Recovering Bistable Systems from Psychological Time Series

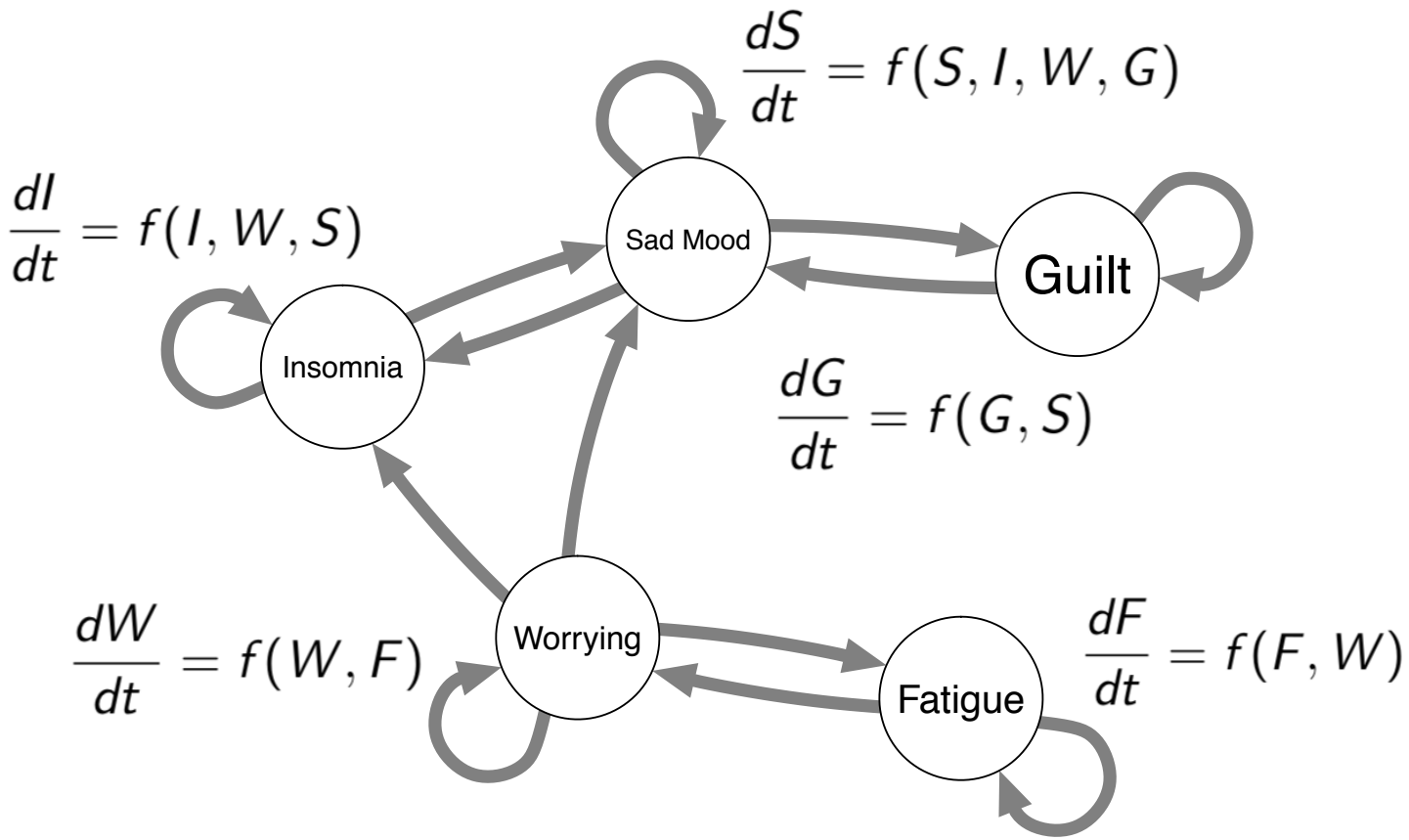
Jonas Haslbeck
University of Amsterdam
www.jonashaslbeck.com

and

Oisín Ryan
Utrecht University
ryanoisin.github.io

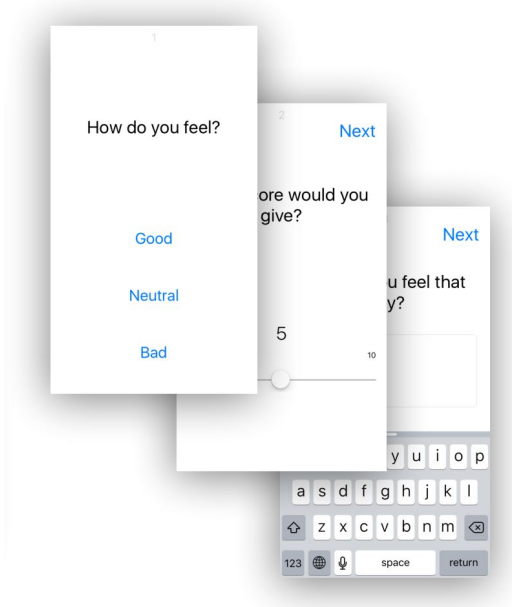
DynaNet Meeting; December 4th, 2019

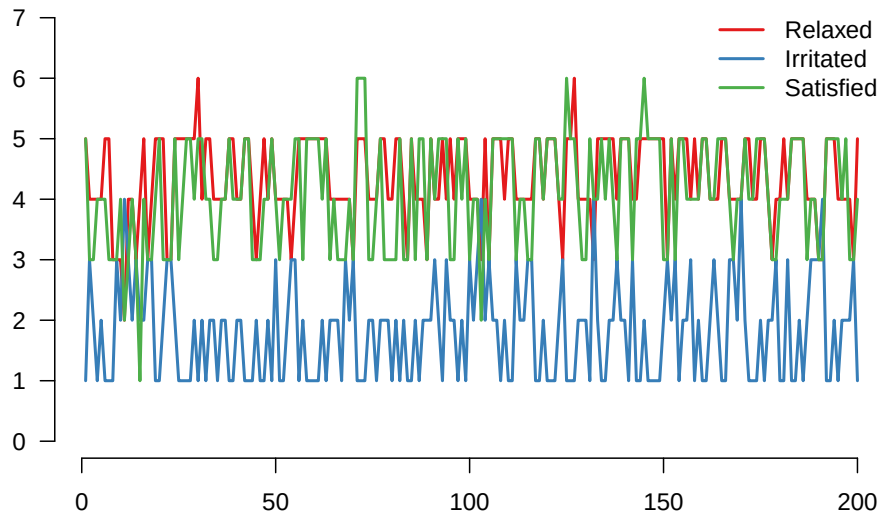
Mental Disorders as Complex Systems



Kendler et al. (2011); Borsboom & Cramer (2013); Borsboom (2017)

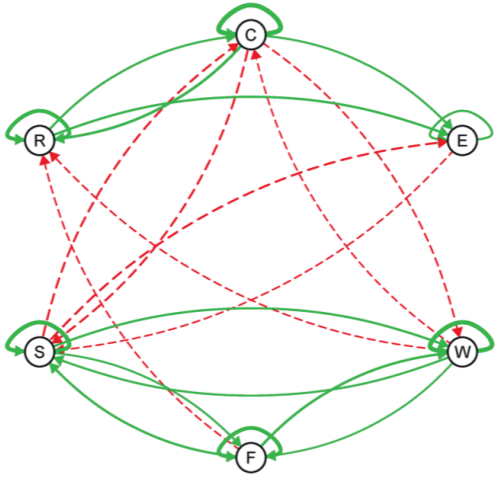
Recovering Dynamical Systems from ESM Time Series



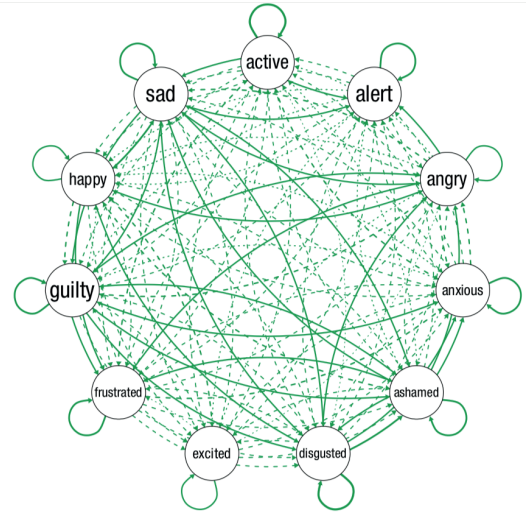


Can we recover (properties of) such systems from ESM time series?

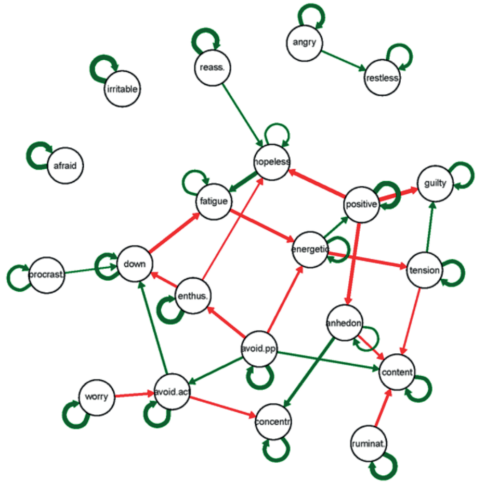
Recovering Dynamical Systems from ESM Time Series



"[...] potentially yield important insights into how the **dynamics of psychopathology** relate to intra- and inter-individual differences." (Bringmann et al., 2013)

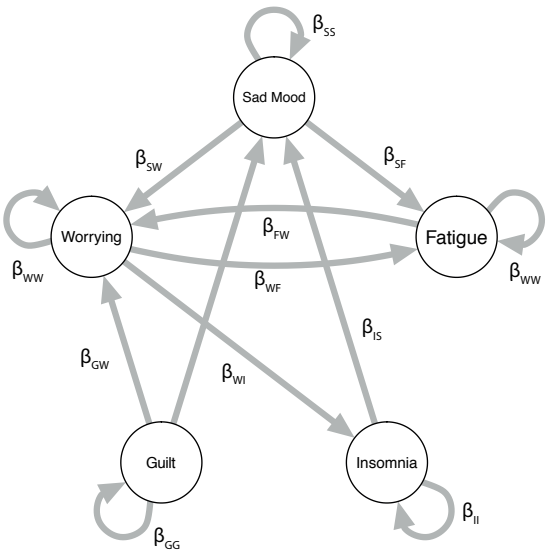


"We can use a network approach to examine the **temporal dynamics of the emotion system** that underlies MDD." (Pe et al., 2015)



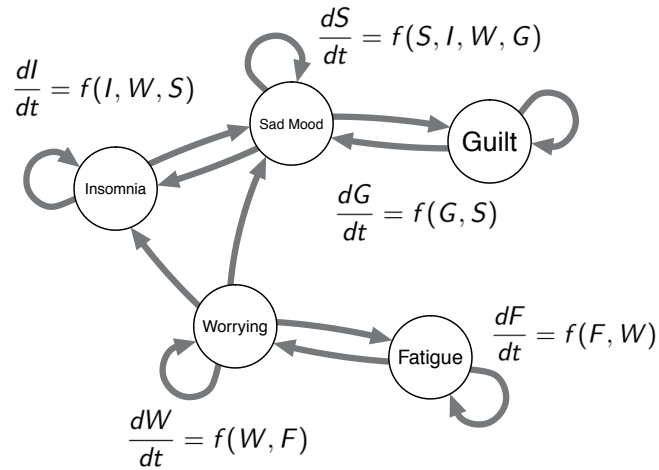
"[…] holds the potential to allow researchers to articulate and **examine complex dynamics** among constituent elements of a psycho-pathological system." (Fisher et al., 2017)

Recovering Dynamical Systems from ESM Time Series



Statistical Time Series Model
(e.g. VAR model)

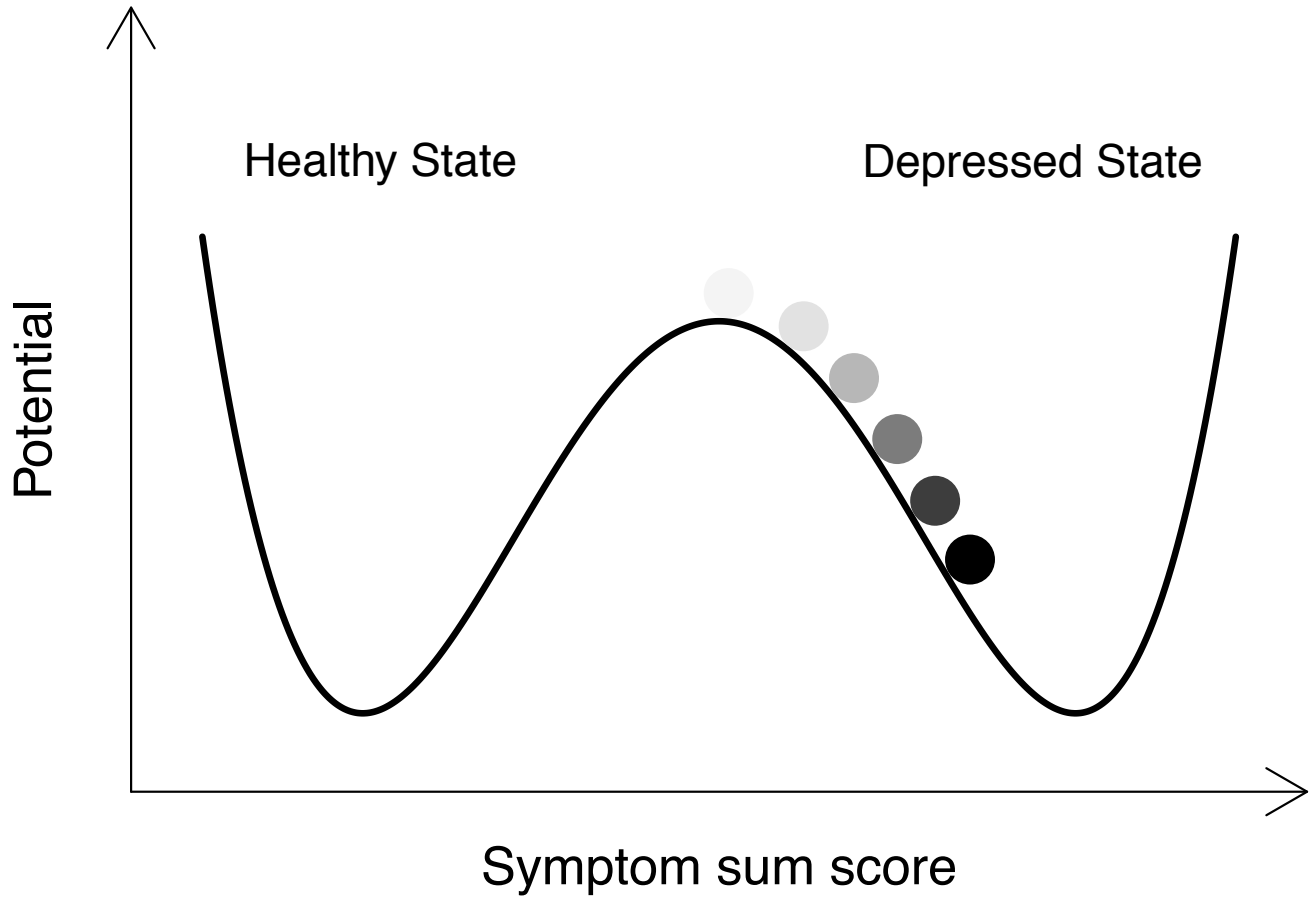
Inference?



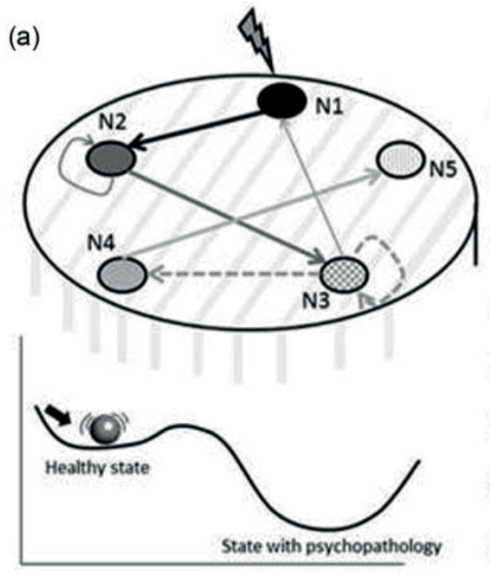
Data Generating System

Can we make straightforward inferences from statistical models to the data generating system?

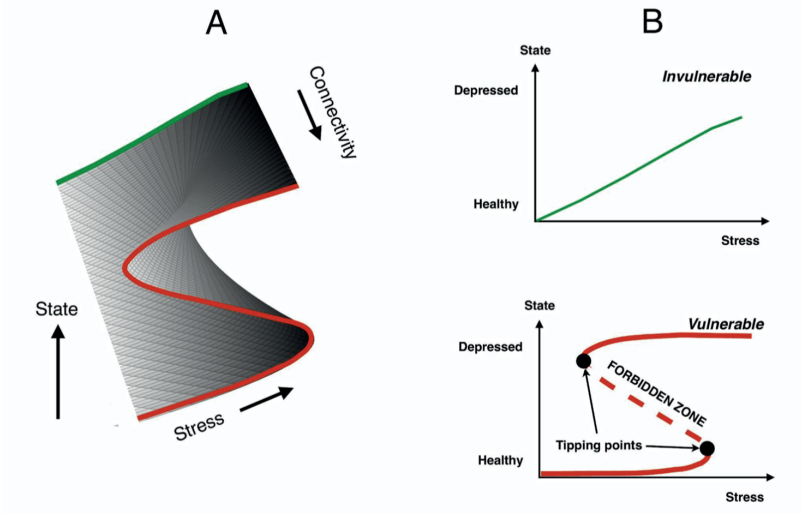
Mental Disorders as Bistable Systems



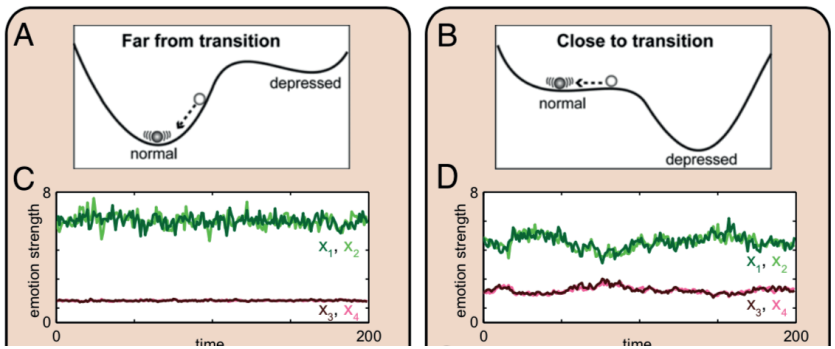
Mental Disorders as Bistable Systems



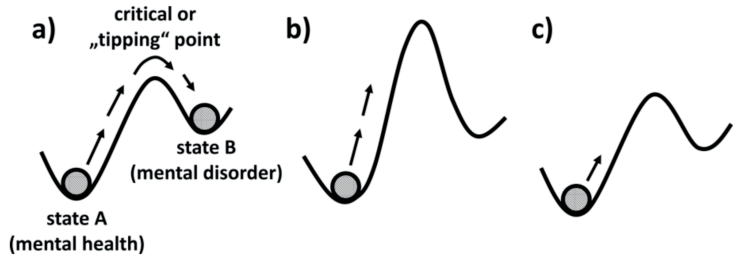
Wichers et al. (2015); Wichers et al. (2018)



Cramer et al. (2016)



van Leemput et al. (2014)



Kalisch et al. (2019)

Bistable System of Emotion Dynamics

Variables: Cheerful (x_1), Content (x_2), Anxious (x_3), and Sad (x_4)

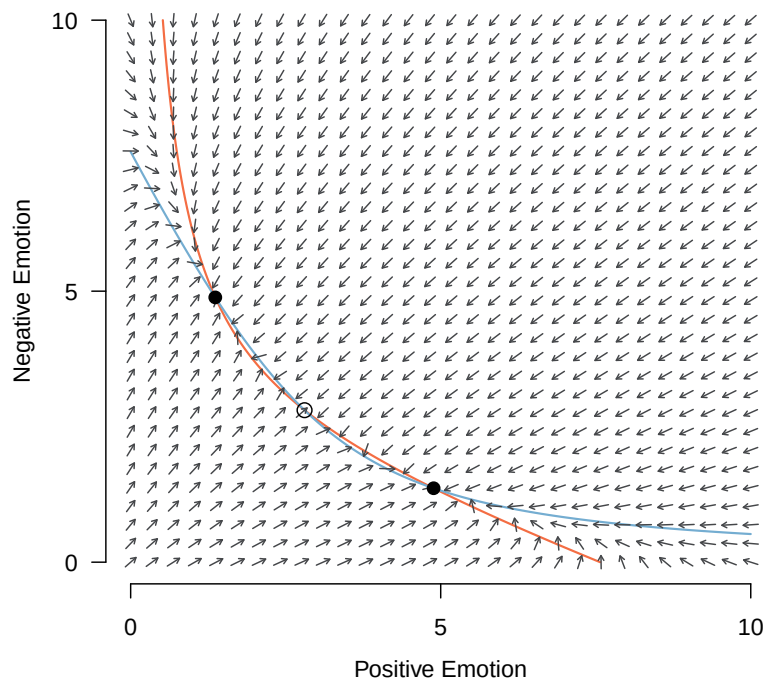
$$\frac{dx_i}{dt} = x_i + \sum_{j=1}^{4} C_{i,j} x_j x_i + 1.6 + \epsilon_i,$$

with

$$C = \begin{bmatrix} -0.2 & 0.04 & -0.2 & -0.2 \\ 0.04 & -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.04 \\ -0.2 & -0.2 & 0.04 & -0.2 \end{bmatrix}$$

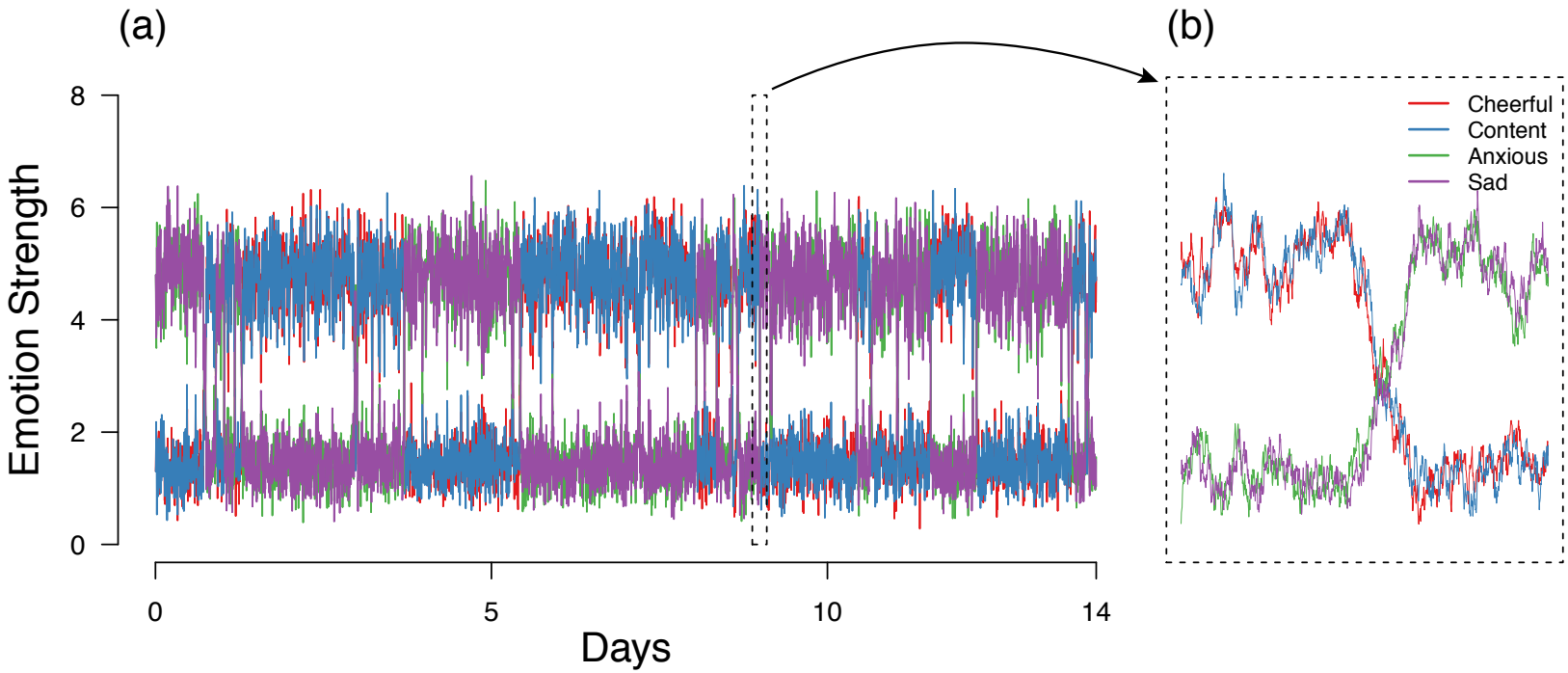
and $\varepsilon_i \sim N(0, \sigma^2)$.

Vectorfield of 2D System:



(Same Lotka-Volterra model with similar parameters as in van Leemput et al., 2014)

Two Weeks of Simulated Data



- 14 days
- Measurement every six seconds
- $14 \times 24 \times 60 \times \frac{60}{6} = 201600$ data points
- Around 17 "switches"

Properties of Bistable System

Global dynamics

1. Bistability (two stable fixed points)
2. Position of fixed points
3. Variability around fixed points
4. Frequency of transitions

Microdynamics

5. Suppressing effects between valences, reinforcing effects within valences
6. Relative size of suppressing/reinforcing effects
7. All parameters are independent of time and independent of variables outside the model

Question 1:

Which inferences can we make from typical statistical models to the true system **with ideal data**?

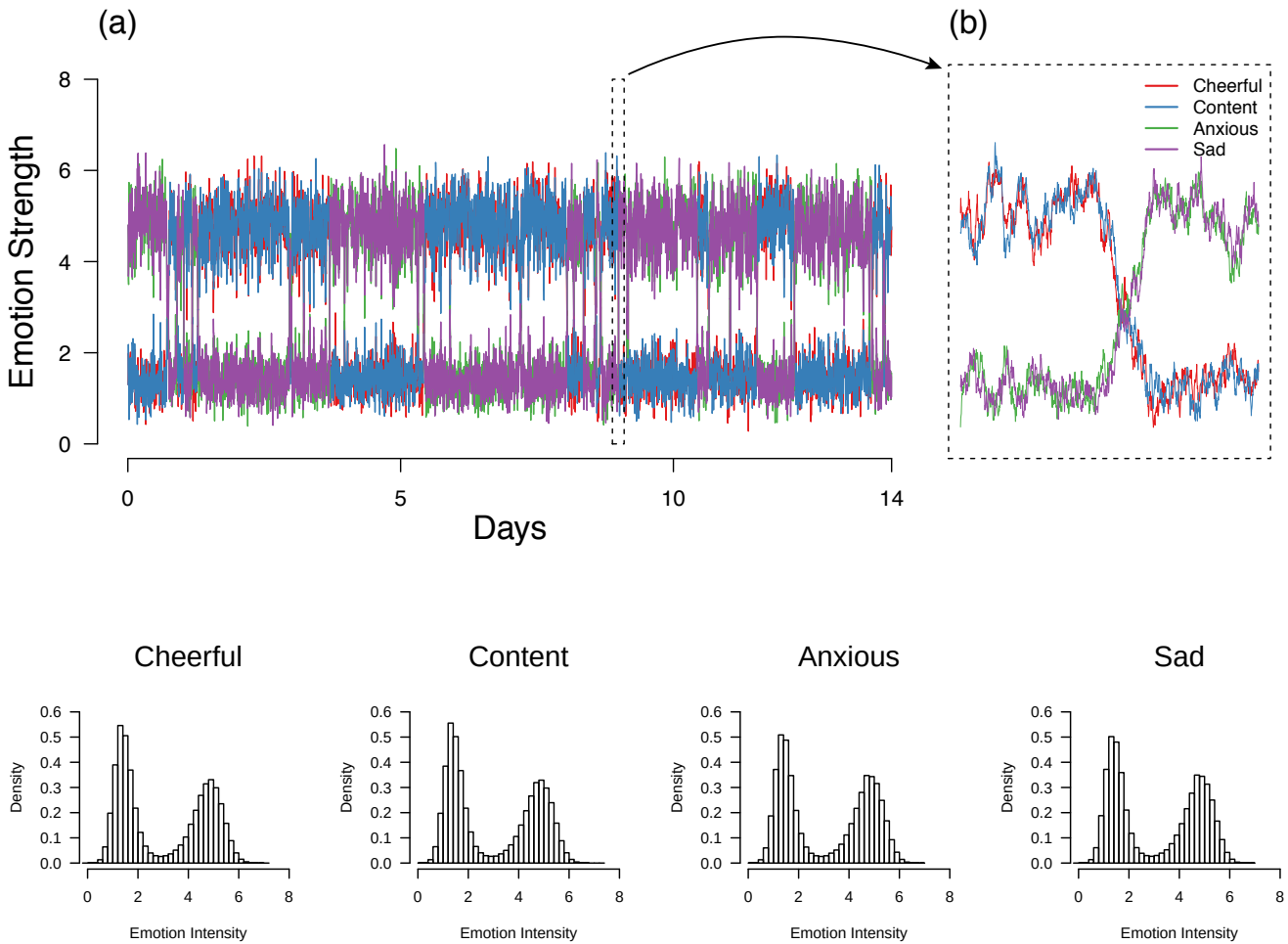
Considered Methods:

- Descriptive Statistics / Data Visualization
- Hidden Markov Model
- Lag-0 / Gaussian Graphical Model
- Lag-1 / VAR model
- Threshold VAR
- Differential Equation Estimation

Considered Methods:

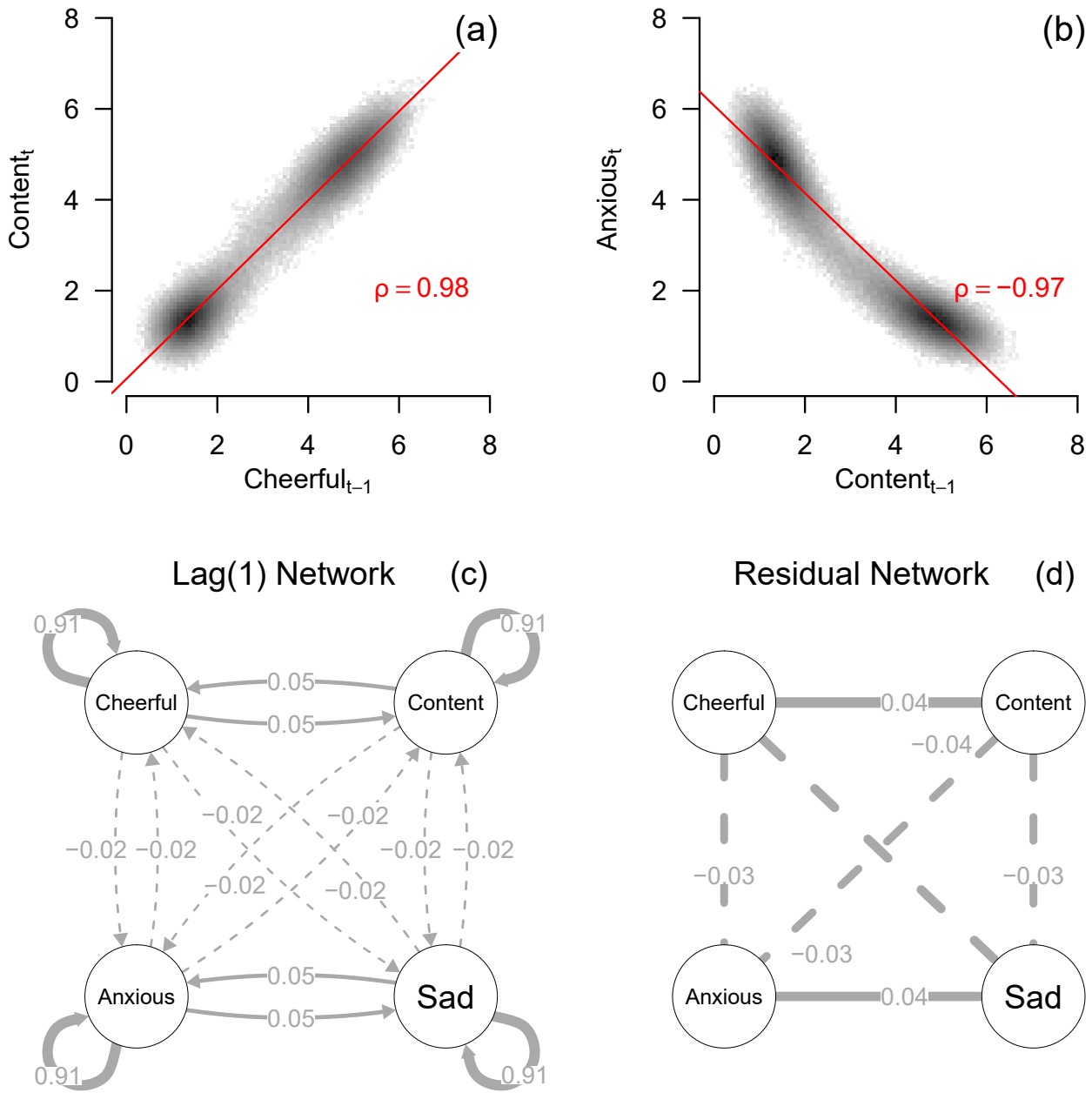
- Descriptive Statistics / Data Visualization
- Hidden Markov Model
- Lag-0 / Gaussian Graphical Model
- Lag-1 / VAR model
- Threshold VAR
- Differential Equation Estimation

Descriptive Statistics and Data Visualization

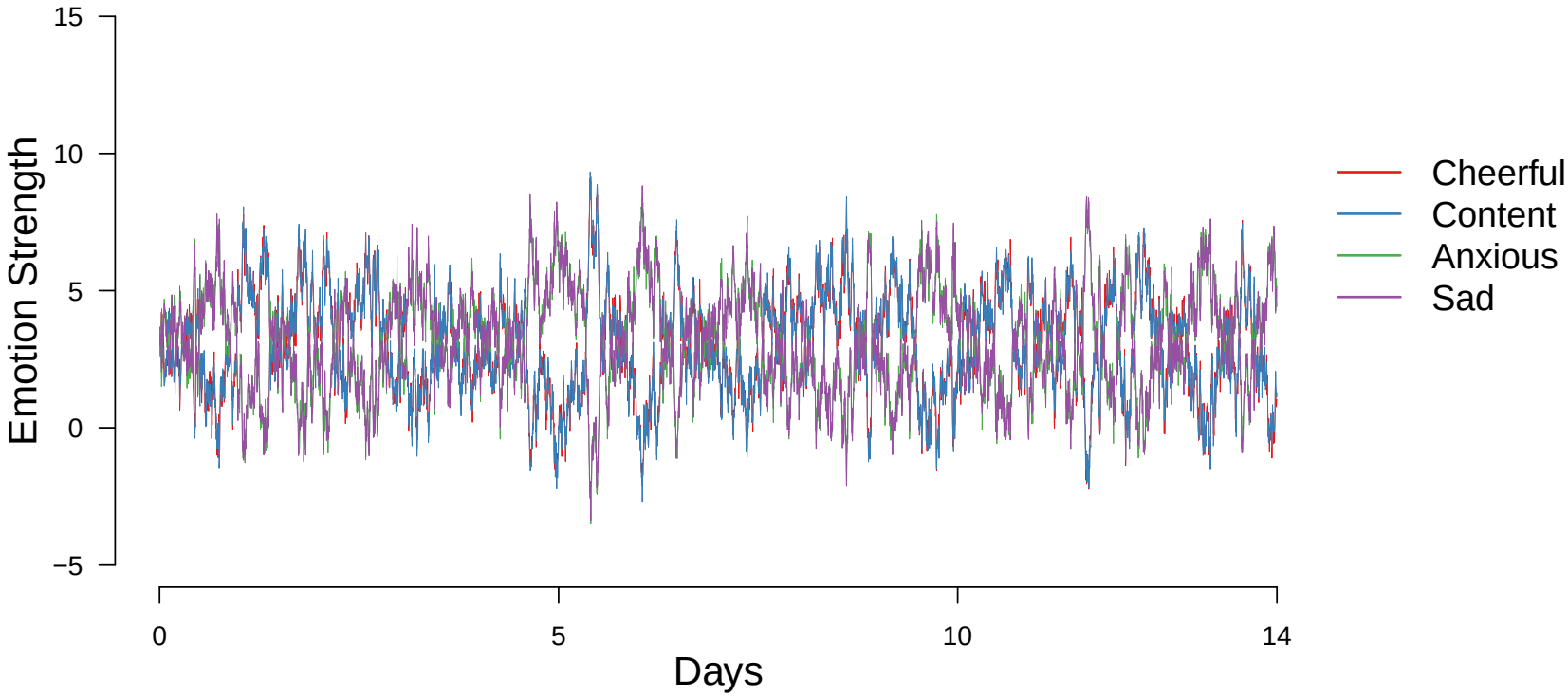


- We can roughly recover global characteristics: bistability, position and variance of fixed points, frequency of switching

Lag-1, VAR(1) and Residual GGM



Data Generated from VAR(1) Model



- Single fixed point around 3 (average of two actual fixed points)

Differential Equation Estimation

Approach:

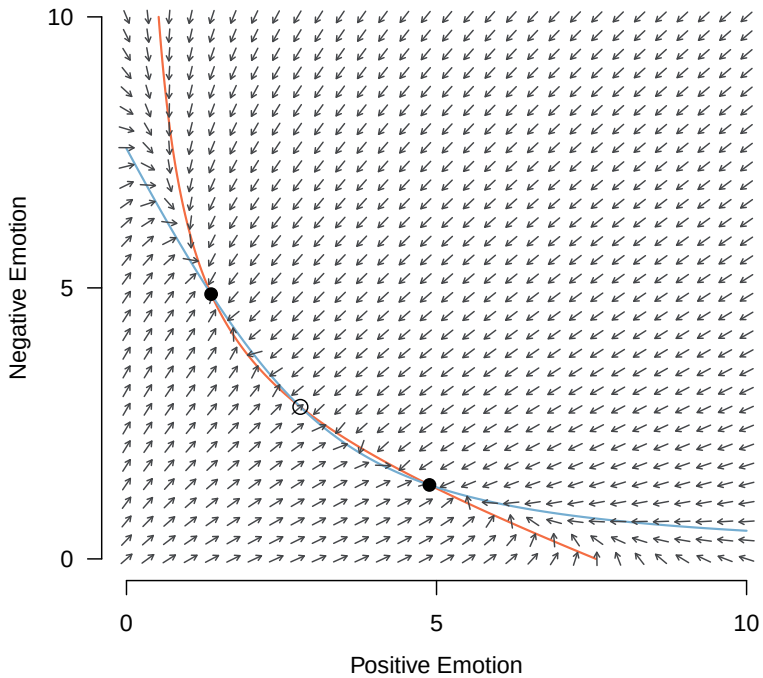
1. Approximate derivatives $\frac{dx_{i,t}}{dt} \approx \frac{x_{i,t+1} - x_{i,t}}{\Delta t}$
2. Regress all variables on the derivative

Model comparison:

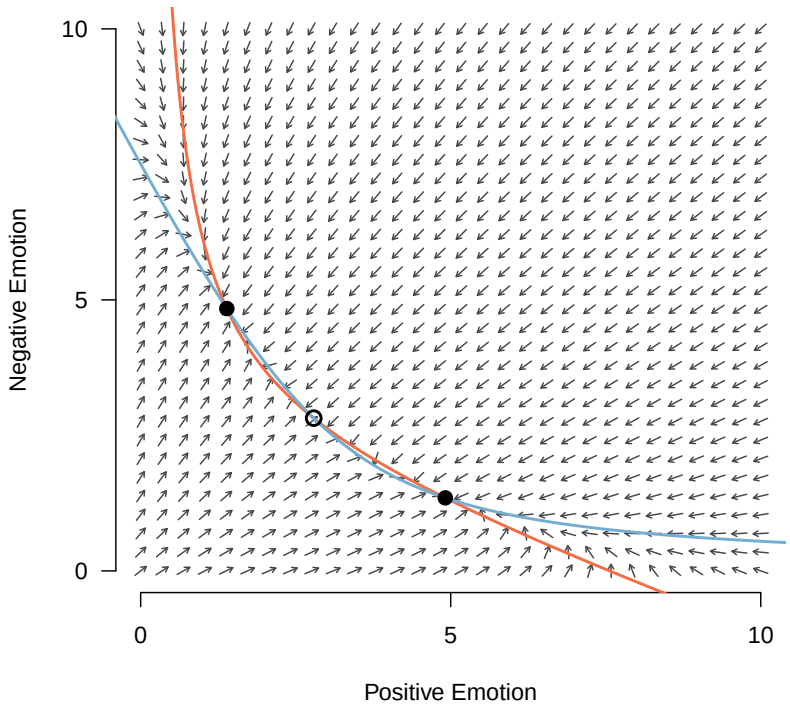
Model	$\frac{dx_{i,t}}{dt} \sim a + r_i x_i + \ldots$	q	R^2
A	$\sum_{j \neq i} r_j x_j$	5	0.04464
B	$\sum_{j \neq i} R_{ij} x_j + \sum_j^p C_{ij} x_j x_i$	9	0.06874
C	$\sum_{j \neq i} R_{ij} x_j + \sum_{(j,k)}^p \beta_j x_j x_k$	15	0.06870

Results: Differential Equation Estimation

True Vectorfield

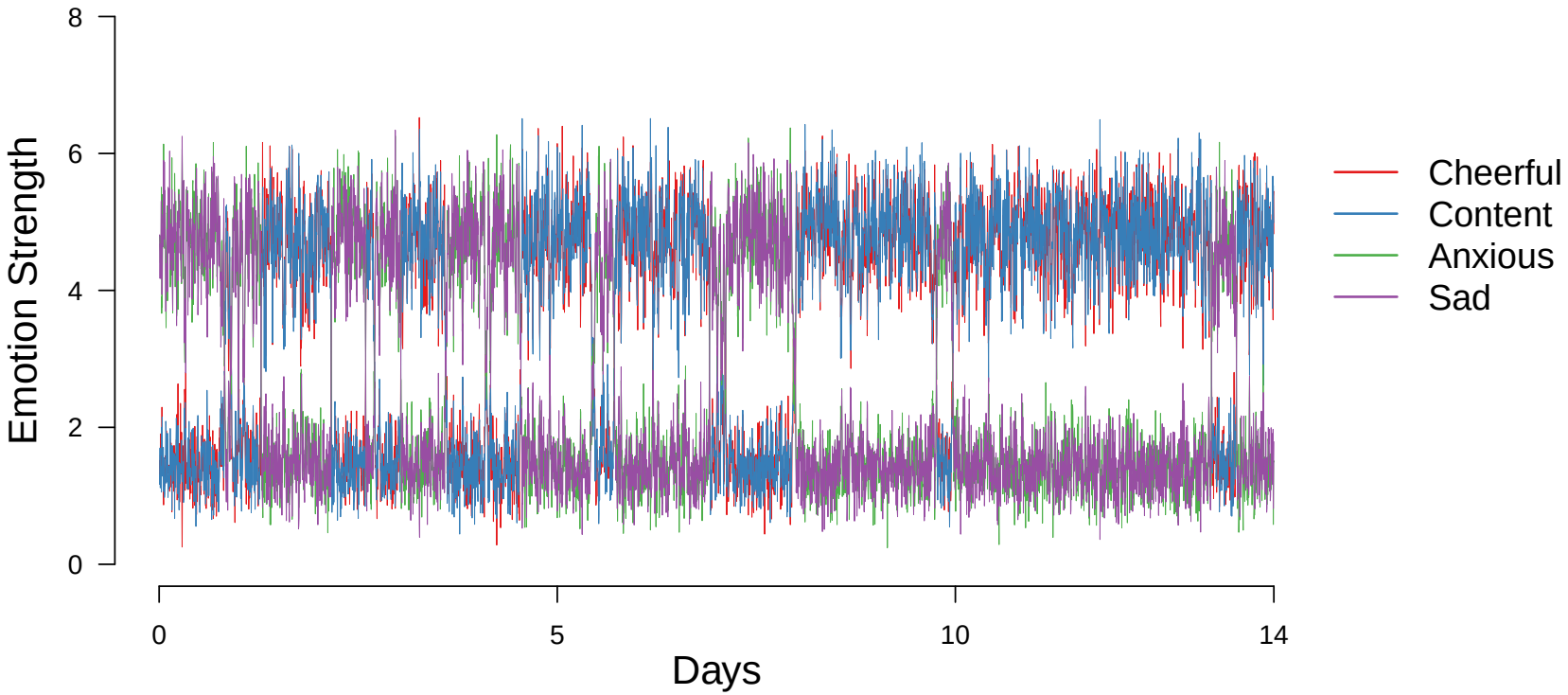


Estimated Vectorfield



- Dynamics of estimated DEs is almost identical to true dynamics.

Data Generated from Estimated DEs



- Very similar behavior compared to original time series

Summary: Ideal Time Series

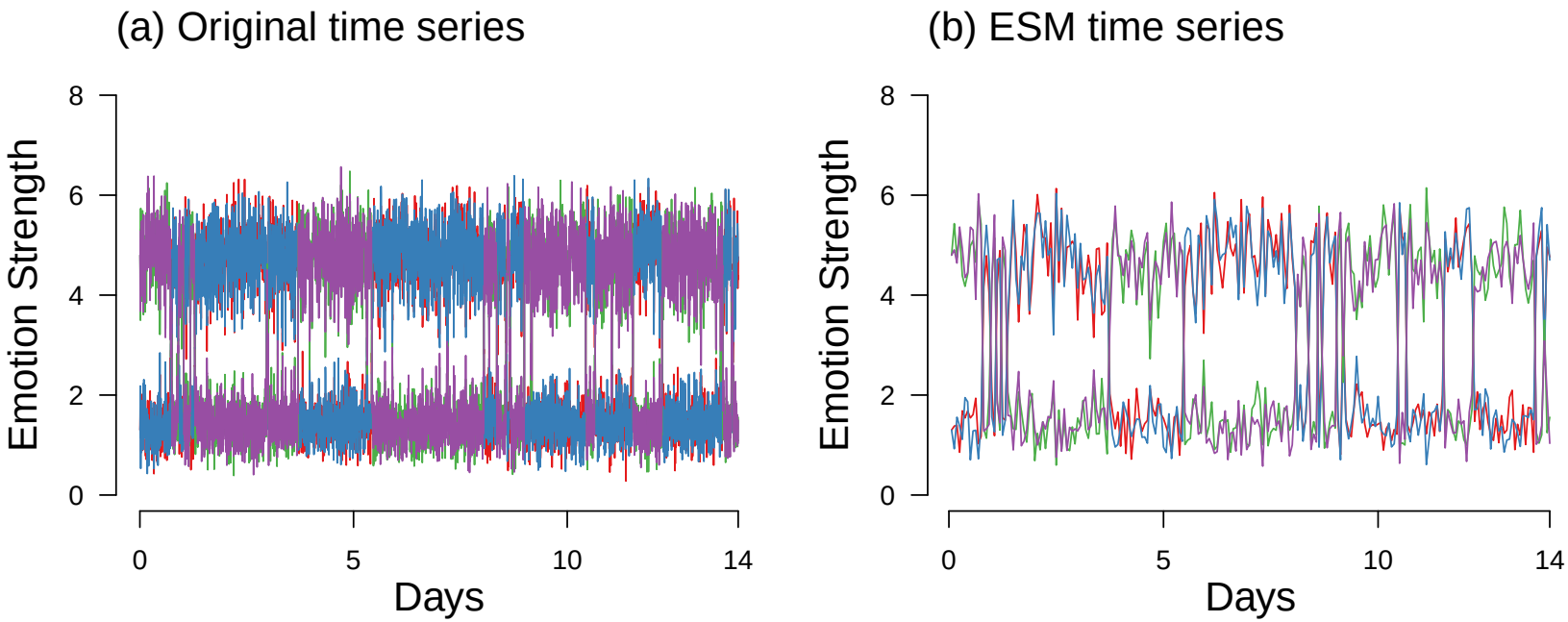
	Bistability (1)	Position (2)	Variance (3)	Transitions (4)	Suppr./Reinf. (5)	Relative Size (6)	Time-constant (7)
Data Visualization	✓	✓	✓	×	×	×	×
HMM	✓*	✓	✓	✓	×	×	×
Lag-0 / GGM	×	×	×	×	✓	×	✓*
Lag-1 / VAR(1)	×	×	×	×	✓	×	✓*
TVAR(1)	✓*	✓	✓	✓	✓	×	×
DE-Estimation	✓	✓	✓	✓	✓	✓	✓*

- Data visualization / HMM recover global dynamics
- Only DE estimation recovers microdynamics

Question 2:

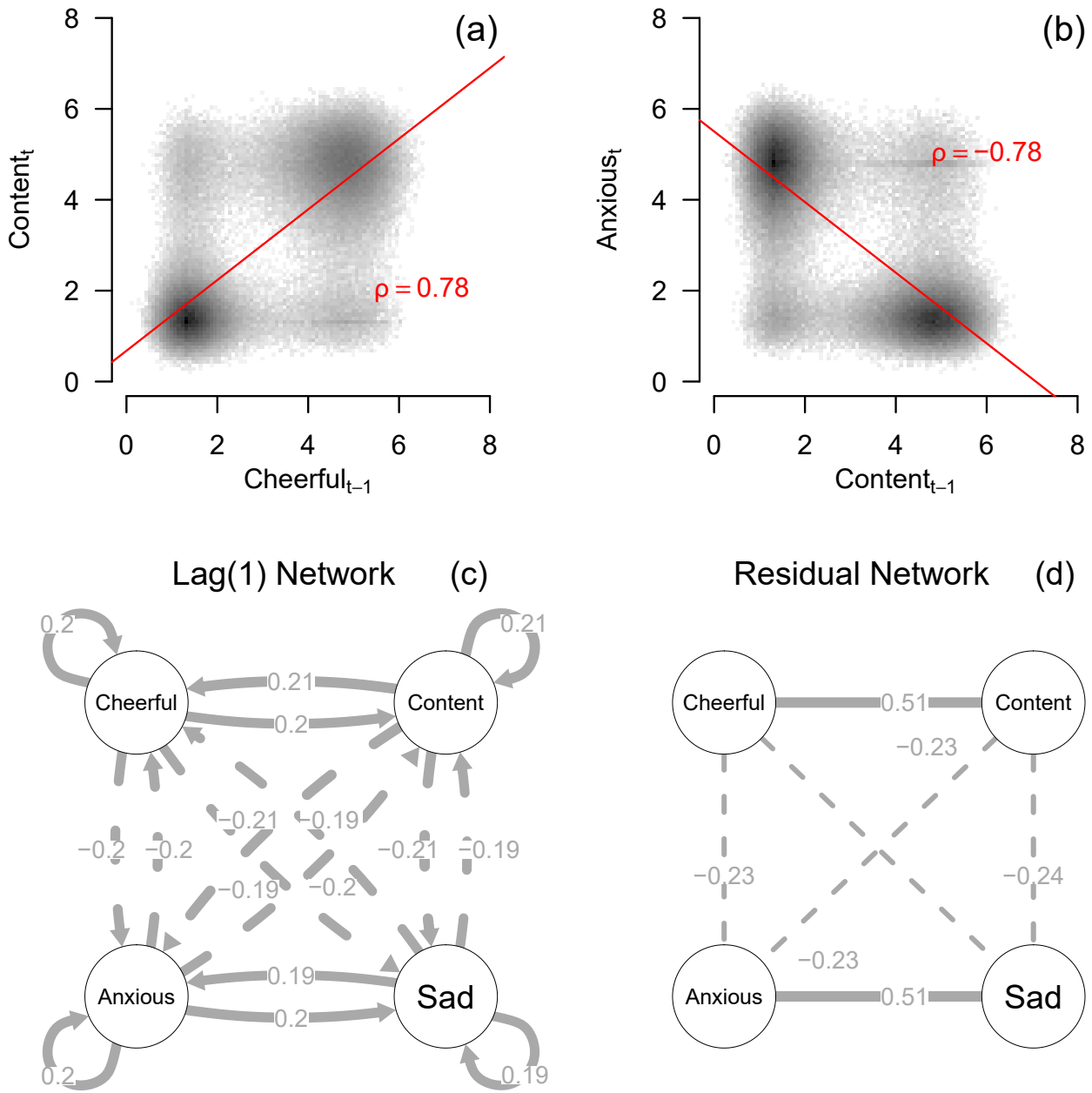
Which inferences can we make from typical statistical models to the true system **with ESM data**?

Emulated ESM Data



- We take a measurement every 90 minutes
- Results in $14 \times 24 \times \frac{60}{90} = 224$ data points in two weeks
- We match the sample size by recording 1800 weeks ($\frac{224}{2} \times 1800 = 201600$)

Lag-1, VAR(1) and Residual GGM



Differential Equation Estimation

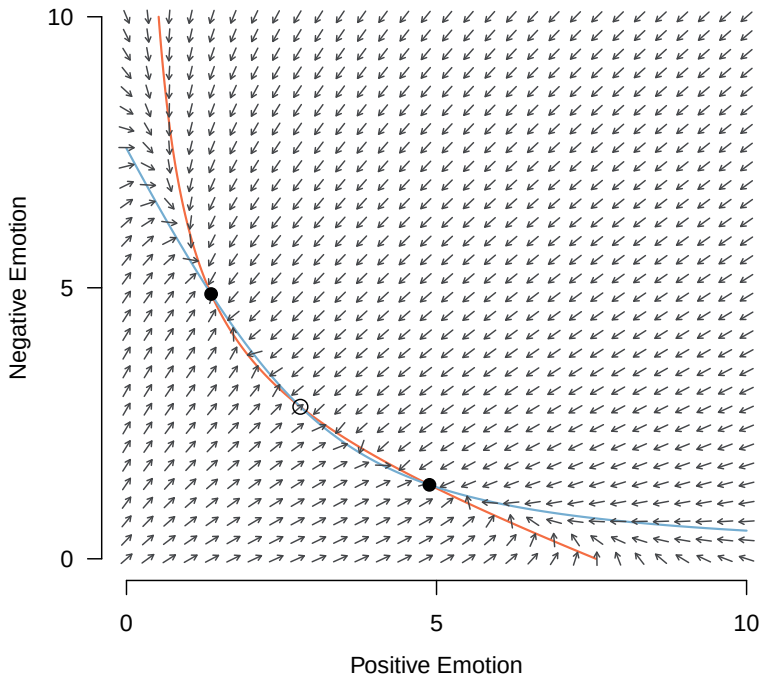
Model Selection:

Model	$\frac{dx_{i,t}}{dt} \sim a + r_i x_i + \ldots$	q	R^2
A	$\sum_{j \neq i} r_j x_j$	5	0.13991
B	$\sum_{j \neq i} R_{ij} x_j + \sum_j^p C_{ij} x_j x_i$	9	0.16827
C	$\sum_{j \neq i} R_{ij} x_j + \sum_{(j,k)}^p \beta_j x_j x_k$	15	0.16928
D	$\sum_{j \neq i} R_{ij} x_j + \sum_{(j,k)}^p \beta_j x_j x_k + \sum_j^p \gamma_j x_j^3$	19	0.19455
E	$\sum_{j \neq i} R_{ij} x_j + \sum_{(j,k)}^p \beta_j x_j x_k + \sum_j^p \gamma_j x_j^3 + \sum_{j \neq k \neq l}^p \zeta_j (x_j x_k x_l)$	23	0.19801
F	$\sum_{j \neq i} R_{ij} x_j + \sum_{(j,k)}^p \beta_j x_j x_k + \sum_j^p \gamma_j x_j^3 + \sum_{(j,k,l)}^p \zeta_j (x_j x_k x_l)$	35	0.19940
G	$\sum_{j \neq i} R_{ij} x_j + \sum_{(j,k)}^p \beta_j x_j x_k + \sum_j^p \gamma_j x_j^3 + \sum_{(j,k,l)}^p \zeta_j (x_j x_k x_l) + \sum_{(j,k,l,m)}^p \eta_j (x_j x_k x_l x_m)$	70	0.20420

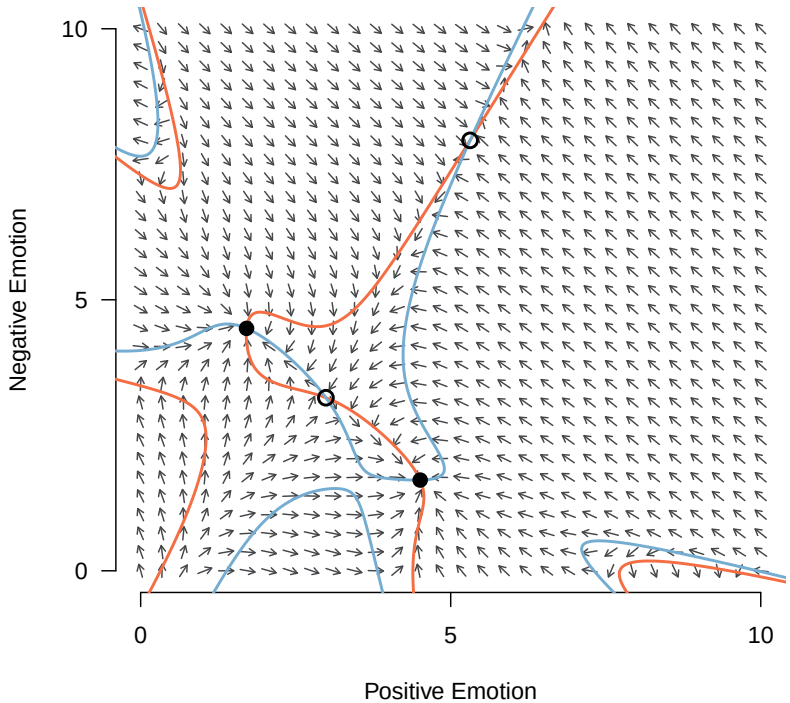
- We stopped evaluating more complex model, becase we added all possible interaction terms

Results: Differential Equation Estimation

True Vectorfield

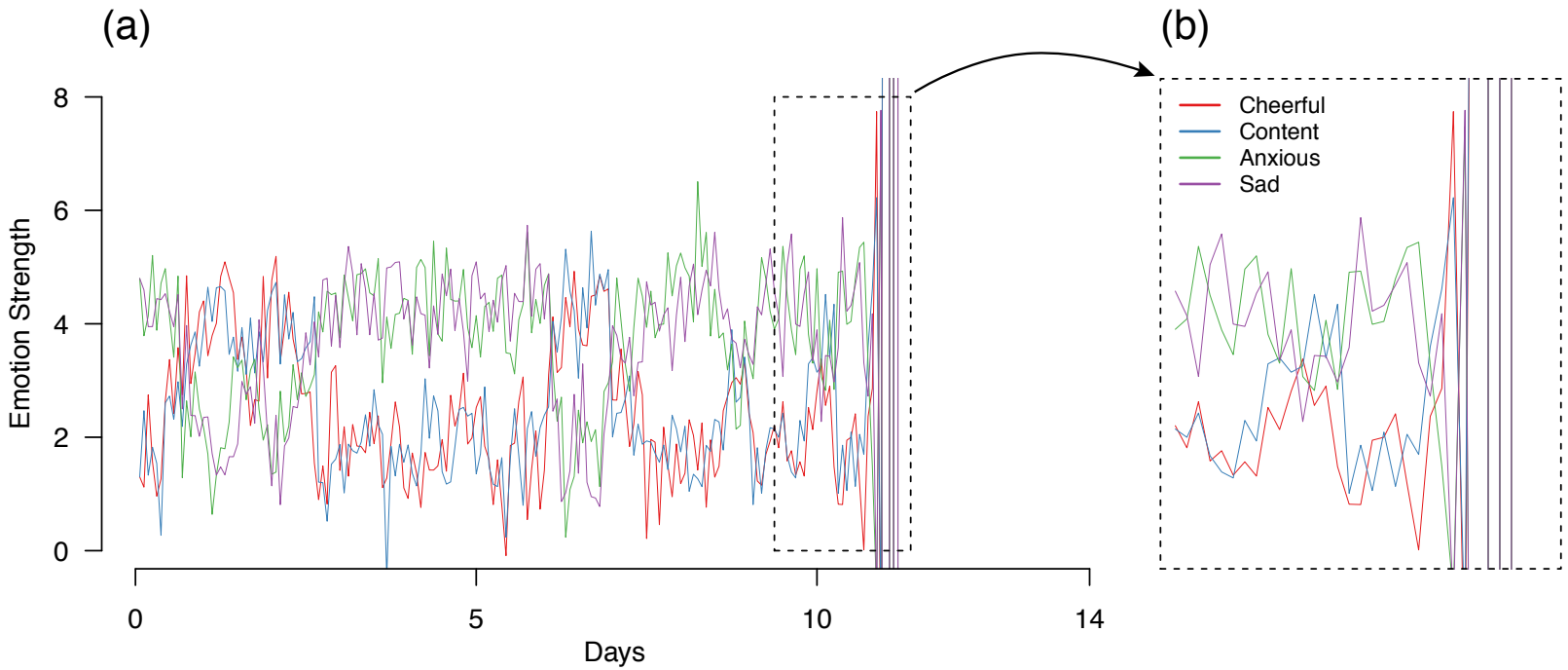


Estimated Vectorfield



- Dynamics of estimated DEs is very different to true dynamics.

Data Generated from Estimated DEs



- System shows bistable behavior but then diverges on day 11

Summary: ESM Time Series

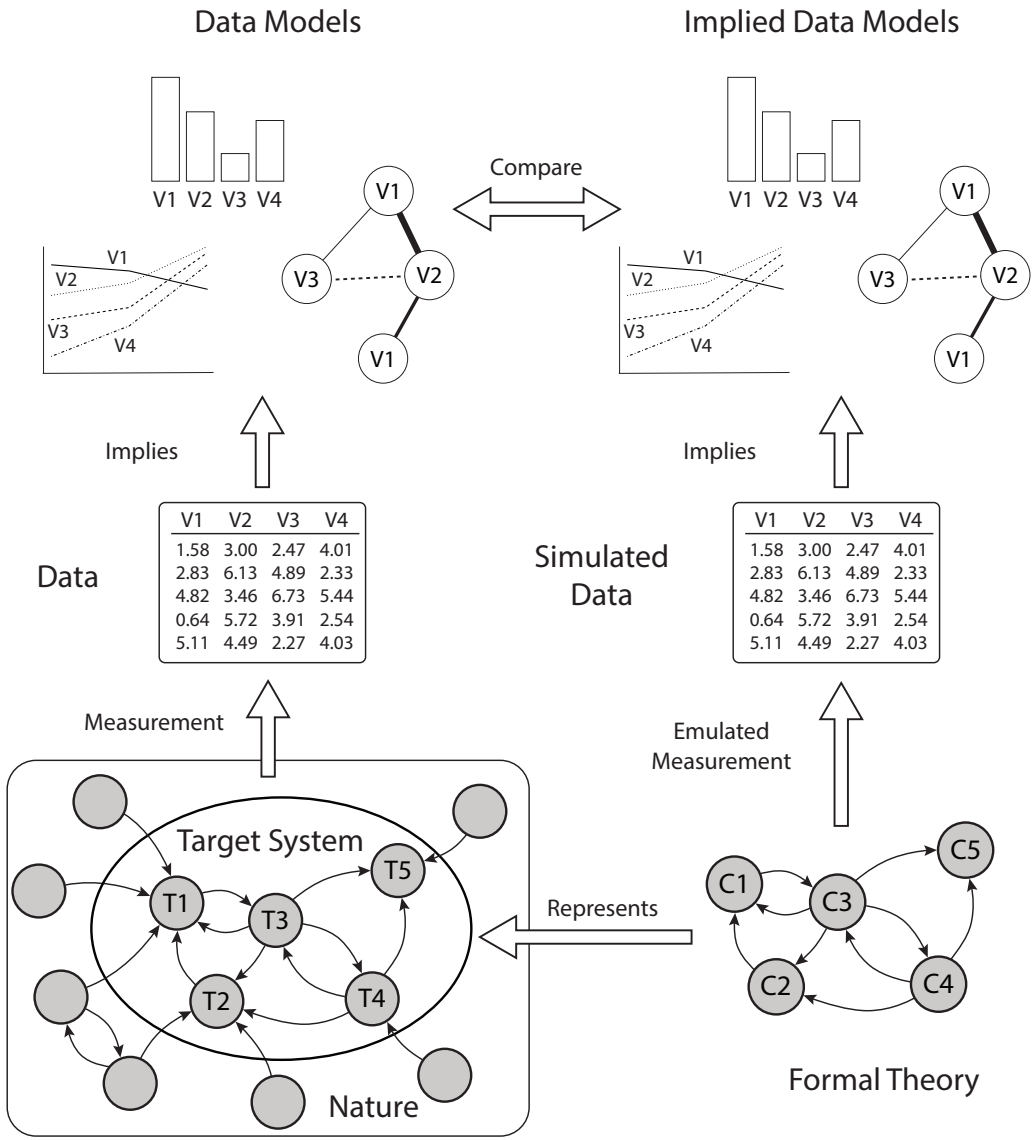
	Bistability (1)	Position (2)	Variance (3)	Transitions (4)	Suppr./Reinf. (5)	Relative Size (6)	Time-constant (7)
Data Visualization	✓	✓	✓	×	×	×	×
HMM	✓*	✓	✓	✓	×	×	×
Lag-0 / GGM	×	×	×	×	×	×	✓*
Lag-1 / VAR(1)	×	×	×	×	×	×	✓*
TVAR(1)	✓*	✓	✓	✓	×	×	×
DE-Estimation	×	×	×	×	×	×	×

- Data visualization / HMM again recover global dynamics
- No method recovers microdynamics (it's not in the data!)

Conclusions

1. Inference: Statistical Model $\rightarrow$ Data Generating System (Q1)
   - Generally unclear how parameters map to properties of data generating system
2. Ideal Time Series vs. ESM Time Series (Q2)
   - Microdynamics essentially lost entirely due to low frequency sampling (we only showed that this *can* be a problem!)
   - Global dynamics always well recovered, since they are not dependencies across time
3. Is there something wrong with using statistical models?
   - No, they correctly summarize the data!
   - But: We typically cannot make straightforward inferences from the model parameters to properties of the system

Way forward? Abductive Approach to Theory Construction



V1	V2	V3	V4
1.58	3.00	2.47	4.01
2.83	6.13	4.89	2.33
4.82	3.46	6.73	5.44
0.64	5.72	3.91	2.54
5.11	4.49	2.27	4.03

V1	V2	V3	V4
1.58	3.00	2.47	4.01
2.83	6.13	4.89	2.33
4.82	3.46	6.73	5.44
0.64	5.72	3.91	2.54
5.11	4.49	2.27	4.03

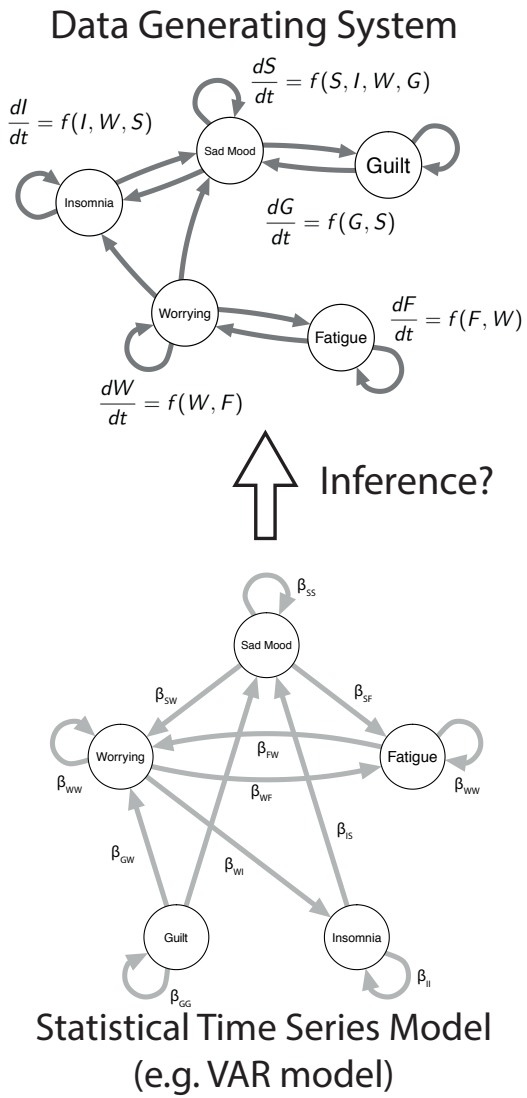
Haslbeck*, Ryan*, Robinaugh*, Waldorp & Borsboom (in preparation) Modeling Psychopathology: From Data Models to Formal Theories

Summary

- Inference from statistical models to properties of systems is not straightforward
- Low sampling frequency may preclude studying certain processes
- Possibly way forward: Abductive approach to formal theory construction

Preprint:

- https://psyarxiv.com/kcv3s

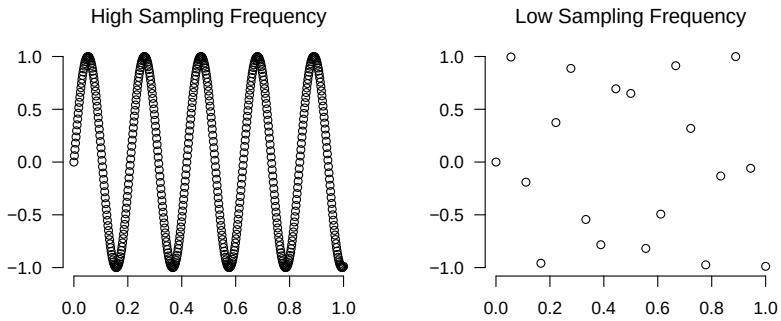


jonashaslbeck@gmail.com | www.jonashaslbeck.com

Extra Slides

What could possibly go wrong?

1. Misspecification / mapping problem
2. Sampling frequency too low



3. Time series too short