What about data?

Most of classic modelling approaches in systems biology (5-20 years ago) make a **limited use of data**, mostly because there was not much usable data available back then.

Kinetic rates were inferred from dedicated experiments (in vitro) and by exploration of biological literature to make educated guesses.
This is a painstakingly time consuming and error prone process, impossible for large models.

With the data revolution we are living in, more and more experimental techniques are capable of producing data that is good to fit dynamic models.

Typically, one needs **time series data.**

Examples of such technologies are flow cytometry, RNAsec, imaging techniques…

Circadian Clock in O. Tauri

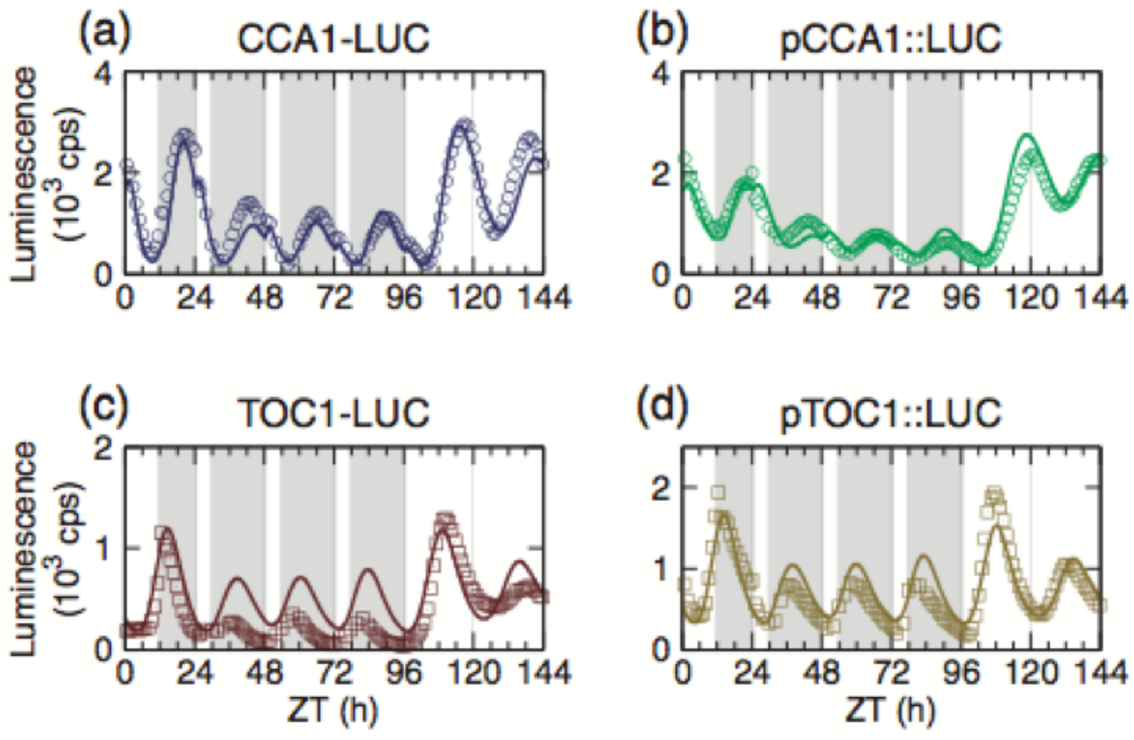
the plant journal

SEB Society for Experimental Biology

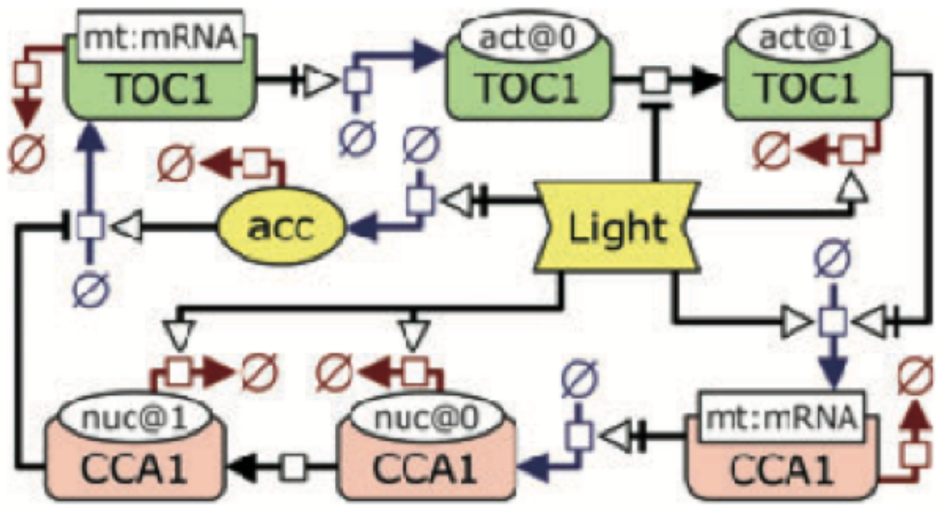
The Plant Journal (2011) **66**, 375–385

doi: 10.1111/j.1365-313X.2011.04489.x

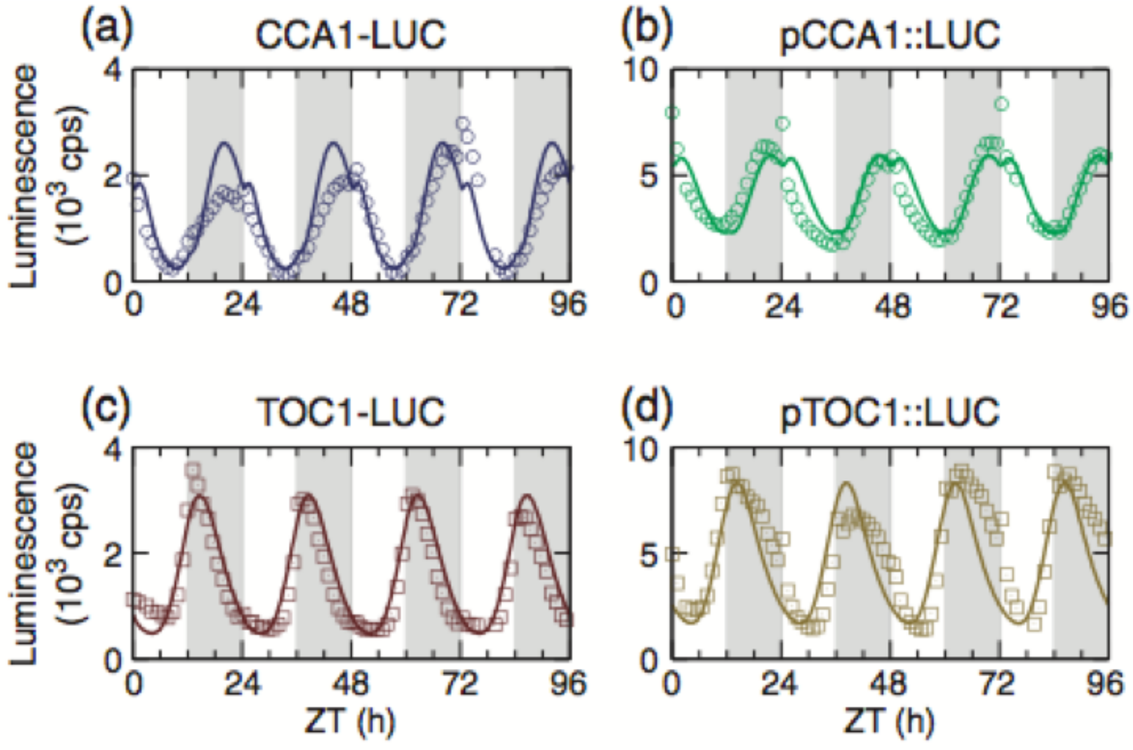
Multiple light inputs to a simple clock circuit allow complex biological rhythms



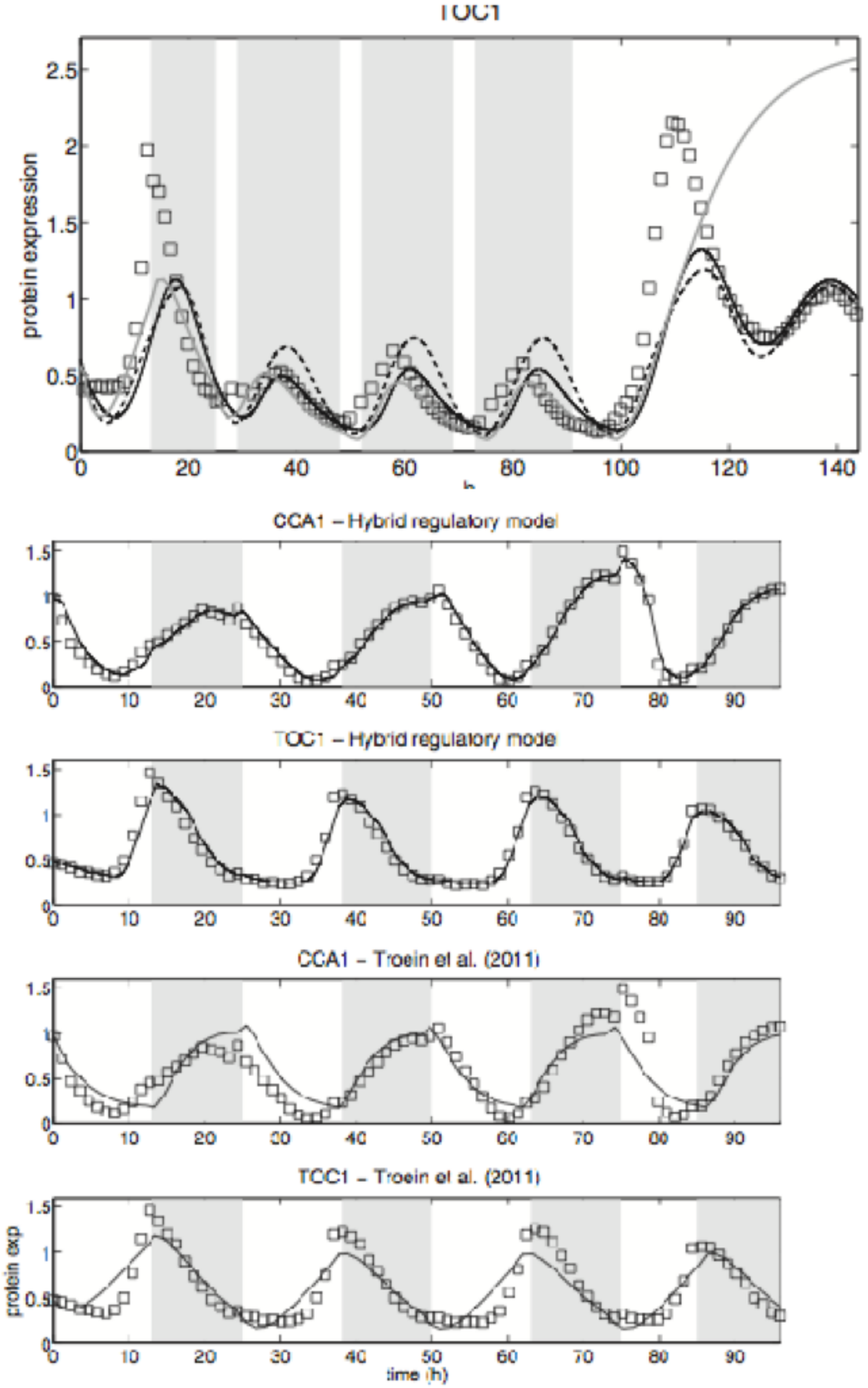
Data: luciferase time series, both transcriptional (LUC attached to CCA1 and TOC1), and translational (LUC attached to promoters of CCA1 and TOC1),



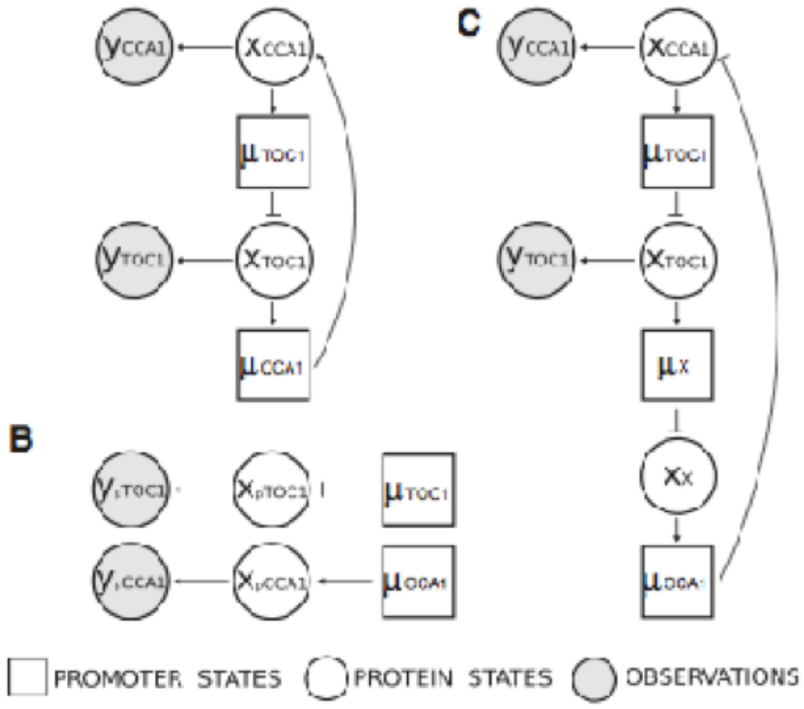
ODE model with 7 variables/ species



Circadian Clock in O. Tauri

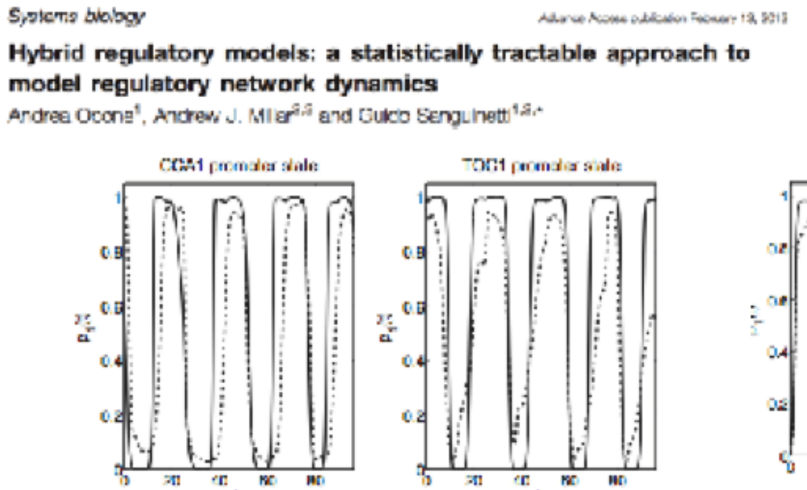


When the ODE model is trained on 12:12 LD data, it fails to predict the behaviour of irregular light patterns.



A switching diffusion stochastic model (2 species) can predict behaviour more accurately.

New protein required to correctly capture the behaviour of both transcriptional and translational data.



Take home messages

Modelling can help **elucidating the role and functioning** of cellular components.
Multi scale modelling can deal with tissues, organs, and so on.
It also tests if current knowledge is **consistent**.

Modelling large scale systems (e.g. whole cell) can provide a cheap **in silico experimentation environment** (e.g. for drug testing)

Modelling is a **key enabling technology** in **synthetic biology**: it allows cheap and fast exploration of the design space.

Modelling requires **time-series data** to estimate model parameters. High quality data is required for proper model identification.