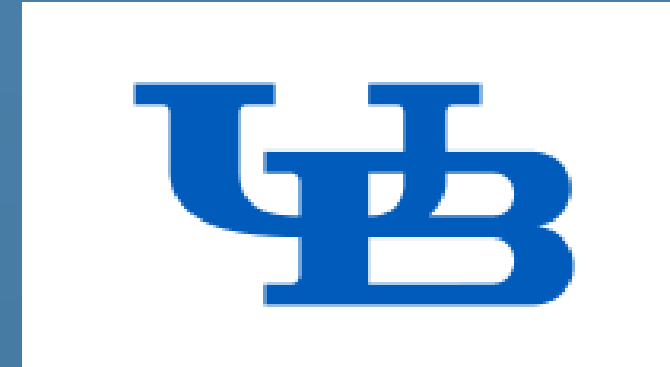




Multi-objective Optimization for Population Pharmacokinetic (PopPK) Model Selection Using NSGA-II and pyDarwin

Xinnong Li (1), Alex Mazur (2), Mark Sale (2), James Craig (2), Keith Nieforth (2), Robert Bies(1)
Institution: (1) University at Buffalo (2) Certara, US

Background & Objectives

The selection of a “good” model typically involves both objective (numerical) and subjective criteria. Subjective criteria, such as interpretation of plots and consideration of biological plausibility. While both of these can in part be described numerically, there still remain aspects of a “good” model that are difficult to quantify formally. This challenge of combining both objective and subjective criteria in solution selection can be addressed by multi-objective optimization (MOO). MOO allows for the simultaneous optimization of multiple criteria. This approach generates a Pareto front, representing a set of non-dominated models where no single solution can be improved in one objective without sacrificing the performance in another objective [1]. An advantage of MOO over single objective optimization is that no, potentially arbitrary, penalties to the OFV are needed. Two objectives of this study are:

- i. Identify a set of non-dominated models by simultaneously optimizing objective function value (OFV) and the number of estimated parameters (n_parms)
- ii. Compare the search results with and without local downhill search to assess its impact on MOO

Methods

This study utilized a 17-DMAG dataset, compromising 66 subjects and 951 observations. The user-defined model search space included 1.57 E6 candidate models, consistent with our previously published study [2]. The search space is shown below:

- Number of compartments (1,2,3)
- Between-subject variability on peripheral volumes and intercompartment clearances
- Effect of weight, age, sex, and serum creatinine on CL
- Effect of weight, age, sex on VC
- Effect of weight on Q
- Effect of weight on Vp
- Between-occasion variability on CL, Q, Vc
- Residual Error Models

NSGA-II algorithm [3] recently added to pyDarwin [4] was used for multi-objective model selection. In this case study, the model selection is driven by two objectives, which are OFV (minus two log likelihood), and n_parms. A local downhill search that is defined as the numerical sum of two objectives was also assessed in this study.

The optimization process ran for 20 generations, with 80 models in each generation. The search was combined with local downhill search, the local search step was implemented every 10 generations. Models were run in parallel, 32 at a time on a Windows server located at SUNY Buffalo.

Results

As the generations progressed, the front that included non-dominated models within each generation shifted toward an improved region, optimizing both OFV and model parsimony. While both objectives improved, the trade-off between them was clearly observed: models with lower OFV generally contained more estimated parameters.

In the search without local downhill search, 17 optimal models were identified on the Pareto front, with OFV ranging from 8034.493 to 9813.408 and n_parms ranging from 5 to 22. The globally best model, as determined by exhaustive search [2] using single objective GA, also appeared on the Pareto Front. All of the optimal models passed the covariance step, while 3 of them (n_parms = 13, 18, 21) failed the convergence step.

Table 1 Model structure of non-dominated models after multi-objective search (without local downhill search)

Total # of parameter	OFV	COM	BSV	BOV	RUV	Covariate
5	9813.4	1	V, CL	-	prop	-
6	9769.9	1	V, CL	-	comb	-
7	8799.4	2	V, CL	-	prop	-
8	8452.4	2	V, CL, V2	-	prop	-
9	8335.4	2	V, CL, V2	Q	prop	-
10	8238.3	3	V, CL, V2	-	prop	-
11	8166.5	3	V, CL, V2	CL	prop	-
12	8134.1	3	V, CL, V2	CL, Q2	prop	-
13	8101.6	3	V, CL, Q3, V2	CL	comb	-
14	8065.6	3	V, CL, V2, Q3	CL, Q2	comb	-
15	8056.9	3	V, CL, V2, Q3	V, CL Q2	comb	-
16	8041.3	3	V, CL, V2, Q3	V, CL Q2	comb	V~WT
17	8041.2	3	V, CL, V2, V3, Q3	V, CL Q2	comb	Q2~WT V~WT
18	8040.8	3	V, CL, V2, Q3	V, CL Q2	comb	V3~WT, V~WT V2~WT
19	8035.1	3	V, CL, V2, V3, Q3	V, CL Q2	comb	Q~WT, V~WT, CL~SCR
21	8034.6	3	V, CL, V2, V3, Q3	V, CL Q2	comb	Q~WT, Q2~WT, V3~WT, V~SEX, CL~SCR
22	8034.5	3	V, CL, V2, V3, Q3	V, CL Q2	comb	Q2~WT, V3~WT, V~SEX, CL~SCR, CL~WT, CL~SEX
Exhaustive Search	8041.3	3	V, CL, V2, Q3	V, CL, Q2	comb	V~WT

Results (continued)

When the local downhill search was implemented, 22 nondominated models were selected within the search space. OFV ranges from 8023.747 to 9813.408, and n_parms ranges from 5 to 26. Similarly, the globally best model was included on the Pareto front. Models identified through the downhill run could perform better, achieving lower OFV with the same level of parsimony. All non-dominated models on the final Pareto front passed the covariance step, while 8 of them failed the convergence step (n_parms = 18, 19, 20, 22, 23, 24, 25, 26), indicating a tendency toward overparameterization. Notably, even though some downhill-searched models have lower OFV, they are more prone to fail the convergence.

These models, selected based on objective criteria, are then presented to the user to further examine a manageable set and select one or more from that set based on subjective criteria such as biological plausibility and diagnostic graphics as the “best” model(s).

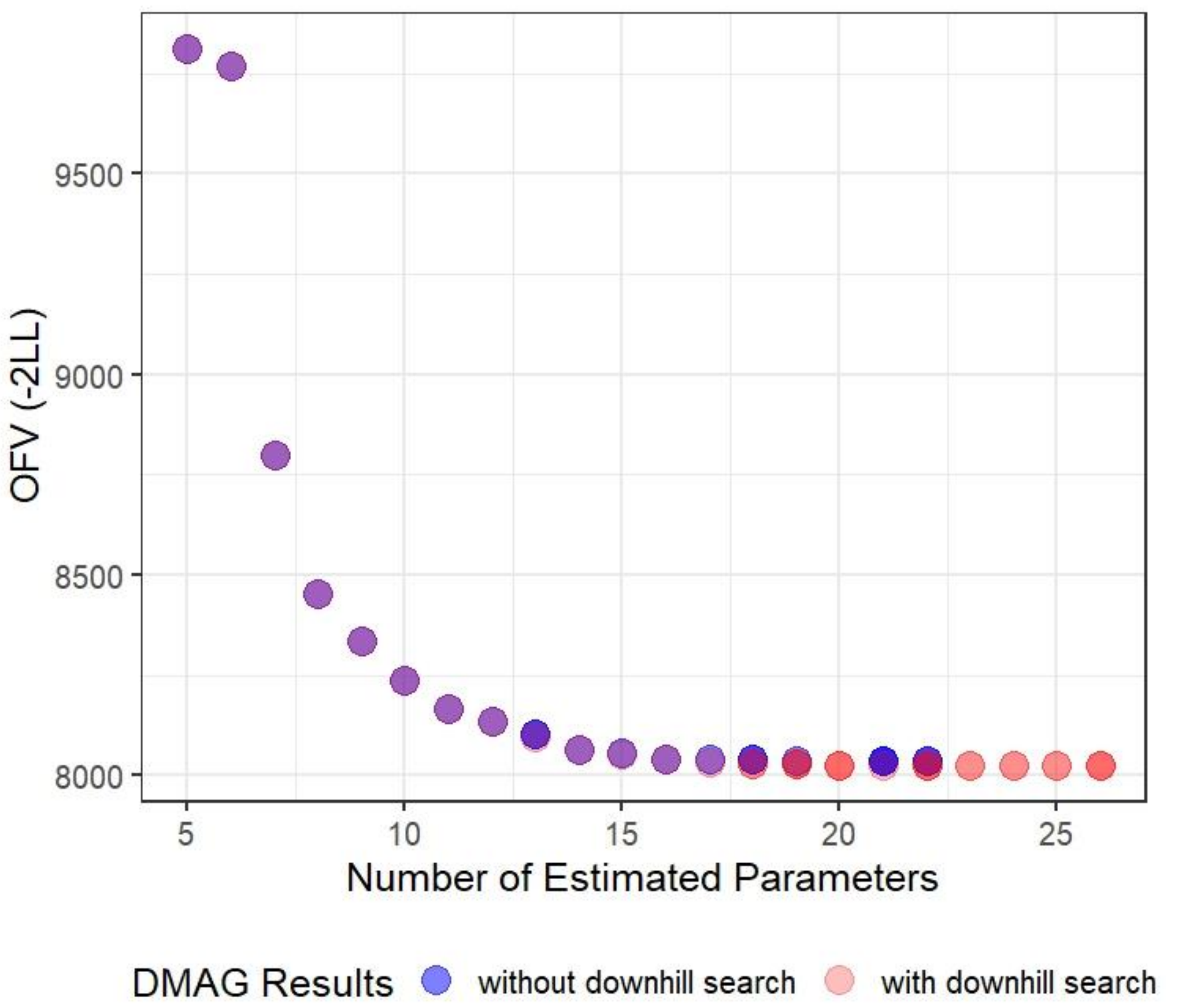


Figure 1 Comparison between MOO search results with and without local downhill search

Conclusions

The multi-objective optimization method successfully identified a set of non-dominated models within the search space by evaluating OFV and parsimony at the same time. Both search strategies (with and without local downhill search) successfully found the globally best model, while the incorporation of downhill search expanded the exploration area, yielding more non-dominated models. Therefore, integrating the local downhill search into the multi-objective optimization is recommended to enhance the model selection. Overall, the application of NSGA-II successfully generated a set of non-dominated PopPK models, offering greater insight and flexibility for decision-making in model selection.

[1] M. Kochenderfer and T. Wheeler. Algorithms for optimization. 2019. The MIT Press.
[2] Li X et al. J Pharmacokinet Pharmacodyn. 2024 Dec;51(6):785-796.
[3] K. Deb et al. IEEE Transactions on Evolutionary Computation, vol. 6, no. 2, pp. 182-197, April 2002.
[4] <https://github.com/certara/pyDarwin>