A Semi-Mechanistic Longitudinal Gaussian Process Regression Model for Estimating Tumor Response to Treatment

Jacqueline Buros, Juho Timonen, Eric Novik

KEYWORDS — Tumor-growth modeling, Gaussian Processes, Semi-Mechanistic, Bayesian, Stan

I. BACKGROUND

A wide variety of semi-mechanistic, mathematical models have been developed to describe tumor growth over time under treatment, and they all share a similar feature: the best model is often specific to the tumor-intrinsic factors driving growth (such as VEGF), and/or to the treatment's mechanism of action (such as immune checkpoint inhibitors).
The use of a treatment- or tumor-specific model has two important limitations:

1. Limits our ability to compare treatment efficacy across subpopulations and treatments
2. Limits our ability to reliably estimate responses to new therapies, whose mechanisms may not yet be well understood

It follows that a goal of tumor-growth modeling efforts should be develop a modeling framework that is both sufficiently *general* as to enable comparisons across treatments and/or sub-populations, and is sufficiently *flexible* to capture treatment response patterns across a variety of treatments and tumor types.
The work presented here is our attempt to satisfy this goal.
We present **`sfgp`**, an open-source software package in R that implements a hierarchical Bayesian semi-mechanistic model with GP-augmented parameters in Stan. We show how this approach is both flexible and general, with good performance for estimating tumor-size responses to treatment across a variety of treatments and indications.

II. USAGE

Installation is easy, from the github repo:

```
remotes::install_github('generable/sfgp')
library(sfgp)
```

Each longitudinal model is an instance of `TSModel` class. The first step is to define the model. We then prepare the data, fit the model, and summarize the results.
In code, this looks like the following:

```
dat <- create_dummy_data()
m1 <- TSModel$new(y ~sf(t, id | arm)) # define model
m1dat <- add_sff_input(dat, m1)       # prepare data
f1 <- m1$fit(m1dat, chains = 1)       # fit model
f1$plot()                             # summarize inferences

# predict at test points
t_pred <- seq(0, 9, by = 0.2)
pred <- f1$predict_time(t_pred, t_var = "t")
pred$plot(plot_y = FALSE) +
  ggplot2::geom_point(data = dat, inherit.aes = F, aes(x = t, y = y))
```

III. IMPLEMENTATION

The likelihood of observation i is

$$\log(y_i + \delta) \sim \mathcal{N}(f_i, \sigma^2), \tag{1}$$

where

$$f_i = f(\mathbf{x}_i) = \sum_{j=1}^{J} f^{(j)}(\mathbf{x}_i) \tag{2}$$

is the expected log tumor size, σ is an unknown parameter, and the δ value is a constant. We use y_i to denote the observed longitudinal measurement, and $\mathbf{x}_i$ the corresponding covariate vector. Typically, $\mathbf{x}_i$ includes a subject identifier id_i, measurement time t_i, treatment arm arm_i and possible other factors or continuous covariates.
The functions $f^{(j)}$, $j = 1, ..., J$ are the additive function components that, together, define the model.
Here are some of the terms implemented:

Term name	Formula	Description	Typical usage
`GroupedOffsetTerm`	`offset(g)`	$f^{\mathrm{BAS}}(\mathrm{g} \mid \mathbf{c}_0)$ where $\mathbf{c}_0$ is an unknown parameter vector with independent priors.	Baseline value per subject (here, g)
`HierOffsetTerm`	`offset(g | h)`	$f^{\mathrm{BAS}}(\mathrm{g}, \mathrm{h} \mid \mathbf{c}_0)$ where $\mathbf{c}_0$ is an unknown parameter vector with hierarchical priors	tumor-growth parameter estimated per subject (here, g), hierarchically per arm (here, h)
Shared `GPTerm`	`gp(x)`	$f^{\mathrm{GP}}(x \mid \alpha, \ell) \sim \mathcal{GP}\big(0, k_{\mathrm{EQ}}(x, x' \mid \alpha, \ell)\big)$	Population-level response over time (here, x).
Group-specific `GPTerm`	`gp(x, z)`	$f^{\mathrm{HSGP}}(x, z \mid \boldsymbol{\xi}_1, ..., \boldsymbol{\xi}_G, \alpha, \ell, B, L) = \begin{cases} f^{\mathrm{HSGP}(x \mid \boldsymbol{\xi}_1, \alpha, \ell, B, L)} \text{ if } z{=}1 \\ \vdots \\ f^{\mathrm{HSGP}(x \mid \boldsymbol{\xi}_G, \alpha, \ell, B, L)} \text{ if } z{=}G \end{cases}$	Group-specific response over time (here, x)
Main `SFTerm`	`sf(x)`	$f^{\mathrm{SF}}\big(x \mid k_g, k_s\big) = \exp\big(k_g x\big) + \exp\big(-k_s x\big) - 1$	Tumor-size response over time (here, x) using Stein-Fojo model
convenience `SFTerm`	`sf(x, id | h)`	$f^{\mathrm{SF}}_{\mathrm{id}}\big(x \mid k_{g,\,\mathrm{id}}, k_{s,\,\mathrm{id}}\big) = \exp\big(k_{g,\,\mathrm{id}} x\big) + \exp\big(-k_{s,\,\mathrm{id}} x\big) - 1$	Hierarchical SF model over time (here, x) by id (here, id) and arm (here, h)
`FormulaSFTerm`	`sff(t | kg ~ ..., ks ~ ...)`		Stein-fojo model where k_g and k_s are functions of covariates (see examples)

IV. EXAMPLE MODELS

i. The SF-only model

In the basic SF-only model using the convenience `SFTerm` we have $\mathbf{x} = \{t, \mathrm{id}, \mathrm{arm}\}$, $J = 2$ and

```
mod1 <- TSModel$new(y ~ sf(t, id | arm))
```

$$\begin{aligned} f^{(1)}(\mathbf{x}) &= f^{\text{log-SF}}\big(\mathrm{t} \mid \mathbf{k}_g, \mathbf{k}_s\big) \\ f^{(2)}(\mathbf{x}) &= f^{\mathrm{BAS}}(\mathrm{id} \mid \mathbf{c}_0) \end{aligned} \tag{3}$$

Notice that in this case, we have

$$\exp(f(\mathbf{x})) = \exp\big(\log \mathrm{c}_{0,\,\mathrm{id}} + \log f^{\mathrm{SF}}\big(t \mid k_{g,\,\mathrm{id}}, \mid k_{s,\,\mathrm{id}}\big)\big) = \mathrm{c}_{0,\,\mathrm{id}}\big(\exp\big(k_{g,\,\mathrm{id}} \cdot t\big) + \exp\big(-k_{s,\,\mathrm{id}} \cdot t\big) - 1\big) \tag{4}$$

which is the original formulation of the SF tumor size model.

ii. An SF+GP model

An example of an SF+GP model with $\mathbf{x} = \{t, \mathrm{id}, \mathrm{arm}\}$, $J = 4$ is

```
mod2 <- TSModel$new(y ~ sf(t, id | arm) + gp(t) + gp(t, arm))
```

It has the terms

$$\begin{aligned} f^{(1)}(\mathbf{x}) &= f^{\text{log-SF}}\big(\mathrm{t} \mid \mathbf{k}_g, \mathbf{k}_s\big) \\ f^{(2)}(\mathbf{x}) &= f^{\mathrm{HSGP}}(\mathrm{t} \mid \boldsymbol{\xi}_{\mathrm{t}}, \alpha_{\mathrm{t}}, \ell_{\mathrm{t}}, B_{\mathrm{t}}, L_{\mathrm{t}}) \\ f^{(3)}(\mathbf{x}) &= f^{\mathrm{HSGP}}\Big(\mathrm{t}, \mathrm{arm} \mid \boldsymbol{\xi}^{(1)}_{\mathrm{t}\times\,\mathrm{arm}}, ..., \boldsymbol{\xi}^{(G_{arm})}_{\mathrm{t}\times\,\mathrm{arm}}, \alpha_{\mathrm{t}\times\,\mathrm{arm}}, \ell_{\mathrm{t}\times\,\mathrm{arm}}, B_{\mathrm{t}\times\,\mathrm{arm}}, L_{\mathrm{t}\times\,\mathrm{arm}}\Big) \\ f^{(4)}(\mathbf{x}) &= f^{\mathrm{BAS}}(\mathrm{id} \mid \mathbf{c}_0) \end{aligned} \tag{5}$$

where G_{arm} is the number of treatment arms.
In this model, we utilize the Stein-Fojo function for the mean response and model the noise using a gaussian process term overall and separately per arm.

V. EXPERIMENTAL SETUP

We use two datasets to evaluate our approach:

1. A simulated dataset comprised of longitudinal tumor-size data for subjects with metastatic melanoma treated with an aPD-L1 checkpoint inhibitor. The details of the simulation have been described previously (R. Dutta, A. Mohan, J. Buros-Novik, G. Goldmacher, O. O. Akala, and B. Topp [1]). A hypothetical early-phase trial is reconstructed by sampling 30 subjects from each of two arms.
2. A sample of subjects with metastatic prostate cancer treated with docetaxel (J. Wilkerson *et al.* [2]). A hypothetical early-phase trial is reconstructed by selecting samples of 30 subjects from two active comparator arms.

For each dataset, we will perform the following as validation:

- Fit the model(s) to the sampled datasets, reserving hold-out subjects from each arm.
- Evaluate model performance using approximate leave-one-observation-out cross-validation (LOOO-CV) (A. Vehtari, A. Gelman, and J. Gabry [3]).
- Summarize distribution of predicted responses for new subjects in each arm
- Compare predictive distribution for new subjects to observed trajectories for held-out subjects.

1) *Sampled subjects:*

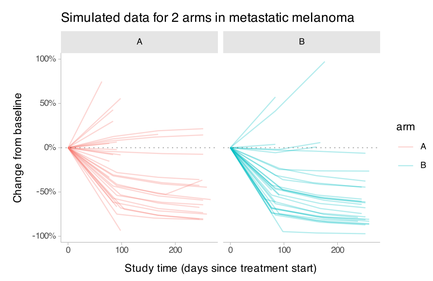


Figure 1

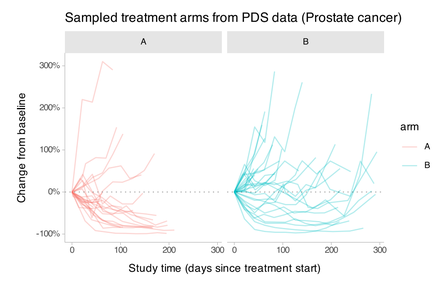


Figure 2

VI. RESULTS

i. Predictive performance

Table 1: Summary of predictive performance using LOO-PSIS

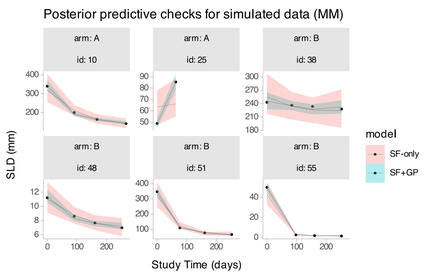
(a) Simulated data

Predictive performance for simulated data (MM)			
model	elpd_loo	elpd_diff	se_diff
SF+GP	230.66	0.00	0.00
SF-only	16.16	−214.50	15.67

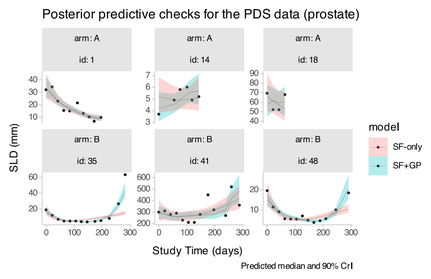
(b) PDS data

Predictive performance for PDS data (prostate)			
model	elpd_loo	elpd_diff	se_diff
SF+GP	−22.59	0.00	0.00
SF-only	−100.33	−77.74	22.29

ii. Predictive checks



(a) Simulated data

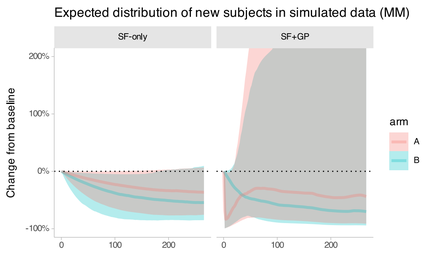


(b) PDS data

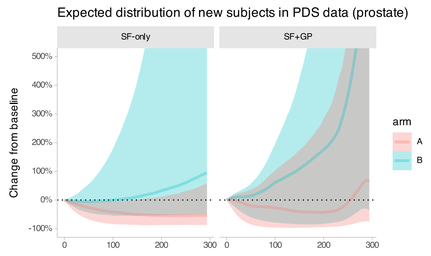
Figure 3: Posterior predictive checks

iii. Expected treatment effects

Here we show summarized treatment effects per arm by simulating new subjects per arm.



(a) Simulated data



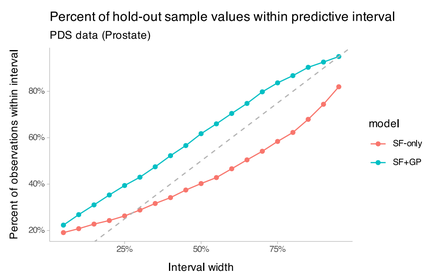
(b) PDS data

Figure 4: Summarized treatment effects

iv. Treatment effect performance

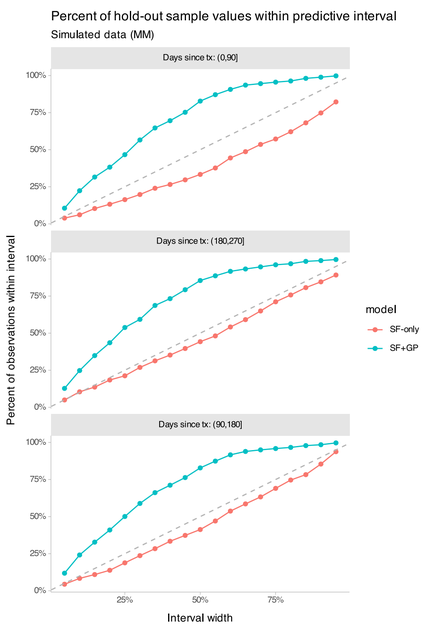
In Figure 5, we summarize the calibration of the simulated treatment effects vs observed data from the hold-out sample. Figure 6 summarizes calibration separately for early and late observations.

(a) Simulated data

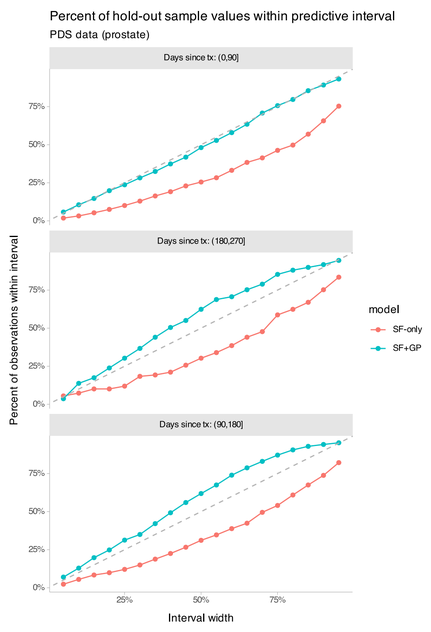


(b) PDS data

Figure 5: Predictive performance for simulated subjects



(a) Simulated data



(b) PDS data

Figure 6: Predictive performance for simulated subjects by time

BIBLIOGRAPHY

[1] R. Dutta, A. Mohan, J. Buros-Novik, G. Goldmacher, O. O. Akala, and B. Topp, "A bootstrapping method to optimize go/no-go decisions from single-arm, signal-finding studies in oncology," *CPT Pharmacometrics Syst. Pharmacol.*, vol. 13, no. 8, pp. 1317–1326, Aug. 2024.

[2] J. Wilkerson *et al.*, "Estimation of tumour regression and growth rates during treatment in patients with advanced prostate cancer: a retrospective analysis," *Lancet Oncol.*, vol. 18, no. 1, pp. 143–154, Jan. 2017.

[3] A. Vehtari, A. Gelman, and J. Gabry, "Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC," *Stat. Comput.*, vol. 27, no. 5, pp. 1413–1432, Sep. 2017.