Linear Dynamical System (LDS)

Robin Vloeberghs

October 4, 2022

1 Overview model

Perceptual decisions are made by comparing incoming sensory evidence to a criterion or category boundary. In a task with two response options (right vs. left), an observer will respond right when the evidence exceeds the criterion, and left otherwise. The criterion with which the sensory evidence is compared can change over time. There are two sources that can induce a shift in criterion:

- Systematic updating (influence previous response, previous confidence, previous stimulus,...)
- Non-systematic updating (slow drifts in criterion due to random fluctations in attention, arousal, motivation,...)

The aim of the model is to disentangle these two sources causing the shifts in criterion. The model assumes that a choice on trial t follows a Bernoulli process:

$$y_t \sim Bern(\theta_t) \tag{1}$$

with mean

$$E[y_t|s_t, c_t] = \theta_t \tag{2}$$

where θ_t represents the chance of a right response, and can be obtained by comparing the stimulus evidence s_t to the decision criterion c_t. The more s_t exceeds c_t, the higher the probability of a right response, and vice versa.

$$E[y_t|s_t, c_t] = \frac{1}{1 + e^{-(\beta_0 + \beta_s s_t + c_t)}} \tag{3}$$

A logistic link function ensures that θ is scaled between 0 and 1. The parameter β_0 captures an overall response tendency (e.g. higher proportion of right responses over the course of the experiment) and the perceptual sensitivity of an observer is represented by β_s. The latter will dictate the steepness of the psychometric function. The decision criterion c_t is influenced by both systematic updating and slow drifts, and can therefore be rewritten as:

$$c_t = \beta_u u_t + \beta_d d_t \tag{4}$$

where u_t represents any systematic updating variable (e.g. previous response, previous confidence,...) and d_t captures the slow drift. These latent drifts are assumed to follow a random walk process.

$$d_t = d_{t-1} + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma_d) \tag{5}$$

Thus formula 4 can be rewritten as:

$$c_t = \beta_u u_t + \beta_d(d_{t-1} + \epsilon_t) \tag{6}$$

Lastly, given all previous steps, we can redefine formula (3) as:

$$E[y_t|s_t, c_t] = \frac{1}{1 + e^{-(\beta_0 + \beta_s s_t + \beta_u u_t + \beta_d(d_{t-1} + \epsilon_t))}} \tag{7}$$

1.1 Model in the SSM package

The mathematical model described in the previous section is the result of many modeling choices made by Gupta (i.e., fixing parameters to a certain value). However, the SSM package in Python in which the model is created and estimated, also allows other model structures. So let's go over this ***general*** model, and discuss the modeling choices that Gupta made, resulting in the model above.

In the SSM package, a linear dynamical system (LDS) is assumed to have two ***levels***. The first level describes the latent process:

$$X_t = b + VU_t + AX_{t-1} + w_t \tag{8}$$

with

$$w_t \sim N(0, \sigma_d) \tag{9}$$

X_t is the latent process (i.e. the slow drift) and is assumed to follow an autoregressive process with lag 1, or AR(1). An AR(1) process has a model structure where $X_t = \beta_0 + \beta_1 X_{t-1} + \epsilon_t$. β_1 can be thought of as a 'choppiness parameter'. The closer to 1, the smoother, or slower the waves will be. Further away from 1, the more abrupt and choppy the time series becomes. When the coefficient becomes negative, the choppiness becomes even more extreme, as the current value X_t will flip sign compared to the previous X_{t-1}. When β_1 is fixed to 1, we call this a random walk.

b is a constant that is added each iteration. U_t is a matrix and contains input variables that can influence the latent process (e.g., stimulus, previous response, previous confidence,...). And lastly, w_t is the error term.

This latent level is then used in the second level, the one that describes the emissions (Bernoulli response outcomes in our case).

$$Y_t = d + FU_t + CX_t \tag{10}$$

where d is the intercept, U_t is a matrix and contains input variables that can influence the emissions (e.g., stimulus, previous response, previous confidence,...), and X_t is our latent process. Finally, Y_t is transformed by a logistic function $\frac{1}{1+e^{(-Y_t)}}$ after which it represents the probability of a 'rightwards' response in a Bernoulli model.

In the model proposed by Gupta, A is fixed to 1, imposing a random walk, and V and b are fixed to 0. The remaining parameters are estimated using the Expectation-Maximization algorithm (EM).
With these modeling choices the whole model can be summarized as:

$$Y_t = d + FU_t + C(X_{t-1} + w_t) \tag{11}$$

with

$$w_t \sim N(0, \sigma_d) \tag{12}$$

2 Simulations

2.1 Parameter recovery: post-correct and post-error

2.1.1 With effects

We simulated 20 datasets with a varying number of trials and estimated the model with 200 iterations for the EM procedure. For each trial a probability of a 'right' response is calculated based on formula 7 (see above) with generative parameters such as perceptual sensitivity (slope for stimulus), bias (intercept), and systematic updating (in this case a win-stay lose-switch strategy). For the latter, a constant is added to the probability of a 'right' response to simulate a shift in criterion. If the constant is positive, this will increase the probability of a right response, if negative, it will decrease the probability. Based on this probability, a coin is tossed (Bernoulli emission), which determines the actual response for that trial. The slow drifts are simulated using an AR(1) process with an AR coefficient of .9995. In the plot below the generative parameter values are indicated by the horizontal line. Note that the stimulus evidence for each trial is drawn from a uniform distribution with range [0,1]. As a result, a bias of -5 will actually be unbiased. To illustrate this, imagine stimulus evidence of .5 (ambiguous trial) and a perceptual sensitivity of 10. When ignoring everything apart from stimulus and bias, we will have 10 * .5 + (-5) = 0. When inserted in the logistic function this will result in a probability of .5. In general, when evidence is between 0 and 1, the bias will be unbiased when bias = -slope evidence/2.

Overall, the fits do not look that great. We see an overestimation for the perceptual sensitivity, even with a large number of trials. Also for the systematic updating parameters PC and PE we see some biases in the fits. Lastly, it seems impossible to correctly estimate sigma. This is probably due to the trade-off between sigma and C. C (formula 10), or b_d (formula 4) weights how much the slow drifts contribute to the emissions. If this value is low, then sigma will be large, and visa versa. According to Gupta, this trade-off is inherent to the current model structure.

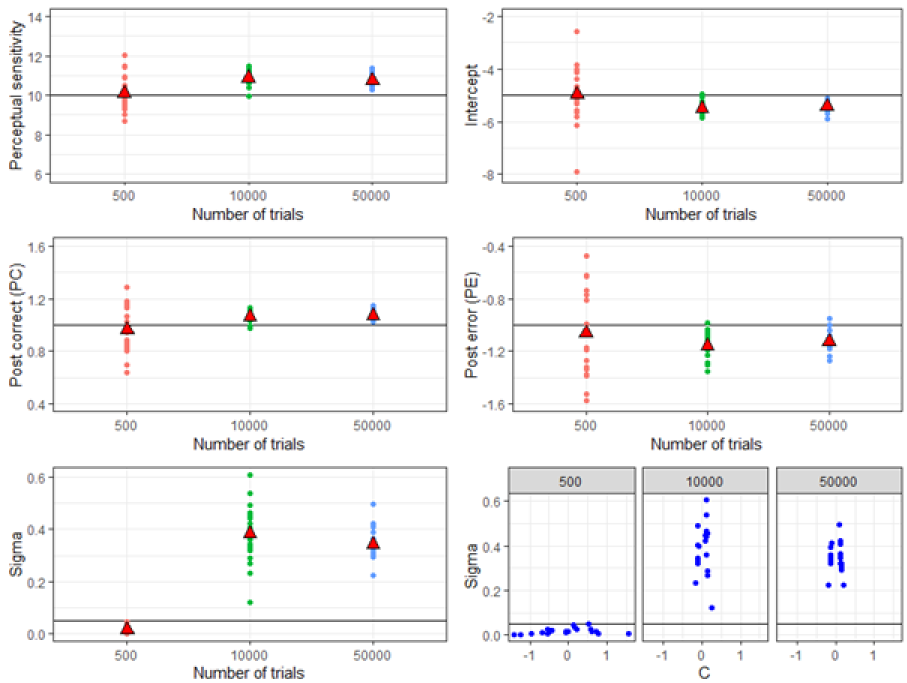

Figure 1: Parameter recovery.

Crucially, can it fit the slow drifts? Below the estimated trajectories for 1 dataset. When we look at the lower two panels it looks good. However, if we zoom in on the first 500 trials we see that the global tendency can be captured but a lot of variation remains uncaptured, even when we have 50000 trials. See param.recovery.originalGupta.estimated.slowdrifts.pdf for all datasets.

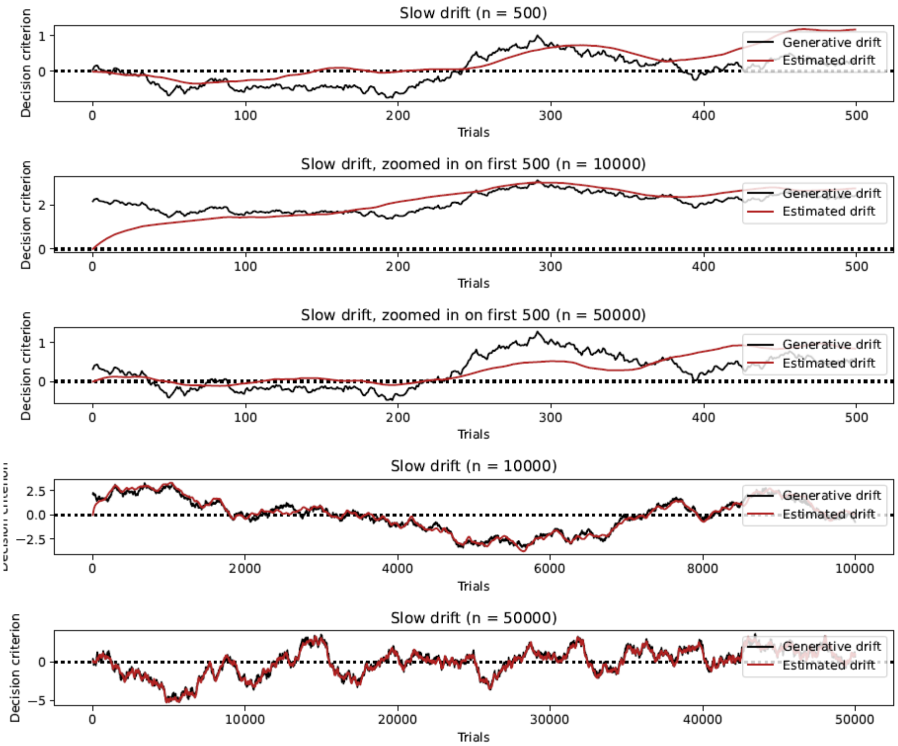

Figure 2: Generative and predicted slow drift trajectories.

2.1.2 Without effects

The same simulation scenario as above, but now without any systematic updating (generative parameters are fixed to 0).

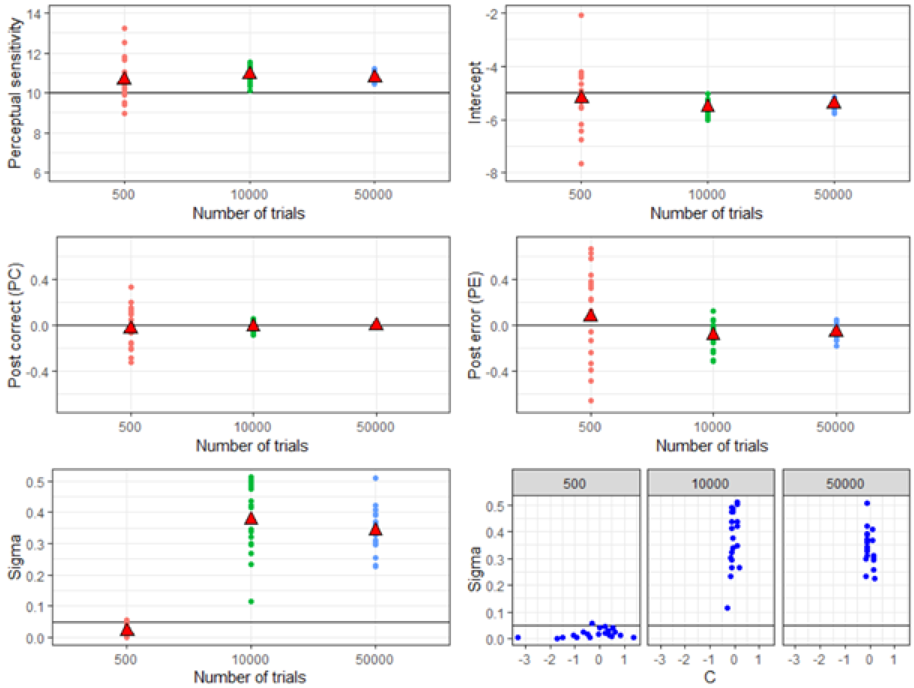

Figure 3: Parameter recovery.

An example of the estimated slow drift trajectory of 1 dataset.

For more examples see param.recovery.originalGupta.estimated.slowdrifts.noSys.pdf

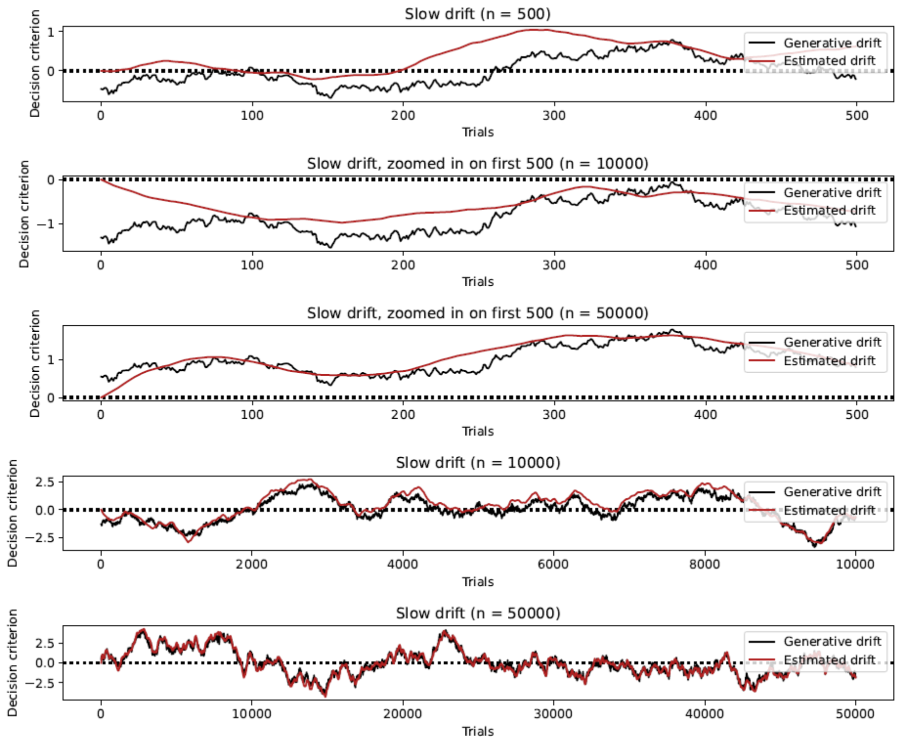

Figure 4: Generative and predicted slow drift trajectories.

2.2 Parameter recovery: previous response, confidence, and stimulus

2.2.1 With effects

Next, I changed the simulation script to allow systematic updating of previous response, previous confidence, previous sign evidence, and previous absolute evidence. Evidence is now drawn from two normal distributions, similar to signal detection theory, with 0 as ambiguous stimulus. As a result, an intercept of 0 will now be unbiased. In addition, confidence is calculated by allowing post-decisional evidence. By doing so, confidence can become negative (i.e., perceived error).

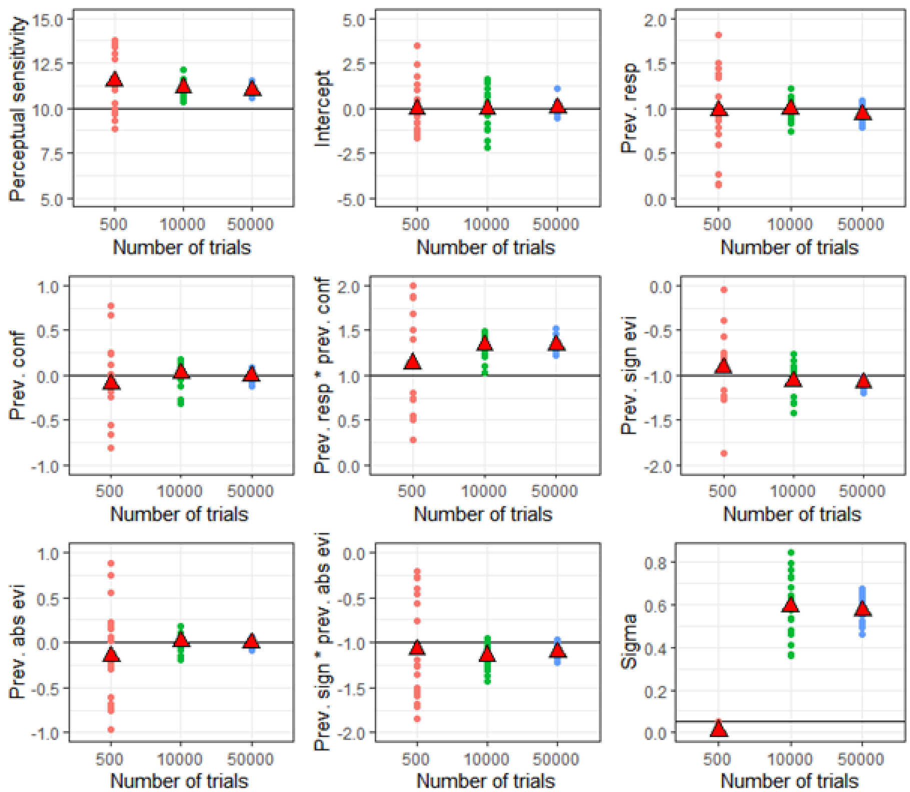

Figure 5: Parameter recovery.

Pdf with estimated slow drift trajectories failed. Will run this again!

2.2.2 Without effects

Same simulation scenario as above, but now all the systematic updating parameters are fixed to 0.
We see some serious biases in the estimates. For example, the estimate for previous response * previous confidence with n = 50000 is 0.225. To put this number in context: the estimate for previous response * previous confidence from the mixed models in the beehives dataset is 0.151.

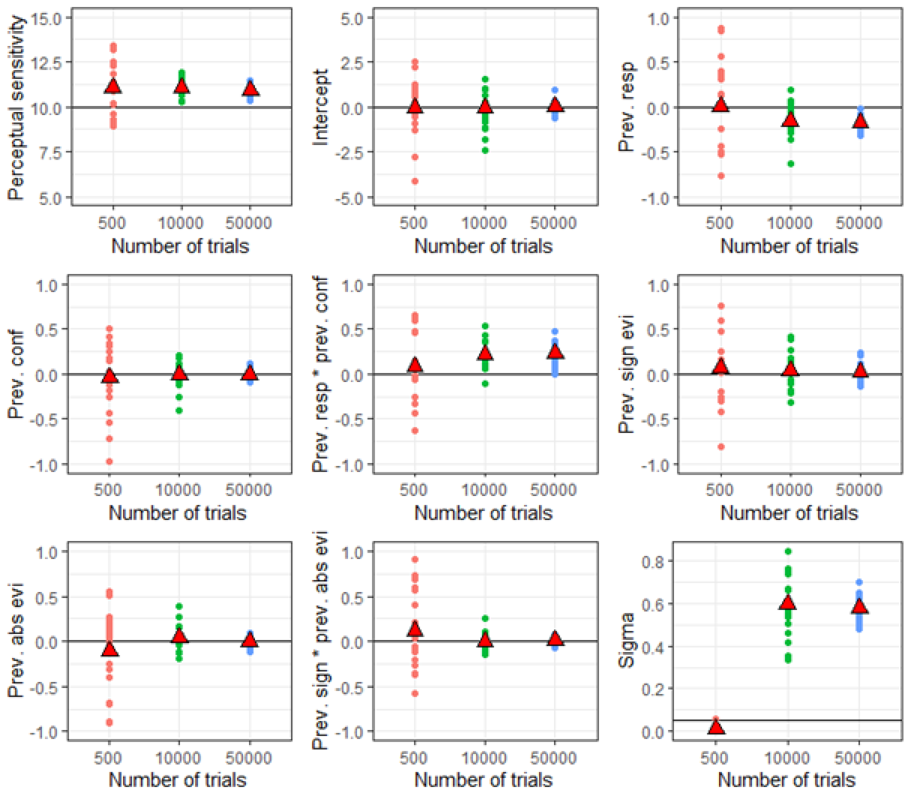

Figure 6: Parameter recovery.

For slow trajectories of all datasets see param.recovery.prevrespconf.noSys.estimated.slowdrifts

2.3 Simulation script in R

For completeness I added this section as well. Simulations were made using the 'old' R script that kickstarted the whole slow drift idea. Data is simulated from a SDT framework where the criterion changes over time due to slow drifts (generated by AR(1) process, similar to Gupta). Responses are made by comparing evidence to criterion. Confidence is calculated by allowing post-decisional evidence accumulation. 10 subjects were simulated with 10000 trials each. No systematic updating is implemented, only slow drifts!

Below the result when slow drifts are not estimated. As expected the model reports systematic updating whereas in fact there is none.

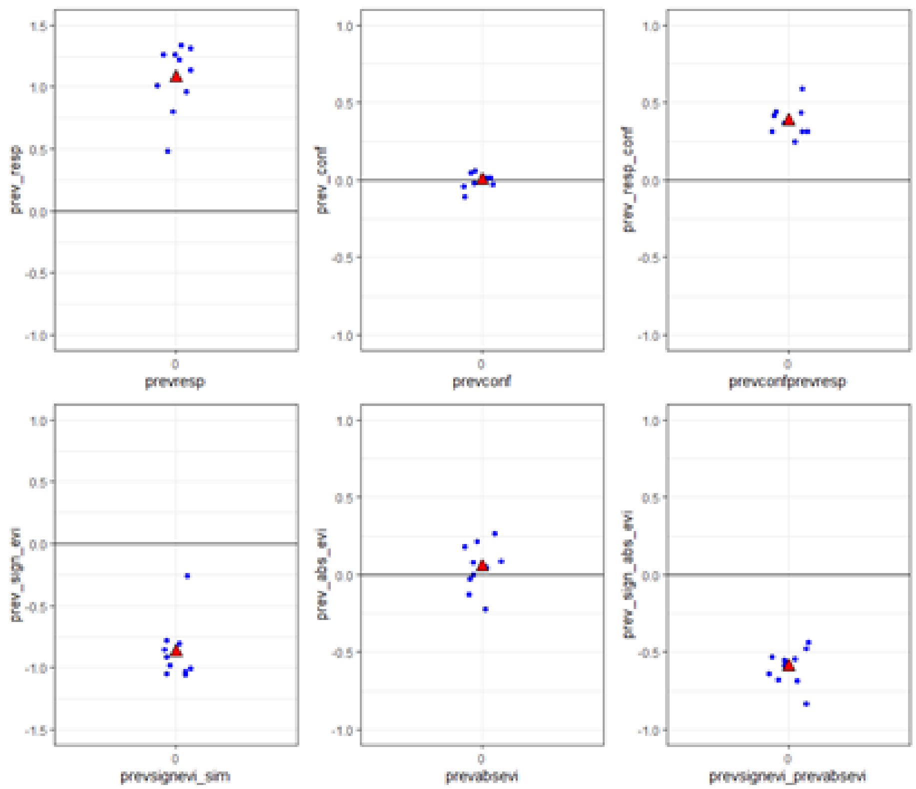

Figure 7: Parameter recovery when slow drifts are not estimated.

Below the result when slow drifts are estimated. Even though the apparent systematic updating is somewhat reduced, the model clearly fails to correctly identify the parameters.

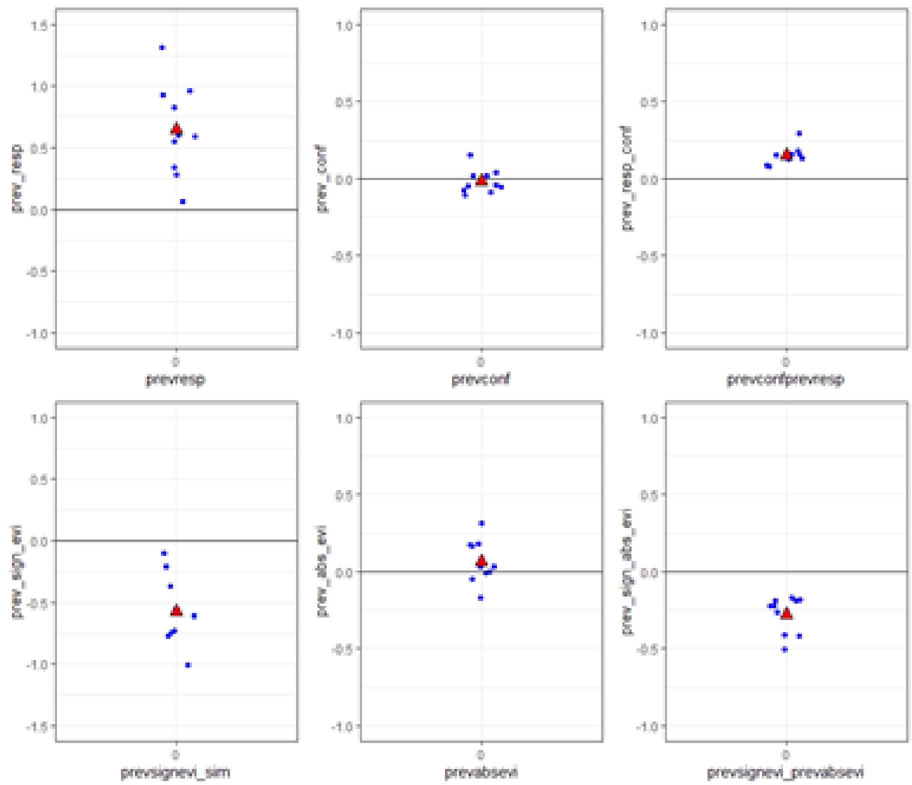

Figure 8: Parameter recovery when slow drifts are estimated.

3 Fitting on beehives

What are the results when we fit the model to the beehives dataset? First of all, fitting the model per subject does not work. If we estimate the same participant 10 times, the estimated slow drifts and parameters change every time so we cannot trust this. If we estimate the data all at the same time, like one super-participant, the fits are stable, but we can't do any hypothesis testing. So can we estimate the slow drifts and add these to the mixed models?

First of all, a model where we predict response based on only drift performs much better than our usual model where response is predicted based on all the systematic updating variables (AIC: Δ -704, BIC: Δ -748).

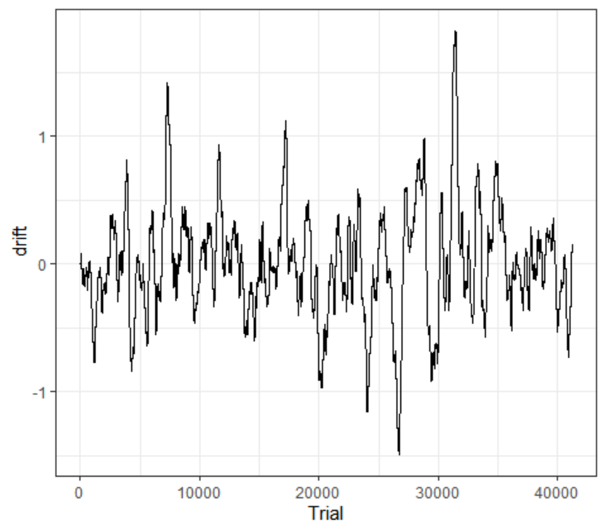

Figure 9: Estimated slow drifts when treating beehives dataset as one super-participant.

Below we see the parameter estimates from our mixed models when the estimated drift is added.

```
Fixed effects:
                                            Estimate Std. Error z value Pr(>|z|)    
(Intercept)                                  0.18036    0.02738   6.588 4.47e-11 ***
evidence_scaled                             15.69747    0.17157  91.493  < 2e-16 ***
prev_abs_evidence_scaled                    -0.05261    0.09791  -0.537 0.591039    
prev_sign_evidence1                         -0.02570    0.02463  -1.044 0.296603    
prev_resp1                                  -0.09178    0.02665  -3.444 0.000574 ***
prev_conf_scaled                             0.04377    0.03605   1.214 0.224724    
drift                                        1.29777    0.03654  35.519  < 2e-16 ***
prev_abs_evidence_scaled:prev_sign_evidence1 -0.16144   0.10197  -1.583 0.113371    
prev_resp1:prev_conf_scaled                  0.12892    0.03811   3.383 0.000718 ***
```

Figure 10: Parameter estimates from mixed model.

Below the hypothesis testing for the parameter estimates using the Wald test. What we see is that the repulsive effect of previous sign evidence, and the interaction previous sign evidence * previous absolute evidence, disappear when drift is added to the model. The effect of previous response * previous confidence still holds. Note that these results stay the same when we fit the LDS model with systematic updating.

```
Analysis of Deviance Table (Type II Wald chisquare tests)

Response: resp
                                            Chisq Df Pr(>Chisq)    
evidence_scaled                         8370.8829  1  < 2.2e-16 ***
prev_abs_evidence_scaled                   0.3758  1  0.5398774    
prev_sign_evidence                         3.7350  1  0.0532834 .  
prev_resp                                  3.4793  1  0.0621410 .  
prev_conf_scaled                           2.1782  1  0.1399808    
drift                                   1261.6158  1  < 2.2e-16 ***
prev_abs_evidence_scaled:prev_sign_evidence   2.5066  1  0.1133713    
prev_resp:prev_conf_scaled                11.4419  1  0.0007181 ***
```

Figure 11: Significance for parameter estimates from mixed model.

If we plot the effect of drift on response, we see that the more positive the drift, the higher the probability of a right response. Note that this is the opposite of what you would expect from SDT. In SDT a positive drift should correspond to an increased probability for left response. However, in the LDS model drift is added to the regression to calculate the probability of a right response. Thus, a higher drift results in a higher probability of a right response. Note that if we would change the sign of the drift, the interpretation and its dynamics is the same as in SDT.

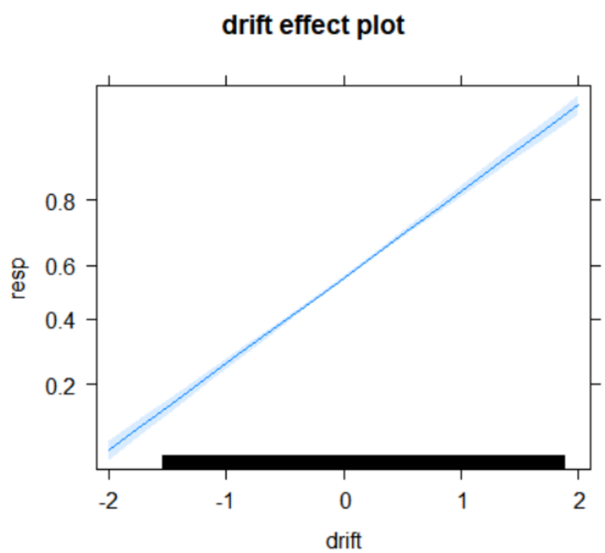

Figure 12: Relationship drift and response.

So only the effect of previous confidence remains. Does this mean that there is an effect of previous confidence? Not so sure about this. In the simulations you can clearly see that the estimated slow drift can capture the global tendency, but it often leaves a lot of variance unexplained (see below). So maybe this effect of previous confidence remains because our estimates are not precise enough.

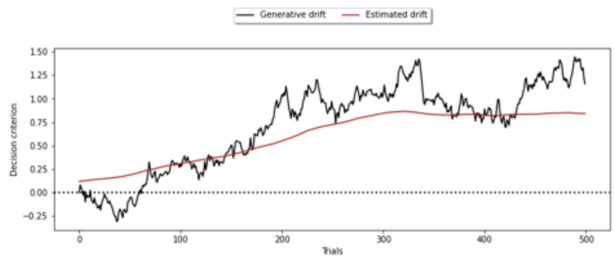

Figure 13: Estimated slow drift trajectory is quite coarse.

An interesting finding though is the relation between confidence and these slow drifts. From the SDT framework where confidence is calculated by comparing evidence to the criterion, you can expect that over - or underconfidence can occur in function of a drifting criterion. This is also what we see in the data. In the figure below the y-axis is confidence and the x-axis is the estimated criterion divided in 20 quantiles with their mean plotted. The stripes are the observed data and the line is the predicted confidence based on a polynomial regression with drift as predict up til power 5. We see that a more extreme criterion is related to higher confidence.

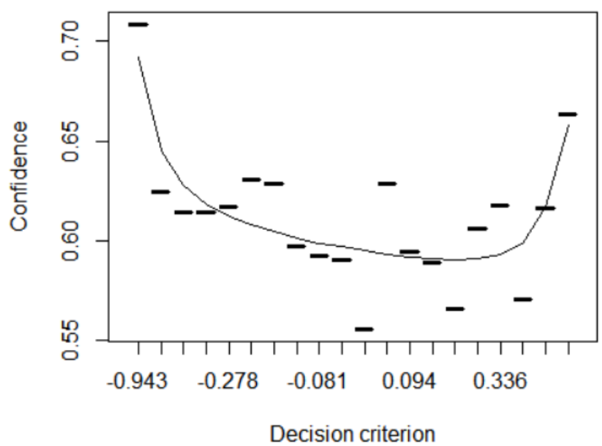

Figure 14: Relationship confidence and criterion quantiles.

In addition, if you predict confidence based on evidence, drift, and their interaction, only the interaction becomes significant (see below).

```
Type III Analysis of Variance Table with Satterthwaite's method
                      Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
drift                 0.3473  0.3473     1 15222  2.0010    0.1572
evidence_scaled       0.1870  0.1870     1 41126  1.0774    0.2993
drift:evidence_scaled 8.1893  8.1893     1 41126 47.1782 6.574e-12 ***
```

Figure 15: Confidence is predicted by an interaction between evidence and criterion, consistent with what can be expected from SDT.

Below a figure that explains the interaction between drift and evidence. To ease the interpretation, I changed the sign of the drift such that it can be interpreted as a criterion in SDT (see earlier). Evidence is scaled between -1 (max. left evidence) and 1 (max. right evidence). If we look at the top right panel when evidence = 1, then we see that confidence is the highest when the criterion is shifted to the left (becomes negative), but becomes lower when shifted to the right (becomes positive). This makes perfect sense from SDT perspective. With right evidence, the more the criterion is shifted to the left, the larger the distance between the evidence and the criterion becomes. Thus, confidence increases.

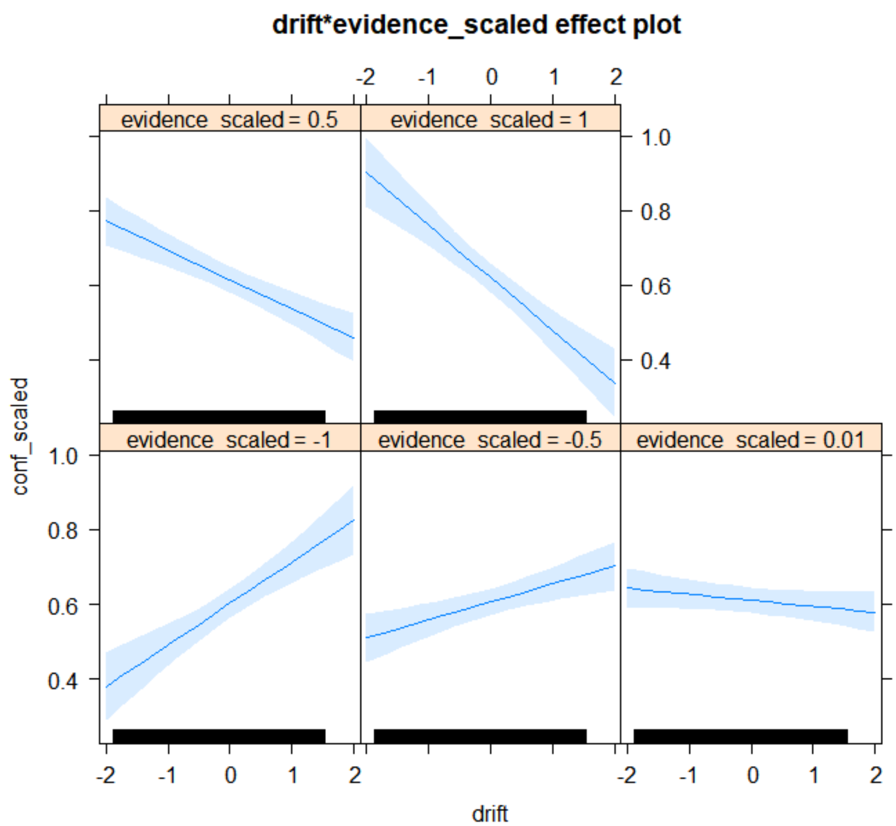

Figure 16: Confidence is predicted by an interaction between evidence and criterion, consistent with what can be expected from SDT.

4 Interim conclusion

- The parameter recovery shows that the model shows some biases in the parameter estimates, especially when simulating with previous response, previous confidence,... Could these biases be related to the slow drifts being estimated too coarse? I.e., the remaining uncapture variance in slow drifts is soaked up by the parameters of the systematic updating, hereby inducing these biases.
- If we estimate the drifts in the beehives dataset and add them to the mixed models, the effect of previous evidence disappears, and the effect of previous confidence is somewhat reduced but stays significant.
- In the beehives dataset confidence is predicted by the interaction between evidence and criterion, consistent with what we would expect from SDT where confidence results from comparing evidence with the criterion.
- We need to find a way to improve the slow drift fit, and figure out how to fit the model to the data without treating it as one super-participant.

5 Random stuff

- Model where AR coefficient is freely estimated does not work yet. The parameter estimates fail miserably.
- If you do not estimate slow drifts with the LDS model, then the parameter estimates are exactly the same as if you run a simple regression model in R per subject.
- If you simulate from a SDT framework with a biased but stationary criterion, the apparent sequential effects will not appear. So a drifting criterion is needed in order to create these artificial effects.
- Can slow drifts be linked with variability in confidence ratings? Seems like it, see figure 16.
- Can slow drifts bias the estimates of d'? Z(H) - Z(FA) with an extreme criterion you have more hits and also more false alarms
- In the DDM there were a lot of problems with autocorrelation is the parameter so I had to use a thinning of 20, which is a lot. Could there be any link with slow drifts, inducing autocorrelation?
- If we manage to get accurate trial-by-trial estimates for the slow drifts, add it to the DDM regression module. Would be interesting to see whether the effects of previous response, previous confidence, and previous stimulus, on the drift criterion (drift bias) disappear, and instead are explained by these slow drifts.