Additional file 1

Rewriting the SDM as a system of linear differential equations

To illustrate the process of rewriting the system dynamics equations, we consider a simple case example (Figure A1). Suppose we are studying the reinforcing feedback loop between sleep quality (S) and mood (M). Experts indicated that perceived stress (P) and inflammatory processes (I) should also be included. Within the base time unit of weeks, perceived stress was assumed to respond significantly faster than the other variables and was therefore treated as an auxiliary. According to the experts, sleep quality was caused by perceived stress and mood, while mood was caused by perceived stress, sleep, and inflammation. Inflammation, in turn, was caused by sleep quality and perceived stress, and perceived stress, in turn, by sleep quality and mood, as asserted by the experts. Figure S1 represents this example visually.

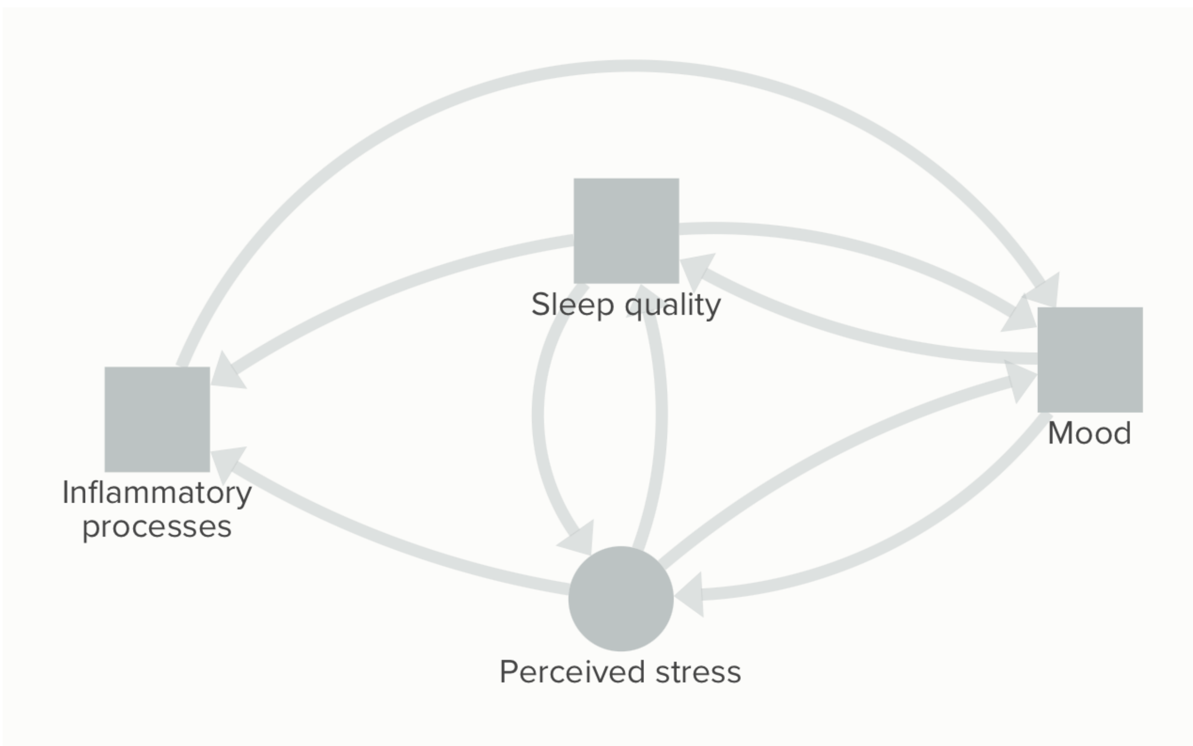

Figure S1 Visual representation of a small system dynamics model consisting of three stocks (squares) and one auxiliary/flow (circle)

Assuming linearity, we can then formulate the following system of differential and algebraic equations:

$$\frac{dS}{dt} = a_{s,p}P + a_{s,m}M + b_s$$

$$\frac{dM}{dt} = a_{m,p}P + a_{m,s}S + a_{m,i}I + b_m$$

$$\frac{dI}{dt} = a_{i,s}S + a_{i,p}P + b_i$$

$$P = b_p + a_{p,s}S + a_{p,m}M$$

where the a parameters are the effects of the variables on each other (e.g., $a_{s,p}$ is the effect of perceived stress, p, on sleep quality, s), and the b parameters correspond to the intercepts. Now we want to obtain a general solution that scales to higher-dimensional problems. Because we are interested in the relative contribution of each variable (including P), we want to find a solution that contains all model parameters (including b_p, $a_{p,s}$, and $a_{p,m}$). First, we rewrite the system without perceived stress:

$$\frac{dS}{dt} = b_s + a_{s,p}\left(b_p + a_{p,s}S + a_{p,m}M\right) + a_{s,m}M$$

$$\frac{dM}{dt} = b_m + a_{m,p}\left(b_p + a_{p,s}S + a_{p,m}M\right) + a_{m,s}S + a_{m,i}I$$

$$\frac{dI}{dt} = b_i + a_{i,s}S + a_{i,p}\left(b_p + a_{p,s}S + a_{p,m}M\right)$$

In matrix form from equation 1, we then get:

$$\frac{d\mathbf{x}}{dt} = \begin{pmatrix} a_{s,p}a_{p,s} & a_{s,p}a_{p,m} + a_{s,m} & 0 \\ a_{m,p}a_{p,s} + a_{m,s} & a_{m,p}a_{p,m} & a_{m,i} \\ a_{i,s}+a_{i,p}a_{p,s} & a_{i,p}a_{p,m} & 0 \end{pmatrix} \cdot \mathbf{x} + \begin{pmatrix} b_s + a_{s,p}b_p \\ b_m + a_{m,p}b_p \\ b_i + a_{i,p}b_p \end{pmatrix}$$

where $\mathbf{x}$ contains S, M, and I, and matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$ contain the coefficients from not just the stocks but also the intercepts.

Additional file 2

Tutorial: Using the D2D approach

There are two ways to use the D2D approach. It is possible to install our *systemdynamics* package in Python 3 (e.g., using PIP), then follow the approach illustrated in the example we have linked here: https://doi.org/10.17605/OSF.IO/4UTGW.

However, for users unfamiliar with Python, it is also possible to use our Streamlit app (also available at https://doi.org/10.17605/OSF.IO/4UTGW), which runs the code under the hood. Note that users may encounter a message saying, 'this app has gone to sleep,' after which they can click on the button saying 'Yes, get this app back up!' Whichever method is used, the D2D approach requires the user to define the CLD using an Excel file that can be exported from or imported into Kumu (kumu.io). The file used for our case example is attached as supplementary material.

Kumu Excel file

The user can export their CLD from Kumu as an Excel file and then add additional information in the format suggested below. This Excel file, generated using Kumu, contains two sheets: 'Elements' and 'Connections'. For our purposes, two additional specifications are needed. Firstly, the Elements sheet should look like Figure S2, where 'Label' contains the variable names; 'Type' contains whether the each variable is an auxiliary, stock, or constant; 'Tags' contains an annotation that determines whether and how strongly a variable will be intervened on (e.g., 1/-1 indicating the direction and strength of intervention) or not (0); and 'Description' asserts which variables are the variable(s) of interest using 'VOI'. The user can replace 1 and -1 with any numerical value (e.g., 0.5), which will then reflect the expected change achievable in the intervention variable.

Label	Type	Tags	Description
Cognitive functioning	stock	0	VOI
Brain atrophy	stock	0	
Neuronal dysfunction	stock	0	
Cerebral endothelial dysfunction	stock	0	
Amyloid beta burden	stock	0	
Daily functioning	stock	0	
Morbidity burden	stock	0	
Depressive symptoms	stock	-1	
Obesity	stock	-1	
Blood pressure	stock	-1	
Tau burden	stock	0	
Engagement in cognitively demanding tasks	auxiliary	1	
Healthy dietary patterns	auxiliary	1	
Physical activity	auxiliary	1	
Sleep quality	auxiliary	1	
Diabetes	constant	-1	

Figure S2 *Example of Elements sheet in Kumu format*

The Connections sheet then looks as follows, where the only required columns are "From," containing the cause variable, "To," containing the effect variable, and "Type," containing the polarity (Figure S3). For instance, for a CLD consisting of only one (balancing) feedback loop with connections A →+ B and B →- A, one row should be added as From: A, To: B, Type: +, and another should be added as From: B, To: A, Type: -. The Kumu format also includes a 'Direction' column, which should indicate 'directed' for each connection, but D2D does not use this column, as it assumes a fully directed CLD. Additionally, D2D can use two columns, LB and UB, which denote the lower and upper bounds of the uniform distribution for a specific connection. These settings will override the connection's polarity as specified under 'Type' and the θ_{max} settings. In this example, the connection from *Sleep quality* on *Amyloid beta burden* would be sampled from a Uniform[0.1, 0.5], whereas the other connections' sampling ranges are determined by their polarities and θ_{max} as detailed in the manuscript.

From	To	Direction	Label	Type	LB	UB
Sleep quality	Amyloid beta burden	directed		-	0.1	0.5
Diabetes	Amyloid beta burden	directed		+		
Cerebral endothelial dysfunctic	Amyloid beta burden	directed		+		
Neuroinflammation	Amyloid beta burden	directed		+		
Head trauma	Amyloid beta burden	directed		+		
ApoE4 carriership	Amyloid beta burden	directed		+		
Social relationships	Blood pressure	directed		-		
Experienced stress	Blood pressure	directed		+		
Physical activity	Blood pressure	directed		-		
Healthy dietary patterns	Blood pressure	directed		-		
Obesity	Blood pressure	directed		+		
Neuronal dysfunction	Brain atrophy	directed		+		
Physical activity	Brain perfusion	directed		+		
Cerebral endothelial dysfunctic	Brain perfusion	directed		-		
Smoking	Cerebral endothelial dysfunctio	directed		+		
Healthy dietary patterns	Cerebral endothelial dysfunctio	directed		-		
Systemic inflammation	Cerebral endothelial dysfunctio	directed		+		
Dyslipidaemia	Cerebral endothelial dysfunctio	directed		+		
Diabetes	Cerebral endothelial dysfunctio	directed		+		
Blood pressure	Cerebral endothelial dysfunctio	directed		+		
Brain perfusion	Cerebral endothelial dysfunctio	directed		-		
Head trauma	Cerebral endothelial dysfunctio	directed		+		
Amyloid beta burden	Cerebral endothelial dysfunctio	directed		+		
ApoE4 carriership	Cerebral endothelial dysfunctio	directed		+		
Oxidative stress	Cerebral endothelial dysfunctio	directed		+		
Depressive symptoms	Cognitive functioning	directed		-		
Sleep quality	Cognitive functioning	directed		+		
Hearing loss	Cognitive functioning	directed		-		
Brain perfusion	Cognitive functioning	directed		+		
Neuronal connectivity	Cognitive functioning	directed		+		
Neuronal dysfunction	Cognitive functioning	directed		-		
Brain atrophy	Cognitive functioning	directed		-		
Neuronal dysfunction	Daily functioning	directed		-		
Cognitive functioning	Daily functioning	directed		+		
Motor function	Daily functioning	directed		+		
Morbidity burden	Daily functioning	directed		-		
Social relationships	Depressive symptoms	directed		-		
Experienced stress	Depressive symptoms	directed		+		
Cognitive functioning	Depressive symptoms	directed		-		
Sleep quality	Depressive symptoms	directed		-		
Physical activity	Depressive symptoms	directed		-		
Healthy dietary patterns	Diabetes	directed		-		

Figure S3 *Example of Connections sheet in Kumu format*

Finally, if the user wishes to incorporate interaction terms, they can include a third sheet called "Interaction terms" (Figure S4). This third sheet works the same as the Connections sheet, but now there are two 'From' columns: "From1" and "From2". To incorporate a quadratic term, one should add the same variable under both the "From1" and "From2" columns. For instance, in the Alzheimer's example, an interaction term could be added where *Sleep quality* and *Depressive symptoms* modify each other's effects on *Cognitive functioning*: higher sleep quality reduces the impact of depressive symptoms, while depressive symptoms diminish the beneficial effect of sleep (Figure S4).

Since higher-order interaction effects are generally smaller than main effects (35), D2D samples from a smaller range for interaction terms, namely [0, $\theta_{max}/2$] for positive polarity, [-$\theta_{max}/2$, 0] for negative polarity, and [-$\theta_{max}/2$, $\theta_{max}/2$] where polarity is not specified.

From1	From2	To	Type
Sleep quality	Depressive symptoms	Cognitive functioning	-

Figure S4 *Example of Interaction terms sheet in Kumu format*

Settings

The D2D method also requires the user to specify several settings. For instance, the *base time unit* ("Quarter-years" in our example on Alzheimer's disease), the *timeframe* (20 in our example), and the maximum parameter value, θ_{max}, for both the stocks and auxiliaries. These settings may vary widely across modeling projects. For instance, if θ_{max} is set too high, the model may reach impossible values (e.g., 100 standard deviations or beyond).

Several default settings can be adjusted if the user wishes, including the number of samples to draw from the uncertain model parameters (N: 100 by default). The seed can be set to any value and determines the uniqueness of the parameter samples that are drawn. Its function is to ensure reproducibility: running a simulation twice with the same seed will produce the same outcome, even when using a random sampler. Since the model uses boxplots and works directly with the samples rather than standard errors, the uncertainty does not scale with the inverse square root of N, making the uncertainty in D2D independent of N. However, for low values of N, the result may look very different with a different seed. We recommend setting N so that changing the seed does not impact the conclusions.