

**Madden, Jin, Lopman, Zufle,
Dalziel, Metcalf, Grenfell & Lau
(2024).**

Apr 17, 2025

Impact

- Too early to tell - published Nov 21, 2024.
- An advance on Lau, Becker, Madden, Waller, Metcalf, Grenfell (2022), “Comparing and linking machine learning and semi-mechanistic models for the predictability of endemic measles dynamics”, [PLoS Comput Biol 18\(9\): e1010251](#).

TSIR as a mechanistic model

- The “time series SIR” approach of [Bjornstad and Grenfell \(2002\)](#) makes a log-linear approximation to disease dynamics conditional on a reconstructed susceptible population.
- This results in a linear regression model, where seasonality or other factors affecting transmission rate can be included as covariates.
- TSIR is very convenient when the assumptions are reasonable. For measles they are quite good.
- TSIR assumes the infection duration and observation interval are the same.

Forecast evaluation

“Our results show that while the TSIR model yields similarly performant short-term (1 to 2 biweeks ahead) forecasts for highly populous cities, our neural network model (SFNN) consistently achieves lower root mean squared error (RMSE) across other forecasting windows.”

- What does that imply for model fit and model specification (for both SFNN and TSIR models)?

SHAP values for XAI

- SHapley Additive exPlanation (SHAP). Or is SHAP just short for Shapley?
- SHAP for a covariate value is the difference between the prediction with and without that covariate.
- That ideally involves integrating out all other covariates, assuming there is a stochastic model for them.
- [Lloyd Shapley](#) developed the Shapley value of a collaborative game as a fair way to distribute gain or cost among players.
- Covariates in a model “collaborate” in the “game” of predicting the next outcome.

SHAP in regression/deep learning

- Interpreted as the difference in
- Incidentally, do you think [this Wikipedia paragraph](#) involved GenAI:

“Shapley value contributions are recognized for their balance of stability and discriminating power, which make them suitable for accurately measuring the importance of service attributes in market research.[17] Several studies have applied Shapley value regression to key drivers analysis in marketing research. Pokryshevskaya and Antipov (2012) utilized this method to analyze online customers’ repeat purchase intentions, demonstrating its effectiveness in understanding consumer behavior.[18] Similarly, Antipov and Pokryshevskaya (2014) applied Shapley value regression to explain differences in recommendation rates for hotels in South Cyprus, highlighting its utility in the hospitality industry.[19] Further validation of the benefits of Shapley value in key-driver analysis is provided by Vriens, Vidden, and Bosch (2021), who underscored its advantages in applied marketing analytics.[20]”

More on SHAP

From [the shap Python package](#):

To evaluate an existing model when only a subset of features are part of the model we integrate out the other features using a conditional expected value formulation. This formulation can take two forms:

$$E[f(X)|X_S = x_S]$$

or

$$E[f(X)|\text{do}(X_S = x_S)].$$

In the first form we know the values of the features in S because we observe them. In the second form we know the values of the features in S because we set them. In general, the second form is usually preferable, both because it tells us how the model would behave if we were to intervene and change its inputs, and also because it is much easier to compute.

The Physics informed neural network (PINN)

- Train a NN with a loss function that combines fitting data and fitting the model.
- This is reminiscent of the likelihood in a POMP model (e.g., Eq. 4,5,6)
- Are we fitting a non-parametric POMP model? A “penalty for how well the solution matches an ODE” is very similar to supposing an SDE.