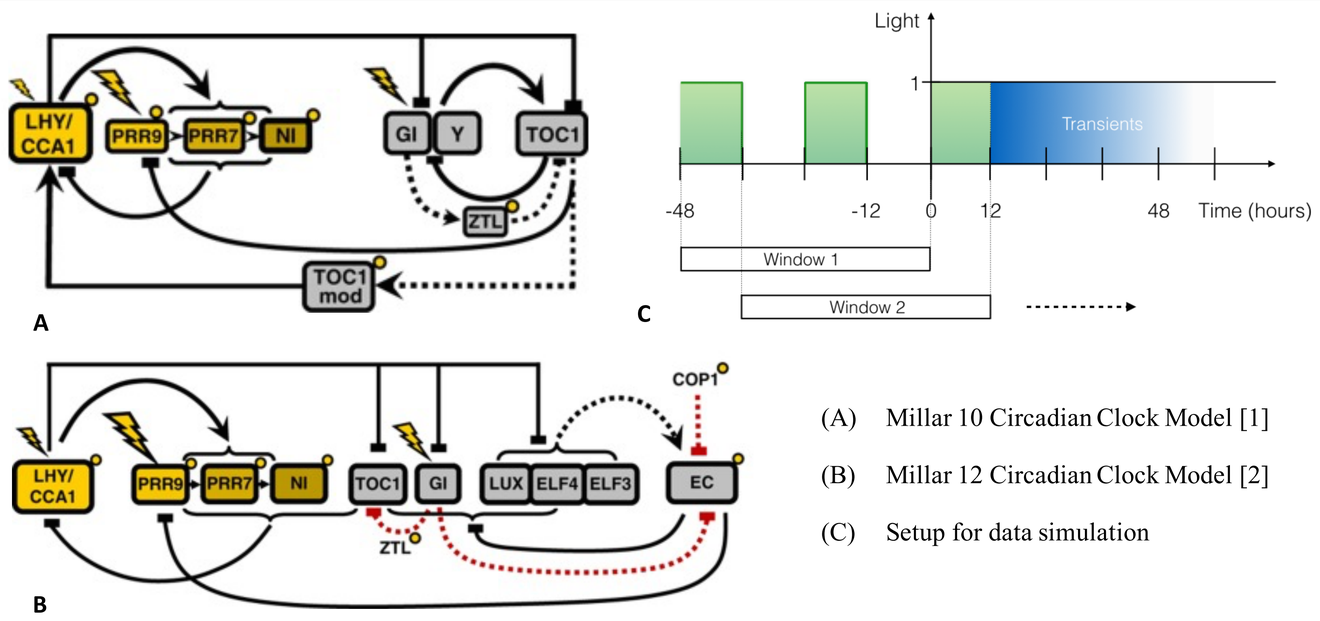
Circadian clocks consist of complex gene regulatory networks (GRN) that coordinate the daily cycle of most organisms. Most of the clock-oriented experiments are performed under constant photoperiodic regime, overlooking the transitory regime that takes place between light/dark cycles and constant light or darkness.

This project:

- Simulates realistic circadian clock time-series data
- Compares performances of network inference resulting from different photoperiodic regimes



Data Simulation



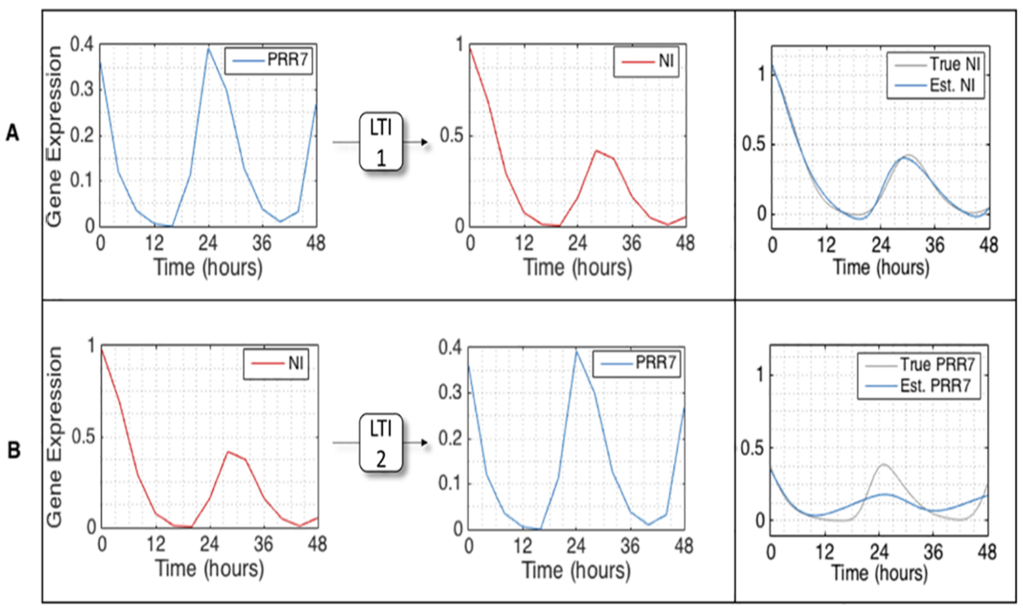
(A) Millar 10 Circadian Clock Model [1]

(B) Millar 12 Circadian Clock Model [2]

(C) Setup for data simulation

Methods

Overview of the network inference tool modus operandi [3]:



(A&B) LTI systems are identified [4] between each pairwise selected time-series.

C

$$\frac{dx}{dt}(t) = Ax(t) + Bu(t) + Ke(t)$$
$$y(t) = Cx(t) + Du(t) + e(t)$$

$x \in R^n$ = states variables
e(t) = noise
u(t), y(t) = input, output

(C) Mathematical expression of a LTI system.

D

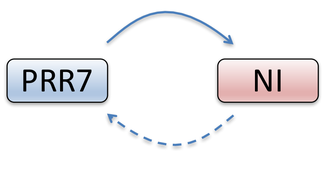
$$\text{fitness} = 100 * \left(1 - \frac{\sum_{k=1}^{N}(y_k - \hat{y}_k)^2}{\sum_{k=1}^{N}(y_k - \bar{y})^2}\right)$$

y_k = kth snapshot of true output
$\hat{y}_k$ = kth snapshot of estimated output
$\bar{y}$ = true output average

(D) Performance index of each LTI system.

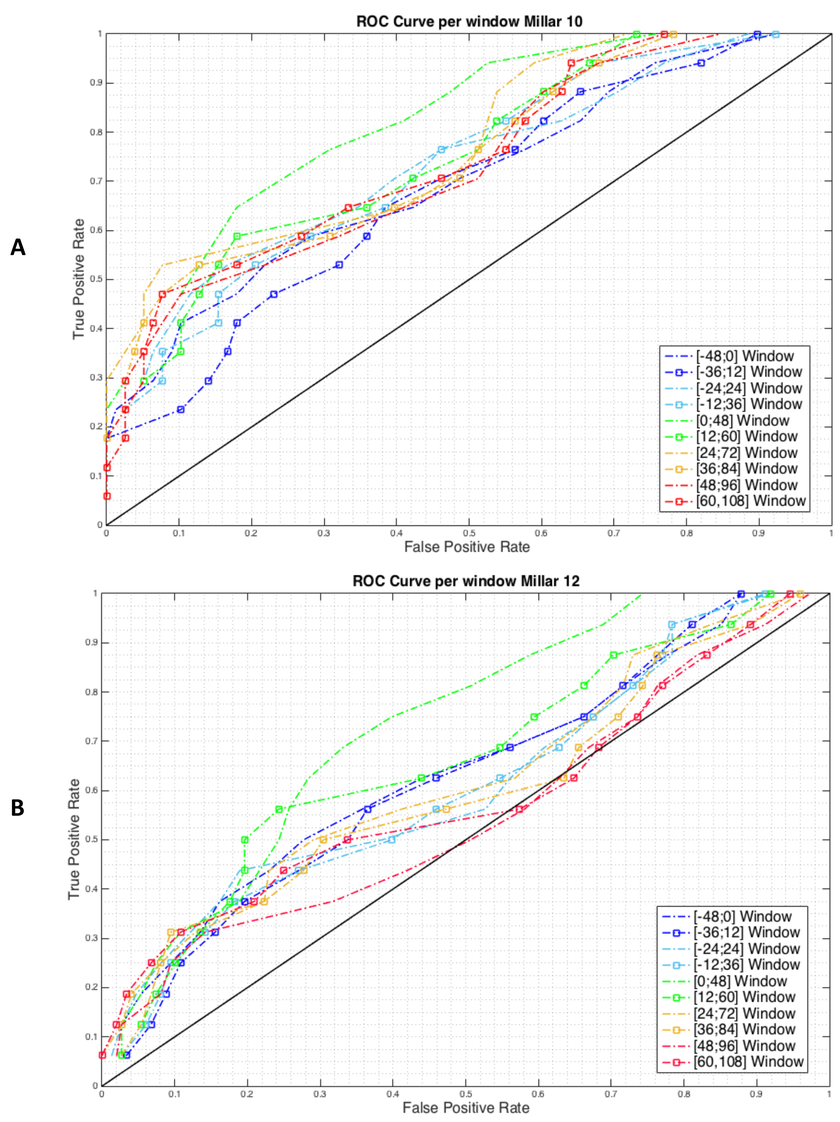
E

I\O	PRR7	NI
PRR7		90
NI	32	



(E) Example of the fitness table corresponding to (A&B). This result suggests that PRR7 regulates NI, which is not reciprocal.

Results



C

Window (hours)	AUROC Millar 10	AUROC Millar 12
-48 to 0	0.70[0.65 - 0.74]	0.64 [0.61 - 0.68]
-36 to 12	0.68 [0.66 - 0.70]	0.62 [0.60 - 0.64]
-24 to 24	0.73 [0.70 - 0.76]	0.62 [0.58 - 0.64]
-12 to 36	0.73 [0.71 - 0.75]	0.61 [0.59 - 0.63]
0 to 48	0.82 [0.78 - 0.84]	0.73 [0.70 - 0.75]
12 to 60	0.75 [0.74 - 0.76]	0.66 [0.63 - 0.68]
24 to 72	0.77 [0.75 - 0.79]	0.62 [0.58 - 0.64]
36 to 84	0.74 [0.73 - 0.77]	0.60 [0.57 - 0.62]
48 to 96	0.73 [0.70 - 0.75]	0.55 [0.52 - 0.58]
60 to 108	0.74 [0.72 - 0.77]	0.59 [0.55 - 0.63]

(A&B) Receiver Operating Characteristic (ROC) curves for both Millar 10 & 12 models. The number of true positives corresponds to the sum of the correctly inferred causal relationships of the network while the number of false positives corresponds to the sum of non-existing links that were wrongly inferred.

(C) Area Under ROC (AUROC)

Conclusion

The results suggest that the dynamics involved in transitory regime provide richer and more informative data sets, and improve dynamical modelling, in comparison to the other regimes.

References:

[1] Pokhilko, A., Hodge, S.K., Stratford, K., Know, K., Edwards, K.D., Thomson, A.W., Mizuno, T., and Millar, A.J. (2010). Data assimilation constrains new connections and components in a complex, eukaryotic circadian clock mdel. Molecular systems biology, 6.

[2] Pokhilko, A., Fernandez, A.P., Edwards, K.D., Southern, M.M. Halliday, K.J., and Millar, A.J (2012). The clock gene circuit in arabidopsis includes a repressilator with additional feedback loops. Molecular systems biology, 8.

[3] Carignano, A., Yuan, Y., Dalchau, N., Webb, A.A.R., and Goncalves, J. (2014). Understanding and prediction biological networks using linear system identification. A Systems Theoretic Approach to Systems and Synthetic Biology I: Models and System Characterizations.

[4] Ljung, L., Prochzka, A., Uhlil, J. Rayner, P., and Kingsbury, N. (1998). System identification in signal analysis and prediction. Birkhuser Boston, 163-173.