**Uncertainty in Statistical and Machine Learning Models**

Traditional statistical models quantify one component of statistical uncertainty in confidence intervals and p-values. These models output a maximum likelihood estimate (MLE) for parameters, surrounded by a confidence interval denoting the spread of uncertainty coming from random chance in sample composition.

Machine learning techniques offer greater predictive power than traditional statistical methods, especially in complex scenarios commonly found in disease forecasting. However, most do not provide a confidence interval for their predictions in the same manner as traditional statistical models. This is a barrier to an actionable model forecast.

It is possible to create estimates of uncertainty consistently across machine learning and traditional statistical models using bootstrap resampling. Bootstrap resampling methods are based on the assumption that a particular sample dataset is representative of a larger population, and runs a particular model many times on repeated subsamples of a dataset. Thousands of iterations of model output are then taken to represent the spread of possible values of a model’s parameters, and used to create confidence intervals. More advanced versions of the bootstrap may apply random noise to a dataset to account for a paucity of data, or may weight the subsamples to account for known sampling bias, or uncertainty from other sources.

**Related Past Work**

The Hotspots II model (figure on Quad Chart), developed by EcoHealth Alliance, uses a boosted regression tree (BRT) model to predict zoonotic disease emergence, based on environmental drivers. A weighted bootstrap resampling regime was used to correct for bias from unequal observation, and to compute 95% confidence estimates incorporating sampling uncertainty and uncertainty in the locations of modeled events.

**An API for Estimating Uncertainty Across Existing Models**

Bootstrap resampling can be applied across model types, which makes it a prime candidate for application to pre-existing disease forecast models. Other relevant techniques include sensitivity analysis, which examines the effect of varying input parameters on model output, and cross-validation, which calculates internal validation statistics on models fit to partitions of datasets.

Our framework (see diagram below) will ingest existing datasets and models in a variety of formats, with API methods for different model types and data formats, and use this to output model predictions in a standardized format. At this point, model predictions of risk and uncertainty can be aggregated or displayed separately for comparison. Our API end-points will allow the framework to be incorporated into existing analytic toolchains, such as the BSVE.

**A GUI for Exploring Model Output**

In addition to the API, we will develop a GUI to allow analysts to interact with model parameters and explore predictions and uncertainty. Building on top of the framework, we will create interface elements to display model predictions individually, side-by-side, and aggregated. Model output will be viewable on a map, in various graphs, and in tabular format, and will be exportable in commonly-used formats.

We will also use the GUI to expose the various parameters used to tune the bootstrap resampling process and other implemented methods.

