Toward optimal drug combination synergy screening.

Haynes Heaton, Emanuel Gonçalves, Patricia Jaaks, Mathew Garnett

Abstract — The use of drug combinations can be beneficial in virtually all disease types by being synergistic in their effect, decreasing the chance for the target to gain resistance, or having non-overlapping side-effect profiles. While the potential benefit of drug combinations is great, the combinatorial nature of such drug combination screenings is costly. Hence it is important to the develop efficient combinatorial drug screens that preclude measuring all possible drug/drug concentration combinations while robustly identifying synergistic associations. To that end, we present a novel data-driven framework for measuring cancer drug synergy which is agnostic to sample layout using gaussian process regression, and for determining a more optimal layout of the sample points in the drug/drug double titration. With more effective sample layouts, fewer data points will be needed to determine drug pair synergy at the same level of confidence. This will allow for the the testing of more drug pairs and will thus find more of these vital synergies needed to effectively treat cancer.

1. Introduction

Synergistic cancer therapeutics are sought to either more effectively treat a cancer or to treat cancer with lower overall side effects[1]. One strategy that is often employed is the use of high throughput screenings of drugs in-vitro with tumor and normal tissue derived cell lines[2][3][4][5]. Generally the structure of this data is a log linear titration of both drugs centered around some expected active dose in either a grid or some subset of that grid. The measured response is a reporter assay for the number of cells to estimate of cell survival. Additionally the drugs are tested alone in a similar dose range. With this data, analysis is then performed to predict synergy versus additivity versus antagonism of the drug pair compared to their expected combined effect calculated with various additive models applied to data of each drug by itself[6][7][8][9][10].

2. Background and Previous Work

2.1. Modelling

The majority of methods developed to measure synergy between two drugs involve first doing a regression on one or both of the combination drug experiment or the expected additive model followed by computing a metric or statistical test between these models. Combenifit[11], CImbinator[12], and CompuSyn[10] all use Maximum Likelihood Estimation (MLE) regression on different parametric sigmoidal surface functions. Whereas Koplev et al.[13] take a global bayesian regression solving for the posterior distribution of parameters of parametric sigmoidal surfaces on all of their data at once. This has the advantage of being capable of providing confidence intervals on the metric calculated based on the underlying uncertainty of the data. All of these rely on the assumption that the form of the function they define can adequately describe the data. In choosing such a model, one must balance the need to minimize the number of free parameters and the ability of that model to adequately describe the data.

2.2. Statistical test or metric evaluation

There have been several metrics and statistical tests designed to determining the confidence of synergism or the degree of synergism in drug combination screens. There is an inherent problem with many if not all of these metrics, including the one we put forward, which is that they cannot generally be compared from drug pair to drug pair. The reason for this is that these metrics rely on the dose ranges chosen for each drug. And as there is no canonical mapping of doses between drugs, the absolute value of these metrics cannot be compared. What can be compared are the metrics for one drug pair on one cell line versus a different cell line. And if the test or metric accurately models the uncertainty, one may be able to determine confi-

dence of synergy. Metrics used have largely been either the integrated volume under the curve between the additive model and the combination experiment model[11][13] or the combination index (CI) of the median effect model[10][12]. There has been no significant comparison of these methods that we can find in the literature.

3. Methods

3.1. Model

Here we put forward a different probabilistic model for characterizing synergistic drug interactions. We then use that model to determine a more optimal sample dose layout. We use Gaussian Process (GP) regression to model both the combination drug experiments as well as the expected additive model. Gaussian Processes[14] are a form of non parametric, probabilistic machine learning. Gaussian Process regressions have few assumptions about the shape of the generating function, have few hyper parameters, and are self regularizing avoiding the overfitting problem. The primary downside of Gaussian Processes is that the regression takes $O(N^3)$ computational cost where N is the number of input samples, but drug combination experiments generally have very few data points. Figures 1 and 2 show examples of these surface fits and their confidence intervals. We believe this creates a more realistic model of both the data and uncertainty than the parametric models used in previous methods.

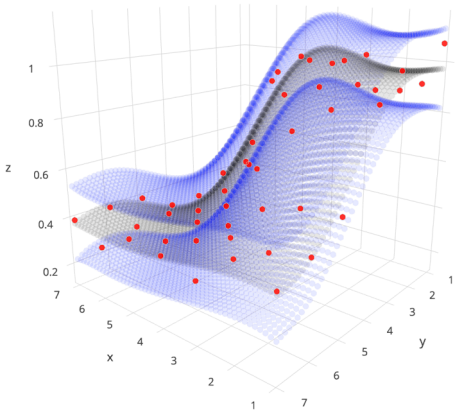

Fig. 1. Mean surface and two standard deviation envelope of GP fit to one combination drug data set.

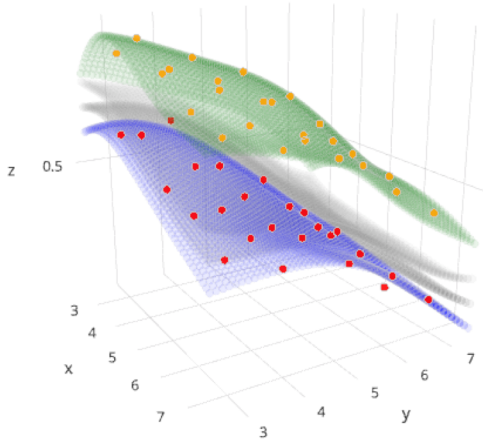

Fig. 2. Mean surface + 1 and 2 sigma envelopes of combination drug experiment vs bliss additive mean surface.

3.2. Statistic

There are many possible metrics we could employ on these models. There has been some previous work in comparing two Gaussian Processes[15] using the log likelihood of the data fit to two separate Gaussian Processes versus a single Gaussian Process. While this is a standard model comparison strategy, in this case it does not take into account the fact that a single model is a simpler solution than two models. It is also not sided which means it will not differentiate between synergy and antagonism. Other metrics one could employ are the Mahalonobis distance or a divergence such as the Kullback Leibler divergence between the two models. While these incorporate the covariance of the data, they are difficult to interpret and are also one sided. We instead use a very simple score taking the differences in means of the functions standardized by each model's variance at that location. We can then take that value integrated over the dose ranges and normalize to the total dose range.

$$\frac{1}{x_1 j - x_1 i}\frac{1}{x_2 l - x_2 k}\int_{x_{1i}}^{x_{1j}}\int_{x_{2k}}^{x_{2l}}\frac{(\mu_B(x_1,x_2)-\mu_C(x_1,x_2))dx_1 dx_2}{\sqrt{\sigma_B(x_1,x_2)^2+\sigma_C(x_1,x_2)^2}} \tag{1}$$

where μ_B and μ_C denote the mean function of the bliss additive model and the combo drug model. This score is similar to a z score but does not take into account the covariance of the data. Instead it is a conservative estimate of the distance between the two models as we ignore the partial independencies in different areas of

the model.

4. Results

We ran this framework on 2531 drug combination screens including both 7x7 and 5x5 grids of dose ranges for three drug combinations and 752 cell lines. We identified 43 drug pair showing synergism (average z-score higher than 2), all but two of which were from the same drug combo. Figure 3 shows the distribution of scores.

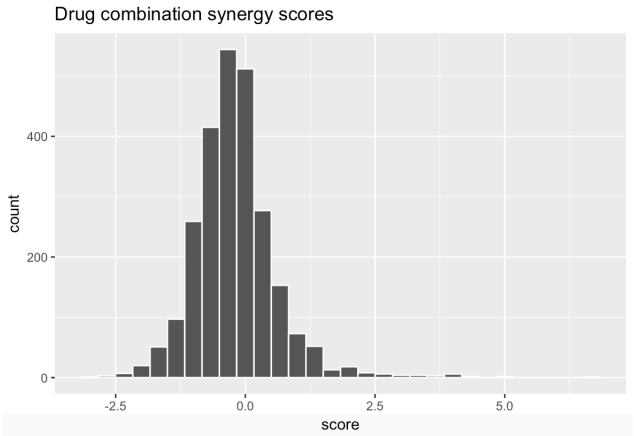

Fig. 3. Synergy score distribution

As a first step toward a more optimal drug concentration layout, we sought to determine which doses were the most informative to synergy. For this we downsample points. We then create a new model with these subsampled points and ask how much less confidence do we have under the new model than we did with the full data set. Figures 4 and 5 show heat maps for the lost confidence for the 5x5 and 7x7 data.

We then compared various dose layouts with the same number of points and compared their confidence differences from the full grid layout for 9 and 16 dose-pair values sampled. To do this we used the GP trained on the full grid and induced sample points at new doses from the normal distribution described by the GP of the full grid at that dose-pair location. We then created a new model on those induced points and measured the same metric versus the bliss additive model. Figures 6 shows the dose layouts we compared.

As shown in figures 7 and 8, in general the star layout performs well as does the anchor-library and the cross. The major inconsistency is with the grid layout.

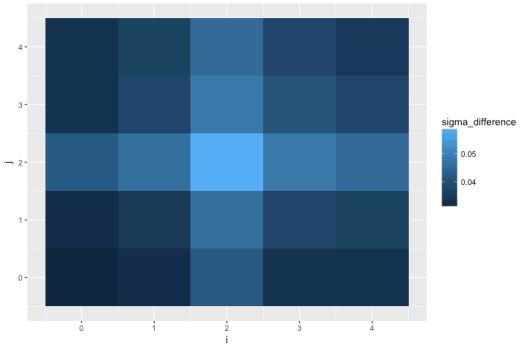

Fig. 4. Relative importance of each data point to synergy confidence in a 5x5 log linear grid.

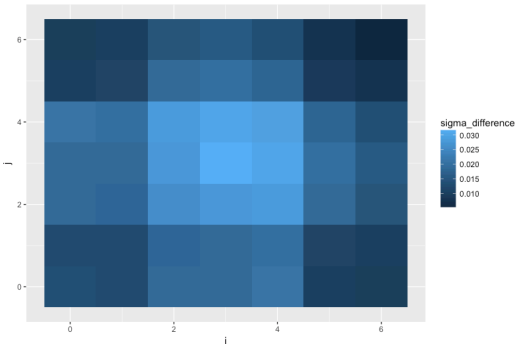

Fig. 5. Relative importance of each data point to synergy confidence in a 7x7 log linear grid.

We believe this is primarily because the sparsity of the 9 sample grid is beyond the length scale of the covariance function which causes it to have low confidence in much of the space whereas the 16 sample grid is dense enough to have decent confidence. While both the diagonal and anti-diagonal perform relatively poorly, the anti-diagonal has somewhat better performance than the diagonal layout.

5. Discussion

Here we have shown a novel model for detecting synergistic drug combinations as well as steps toward finding more optimal drug titration layouts to detect these vital synergies with the most confidence for a given number of samples taken. These methods will have some inherent drawbacks. They rely on the experiment designer to choose dose ranges well. There will always be a trade off between exploration of the dose ranges versus exploitation of the region with the highest likelihood to have the largest synergistic signal. Further work will include applying this method to a larger dataset created by Merck[16] which may tell

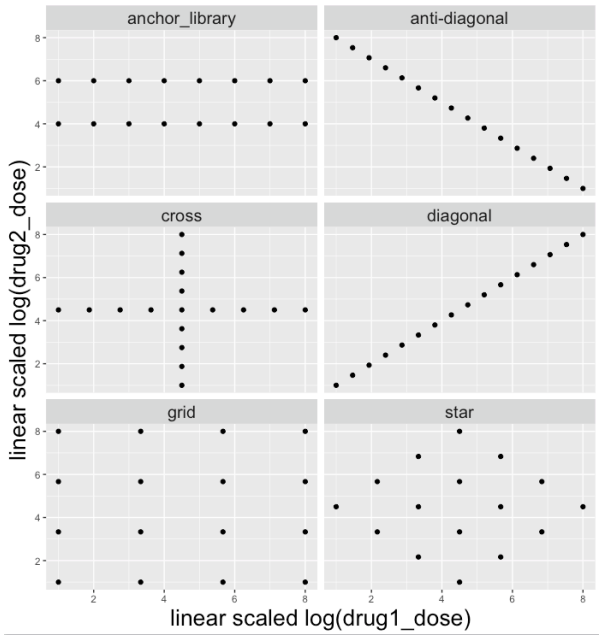

Fig. 6. Dose schedules for comparison for n=16 samples. Note the cross has 17 samples in this case. n=9 are the same shapes without the anchor-library layout.

16 points	mean metric difference	IQR
diagonal	1.8	1.2-2.4
anti-diagonal	1.7	1.1-2.2
cross	1.5	1.0-2.0
star	0.7	0.4-0.9
grid	0.9	0.6-1.2
anchor-library	1.2	0.7-1.6

Fig. 7. Results for 16 dose sample points

9 points	mean metric difference	IQR
diagonal	1.9	1.2-2.4
anti-diagonal	1.8	1.1-2.3
cross	1.6	1.4-2.1
star	1.4	0.9-1.7
grid	2.0	1.5-2.4

Fig. 8. Results for 9 dose sample points.

us more about how generalizable the conclusions are. We also plan on simulating a shift of the chosen doses to show how bad dose choices will favor layouts which explore more of the space whereas good dose choices will favor layouts with many points toward the center of the ranges.

References

[1] Lehar, J. *et al.* Synergistic drug combinations tend to improve therapeutically relevant selectivity. *Nat. Biotechnol.* **27**, 659–666 (2009).

[2] Held, M. A. *et al.* Genotype-selective combination therapies for melanoma identified by high-throughput drug screening. *Cancer Discov* **3**, 52–67 (2013).

[3] Mathews Griner, L. A. *et al.* High-throughput combinatorial screening identifies drugs that cooperate with ibrutinib to kill activated B-cell-like diffuse large B-cell lymphoma cells. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2349–2354 (2014).

[4] Axelrod, M. *et al.* Combinatorial drug screening identifies compensatory pathway interactions and adaptive resistance mechanisms. *Oncotarget* **4**, 622–635 (2013).

[5] Earl, H., Molica, S. & Rutkowski, P. Spotlight on landmark oncology trials: the latest evidence and novel trial designs. *BMC Med* **15**, 111 (2017).

[6] Bliss, C. I. The toxicity of poisons applied jointly1. *Annals of Applied Biology* **26**, 585?615 (1939).

[7] Loewe, S. The problem of synergism and antagonism of combined drugs. *Arzneimittelforschung* **3**, 285–290 (1953).

[8] Greco, W. R., Bravo, G. & Parsons, J. C. The search for synergy: a critical review from a response surface perspective. *Pharmacol. Rev.* **47**, 331–385 (1995).

[9] Foucquier, J. & Guedj, M. Analysis of drug combinations: current methodological landscape. *Pharmacol Res Perspect* **3**, e00149 (2015).

[10] Chou, T. C. Drug combination studies and their synergy quantification using the Chou-Talalay method. *Cancer Res.* **70**, 440–446 (2010).

[11] Di Veroli, G. Y. *et al.* Combenefit: an interactive platform for the analysis and visualization of drug combinations. *Bioinformatics* **32**, 2866–2868 (2016).

[12] Flobak, A., Vazquez, M., Lægreid, A. & Valencia, A. CImbinator: a web-based tool for drug synergy analysis in small- and large-scale datasets. *Bioinformatics* **33**, 2410–2412 (2017).

[13] Koplev, S. *et al.* Dynamic Rearrangement of Cell States Detected by Systematic Screening of Sequential Anticancer Treatments. *Cell Rep* **20**, 2784–2791 (2017).

[14] Rasmussen, C. & Williams, C. *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning (MIT Press, Cambridge, MA, USA, 2006).

[15] Stegle, O. *et al.* A robust Bayesian two-sample test for detecting intervals of differential gene expression in microarray time series. *J. Comput. Biol.* **17**, 355–367 (2010).

[16] O'Neil, J. *et al.* An Unbiased Oncology Compound Screen to Identify Novel Combination Strategies. *Mol. Cancer Ther.* **15**, 1155–1162 (2016).