A Next-Generation Modular and Compositional Framework for System Dynamics Modeling and Beyond

Xiaoyan Li[1,4,5]

Supervisor: Nathaniel Osgood[4]

Joint work with J. Baez[1,2,3], E. Patterson[1], S. Libkind[1],
A. Alegre[4], T. Purdy[4], E. Redekopp[4] & N. Meadows[4]

[1]Topos Institute [2]University of Edinburgh [3]University of California, Riverside
[4]University of Saskatchewan [5]University of Lethbridge

2025 February 27th

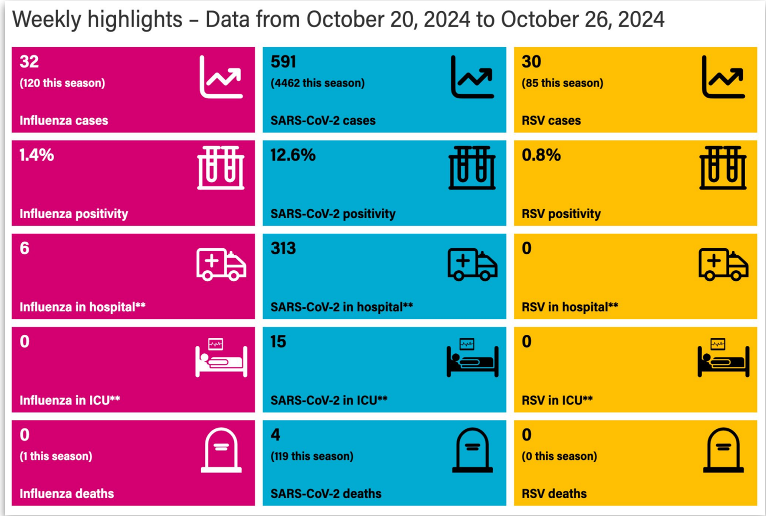
Outline

- **Application**: A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

- **Theory (ACT)**: A Modular and Compositional Framework for System Dynamics Modeling

1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

Traditional Surveillance data[1]:



Characteristics of Big Data:

Large Volume

High Velocity

High Variety

Big data[2]:

1. cite from publicly accessed data from "Alberta Respiratory virus dashboard": https://www.alberta.ca/stats/dashboard/respiratory-virus-dashboard.htm?data=highlights#highlights
2. cite from "What 'Data Never Sleeps 9.0' proves about the pandemic": https://www.domo.com/blog/what-data-never-sleeps-9-0-proves-about-the-pandemic/

1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

Methodology:

- Using **Dynamic Models** simulate the system
- Using Machine Learning Algorithms (**Bayesian models**) and (large scale) **empirical datasets** to train the machine learning dynamic models
- **Project** forward based on the trained models
- Perform **counterfactual analysis** to help **decision making**.

1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

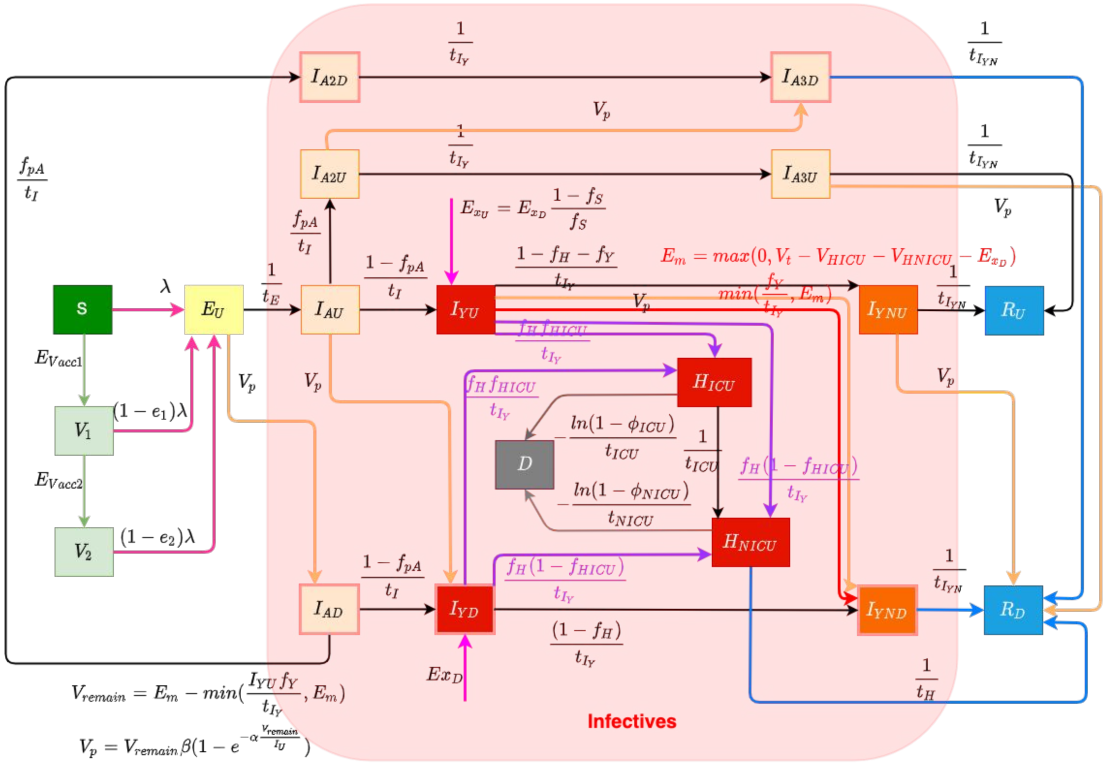
Bayesian Machine Learning & Dynamic Models

- **MCMC:** Sample from $p_M(\theta|y_{1:T})$: posteriors of ***deterministic*** dynamic model static parameters, scenario results, and incremental scenario gains.

- **Particle Filtering/SMC:** Sample from $p_{\theta,M}(x_{1:T}|y_{1:T})$: posteriors of ***stochastic*** dynamic model latent states stochastically evolving parameters, scenario results, and incremental scenario gains.

- **Particle MCMC (PMCMC):** Sample from $p_M(\theta,x_{1:T}|y_{1:T})$: posteriors of ***stochastic*** dynamic model latent states, stochastically evolving parameters, scenario results, and incremental scenario gains *and static parameters*.

1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention



An example of a COVID-19 dynamic model applied daily for 17 jurisdictions in Canada:

- Saskatchewan
- All other Canadian provinces (for PHAC)
- Weekly for First Nations Reserves (for FNIHB)

The mathematical structure of the COVID-19 dynamic model employed in particle filtering

1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

Associated ODEs

Stocks:

$$\frac{d\mathbf{S}}{dt} = -\lambda S$$

$$\frac{d\mathbf{E_U}}{dt} = \lambda S - \frac{E_U}{t_E} - V_p \frac{E_U}{I_U}$$

$$\frac{d\mathbf{E_D}}{dt} = V_p \frac{E_U}{I_U} - \frac{E_D}{t_E}$$

$$\frac{d\mathbf{I_{AU}}}{dt} = \frac{E_U}{t_E} - \frac{I_{AU}}{t_I} - V_p \frac{I_{AU}}{I_U}$$

$$\frac{\mathbf{I_{AD}}}{dt} = \frac{E_D}{t_E} + V_p \frac{I_{AU}}{I_U} - \frac{I_{AD}}{t_I}$$

$$\frac{d\mathbf{I_{A2U}}}{dt} = f_{pA} \frac{I_{AU}}{t_I} - \frac{I_{A2U}}{t_{I_Y}} - V_p \frac{I_{A2U}}{I_U}$$

$$\frac{d\mathbf{I_{A2D}}}{dt} = f_{pA} \frac{I_{AD}}{t_I} + V_p \frac{I_{A2U}}{I_U} - \frac{I_{A2D}}{t_{I_Y}}$$

$$\frac{d\mathbf{I_{A3U}}}{dt} = \frac{I_{A2U}}{t_{I_Y}} - V_p \frac{I_{A3U}}{I_U} - \frac{I_{A3U}}{t_{I_{YN}}}$$

$$\frac{d\mathbf{I_{A3D}}}{dt} = \frac{I_{A2D}}{t_{I_Y}} + V_p \frac{I_{A3U}}{I_U} - \frac{I_{A3D}}{t_{I_{YN}}}$$

$$\frac{d\mathbf{I_{YU}}}{dt} = Ex_D \frac{1-f_S}{f_S} + (1.0 - f_{pA}) \frac{I_{AU}}{t_I} - \frac{I_{YU}}{t_{I_Y}} - V_p \frac{I_{YU}}{I_U} - min(\frac{I_{YU} f_{NH}}{t_{I_Y}}, E_m)$$

$$\frac{d\mathbf{I_{YD}}}{dt} = Ex_D + (1 - f_{pA}) \frac{I_{AD}}{t_I} + V_p \frac{I_{YU}}{I_U} - \frac{I_{YD}}{t_{I_Y}}$$

$$\frac{d\mathbf{H_{ICU}}}{dt} = I_{YU} \frac{f_H f_{HICU}}{t_{I_Y}} + I_{YD} \frac{f_H f_{HICU}}{t_{I_Y}} - \frac{H_{ICU}}{t_{ICU}} - (H_{ICU} \frac{-ln(1-\phi_{ICU})}{t_{ICU}})$$

$$\frac{d\mathbf{H_{NICU}}}{dt} = I_{YU} \frac{f_H(1-f_{HICU})}{t_{I_Y}} + I_{YD} \frac{f_H(1-f_{HICU})}{t_{I_Y}} + \frac{H_{ICU}}{t_{ICU}} - (H_{NICU} \frac{-ln(1-\phi_{NICU})}{t_{NICU}}) - \frac{H_{NICU}}{t_H}$$

$$\frac{d\mathbf{I_{YNU}}}{dt} = I_{YU} \frac{1-f_H}{t_{I_Y}} - \frac{I_{YNU}}{t_{I_{YN}}} - V_p \frac{I_{YNU}}{I_U}$$

$$\frac{d\mathbf{I_{YND}}}{dt} = f_{I_r} \frac{I_{YNU}}{t_{I_{YN}}} + V_p \frac{I_{YNU}}{I_U} + \frac{I_{YD}}{t_{I_Y}} + min(\frac{I_{YU} f_{NH}}{t_{I_Y}}, E_m) - \frac{I_{YND}}{t_{I_{YN}}}$$

$$\frac{d\mathbf{R_U}}{dt} = (1 - f_{I_r}) \frac{I_{YNU}}{t_{I_{YN}}} + \frac{I_{A3U}}{t_{I_{YN}}}$$

$$\frac{d\mathbf{R_D}}{dt} = \frac{I_{YND}}{t_{I_{YN}}} + \frac{H_{ICU}}{t_H} + \frac{I_{A3D}}{t_{I_{YN}}}$$

$$\frac{d\mathbf{D}}{dt} = (H_{NICU} \frac{-ln(1-\phi_{NICU})}{t_{NICU}}) + (H_{ICU} \frac{-ln(1-\phi_{ICU})}{t_{ICU}})$$

1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

Dynamic Parameters:

$$I_U = E_U + I_{AU} + I_{A2U} + I_{A3U} + I_{YU} + I_{YNU}$$

$$t_{I_{YN}} = t_R - t_{I_Y}$$

$$N = S + E_U + E_D + I_{AU} + I_{AD} + I_{A2U} + I_{A2D} + I_{A3U} + I_{A3D} + I_{YU} + I_{YD} + I_{YNU} + I_{YND} + R_U + R_D$$
$$+ H_{ICU} + H_{NICU} + D$$

$$\lambda = c\beta \frac{(I_{AU} + I_{A2U} + I_{A3U}) + \rho_U(I_{YU} + I_{YNU}) + \rho_D(I_{AD} + I_{A2D} + I_{A3D} + I_{YD} + I_{YND})}{(S + E_U + I_{AU} + I_{A2U} + I_{A3U}) + \rho_U(I_{YU} + I_{YNU}) + \rho_D(E_D + I_{AD} + I_{A2D} + I_{A3D} + I_{YD} + I_{YND})}$$

$$E_m = max(0, V_t - V_{HICU} - V_{HNICU} - E_{x_D})$$

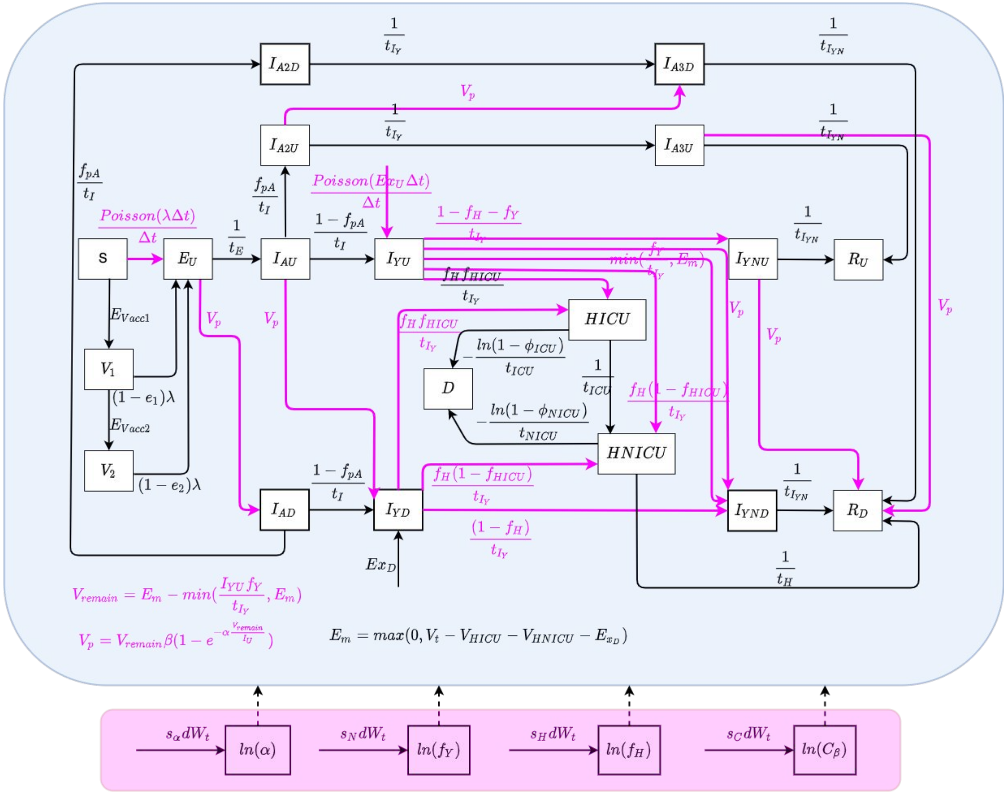
$$V_{remain} = E_m - min(\frac{f_{NH}}{t_{I_Y}}, E_m)$$

$$V_p = V_{remain}\beta(1 - e^{-\alpha\frac{V_{remain}}{I_U}})$$

$$R_0 = C\beta(t_I + t_{I_Y} + t_{I_{YN}})$$

$$R^*(t) = \frac{SR_0}{(S + E_U + I_{AU} + I_{A2U} + I_{A3U} + R_U) + \rho_U(I_{YU} + I_{YNU}) + \rho_D(E_D + I_{AD} + I_{A2D} + I_{A3D} + I_{YD} + I_{YND} + R_D)}$$

1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention



State Space: $[S, E_U, E_D, I_{AU}, I_{AD}, I_{A2U}, I_{A2D}, I_{A3U}, I_{A3D}, I_{AU}, I_{AD}, I_{YU}, I_{YD}, I_{YNU}, I_{YND}, H_{ICU}, H_{NICU}, R_U, R_D, D, C_\beta, \alpha, f_{NH}, f_H]^T$

1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

1. Daily count of new reported incident confirmed or suspected cases

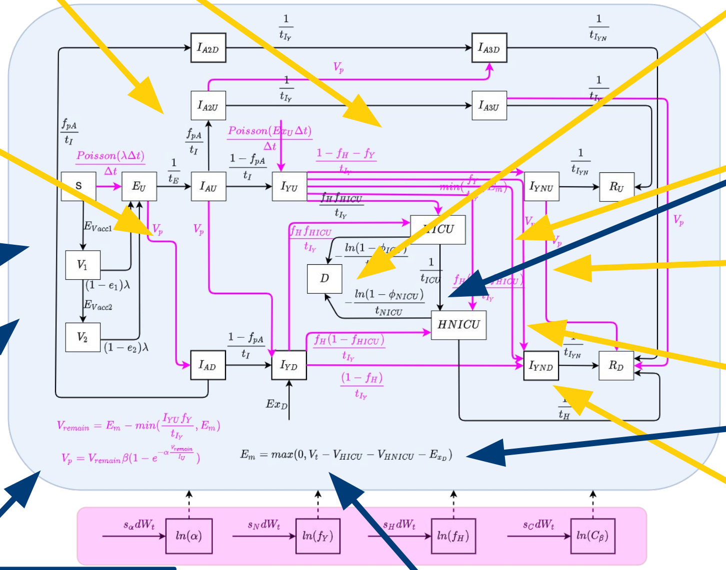
2. Cumulative reported incident confirmed or suspected cases from the inception of the pandemic

3. Weekly average virus SARS-CoV-2 concentration in wastewater

4. Daily count of persons who received their first vaccination dose

5. Daily count of persons who received their second vaccination dose

6. Daily count of persons undergoing PCR (nasopharyngeal swab)-based testing



7. Cumulative reported deaths from COVID-19

8. Daily COVID-19 patients admitted into hospital for non-ICU care

9. Daily midnight census (count) of COVID-19 patients in the ICU

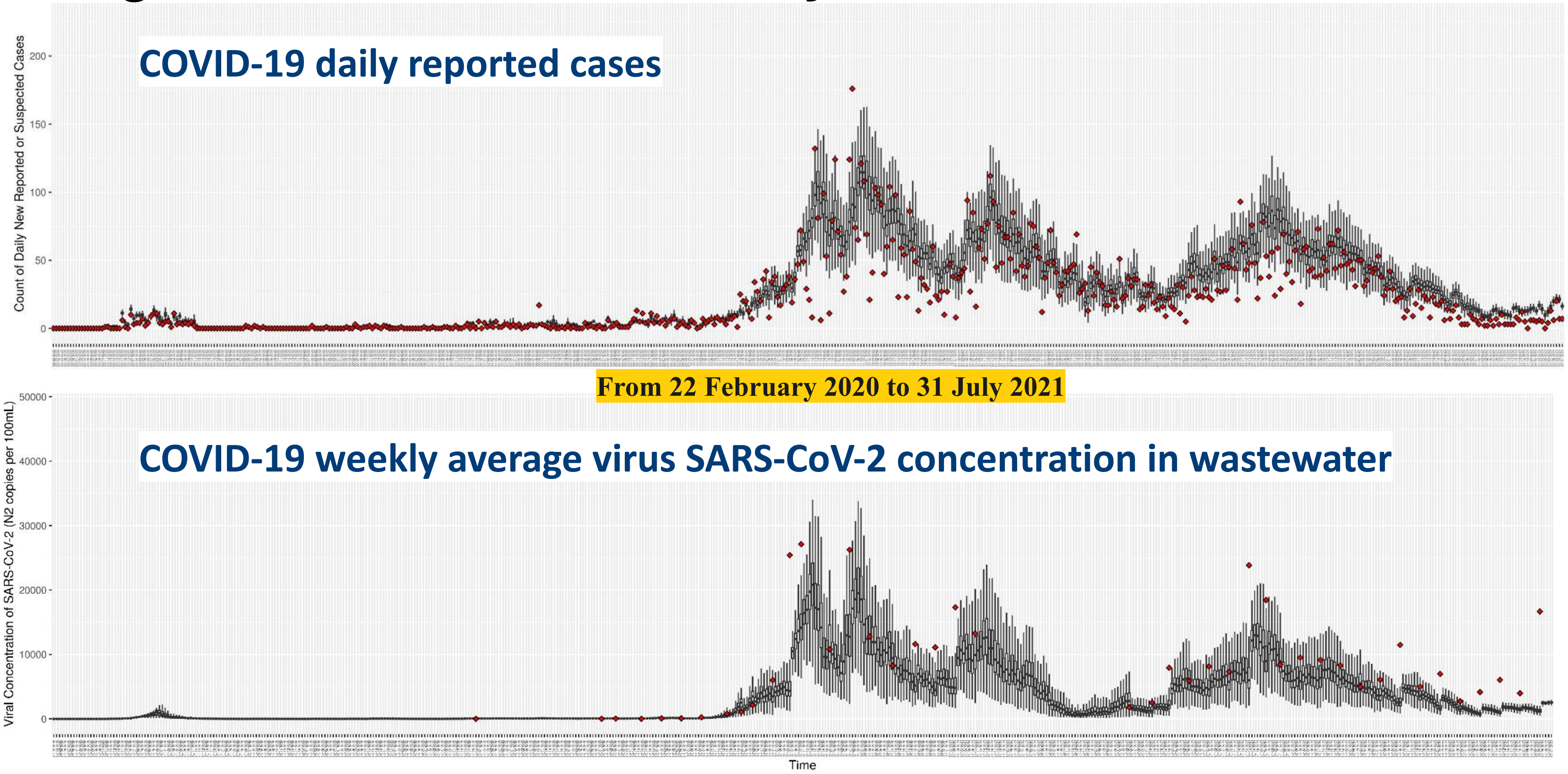
10. Daily count COVID-19 patients admitted into hospital for ICU care

11. Daily midnight census (count) of COVID-19 patients in the hospital for non-ICU care

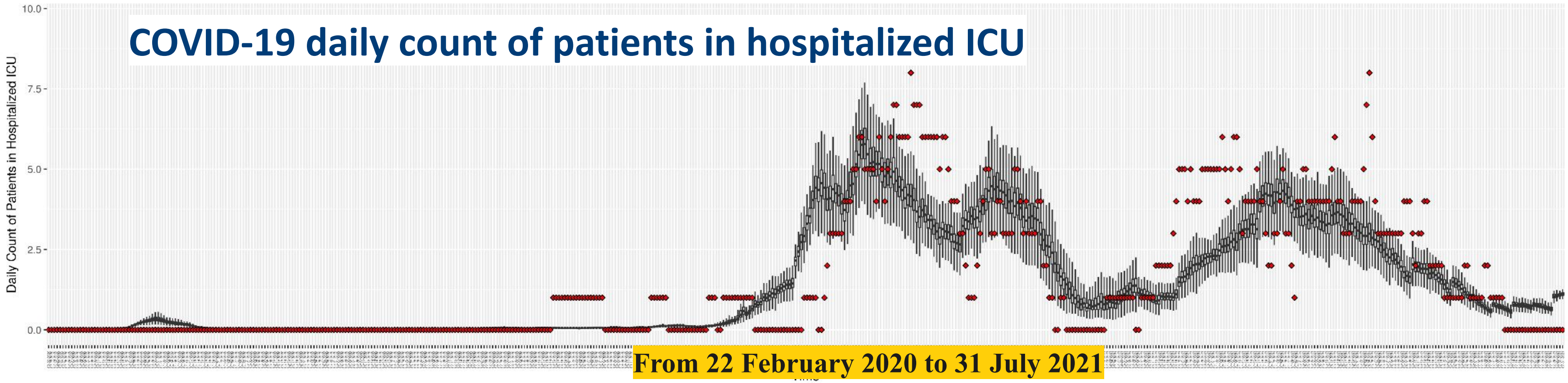
12. Daily new likely exogenous cases

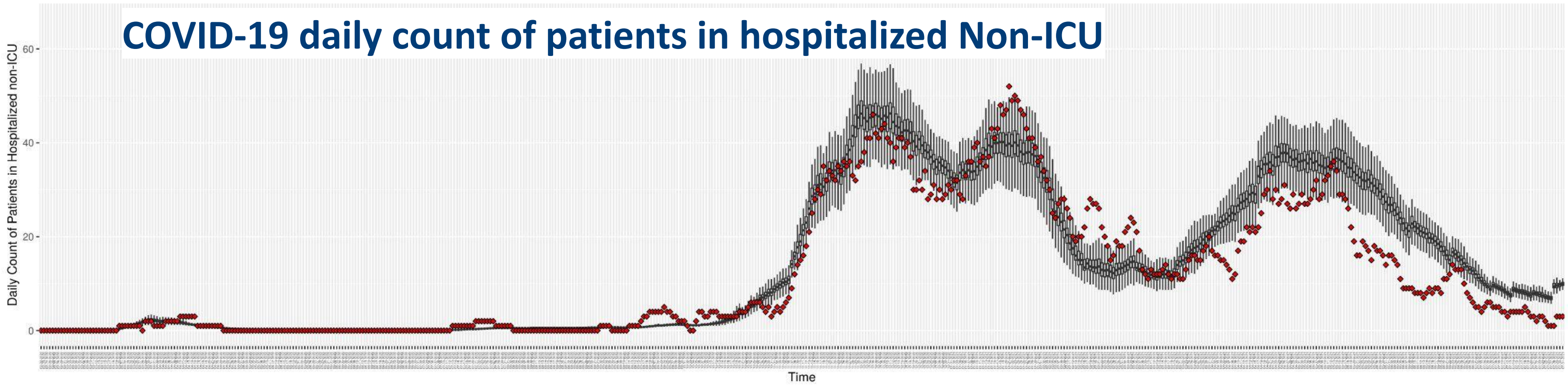
12 empirical datasets: used to **drive the model**, and used in the PF likelihoods.

1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention



1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

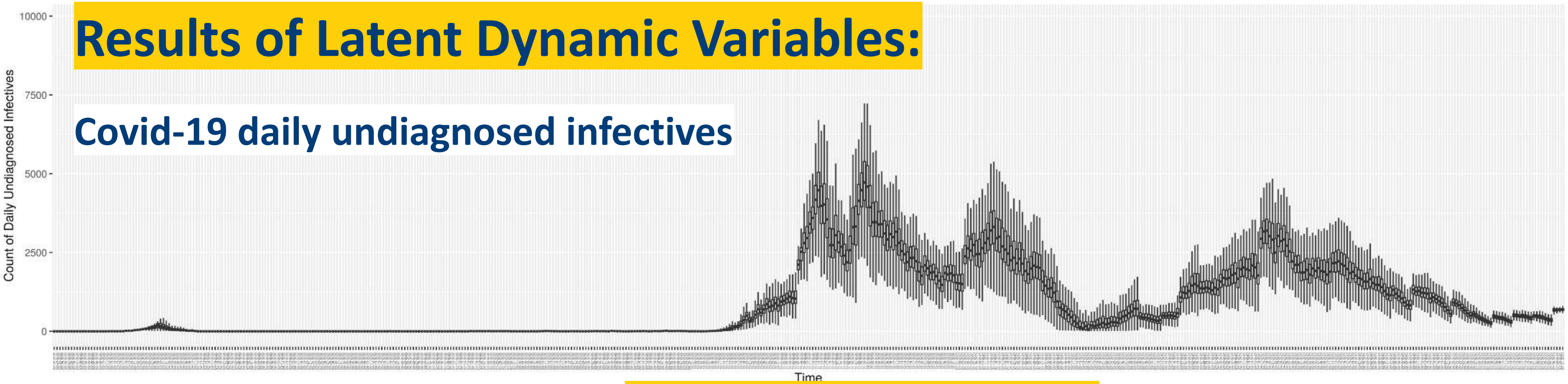




1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

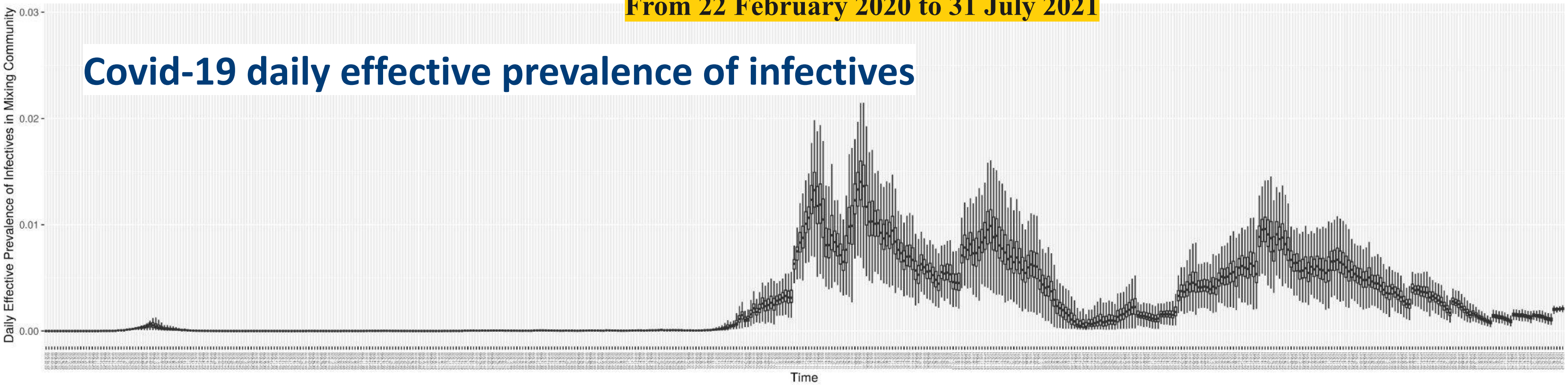
Results of Latent Dynamic Variables:

Covid-19 daily undiagnosed infectives

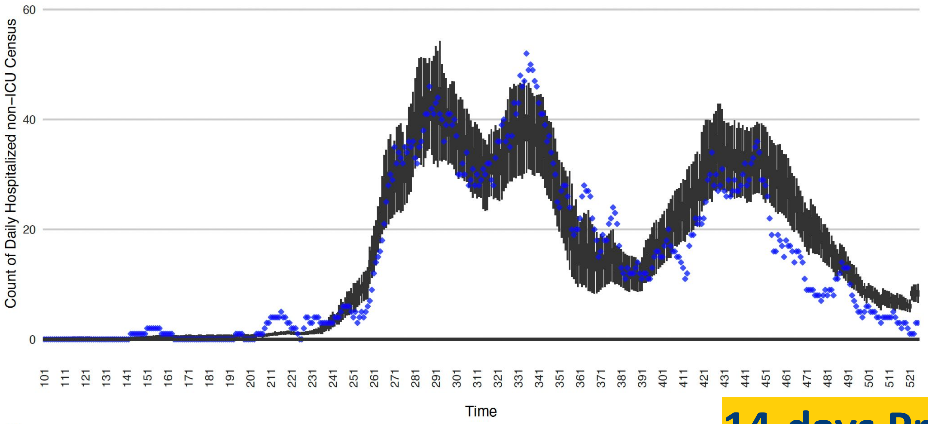


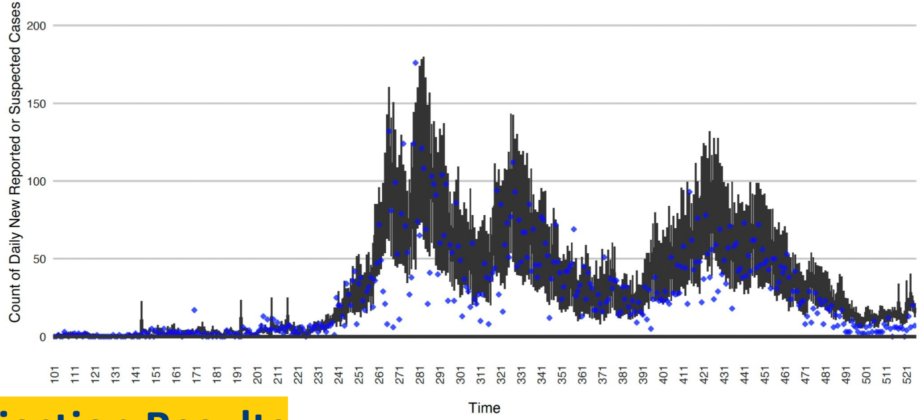
From 22 February 2020 to 31 July 2021

Covid-19 daily effective prevalence of infectives

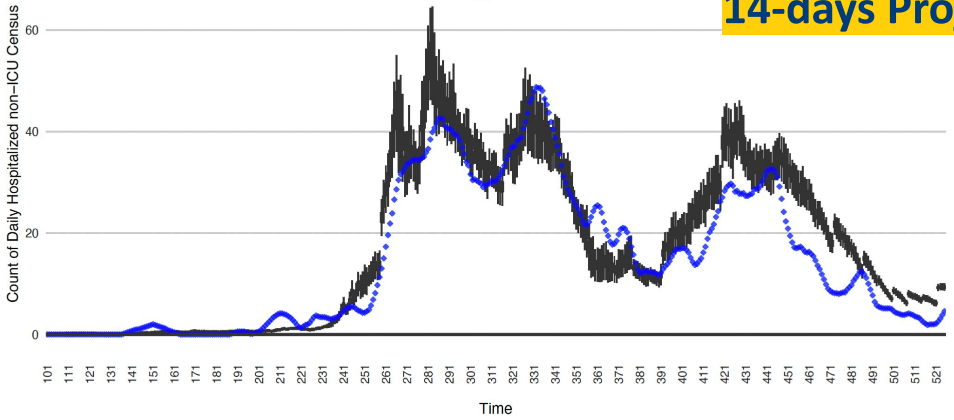


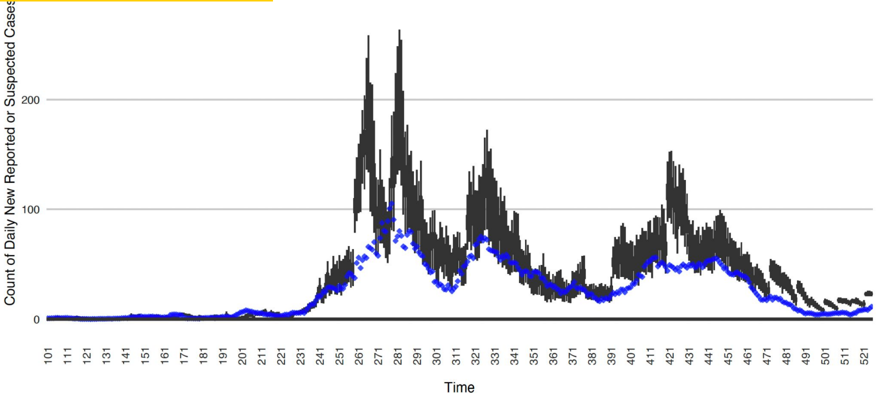
1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention



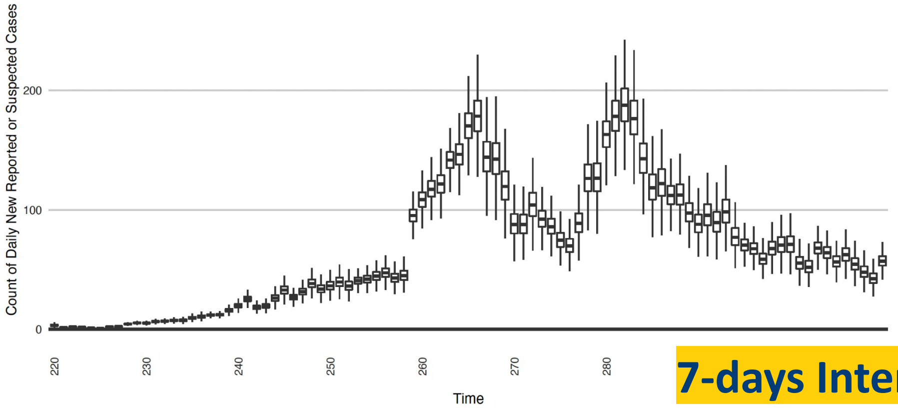


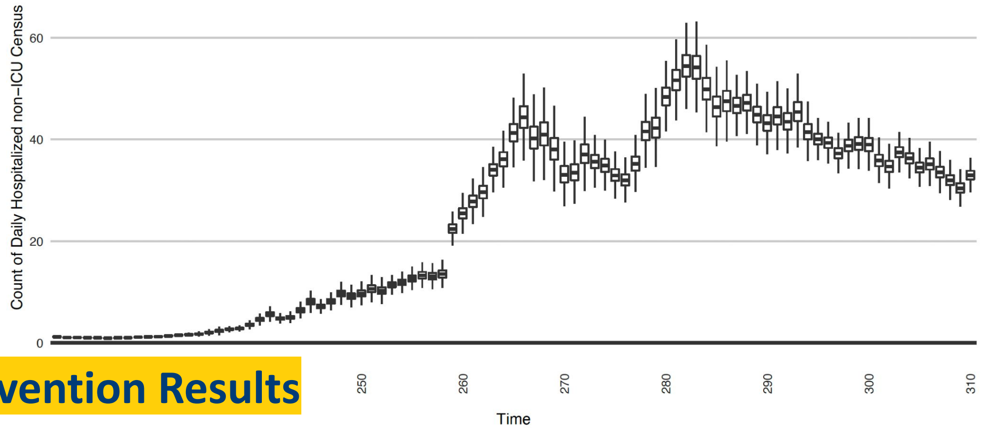
14-days Projection Results





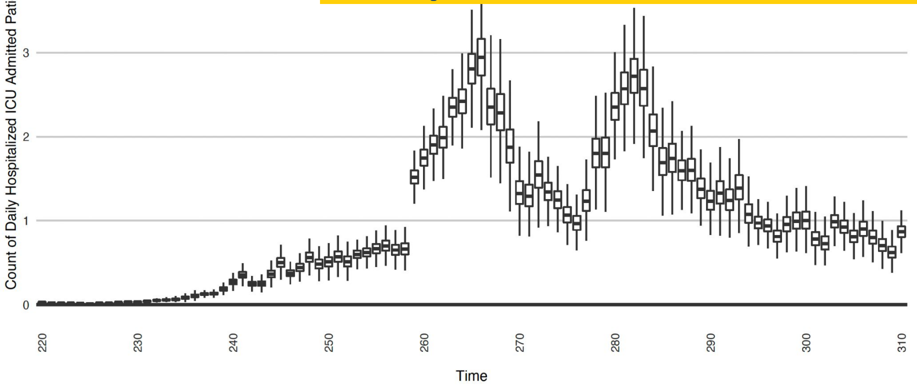
1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

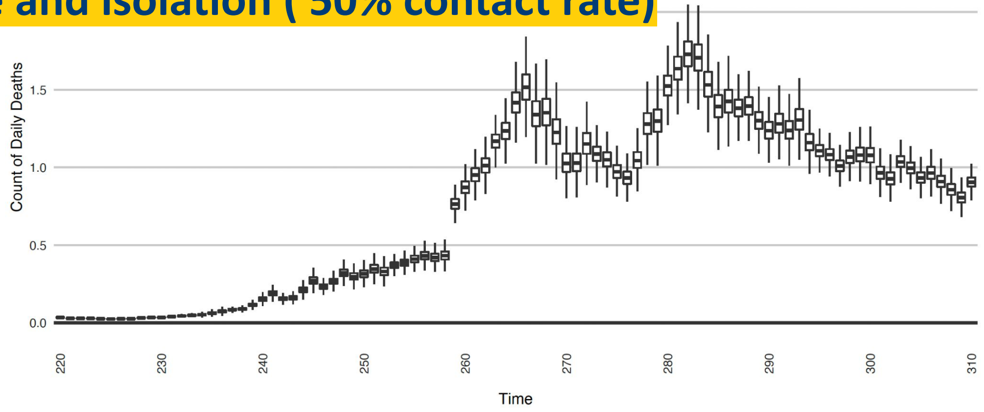




7-days Intervention Results

Example Scenario: Quarantine and isolation (50% contact rate)





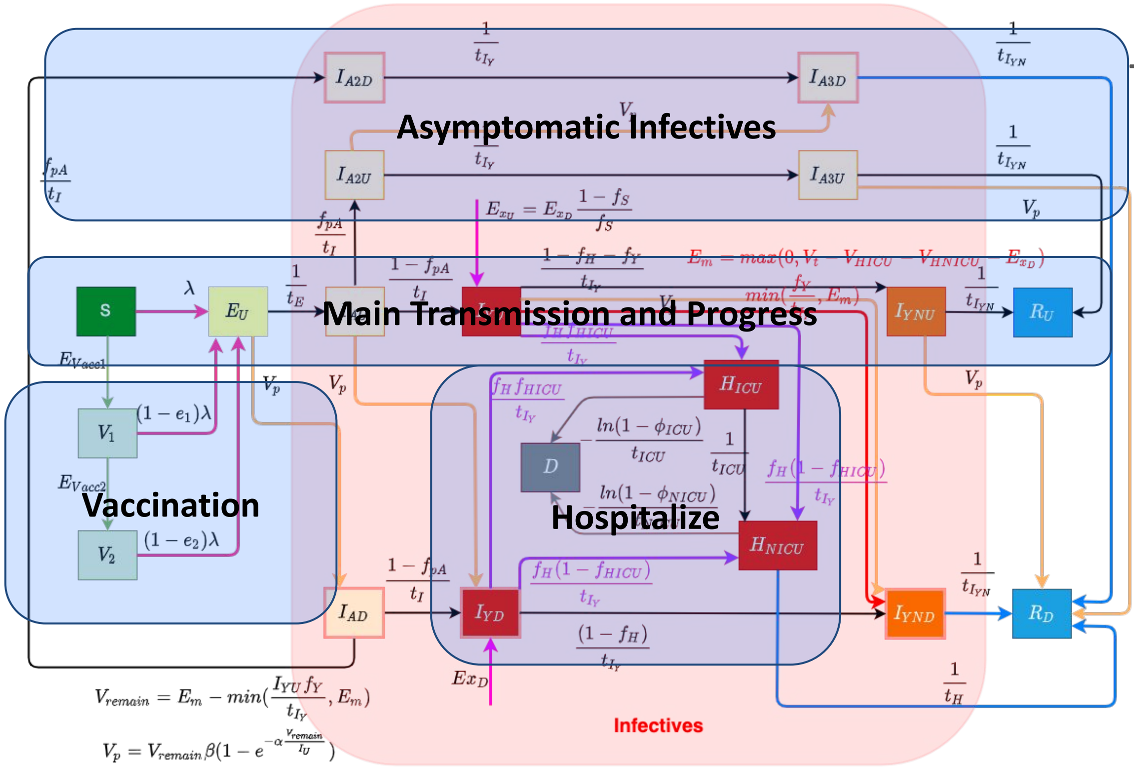
2. A Modular and Compositional Framework for System Dynamics Modeling

Motivations:

Limitations of current modeling tools:

- Lack of Explicit Representation of Mathematical Structures
- Insufficient Representation of Mathematical Relationships
- No Clear Separation Between Diagrammatic Syntax and Semantics
- Inability to Represent Models Modularly
- Limited Support for Submodel Composition (Reusability)

2. A Modular and Compositional Framework for System Dynamics Modeling

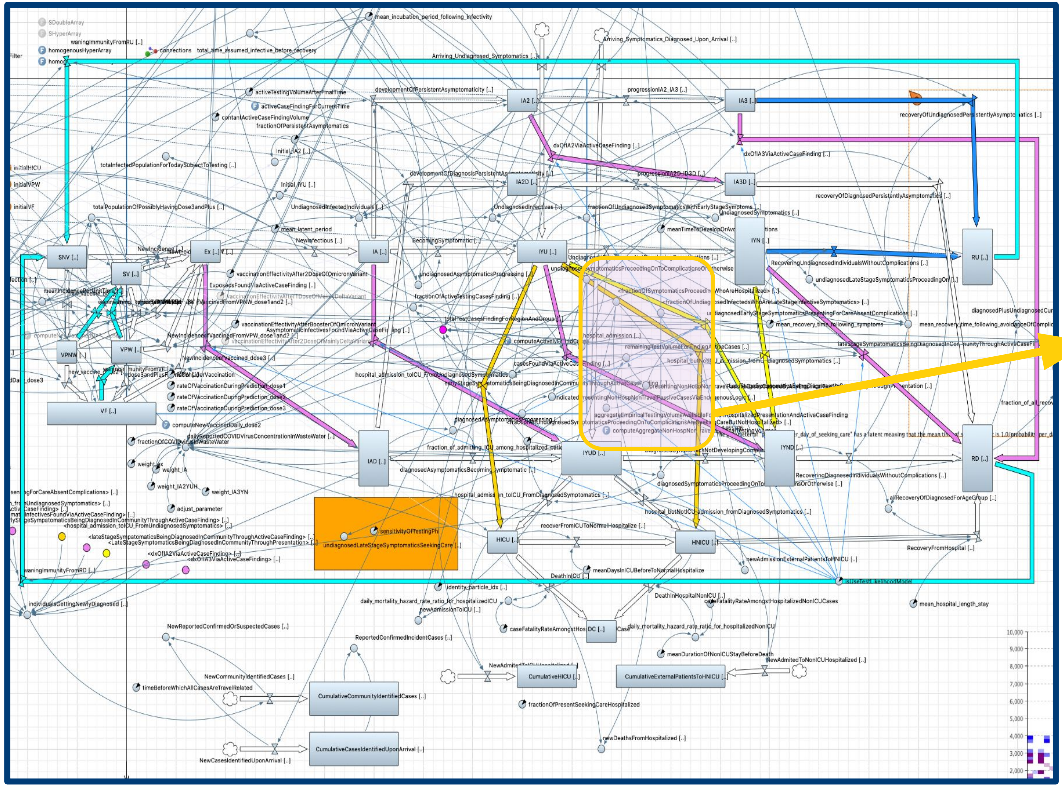


Three models with same structure:

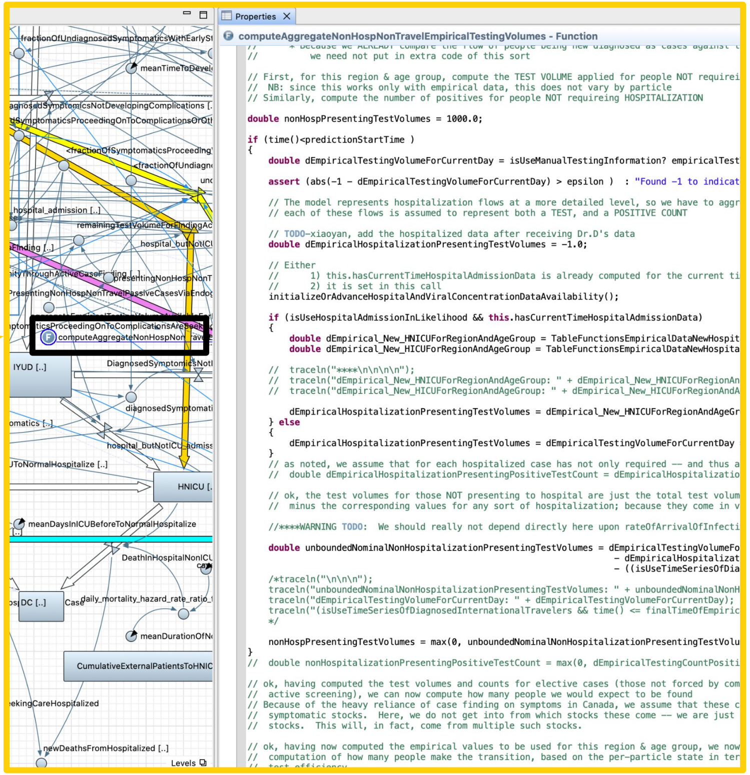
1. Aggregate population
2. Age Stratified
3. Age & Region Stratified

2. A Modular and Compositional Framework for System Dynamics Modeling

The PF COVID-19 Model Graphical Interface



An Example of A Function



2. A Modular and Compositional Framework for System Dynamics Modeling

Research Question:

Traditional Modeling Methods

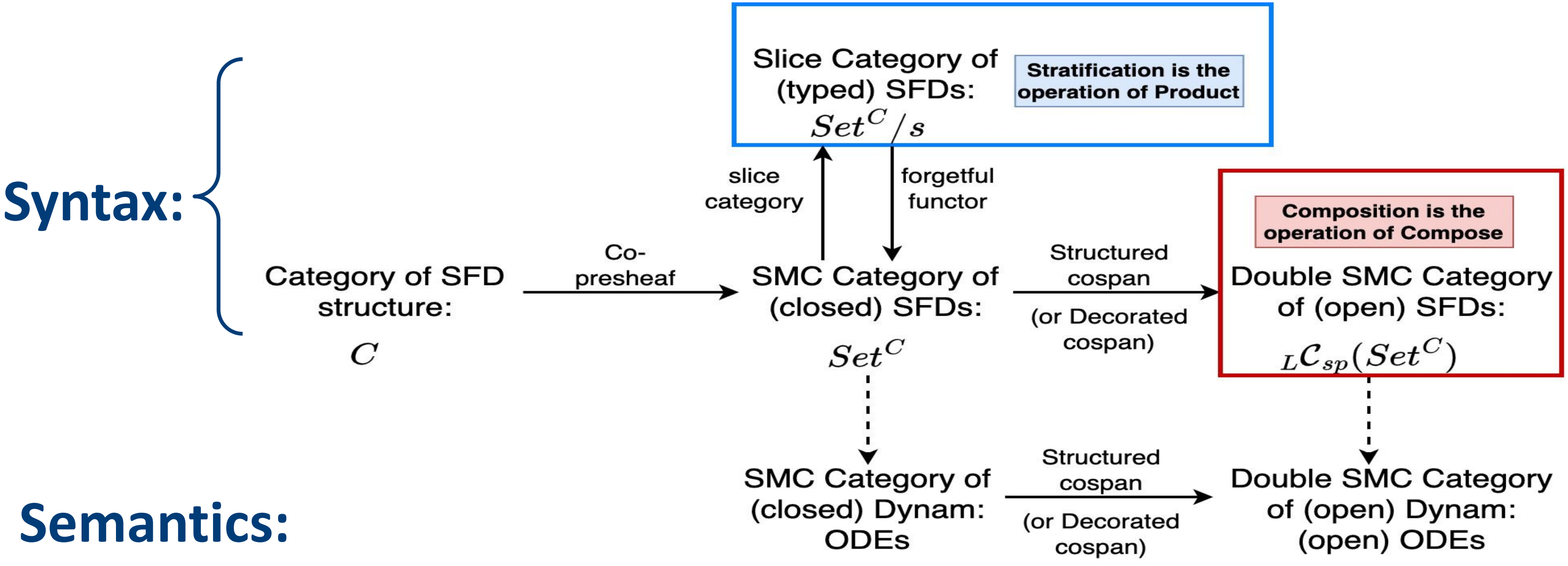
?

A Next-generation Modular Modeling Framework of Constructing Models by Composition

2. A Modular and Compositional Framework for System Dynamics Modeling

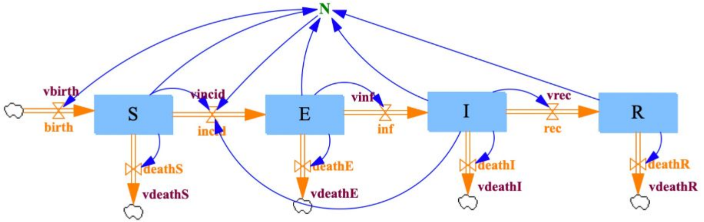
Methods: Applied Category Theory

Mathematical construction of **composition** and **stratification**:



2. A Modular and Compositional Framework for System Dynamics Modeling

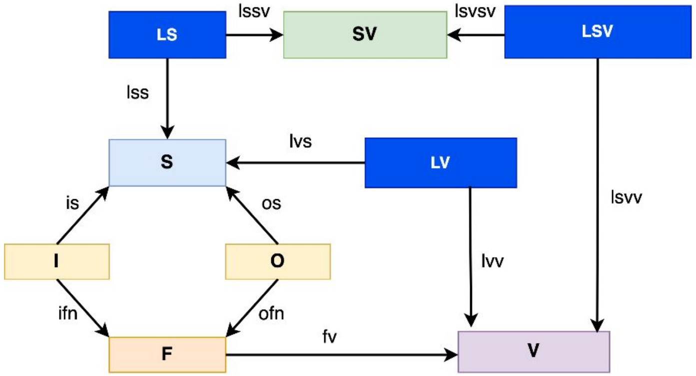
Category of Closed Stock and Flow Diagrams: ($\mathbf{FinSet}^{\mathbf{C}}$, $\boldsymbol{\varphi}$)



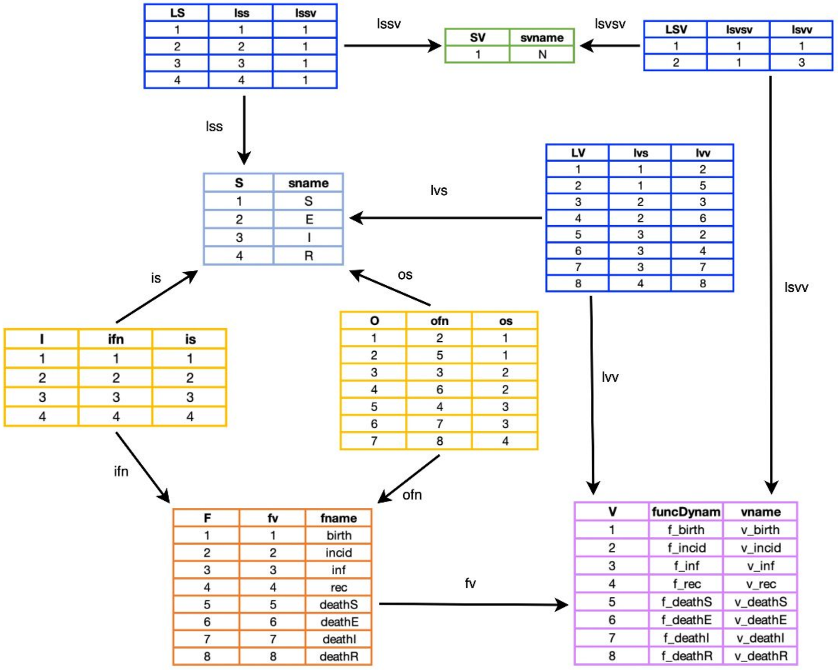
$f_{\text{birth}} = \mu * N$ $\quad f_{\text{incid}} = \beta * S * I/N$ $\quad f_{\text{inf}} = E/t_{\text{latent}}$ $\quad f_{\text{rec}} = I/t_{\text{recovery}}$

$f_{\text{deathS}} = \delta * S$ $\quad f_{\text{deathE}} = \delta * E$ $\quad f_{\text{deathI}} = \delta * I$ $\quad f_{\text{deathR}} = \delta * R$

(a)



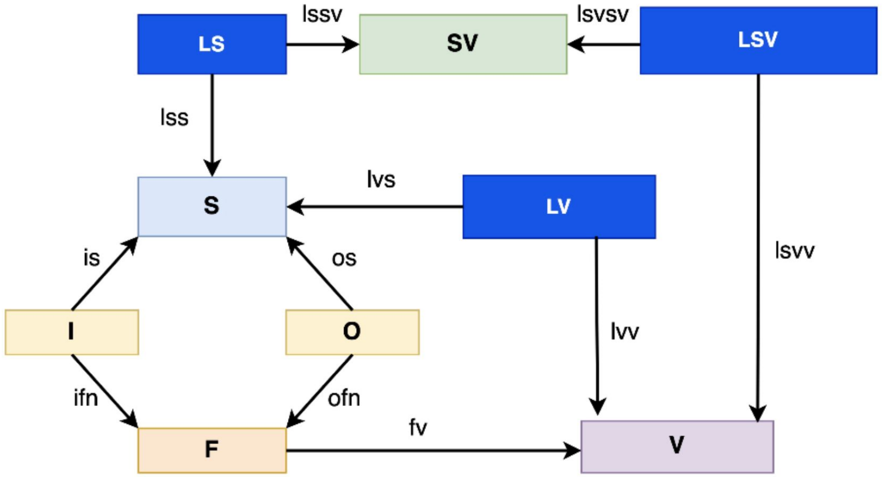
(b)



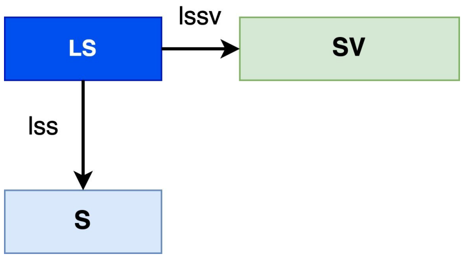
(c)

2. A Modular and Compositional Framework for System Dynamics Modeling

Constructing Category of Open Stock and Flow Diagrams (Structured/Decorated Cospan):



Schema of SFD (C)

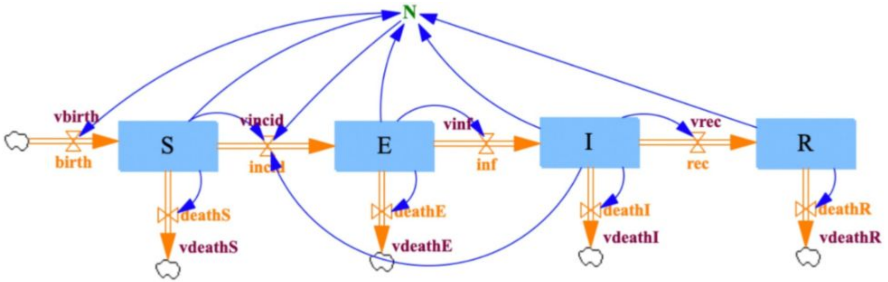


Schema of interface (H)

For functor L: $\mathrm{FinSet}^H \rightarrow \mathrm{FinSet}^C$, there is a double category ${}_L\mathrm{Csp}(\mathrm{FinSet}^C)$

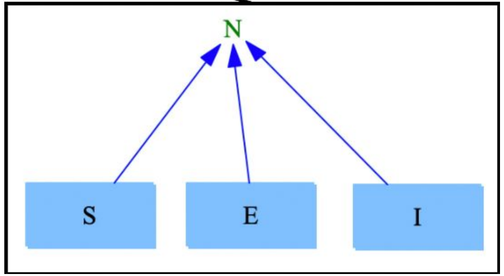
2. A Modular and Compositional Framework for System Dynamics Modeling

A horizontal 1-cell from object a to b (in FinSet^{H}) is a diagram in FinSet^{C}:



$f_{\text{birth}} = \mu * N$ $\quad f_{\text{incid}} = \beta * S * I/N$ $\quad f_{\text{inf}} = E/t_{\text{latent}}$ $\quad f_{\text{rec}} = I/t_{\text{recovery}}$

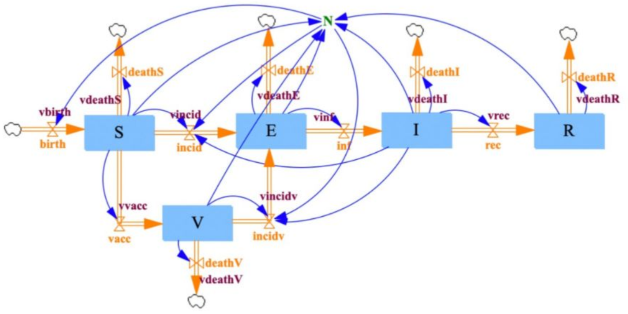
$f_{\text{deathS}} = \delta * S$ $\quad f_{\text{deathE}} = \delta * E$ $\quad f_{\text{deathI}} = \delta * I$ $\quad f_{\text{deathR}} = \delta * R$



$$L(a) \xrightarrow{i} x \xleftarrow{o} L(b)$$

2. A Modular and Compositional Framework for System Dynamics Modeling

$$\begin{tikzcd}
 & & x +_{L(b)} y & & \\
 & x \arrow[ur, "j_x"] & & y \arrow[ul, "j_y"'] & \\
L(a) \arrow[ur, "i_1"] & & L(b) \arrow[ul, "o_1"'] \arrow[ur, "i_2"] & & L(c) \arrow[ul, "o_2"']
\end{tikzcd}$$

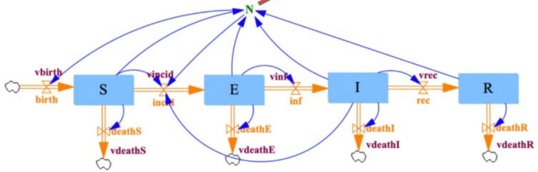


$f_{\text{birth}} = \mu * N$ $\quad f_{\text{incid}} = \beta * S * I/N$ $\quad f_{\text{inf}} = E/t_{\text{latent}}$ $\quad f_{\text{rec}} = I/t_{\text{recovery}}$ $\quad f_{\text{vacc}} = \alpha * S$

$f_{\text{deathS}} = \delta * S$ $\quad f_{\text{deathE}} = \delta * E$ $\quad f_{\text{deathI}} = \delta * I$ $\quad f_{\text{deathR}} = \delta * R$ $\quad f_{\text{deathV}} = \delta * V$

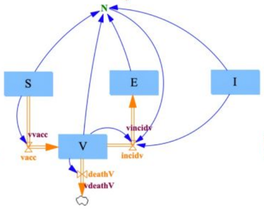
$f_{\text{incidv}} = \beta * (1 - e) * V * I/N$

Composition



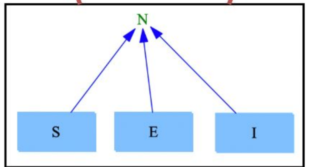
$f_{\text{birth}} = \mu * N$ $\quad f_{\text{incid}} = \beta * S * I/N$ $\quad f_{\text{inf}} = E/t_{\text{latent}}$ $\quad f_{\text{rec}} = I/t_{\text{recovery}}$

$f_{\text{deathS}} = \delta * S$ $\quad f_{\text{deathE}} = \delta * E$ $\quad f_{\text{deathI}} = \delta * I$ $\quad f_{\text{deathR}} = \delta * R$

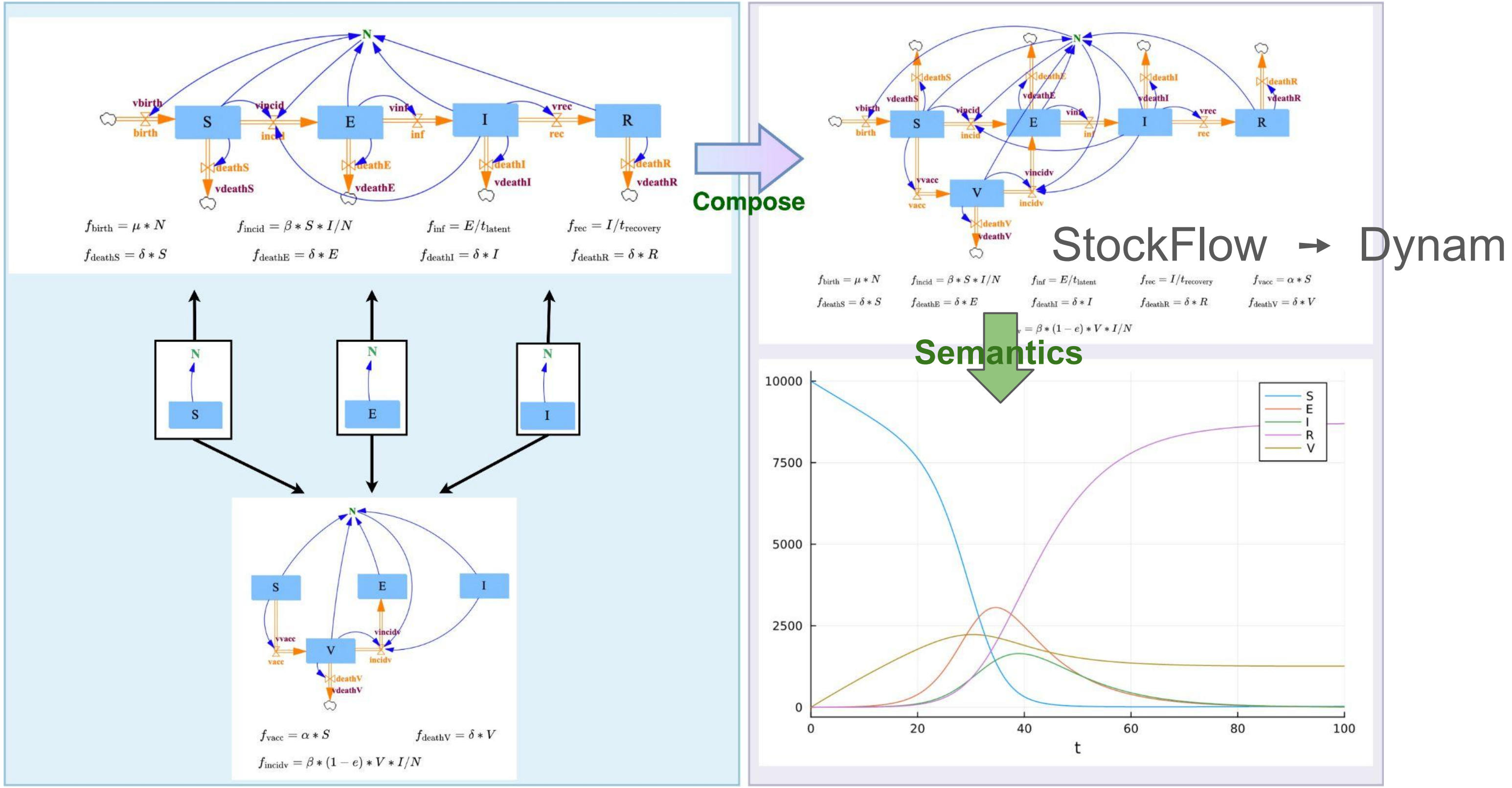


$f_{\text{vacc}} = \alpha * S$ $\quad f_{\text{deathV}} = \delta * V$

$f_{\text{incidv}} = \beta * (1 - e) * V * I/N$

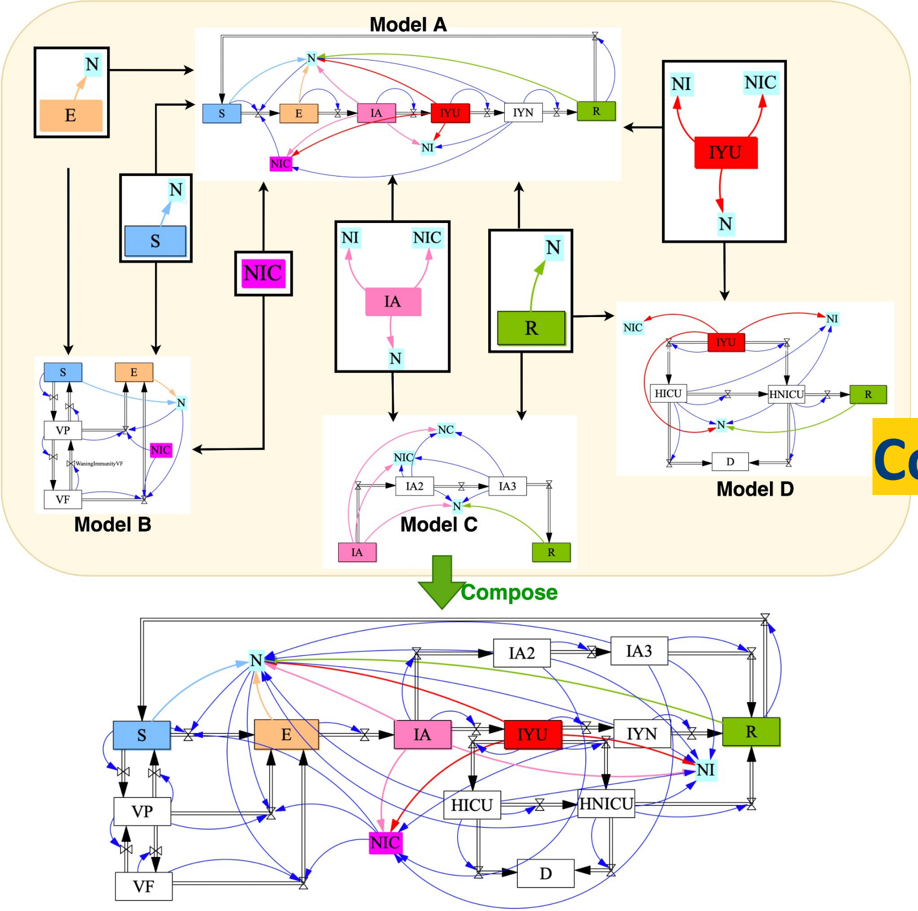


2. A Modular and Compositional Framework for System Dynamics Modeling



$$d(\mathrm{Stock}) / dt = \mathrm{sum}(\varphi_{\mathrm{Inflows}}) - \mathrm{sum}(\varphi_{\mathrm{Outflows}})$$

2. A Modular and Compositional Framework for System Dynamics Modeling

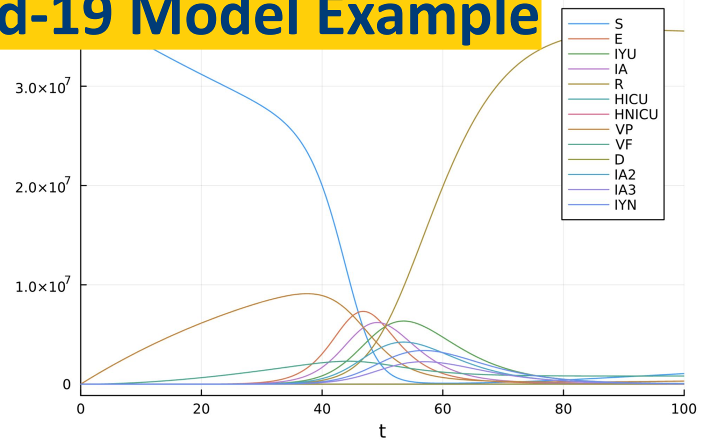


$$N_{IC} = I_A + I_{YU} + I_{YN} + I_{A2} + I_{A3}$$
$$N = S + E + V_P + V_F + I_A + I_{YU} + I_{YN} + I_{A2} + I_{A3} + H_{ICU} + H_{NICU} + R$$
$$N_I = I_A + I_{YU} + I_{YN} + I_{A2} + I_{A3} + H_{ICU} + H_{NICU}$$

$$\dot{S} = \frac{R}{t_w} + \frac{V_P}{t_w} - \frac{\beta S N_{IC}}{N} - r_v S$$
$$\dot{I_A} = \frac{E}{t_l} - \frac{I_A}{t_i}$$
$$\dot{I_{A2}} = f_{ia}\frac{I_A}{t_i} - \frac{I_{A2}}{t_d}$$
$$\dot{I_{YU}} = (1 - f_{ia})\frac{I_A}{t_l} - \frac{I_{YU}}{t_d}$$
$$\dot{V_F} = r_v V_P - \frac{V_F}{t_w} - \frac{\beta(1-e_f)N_{IC}V_F}{N}$$
$$\dot{H}_{ICU} = \frac{f_H f_{ICU} I_{YU}}{t_d} - \frac{H_{ICU}}{t_{ICU}} - r_D H_{ICU}$$

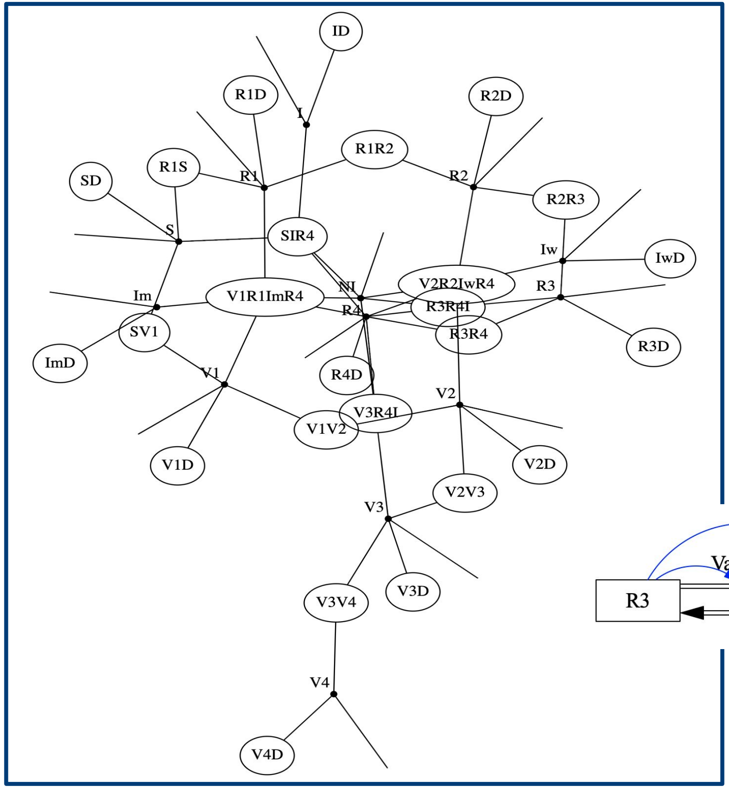
$$\dot{E} = \frac{\beta S N_{IC}}{N} + \frac{\beta(1-e_p)N_{IC}V_P}{N} + \frac{\beta(1-e_f)N_{IC}V_F}{N} - \frac{E}{t_l}$$
$$\dot{R} = \frac{I_{YN}}{t_r} + \frac{I_{A3}}{t_r} + \frac{H_{NICU}}{t_H} - \frac{R}{t_w}$$
$$\dot{I_{A3}} = \frac{I_{A2}}{t_d} - \frac{I_{A3}}{t_r}$$
$$\dot{I_{YN}} = (1 - f_H)\frac{I_{YU}}{t_d} - \frac{I_{YN}}{t_r}$$
$$\dot{D} = r_{DN} H_{NICU} + r_D H_{ICU}$$
$$\dot{V_P} = r_v S + \frac{V_F}{t_w} - \frac{V_P}{t_w} - r_v V_P - \frac{\beta(1-e_p)N_{IC}V_P}{N}$$
$$\dot{H}_{NICU} = \frac{H_{ICU}}{t_{ICU}} + \frac{f_H(1-f_{ICU})I_{YU}}{t_d} - \frac{H_{NICU}}{t_H} - r_{DN} H_{NICU}$$

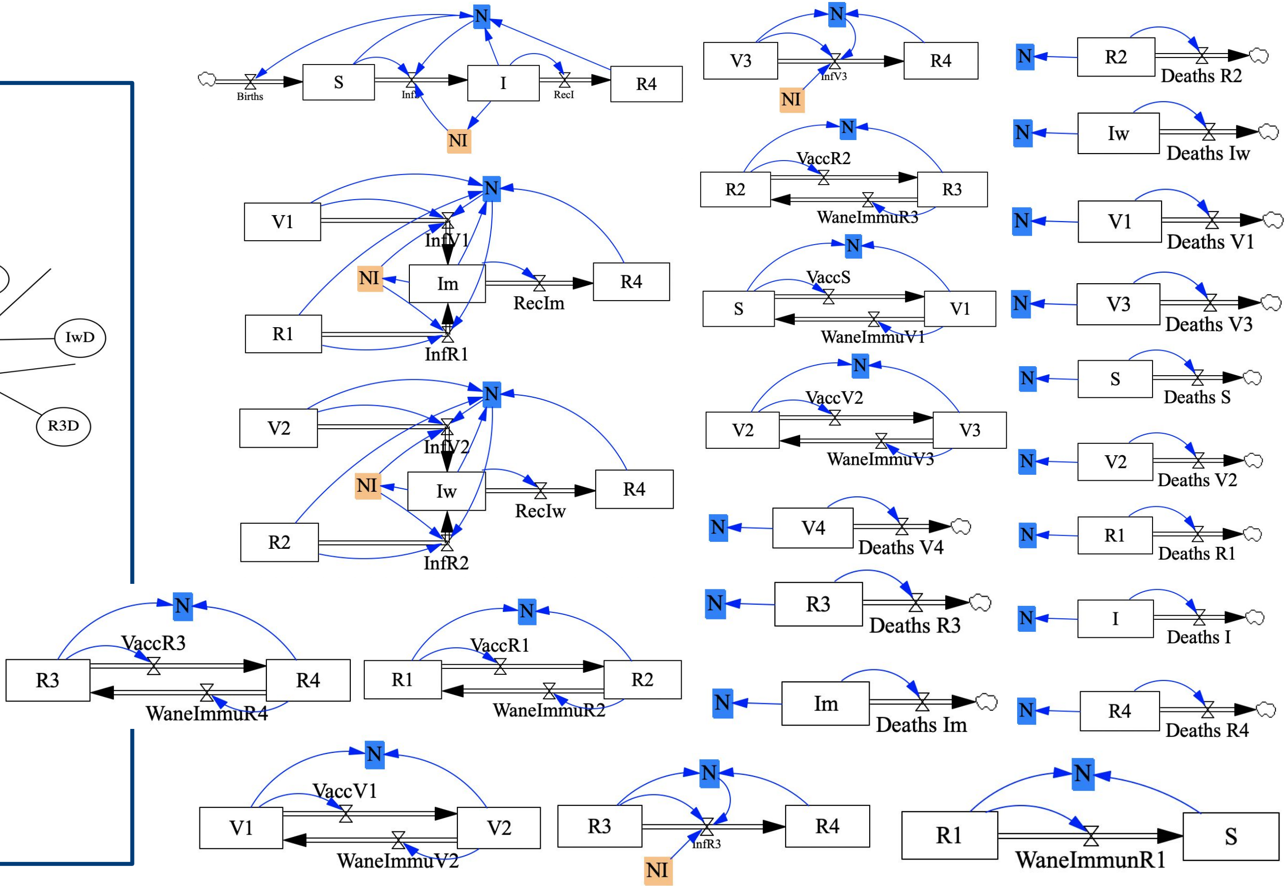
Covid-19 Model Example



2. A Modular and Compositional Framework for System Dynamics Modeling

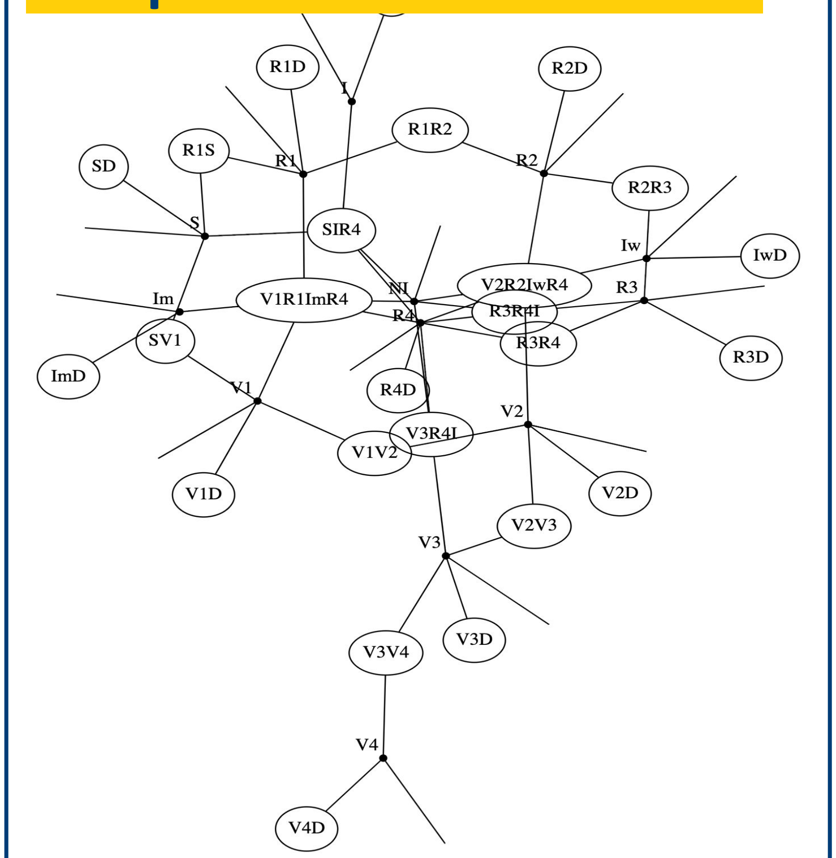
Operadic Algebra:

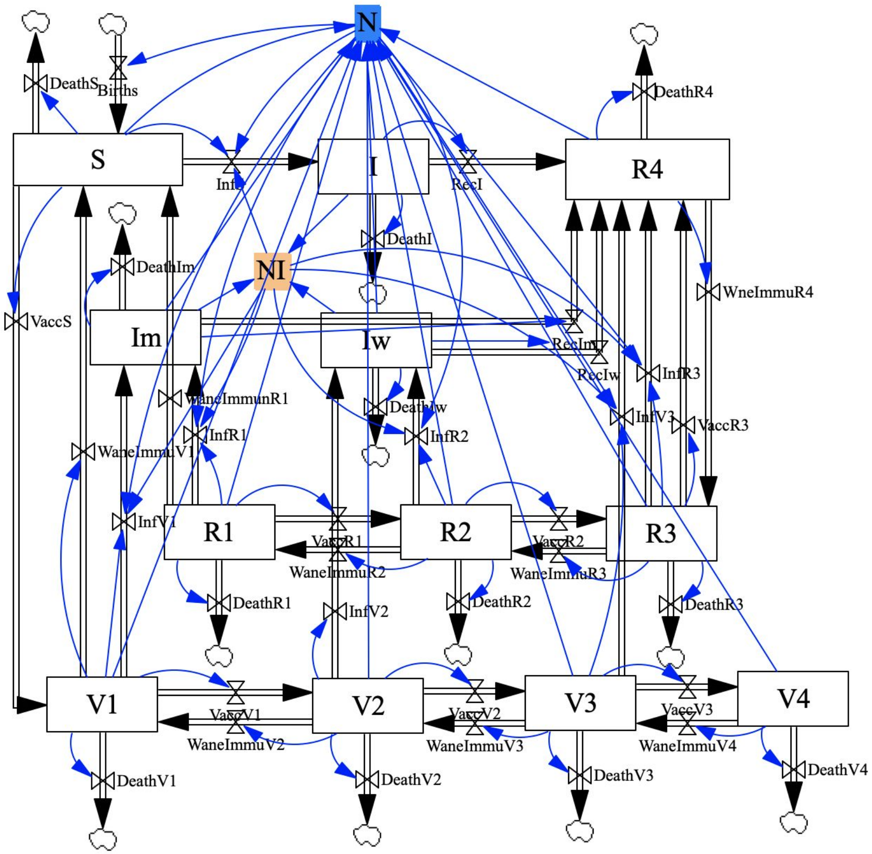




2. A Modular and Compositional Framework for System Dynamics Modeling

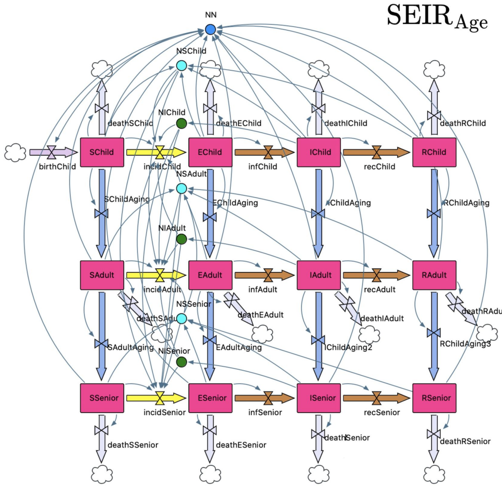
Composed Pertussis Model:



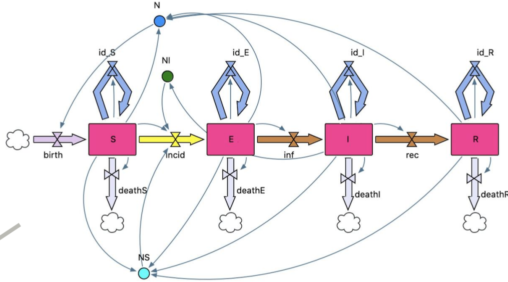


2. A Modular and Compositional Framework for System Dynamics Modeling

A Framework of Stratification

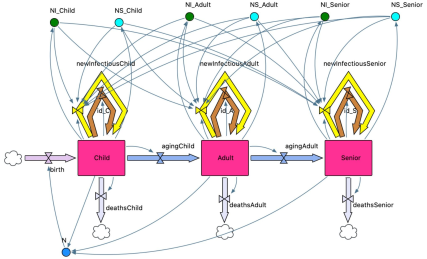


example

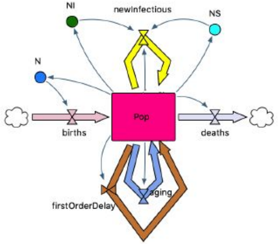


example

$$
\begin{tikzcd}
S_{\text{stratified}} \arrow[r, "p_1"] \arrow[d, "p_2"'] \arrow[dr, phantom, "\text{Pullback Square}"] & S_{\text{aggregate}} \arrow[d, "t_{\text{aggregate}}"] \\
S_{\text{strata}} \arrow[r, "t_{\text{strata}}"'] & S_{\text{type}}
\end{tikzcd}
$$

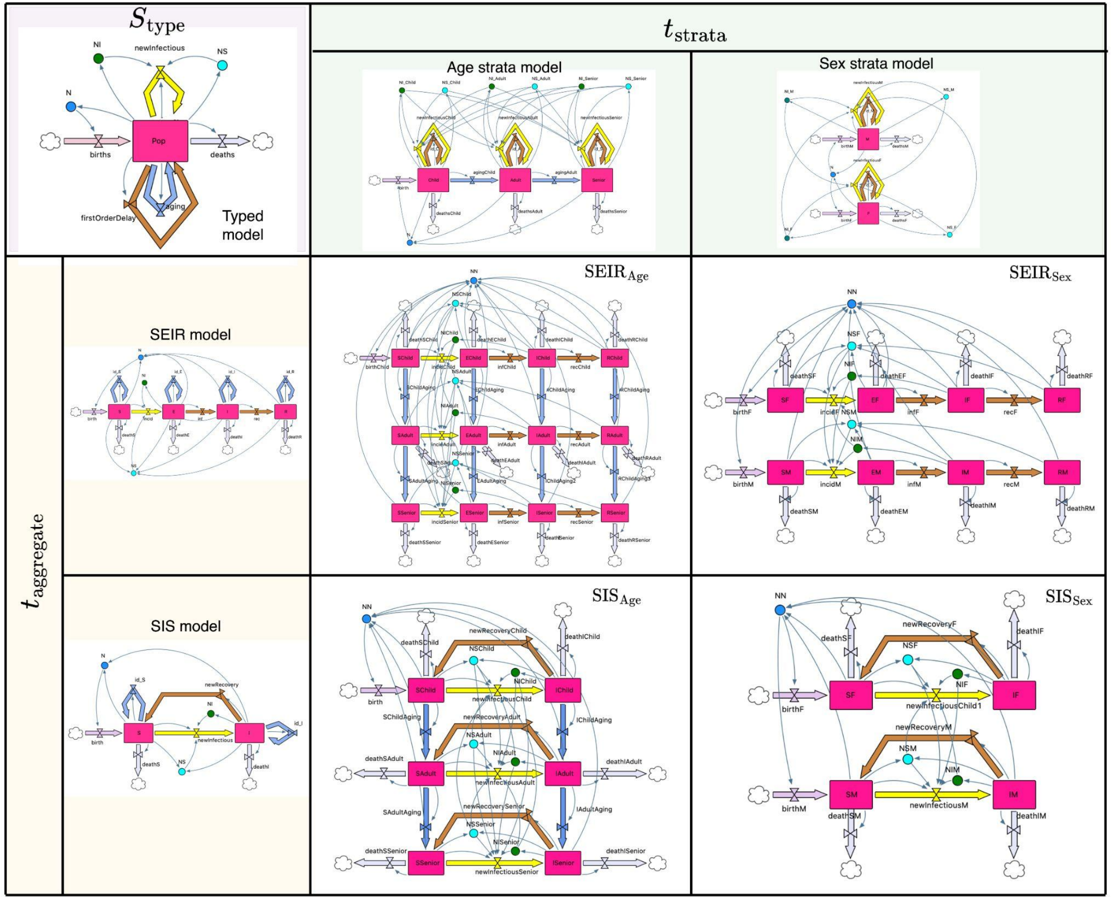


example



example

2. A Modular and Compositional Framework for System Dynamics Modeling



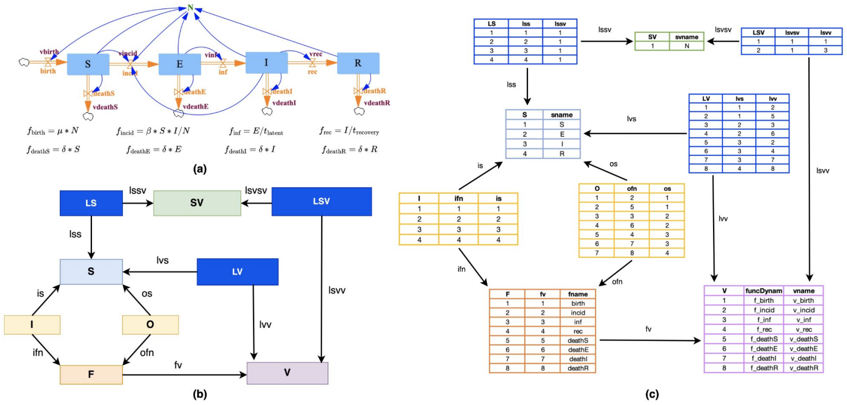
Examples of constructing stratified models

2. A Modular and Compositional Framework for System Dynamics Modeling

System Dynamics Modeling

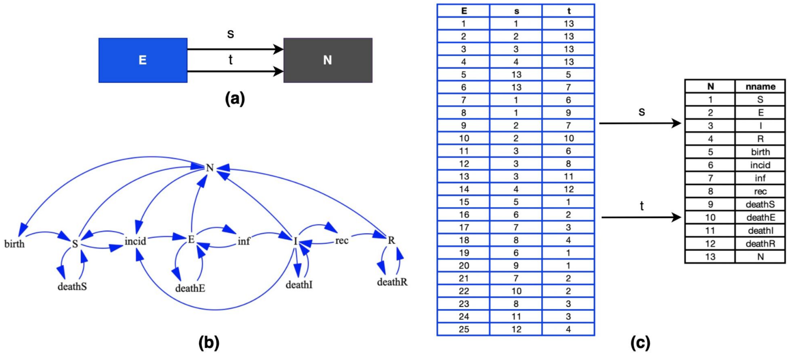
- Stock and Flow diagrams
- System Structure Diagrams
- Causal Loop Diagrams

$$\begin{tikzcd} StockFlow \arrow[rr, "B"] \arrow[dr, "A"'] & & CausalLoop \\ & SystemStructure \arrow[ur, "C"'] & \end{tikzcd}$$



Stock and Flow diagram

functor



Causal Loop Diagram

Take-Home Messages

- Impactful modeling projects are best contributed by interdisciplinary teams
- System Dynamics is a diagram- and stakeholder-focused tradition offering high potential for transparency and team support, but greatly limited by extant tools
- ACT empowers teams via transparency, modularity, composition, abstraction interconnections between diagram types, cleaner stratification & enhanced semantic flexibility
- ACT-based systems can support System Dynamics modeling in teams without end-user familiarity with category theory

Q&A

Thank you!