Understanding the Time Course of Interventions with Continuous Time Dynamic Models

Charles C Driver[1] and Manuel C Voelkle[1,2]

[1]Max Planck Institute for Human Development, Berlin
[2]Humboldt University, Berlin

How long does a treatment take to reach maximum effect? Is the effect maintained, does it dissipate, or perhaps even reverse? Do certain sorts of people respond faster or stronger than others? Is the treatment more effective in the long run for those that respond quickly? We describe a continuous time dynamic modelling approach for addressing such questions, with discussion and example code for: Simple impulse effects; persistent changes in level; treatments where the effect may reverse in direction over time; treatments that change a trend; assessing mediation in treatment effects; examining individual differences in treatment effects, duration and shape, and correlates of such individual differences.

1 Introduction

Time has often been given only limited consideration when assessing experimental interventions or treatments, but as all such effects unfold over time, a genuine understanding of the system in question is only possible with the full consideration of the time course. How long does a treatment take to have maximum effect? Is the effect then maintained, does it dissipate, or perhaps even reverse? Do certain sorts of people respond faster or stronger than others? Is the treatment more effective in the long run for those that respond quickly? These are the sorts of questions we should be able to answer if we truly understand an intervention and the system we apply it to.

The randomised controlled trial is recognised as something of a gold standard for the analysis of interventions, for good reason. Yet still, when it comes to the direct knowledge of the world such a trial can provide to us, we learn only about what happened to some subjects at a particular moment, or moments, of observation. Going beyond such experimental knowledge in order to produce useful, powerful, and predictive inferences about the world requires some assumptions about regularity and stability of the universe.

For some experimental manipulations it may be sufficient that we learn that there tends to be an effect while the manipulation is occurring. In such cases we need only the most basic assumptions of universal regularity for the knowledge to be useful at later occasions – similar antecedents result in similar consequences. However there are many effects which are only of interest because we assume they persist in some sense outside the bounds of our observation window – a treatment for depression would not be very useful if it only improved a persons depression while being evaluated! As such, whether we are explicitly conscious of it or not, we typically rely on some model of temporal regularity.

In many cases such a model of temporal regularity may be vague and implicit – something like 'there is some sort of continuity in the observed effects even when we are not observing the subjects'. This

stance may be adequate if we know enough about the nature of our system at the time scale we are interested in. For instance, if we assess the effect of a treatment on the level of depression 90 days later, we would also expect to know quite a lot about the effect of the treatment at 89 and 91 days, even though no observations were made on those days. But what about the effect after 30 days, or 180 days? Probably, most people would agree that additional observations would be necessary. With multiple observations we could then interpolate between and beyond them, but how should this be done? Simple linear interpolation between the strength of an average treatment effect (across subjects) at various occasions can be adequate for some situations and research questions, but we can also do much better. In this work we endeavour to show that adopting a dynamic systems approach can not only yield improved estimates of the effect of an intervention at unobserved times, but can also help us better understand the nature of the intervention and system more fully, and improve possibilities for personalised treatments.

Specifically, in this chapter we adopt a continuous time dynamic modelling approach to the problem, based on linear stochastic differential equations. With this approach, variability within a subject over time is partitioned into: stochastic inputs at the system level (latent process variance), deterministic changes based on earlier states of the system; stochastic inputs at the measurement level (measurement error); deterministic inputs of unknown origin (i.e., trends); and then finally, deterministic inputs of known cause – an intervention. In broad terms, the approach differs from what could be done using latent growth curves by the inclusion of the system noise and dynamics component. Thus, rather than sketching only a *description* of change over time, to the extent possible the generating *process* is also considered, even if only in very broad terms. This can lead to more informative inferences, dependent of course on the quality of the data, assumptions, and modelling. For an introduction to continuous time models in the field of psychology see Voelkle, Oud, Davidov and Schmidt (2012), and for more background see Oud and Jansen (2000). The text from Gardiner (1985) gives a detailed treatment of stochastic differential equations in general.

Interventions in a system dynamics context are already considered in fields such as pharmacodynamics and pharmacokinetics for drug discovery and testing. There they endeavour to describe processes on the causal path between blood plasma concentration and effect, for multiple subjects (see Danhof, de Lange, Della Pasqua, Ploeger & Voskuyl, 2008; Donnet & Samson, 2013, for general and estimation-focused overviews, respectively). Single-subject analysis, in which many measurements over time are required, have been more commonly undertaken with such an approach (see for instance the insulin dose and glucose monitoring work of Wang et al., 2014). In the realm of psychology, Boker, Staples and Hu (2016, 1-2) discuss how certain inputs can lead to changes in the equilibrium of a process. With respect particularly to continuous time models, Oud and Jansen (2000), Singer (1993, 1998) have detailed model specifications that included time-varying parameters and exogenous inputs.

In the remainder of this work we will first describe the continuous time dynamic model used here in more depth, then consider why and how various deterministic input effects may be modelled, as well as mediation – how we may be able to use interventions to understand relations between the processes – and individual differences. We finish with a brief example demonstrating how instead of being a main focus, interventions can also be used to aid in system identification. Throughout the work we will provide examples using the ctsem (Driver, Oud & Voelkle, 2017) software package for R (R Core Team, 2014), which interfaces to both the OpenMx (Neale et al., 2016) and Stan (Carpenter et al., 2017) software.

2 The model

The continuous time dynamic model we are interested in here is comprised of a latent dynamic model, and a measurement model. We have previously described different approaches and software for estimating such models in either single or multiple-subject contexts, for a maximum likelihood based specification with mixed effects see (Driver et al., 2017), and for fully random-effects with a hierarchical Bayesian approach see (Driver & Voelkle, In Press). Note that while various elements in the model depend on time, the fundamental parameters of the model as described here are time-invariant. Note also that while subject specific subscripts on the parameters are possible, for simplicity they are not included at this point.

2.1 Latent dynamic model

The dynamic system is described by the following linear stochastic differential equation:

$$\mathrm{d}\boldsymbol{\eta}(t) = \Big(\boldsymbol{A}\boldsymbol{\eta}(t) + \boldsymbol{b} + \boldsymbol{M}\boldsymbol{\chi}(t)\Big)\mathrm{d}t + \boldsymbol{G}\mathrm{d}\boldsymbol{W}(t) \tag{1}$$

Vector $\boldsymbol{\eta}(t) \in \mathbb{R}^{v}$ represents the state of the latent processes at time t. The matrix $\boldsymbol{A} \in \mathbb{R}^{v \times v}$ is often referred to as the drift matrix, with auto effects on the diagonal and cross effects on the off-diagonals characterizing the temporal dynamics of the processes. Negative values on the auto effects are typical of non-explosive processes, and imply that as the latent state becomes more positive, a stronger negative influence on the expected change in the process occurs – in the absence of other influences the process tends to revert to a baseline. A positive cross-effect in the first row and second column would imply that as the second process becomes more positive, the expected change in the first process also becomes more positive. The expected change for a given interval of time can be calculated, and is shown in Equation 4.

The continuous time intercept vector $\boldsymbol{b} \in \mathbb{R}^{v}$, provides a constant fixed input to the latent processes $\boldsymbol{\eta}$. In combination with $\boldsymbol{A}$, this determines the long-term level at which the processes fluctuate around. Without the continuous time intercept the processes (if mean reverting) would simply fluctuate around zero.

Time dependent predictors $\boldsymbol{\chi}(t)$ represent exogenous inputs (such as interventions) to the system that may vary over time, and are independent of fluctuations in the system. Equation 1 shows a generalized form for time dependent predictors, that could be treated a variety of ways depending on the predictors assumed time course (or shape). We use a simple impulse form shown in Equation 2, in which the predictors are treated as impacting the processes only at the instant of an observation occasion u. Such a form has the virtue that many alternative shapes are made possible via augmentation of the system state matrices, as we will describe throughout this work.

$$\boldsymbol{\chi}(t) = \sum_{u \in \boldsymbol{U}} \boldsymbol{x}_u \delta(t - t_u) \tag{2}$$

Here, time dependent predictors $\boldsymbol{x}_u \in \mathbb{R}^{l}$ are observed at measurement occasions $u \in \boldsymbol{U}$, where $\boldsymbol{U}$ is the set of measurement occasions from 1 to the number of measurement occasions, with $u = 1$ treated as occurring at $t = 0$. The Dirac delta function $\delta(t - t_u)$ is a generalized function that is ∞ at 0 and 0 elsewhere, yet has an integral of 1, when 0 is in the range of integration. It is useful to model an impulse to a system, and here is scaled by the vector of time dependent predictors

$\boldsymbol{x}_u$. The effect of these impulses on processes $\boldsymbol{\eta}(t)$ is then $\boldsymbol{M} \in \mathbb{R}^{v \times l}$. Put simply, the equation means that when a time dependent predictor is observed at occasion u, the system processes spike upwards or downwards by $\boldsymbol{M}\boldsymbol{x}_u$.

$\boldsymbol{W}(t) \in \mathbb{R}^v$ represents v independent Wiener processes, with a Wiener process being a random-walk in continuous time. $\mathrm{d}\boldsymbol{W}(t)$ is meaningful in the context of stochastic differential equations, and represents the stochastic error term, an infinitesimally small increment of the Wiener process. Lower triangular matrix $\boldsymbol{G} \in \mathbb{R}^{v \times v}$ represents the effect of this noise on the change in $\boldsymbol{\eta}(t)$. $\boldsymbol{Q}$, where $\boldsymbol{Q} = \boldsymbol{G}\boldsymbol{G}^\top$, represents the variance-covariance matrix of this diffusion process in continuous time. Intuitively, one may think of $\mathrm{d}\boldsymbol{W}(t)$ as random fluctuations, and $\boldsymbol{G}$ as the effect of these fluctuations on the processes. $\boldsymbol{G}\mathrm{d}\boldsymbol{W}(t)$ then simply represents unknown changes in the direction of $\boldsymbol{\eta}$, which are distributed according to a multivariate-normal with continuous time covariance matrix $\boldsymbol{Q}$.

2.2 Discrete time solution of latent dynamic model

To derive expectations for discretely sampled data, the stochastic differential Equation 1 may be solved and translated to a discrete time representation, for any observation $u \in \boldsymbol{U}$:

$$\boldsymbol{\eta}_u = \boldsymbol{A}^*_{\Delta t_u}\boldsymbol{\eta}_{u-1} + \boldsymbol{b}^*_{\Delta t_u} + \boldsymbol{M}\boldsymbol{x}_u + \boldsymbol{\zeta}_u \quad \boldsymbol{\zeta}_u \sim \mathrm{N}(0_v, \boldsymbol{Q}^*_{\Delta t_u}) \tag{3}$$

The * notation is used to indicate a term that is the discrete time equivalent of the original, for the time interval Δt_u (which is the time at u minus the time at $u-1$). $\boldsymbol{A}^*_{\Delta t_u}$ then contains the appropriate auto and cross regressions for the effect of latent processes $\boldsymbol{\eta}$ at measurement occasion $u-1$ on $\boldsymbol{\eta}$ at measurement occasion u. $\boldsymbol{b}^*_{\Delta t_u}$ represents the discrete time intercept for measurement occasion u. Since $\boldsymbol{M}$ is conceptualized as the effect of instantaneous impulses $\boldsymbol{x}$ (which only occur at occasions $\boldsymbol{U}$ and are not continuously present as for the processes $\boldsymbol{\eta}$), its discrete time form matches the general continuous time formulation in Equation 1. $\boldsymbol{\zeta}_u$ is the zero mean random error term for the processes at occasion u, which is distributed according multivariate normal with covariance $\boldsymbol{Q}^*_{\Delta t_u}$. The recursive nature of the solution means that at the first measurement occasion $u = 1$, the system must be initialized in some way, with $\boldsymbol{A}^*_{\Delta t_u}\boldsymbol{\eta}_{u-1}$ replaced by $\boldsymbol{\eta}_{t0}$, and $\boldsymbol{Q}^*_{\Delta t_u}$ replaced by $\boldsymbol{Q}^*_{t0}$. These initial states and covariances are later referred to as T0MEANS and T0VAR respectively.

Unlike in a purely discrete time model, where the various discrete-time effect matrices described above would be unchanging, in a continuous time model the discrete time matrices all depend on some function of the continuous time parameters and the time interval Δt_u between observations u and $u-1$, these functions look as follows:

$$\boldsymbol{A}^*_{\Delta t_u} = e^{\boldsymbol{A}(t_u - t_{u-1})} \tag{4}$$

$$\boldsymbol{b}^*_{\Delta t_u} = \boldsymbol{A}^{-1}(\boldsymbol{A}^*_{\Delta t_u} - \boldsymbol{I})\boldsymbol{b} \tag{5}$$

$$\boldsymbol{Q}^*_{\Delta t_u} = \boldsymbol{Q}_\infty - \boldsymbol{A}^*_{\Delta t_u}\boldsymbol{Q}_\infty(\boldsymbol{A}^*_{\Delta t_u})^\top \tag{6}$$

Where the asymptotic diffusion $\boldsymbol{Q}_\infty = \mathrm{irow}\big(-\boldsymbol{A}^{-1}_{\#}\,\mathrm{row}(\boldsymbol{Q})\big)$ represents the latent process variance as t approaches infinity, $\boldsymbol{A}_{\#} = \boldsymbol{A} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{A}$, with $\otimes$ denoting the Kronecker-product, row is an operation that takes elements of a matrix row wise and puts them in a column vector, and irow is the inverse of the row operation.

2.3 Measurement model

While non-Gaussian generalisations are possible, for the purposes of this work the latent process vector $\boldsymbol{\eta}(t)$ has the linear measurement model:

$$\boldsymbol{y}(t) = \boldsymbol{\Lambda}\boldsymbol{\eta}(t) + \boldsymbol{\tau} + \boldsymbol{\epsilon}(t) \quad \text{where } \boldsymbol{\epsilon}(t) \sim \mathrm{N}(0_c, \boldsymbol{\Theta}) \tag{7}$$

$\boldsymbol{y}(t) \in \mathbb{R}^c$ is the vector of manifest variables, $\boldsymbol{\Lambda} \in \mathbb{R}^{c \times v}$ represents the factor loadings, and $\boldsymbol{\tau} \in \mathbb{R}^c$ the manifest intercepts. The residual vector $\boldsymbol{\epsilon} \in \mathbb{R}^c$ has covariance matrix $\boldsymbol{\Theta} \in \mathbb{R}^{c \times c}$.

2.4 Between subjects model and estimation

We will not go into detail on between subjects and estimation aspects here, as they can be handled in various ways. A frequentist approach with random-effects on the continuous time intercepts and manifest intercepts, and fixed-effects on other model parameters, is presented in Driver et al., 2017, which also describes the R (R Core Team, 2014) package ctsem which can be used for estimating these models. Driver and Voelkle, In Press extends the earlier work and ctsem software to fully random-effects with a hierarchical Bayesian approach. The majority of this present work uses the frequentist approach for the sake of speed, but for an example of the Bayesian approach see Section 4.5

3 Shapes of input effects

When we speak of the effect of an input, we mean the effect of some observed variable that occurs at a specific moment in time. So for instance while a persons‘ fitness is more easily thought of as constantly present and could thus be modelled as a latent process, an event such as ‘going for a run’ is probably easier to consider as a single event in time. We do not propose strict guidelines here, it would also be possible to model a ‘speed of running’ latent process, and it is also clear that all events take some time to unfold. However, we suggest it is both reasonable and simpler to model events occurring over short time scales (relative to observation intervals) as occurring at a single moment in time. Nevertheless, there may be cases when events should be thought of as persisting for some finite span of time, and this may also be approximated using the approaches we will outline.

So, although we speak of an input as occurring only at a single moment in time, the *effects* of such an input on the system of interest will persist for some time, and may exhibit a broad range of shapes. While for the sake of clarity we will discuss the shape of an input effect on a process that is otherwise a flat line, what we really mean by ‘shape of the effect’ is the *difference* between the expected value of the process conditional on an input, and the expected value of the process without any such input. The shape of the effect could then be relative to a randomly fluctuating process, an oscillation, exponential trend, etc. Some examples of possible input effect shapes are shown in Figure 1.

3.1 Basic impulse effect

As formalised by Equation 2, the basic form of the effect, and fundamental building block for more complex effects, is that an input at a singular moment in time causes an impulse in the system at

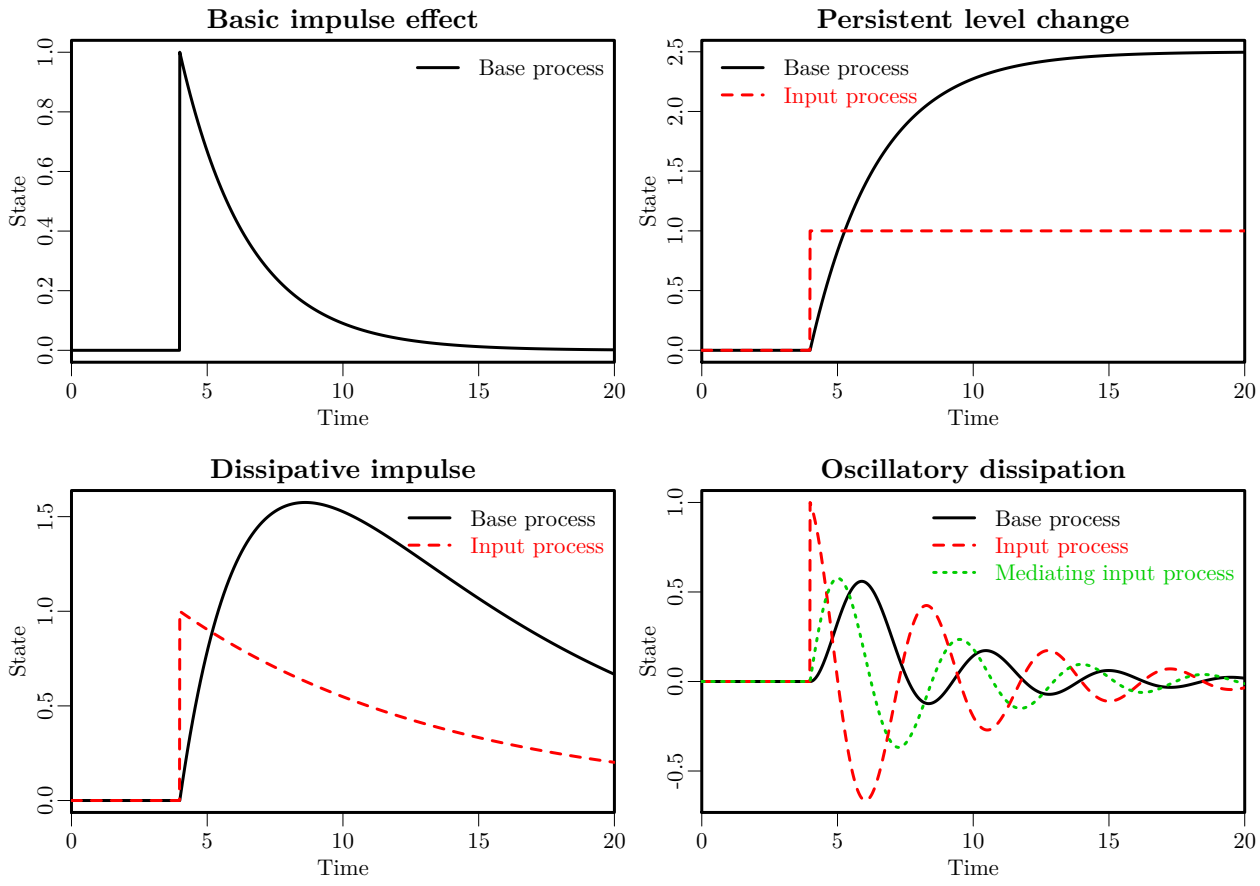

Figure 1: Possible shapes of effects resulting from impulse input affecting a mean-reverting process, given various configurations of the state matrices. For the basic impulse effect no additional processes need to be modelled, while the other examples require either one or two additional processes, that are not directly observed.

that moment, which then dissipates according to the temporal dependencies (drift matrix) in the system. The effect of such an impulse on a first order, mean-reverting and non-stochastic process (i.e., a flat line with any changes driven by deterministic inputs) is shown in the top left of Figure 1. The effect will take on a similar shape for any mean-reverting process, it is just not as easy to see when the process is stochastic and or oscillating.

Effects we might plausibly model in such a way are those where the observed input is expected to have a sudden effect on the system, and that effect dissipates in a similar way to the other, unobserved, inputs on the system that are modelled via the stochastic term. An example of such could be the effect of encountering a friend in the street on ones‘ mood. In such a situation, mood may rapidly rise, then decline back to some baseline in much the same way as would occur for other random mood shifting events throughout the day.

Though the equivalence is perhaps not obvious at first, when time-varying covariates are included in a discrete time cross-lagged or latent change score model, and no temporal dependencies in the covariates are modelled, it is just such an impulse effect that is being instantiated. As such, the approaches and thinking we outline in this paper can also be used in the discrete-time case, though care should be taken to consider divergences between the continuous and discrete-time approaches, particularly when models greater than first order (such as those with oscillations) are considered.

An example R script to simulate data from such a model,then fit, summarise, and plot results (remove hashes in front of plot lines to use), is as follows. [1]

```
install.packages("ctsem")
library(ctsem)



nlatent=1 #number of latent processes
nmanifest=1 #number of manifest variables
tpoints=30 #number of measurement occasions
ntdpred=1 #number of time dependent predictors
TDPREDMEANS=matrix(0,ntdpred*tpoints,1)
TDPREDMEANS[floor(tpoints/2)]=1 #input after 50% of observations

genm=ctModel(Tpoints=tpoints,
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=matrix(c(1), nrow=nmanifest, ncol=nlatent),
  DRIFT=matrix(c(-.2),  nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c(.8), nrow=nlatent, ncol=nlatent),
  MANIFESTVAR=matrix(c(.8), nrow=nmanifest, ncol=nmanifest),
  TDPREDEFFECT=matrix(c(2), nrow=nlatent, ncol=ntdpred),
  CINT=matrix(c(0), nrow=nlatent, ncol=1),
  TDPREDMEANS=TDPREDMEANS,
  MANIFESTMEANS=matrix(c(0), nrow=nmanifest, ncol=1))

dat=ctGenerate(ctmodelobj=genm, n.subjects=50, burnin=50)
#ctIndplot(datawide=dat,n.subjects=10,n.manifest=1,Tpoints=tpoints)

fitm=ctModel(Tpoints=tpoints, type="omx",
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=matrix(c(1), nrow=nmanifest, ncol=nlatent),
```

[1] For the sake of simplicity we generate and fit data without stable between-subjects differences, but in real world analyses of multiple subjects it may be advisable to account for such effects. With ctsem this can be done either via the MANIFESTTRAITVAR or TRAITVAR matrices in frequentist configuration, or by allowing individually varying parameters with the Bayesian approach – discussed briefly in Section /refsec:individualdifferences.

```
  DRIFT=matrix(c("drift11"), nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c("diffusion11"), nrow=nlatent, ncol=nlatent),
  MANIFESTVAR=matrix(c("merror11"), nrow=nmanifest, ncol=nmanifest),
  MANIFESTMEANS=matrix(c("mmeans_Y1"), nrow=nmanifest, ncol=nmanifest),
  TDPREDEFFECT=matrix(c("tdpredeffect21"), nrow=nlatent, ncol=ntdpred))

fit=ctFit(dat, fitm)
summary(fit)
#ctKalman(fit, timestep=.1, subjects = 1:2, plot=TRUE, kalmanvec=c('y','ysmooth'))
```

The matrix forms of the model equations for a basic impulse affecting a first order process are as follows, with underbraced notations denoting the symbol used to represent the matrix in earlier formulas, and where appropriate also the matrix name in the R specification.

$$\underbrace{\mathrm{d}\left[\eta_1\right](t)}_{\mathrm{d}\boldsymbol{\eta}(t)} = \Bigg(\underbrace{\left[\mathrm{drift11}\right]}_{\boldsymbol{A}}\underbrace{\left[\eta_1\right](t)}_{\boldsymbol{\eta}(t)} + \underbrace{\left[0\right]}_{\boldsymbol{b}} + \underbrace{\left[\mathrm{tdpredeffect21}\right]}_{\boldsymbol{M}}\underbrace{\left[\chi_1\right]}_{\boldsymbol{\chi}(t)}\Bigg)dt +$$

$$\underbrace{\left[\mathrm{diffusion11}\right]}_{\boldsymbol{G}}\underbrace{\mathrm{d}\left[W_1\right](t)}_{\mathrm{d}\boldsymbol{W}(t)}$$

$$\underbrace{\left[Y_1\right](t)}_{\boldsymbol{Y}(t)} = \underbrace{\left[1\right]}_{\boldsymbol{\Lambda}}\underbrace{\left[\eta_1\right](t)}_{\boldsymbol{\eta}(t)} + \underbrace{\left[\mathrm{mmeans\_Y1}\right]}_{\boldsymbol{\tau}} + \underbrace{\left[\epsilon_1\right](t)}_{\boldsymbol{\epsilon}(t)} \quad \text{where } \underbrace{\left[\epsilon_1\right](t)}_{\boldsymbol{\epsilon}(t)} \sim \mathrm{N}\Bigg(\left[0\right], \underbrace{\left[\mathrm{merror11}\right]}_{\boldsymbol{\Theta}}\Bigg)$$

LAMBDA: $\boldsymbol{\Lambda}$ — MANIFESTMEANS: $\boldsymbol{\tau}$ — MANIFESTVAR: $\boldsymbol{\Theta}$

3.2 Level change effect

In contrast to the impulse effect, some inputs may result in a stable change in the level of a process. Such a change may occur near instantaneously, or more gradually. The top right of Figure 1 shows the more gradual change. Often, we would hope that treatments may generate such an effect. Consider for instance the effect of an exercise intervention on fitness. In the intervention condition, subjects are encouraged to increase the amount of exercise they do throughout the week. If the intervention is successful, we then wouldn't necessarily expect to see an immediate improvement in fitness, but would hope that people had begun exercising more in general, which would slowly increase their fitness towards some new level.

There are various ways one could model such an effect. An intuitive approach that may spring to mind would be to code the input variable as 0 for both treatment and control prior to the intervention, then when treatment begins, code the input variable of the treated group as 1 for all further observations. This is somewhat problematic however, as it will result in a series of impulses of the same strength occurring at the observations, with the process declining as normal after an observation, then jumping again with the next impulse. This is not at all representative of a consistent change in level when we consider the process at both observed and unobserved times, that is, *between the observations*. Nevertheless, if observations are equally spaced in time, it is potentially adequate for estimating the extent to which the treatment group exhibited a persistent

change in their level. For instance, as illustrated in Figure 2, in the case of unequal observation intervals, it is clear that both the underlying model and resulting predictions will be very wrong when taking such an input variable coding approach. With the equal observation time intervals on the right of Figure 2 however, we see that although the underlying model is incorrect, predictions at the times when an observation occurs (and the input variable is coded as 1) do gradually rise towards a new level.

While possible in specific circumstances to model the level change via input variable coding, we would in general argue for a more adequate model specification, which is easily achievable. The approach we will take throughout this work is to model additional, *unobserved*, processes, which have no stochastic component. This might also be referred to as augmenting the state matrices. These additional processes are affected by the impulse of the input variable, and in turn affect the actual process of interest, which we will refer to as the *base process*. What have we gained by including such an intermediary? This intermediary, or *input process*, is not restricted to an impulse shape, but can exhibit a time course characterised by its own auto effect parameter.

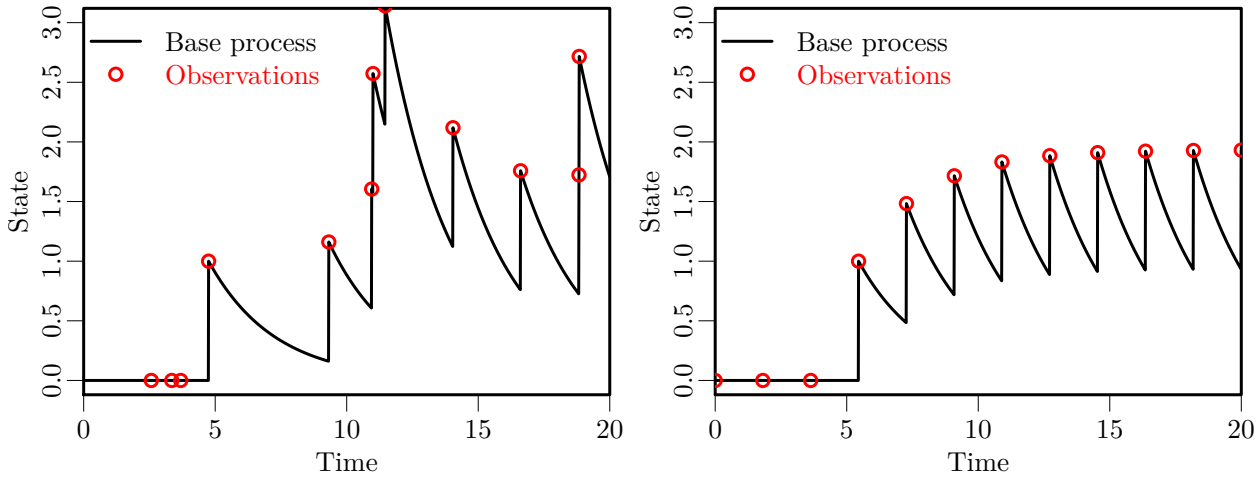

Figure 2: Modelling a persistent level change via input variable coding. While strictly speaking the underlying model is incorrect, nevertheless with equal time intervals (right panel) a level change effect may be captured by a series of impulses, though this is only accurate at the exact moments of observation. With unequal time intervals (left panel), the approach is not even heuristically useful.

One way we might think of such an input process is that of directly characterising the effect we are interested in. Consider for instance the effect of room temperature on attention, and an experimental situation in which room temperature was changed at some point via a thermostat. We do not have direct experimental control or measurements of temperature, rather, we simply know that the thermostat was changed at some point. So we include 'change of thermostat' as our observed input variable, which generates a gradual rise in an unobserved 'temperature' process, which has some effect on attention that we are interested in.

Alternatively, we may have directly observed the manipulation we are interested in, but its effect on the process of interest may occur via some unobserved mediating process. Consider the example of fitness change in response to a cognitive behavioural intervention to motivate exercise. We have a direct observation of 'intervention took place', but know that the subjects must actually exercise for any change in fitness to occur. Thus, we could include a dummy coded input variable of 'intervention', which generates an instantaneous and persistent change on some unobserved input

process that we could think of as something like 'amount of weekly exercise', which in turn has an effect on our measure of 'fitness'. We do not propose any strong interpretation of the unobserved input process, as there will always be many more mediators at many more time scales than we can model. Rather, we propose to model the input process sufficiently such that the major features of the causal chain, at the time scale we are interested in, can be adequately captured. So, although we will not necessarily know what the causal links are, with such a model we can aim at least to understand that a) there appears to be some causal chain between our input and our process of interest, and b) in the vicinity of the values of the context we have observed, the time course of the effect is likely to be similar to our estimated results.

Now, how to explicitly formulate such a level change model? Using ctsem, we configure our model just as for the impulse specification, but include an additional latent process. For this process, we fix all input effects that do not stem from the input predictor to 0. These input effects we fix include the stochastic effects of the DIFFUSION and T0VAR matrices as well as intercept elements from the T0MEANS and CINT matrices – these are fixed to 0 because the input process should represent only the effect of our observed input variable. Because the process is unobserved, we must identify its scale in some way, which can be done either by fixing the effect of the input variable on the process (To any non-zero value, though 1 would seem sensible), or by similarly fixing the effect of the input process on our process of interest. The following script demonstrates simulating and fitting such a model. Note that because zeroes on diagonals can cause problems for matrix inversions, where necessary very small deviations from zero are used instead, which has a negligible effect on the model at the time scale of interest.

```
nlatent=2 #number of latent processes
nmanifest=1 #number of manifest variables
tpoints=30 #number of measurement occasions
ntdpred=1 #number of time dependent predictors
TDPREDMEANS=matrix(0,ntdpred*tpoints,1)
TDPREDMEANS[floor(tpoints/2)]=1 #intervention after 50% of observations

genm=ctModel(Tpoints=tpoints,
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=matrix(c(1, 0), nrow=nmanifest, ncol=nlatent),
  DRIFT=matrix(c(-.4, 0, 1, -0.00001), nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c(.5, 0, 0, 0), nrow=nlatent, ncol=nlatent),
  MANIFESTVAR=matrix(c(.5), nrow=nmanifest, ncol=nmanifest),
  TDPREDEFFECT=matrix(c(0, .4), nrow=nlatent, ncol=ntdpred),
  CINT=matrix(c(0, 0), nrow=nlatent, ncol=1),
  TDPREDMEANS=TDPREDMEANS,
  MANIFESTMEANS=matrix(c(0), nrow=nmanifest, ncol=1))

dat=ctGenerate(ctmodelobj=genm, n.subjects=50, burnin=0)
#ctIndplot(datawide=dat,n.subjects=10,n.manifest=1,Tpoints=tpoints)

fitm=ctModel(Tpoints=tpoints, type="omx",
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=matrix(c(1, 0), nrow=nmanifest, ncol=nlatent),
  DRIFT=matrix(c("drift11", 0, 1, -0.0001), nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c("diffusion11", 0, 0, 0), nrow=nlatent, ncol=nlatent),
  MANIFESTVAR=matrix(c("merror11"), nrow=nmanifest, ncol=nmanifest),
  MANIFESTMEANS=matrix(c("mmeans_Y1"), nrow=nmanifest, ncol=nmanifest),
  T0MEANS=matrix(c("t0m1",0), ncol=1),
  T0VAR=matrix(c("t0var11",0,0,0), nrow=nlatent, ncol=nlatent),
  TDPREDEFFECT=matrix(c(0, "tdpredeffect21"), nrow=nlatent, ncol=ntdpred))
```

```
fit=ctFit(dat, fitm)
summary(fit)
#ctKalman(fit, timestep=.1, subjects = 1:2, plot=TRUE, kalmanvec=c('y','ysmooth'))
```

The matrix forms for a level change intervention affecting a first order process are as follows, with underbraced notations denoting the symbol used to represent the matrix in earlier formulas, and where appropriate also the matrix name in the ctsem specification.

$$\underbrace{\mathrm{d}\begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix}(t)}_{\mathrm{d}\boldsymbol{\eta}(t)} = \left(\underbrace{\begin{bmatrix}\mathrm{drift11} & 1\\ 0 & -1e-04\end{bmatrix}}_{\boldsymbol{A}\ \text{DRIFT}}\underbrace{\begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix}(t)}_{\boldsymbol{\eta}(t)} + \underbrace{\begin{bmatrix}0\\ 0\end{bmatrix}}_{\boldsymbol{b}\ \text{CINT}} + \underbrace{\begin{bmatrix}0\\ \mathrm{tdpredeffect21}\end{bmatrix}}_{\boldsymbol{M}\ \text{TDPREDEFFECT}}\underbrace{[\chi_1]}_{\boldsymbol{\chi}(t)}\right)dt +$$

$$\underbrace{\begin{bmatrix}\mathrm{diffusion11} & 0\\ 0 & 0\end{bmatrix}}_{\boldsymbol{G}\ \text{DIFFUSION}}\underbrace{\mathrm{d}\begin{bmatrix}W_1\\ W_2\end{bmatrix}(t)}_{\mathrm{d}\boldsymbol{W}(t)}$$

$$\underbrace{[Y_1](t)}_{\boldsymbol{Y}(t)} = \underbrace{\begin{bmatrix}1 & 0\end{bmatrix}}_{\boldsymbol{\Lambda}\ \text{LAMBDA}}\underbrace{\begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix}(t)}_{\boldsymbol{\eta}(t)} + \underbrace{[\mathrm{mmeans\_Y1}]}_{\boldsymbol{\tau}\ \text{MANIFESTMEANS}} + \underbrace{[\epsilon_1](t)}_{\boldsymbol{\epsilon}(t)} \quad \text{where } \underbrace{[\epsilon_1](t)}_{\boldsymbol{\epsilon}(t)} \sim \mathrm{N}\left([0],\ \underbrace{[\mathrm{merror11}]}_{\boldsymbol{\Theta}\ \text{MANIFESTVAR}}\right)$$

4 Various extensions

So far we have discussed two possible extremes in terms of effect shape, the sudden and singular impulse, and the slower but constant level change. These are both somewhat restrictive in terms of the shape of the effects implied by the model, so in many cases it may be worthwhile to further free the possible shape of effects, either as a comparison for the more restrictive model, or directly as the model of interest.

4.1 Dissipation

The most obvious and simplest relaxation is to take the level change model, and free the auto effect (diagonal of the drift matrix) for the input process. Then, the extent to which the effect of the input persists is directly estimated, rather than assumed to persist forever. As the auto effect takes on more negative values the input effect approaches the basic impulse type, and as the auto effect nears 0 the resulting input process approaches the level change type. Values greater than 0 would suggest that the input effect is compounding on itself with time, generating an explosive process. Such explosive processes tend to be highly unrealistic for forecasting much further into the future, but may be an adequate characterisation over the time range considered. Note that, although with a highly negative auto effect the input process approaches the impulse shape, the value of the effect strength parameter will need to be much larger in order to match the impulse form with no mediating input input process. This is shown in the top two plots of Figure 3, while the bottom two show a slower dissipation on the left, and an explosive effect on the right.

Such a dissipative input is probably a very reasonable starting point for modelling the effect of an intervention intended to have long-term consequences, but where it is unclear if the consequences

really do persist forever. Were we to use the simple impulse form, our estimated effect would only represent the magnitude of short term changes, and may not represent the bulk of the effect. With the level-change form of predictor, the estimated effect captures changes that persist for all later observations, and thus if the effect actually declines over time, may underestimate the size of the effect at the short and medium term. Instead of these two extremes, it may instead be the case that the intervention was somewhat successful, with change persisting for some time but slowly dissipating. As the script for simulation and fitting of such a model is very similar to that for the level-change model (with the input process auto-effect simply given a parameter name rather than a fixed value of near zero), we do not repeat it here.

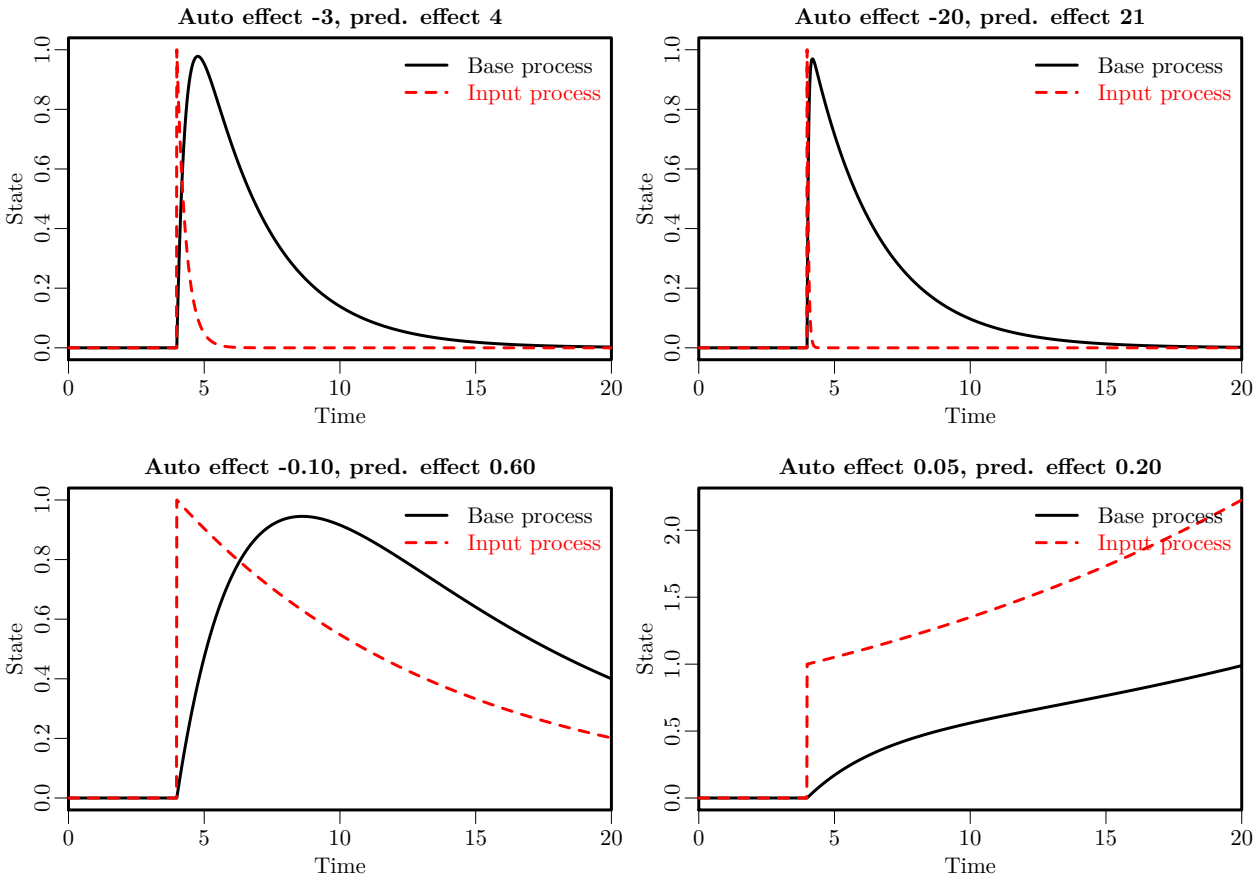

Figure 3: Dissipating input effect, with different parameter values of the input process.

4.2 Multiple time-scales and oscillations

An additional flexibility that one could also consider is to take either the level-change or dissipation type model specification, and free the direct effect of the input (via the TDPREDEFFECT matrix) on the base process itself. This will then allow for estimating both an instantaneous impulse that dissipates according to the base process, and a more persistent input process, allowing for short term effects to differ markedly from longer term, potentially even in the opposite direction. Some real-world examples of such could be the effect of various drugs, for which short-term effects are often very different to longer-term effects, or, at a different time-scale, perhaps the effect of harshly disciplining a child – the child's behaviour may temporarily improve but with negative consequences

later. A mild expansion of this approach could involve the specification of two independent input processes, each with a distinct effect on the base process (thus ensuring that dynamics of the base process in general are not confounded with dynamics in response to the intervention). Including parameters to try and tease apart different time-scale effects will likely make interpretation and hypothesis testing somewhat more complex, and empirical under identification may present an issue for convergence (at least when using the frequentist approach). An approach to mitigate these difficulties may be to revert to the level-change form (wherein the long-term persistence is fixed) rather than estimating the persistence – at least to attain initial estimates.

For more complex shapes of effects due to an intervention, we will need to change from a simple first order input process, to higher order configurations. In the vector form used in this work and the relevant estimation software, higher order processes are always modelled via additional first order processes. Thus, whether one thinks of a single higher order system or multiple first order systems that interact, makes no particular difference. A damped oscillation (as in Voelkle & Oud, 2013) is probably the simplest higher-order model, and could be used in similar circumstances to those of the above multiple time-scales example, though the model will have somewhat different properties. In the multiple time-scales model above, a strong interpretation of the model parameters would suggest that there were two distinct and independent effect processes unfolding, one short and one longer term. This is in contrast to the damped oscillation form, in which a single effect unfolds in a more complex fashion, requiring two unobserved processes that are coupled together.

Figure 4 plots an abstract example of the oscillation generating input effect, as well as an input effect comprised of two independent input processes – one fast and negative, the other slow and positive. The following R code generates and fits a system with a single first order process of interest (our base process), which is impacted upon by an input effect that generates a dissipating oscillation. Figure 5 shows the fit of this oscillation model for one of the generated subjects.

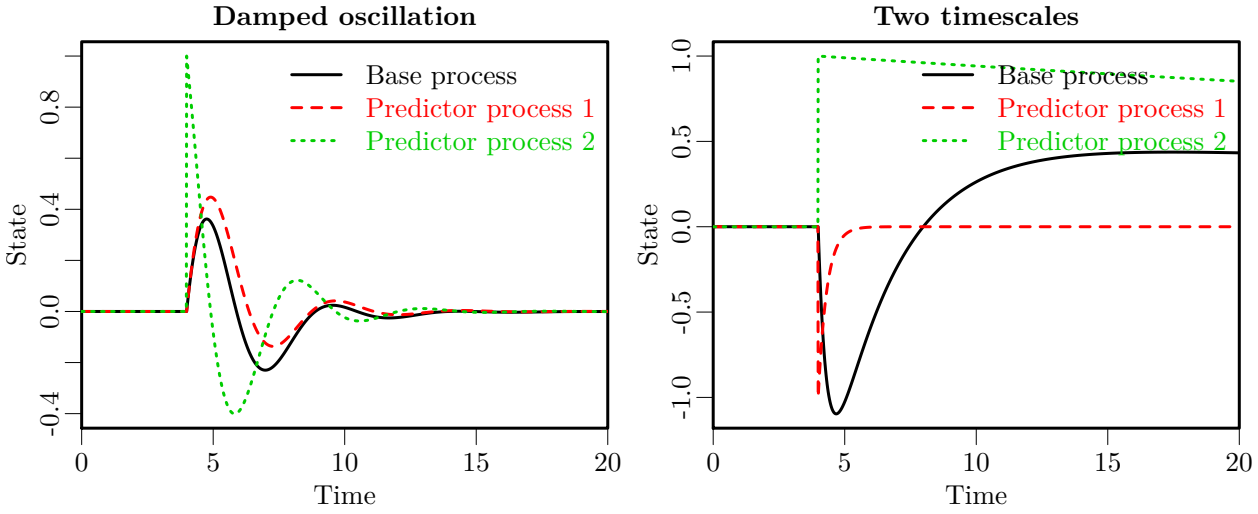

Figure 4: Oscillating input effects. On the left the input process is a classic damped linear oscillator, involving two processes that are deterministically linked, and the direct effect on the base process occurring via input process 1. On the right, the two input processes are independent, with each having a direct effect on the base process. A simpler variant of the latter may neglect to include the short-term input process, instead having the input directly impact the base process.

```
nlatent=3
nmanifest=1
tpoints=30
ntdpred=1
TDPREDMEANS=matrix(0,ntdpred*(tpoints),1)
TDPREDMEANS[floor(tpoints/2)]=1

genm=ctModel(Tpoints=tpoints,
  n.latent=nlatent,
  n.manifest=nmanifest,
  n.TDpred=ntdpred,
  LAMBDA=matrix(c(1, 0, 0), nrow=nmanifest, ncol=nlatent),
  DRIFT=matrix(c(
    -.3, 1, 0,
    0, 0,1,
    0,-2,-.1), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c(
    1, 0, 0,
    0,0,0,
    0,0,0), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  MANIFESTVAR=matrix(c(.5), nrow=nmanifest, ncol=nmanifest),
  TDPREDEFFECT=matrix(c(0, 0, 4), nrow=nlatent, ncol=ntdpred),
  CINT=matrix(c(0), nrow=nlatent, ncol=1),
  TDPREDMEANS=TDPREDMEANS,
  MANIFESTMEANS=matrix(c(0), nrow=nmanifest, ncol=1))

dat=ctGenerate(ctmodelobj=genm, n.subjects=100, burnin=50)

#ctIndplot(datawide=dat,n.subjects=10,n.manifest=1,Tpoints=tpoints)

fitm=ctModel(Tpoints=tpoints,
  n.latent=nlatent,
  n.manifest=nmanifest,
  n.TDpred=ntdpred,
  LAMBDA=matrix(c(1, 0, 0), nrow=nmanifest, ncol=nlatent),
  DRIFT=matrix(c(
    "drift11", 1, 0,
    0, 0,1,
    0,"drift32","drift33"), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c(
    "diffusion11", 0, 0,
    0,0,0,
    0,0,0), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  T0VAR=matrix(c(
    "t0var11", 0, 0,
    0,0,0,
    0,0,0), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  TDPREDEFFECT=matrix(c(0, 0, "tdpredeffect31"), nrow=nlatent, ncol=ntdpred),
  CINT=matrix(c(0), nrow=nlatent, ncol=1),
  T0MEANS=matrix(c("t0mean1",0,0),nrow=nlatent,ncol=1),
  MANIFESTVAR=matrix(c("merror11"), nrow=nmanifest, ncol=nmanifest),
  MANIFESTMEANS=matrix(c("mmeans_Y1"), nrow=nmanifest, ncol=1))

fit=ctFit(dat, fitm)

summary(fit)

#ctKalman(fit, timestep=0.1, subjects = 1, plot=TRUE, kalmanvec=c('y', 'etasmooth'))
```

```
ctKalman(fit, timestep=0.1, subjects = 1, plot=TRUE, kalmanvec=c('y', 'etasmooth'))
```

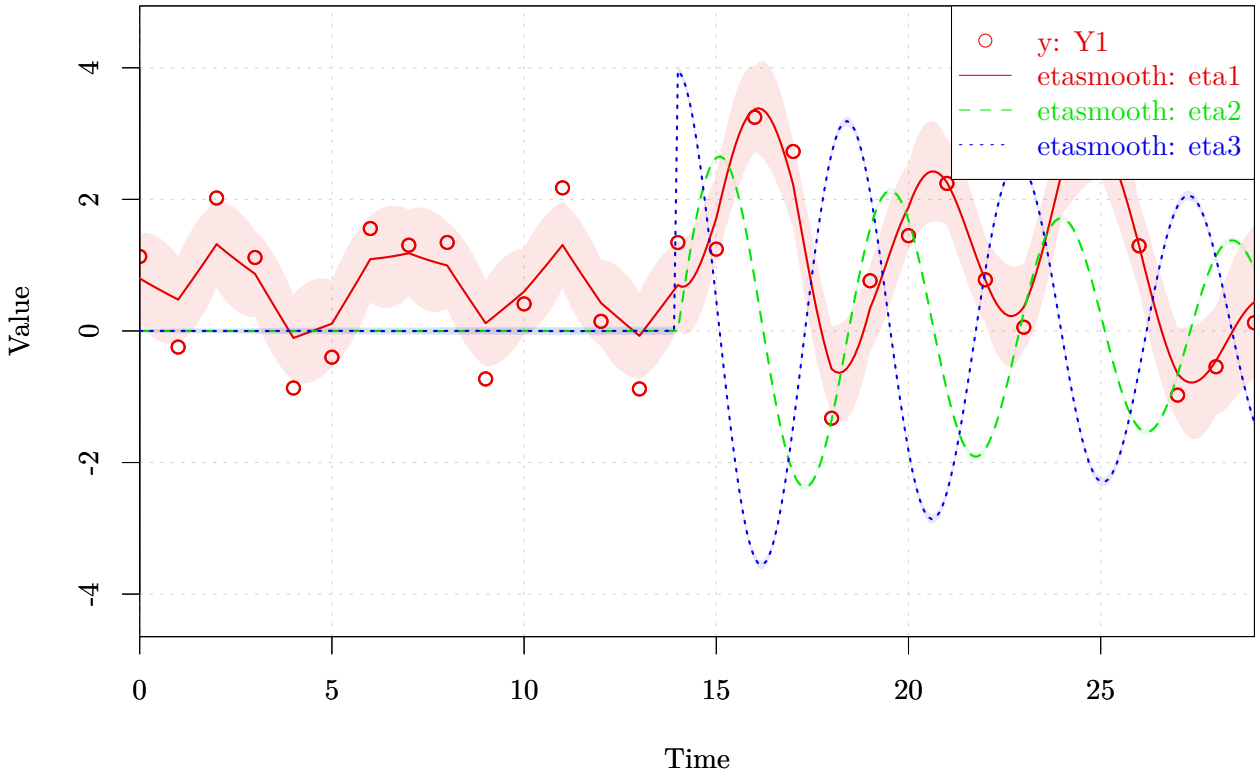

Figure 5: Estimates of latent mean and covariance of base and input processes, for one subject, along with observed data. Estimates are from the Kalman smoother, so conditional on the fitted model and all time points. R code to obtain plots provided at top.

The matrix forms for an intervention effect that first rises and then oscillates back to equilibrium, and which affects a first order process, are as follows. Underbraced notations denoting the symbol used to represent the matrix in earlier formulas, and where appropriate also the matrix name in the ctsem specification.

$$\underbrace{\mathrm{d}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\end{bmatrix}(t)}_{\mathrm{d}\boldsymbol{\eta}(t)} = \Bigg(\underbrace{\begin{bmatrix}\text{drift11} & 1 & 0\\ 0 & 0 & 1\\ 0 & \text{drift32} & \text{drift33}\end{bmatrix}}_{\underset{\text{DRIFT}}{\boldsymbol{A}}}\underbrace{\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\end{bmatrix}(t)}_{\boldsymbol{\eta}(t)} + \underbrace{\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}}_{\underset{\text{CINT}}{\boldsymbol{b}}} + \underbrace{\begin{bmatrix}0\\ 0\\ \text{tdpredeffect31}\end{bmatrix}}_{\underset{\text{TDPREDEFFECT}}{\boldsymbol{M}}}\underbrace{\begin{bmatrix}\chi_1\end{bmatrix}}_{\boldsymbol{\chi}(t)}\Bigg)dt +$$

$$\underbrace{\begin{bmatrix}\text{diffusion11} & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}}_{\underset{\text{DIFFUSION}}{\boldsymbol{G}}}\underbrace{\mathrm{d}\begin{bmatrix}W_1\\ W_2\\ W_3\end{bmatrix}(t)}_{\mathrm{d}\boldsymbol{W}(t)}$$

$$\underbrace{\begin{bmatrix}Y_1\end{bmatrix}(t)}_{\boldsymbol{Y}(t)} = \underbrace{\begin{bmatrix}1 & 0 & 0\end{bmatrix}}_{\underset{\text{LAMBDA}}{\boldsymbol{\Lambda}}}\underbrace{\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\end{bmatrix}(t)}_{\boldsymbol{\eta}(t)} + \underbrace{\begin{bmatrix}\text{mmeans\_Y1}\end{bmatrix}}_{\underset{\text{MANIFESTMEANS}}{\boldsymbol{\tau}}} + \underbrace{\begin{bmatrix}\epsilon_1\end{bmatrix}(t)}_{\boldsymbol{\epsilon}(t)} \quad \text{where } \underbrace{\begin{bmatrix}\epsilon_1\end{bmatrix}(t)}_{\boldsymbol{\epsilon}(t)} \sim \mathrm{N}\Bigg(\begin{bmatrix}0\end{bmatrix}, \underbrace{\begin{bmatrix}\text{merror11}\end{bmatrix}}_{\underset{\text{MANIFESTVAR}}{\boldsymbol{\Theta}}}\Bigg)$$

4.3 Trends

So far we have been using models where the base process is assumed to be stationary over the time window we are interested in (independent of any inputs). This means that given knowledge of only the observation times (and not any of the values), our expectations for the unobserved process states will always be the same, that is, neither expectation nor uncertainty regarding our processes is directly dependent on time. However, cases such as long term development, as for instance when observing from childhood to adulthood, are likely to exhibit substantial trends. If unaccounted for, such trends are likely to result in highly non-Gaussian prediction errors, violating the assumptions of our model. Furthermore, there may be cases where influencing such a trend via an intervention is of interest, and we thus need to be able to incorporate a long term trend in our model, and include any potential effects of the input on the trend.

Let us consider an example of the influence of a health intervention on reading ability, in children of the developing world. Such a health intervention might consist of a short period involving health checks, treatment, and education, with the idea that better health may facilitate learning, both at school and elsewhere. For an observation window of years with a limited observation frequency, the intervention period can reasonably be treated as a singular event.

To specify such a model, we specify our base process as usual, capturing short-term fluctuations in our process of interest, reading ability. We then need to include an additional process that captures the slow trend component. The initial state (T0MEANS), temporal effects between the trend process and other processes, and stochastic input (DIFFUSION) of this trend process are fixed to zero (or near zero), and in contrast to earlier models, we need to include a non-zero continuous-time intercept parameter, to capture the unknown trend size. [2] Other components are

[2] In a model with between subjects differences in the trend, variability in this parameter can be accommodated via the TRAITVAR matrix (for frequentist ctsem) or by simply setting the parameter to individually varying (in the Bayesian approach).

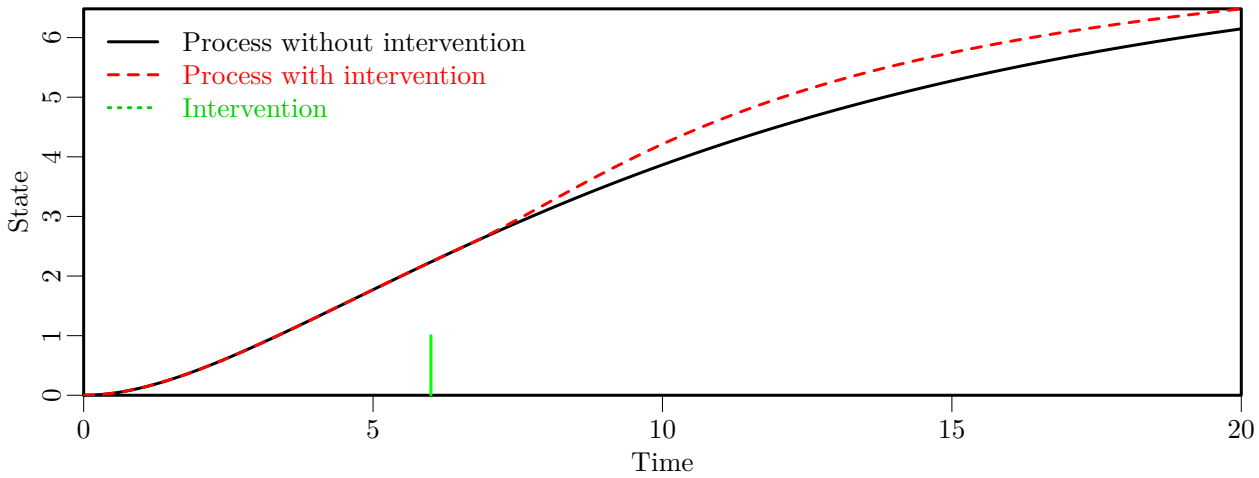

Figure 6: Input effects on a developmental trend. The intervention at time 6 increases the trend slope temporarily.

estimated as usual. Then we include an input effect onto some form of input process, and have this input process affect the trend process. In this case, our measurement model must reflect that our measurements are a summation of both base and trend processes, so two elements of the LAMBDA matrix of factor loadings are now fixed to 1.00, in contrast to prior examples where only one latent process ever directly influenced the indicators.

```
nlatent=3
nmanifest=1
tpoints=30
ntdpred=1
TDPREDMEANS=matrix(0,ntdpred*(tpoints),1)
TDPREDMEANS[floor(tpoints/2)]=1

genm=ctModel(Tpoints=tpoints,
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=matrix(c(1, 1, 0), nrow=nmanifest, ncol=nlatent),
  DRIFT=matrix(c(
    -.5, 0, 0,
    0, -.03,1,
    0,0,-.1), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c(
    3, 0, 0,
    0,0,0,
    0,0,.0001), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  MANIFESTVAR=matrix(c(1), nrow=nmanifest, ncol=nmanifest),
  TDPREDEFFECT=matrix(c(0, 0, 4), nrow=nlatent, ncol=ntdpred),
  CINT=matrix(c(0,1,0), nrow=nlatent, ncol=1),
  TDPREDMEANS=TDPREDMEANS,
  MANIFESTMEANS=matrix(c(0), nrow=nmanifest, ncol=1))

dat=ctGenerate(ctmodelobj=genm, n.subjects=50, burnin=5)
#ctIndplot(datawide=dat,n.subjects=10,n.manifest=1,Tpoints=tpoints)

fitm=ctModel(Tpoints=tpoints,
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=matrix(c(1, 1, 0), nrow=nmanifest, ncol=nlatent),
```

```
  DRIFT=matrix(c(
    "drift11", 0, 0,
    0, "drift22",1,
    0,0,"drift33"), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c(
    "diffusion11", 0, 0,
    0,0,0,
    0,0,.0001), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  T0VAR=matrix(c(
    "t0var11", 0, 0,
    0,0,0,
    0,0,.0001), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  TDPREDEFFECT=matrix(c(0, 0, "tdpredeffect31"), nrow=nlatent, ncol=ntdpred),
  CINT=matrix(c(0,"cint2",0), nrow=nlatent, ncol=1),
  T0MEANS=matrix(c("t0mean1",0,0),nrow=nlatent,ncol=1),
  MANIFESTVAR=matrix(c("merror11"), nrow=nmanifest, ncol=nmanifest),
  MANIFESTMEANS=matrix(c("mmeans_Y1"), nrow=nmanifest, ncol=1))

fit=ctFit(dat, fitm)
summary(fit)
```

The matrix forms for an intervention effect on the long-term trend of a process are as follows. Underbraced notations denote the symbol used to represent the matrix in earlier formulas, and where appropriate also the matrix name in the ctsem specification.

$$\underbrace{\mathrm{d}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\end{bmatrix}(t)}_{\mathrm{d}\boldsymbol{\eta}(t)} = \Bigg(\underbrace{\begin{bmatrix}\text{drift11} & 0 & 0\\ 0 & \text{drift22} & 1\\ 0 & 0 & \text{drift33}\end{bmatrix}}_{\substack{\boldsymbol{A}\\ \text{DRIFT}}}\underbrace{\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\end{bmatrix}(t)}_{\boldsymbol{\eta}(t)} + \underbrace{\begin{bmatrix}0\\ \text{cint2}\\ 0\end{bmatrix}}_{\substack{\boldsymbol{b}\\ \text{CINT}}} + \underbrace{\begin{bmatrix}0\\ 0\\ \text{tdpredeffect31}\end{bmatrix}}_{\substack{\boldsymbol{M}\\ \text{TDPREDEFFECT}}}\underbrace{\begin{bmatrix}\chi_1\end{bmatrix}}_{\boldsymbol{\chi}(t)}\Bigg)dt +$$

$$\underbrace{\begin{bmatrix}\text{diffusion11} & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 1e-04\end{bmatrix}}_{\substack{\boldsymbol{G}\\ \text{DIFFUSION}}}\underbrace{\mathrm{d}\begin{bmatrix}W_1\\ W_2\\ W_3\end{bmatrix}(t)}_{\mathrm{d}\boldsymbol{W}(t)}$$

$$\underbrace{\begin{bmatrix}Y_1\end{bmatrix}(t)}_{\boldsymbol{Y}(t)} = \underbrace{\begin{bmatrix}1 & 1 & 0\end{bmatrix}}_{\substack{\boldsymbol{\Lambda}\\ \text{LAMBDA}}}\underbrace{\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\end{bmatrix}(t)}_{\boldsymbol{\eta}(t)} + \underbrace{\begin{bmatrix}\text{mmeans\_Y1}\end{bmatrix}}_{\substack{\boldsymbol{\tau}\\ \text{MANIFESTMEANS}}} + \underbrace{\begin{bmatrix}\epsilon_1\end{bmatrix}(t)}_{\boldsymbol{\epsilon}(t)} \quad \text{where } \underbrace{\begin{bmatrix}\epsilon_1\end{bmatrix}(t)}_{\boldsymbol{\epsilon}(t)} \sim \mathrm{N}\Bigg(\begin{bmatrix}0\end{bmatrix}, \underbrace{\begin{bmatrix}\text{merror11}\end{bmatrix}}_{\substack{\boldsymbol{\Theta}\\ \text{MANIFESTVAR}}}\Bigg)$$

4.4 Mediation

Throughout this work, we have modelled a range of different shapes of input effects by including additional processes in our system model, and these processes have not been directly measured – regression strengths (i.e., elements of the LAMBDA matrix in ctsem) directly from these additional processes to data have been zero, in most cases. One possible interpretation of these unobserved processes is that they represent some aggregate over all mediating processes that occur between the measured input effect and our measured process of interest. While such processes can simply be

```
ctKalman(fit, timestep=.1, subjects = 1:2, plot=TRUE, plotcontrol=list(ylim=c(-5,50)),
  kalmanvec=c('y','etasmooth'), legendcontrol=list(x='topleft',bg='white'))
ctKalman(fit, timestep=.1, subjects = 1:2, plot=TRUE, kalmanvec=c('y','ysmooth'),
  plotcontrol=list(ylim=c(-5,50)), legend=FALSE)
```

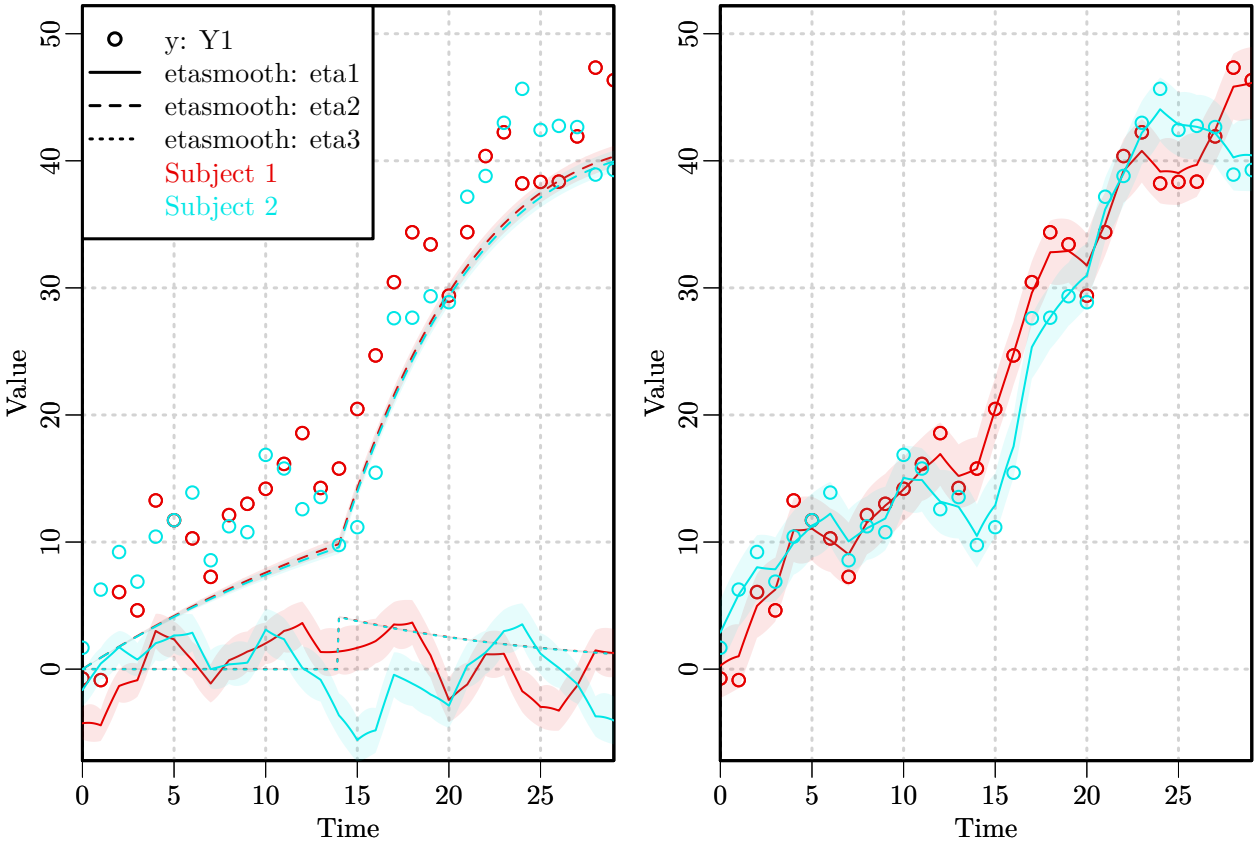

Figure 7: Estimated trend with intervention model for two individuals. On the left, the individual latent processes – base, trend, and predictor – and observed data points are shown. On the right, the latent processes are combined with the measurement model. Estimates are from the Kalman smoother, so conditional on the fitted model and all time points. R code to obtain plots provided at top.

left unobserved, understanding the mediators of effects is a common goal of psychological research, and the framework outlined here offers possibilities for such.

Let us consider again the example of an experimental intervention to improve fitness levels in a group of patients. A successful intervention is unlikely to generate a sudden increase in fitness, rather, it could be expected to gradually rise towards some new level, for which we have already discussed a modelling approach. However, suppose we had also observed some measure of the amount of daily exercise. Conditional on daily exercise, it seems unlikely that the intervention would have any further effect on fitness. That is, we assume that daily exercise mediates the effect of the intervention on fitness. We can test such a theory by comparing a model which includes both mediated and direct effects on fitness, with one that includes only the mediated effect. The following R script provides an example of such a model, and compares a model with effects from the input to all processes, to a restricted model where the input indirectly affects fitness via exercise. In order to provide a fair comparison, the full model for fitting contains an unobserved input process that is impacted by the input. In our example however, the data does not support such a model. Such an over-specified model can make attaining convergence somewhat trickier, as multiple parameters must be estimated where only one would suffice. To minimise these difficulties, we fit the full model using the parameter estimates of the more restricted model as starting values. The mxCompare function from OpenMx is used to compare the two fits, and will show there is no significant difference between the model with the intervention directly affecting fitness, and the restricted model that contains only the intervention effect on exercise rate.

```
set.seed(4)
nlatent=3
nmanifest=2
tpoints=30
ntdpred=1
TDPREDMEANS=matrix(0,ntdpred*(tpoints),1)
TDPREDMEANS[floor(tpoints/2)]=1

genm=ctModel(Tpoints=tpoints,
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=matrix(c(1, 0, 0, 1, 0, 0), nrow=nmanifest, ncol=nlatent),
  DRIFT=matrix(c(
    -.2, .1, 0,
    0, -.3,1,
    0,0,-.0001), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c(
    1, 0, 0,
    3,2,0,
    0,0,.0001), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  MANIFESTVAR=matrix(c(1,0,0,1), nrow=nmanifest, ncol=nmanifest),
  TDPREDEFFECT=matrix(c(0, 0, 5), nrow=nlatent, ncol=ntdpred),
  CINT=matrix(c(0), nrow=nlatent, ncol=1),
  TDPREDMEANS=TDPREDMEANS,
  MANIFESTMEANS=matrix(c(0), nrow=nmanifest, ncol=1))

dat=ctGenerate(ctmodelobj=genm, n.subjects=50, burnin=50)
#ctIndplot(datawide=dat,n.subjects=10,n.manifest=2,Tpoints=tpoints)

nlatent=4 #because for our fit we include extra input process

fullm=ctModel(Tpoints=tpoints, #type="stanct",
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=matrix(c(1, 0, 0, 0, 0, 1, 0, 0), nrow=nmanifest, ncol=nlatent),
```

```
  DRIFT=matrix(c(
    "drift11", 1, "drift13", 0,
    0, "drift22", 0, 0,
    0, 0, "drift33", 1,
    0, 0, 0, "drift44"), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c(
    "diffusion11", 0, 0, 0,
    0, .0001, 0, 0,
    "diffusion31", 0, "diffusion33", 0,
    0, 0, 0, .0001), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  T0VAR=matrix(c(
    "t0var11", 0, 0, 0,
    0, 0, 0, 0,
    "t0var31", 0, "t0var33", 0,
    0, 0, 0, 0), byrow=TRUE, nrow=nlatent, ncol=nlatent),
  TDPREDEFFECT=matrix(c("tdpredeffect11", "tdpredeffect21",
    "tdpredeffect31", "tdpredeffect41"), nrow=nlatent, ncol=ntdpred),
  CINT=matrix(c(0), nrow=nlatent, ncol=1),
  T0MEANS=matrix(c("t0mean1", 0, "t0mean3",0),nrow=nlatent,ncol=1),
  MANIFESTVAR=matrix(c("merror11",0,0,"merror22"), nrow=nmanifest, ncol=nmanifest),
  MANIFESTMEANS=matrix(c("mmeans_fit","mmeans_ex"), nrow=nmanifest, ncol=1))

mediationm=fullm
mediationm$TDPREDEFFECT[1:3,1]=0

mediationfit=ctFit(dat, mediationm)

fullfit=ctFit(dat, fullm,carefulFit=FALSE, #because we specify start values
omxStartValues = omxGetParameters(mediationfit$mxobj))

mxCompare(base = fullfit$mxobj, comparison = mediationfit$mxobj)
```

The matrix forms for the full (not yet restricted) mediation model, with an intervention affecting a measured first order process that in turn affects another measured first order process, are as follows. Underbraced notations denote the symbol used to represent the matrix in earlier formulas, and where appropriate also the matrix name in the ctsem specification.

$$\underbrace{\mathrm{d}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\end{bmatrix}(t)}_{\mathrm{d}\boldsymbol{\eta}(t)} = \Bigg(\underbrace{\begin{bmatrix}\text{drift11} & 1 & \text{drift13} & 0\\ 0 & \text{drift22} & 0 & 0\\ 0 & 0 & \text{drift33} & 1\\ 0 & 0 & 0 & \text{drift44}\end{bmatrix}}_{\boldsymbol{A}\ \text{DRIFT}}\underbrace{\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\end{bmatrix}(t)}_{\boldsymbol{\eta}(t)} + \underbrace{\begin{bmatrix}0\\0\\0\\0\end{bmatrix}}_{\boldsymbol{b}\ \text{CINT}} + \underbrace{\begin{bmatrix}\text{tdpredeffect11}\\ \text{tdpredeffect21}\\ \text{tdpredeffect31}\\ \text{tdpredeffect41}\end{bmatrix}}_{\boldsymbol{M}\ \text{TDPREDEFFECT}}\underbrace{\begin{bmatrix}\chi_1\end{bmatrix}}_{\boldsymbol{\chi}(t)}\Bigg)dt +$$

$$\underbrace{\begin{bmatrix}\text{diffusion11} & 0 & 0 & 0\\ 0 & 1e-04 & 0 & 0\\ \text{diffusion31} & 0 & \text{diffusion33} & 0\\ 0 & 0 & 0 & 1e-04\end{bmatrix}}_{\boldsymbol{G}\ \text{DIFFUSION}}\underbrace{\mathrm{d}\begin{bmatrix}W_1\\ W_2\\ W_3\\ W_4\end{bmatrix}(t)}_{\mathrm{d}\boldsymbol{W}(t)}$$

$$\underbrace{\begin{bmatrix}Y_1\\ Y_2\end{bmatrix}(t)}_{\boldsymbol{Y}(t)} = \underbrace{\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}}_{\boldsymbol{\Lambda}\ \text{LAMBDA}}\underbrace{\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\end{bmatrix}(t)}_{\boldsymbol{\eta}(t)} + \underbrace{\begin{bmatrix}\text{mmeans\_fit}\\ \text{mmeans\_ex}\end{bmatrix}}_{\boldsymbol{\tau}\ \text{MANIFESTMEANS}} + \underbrace{\begin{bmatrix}\epsilon_1\\ \epsilon_2\end{bmatrix}(t)}_{\boldsymbol{\epsilon}(t)}$$

$$\text{where } \underbrace{\begin{bmatrix}\epsilon_1\\ \epsilon_2\end{bmatrix}(t)}_{\boldsymbol{\epsilon}(t)} \sim \mathrm{N}\Bigg(\begin{bmatrix}0\\0\end{bmatrix}, \underbrace{\begin{bmatrix}\text{merror11} & 0\\ 0 & \text{merror22}\end{bmatrix}}_{\boldsymbol{\Theta}\ \text{MANIFESTVAR}}\Bigg)$$

4.5 Individual differences

While understanding how the *average* effect of an intervention develops over time is useful, it has long been observed that any such average may not be representative of the development in any single individual – the exercise intervention we have discussed may be more effective for those who live near a park or recreational space, for instance, as the barriers to following the intervention guidelines are lower. An extreme approach to such a problem is to treat individuals as entirely distinct, but this requires very many observations per subject if even moderately flexible dynamic models are to be fitted, and also raises questions as to how one should treat individuals for which no observations exist. A benefit to this extreme view is that modelling is simplified, and the previously discussed approaches suffice.

A more flexible approach to individual differences is that of random effects (or hierarchical models). These approaches treat individuals as somewhat similar, and estimate the extent of this similarity. This allows for situations where some individuals have been observed many times and others very few, with the resulting model for those observed only few times relying more on the average model across all individuals. For more extended discussion on such models in this context see Driver and Voelkle, In Press. Both frequentist and Bayesian approaches for random effects of observed input variables on latent processes are relatively straightforward, however random effects on the parameters of any unobserved input processes are more complicated in the frequentist case. As such, we demonstrate the case of individual differences using the Bayesian formulation of the ctsem software, which can take longer to fit. In this case we have specified a minimal number of iterations

and it takes roughly 5-10 minutes on a modern PC – for many problems more iterations will be necessary.

For this example, we will look how the strength and persistence of an intervention varies in our sample, and relate this variation to an observed covariate. For this we will use the dissipative predictor model developed in Section 4.1, but allow for variation in the strength (parameters of the TDPREDEFFECT matrix) and persistence (the drift auto-effect parameter of the unobserved input process).

```
nlatent=2 #number of latent processes
nmanifest=1 #number of manifest variables
tpoints=30 #number of measurement occasions
ntdpred=1 #number of time dependent predictors
nsubjects=30 #number of subjects
TDPREDMEANS=matrix(0,ntdpred*tpoints,1)
TDPREDMEANS[floor(tpoints/2)]=1 #intervention after 50% of observations

genm=ctModel(Tpoints=tpoints,
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=matrix(c(1, 0), nrow=nmanifest, ncol=nlatent),
  DRIFT=matrix(c(-.2, 0, 1, -.2), nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c(.1, 0, 0, 0.00001), nrow=nlatent, ncol=nlatent),
  MANIFESTVAR=matrix(c(.1), nrow=nmanifest, ncol=nmanifest),
  TDPREDEFFECT=matrix(c(0, .2), nrow=nlatent, ncol=ntdpred),
  CINT=matrix(c(0, 0), nrow=nlatent, ncol=1),
  TDPREDMEANS=TDPREDMEANS,
  MANIFESTMEANS=matrix(c(0), nrow=nmanifest, ncol=1))

library(plyr)
dat=aaply(1:nsubjects, 1, function(x){ #generate data w random parameter in DRIFT
  tempm=genm
  stdage=rnorm(1)
  tempm$DRIFT[2, 2] = -exp(rnorm(1, -2, .5) + stdage * .5)
  cbind(ctGenerate(ctmodelobj=tempm, n.subjects=1, burnin=50), stdage)
})

#convert to long format used by Bayesian ctsem
datlong=ctWideToLong(datawide = dat, Tpoints = tpoints, n.manifest = nmanifest,
  n.TDpred = ntdpred,n.TIpred = 1,manifestNames = c("Y1"),
  TDpredNames = c("TD1"), TIpredNames=c("stdage"))
datlong=ctDeintervalise(datlong) #convert intervals to abs time

fitm=ctModel(Tpoints=tpoints, type="stanct", n.latent=nlatent, n.manifest=nmanifest,
  n.TDpred=ntdpred, n.TIpred=1, TIpredNames = "stdage",
  LAMBDA=matrix(c(1, 0), nrow=nmanifest, ncol=nlatent),
  DRIFT=matrix(c("drift11", 0, 1, "drift22"), nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c("diffusion11", 0, 0, 0.001), nrow=nlatent, ncol=nlatent),
  MANIFESTVAR=matrix(c("merror11"), nrow=nmanifest, ncol=nmanifest),
  MANIFESTMEANS=matrix(c("mmeans_Y1"), nrow=nmanifest, ncol=nmanifest),
  TDPREDEFFECT=matrix(c(0, "tdpredeffect21"), nrow=nlatent, ncol=ntdpred))

#only the persistence and strength of the predictor effect varies across individuals
fitm$pars$indvarying[-c(8,18)] = FALSE
#and thus standardised age can only affect those parameters
fitm$pars$stdage_effect[-c(8,18)] = FALSE

fit=ctStanFit(datlong, fitm, iter = 200, chains=3)
summary(fit)
```

```
ctKalman(fit, subjects = 1:3, timestep=.01, plot=TRUE,
  kalmanvec='etasmooth', errorvec='etasmoothcov', legendcontrol = list(x = "topleft"))
```

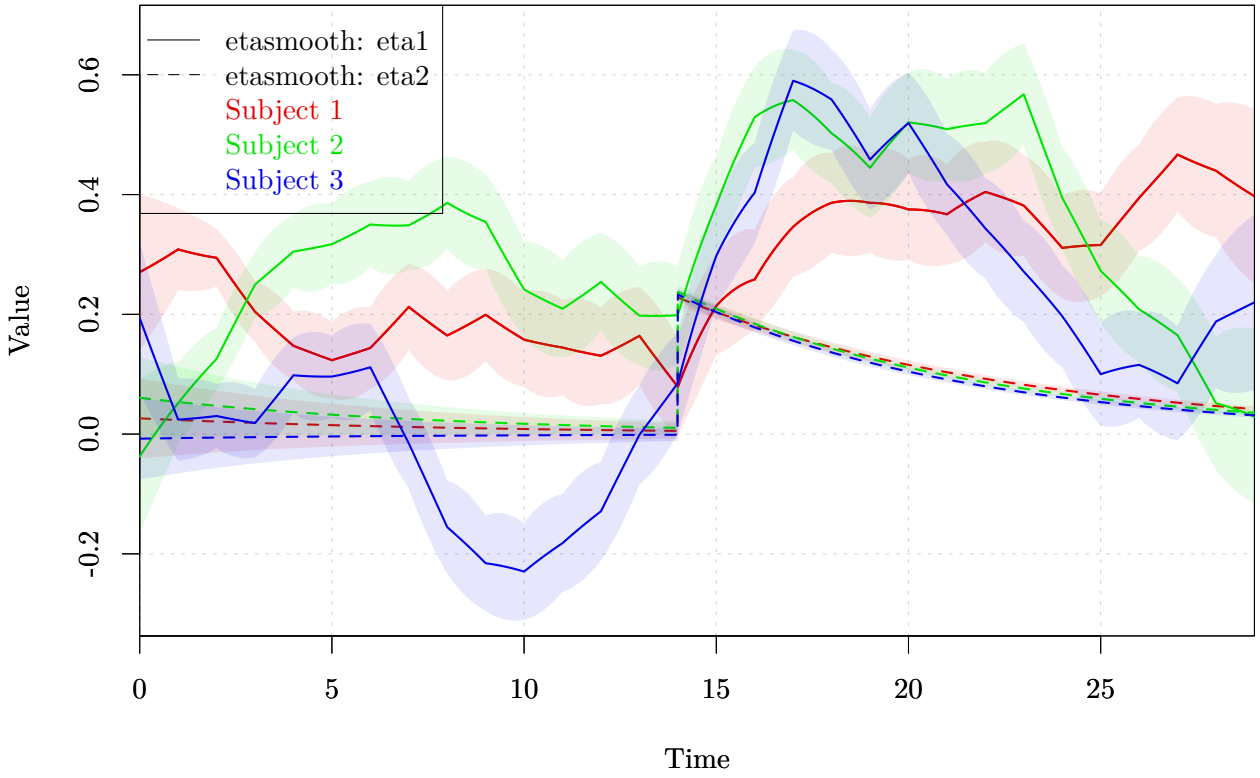

Figure 8: Posterior estimates of the latent mean and covariance of base and input processes, for three subjects. Note the individual differences in the time spans of the input process (eta2). R code to obtain plots at top.

Now, the summary of fit gives an estimated population standard deviation of the persistence and strength of the predictor effect (under *$popsd*), and also an estimate of the effect of standardised age on these parameters (under *$tipreds*). In this case there is no genuine effect of age on the effect strength, but it is important to allow for this effect because of the strong dependence between the strength and persistence parameters – it may be difficult to distinguish one rising from the other lowering in some datasets.

4.6 System identification via interventions

In addition to focusing on how interventions unfold over time, the concepts and modelling procedures we have discussed in this work can also be applied in a context wherein interventions are used primarily with the goal for developing an understanding of the underlying system – the intervention effect itself is not of primary interest. Figure 9 shows a simple case where this may be useful, where the natural state of the bivariate system exhibits limited variation, and measurement error is relatively high. In this context, estimates of the system dynamics can have high uncertainty. By intervening on the system, first on one process and then later on the other, shown on the right of Figure 9, the relation between the two processes becomes much more evident, and more easily estimable. An R script is provided, first fit using data without intervention effects, and then with intervention effects – in general the estimates of the dynamics should become more accurate, with

lower standard errors, when the interventions are included.

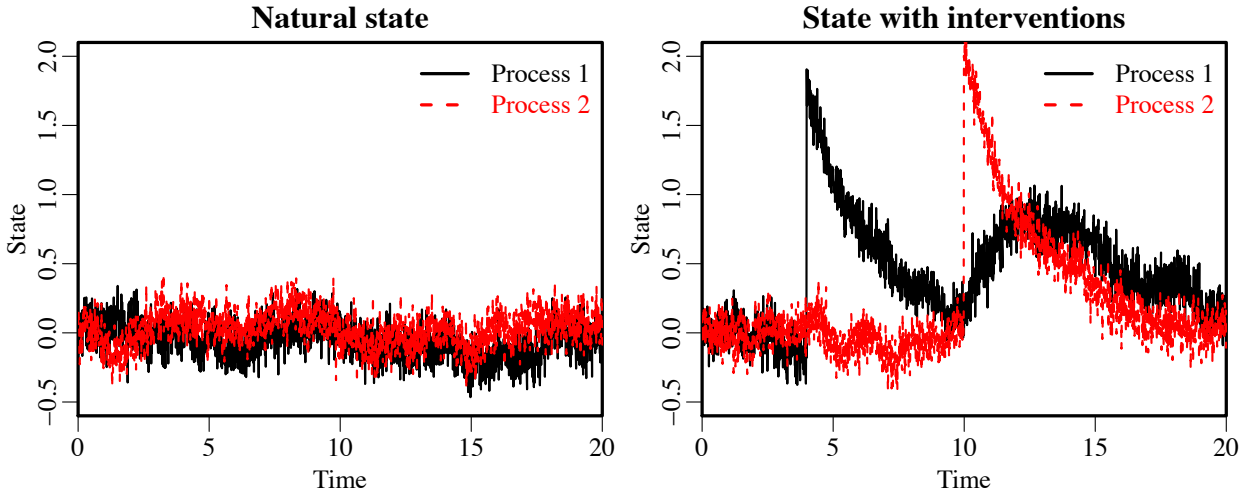

Figure 9: Measurement error combined with limited natural variation in the system can make determining the relationships difficult. On the left, it is not clear which process is driving the other, but with interventions on the right, it is clear that changes in process 2 lead to changes in process 1, and not vice versa.

```
set.seed(1)
nlatent=2 #number of latent processes
nmanifest=2 #number of manifest variables
tpoints=10 #number of measurement occasions
ntdpred=2 #number of time dependent predictors
TDPREDMEANS=matrix(0,ntdpred*tpoints,1)
TDPREDMEANS[c(3,tpoints+6)]=1 #intervention 1 at t3 and intervention 2 at t6

genm=ctModel(Tpoints=tpoints,
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=diag(1,2),
  DRIFT=matrix(c(-.4, 0, .1, -.3), nrow=nlatent, ncol=nlatent),
  DIFFUSION=matrix(c(1, 0, 0, 1), nrow=nlatent, ncol=nlatent),
  MANIFESTVAR=matrix(c(1, 0, 0, 1), nrow=nmanifest, ncol=nmanifest),
  TDPREDEFFECT=matrix(c(0, 0, 0, 0), nrow=nlatent, ncol=ntdpred),
  CINT=matrix(c(0, 0), nrow=nlatent, ncol=1),
  TDPREDMEANS=TDPREDMEANS,
  MANIFESTMEANS=matrix(c(0), nrow=nmanifest, ncol=1))

dat=ctGenerate(ctmodelobj=genm, n.subjects=20, burnin=20)

fitm=ctModel(Tpoints=tpoints, type="omx",
  n.latent=nlatent, n.manifest=nmanifest, n.TDpred=ntdpred,
  LAMBDA=diag(1,2))

fit=ctFit(dat, fitm)
summary(fit)

#now with input effects
genm$TDPREDEFFECT[c(1,4)] = 10
interventiondat=ctGenerate(ctmodelobj=genm, n.subjects=50, burnin=20)
```

```
interventionfit=ctFit(interventiondat, fitm)
summary(interventionfit)
```

5 Discussion

The effect of some intervention or event on an ongoing process can manifest in many forms over a range of time scales. We have shown that using an impulse (the Dirac delta) as an exogenous input effect in continuous time dynamic models, allows a wide variety of possible shapes to be estimated by including additional unobserved mediating processes. This allows for changes in subjects‘ baseline levels or trends, that can happen either instantaneously or gradually. Such changes can dissipate very rapidly, persist for the entire observation window, or even build on themselves over time. Changes at different time scales need not even be in the same direction, as we have shown how one may model oscillations back to equilibrium, or alternatively an initial change with a slower recovery to a new equilibrium. Such an approach can be used both to formulate and test hypotheses regarding the response over time of individuals to some intervention or stimulus, or in a more exploratory approach using a flexible initial specification or iterative model development. We have demonstrated possibilities of formalising and testing mediation models, as well as for examining relations between individuals‘ specific traits and their response to an intervention – a key component for personalised approaches.

While the approach put forward here relies on an impulse input that may then be transformed to a range of shapes via the drift matrix, alternative treatments of input variables in continuous time have been proposed by Oud and Jansen (2000), Singer (1998). In such approaches, the input is conceived as occurring over some time frame, a polynomial describes the temporal evolution, and some approximation of the integral over time must be computed. Because such an approach does not extend the state matrices, it may in some cases be computationally faster.

An aspect which has been implicit throughout the work so far is that of causal inference. We do not think it controversial to state that, to the extent that the model is accurately specified and the observed input effects are exogenous to the system processes, causal interpretations of the input effects and their time course may be reasonable. While a fully explicated formal treatment of causality is beyond the scope of this work, we will briefly discuss exogeneity of effects:

When input effects are not exogenous, as for instance when they are used to model events in a persons‘ life that the person can have some influence over – as for example when ending a relationship – it may still be helpful to model input effects as discussed, but interpretation is far less clear. In such a case, the observed response to a particular type of event may still be interesting, but the response cannot be assumed to be due to the event specifically, as it may instead be due to antecedents that gave rise to the event. Finding that those who choose to end a relationship become happier, does not imply that ending relationships is a useful intervention strategy for people in general!

Amongst the benefits of the approach we have been discussing, there are also some limitations to be considered: The timing of the start of the input effect must be known; the value of the input variables must be known, non-linearities in the input effect are not directly modelled; and the input variable only influences states of the processes, not parameters of the model. In the following paragraphs we elaborate on these limitations.

In the approach we have put forward, the timing of all observations, including observations of input

variables, are regarded as known, and must be input as data. This could be troublesome if a) the timing of the intervention or event in question is simply not well measured, or b) there is some genuine lag time during which the effect of the input or intervention on the process of interest is truly zero, before suddenly taking effect. In both cases, when the possible error is small relative to the observation time intervals, there are unlikely to be substantial ramifications for model fitting and inference. The latter case may be somewhat trickier to determine, however so far as we can imagine, should be limited to situations involving physical stimuli and very short time scales. For instance, a loud noise must travel from the source to the person, and the person must engage in some low-level processing of the noise, before any startle response would occur. For such situations, if the time-lag is reasonably well known, the timing of the input variable can simply be adjusted. In other cases, extending the model to include a measurement model of the observation timing would seem to make sense, though we have not experimented with such.

Just as we may be uncertain about the exact timing, we may also be uncertain about the exact values of the input variables. This latter case is more straightforward however, simply requiring a measurement model on the input variables. This is not explicitly available within ctsem at time of writing, but could be achieved by treating the input variable as an indicator of a regular process in the system, and fixing the variance and covariance of system noise (Diffusion) for this process to near zero. This configuration would result in a very similar model to the dissipating input effect model we have described, wherein the auto-effect could be freely estimated, or fixed if one wished explicitly for an impulse or level change effect. The strength of the intervention effect would then need to be captured either by the factor loading, or the temporal effect in the drift matrix.

Throughout this paper, we have discussed a model in which the input effect is independent of the current system state. So, although the input effect parameters may depend in some way on the stable characteristics of the subject (either their model parameters or included covariates), this effect does not change depending on whether the subject is higher or lower on any processes. Taking the results from such a model at face value could imply that the joy of receiving a chocolate bar as a gift is independent of whether it is three o'clock in the afternoon, or three in the morning. So, some care in application and interpretation is certaintly warranted. This independence of effect from the system state may become more problematic in situations with repeated administration of an intervention. In these situations, one approach would be to include the repeated input effect as a distinct input variable with its own set of parameters, rather than another occurrence of the same input variable.

Such differential response to repeated inputs is similar to another potential issue, that of non-linear response to dosage. An example of this could be caffeine, wherein a moderate dose may improve performance on some tasks but too high a dose leads to decrements. This non-linearity is a common modelling problem and not specific to our continuous time approach to interventions. Possible ways of tackling the issue in this context include the addition of quadratic and higher order polynomial versions of the input variables, or binning together ranges of dosage levels and treating these as distinct input variables.

The final limitation we will discuss is that input variables only affect process states, and may not alter the model parameters themselves. While a persistent level change effect is equivalent to a persistent change in the continuous intercept parameter matrix, there are no such analogues for the other parameter matrices such as the temporal dynamics or system noise. In the case that an intervention substantially changes the dynamics or measurement properties of the system, model parameters estimated using this approach will represent some average over the observation window. There is no cause (that we can see) to expect bias in either state estimates or the expectation given

some intervention, but estimates of uncertainty may be inaccurate. In situations where one is concerned about the possibility of a change in the model parameters induced by an intervention, an approach to test for this could be to include a comparison model wherein the parameters are allowed to change in an instantaneous fashion when the intervention occurs, as with interrupted time-series approaches. This is relatively simple to implement, and would correspond to an instantaneous level-change type effect on the model parameters themselves. A more realistic though more complex approach is to include any relevant model parameters in the system state equations, which would require non-linear filtering techniques.

The continuous-time dynamic modelling approach allows for inference with regards to interventions in two important domains. The primary domain is that of the effect of the intervention on any processes of interest – how long do the effects take to manifest, do they vary in direction over time, and are there individual differences in the effects? The second domain is that of the processes themselves – by generating a substantial impulse of known cause, timing and quantity on the processes, this may enable us to better estimate any causal and mediating relations between processes.

So, just how long does a treatment take to reach maximum effect, how long does the effect last, and what sorts of people is it most effective for? By adopting a modelling framework such as we have proposed, and developing an improved understanding of the timing of interventions and their effects, we may be able to better answer such questions.

References

Boker, S. M., Staples, A. D. & Hu, Y. (2016). Dynamics of change and change in dynamics. *Journal of Person-Oriented Research*, *2*, 34–55. doi:10.17505/jpor.2016.05

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1). doi:10.18637/jss.v076.i01

Danhof, M., de Lange, E. C. M., Della Pasqua, O. E., Ploeger, B. A. & Voskuyl, R. A. (2008). Mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modeling in translational drug research. *Trends in Pharmacological Sciences*, *29*(4), 186–191. doi:10.1016/j.tips.2008.01.007

Donnet, S. & Samson, A. (2013). A review on estimation of stochastic differential equations for pharmacokinetic/pharmacodynamic models. *Advanced Drug Delivery Reviews*. Mathematical modeling of systems pharmacogenomics towards personalized drug delivery, *65*(7), 929–939. doi:10.1016/j.addr.2013.03.005

Driver, C. C., Oud, J. H. L. & Voelkle, M. C. (2017). Continuous time structural equation modeling with r package ctsem. *Journal of Statistical Software*, *77*(5). doi:10.18637/jss.v077.i05

Driver, C. C. & Voelkle, M. C. (In Press). Hierarchical Bayesian continuous time dynamic modeling. *Psychological Methods*. Retrieved from https://www.researchgate.net/publication/310747801_Hierarchical_Bayesian_Continuous_Time_Dynamic_Modeling

Gardiner, C. W. (1985). *Handbook of stochastic methods*. Springer Berlin. Retrieved from http://tocs.ulb.tu-darmstadt.de/12326852.pdf

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., ... Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81*(2), 535–549. doi:10.1007/s11336-014-9435-8. pmid: 25622929

Oud, J. H. L. & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika*, *65*(2), 199–215. doi:10.1007/BF02294374

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Singer, H. (1993). Continuous-time dynamical systems with sampled data, errors of measurement and unobserved components. *Journal of Time Series Analysis*, *14*(5), 527–545. 00046. doi:10.1111/j.1467-9892.1993.tb00162.x

Singer, H. (1998). Continuous panel models with time dependent parameters. *Journal of Mathematical Sociology*, *23*(2), 77–98. doi:10.1080/0022250X.1998.9990214

Voelkle, M. C. & Oud, J. H. L. (2013). Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes. *British Journal of Mathematical and Statistical Psychology*, *66*(1), 103–126. doi:10.1111/j.2044-8317.2012.02043.x

Voelkle, M. C., Oud, J. H. L., Davidov, E. & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods*, *17*(2), 176–192. doi:10.1037/a0027543. pmid: 22486576

Wang, Q., Molenaar, P., Harsh, S., Freeman, K., Xie, J., Gold, C., ... Ulbrecht, J. (2014). Personalized state-space modeling of glucose dynamics for type 1 diabetes using continuously monitored glucose, insulin dose, and meal intake: An extended kalman filter approach. *Journal of Diabetes Science and Technology*, *8*(2), 331–345. doi:10.1177/1932296814524080. pmid: 24876585