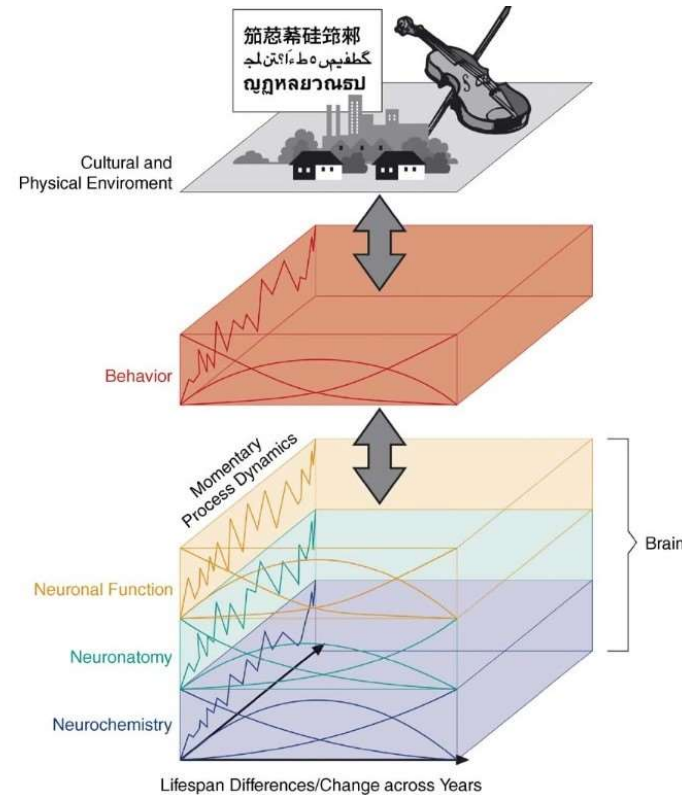# Continuous Time Systems – Generative Models and the Infinite Rabbit Hole

Charles Driver

May 2021

Motivation

- The world is complex, our fancy models are massive simplifications!

- The statistical modelling frameworks we use inevitably guide our thinking.

- Thinking harder about how the processes actually behave can help interpreting and critiquing our models.

- Using more flexible / appropriate modelling frameworks helps (sometimes forces!) us to to do this harder thinking.

- Lindenberger, Li, & Bäckman, (2006), neatly demonstrates some of the complexity, for typical cases in developmental psychology / neuroscience.

Overview

- Dynamic systems

  - General
  - Discrete time
  - State space
  - Continuous time

- Mapping statistical models to theories of change

  - Problems of the discrete time representation
  - Non-linearity -- time and subject heterogeneity

- Estimation

- ctsem software

Dynamic Systems

- Most things scientists are interested in are 'dynamic systems'.

- A dynamic systems *model* is then a formal representation of such.

    - Simplest model of change / 'dynamics'? Probably t-test!

    - Does ice cream make people happy?

        - 2 groups, experimental and control.
        - Give experimental group ice cream.
   - Ask about happiness.

- Perhaps a little more nuance is needed?
  - *When* do people become happier?
  - For how *long*?
  - What does happy really *mean* in this case?
  - Does the effect change with temperature, flavour, amount, age,...?

Starting to build...

- Mapping fuzzy theory to dynamic system highlights gaps, helps ensure coherent, testable, incrementally improvable theory.

- Theory exploration tool -- how do systems interact / generate dynamics, what needs to be specified that was otherwise implicit?

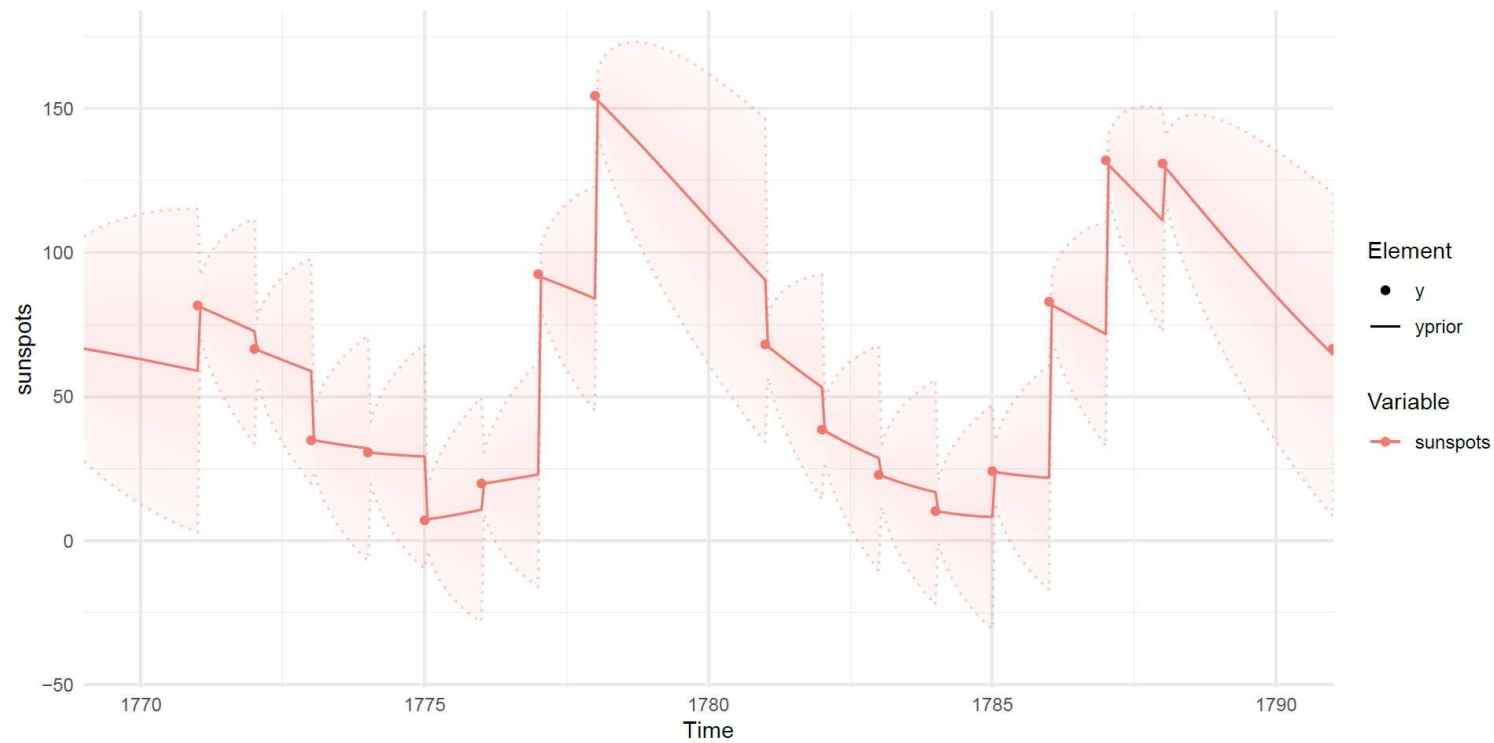- Theory quantification / testing -- fit model to data.
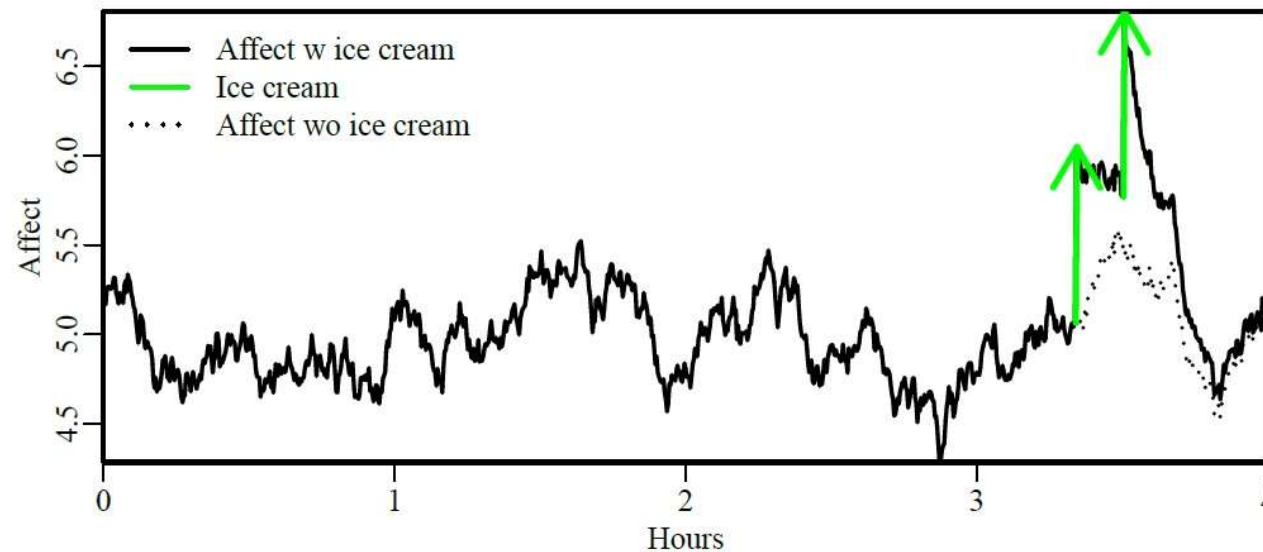
- With a few extra pieces, we get more interesting models that can tell us what may happen in a broader range of circumstances…
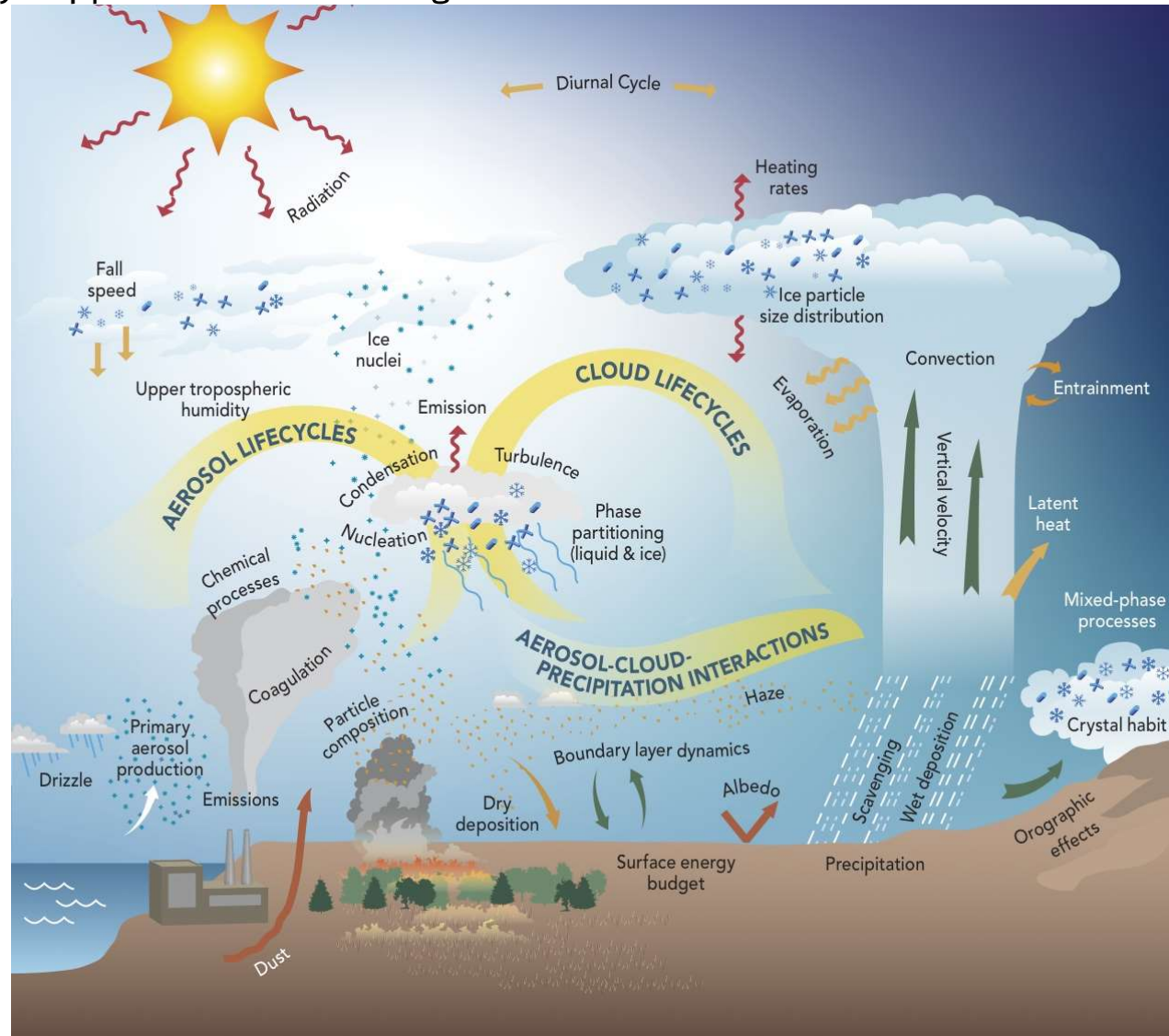
How do we start building?

- Interrelated systems

- Measurement

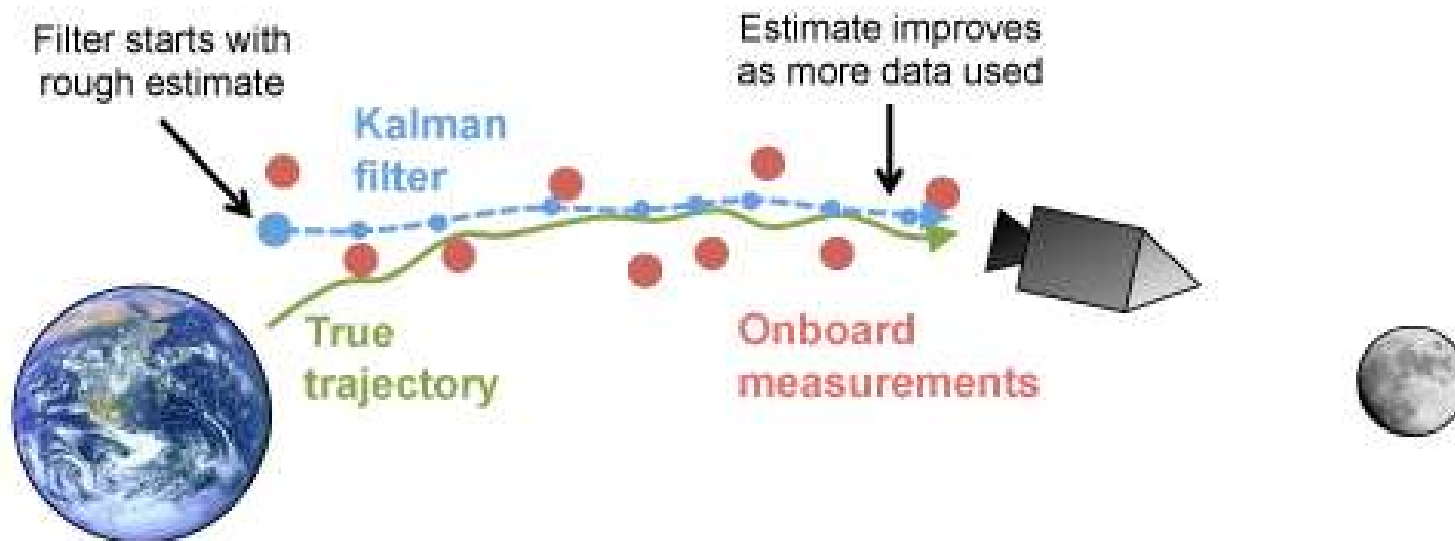- Likely complex dependencies between the two!

Why 'time' and 'dynamics' ?

- Why should we care about how a system develops over time -- isn't time just like any other variable?

- Time is a proxy for 'all the stuff in the universe that tends to happen'.

  - You eat the ice cream, start digesting, chemicals start bouncing around a little differently inside you, behaviour changes slightly, other people respond a little differently, and you continue experiencing more of the 'stuff of the universe happening near you'.

  - A model for how something changes over time, is then a general model for how this thing tends to change when we don't do anything specific.

Example uses?

- Non psychology:

    - Weather and climate

    - Finance

    - Landing on the moon

    - What happens when we shine two lasers at each other?

Filter starts with rough estimate

Kalman filter

Estimate improves as more data used

True trajectory

Onboard measurements

Forecasting

- Dynamic systems approaches provide excellent scope for predictions of the future. Such predictions could be based on:

- Earlier time points -- where do we think the red line will go?


Time
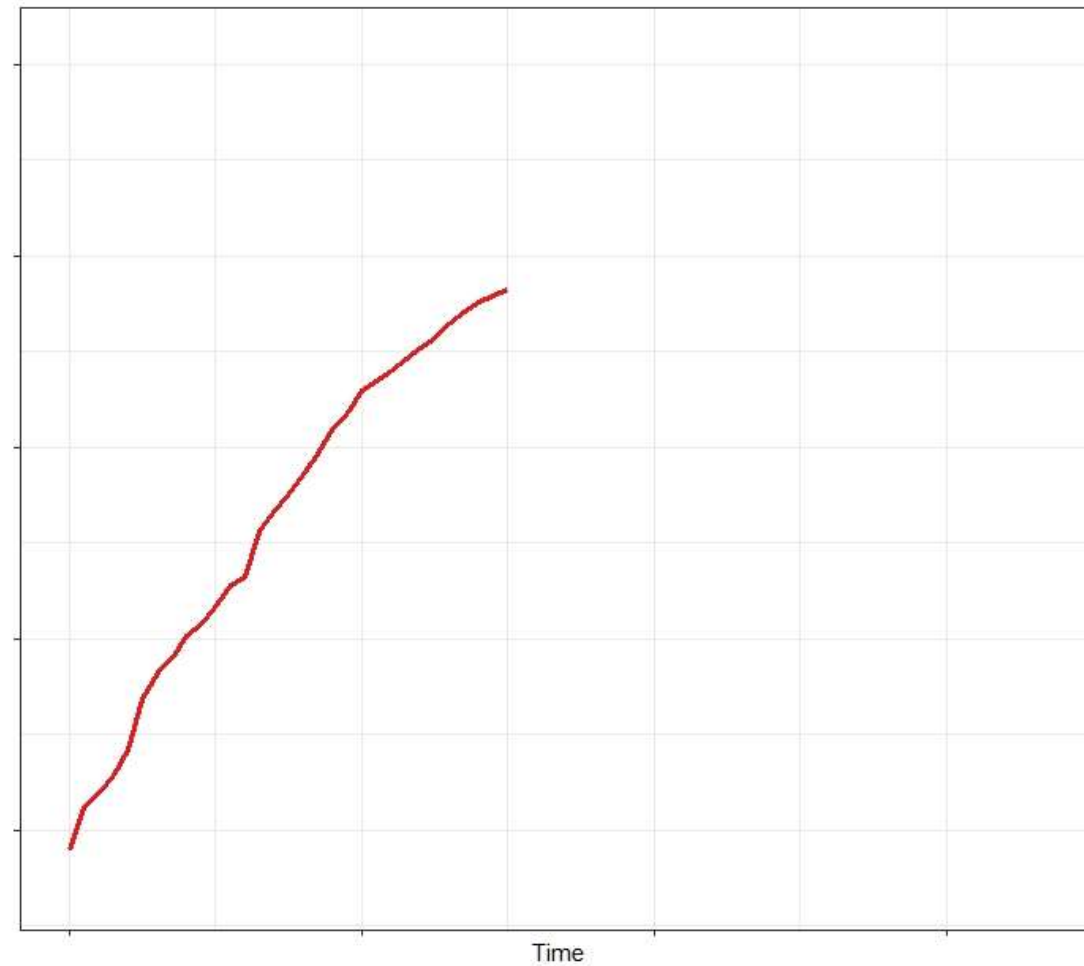
Forecasting 2

- In psychology, given that there are always similarities between people (compared to say, a person and a block of cheese), we can also leverage knowledge of how other subjects behave -- in this case we are probably more confident of where the red line will go now.
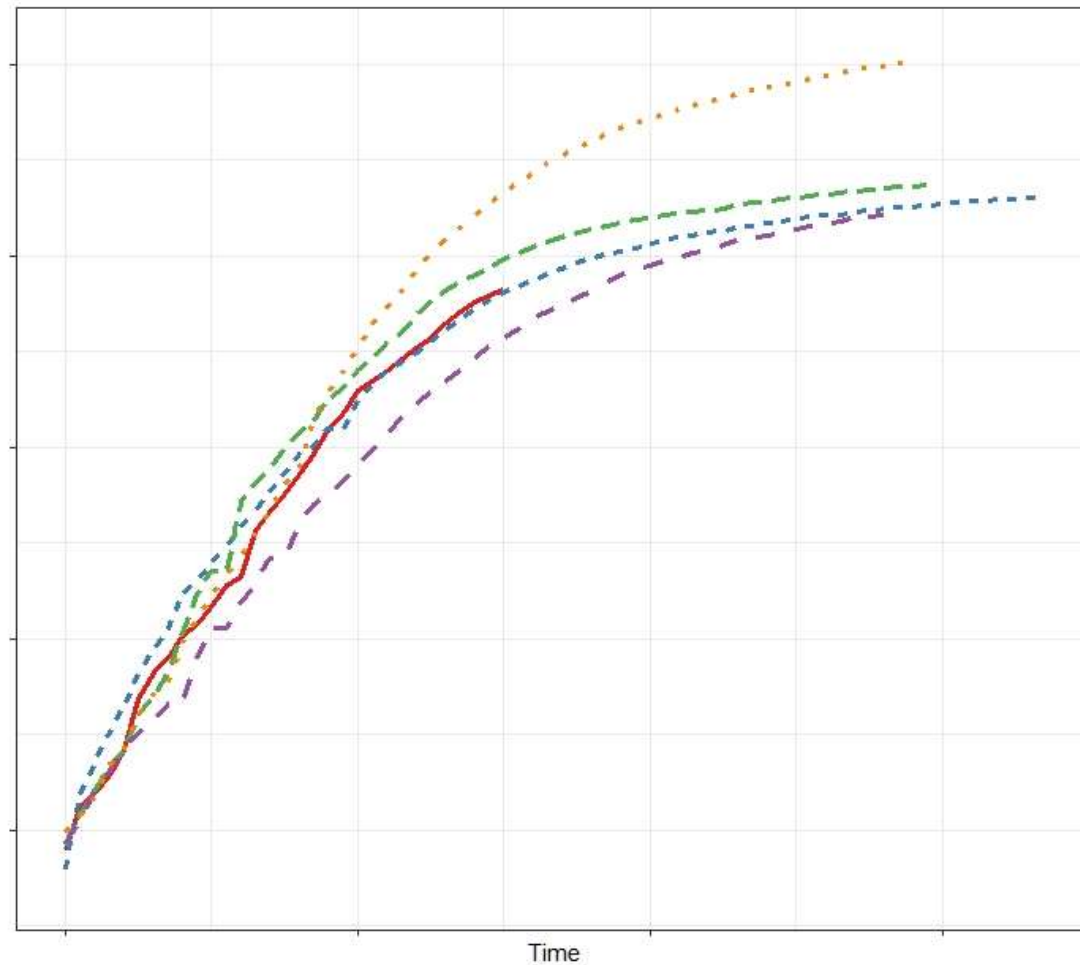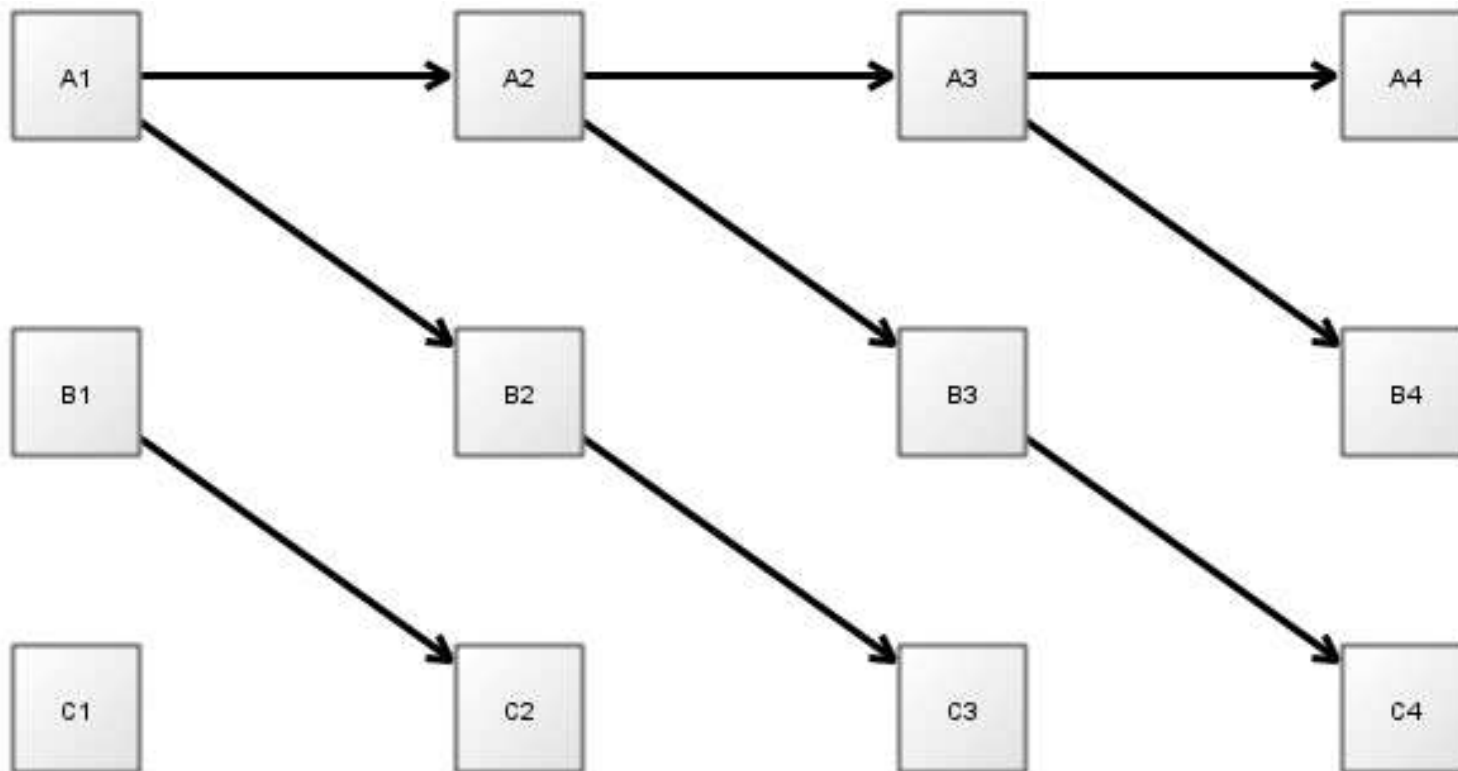
Forecasting 3

- Such knowledge of other subjects behaviour, and a specific subjects past, can lead to improved predictions of the future.
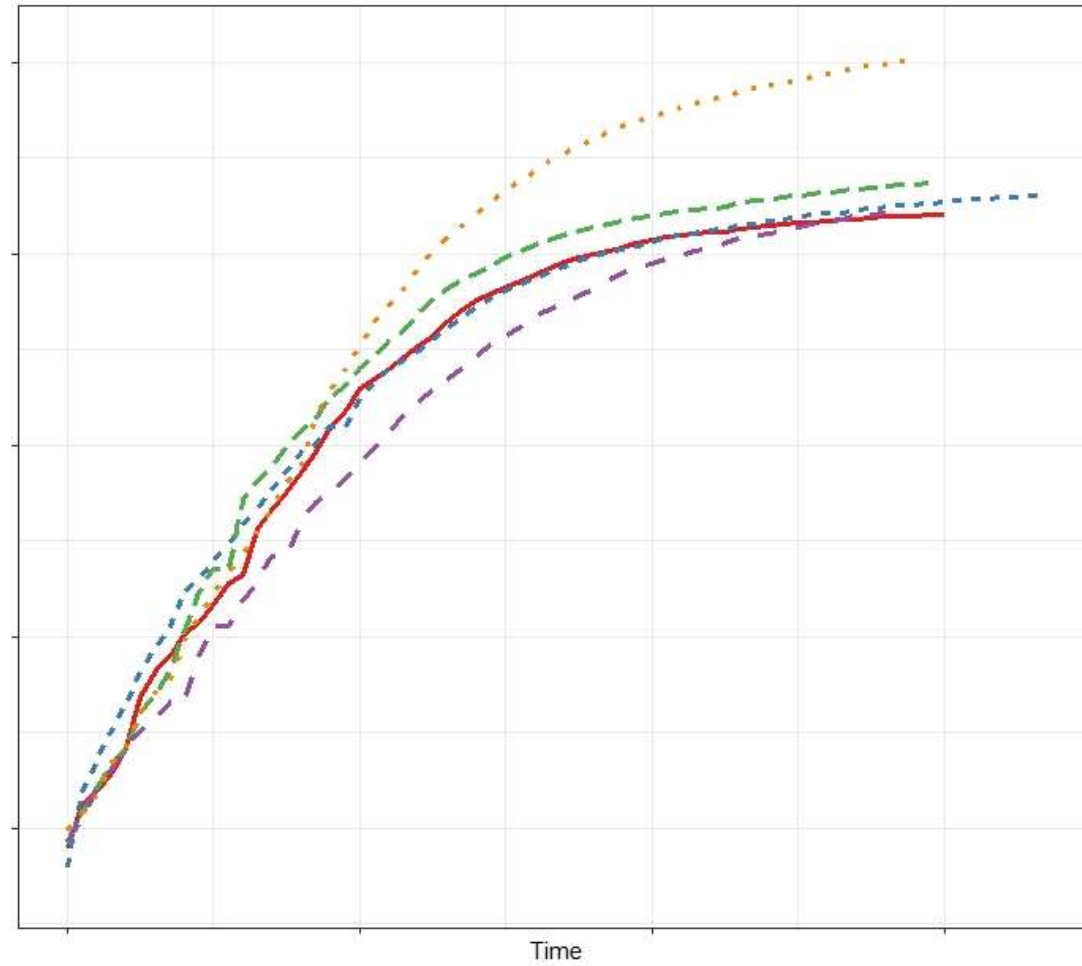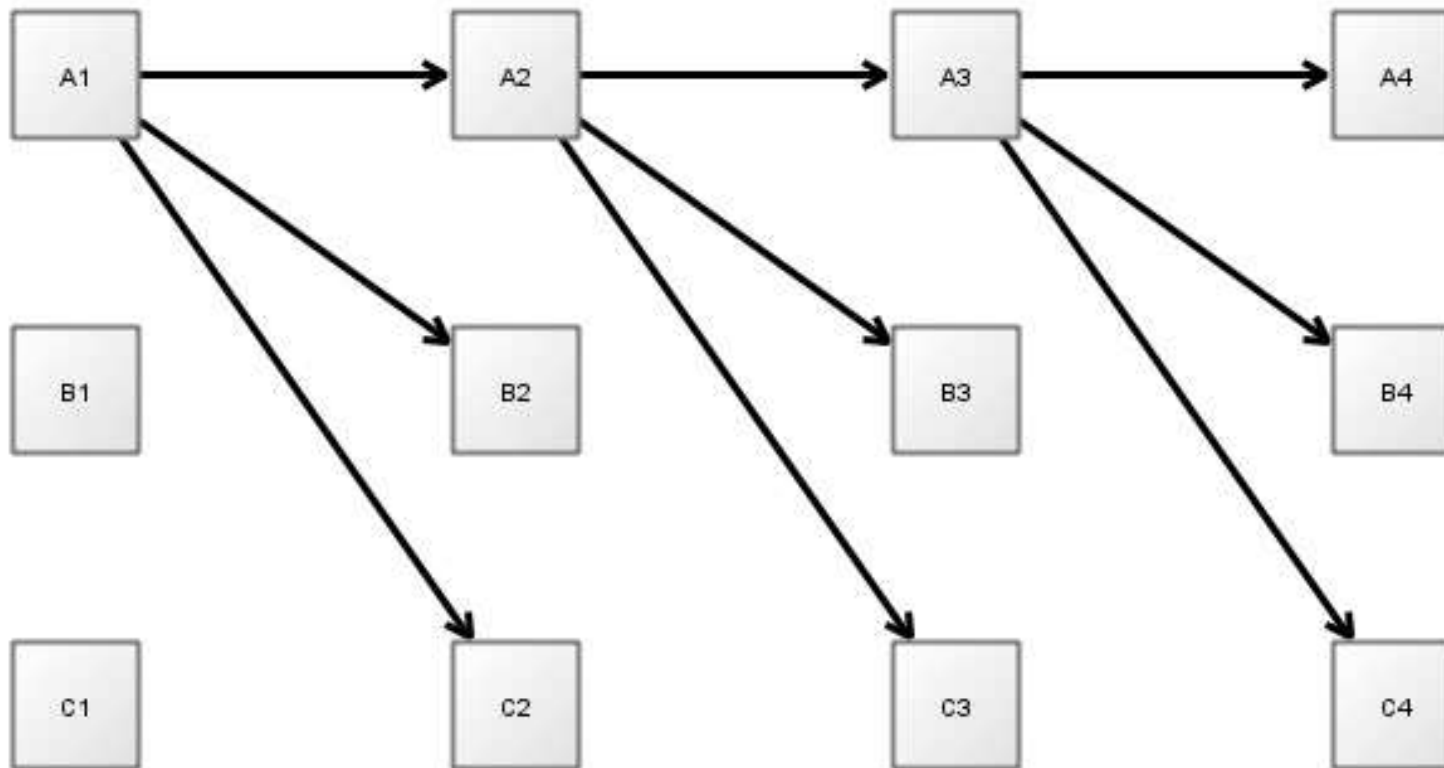
The model - reality link

- Models that reflect underlying processes not that important for prediction, but crucial for inference and understanding of structure.

- The 2 day observation frequency model does just fine for prediction, but if we use the model to infer anything about reality, we get it badly wrong.

# But... I think my process changes all the time?

Structural equation modelling (SEM)

- + Nice software.
- + Easy missing data handling.
- + Nice simple recipes for model checking.
- - Usually not well designed for lots of repeated measurements.
- - Either restrictive or highly complex to specify genuine causal models over time.
- - Very inflexible in terms of anything non-linear, interactions, etc.
- - The 'nice simple recipes' strong assumptions are generally not checked, difficult to check.
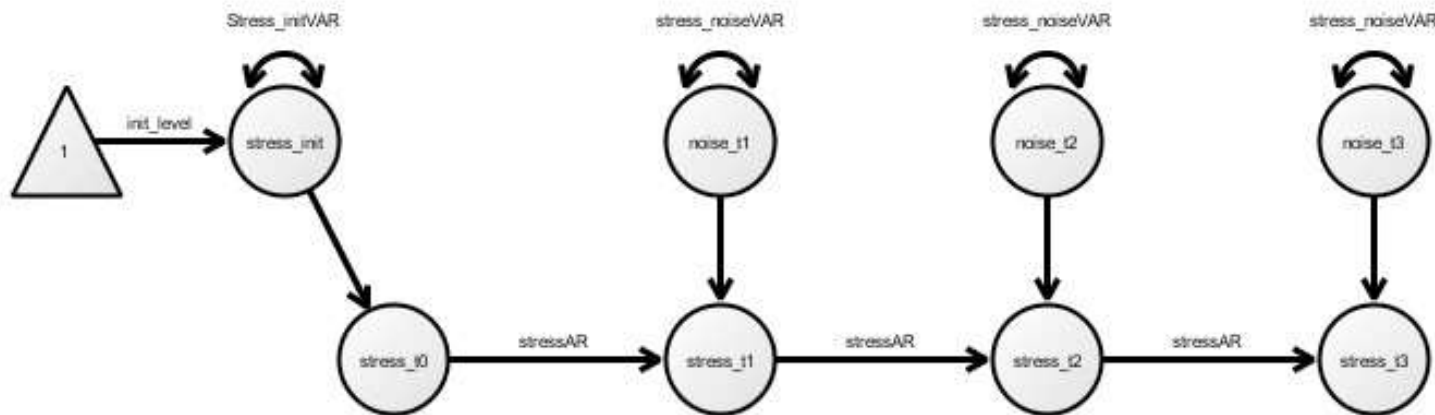
SEM -- time interval

- Spacing represents length of time.
- Dynamic relationship parameter (e.g. autoregressive effect) is inherently a 'discrete time' form.
- Representing effect over different time intervals using same parameter results in some weighted mixture of 'true' parameters.
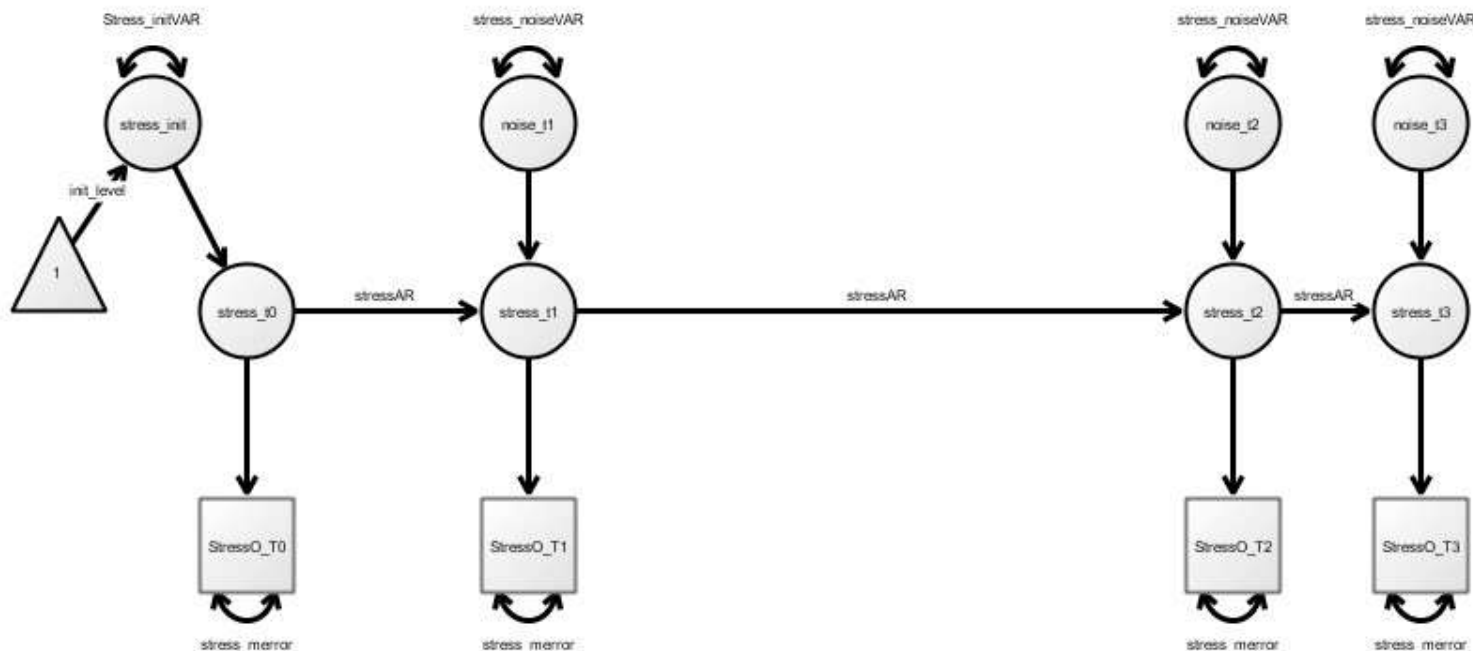
Continuous time

- Instead of specifying temporal regressions directly, we need to specify and estimate the underlying continuous time, or differential equation, system.
- Solves the time interval problem, and the direct effect problem.
- Classic discrete time form:
  - $x_u = Ax_{u-1} + b + g\xi_u$
- Stochastic differential equation (continuous time) form:
  - $\dot{x}_t = Ax_t + b + g\xi(t)$
- Discrete time regression coefficient *based* on ct model:
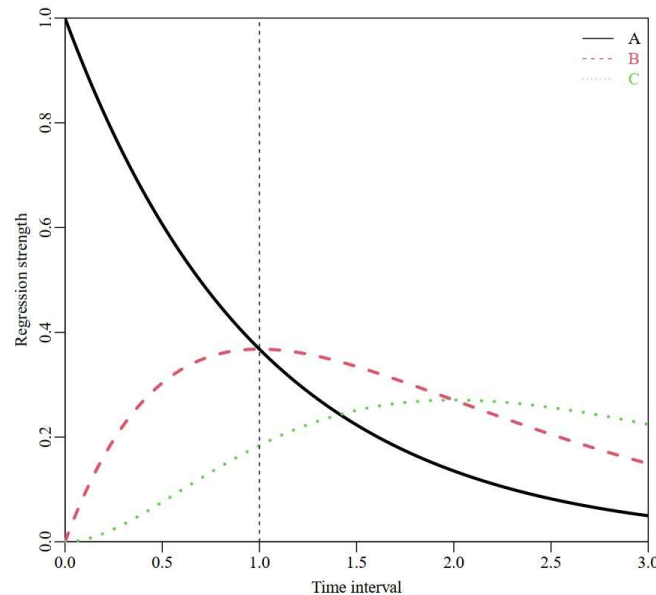
Differential equations

$$\dot{y}(t) = ay(t) + b$$

- Approx. solution
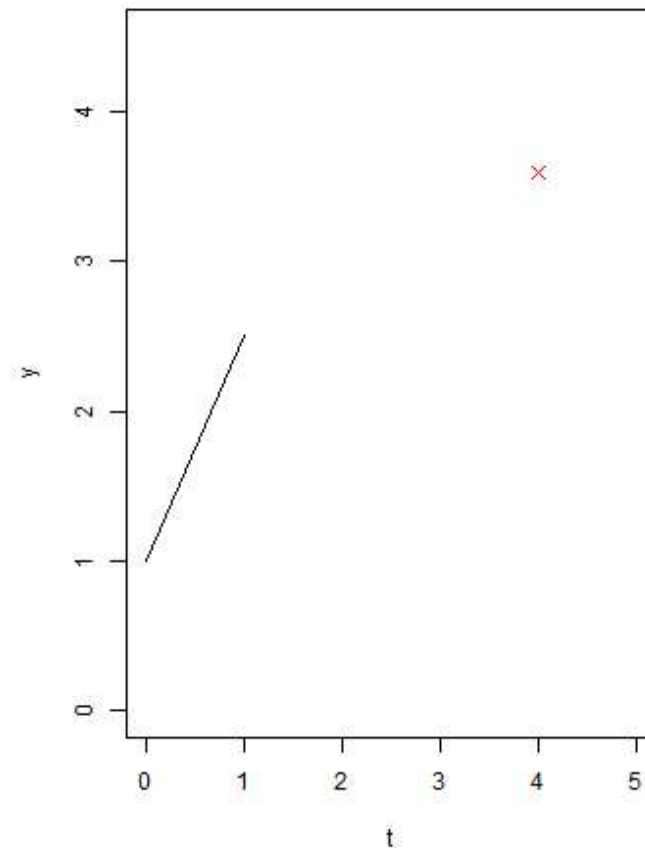 $$y_u = y_{u-1} + ay_{u-1} + b$$

- Exact solution:
  $$y(t) = e^{at}y0 + (e^{at} - 1) * (1/a) * b$$

```
y0 = 1
t= 4
b= 2
a = -0.5
y= exp(a * t) * y0 +
   (exp(a*t)-1)*(1/a)*b
```
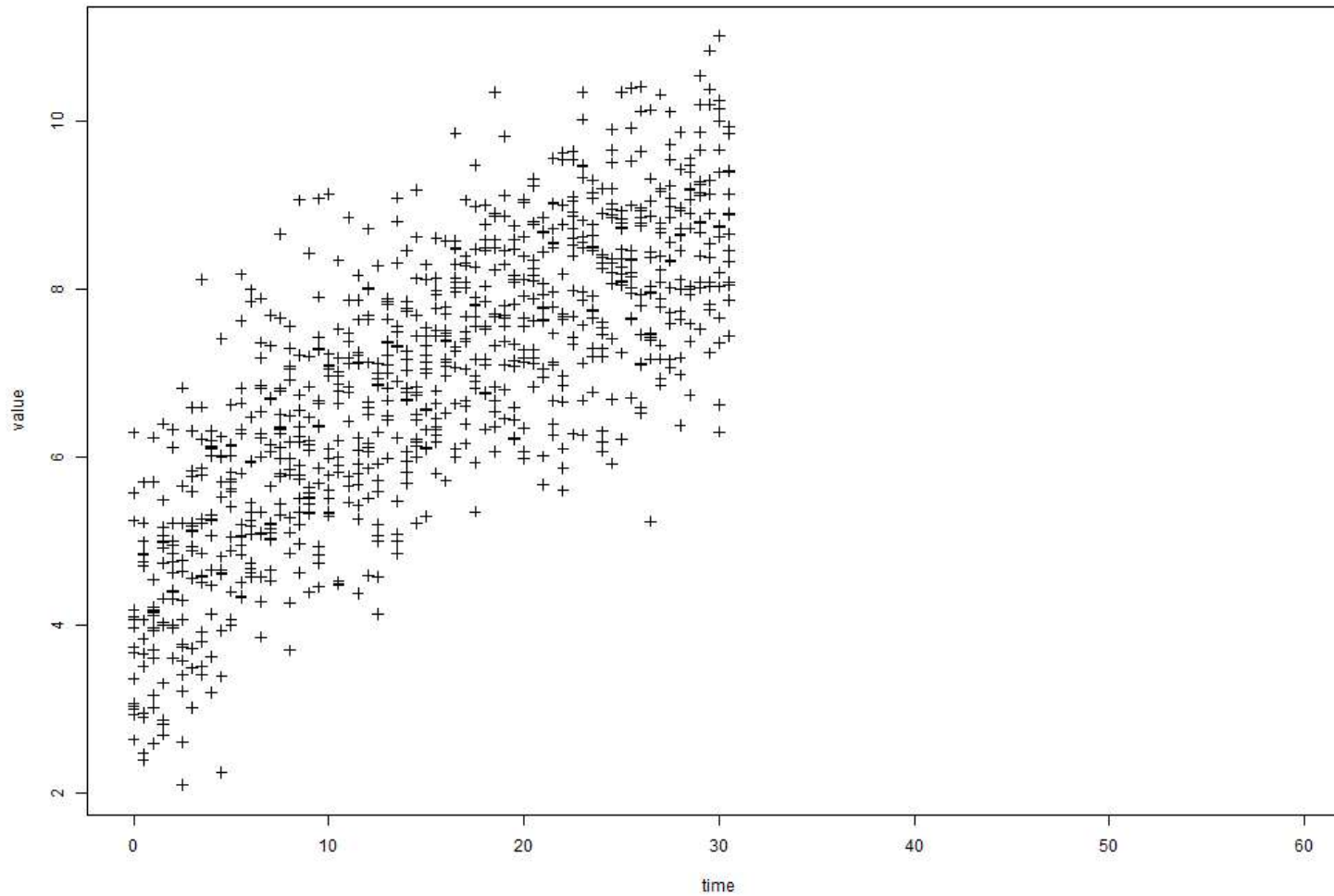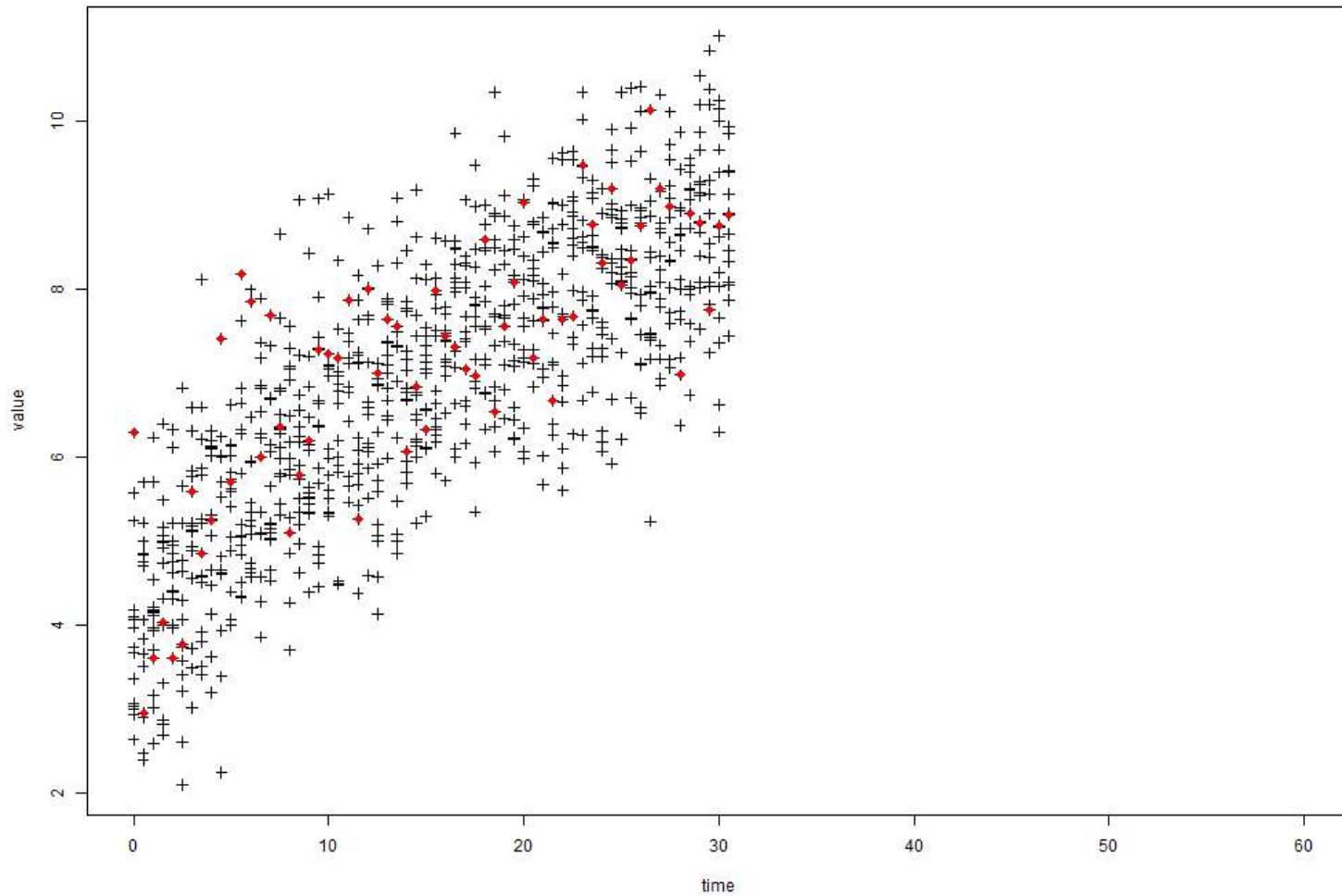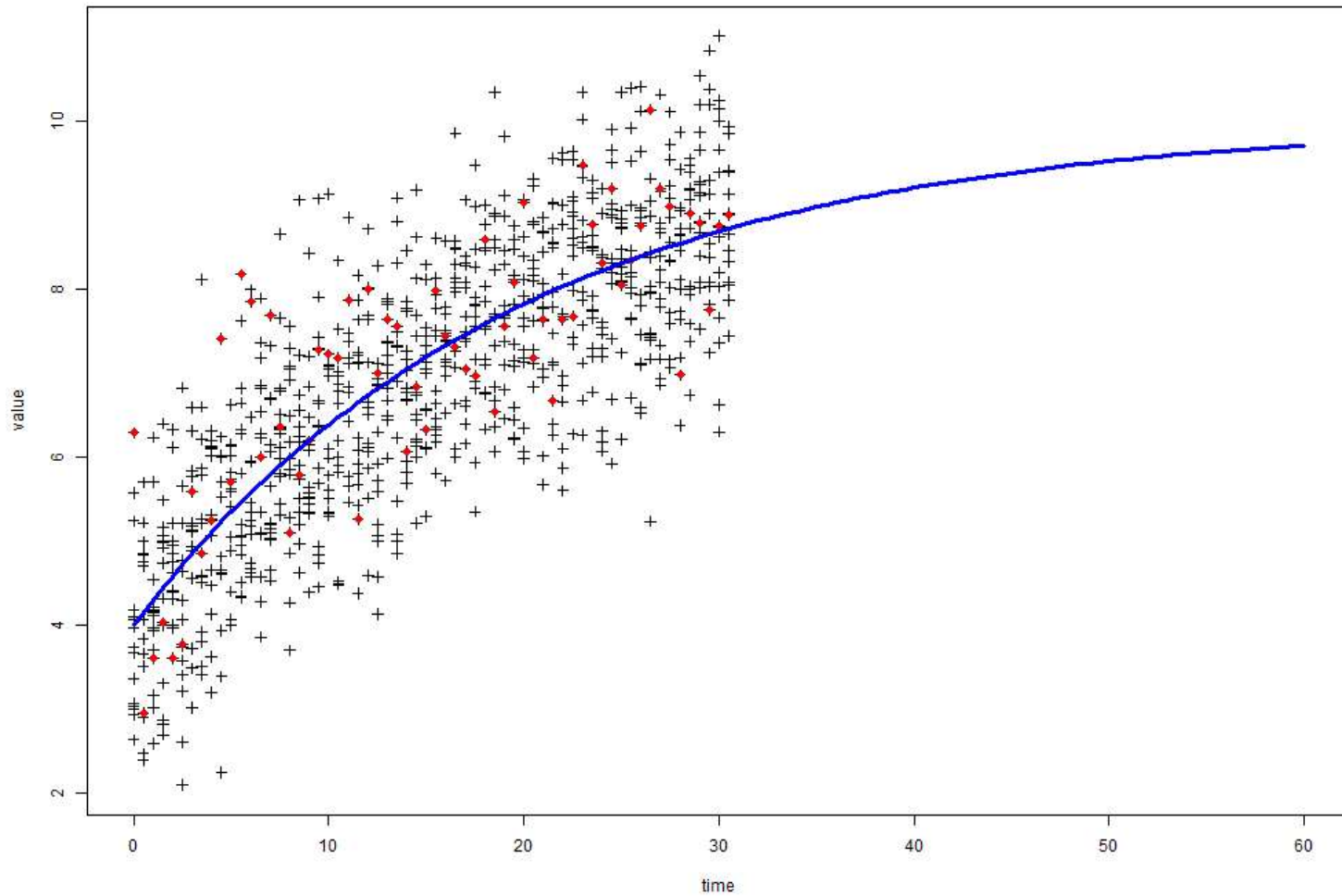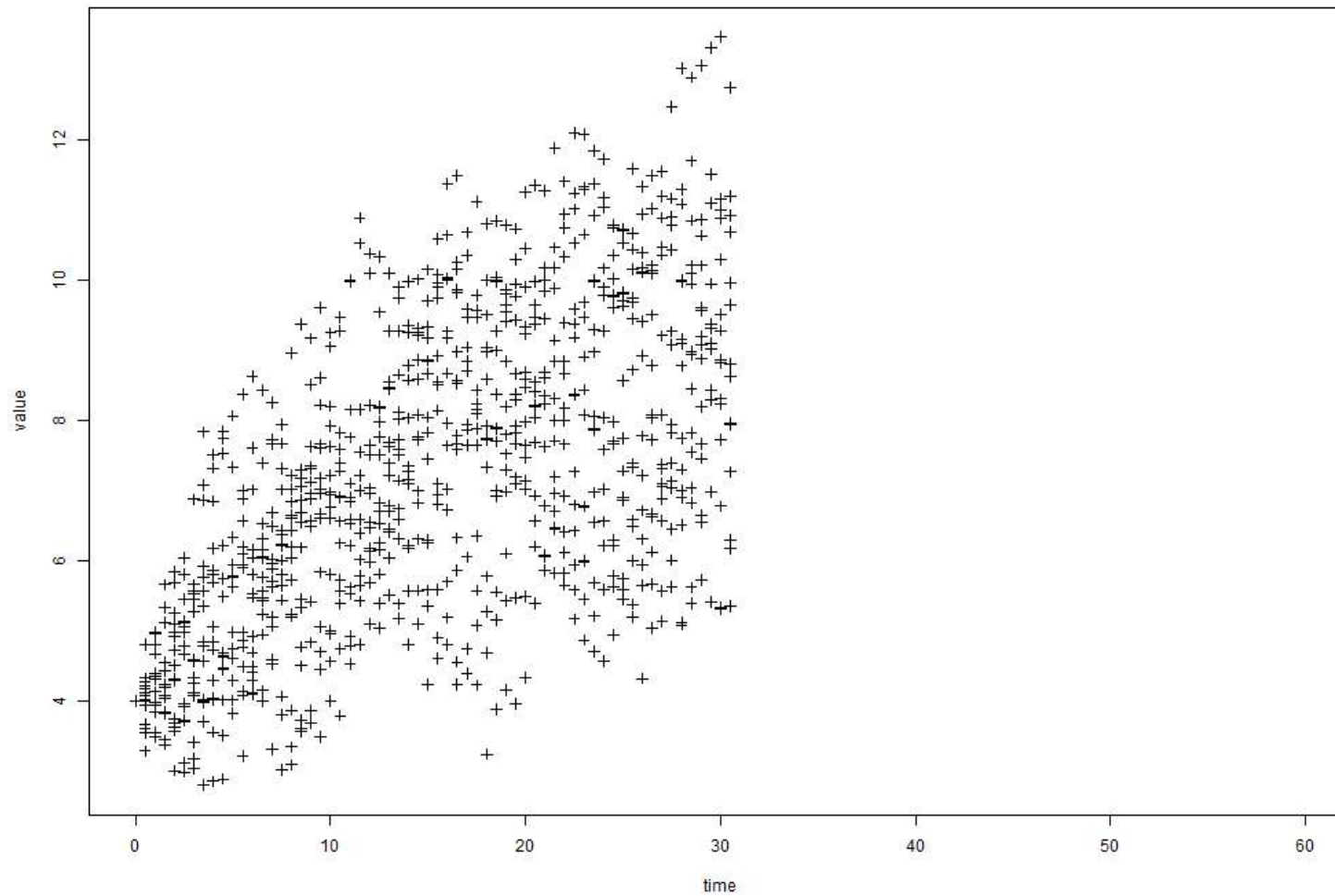
```
## [1] 3.593994
```

Uncertainty, fluctuations, error!
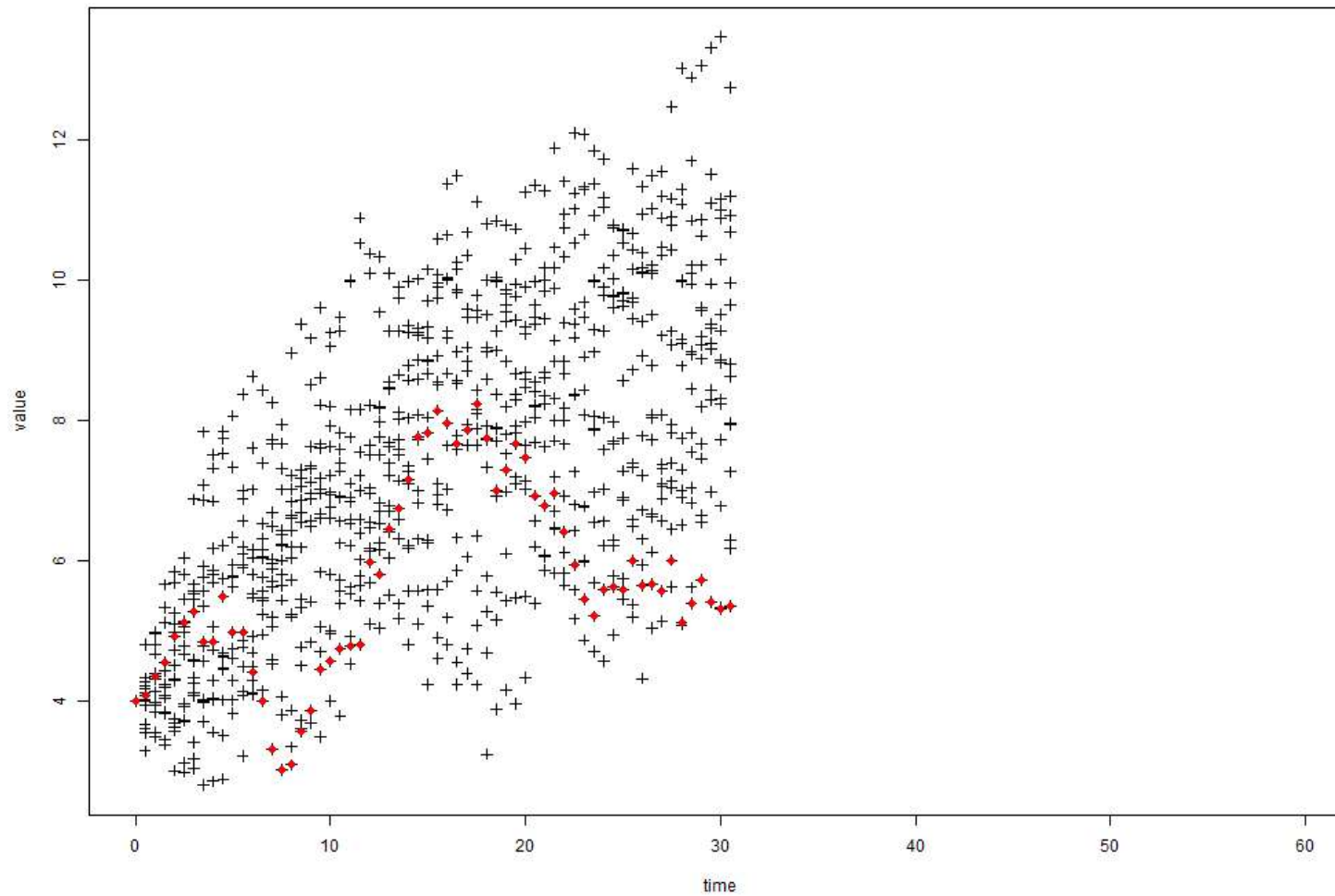
Uncertainty, fluctuations, error!
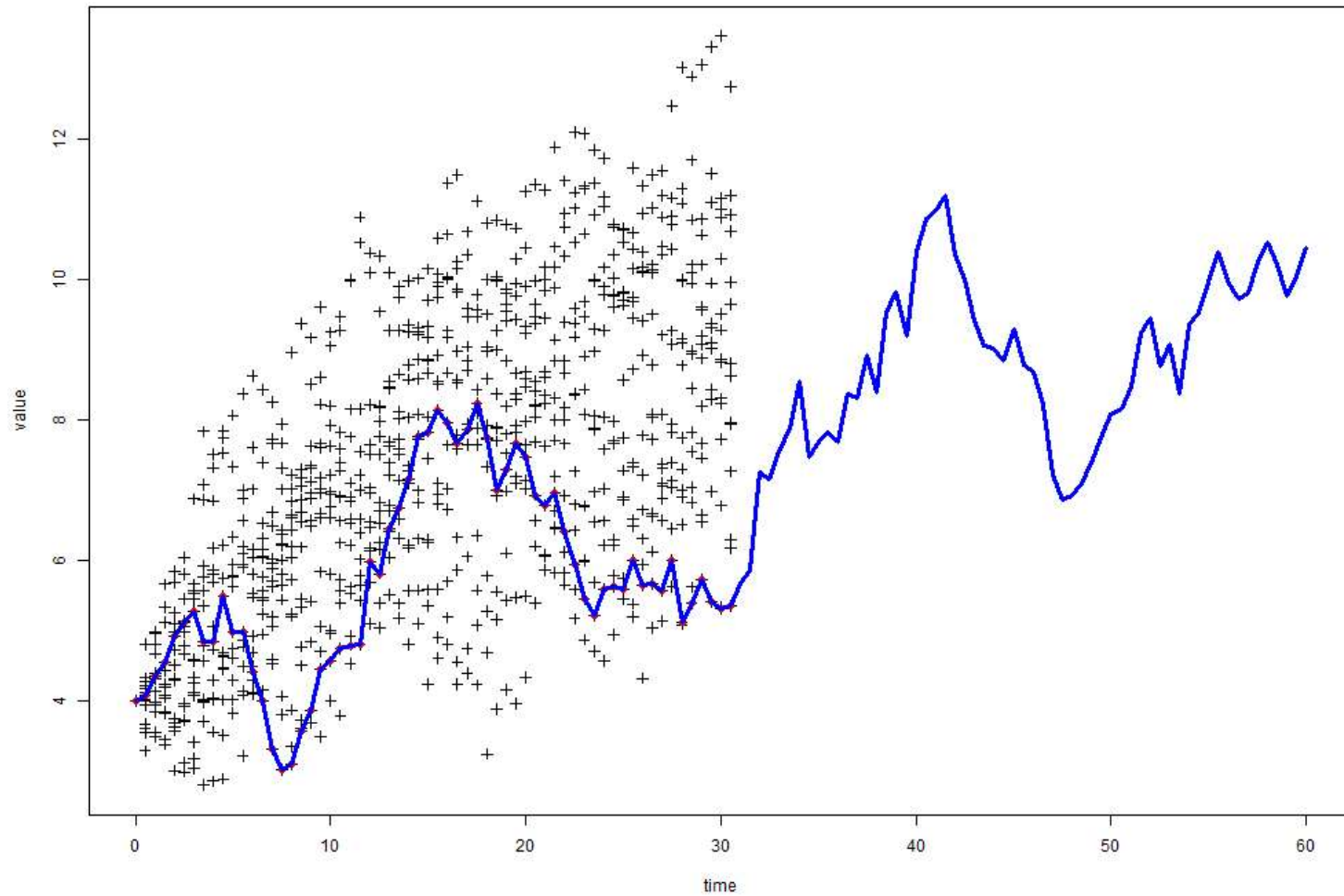
Uncertainty, fluctuations, error!

Uncertainty, fluctuations, error!

Uncertainty, fluctuations, error!

Uncertainty, fluctuations, error!

Different sorts of noise

- Noise is everywhere!

  - Sometimes, although our model couldn't predict the noise, including the noise as information can improve our predictions. This can be called process, or system, noise.

- Sources of such process noise can be things like:

  - genuine but unpredictable changes in our processes.
  - model misspecification.

- As we build better models, or include additional data, what we once had to accept as unpredictable can become predictable.
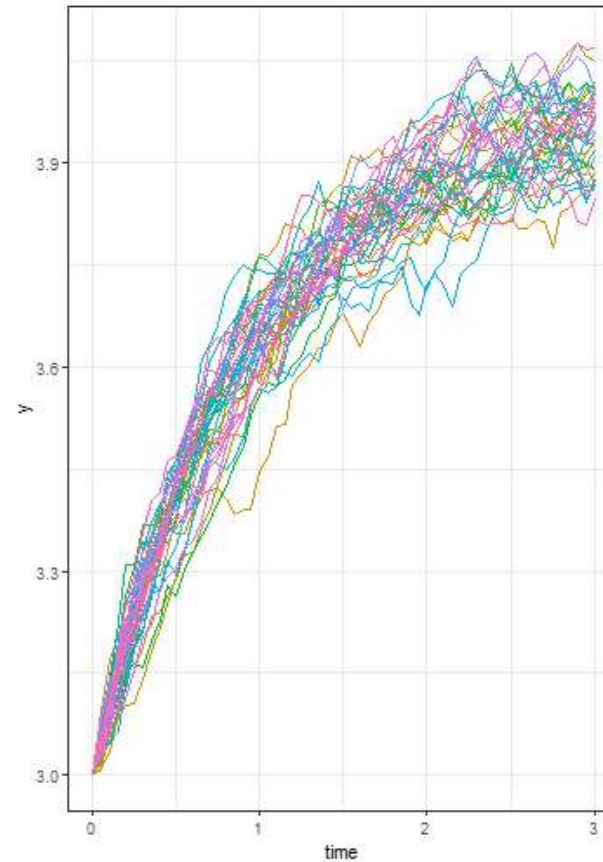
Process noise

- Ordinary differential equation:
  $\dot{y} = ay + b$

- Stochastic differential equation:
  $\dot{y}_t = ay_t + b + g\xi(t)$

- Heuristic form -- $\xi(t)$ represents Gaussian noise.

- Exact solution no longer exists!

- Expectations:

$$E(y_t) = e^{at}y0 + (e^{at} - 1) * (1/a) * b$$
$$cov(y_t, y0) = \frac{g^2}{2a}(e^{at} - e^{-at})$$

- "Euler method" for specific Δt:
  $y_u = y_{u-1} + (ay_{u-1} + b) * \Delta t + g\xi_u$
  $g\xi_u \sim N(0, \Delta t)$

Multiple processes

- How should we think of processes like this?

    - Alternative indicators of the same construct?
    - Different constructs strongly influenced by a common cause?
    - Different constructs where one or both are causing the other?

- For prediction, it often doesn't matter much!

- For scientific inference (ie causality) it is critical.
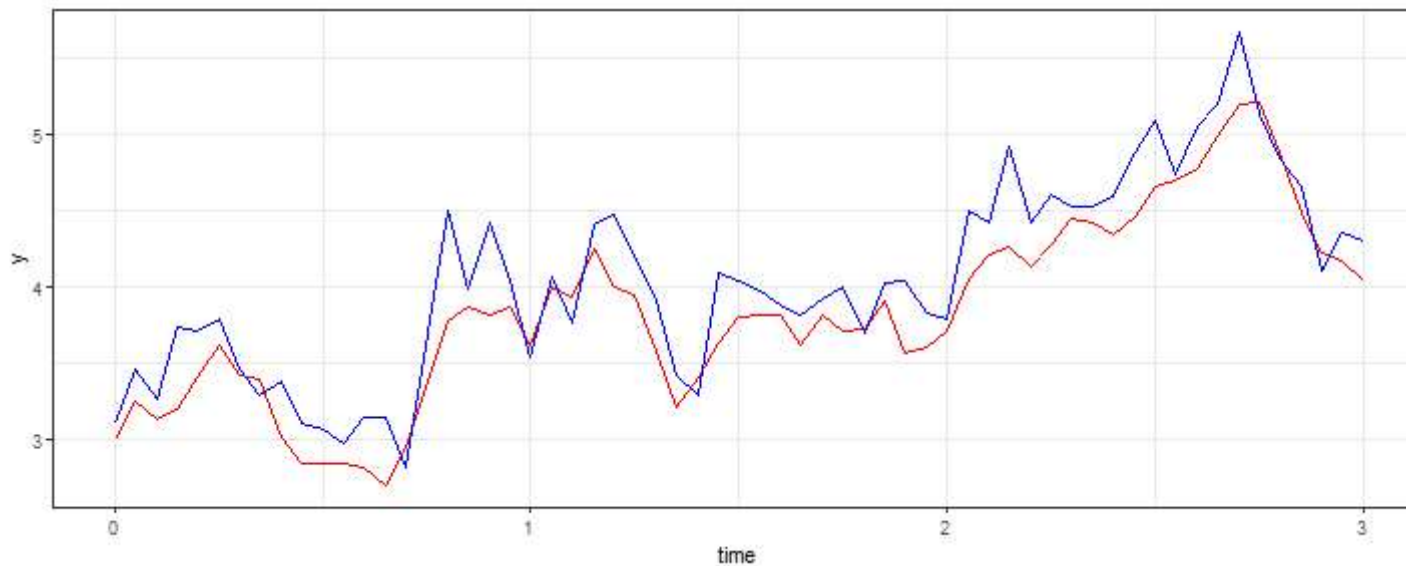
Critical, and often terribly done

GLOBAL WARMING VS PIRATES
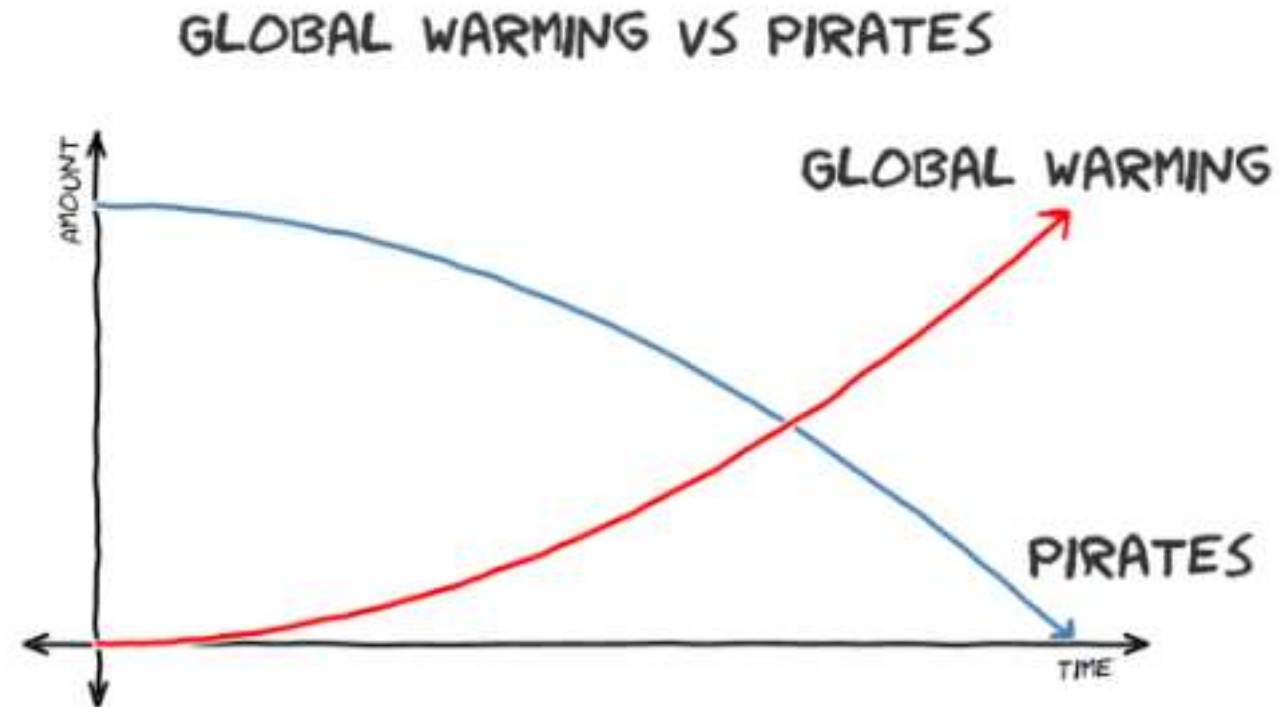
Pirates

State Space Representation

- Single latent process, 2 indicators:

Process change:
$$\underbrace{\mathrm{d}\,[\text{eta1}]\,(t)}_{\mathrm{d}\boldsymbol{\eta}(t)} = \left(\underbrace{[-1]}_{\mathbf{A}} \underbrace{[\text{eta1}]\,(t)}_{\boldsymbol{\eta}(t)} + \underbrace{[0]}_{\mathbf{b}} \right) \mathrm{d}t + \underbrace{\{[1]\}}_{\mathbf{G}} \underbrace{\mathrm{d}\,[W_1]\,(t)}_{\mathrm{d}\mathbf{W}(t)}$$

Observations:
$$\underbrace{\begin{bmatrix} \text{Y1} \\ \text{Y2} \end{bmatrix}(t)}_{\mathbf{Y}(t)} = \underbrace{\begin{bmatrix} 1 \\ 0.97 \end{bmatrix}}_{\boldsymbol{\Lambda}} \underbrace{[\text{eta1}]\,(t)}_{\boldsymbol{\eta}(t)} + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{\boldsymbol{\tau}} + \underbrace{\begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}}_{\boldsymbol{\Theta}} \underbrace{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}(t)}_{\boldsymbol{\epsilon}(t)}$$

Latent noise per time step:
$$\Delta[W_{j\in[1,2]}](t-u) \sim \mathrm{N}(0, t-u)$$

Observation noise:
$$[\epsilon_{j\in[1,2]}](t) \sim \mathrm{N}(0,1)$$

Heterogeneity

- Qualities of an observed process in a specific time and context might be reasonably represented by a relatively simple process.

- That same representation may not work well if applied to a different person, context, or time.

What *are* individual differences?

- Could be genuine, fixed differences, e.g. genes / country of birth.

- Or, could be differences in some slower changing process:

  - Differences between subjects looks fairly stable at shorter time scale.
 - But changes as we expand our time window.
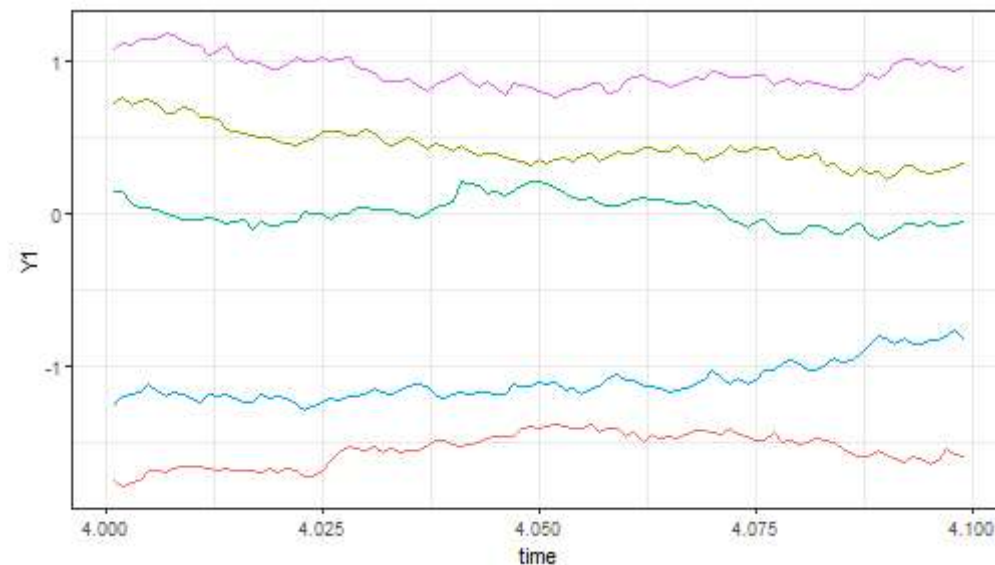
More on individual differences...

- Differences need not be in means, but in co/variances, temporal dependencies, intervention persistence, etc!

- 'Random' effects via Bayesian sampling ideal, but very slow for moderate to large models / data -- many dimensions, many parameters!

- Combining 'fixed' effects via covariates and random effects integrated out via Kalman filter is usually a tractable middle ground.

- With large models and many covariates, the line between machine learning and classical modelling starts to blur...

Interventions / shocks

- Random shocks that we can measure (e.g. time and magnitude of event) can be very useful for determining structure of the system. Here an intervention increases exercise motivation, which in turn increases fitness.
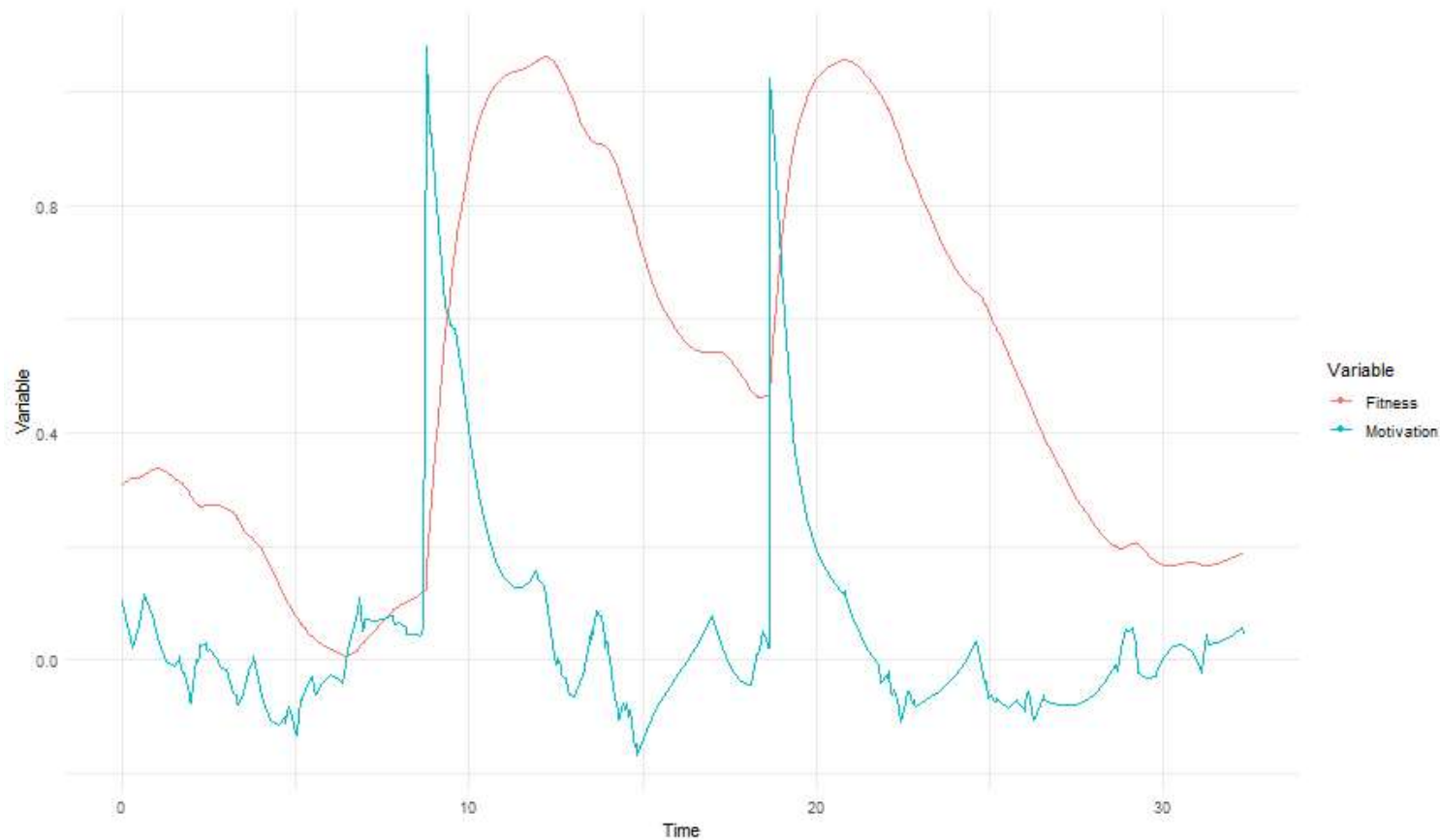
More complex changes...

- We could also think about processes or interventions that don't simply generate a shift up or down, but actually change parameters in the system.

- e.g. operant conditioning, CBT, shift to working from home...

- These take us into the realm of *non-linear* stochastic differential equations...
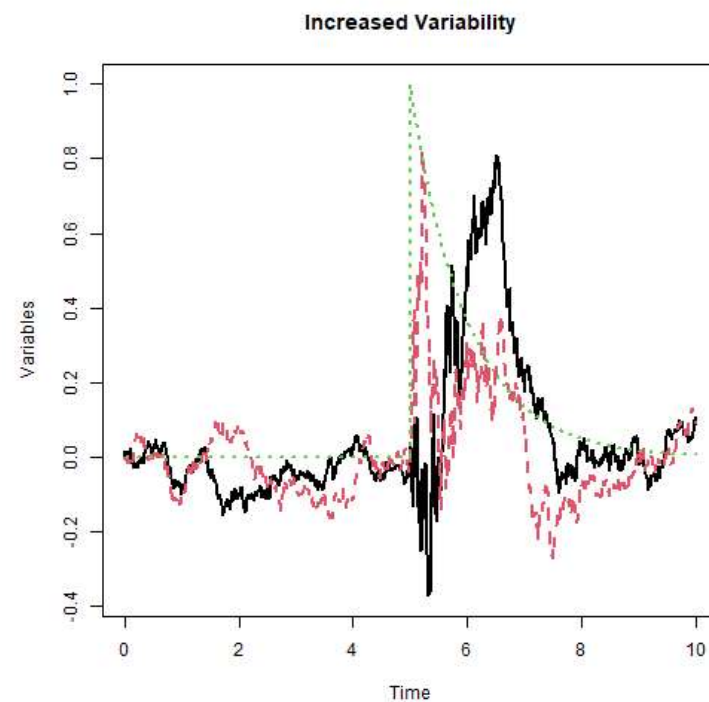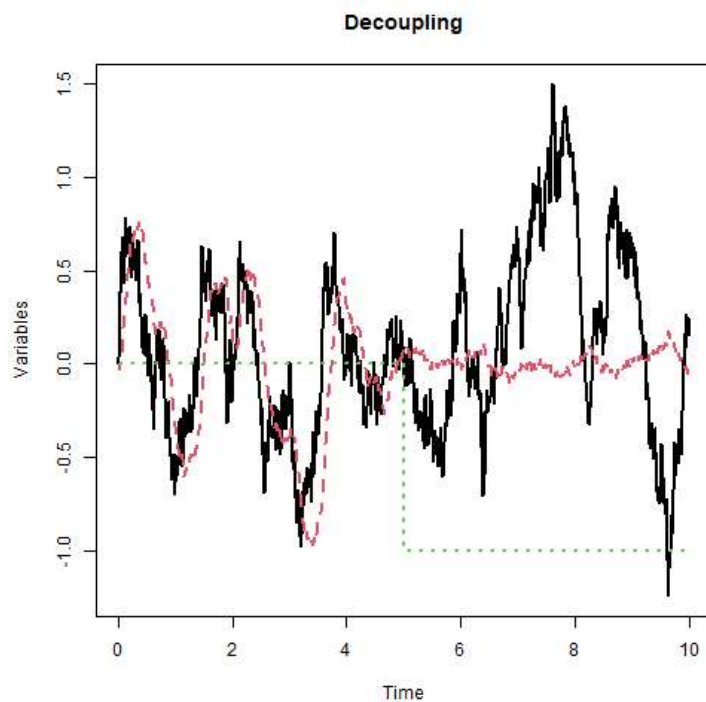
ctsem

- R package: Continuous Time Structural Equation Modelling

- Non-linear, hierarchical, frequentist and Bayesian and some things in the middle, continuous time state space modelling.

- Measurement models, time varying parameters, priors / regularization, interventions, individual differences on all parameters, etc!

Process change:

$$\mathrm{d} \underbrace{\begin{bmatrix} \text{eta1} \\ \text{eta2} \end{bmatrix}(t)}_{\mathrm{d}\boldsymbol{\eta}(t)} = \left(\underbrace{\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \text{eta1} \\ \text{eta2} \end{bmatrix}(t)}_{\boldsymbol{\eta}(t)} + \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\mathbf{b}} \right) \mathrm{d}t + \underbrace{\left\{ \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \right\}}_{\mathbf{G}} \mathrm{d} \underbrace{\begin{bmatrix} W_1 \\ W_2 \end{bmatrix}(t)}_{\mathrm{d}\mathbf{W}(t)}$$

Observations:

$$\underbrace{\begin{bmatrix} \text{Y1} \\ \text{Y2} \end{bmatrix}(t)}_{\mathbf{Y}(t)} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{\boldsymbol{\Lambda}} \underbrace{\begin{bmatrix} \text{eta1} \\ \text{eta2} \end{bmatrix}(t)}_{\boldsymbol{\eta}(t)} + \underbrace{\begin{bmatrix} \text{mm\_Y1} \\ \text{mm\_Y2} \end{bmatrix}}_{\boldsymbol{\tau}} + \underbrace{\begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}}_{\boldsymbol{\Theta}} \underbrace{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}(t)}_{\boldsymbol{\epsilon}(t)}$$
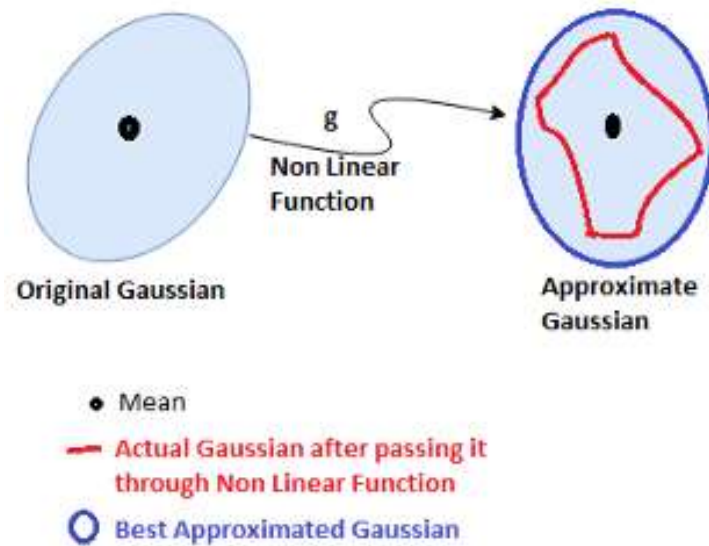
Latent noise per time step:

$$\Delta\big[W_{j\in[1,2]}\big](t-u) \sim \mathrm{N}(0, t-u)$$

Observation noise:

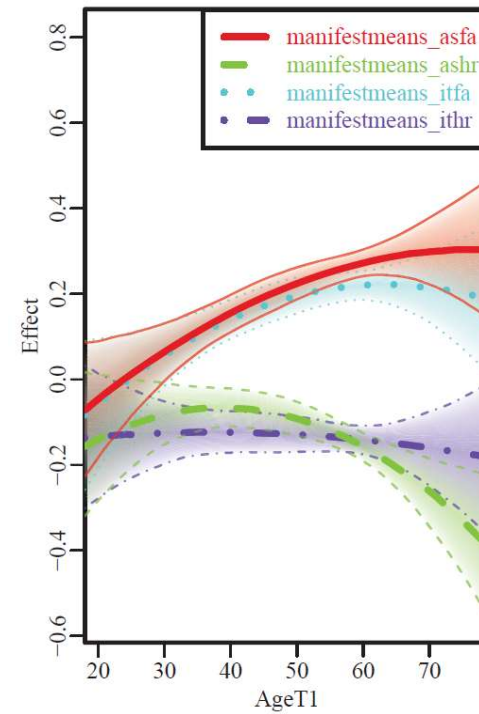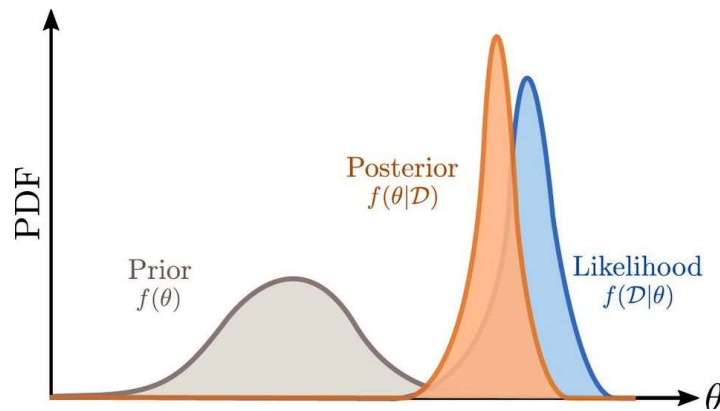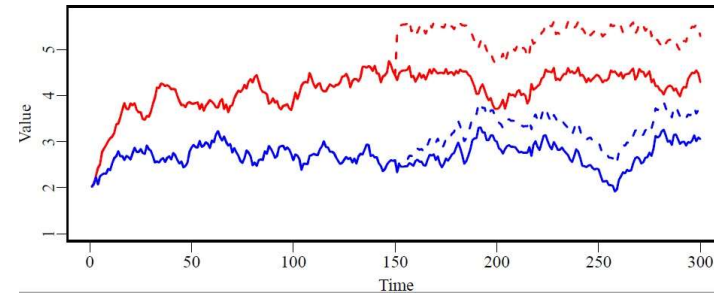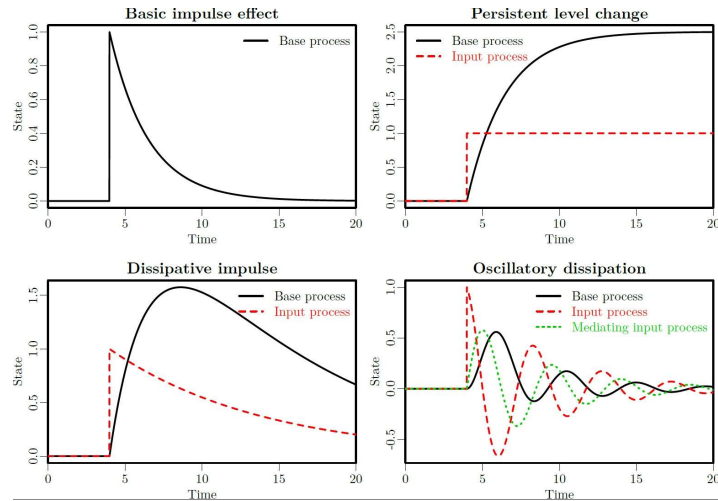$$\big[\epsilon_{j\in[1,2]}\big](t) \sim \mathrm{N}(0,1)$$

Estimation

- Extended Kalman filter:

- With some combination of BFGS / stochastic optimization, importance sampling, or dynamic HMC from Stan.

  - Giving maximum likelihood, maximum a posteriori, or 'full Bayesian' posterior estimation.

- If using the original SEM oriented version and it's slow, use the newer approach!

Going further…

Summary

- The world is complex! Models are vast simplifications.

    - If we ignore the complexity, does it just go away?

- Don't fall into the trap of thinking reality matches your model.

    - Better think of a model as a means to examine and usefully simplify aspects of reality.

- With ctsem, I've tried to reduce some of the typical specification limitations for dynamic systems modelling in social sciences, and allow a flexible and 'genuinely causal' specification.

- A few blog posts on dynamic systems modelling here: https://cdriver.netlify.app/

- If you want to dive in but get lost and confused, post up here: https://github.com/cdriveraus/ctsem/discussions

University of Zurich^{UZH}

FNSNF

SWISS NATIONAL SCIENCE FOUNDATION