INTERPRETABILITY OF EPIDEMIOLOGICAL MODELS: THE CURSE OF NON-IDENTIFIABILITY

Ayush Deva, Siddhant Shingi, Avtansh Tiwari, Nayana Bannur, Sansiddh Jain, Jerome White, Alpan Raval, Srujana Merugu
Wadhwani Institute for Artificial Intelligence

ICLR

Motivation

Identifiability is a measure of how precisely we can learn and infer the true parameters of a model. Lack of identifiability in epidemiological model leads to high variance in parameter estimation and reduced accuracy in long term forecasts as show in the figure.

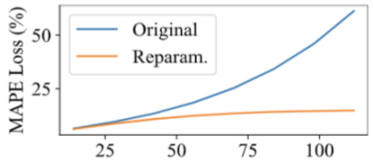
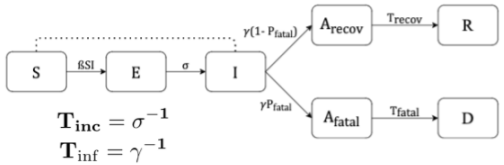


Fig 1. Forecast horizon (days)

Dynamical Systems and SEIARD Model

The proposed notions apply to any dynamical system characterized by $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{\theta})$ where $\boldsymbol{x}$ denotes the system state and $\boldsymbol{\theta}$ the parameters. The observation at time t can be written as $y(t) = h(\boldsymbol{x}_0, \theta, t)$, where $\boldsymbol{x}_0$ is the initial state. SEIR-like models are an example, such as the one shown below (SEIARD).



$\mathbf{T_{inc}} = \sigma^{-1}$
$\mathbf{T_{inf}} = \gamma^{-1}$

Profile Likelihood (PL)

We consider a parameter-wise loss function, called the profile-likelihood(PL) of θ_i, defined as:

$$\mathscr{L}_D^i(\theta_i) := \min_{\boldsymbol{\theta}_{-i}} \mathscr{L}_D(\theta_i, \boldsymbol{\theta}_{-i}) \quad \text{here } \boldsymbol{\theta}_{-i} = \boldsymbol{\theta} \setminus \theta_i$$

Notions of Identifiability

Based on what it depends on (as shown in the table below), we propose and differentiate three notions of identifiability:

1. **Structural Identifiability :** A model parameter is structurally identifiable if no two distinct values lead to the same output.
2. **Statistical Identifiability** : A model parameter is statistically identifiable if its profile likelihood is strictly convex.
3. **Practical Identifiability :** Practical identifiability is defined in terms of tightness of the confidence interval of the estimated parameters. Unlike others, it isn't a binary notion.

	Structural	Statistical	Practical
Model Form	✓	✓	✓
Loss Function		✓	✓
Observation Interval		✓	✓
Noisy Data			✓
Fitting Method			✓

Experimental Results on SEIARD

- The PL curve for β (Fig 2: orange curve) is not convex and has a wide confidence interval (Fig 2 & 4 : orange)
- Correlation matrix between parameters (Fig 3) shows strong correlation between some parameters. To decouple them, we fix T_{inf}, T_{inc}, T_{fatal} (Reparam. model). Shown in blue - the PL curve for β is now convex and has tighter confidence bound - depicting increase in identifiability.
- On increasing the data used for training, we see a qualitative increase in convexity of PL curves (Fig 5).

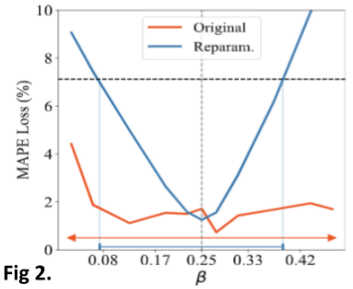


Fig 2.

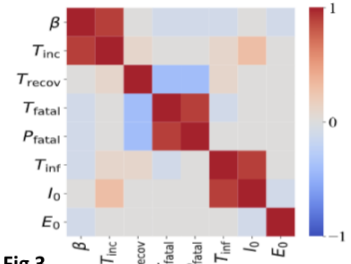


Fig 3.

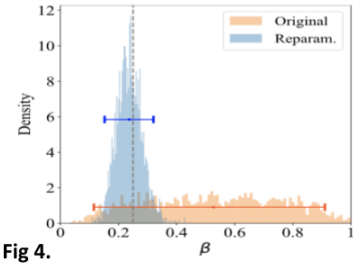


Fig 4.

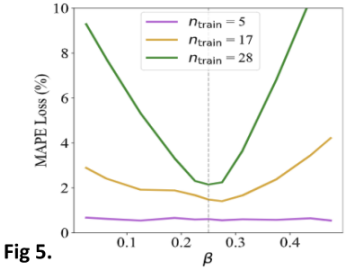


Fig 5.

References

1. Andreas Raue, Clemens Kreutz, Thomas Maiwald, Julie Bachmann, Marcel Schilling, Ursula Klingmuller, and Jens Timmer. *Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood.* Bioinformatics, 25(15):1923–1929, 2009
2. Franz-Georg Wieland, Adrian L Hauber, Marcus Rosenblatt, Christian T ̈onsing, and Jens Timmer. *On structural and practical identifiability.*arXiv preprint arXiv:2102.05100, 2021
3. Alejandro F Villaverde. *Observability and structural identifiability of nonlinear biological systems.*Complexity,2019, 2019.

IGIB INSTITUTE OF GENOMICS & INTEGRATIVE BIOLOGY Genomics Knowledge Partner
USAID FROM THE AMERICAN PEOPLE
WADHWANI AI
BILL & MELINDA GATES foundation
fondation BOTNAR