# Machine Learning Based Model Selection with pyDarwin in Pirana

Mark Sale, Keith Nieforth, James Craig Certara, USA Inc.



# Machine Learning Based Model Selection with *pyDarwin* in *Pirana:* Automated NONMEM PK/PD Model Selection



# **Background & Objective**

pyDarwin[1] is a recently released open-source Python package for automated NONMEM Pharmacokinetic/Pharmacodynamic (PK/PD) model selection developed under a grant from, and in collaboration with FDA[2]. It offers a general solution for searching a wide range of PK and PD models within a user defined model search space and permits customization of model fitness criteria to allow prioritization of desired model performance characteristics in the model selection process.

Pirana is a PMX analysis workbench that integrates separate software tools under a common user interface, lending organization and efficiency to the PMX workflow while preserving the flexible nature of the underlying toolsets[3]. Originally developed for use with NONMEM, Perl speaks NONMEM (PsN), and R/xpose package, Pirana has been enhanced to support command line modeling in R with Certara's NLME-Engine (RsNLME) [4], and now supports machine learning based model selection through a graphical user interface to pyDarwin.

### Results

The demonstration provides a comprehensive overview of the theory and execution of a machine learning based PMX model search.

The python command line interface to pyDarwin is logical and well organized but complex, requiring the editing of three text files; one to define the basic NONMEM control file structure, one to define the search space and one to define the search options and parameters. In addition, some knowledge of the underlying Python language.

The availability of pyDarwin in the Pirana workbench provides a robust and efficient method to perform PK/PD analysis within a graphical user interface, integrating with existing Pirana features to view estimation results and generate diagnostics across both the individual candidate models and the final best model selected from the automated search.

## **Additional Figures**

### General Setup and Algorithm Selection

📒 Darwin Search properties				_		×
General	General					
Data ⊡ Model template	Name of the search	search_1				
-ADVAN1	Author of the search	Certara	?			
-ADVAN2 -ADVAN3	Algorithm	GA 🗖	?			
-ADVAN4 -ADVAN11	Number of generations	GA GP	?			
ADVAN12	Number of models in every generation	RF	?			
Sigmas	Seed value for random number generator	EX	?			
Downhill step Penalties	✓ Use system options	?				
Postprocessing	Performance					
Directories	Model run timeout (seconds)	1200	?			
Model cache	Model run priority	Below normal	?			
Custom options	Size of exhaustive search model batch	100	?			
	Number of parallel processes to perform the "ask" step	4	?			
	Help		Save		Cancel	

### Structural and Omega Parameters Search

📒 Darwin Search properties

- 🗆 X

# Methods

A basic introduction to machine learning model selection will be provided covering general theory and an overview of available search algorithms in pyDarwin. This will be followed by a demonstration of the Pirana software covering the following topics:

- Overview of Pirana user interface demonstrating how to create, execute, postprocess and generate report objects for a single model run
- Comparison of machine learning algorithms for model selection, which include: Genetic Algorithm, Bayesian Optimization, Random Forest, Gradient Boosted Random Tree,

### Pirana and pyDarwin PMX Workflow



# Discussion

Machine learning based pharmacometric (PMX) model development may offer advantages over manual approaches by providing an efficient, objective and more robust method for model selection.

General	ADVAN2	setup								
Data	Thetas							Omegas		
Model template	Parameter	Presence	Theta	Lower	Initial	Upper	Frozen	Presence	Initial	Frozen
	CL	Present 🗕	CL	0	0.05		Г	Present 📼	0.3	Г
-CLCR	Ка	Present 🖃	Ка	0	1		Г	Searched 📼	0.3	Г
	V	Present 🗕	v	0	0.1		Г	Present 📼	0.3	Г
	ALAG1	Searched 📼	ALAG1	0	0.1		Г	Searched 📼	0.3	Г
-CLCR	F1	None 📼	F1	0	0.5		Г	None 🗖	1	Г
⊂SEX Template extras	ALAG2	None 📼	ALAG2	0	1		Г	None 📼	1	Г
Sigmas	F2	None 🗖	F2	0	0.5		Г	None 🗖	1	Г
Directories GA setup Model cache Custom options										
	Help								Sa	ve Cancel

### **Covariate Parameters Search**

General	ADVAN4	-SEX setu	1p							
Data	Thetas									
Model template	Parameter	Category	Presence	Relation	Theta	Lower	Initial	Upper	Frozen	
⊟-ADVAN2	CL	1	Searched 🗔	exponential	CL SEX 1		0.1		Г	
	Ka	1	Searched	exponential	Ka SEV 1		0.1			
L	Ka	1	Searched -	exponential	Ka_SEA_I		0.1		-	
I ⊡-ADVAN4	Q	1	Searched 🗔	exponential	Q_SEX_1		0.1		Г	
-wt	V2	1	Searched 📼	exponentia⊨	V2_SEX_1		0.1		Γ	
-CLCR	V3	1	Searched 💳	exponentia⊨	V3_SEX_1		0.1		Г	
	ALAG1	1	Searched -	exponential	ALAG1 SEX 1		0.1		Г	
Template extras		-	Scarched -	exponential			0.1		_	
Downbill step	F1	1	None —	exponential	F1_SEX_1		0.1		E	
Penalties	ALAG2	1	None -	exponentia	ALAG2_SEX_1		0.1		Г	
Postprocessing	F2	1	None 🗕	exponentia	F2_SEX_1		0.1		Г	
Directories	ALAG3	1	None	exponentia	ALAG3 SEX 1		0.1		Г	
GA setup	E2	1	Nono	ovpopoptiat	E2 CEV 1		0.1		_	
Model cache	F3	1	None	exponential	F3_SEA_I		0.1		1	
	Ush								-	Canad

Particle Swarm Optimization and full Exhaustive Search

- Setup of the model search space, including specification of structural and omega model parameters, and covariate-parameter relationships to include in the search
- Creation and execution of the automated model selection with the pyDarwin interface in Pirana, including methods to visualize model fitness across iterations of the search
- Exploration of model diagnostics using the final best model selected from pyDarwin's automated model selection

The graphical user interface in Pirana greatly simplifies execution of pyDarwin model searches and enables the usability of machine learning based model selection for PMX workflows[3].

# References

[1] https://certara.github.io/pyDarwin/html/index.html[2] https://grants.nih.gov/grants/guide/rfa-files/rfa-fd-21-027.html

[3] Keizer RJ, Karlsson MO and Hooker A. Modeling and Simulation Workbench for NONMEM: Tutorial on Pirana, PsN, and Xpose. CPT: Pharmacometrics & Systems Pharmacology (2013) 2, e50; doi:10.1038/psp.2013.24.
[4] https://certara.github.io/R-Certara/

### Fitness vs Iteration Plot





Want to learn more? << Scan Here