Surrogate Modeling with Machine Learning for Faster Virtual Patient Cohort Generation in QSP Models

Renee Myers[1]*, Christina Friedrich[1]*, Tongli Zhang[2] , Michael Weis[1]. 1. Rosa & Co., San Carlos, CA 94070, USA 2. University of Cincinnati

Problem Statement

Introduction

- Virtual patients (VPs) are used within QSP modeling to explore the impact of variability and uncertainty on clinical response
- Combinations of parameter values that produce physiological solutions, i.e., valid VPs, are difficult to determine a priori, because of complex interdependencies in QSP models
- Sampling parameters, simulating protocols of interest, and then filtering out any non-physiological solutions is a common approach
  - Number of plausible VPs may be only a fraction of the generated VPs (<5%)
  - Requires time-consuming simulations to build a cohort of VPs that fully characterizes variability
- Machine learning (ML) surrogate models are a promising approach for improving efficiency of VP generation
  - Infers valid parameter combinations from preliminary QSP model simulations and "pre-filters" parameter sets based on **predicted** response values

Objectives

- Use machine learning surrogate models to improve efficiency of VP cohort generation in a psoriasis QSP model

Conclusion

Use of surrogate models significantly increases efficiency over random parameter sampling for generation of VP cohorts with no loss of diversity.

Design

- Use existing psoriasis QSP model constructed in MATLAB®SimBiology®
  - 11 constrained model species, 5 varied parameters
- Use Gaussian Process Regression (GPR) as regression function to generate surrogate models
  - Evaluated several ML models using MATLAB® Regression Learner app; compared Root Mean Squared Error (RMSE) for each model type
- Train surrogate models with data (10,000 simulations) at a 70/30 partitioned split
  - Use QSP model to generate data for the regression function

Methods

Surrogate Modeling

- Surrogate model: Statistical model that approximates a complex, higher order model
  - Surrogate model is trained using input/output data obtained through simulation of the original model
- Surrogate models can be used to "replace" a mechanistic model that might be slow to simulate, yielding very fast predictions
- Surrogate models are a smarter approach to VP cohort generation
  - Simulation of QSP models: Computationally expensive; potential for many parameter sets to be rejected after filtering
  - Surrogate modeling approach: Instantaneously predict model output for a parameter set, accept/reject parameter sets based on filtering criteria

Approaches to VP Cohort Generation

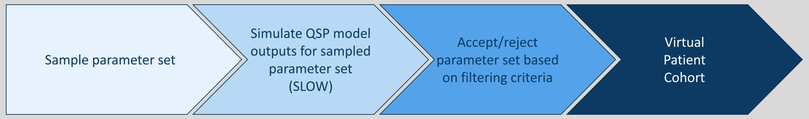

Figure 1: Original workflow for generating virtual patient cohort

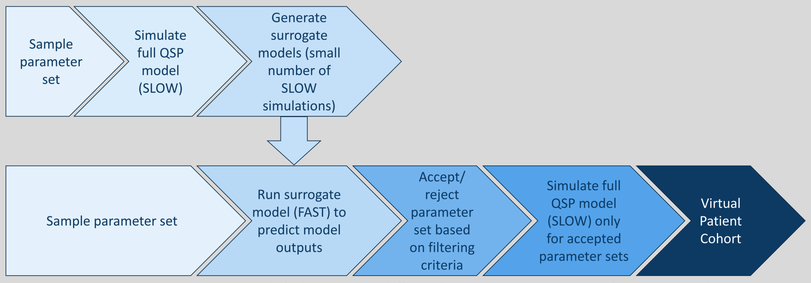

Figure 2: Surrogate model-driven workflow for generating virtual patient cohort

Efficiency of Method

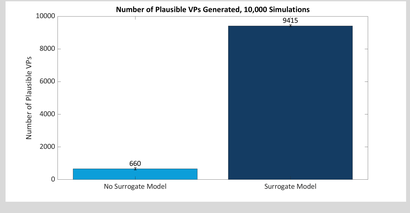

Figure 3: Number of plausible VPs generated by both the original and surrogate modeling approaches. Error bars represent standard deviation from 5 trials of 10,000 simulations each.

- **Result:** Surrogate modeling approach generated **14x** as many plausible VPs as the original approach for the same amount of computational time once the surrogate models were constructed

Results

Distribution of Model Observables

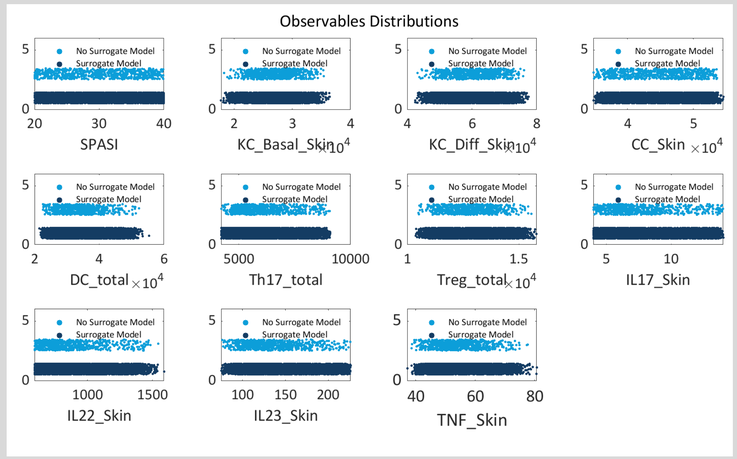

Figure 4: Distribution of model observables after filtering at steady state (60 weeks).

Distribution of Sampled Parameters

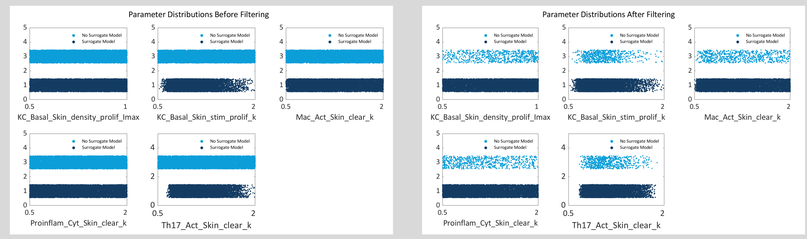

Figure 5: Distribution of sampled parameters before filtering (left) and after filtering (right) for both the original and surrogate modeling approaches.

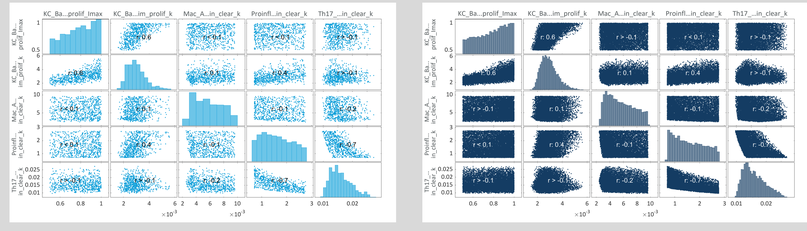

Figure 6: Correlation matrix for sampled parameters after final filtering step for (left) virtual cohort created using original workflow and (right) parameters sampled using surrogate modeling approach.

- **Results:** Model outcomes and parameters were similarly distributed in both methods. Relationships between parameters were also preserved.
- **Significance:** The use of surrogate models does not diminish sampling or outcome variability, and recapitulates relationships between parameters consistent with the full QSP model.

For more information about this work, please contact:
Christina Friedrich
cfriedrich@rosaandco.com
Rosa & Co LLC