

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

**Xiaoyan Li**<sup>1,4,5</sup>

Supervisor: Nathaniel Osgood<sup>4</sup>

Joint work with J. Baez<sup>1,2,3</sup>, E. Patterson<sup>1</sup>, S. Libkind<sup>1</sup>,  
A. Alegre<sup>4</sup>, T. Purdy<sup>4</sup>, E. Redekopp<sup>4</sup> & N. Meadows<sup>4</sup>

<sup>1</sup>Topos Institute <sup>2</sup>University of Edinburgh <sup>3</sup>University of California, Riverside  
<sup>4</sup>University of Saskatchewan <sup>5</sup>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<sup>1</sup>:

Weekly highlights - Data from October 20, 2024 to October 26, 2024

32 (120 this season) Influenza cases	591 (4462 this season) SARS-CoV-2 cases	30 (85 this season) RSV cases
1.4% Influenza positivity	12.6% SARS-CoV-2 positivity	0.8% RSV positivity
6 Influenza in hospital**	313 SARS-CoV-2 in hospital**	0 RSV in hospital**
0 Influenza in ICU**	15 SARS-CoV-2 in ICU**	0 RSV in ICU**
0 (1 this season) Influenza deaths	4 (119 this season) SARS-CoV-2 deaths	0 (0 this season) RSV deaths

## Characteristics of Big Data:

Large Volume

High Velocity

High Variety

## Big data<sup>2</sup>:



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:

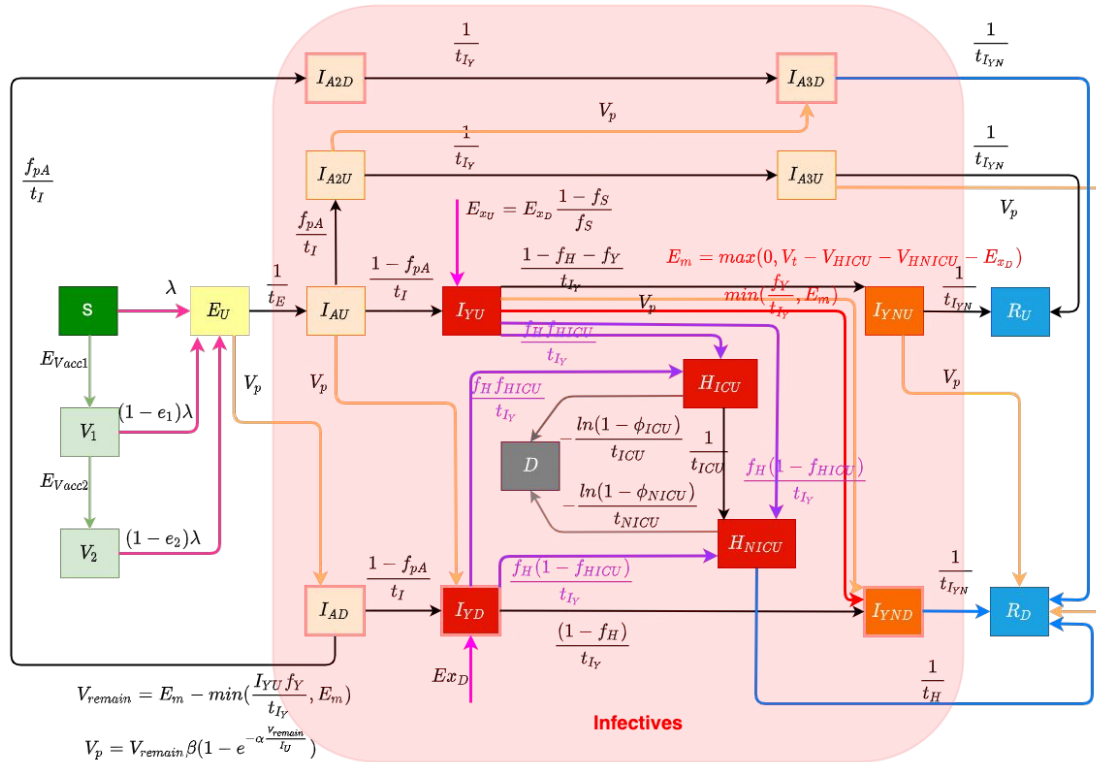
- Using **Dynamic Models** simulate the system
- Using Machine Learning Algorithms (**Bayesian models**) and (large scale) **empirical datasets** to train the machine learning dynamic models
- **Project** forward based on the trained models
- Perform **counterfactual analysis** to help **decision making**.

# 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:

$$\frac{dS}{dt} = -\lambda S$$

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

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

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

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

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

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

$$\frac{dI_{A3U}}{dt} = \frac{I_{A2U}}{t_{I_Y}} - V_p \frac{I_{A3U}}{I_U} - \frac{I_{A3U}}{t_{I_{Y_N}}}$$

$$\frac{dI_{A3D}}{dt} = \frac{I_{A2D}}{t_{I_Y}} + V_p \frac{I_{A3U}}{I_U} - \frac{I_{A3D}}{t_{I_{Y_N}}}$$

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

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

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

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

$$\frac{dI_{YNU}}{dt} = I_{YU} \frac{1 - f_H}{t_{I_Y}} - \frac{I_{YNU}}{t_{I_{Y_N}}} - V_p \frac{I_{YNU}}{I_U}$$

$$\frac{dI_{YND}}{dt} = f_{I_r} \frac{I_{YNU}}{t_{I_{Y_N}}} + V_p \frac{I_{YNU}}{I_U} + \frac{I_{YD}}{t_{I_Y}} + \min\left(\frac{I_{YU} f_{NH}}{t_{I_Y}}, E_m\right) - \frac{I_{YND}}{t_{I_{Y_N}}}$$

$$\frac{dR_U}{dt} = (1 - f_{I_r}) \frac{I_{YNU}}{t_{I_{Y_N}}} + \frac{I_{A3U}}{t_{I_{Y_N}}}$$

$$\frac{dR_D}{dt} = \frac{I_{YND}}{t_{I_{Y_N}}} + \frac{H_{ICU}}{t_H} + \frac{I_{A3D}}{t_{I_{Y_N}}}$$

$$\frac{dD}{dt} = \left(H_{NICU} \frac{-\ln(1 - \phi_{NICU})}{t_{NICU}}\right) + \left(H_{ICU} \frac{-\ln(1 - \phi_{ICU})}{t_{ICU}}\right)$$

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

## Dynamic Parameters:

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

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

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

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

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

$$V_{remain} = E_m - \min\left(\frac{f^{NH}}{t_{I_Y}}, E_m\right)$$

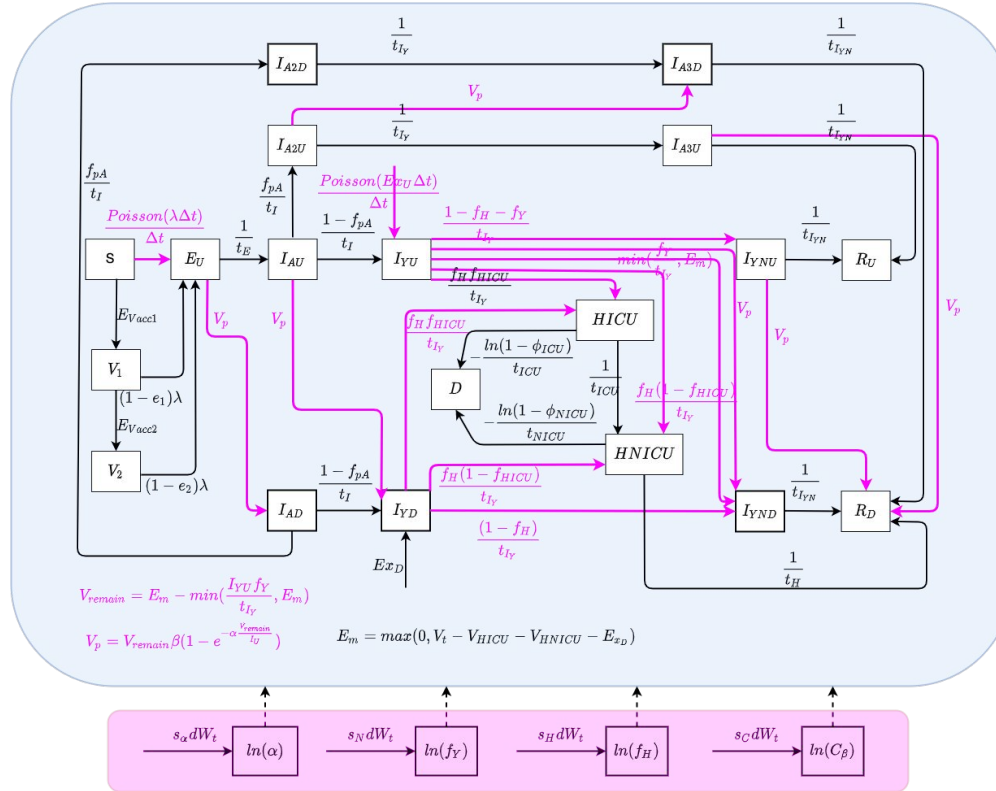
$$V_p = V_{remain} \beta \left(1 - e^{-\alpha \frac{V_{remain}}{I_U}}\right)$$

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

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



# 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

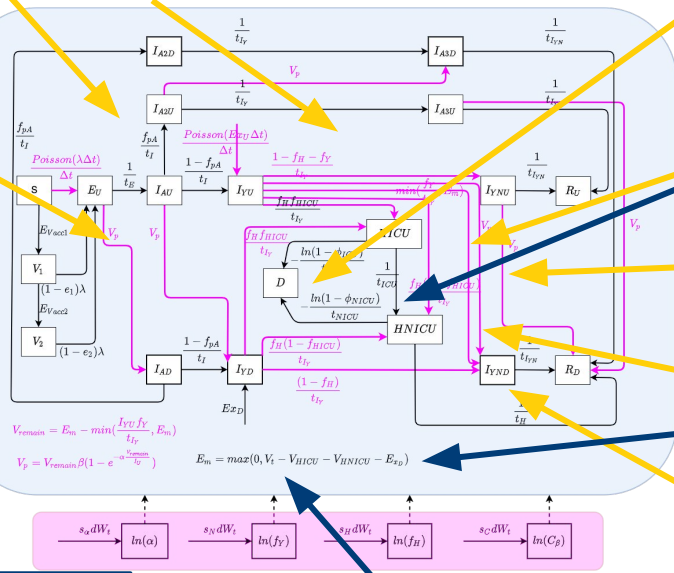
8. Daily COVID-19 patients admitted into hospital for non-ICU care

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

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

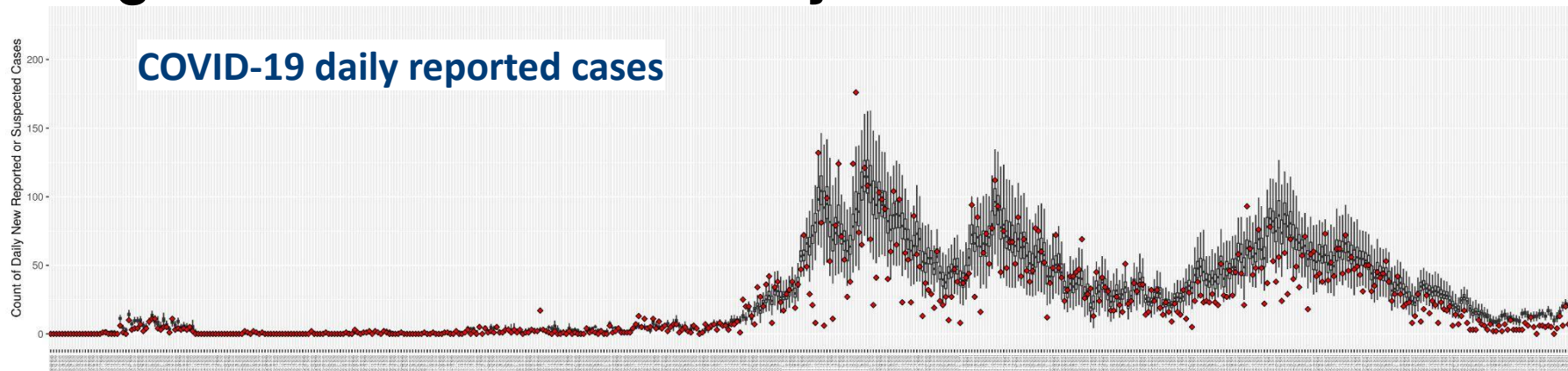
11. Daily midnight census (count) of COVID-19 patients in the hospital for non-ICU care

12. Daily new likely exogenous cases

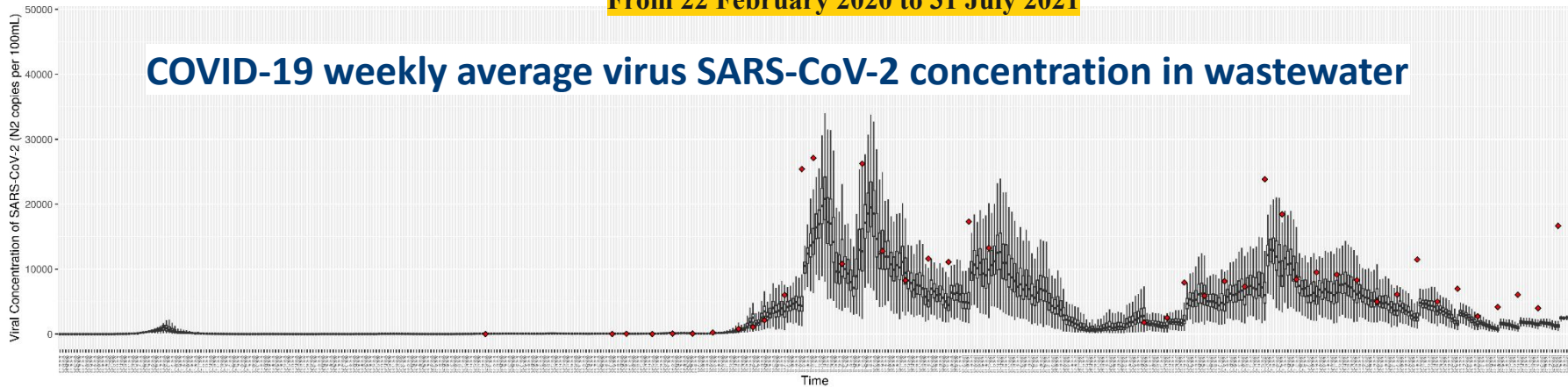


12 empirical datasets: used to **drive the model**, and used in the PF likelihoods.

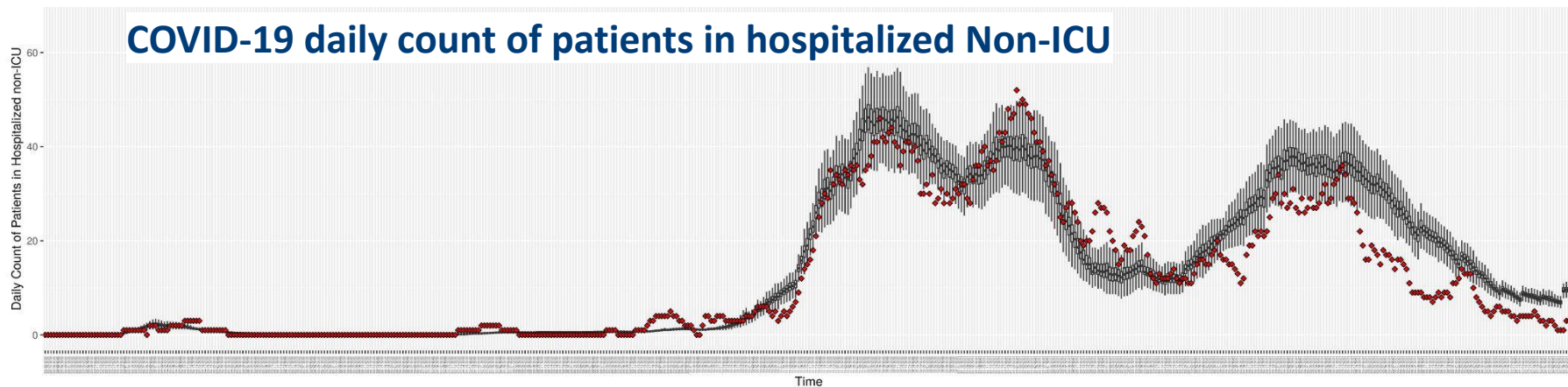
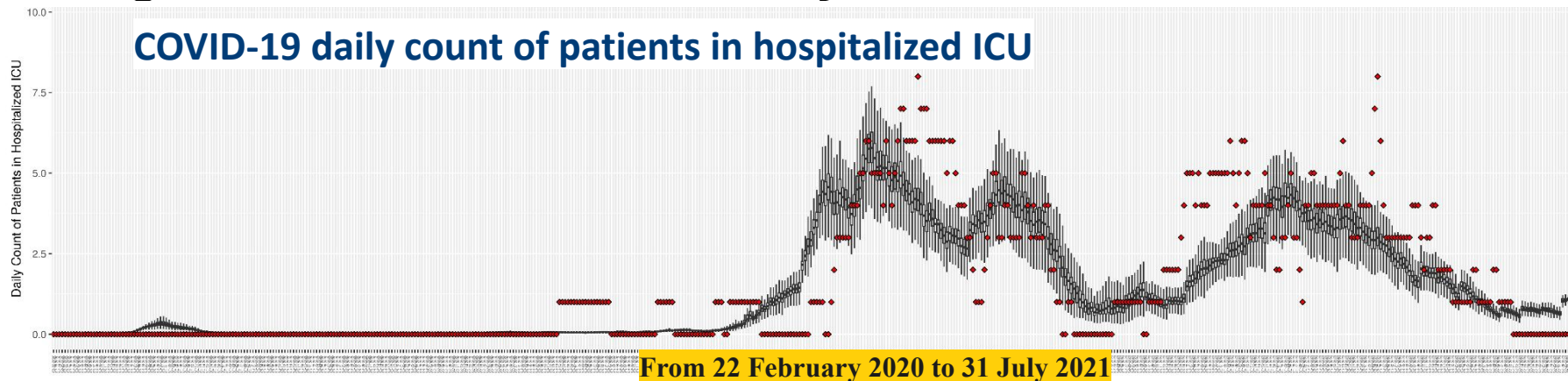
# 1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention



From 22 February 2020 to 31 July 2021

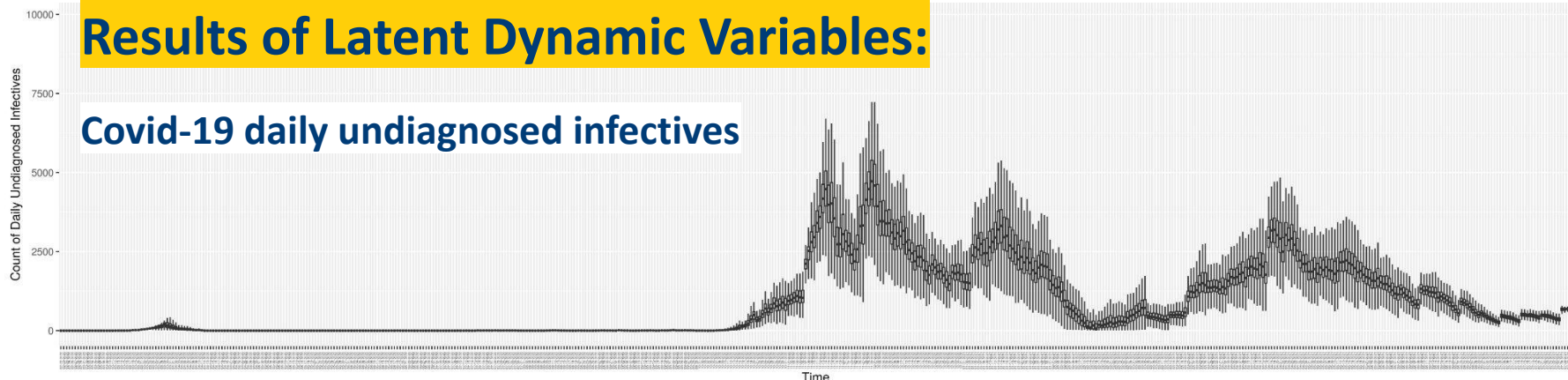


# 1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

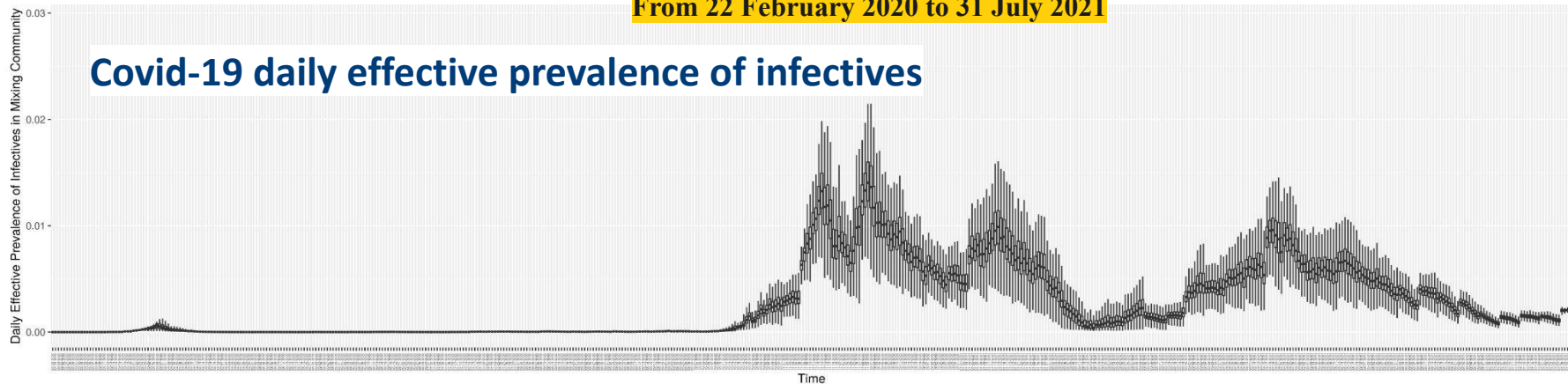


# 1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention

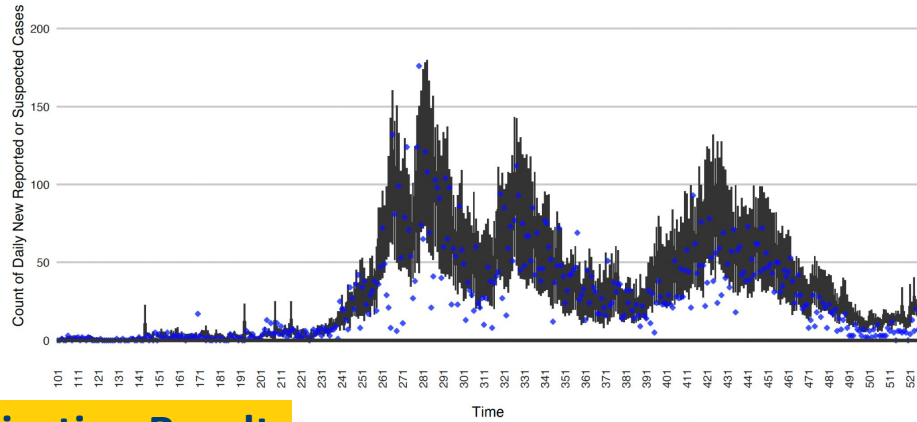
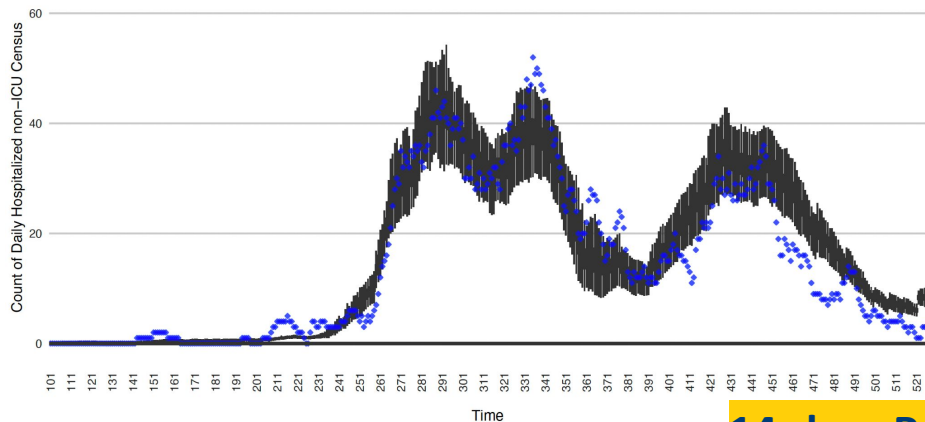
## Results of Latent Dynamic Variables:



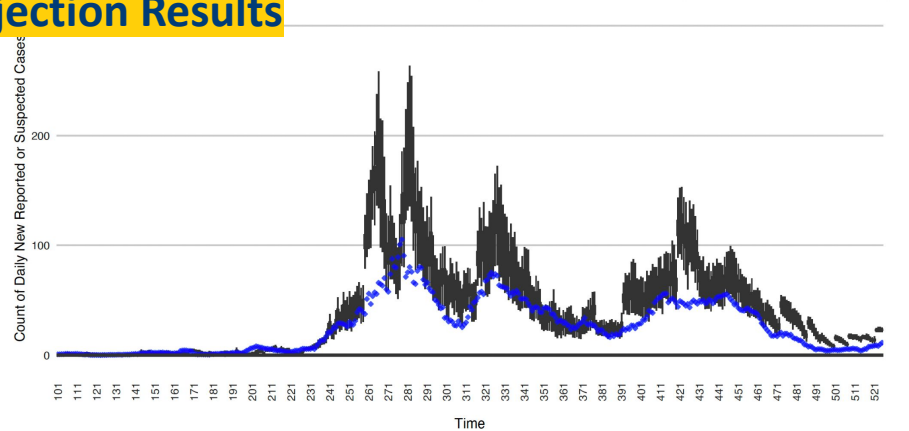
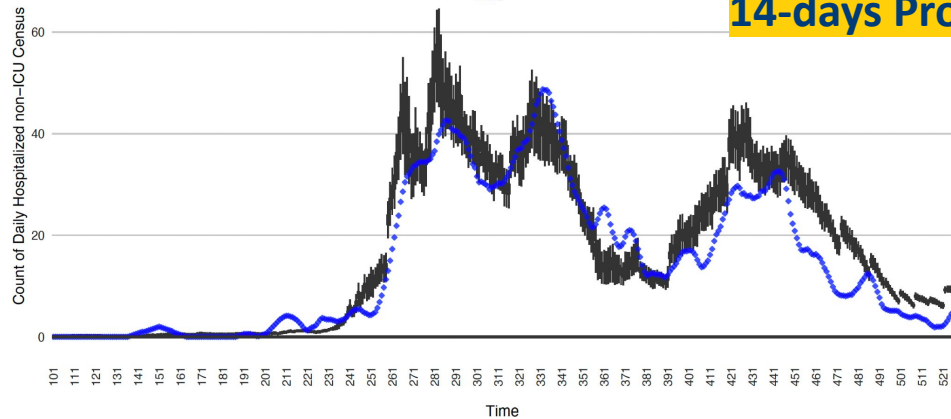
From 22 February 2020 to 31 July 2021



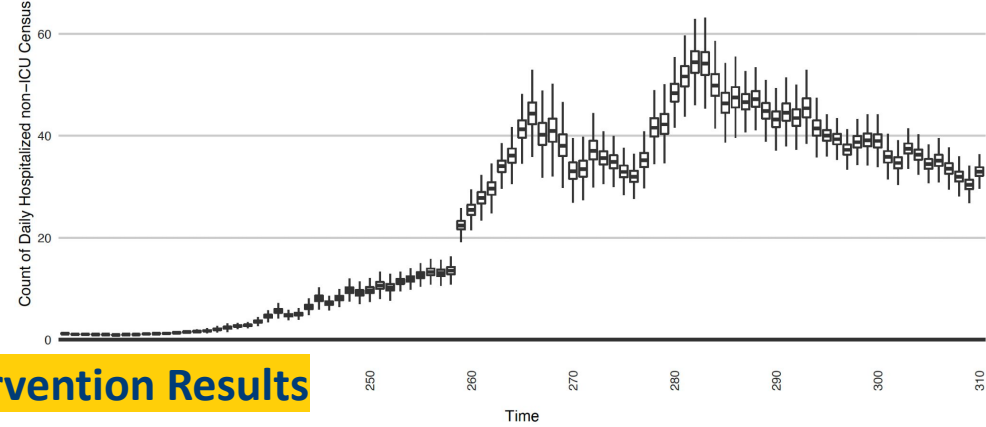
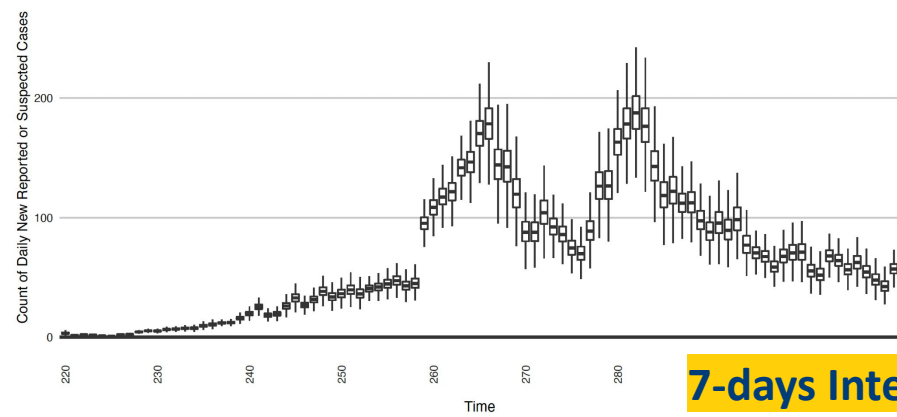
# 1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention



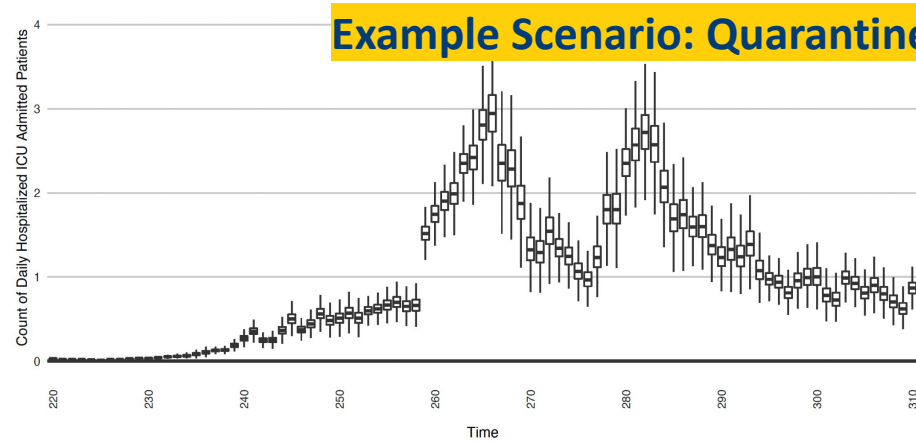
**14-days Projection Results**



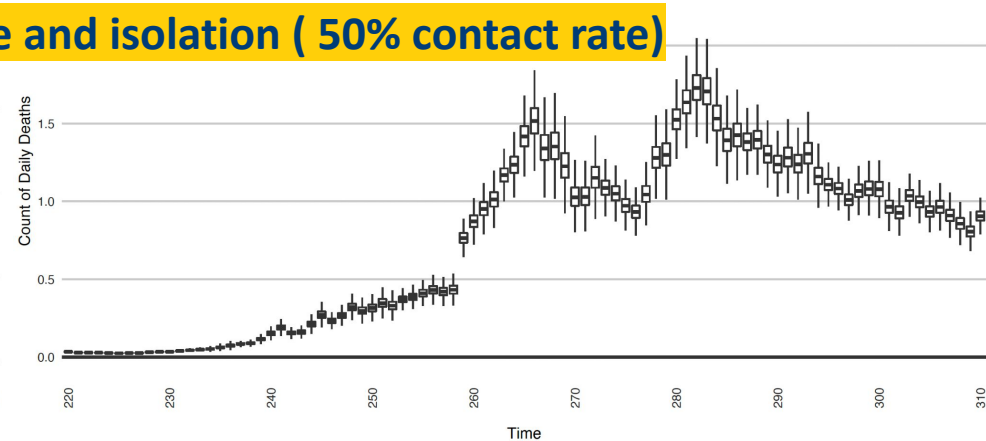
# 1. A Bayesian Machine Learning Framework for Large-Scale Time Series Projection and Intervention



**7-days Intervention Results**



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



## 2. A Modular and Compositional Framework for System Dynamics Modeling

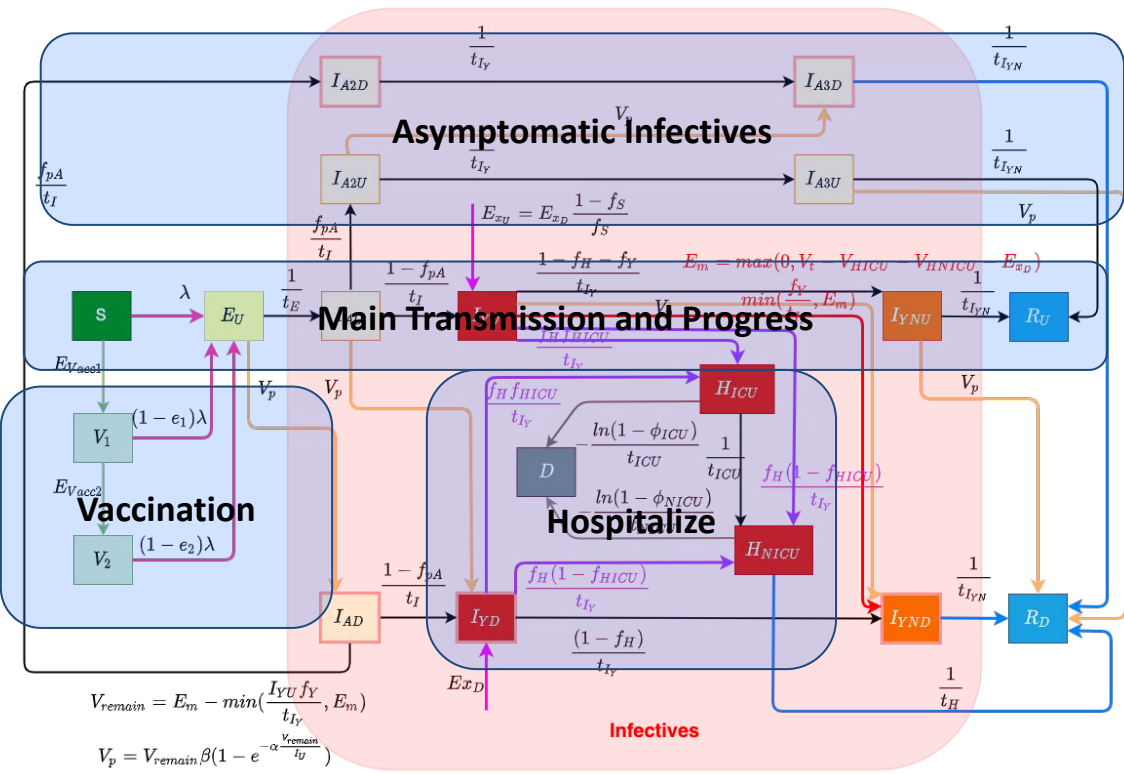
### Motivations:

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

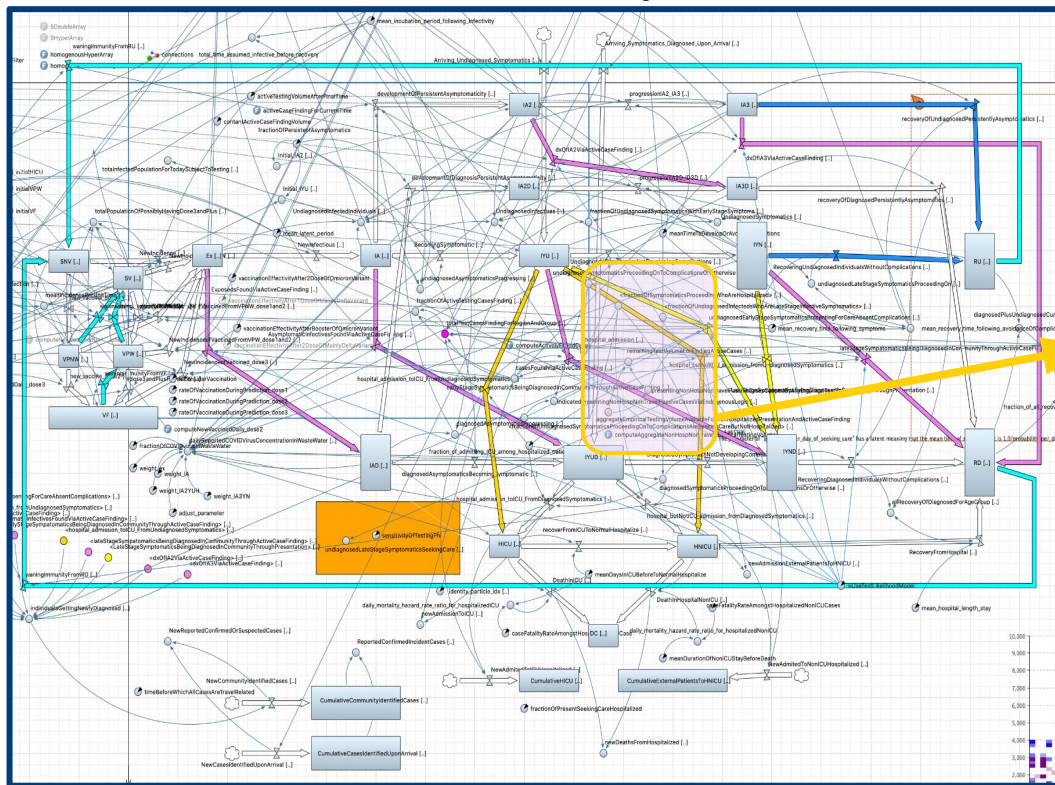


Three models with same structure:

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

```
computeAggregateNonHospNonTravelEmpiricalTestingVolumes - Function
// calculate the number of people using non hospital or travel testing
// we need not put in extra code of this sort

// First, for this region & age group, compute the TEST VOLUME applied for people NOT required
// NB: since this works only with empirical data, this does not vary by particle
// Similarly, compute the number of positives for people NOT requiring HOSPITALIZATION

double nonHospPresentingTestVolumes = 1000.0;
if (time() < predictionStartTime)
{
    double dEmpiricalTestingVolumeForCurrentDay = isUseManualTestingInformation? empiricalTest
    assert (abs(1 - dEmpiricalTestingVolumeForCurrentDay) > epsilon) : "Found -1 to indicate
    // The model represents hospitalization flows at a more detailed level, so we have to aggreg
    // each of these flows is assumed to represent both a TEST, and a POSITIVE COUNT
    // TODO-xiaoyan, add the hospitalized data after receiving Dr.D's data
    double dEmpiricalHospitalizationPresentingTestVolumes = -1.0;
    // Either
    // 1) this hasCurrentTimeHospitalAdmissionData is already computed for the current ti
    // 2) it is set in this call
    initializeAdvancedHospitalAndViralConcentrationDataAvailability();
    if (isUseHospitalAdmissionLikelihood & this.hasCurrentTimeHospitalAdmissionData)
    {
        double dEmpirical_New_HICUFORRegionAndAgeGroup = TableFunctionsEmpiricalDataNewHospit
        double dEmpirical_New_HICUFORRegionAndAgeGroup = TableFunctionsEmpiricalDataNewHospit
        // traceIn("****n\n\n");
        // traceIn("dEmpirical_New_HICUFORRegionAndAgeGroup: " + dEmpirical_New_HICUFORRegionAn
        // traceIn("dEmpirical_New_HICUFORRegionAndAgeGroup: " + dEmpirical_New_HICUFORRegionAndA
    }
    dEmpiricalHospitalizationPresentingTestVolumes = dEmpirical_New_HICUFORRegionAndAgeGr
    } else
    {
        dEmpiricalHospitalizationPresentingTestVolumes = dEmpiricalTestingVolumeForCurrentDay
    }
    // as noted, we assume that for each hospitalized case has not only required — and thus a
    // double dEmpiricalHospitalizationPresentingPositiveTestCount = dEmpiricalHospitalizatio
    // the test volumes for those NOT presenting to hospital are just the total test volum
    // minus the corresponding values for any sort of hospitalization; because they come in v
    //****WARNING 1000: We should really not depend directly here upon rateOfArrivalOfInfecti
    double unboundedNominalNonHospitalizationPresentingTestVolumes = dEmpiricalTestingVolumeFo
    - ((isUseTimeSeriesOfDiagnosedInternationalTravelers & time()) <= finalTimeOfEmpiric
    /> traceIn("n\n\n");
    traceIn("unboundedNominalNonHospitalizationPresentingTestVolumes: " + unboundedNominalNon
    traceIn("dEmpiricalTestingVolumeForCurrentDay: " + dEmpiricalTestingVolumeForCurrentDay);
    traceIn("isUseTimeSeriesOfDiagnosedInternationalTravelers & time() <= finalTimeOfEmpiric
}

nonHospPresentingTestVolumes = max(0, unboundedNominalNonHospitalizationPresentingTestVolu
}
double nonHospHospitalizationPresentingPositiveTestCount = max(0, dEmpiricalTestingCountPositi
// ok, having computed the test volumes and counts for elective cases (those not forced by co
// active screening), we can now compute how many people we would expect to be found
// Because of the heavy reliance of case finding on symptoms in Canada, we assume that these c
// symptomatic stocks. Here, we do not get into from which stocks these come — we are just
// stocks. This will, in fact, come from multiple such stocks.
// ok, having now computed the empirical values to be used for this region & age group, we now
// computation of how many people make the transition, based on the per-particle state in ter
```

## 2. A Modular and Compositional Framework for System Dynamics Modeling

Research Question:



Traditional Modeling Methods

?

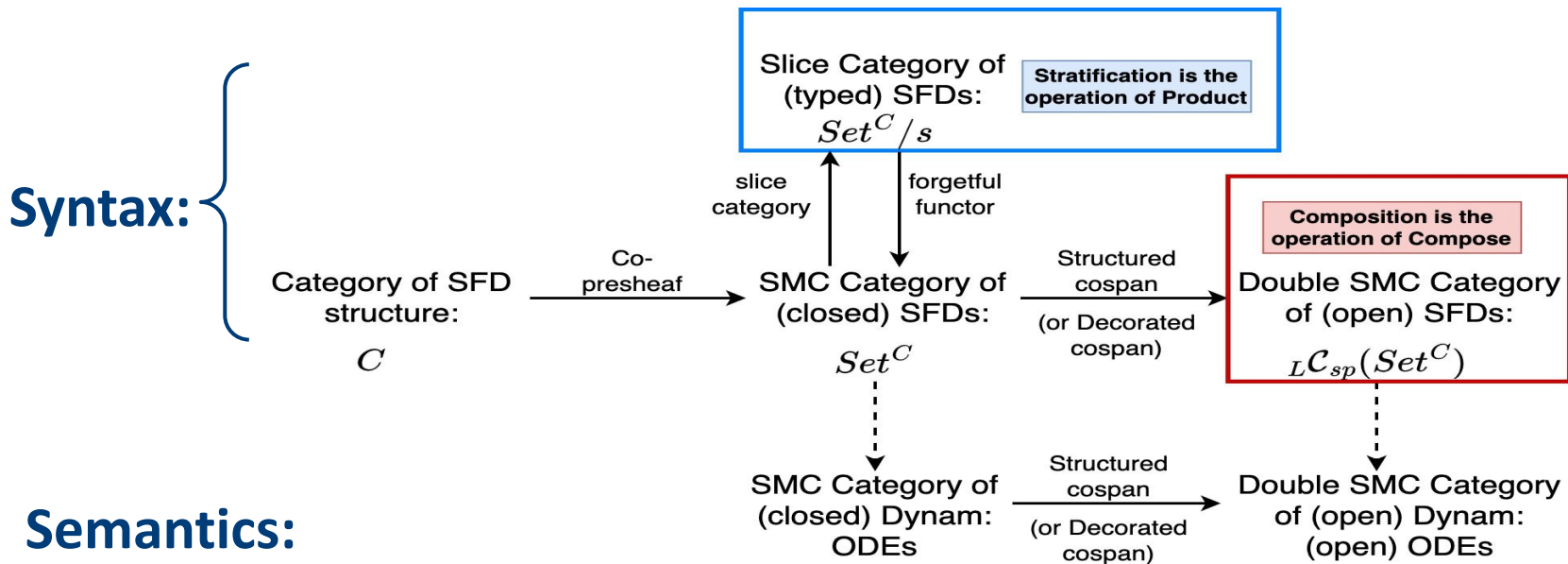


A Next-generation Modular Modeling Framework of Constructing Models by Composition

## 2. A Modular and Compositional Framework for System Dynamics Modeling

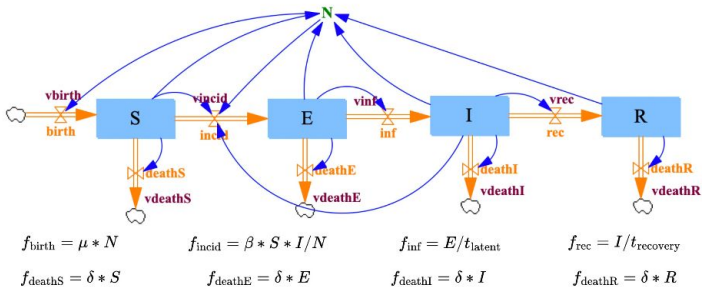
### Methods: Applied Category Theory

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

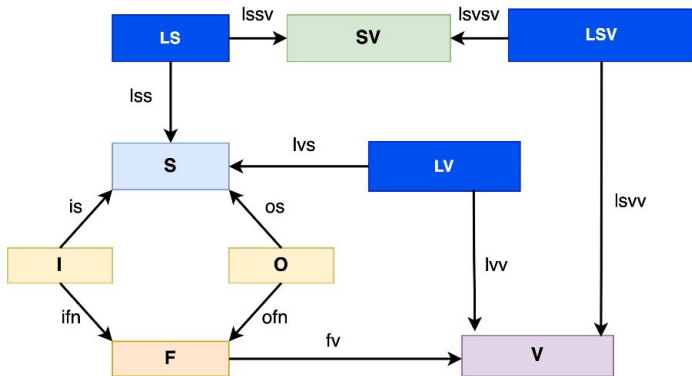


# 2. A Modular and Compositional Framework for System Dynamics Modeling

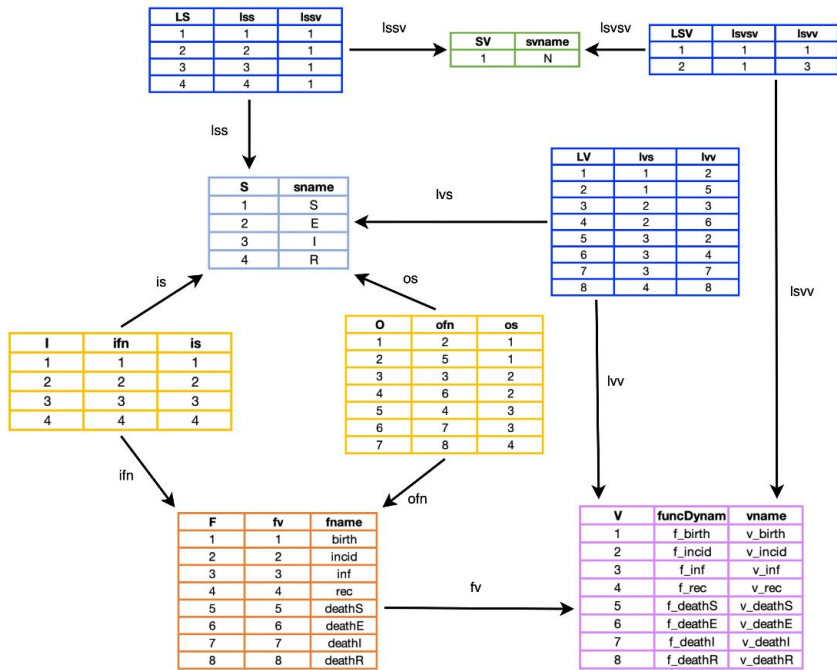
## Category of Closed Stock and Flow Diagrams: (FinSet<sup>C</sup>, φ)



(a)



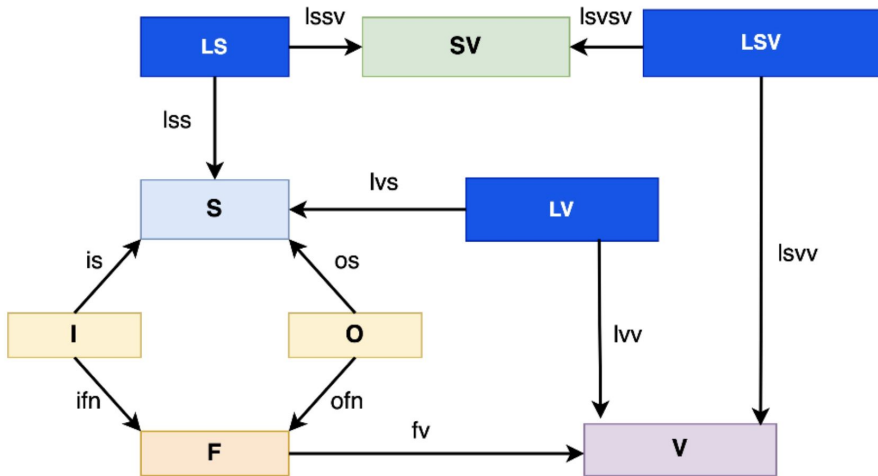
(b)



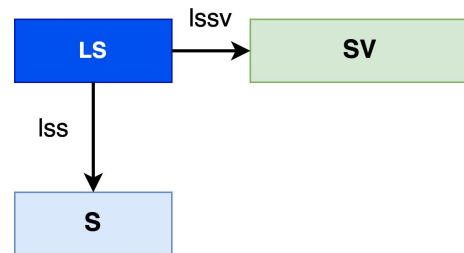
(c)

## 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)

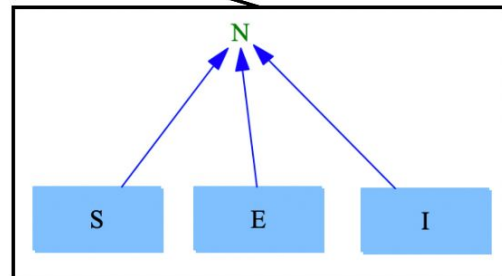
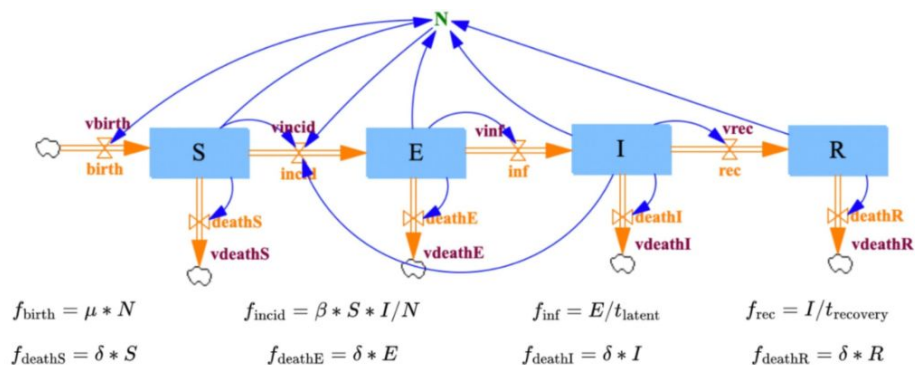


Schema of interface (H)

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

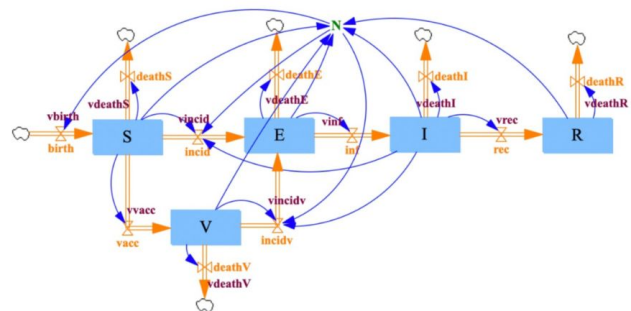
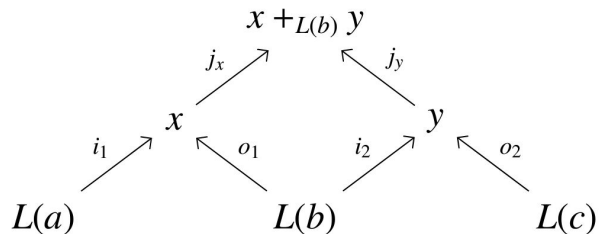
## 2. A Modular and Compositional Framework for System Dynamics Modeling

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



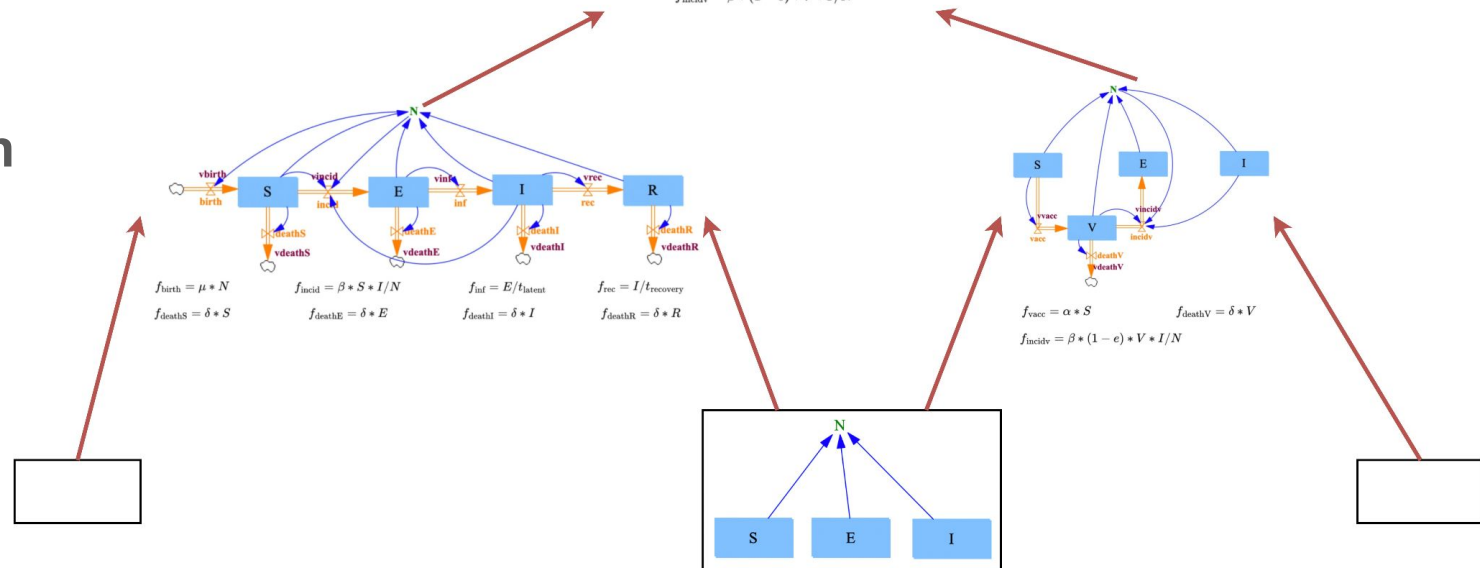
$$L(a) \xrightarrow{i} x \xleftarrow{o} L(b)$$

# 2. A Modular and Compositional Framework for System Dynamics Modeling



$$\begin{aligned}
 f_{\text{birth}} &= \mu * N & f_{\text{incid}} &= \beta * S * I / N & f_{\text{inf}} &= E / t_{\text{latent}} & f_{\text{rec}} &= I / t_{\text{recovery}} & f_{\text{vacc}} &= \alpha * S \\
 f_{\text{deathS}} &= \delta * S & f_{\text{deathE}} &= \delta * E & f_{\text{deathI}} &= \delta * I & f_{\text{deathR}} &= \delta * R & f_{\text{deathV}} &= \delta * V \\
 f_{\text{incidV}} &= \beta * (1 - \epsilon) * V * I / N
 \end{aligned}$$

## Composition

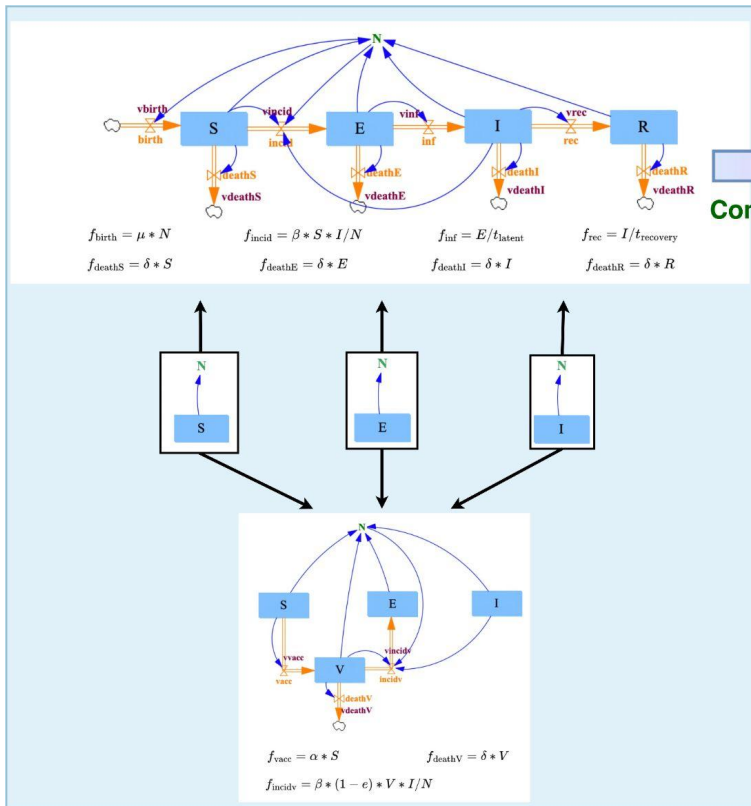


$$\begin{aligned}
 f_{\text{birth}} &= \mu * N & f_{\text{incid}} &= \beta * S * I / N & f_{\text{inf}} &= E / t_{\text{latent}} & f_{\text{rec}} &= I / t_{\text{recovery}} \\
 f_{\text{deathS}} &= \delta * S & f_{\text{deathE}} &= \delta * E & f_{\text{deathI}} &= \delta * I & f_{\text{deathR}} &= \delta * R
 \end{aligned}$$

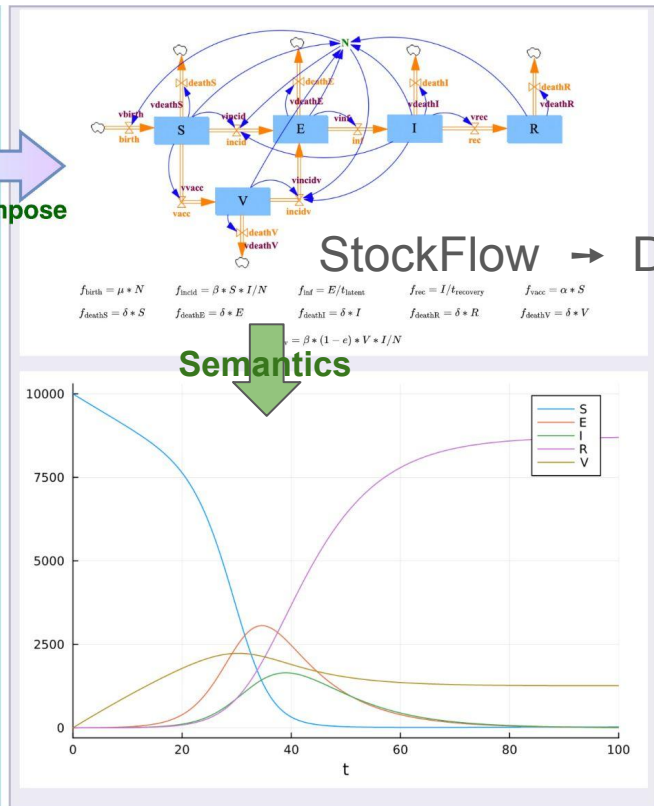
$$\begin{aligned}
 f_{\text{vacc}} &= \alpha * S & f_{\text{deathV}} &= \delta * V \\
 f_{\text{incidV}} &= \beta * (1 - \epsilon) * V * I / N
 \end{aligned}$$



# 2. A Modular and Compositional Framework for System Dynamics Modeling

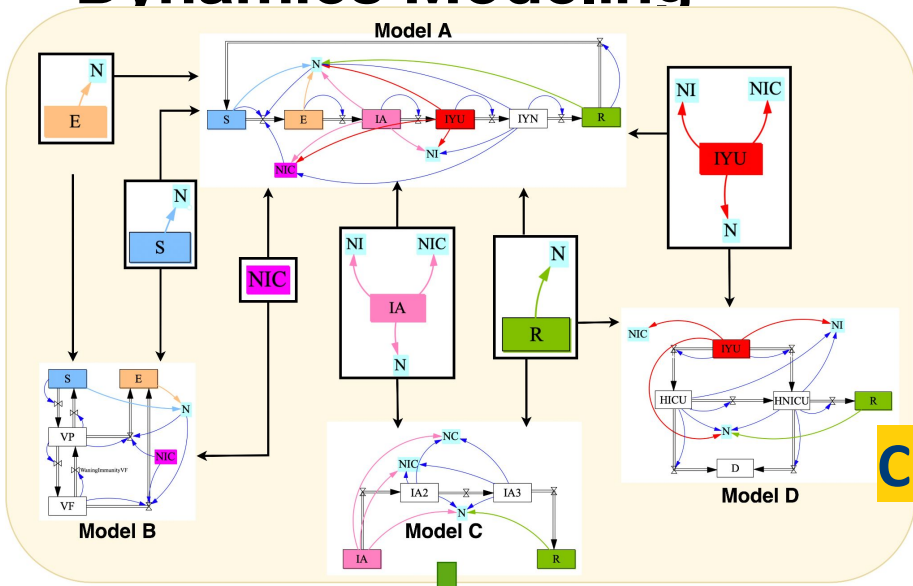


Compose

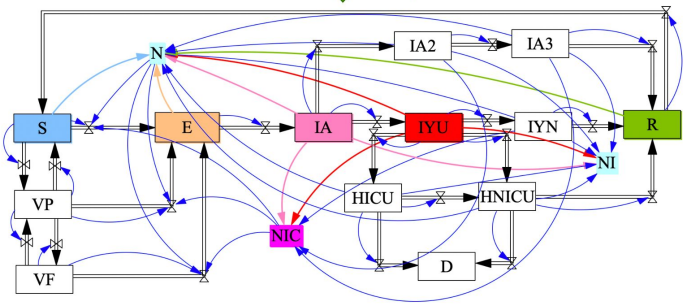


$$d(\text{Stock}) / dt = \text{sum}(\varphi_{\text{Inflows}}) - \text{sum}(\varphi_{\text{Outflows}})$$

# 2. A Modular and Compositional Framework for System Dynamics Modeling



Compose



$$N_{IC} = I_A + I_{YU} + I_{YN} + I_{A2} + I_{A3}$$

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

$$N_I = I_A + I_{YU} + I_{YN} + I_{A2} + I_{A3} + H_{ICU} + H_{NICU}$$

$$\dot{S} = \frac{R}{t_w} + \frac{V_P}{t_w} - \frac{\beta S N_{IC}}{N} - r_v S$$

$$\dot{E} = \frac{\beta S N_{IC}}{N} + \frac{\beta(1-e_p)N_{IC}V_P}{N} + \frac{\beta(1-e_f)N_{IC}V_F}{N} - \frac{E}{t_i}$$

$$\dot{I}_A = \frac{E}{t_i} - \frac{I_A}{t_i}$$

$$\dot{I}_{YU} = \frac{I_{YU}}{t_r} + \frac{I_{A3}}{t_r} + \frac{H_{NICU}}{t_H} - \frac{R}{t_w}$$

$$\dot{I}_{A2} = f_{ia} \frac{I_A}{t_i} - \frac{I_{A2}}{t_d}$$

$$\dot{I}_{A3} = \frac{I_{A2}}{t_d} - \frac{I_{A3}}{t_r}$$

$$\dot{I}_{YN} = (1-f_{ia}) \frac{I_A}{t_i} - \frac{I_{YN}}{t_d}$$

$$\dot{I}_{YN} = (1-f_H) \frac{I_{YU}}{t_d} - \frac{I_{YN}}{t_r}$$

$$\dot{V}_F = r_v V_P - \frac{V_F}{t_w} - \frac{\beta(1-e_f)N_{IC}V_F}{N}$$

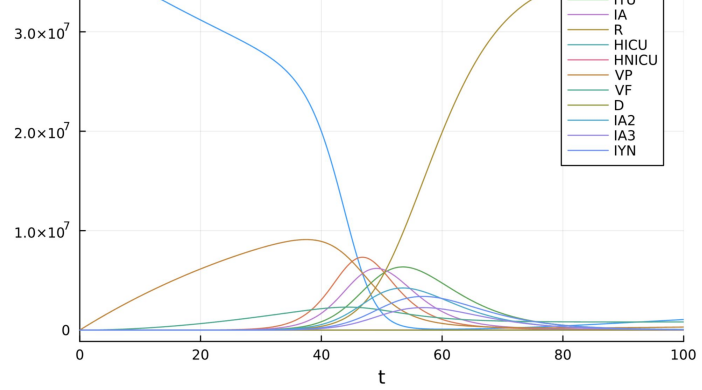
$$\dot{D} = r_{DN} H_{NICU} + r_D H_{ICU}$$

$$\dot{H}_{ICU} = \frac{f_H f_{ICU} I_{YU}}{t_d} - \frac{H_{ICU}}{t_{ICU}} - r_D H_{ICU}$$

$$\dot{V}_P = r_v S + \frac{V_F}{t_w} - \frac{V_P}{t_w} - r_v V_P - \frac{\beta(1-e_p)N_{IC}V_P}{N}$$

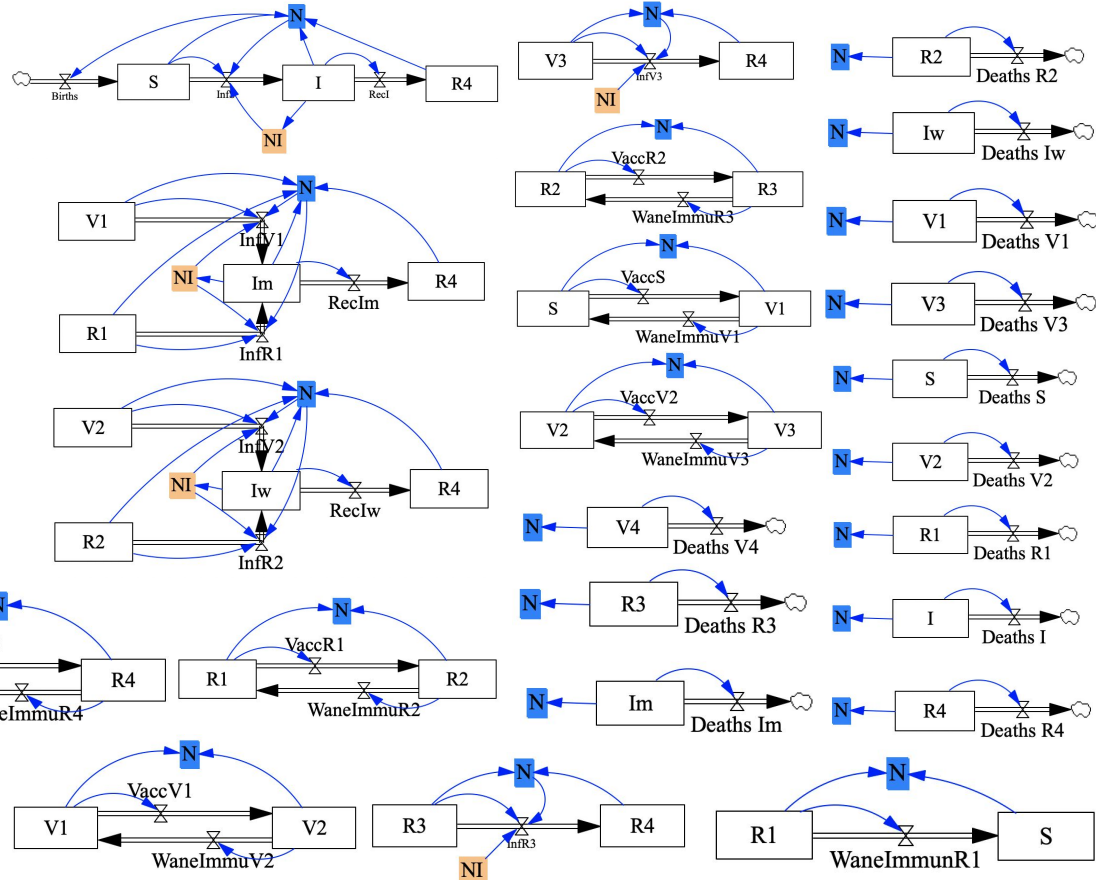
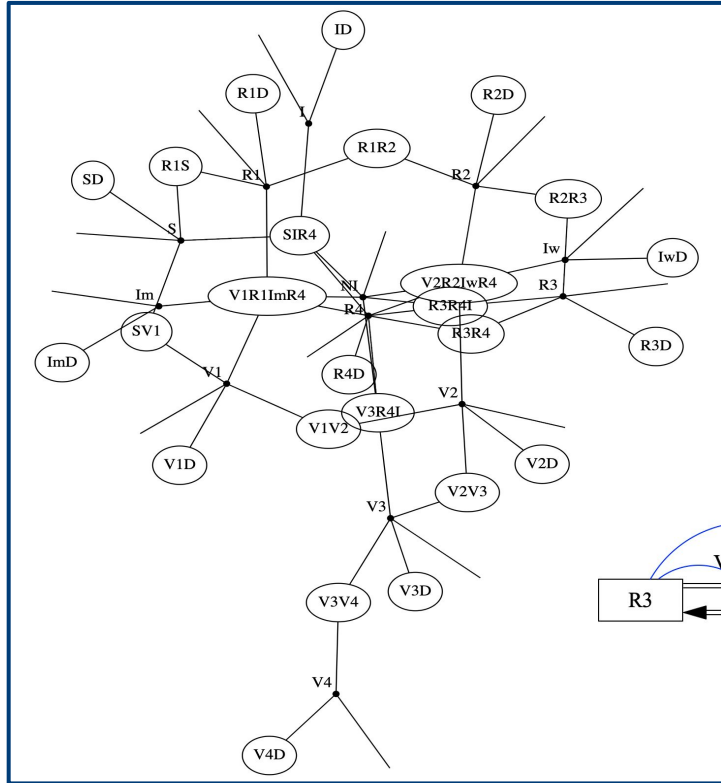
$$\dot{H}_{NICU} = \frac{H_{ICU}}{t_{ICU}} + \frac{f_H(1-f_{ICU})I_{YU}}{t_d} - \frac{H_{NICU}}{t_H} - r_{DN} H_{NICU}$$

## Covid-19 Model Example



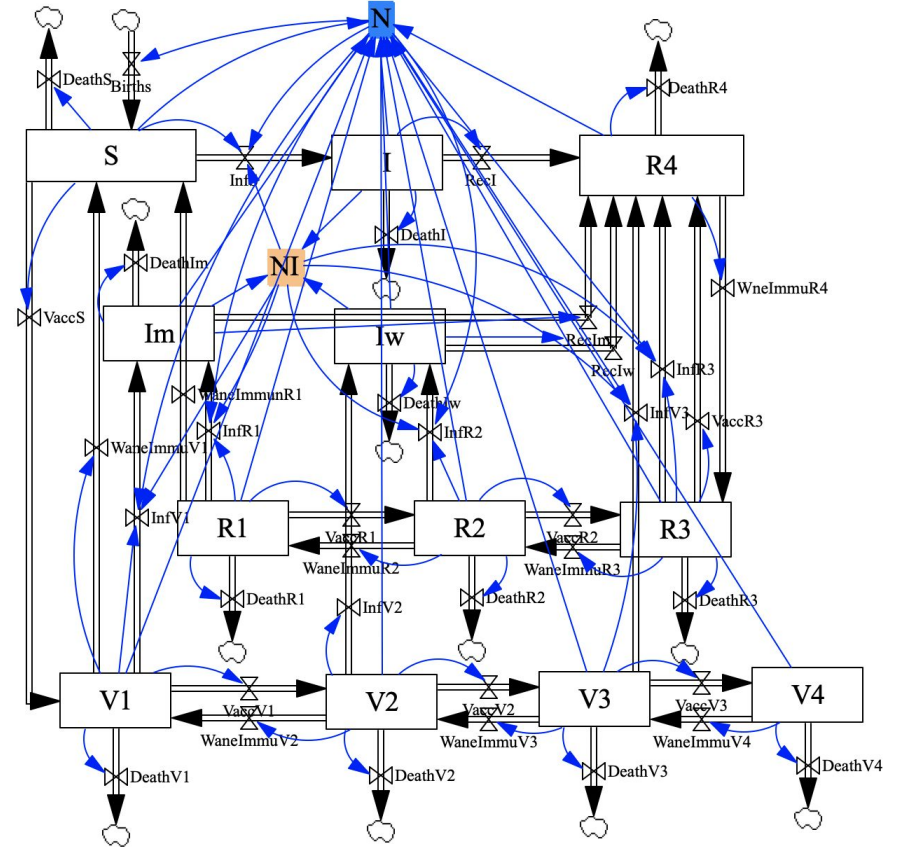
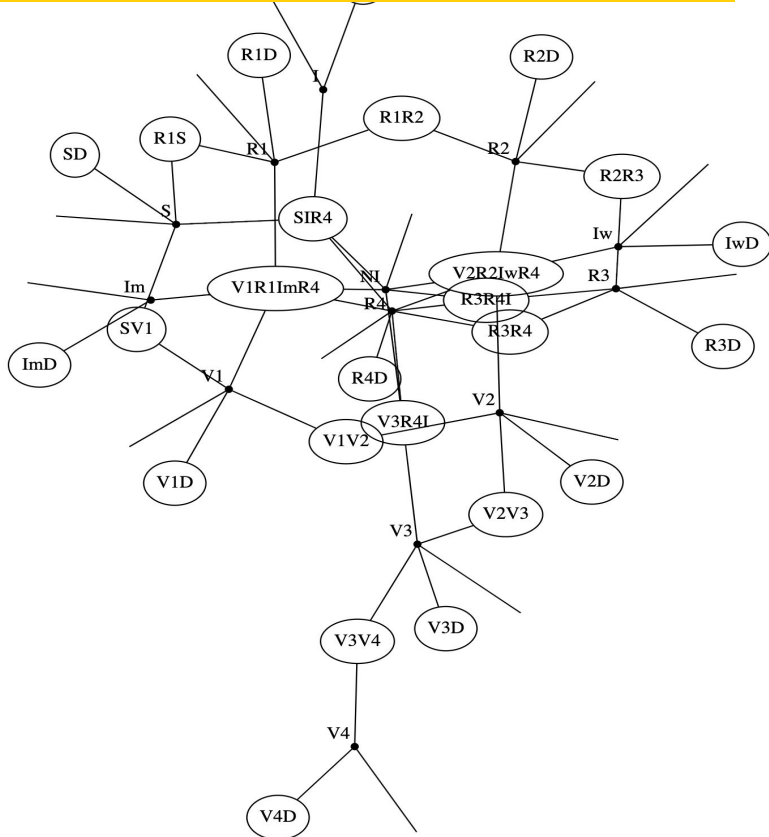
# 2. A Modular and Compositional Framework for System Dynamics Modeling

## Operadic Algebra:



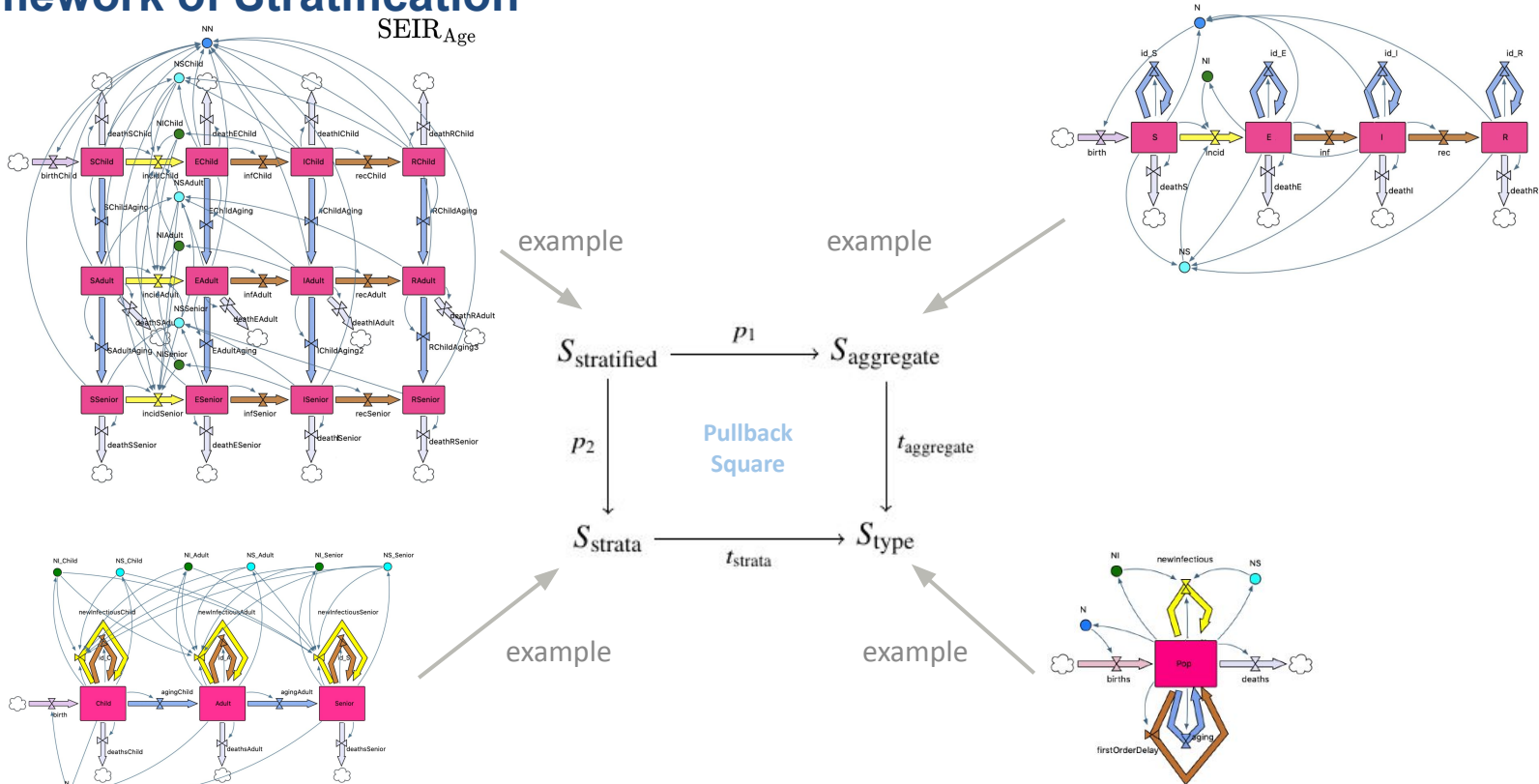
# 2. A Modular and Compositional Framework for System Dynamics Modeling

## Composed Pertussis Model:

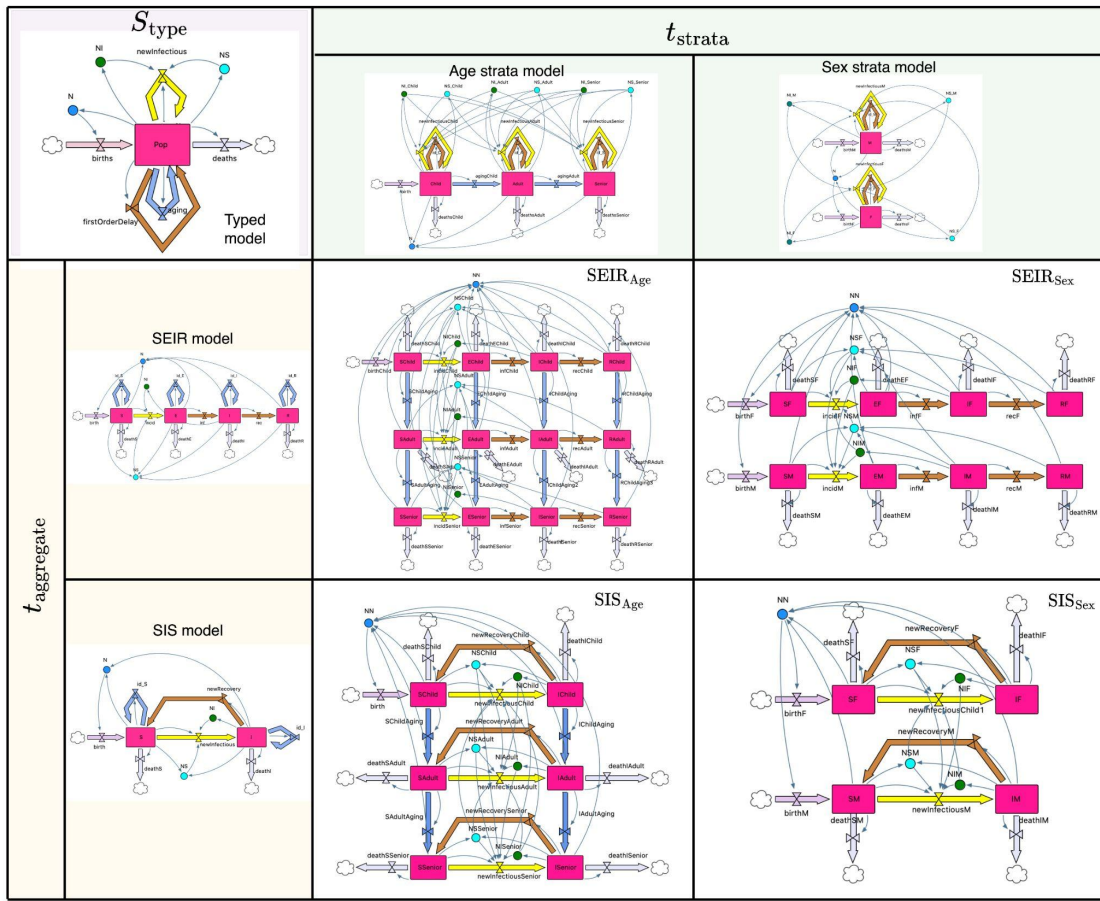


# 2. A Modular and Compositional Framework for System Dynamics Modeling

## A Framework of Stratification



# 2. A Modular and Compositional Framework for System Dynamics Modeling

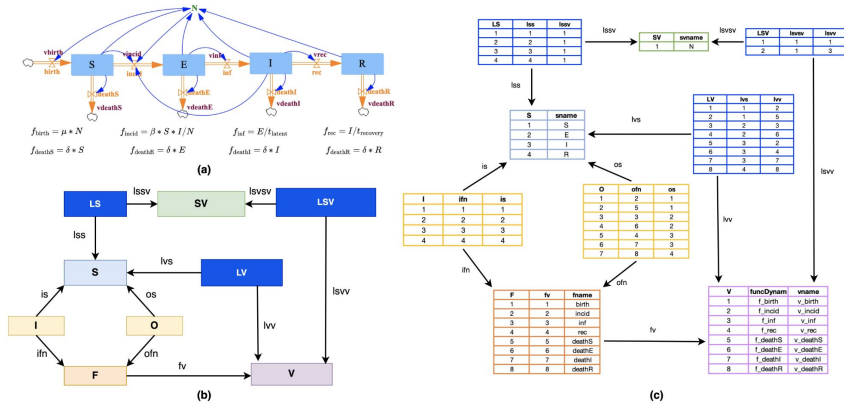


Examples of constructing stratified models

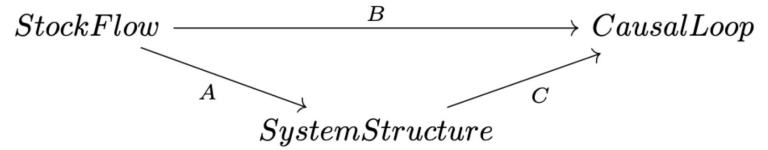
# 2. A Modular and Compositional Framework for System Dynamics Modeling

## System Dynamics Modeling

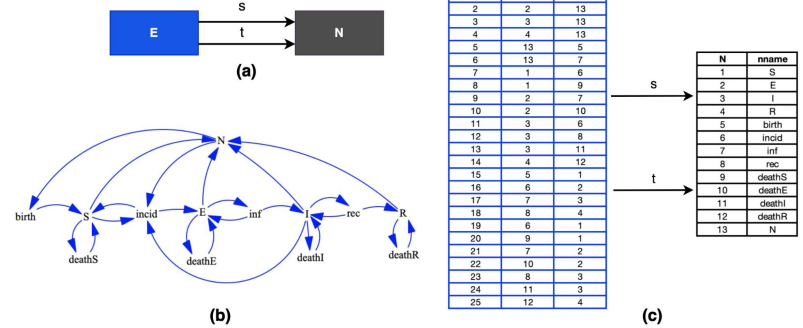
- Stock and Flow diagrams
- System Structure Diagrams
- Causal Loop Diagrams



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!