Observable Neural ODEs for Identifiable Causal Forecasting in Continuous Time

Jennifer Wendland[1] Nicolas Freitag[1] **Maik Kschischo**[1]

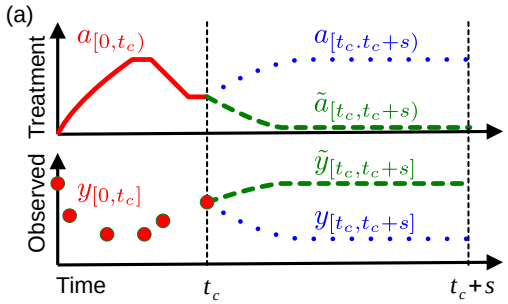
[1]AI in Healthcare,
Institute for Computer Science, University of Koblenz





What would happen to a patient under different treatment options?

The treatment decision requires forecasts under alternative interventions.



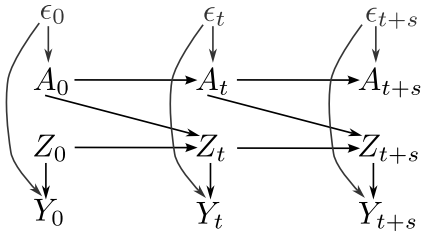
- Decision time point t_c.
- **Given the patient's history up to t_c, how would the outcome differ under treatment plans a or $\tilde{a}$?**

$\boldsymbol{Y}_t$: measured patient condition (lab, vital, images,...) at time t

$\boldsymbol{A}_t$: treatment at time t

Observable Neural ODE (ObsNODE)

- Latent disease state $\boldsymbol{Z}_t$
- Unobserved random disturbances: $\boldsymbol{\epsilon}_t$

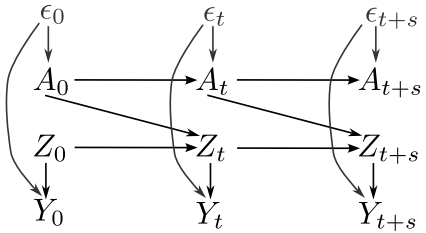


ObsNODE

$$\begin{aligned} \text{initial condition:} \quad & \boldsymbol{Z}_0 \sim p_\theta(\boldsymbol{Z}_0) \\ \text{latent disease state:} \quad & \dot{\boldsymbol{Z}}_t = \boldsymbol{f}_\theta(\boldsymbol{Z}_t, \boldsymbol{A}_t) \\ \text{observations:} \quad & \boldsymbol{Y}_t \sim \boldsymbol{q}_\theta(\boldsymbol{Z}_t, \boldsymbol{\epsilon_t}) \end{aligned}$$

Observable Neural ODE (ObsNODE)

- Latent disease state $\boldsymbol{Z}_t$
- Unobserved random disturbances: $\boldsymbol{\epsilon}_t$



ObsNODE

$$\begin{aligned} \text{initial condition:} \quad & \boldsymbol{Z}_0 \sim p_\theta(\boldsymbol{Z}_0) \\ \text{latent disease state:} \quad & \dot{\boldsymbol{Z}}_t = \boldsymbol{f}_\theta(\boldsymbol{Z}_t, \boldsymbol{A}_t) \\ \text{observations:} \quad & \boldsymbol{Y}_t \sim \boldsymbol{q}_\theta(\boldsymbol{Z}_t, \boldsymbol{\epsilon}_t) \end{aligned}$$

Observability condition:

- Given the treatment input, the latent state can be reconstructed from the observation trajectory over a suitable interval.
- Observability depends on the structure of the ODE and the output map.

Causal identifiability and observability

Assumption

1. Causal dependency structure of the latent state space model (previous graph) including temporally independent disturbances
2. Consistency and overlap.

Theorem

The conditional potential outcome distribution (CPOD) is identifiable, if the latent state space model is observable. Adjustment formula:

$$P(\boldsymbol{Y}_{t+s}(\boldsymbol{a}_{[t,t+s)})|\mathcal{H}_{[0,t]}) = \int_{\boldsymbol{Z}_{t+s}} \int_{\boldsymbol{Z}_t} \underbrace{p(\boldsymbol{z}_t|\mathcal{H}_{[0,t]})}_{\text{state estimate}} \underbrace{p(\boldsymbol{z}_{t+s}|\boldsymbol{a}_{[t,t+s)}, \boldsymbol{z}_t)}_{\text{propagate state under treatment}} \underbrace{p(\boldsymbol{y}_{t+s}|\boldsymbol{z}_{t+s})}_{\text{map state to outcome}}$$

$$\mathcal{H}_{[0,t]} = \{\textit{current observation, previous observations and treatments}\}$$

This remains valid in the presence of the hidden disturbance $\boldsymbol{\epsilon_t}$ and for dynamic treatment regimes.

Observable Neural ODE (ObsNODE)

Observability by construction

- To ensure **observability**, we can restrict the Neural ODE

$$\dot{\boldsymbol{Z}}_t = \boldsymbol{f}_\theta(\boldsymbol{Z}_t, \boldsymbol{A}_t)$$
$$\boldsymbol{Y}_t = \boldsymbol{h}_\theta(\boldsymbol{Z}_t) \qquad \text{(noiseless output)}$$

  to an **observable normal form** (Bernard, 2019).
- This special architecture guarantees observability by design .

Selected test results

Semi-synthetic data: Known confounding between treatment and outcome using MIMIC-IV as background.

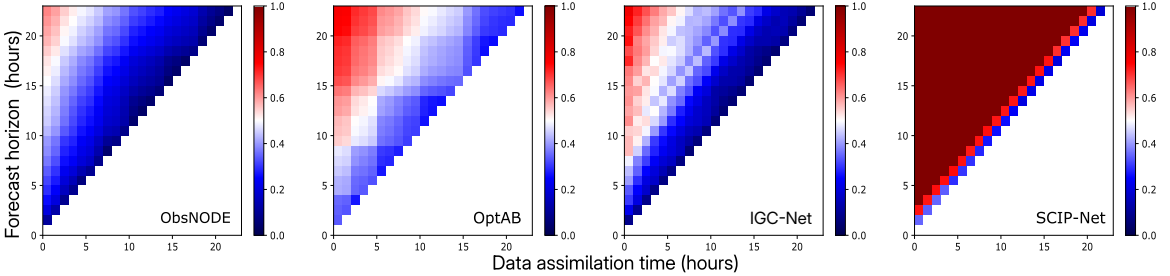


Figure: Comparison of RMSE (lower is better) to models from Wendland et al. 2024, Hess et al. 2026 and 2025 for different t_c and forecast horizons.

Observability provides a sufficient condition for causal identification under our structural assumptions, and ObsNODE enforces it by construction.