**SIAM 2019 Abstract**

*A dynamical systems approach to transforming disparate timescales in data driven equation-free modeling of disease dynamics*

Mollie M. Van Gordon and Joshua L. Proctor

Effective prediction of disease dynamics often requires knowledge of both endogenous and exogenous factors. Incorporating diverse types of information, however, introduces covariates at different observational timescales. Despite advances in time dependency and discretization in data driven equation-free modeling, dissimilarity in time intervals has remained a challenge. We present a method that addresses this limitation, judiciously unifying timescales of input data. To demonstrate, we develop a malaria transmission model. Relevant remote sensing data includes weekly land cover data alongside sub-daily precipitation and temperature. Plasmodium incidence is reported at a weekly to monthly frequency. Our approach rigorously connects these timescales using principles from non-autonomous dynamical systems and a recently developed method for sparse identification of nonlinear dynamics (SINDy). We employ this method in multiple layers: identifying the disease transmission model as well as optimizing the timescale transformation itself. This layered SINDy approach identifies the most appropriate functional form for converting inputs to a unified timescale. Incorporation of environmental and epidemiological data initially at different timescales expands the options for feature selection and increases the model’s predictive power. Applications of this approach are broadly relevant to disease modeling as well as any data driven equation-free modeling incorporating diverse input data.

*[too long]*

Effective prediction of disease dynamics often requires knowledge of both endogenous and exogenous factors. Incorporating diverse types of information, however, introduces covariates with different observational timescales. Despite advances in time dependency and discretization in data driven equation-free modeling, dissimilarity in covariate time intervals has remained a challenge. We present here a method that addresses this limitation, judiciously unifying timescales of input data for modeling diseases such as malaria. To demonstrate our method, we develop a model of malaria transmission with environmental and epidemiological variables at different time scales. Relevant remote sensing data includes weekly land cover data alongside sub-daily precipitation, temperature, and humidity. Plasmodium incidence is reported by health facilities at a weekly to monthly frequency. Our approach rigorously connects the timescales of these inputs using principles from non-autonomous dynamical systems and a recently developed method for working with time series data called sparse identification of nonlinear dynamics (SINDy). We employ this method in multiple layers: identifying the disease transmission model as well as optimizing the timescale transformation itself. This layered SINDy approach identifies the most appropriate functional form for converting inputs to a unified timescale. Incorporation of environmental and epidemiological data initially at different timescales expands the options for feature selection and increases the model’s predictive power. Applications of this approach are broadly relevant to disease modeling as well as any data driven equation-free modeling incorporating diverse input data.