A general workflow for parameter estimation to help establish confidence in model predictions

Sietse Braakman, PhD
Computational Biology
MathWorks

Topics that aren’t explicitly covered

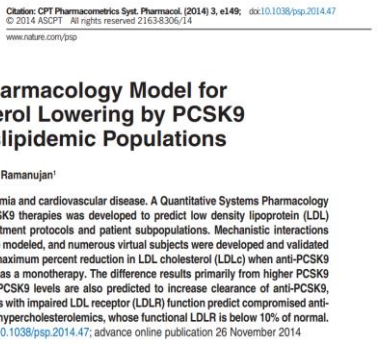
- Non-linear mixed effects estimation
- Virtual population approaches

Case study: calibrate PCSK9 model to alternate data

Scenario:

- I'm a modeler at a pharmaceutical company
- My company is developing a new therapy for hyperlipidemia and associated cardiac disease
- I am tasked to do a feasibility study using modeling
- First aim is to determine whether our therapy can match anti-PCSK9 efficacy

Recalibrate model and compare predictions from both compounds for efficacy



Citation: CPT Pharmacometrics Syst. Pharmacol. (2014) 3, e149; doi:10.1038/psp.2014.47

www.nature.com/psp

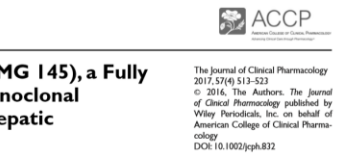
ORIGINAL ARTICLE

A Mechanistic Systems Pharmacology Model for Prediction of LDL Cholesterol Lowering by PCSK9 Antagonism in Human Dyslipidemic Populations

K Gadkar[1], N Budha[1], A Baruch[1], JD Davis[1], P Fielder[1] and S Ramanujan[1]

PCSK9 is a promising target for the treatment of hyperlipidemia and cardiovascular disease. A Quantitative Systems Pharmacology model of the mechanisms of action of statin and anti-PCSK9 therapies was developed to predict low density lipoprotein (LDL) changes in response to anti-PCSK9 mAb for different treatment protocols and patient subpopulations. Mechanistic interactions and cross-regulation of LDL, LDL receptor, and PCSK9 were modeled, and numerous virtual subjects were developed and validated against clinical data. Simulations predict a slightly greater maximum percent reduction in LDL cholesterol (LDLc) when anti-PCSK9 is administered on statin background therapy compared to as a monotherapy. The difference results primarily from higher PCSK9 levels in patients on statin background. However, higher PCSK9 levels are also predicted to increase clearance of anti-PCSK9, resulting in a faster rebound of LDLc. Simulations of subjects with impaired LDL receptor (LDLR) function predict compromised anti-PCSK9 responses in patients such as homozygous familial hypercholesterolemics, whose functional LDLR is below 10% of normal.
CPT Pharmacometrics Syst. Pharmacol. (2014) **3**, e149; doi:10.1038/psp.2014.47; advance online publication 26 November 2014

Model



Special Populations

ACCP

The Journal of Clinical Pharmacology 2017,57(4) 513-523

DOI: 10.1002/jcph.832

Evaluation of Evolocumab (AMG 145), a Fully Human Anti-PCSK9 IgG2 Monoclonal Antibody, in Subjects With Hepatic Impairment

John P. Gibbs, PhD[1,5], J. Greg Slatter, PhD[1,6], Ogo Egbuna, MD, MSc[1], Michelle Geller, MD[2], Lisa Hamilton, MSc[3], Clapton S. Dias, PhD[1,7], Ren Y. Xu, MSc, MBA[1,8], Jessica Johnson, BS[1], Scott M. Wasserman, MD[4], and Maurice G. Emery, PharmD, PhD[1]

Abstract

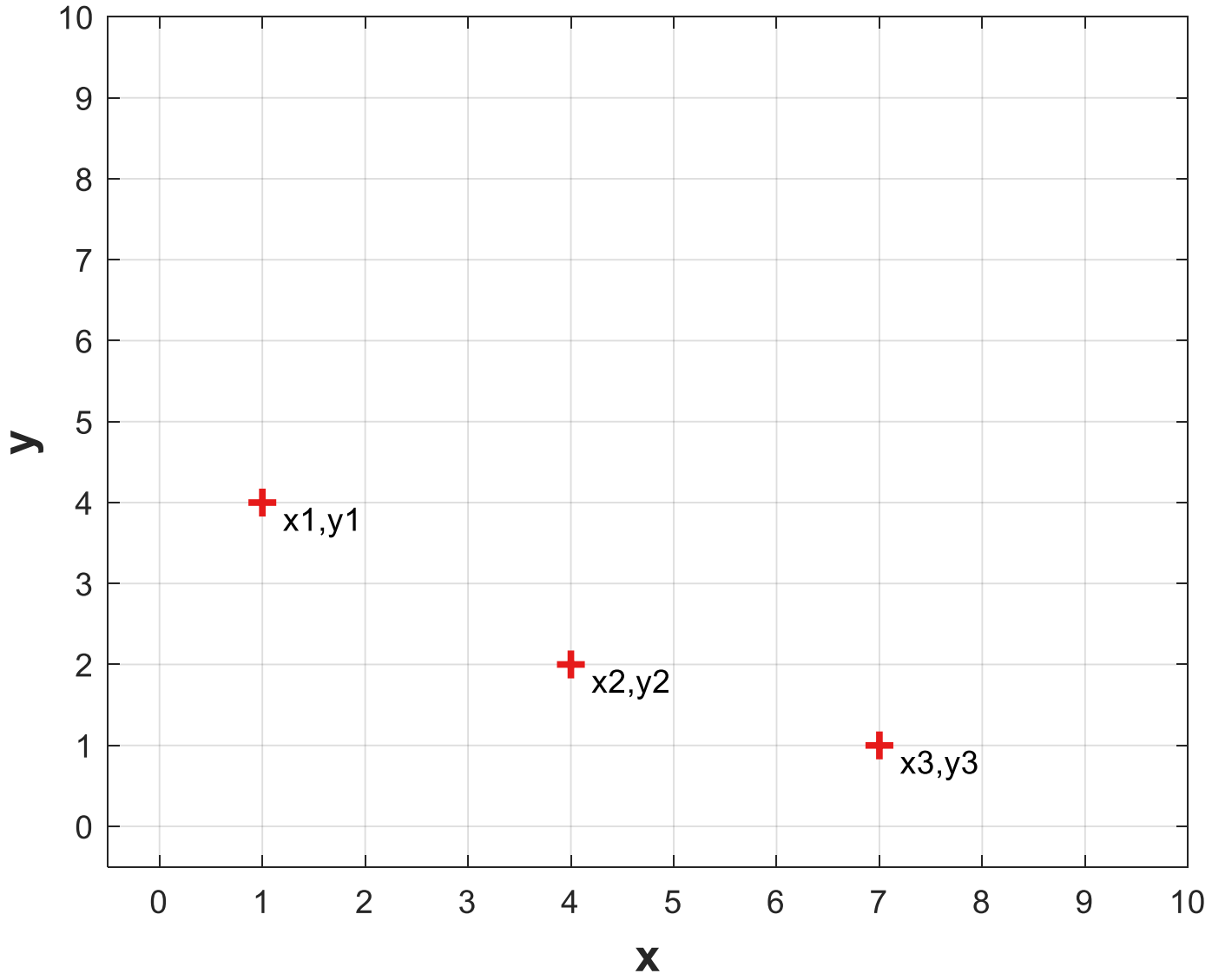
Evolocumab binds PCSK9, increasing low-density lipoprotein cholesterol (LDL-C) receptors and lowering LDL-C. Target-mediated evolocumab elimination is attributable to PCSK9 binding. As circulating PCSK9 and LDL-C levels are primarily regulated by the liver, we compared evolocumab pharmacokinetics, pharmacodynamics, and safety in individuals with and without hepatic impairment. An open-label, parallel-group study evaluated the pharmacokinetics of evolocumab in hepatic-impaired (Child-Pugh Class A or B) or healthy adults. Participants were classified as having no, mild, or moderate hepatic impairment (n = 8/group) and received a single 140-mg evolocumab dose. Assessments of unbound evolocumab and PCSK9 were made predose and postdose. Adverse events were monitored throughout the study. No significant association was observed between baseline PCSK9 and increasing level of hepatic impairment. No difference in extent and time course of PCSK9 or LDL-C reduction was observed despite an apparent decrease in mean unbound evolocumab exposure with increasing hepatic impairment (Jonckheere-Terpstra trend test; maximum serum concentration $P = .18$; area under the curve $P = .09$). Maximum reductions were observed in moderately impaired subjects vs healthy individuals: mean maximum serum concentration –34%; mean area under the concentration-time curve (AUC) –47%. On average, unbound PCSK9 serum concentrations fell by >80% at 4 hours after a single evolocumab dose. Mean (95% confidence interval) maximum LDL-C reductions in the healthy, mild, and moderate groups were –57% (–64% to –48%), –70% (–75% to –63%), and –53% (–61% to –43%), respectively. No safety risks were identified. These results support evolocumab use without dose adjustment in patients with active liver disease and mild or moderate hepatic impairment.

Data from healthy subjects

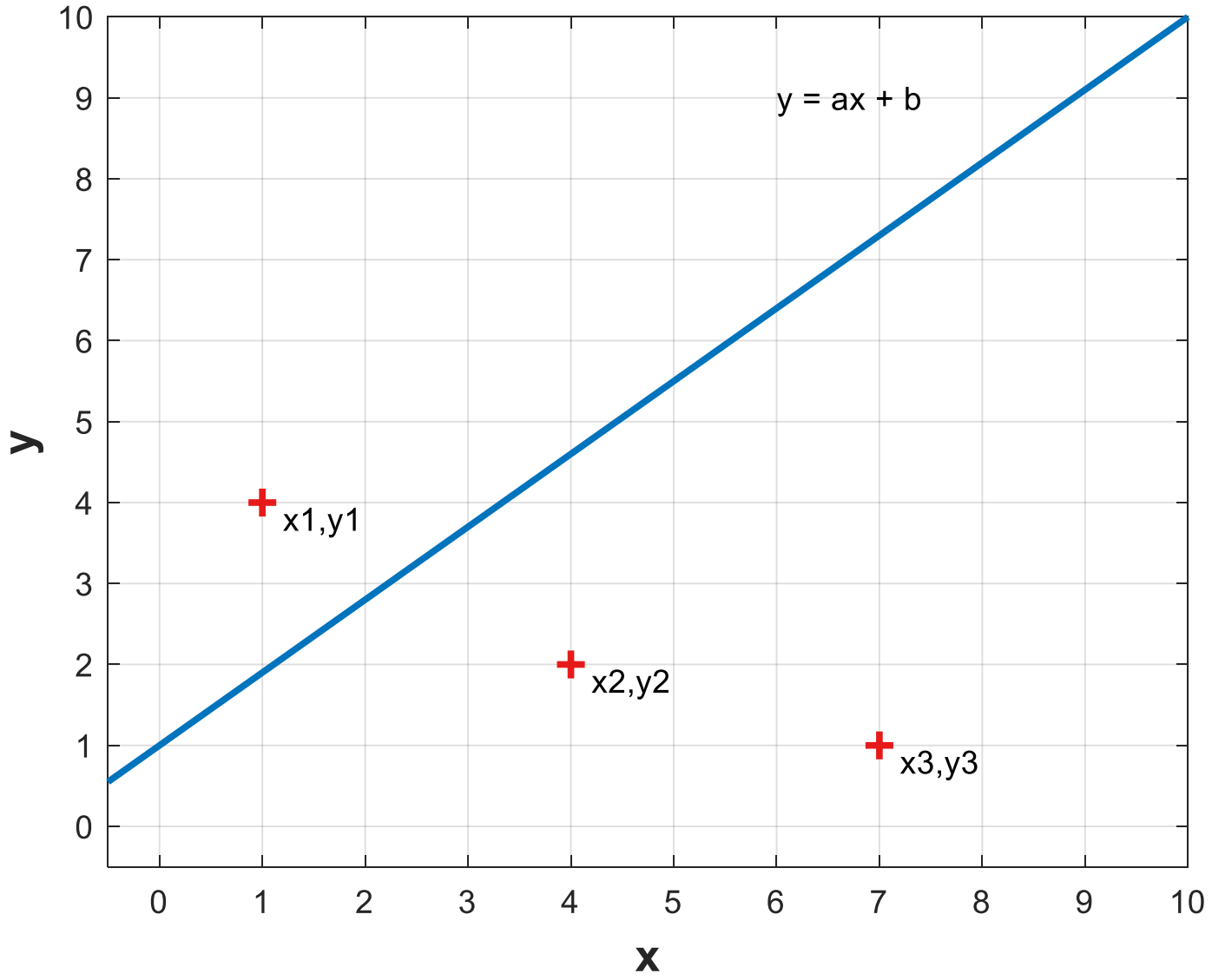
Agenda: Parameter Estimation

- Central concepts
  - Objective function
  - Optimization
- Mechanics of parameter estimation
  - Setting up an optimization
  - Pitfalls
- Workflow to guide estimation process
  - Sensitivity analysis
  - Identifiability analysis
  - Confidence intervals

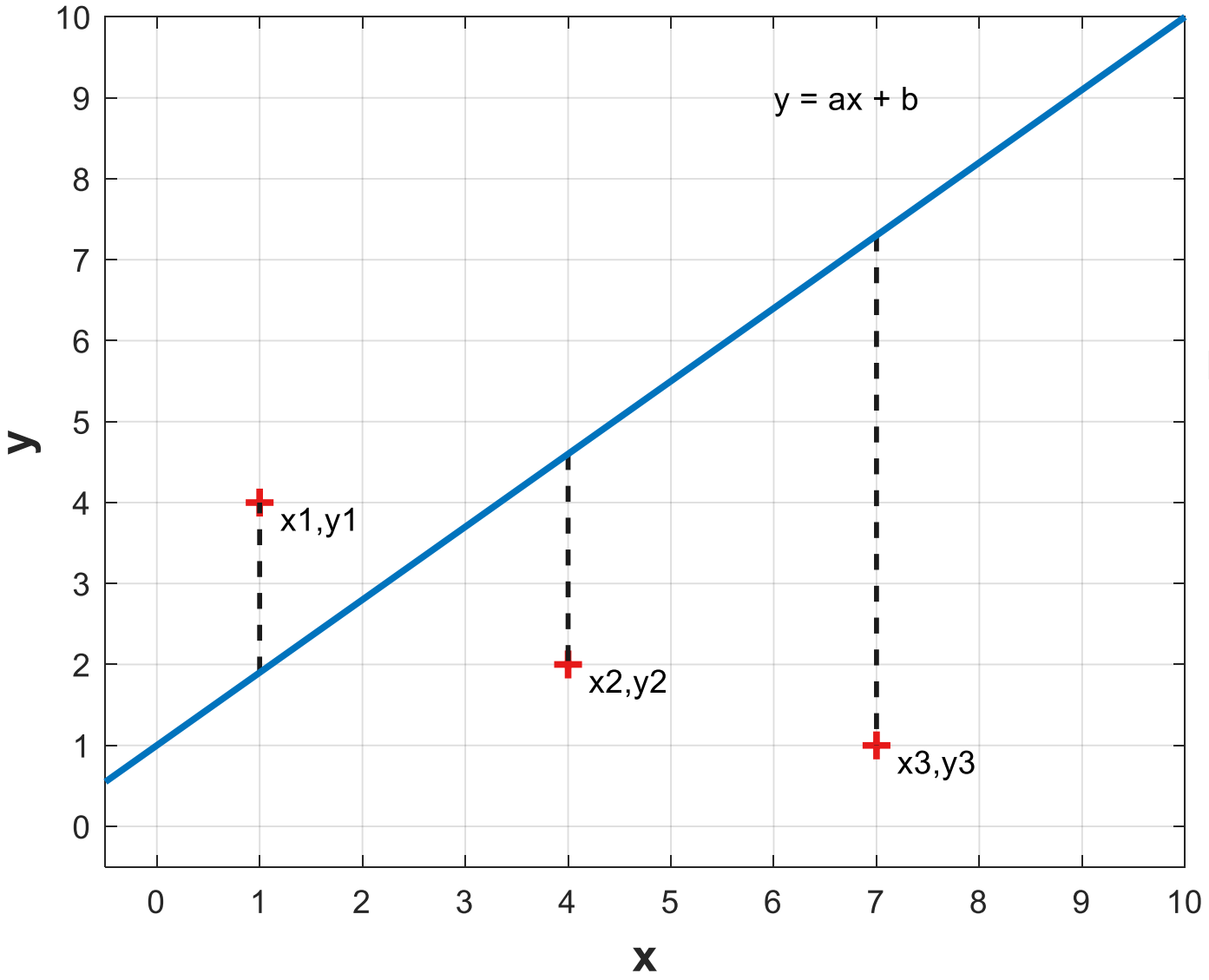
1. Start with a set of data points



2. Define an appropriate model



3. Introduce an objective function

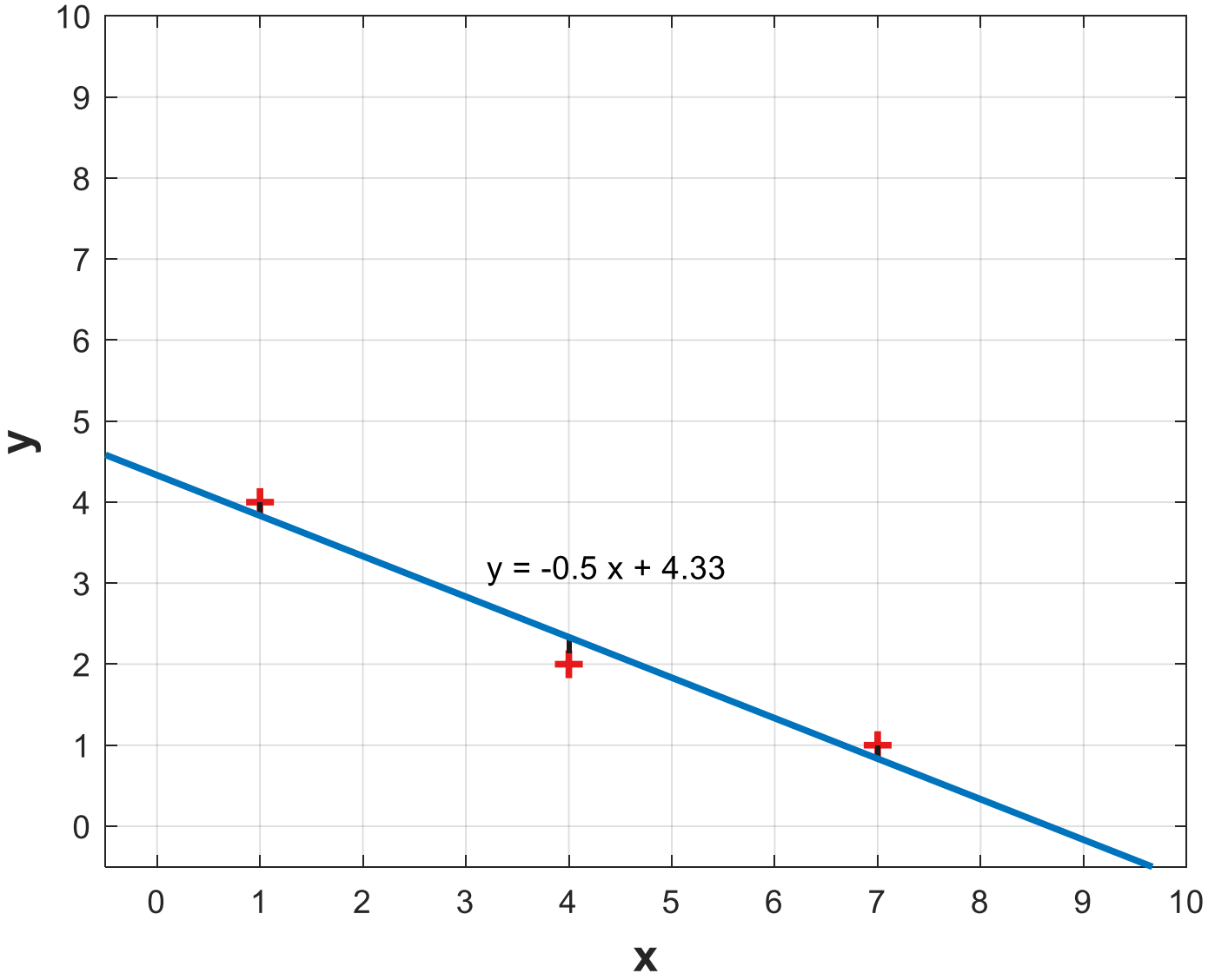


Example of a simple objective function:

$$O(a,b)=\sum_{i=1}^{n}\left(y_i-y(x_i)\right)^2$$

Alternative names: cost or loss function

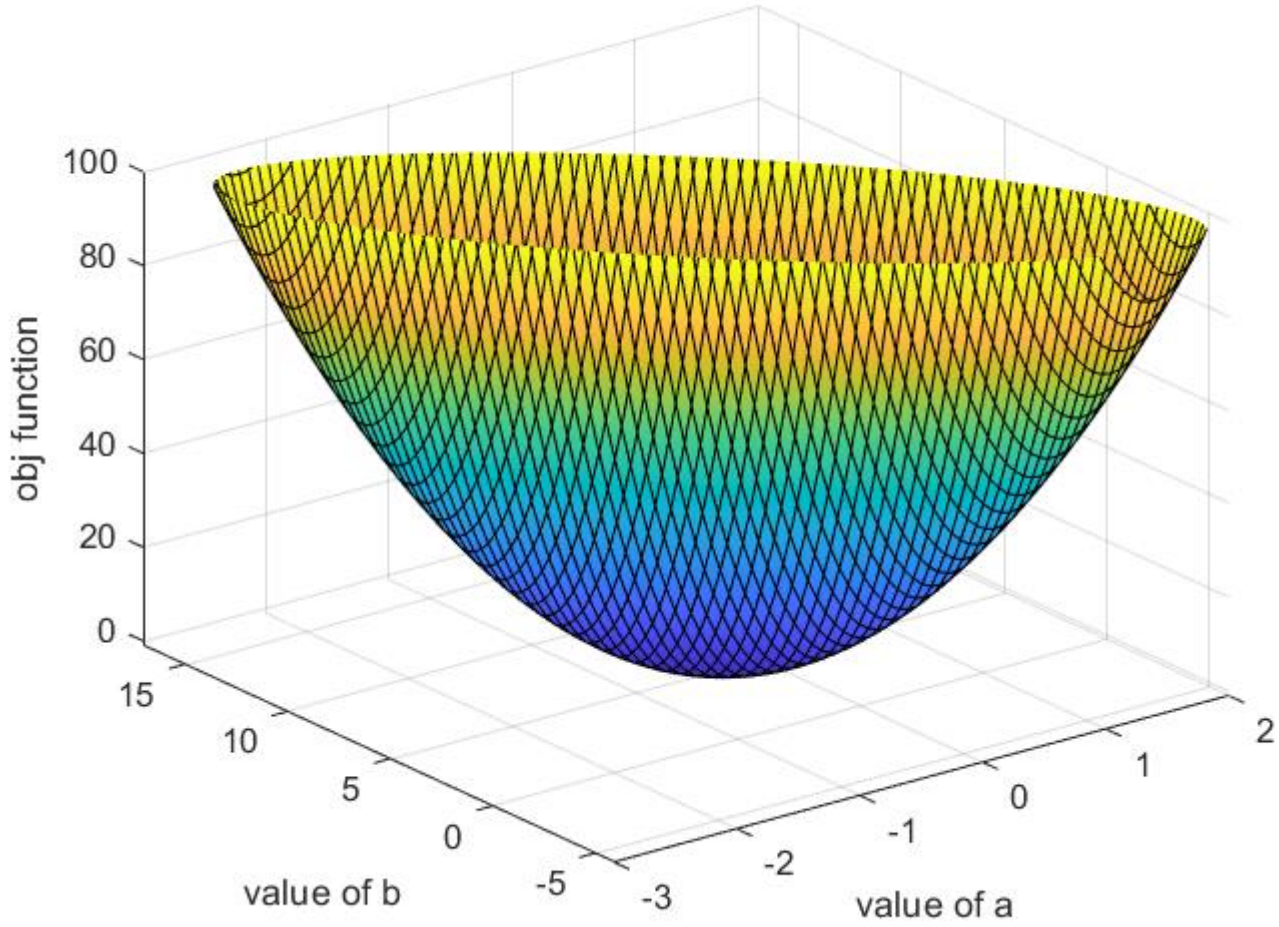
4. Minimize value of objective function to obtain estimates



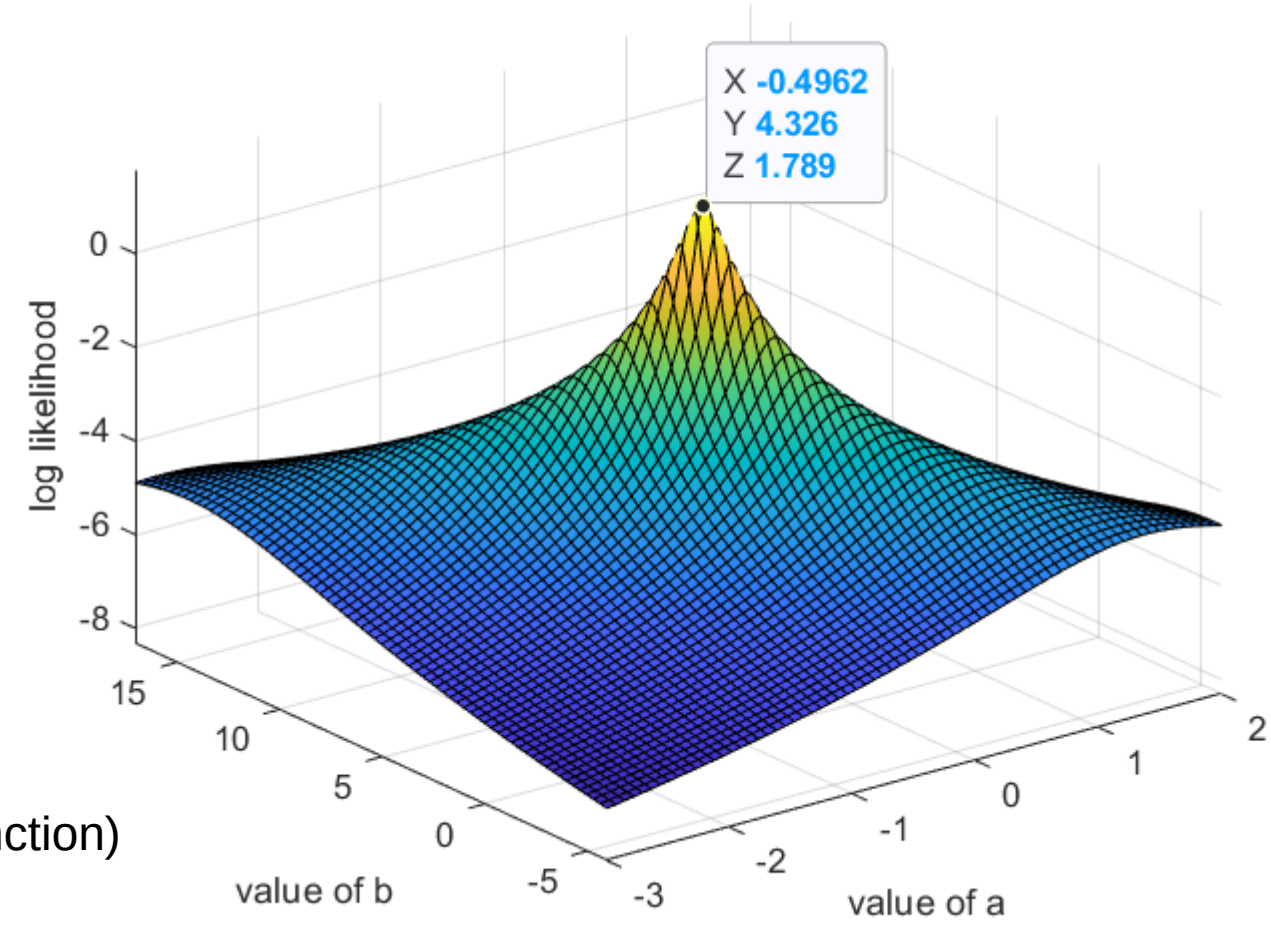
Find a,b for which the value of O(a,b) is minimal

$$O(a,b)=\sum_{i=1}^{n}\left(y_i-y(x_i)\right)^2$$

What does an objective function look like?



Using the log likelihood function



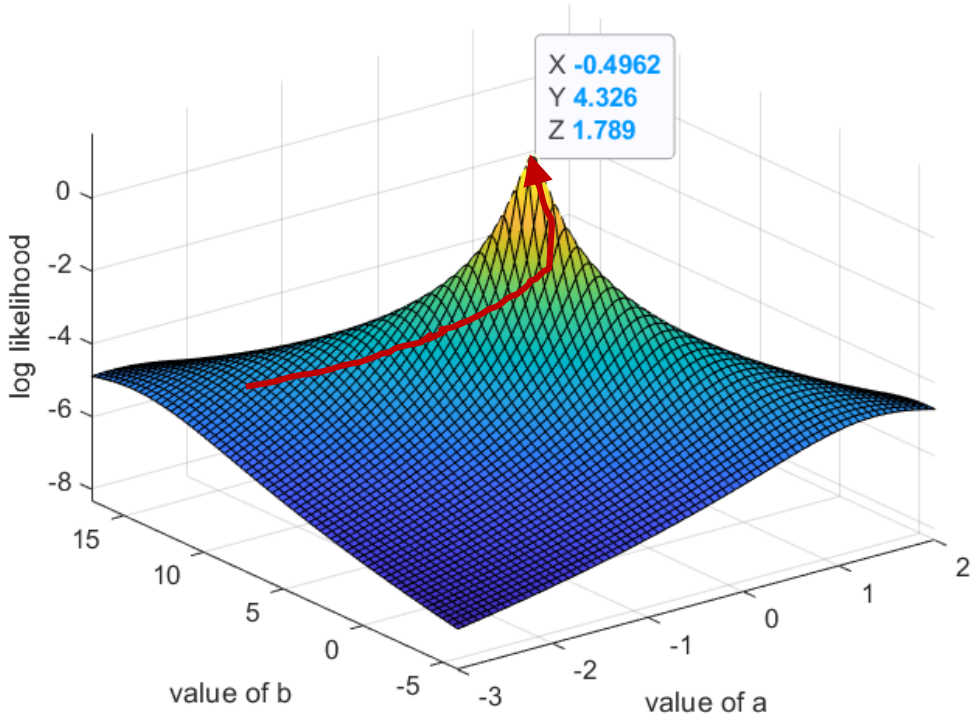
LL = -log(objective function)

Taking the logarithm makes calculations easier

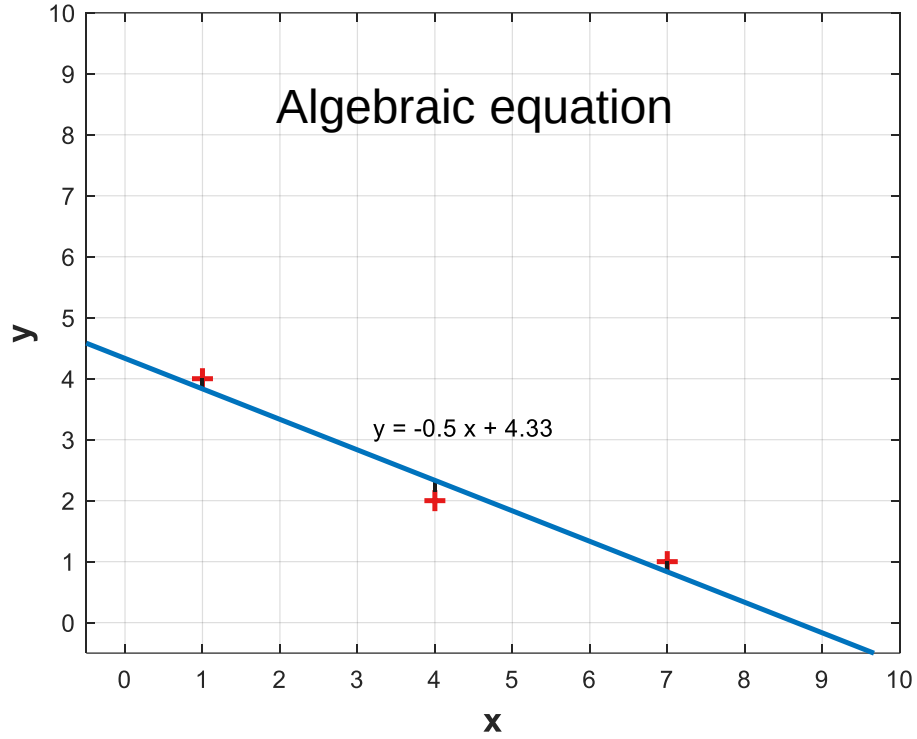
The coordinates of a,b at the maximum of the log likelihood function represent the optimal parameter values

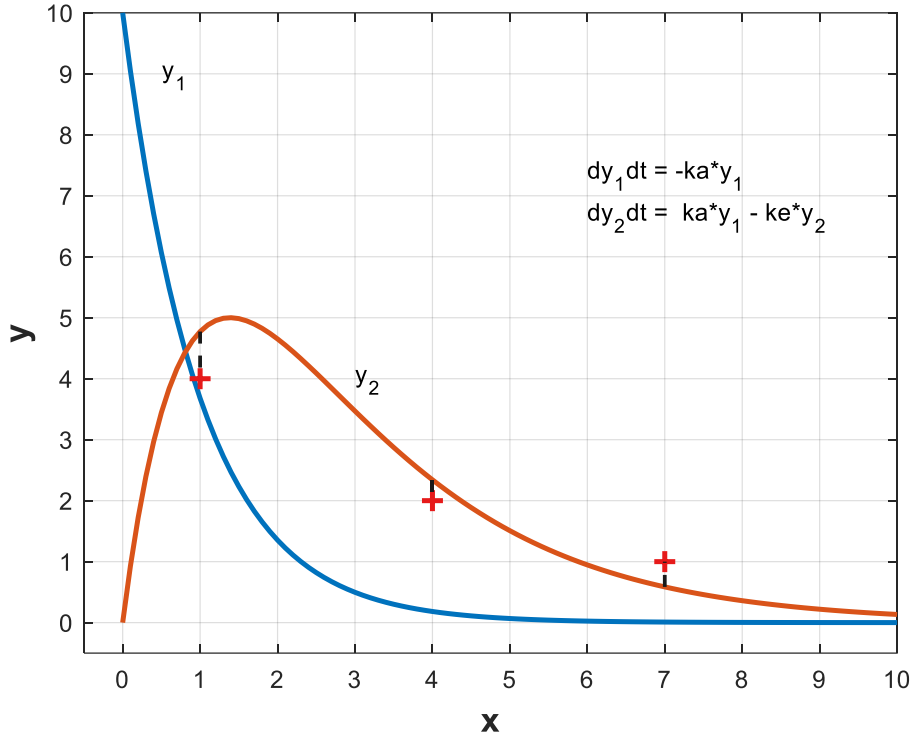
Finding the maximum value of loglikelihood

- Take the maximum value of the surface
  - Calculate entire ‘surface’
  - Computationally expensive
- Use an optimization algorithm
  - E.g. following a gradient
  - Calculate only along optimization path



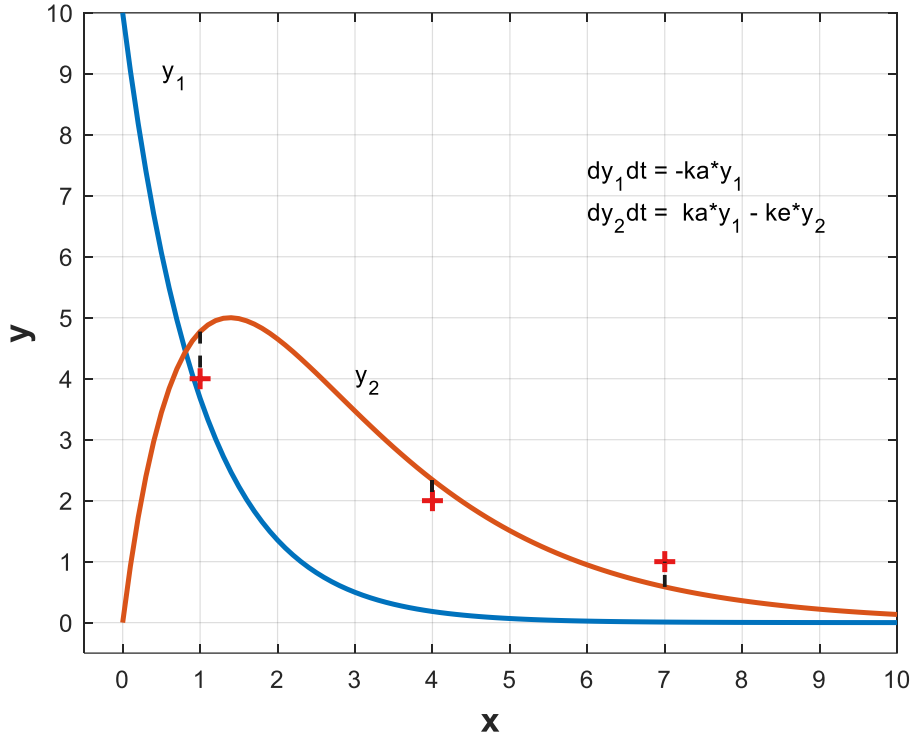
Applying these concepts to ODE models





Applying these concepts to ODE models

1. Solve the set of ODEs as a function of
   - Parameter values – ka, ke
   - Initial conditions – $y_1(0)$, $y_2(0)$
   - Dose schedules
2. Calculate the objective function
   - Decide on appropriate objective function
     - Which observations/data to include
     - How the data maps to the model
   - Take into account the residual error model
     - Constant, proportional, combined, exponential



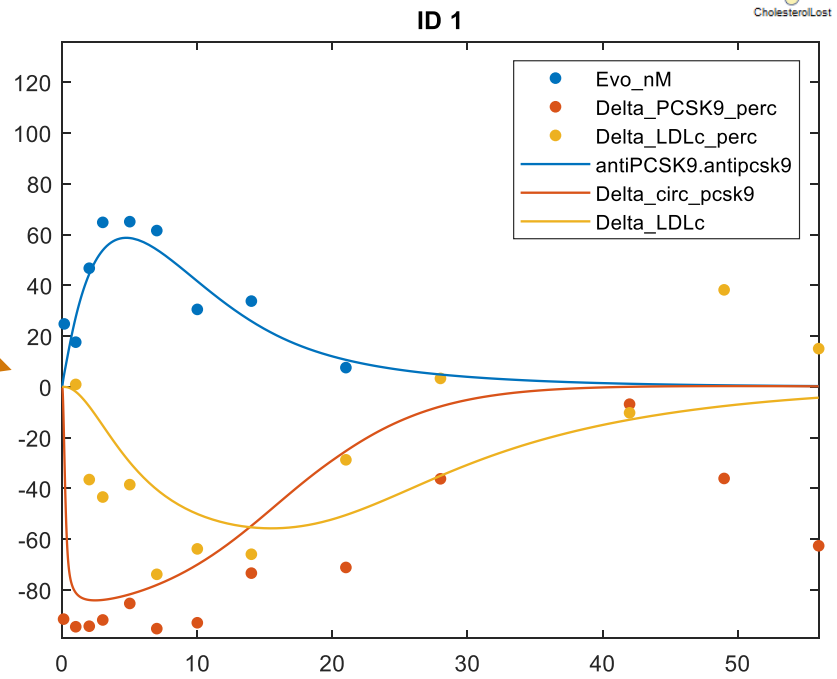
Agenda: Parameter Estimation

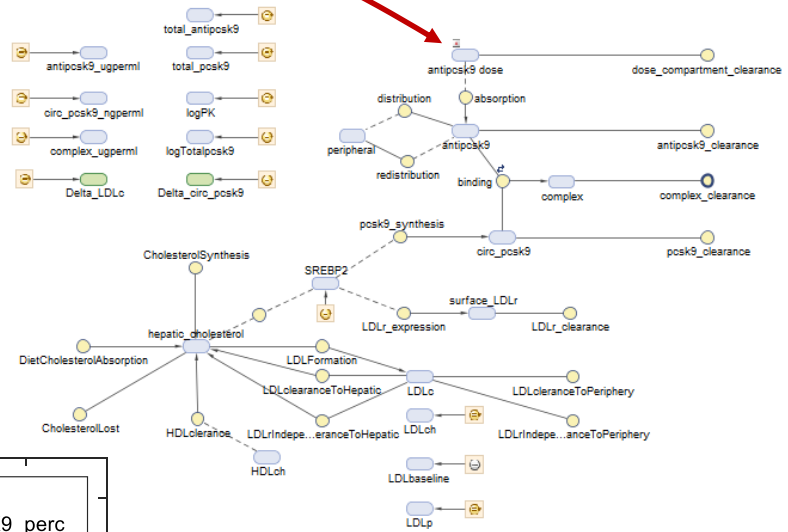
- Central concepts
  - Objective function
  - Optimization
- Mechanics of parameter estimation
  - Setting up an optimization
  - Pitfalls
- Workflow to guide estimation process
  - Sensitivity analysis
  - Identifiability analysis
  - Confidence intervals

Compare data with simulation, not a fit

ID	Time	Evo	PCSK9	LDLc	Dose
1	0				140
1	1	3	-90	-10	
1	3	15	-95	-30	
…					
2	0				140

data

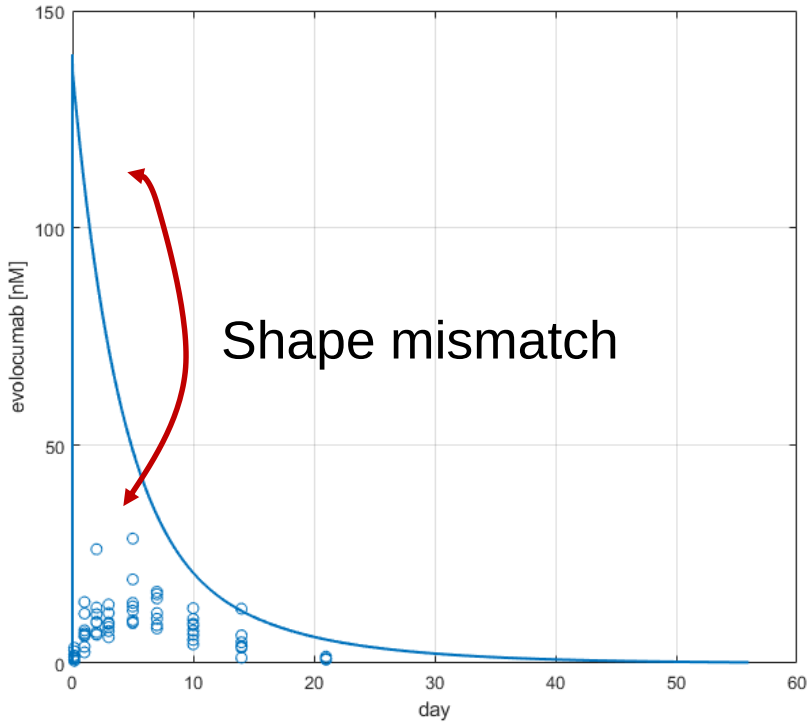




simulation

Pitfall 1
Mapping of dose or response column to model

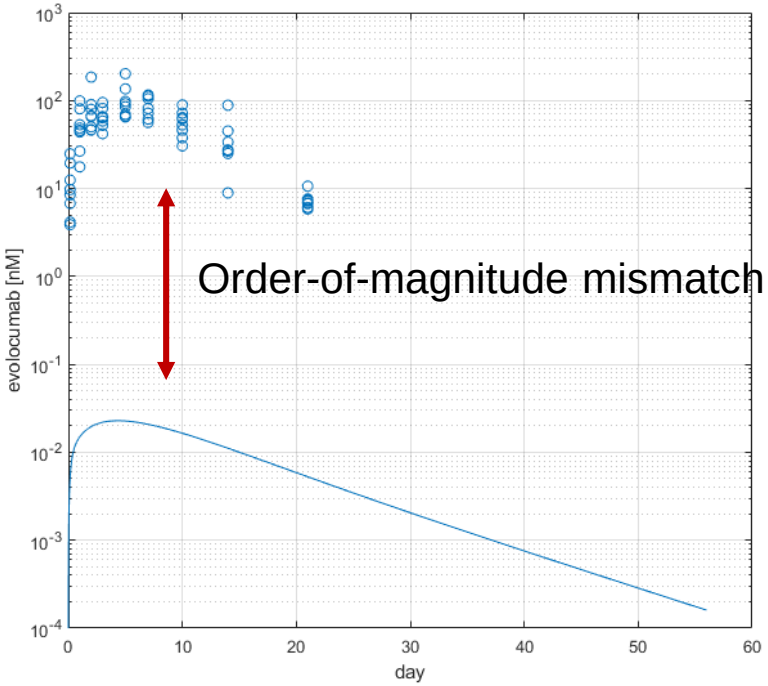
- Adding a dose or response to the wrong state (ODE) in the model
- Symptom:



- Solution:
  - Determine what the protocol for the experimental data was
  - For dose: use the right route of administration/type of dosing
  - For response mismatches:
    - Determine the equivalent state in your model to your data column
    - Or create an equivalent state in your model
    - Map the data column to the correct state in your model

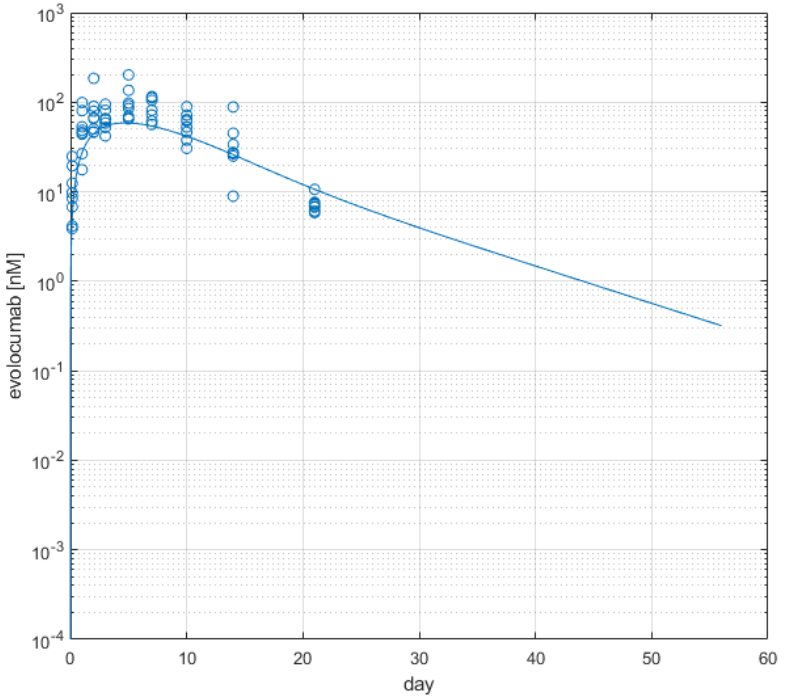
Pitfall 2
Units

- Inconsistent units between dataset and model
- Symptom:



Solution:

- Use unit conversion (SimBiology)
- Convert your units manually, either in your model or dataset



Pitfall 3
Data incompatible with error model

Gibbs_JCP16_evo

	ID	Time_day	Evo_ug_mL	Delta_PCSK9_perc	Delta_LDLc_perc	Dose_mg	Evo_nM
	group	independent	dependent	dependent	dependent	CLASSIFICATION	dependent
	dimensionless	day	dimensionless	dimensionless	dimensionless	milligram	nanomole/liter
							Evo_ug_mL*1e6/1418001800
1	1	0	NaN	NaN	NaN	140	NaN
2	1	0	NaN	0	NaN	NaN	NaN
3	1	0	NaN	NaN	0	NaN	NaN
4	1	0.1000	NaN	-91.4646	NaN		
5	1	0.1500	3.5172	NaN	NaN		

With geometric mean: $f_{gm} = \left(\prod_i^N |f_i| \right)^{1/N}$

If one observation is zero, $f_{gm} = 0$

Proportional error model: $\sum_i^N \frac{(y_i - f_i)^2}{f_i^2 / f_{gm}^2}$

Results in divide by zero, leading to NaN (not a number)

[Link to error models]

ERROR MODEL

Use the same error model for each response : proportional

ESTIMATION METHOD

Use a global solver

lsqnonlin fminunc lsqcurvefit

fminsearch fmincon

ALGORITHM SETTINGS

Click to view options

ADVANCED ALGORITHM SETTINGS

Click to view advanced options

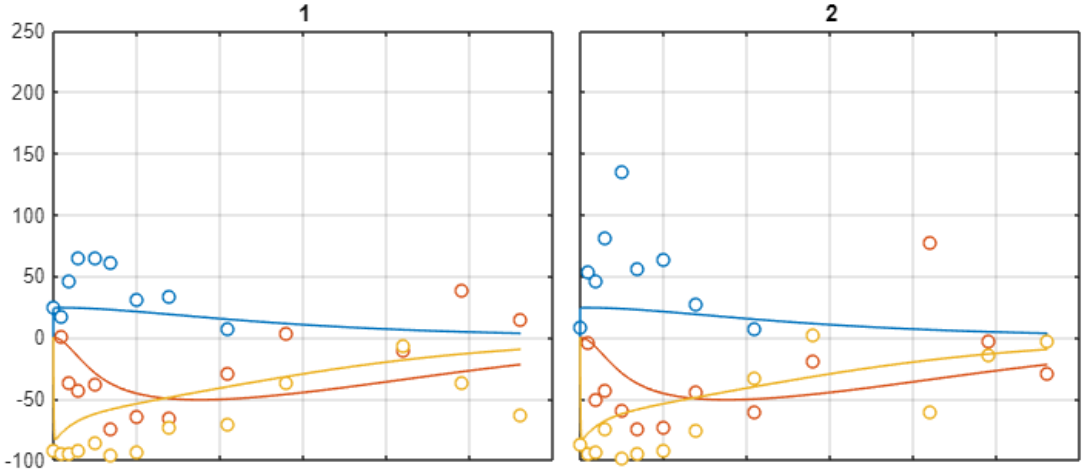
MESSAGES

MESSAGES

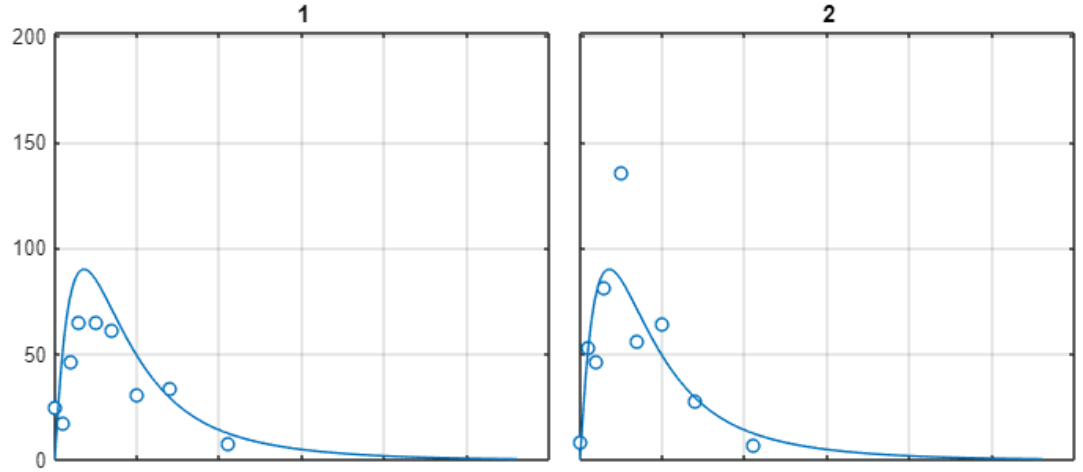
1 Optimization cannot proceed because the initial value of the negative log-likelihood is NaN. This can occur when the initial parameter estimates are extremely poor or when selecting an error model that is inappropriate for your data. Ensure that initial parameter estimates result in reasonable simulation results. When using the proportional error model, ensure that simulation results are never zero. When using the exponential error model, ensure that simulation results are always positive.

Pitfall 4
Objective function not well-posed for parameters being estimated

- Example: Objective function takes into account three response types (anti-PCSK9, PCSK9 and LDLc) but only PK parameters are estimated
  - Because PK parameters are primarily associated with anti-PCSK9 response, the model does not have degrees of freedom to get a good fit with all three responses
  - The PCSK9 and LDLc responses penalize the fit unnecessarily
- Solution: “it depends” remove PCSK9 and LDLc responses or add weights



All three responses



Only anti-PCSK9 response

Pitfall 5:
Ambiguity in observations at time of dose

ID	Time	Dose_IV	Conc_central
1	0	10	0
1	2		3.1
1	5		1.3
1	10		0.4
2	0	100	0

Was the concentration in the central compartment 0 before the dosing or after the dosing happened?

ID	Time	Dose_IV	Conc_central
1	0		0
1	0.1	10	
1	2		3.1
1	5		1.3
1	10		0.4

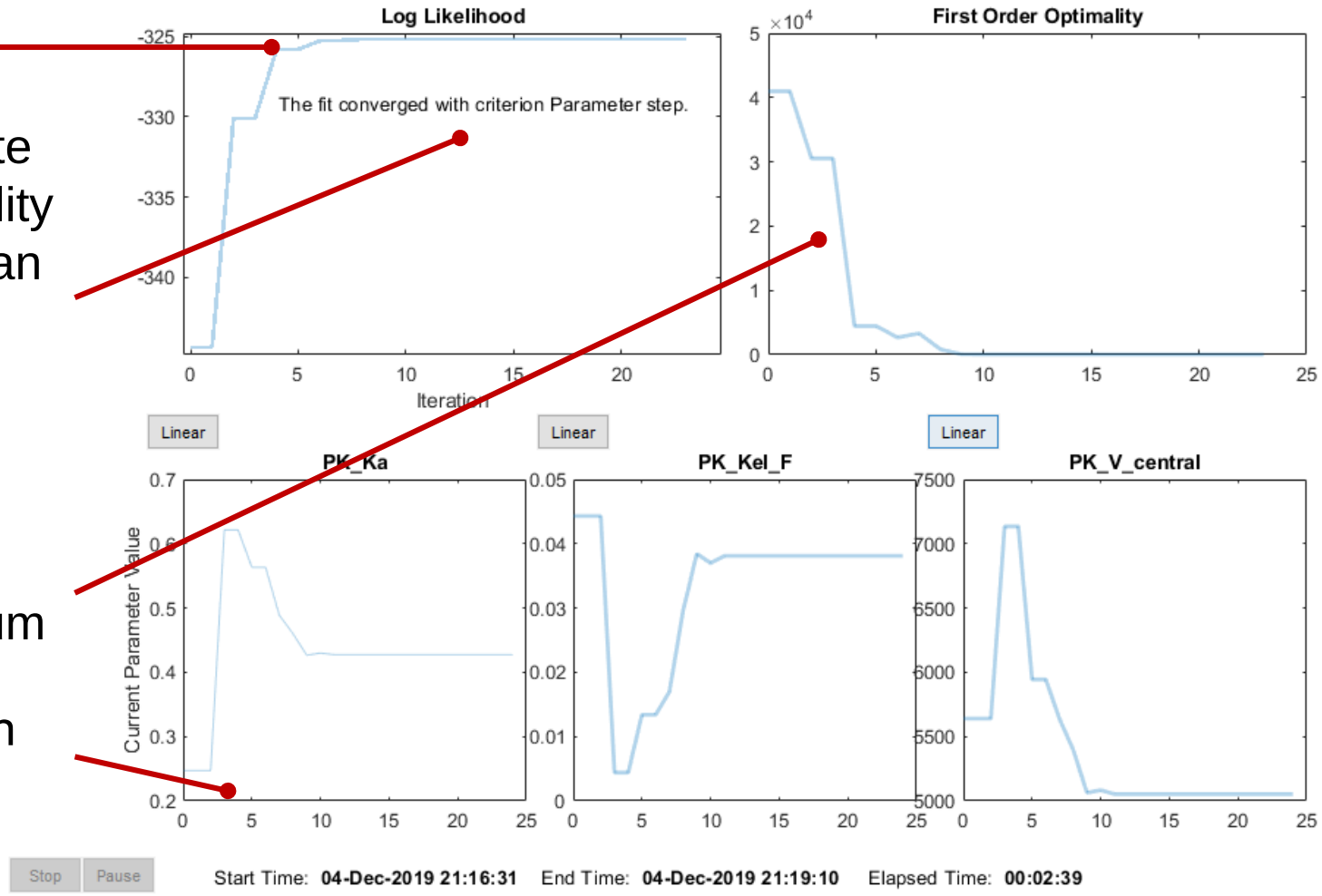
Progress plot

Maximize Log Likelihood

Aim for the optimization to terminate on log likelihood, first order optimality and/or parameter values, rather than maximum number of iteration.
Max iterations is no guarantee the optimization found a maximum!

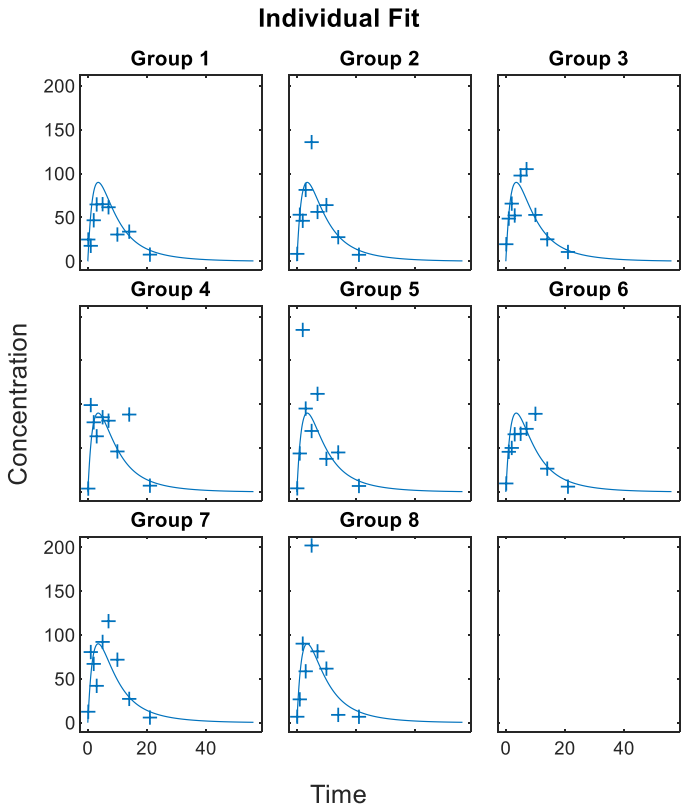
Minimize gradient at solution: the derivative should be 0 at a maximum

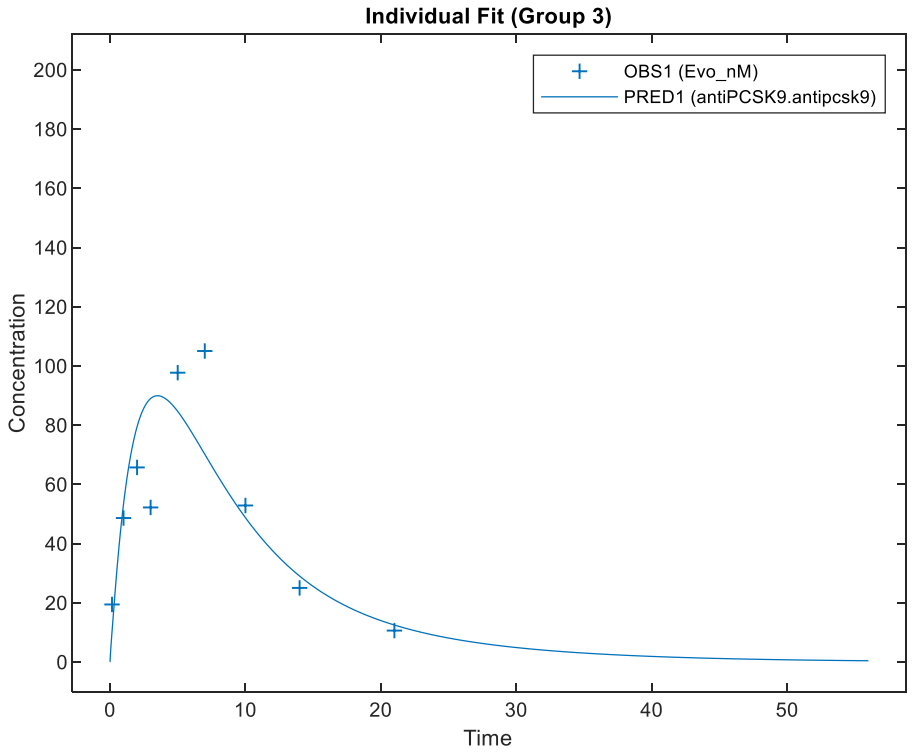
Ensure the optimization doesn't run into upper or lower bounds



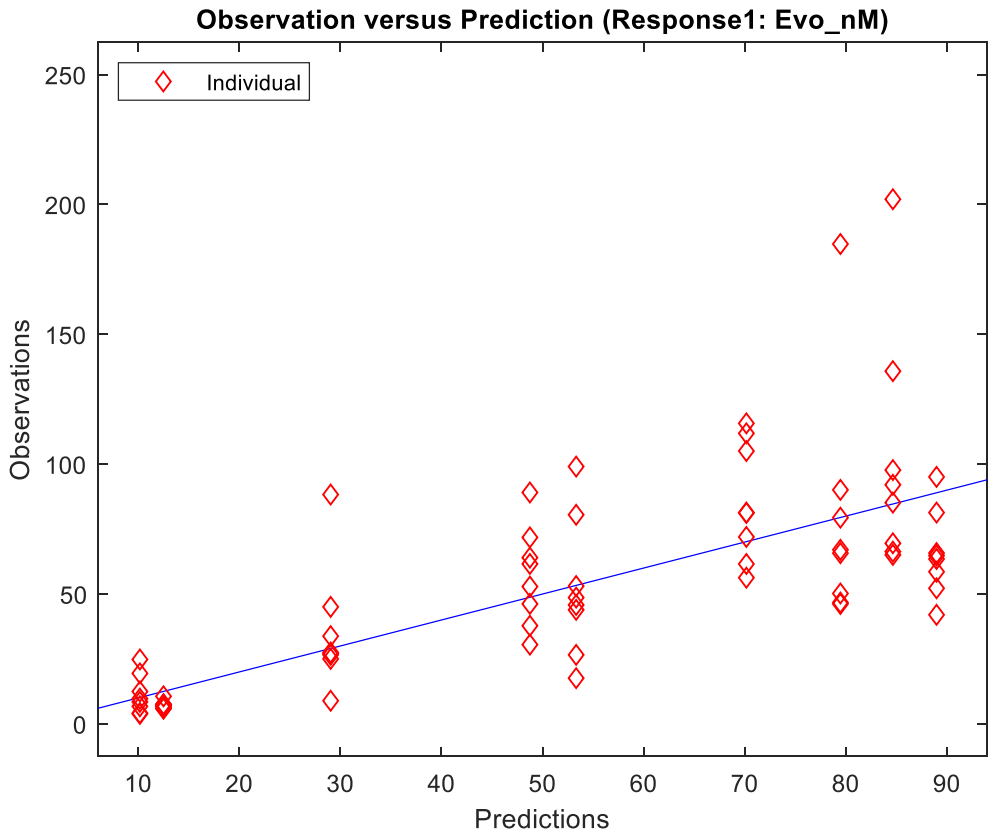
Fit plot

- Did the final fit represent the data well?
- What can we do beyond eye-balling?



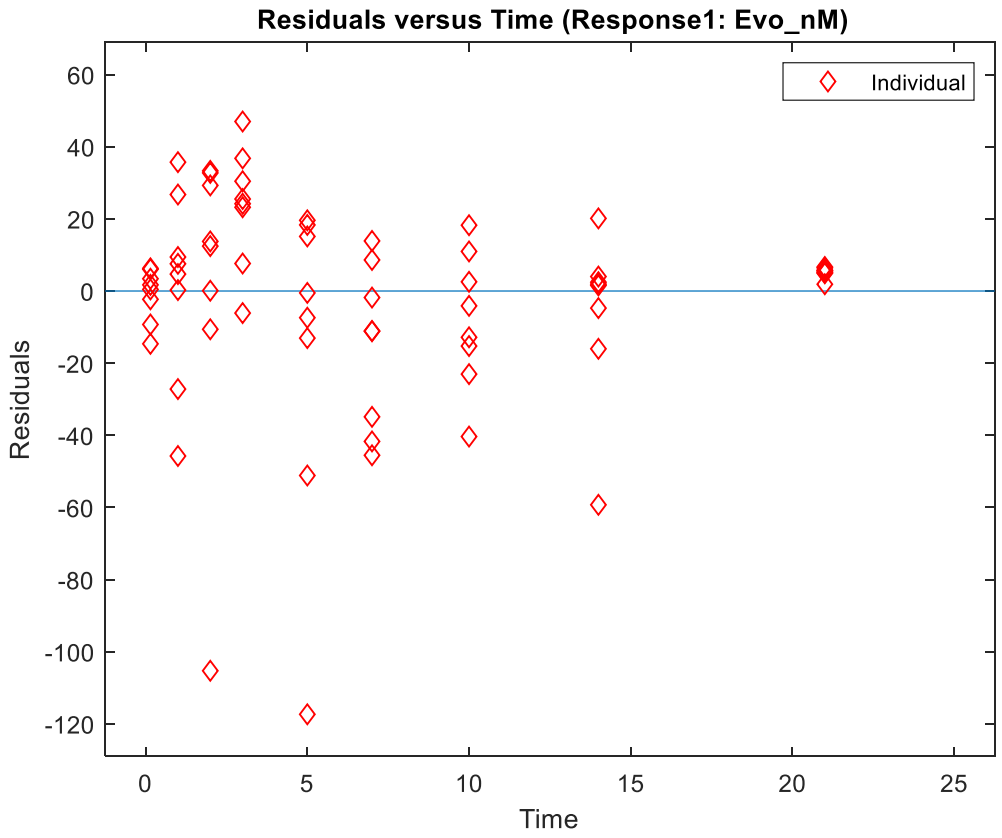


Observation vs prediction plot



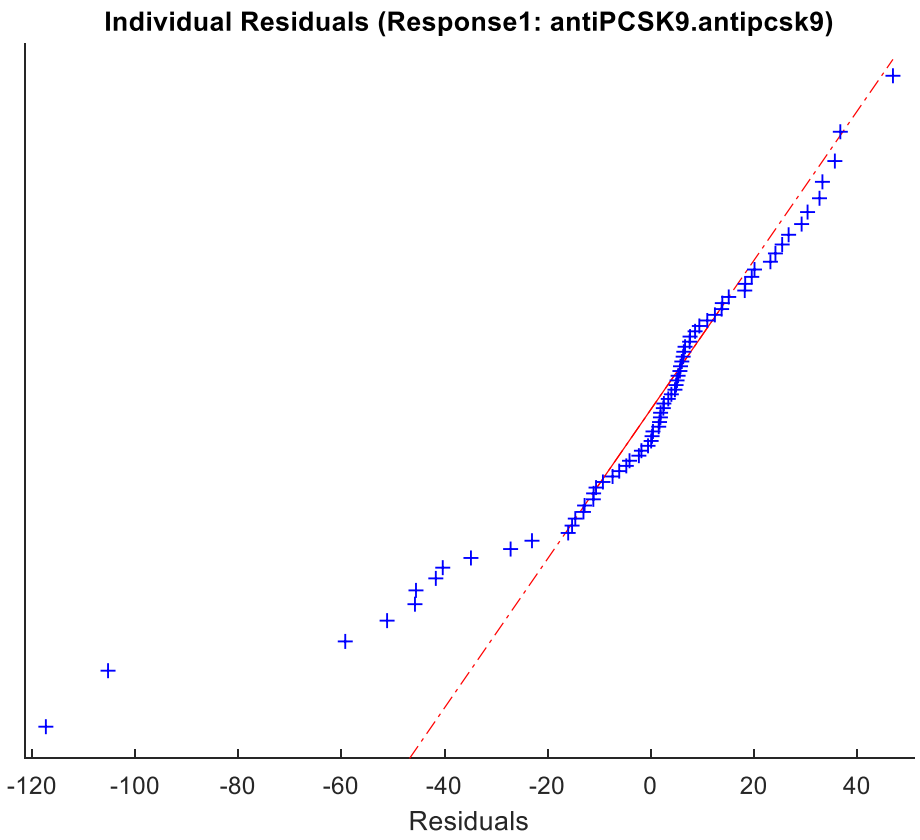
- Blue line is the unity line (y = x)
- For perfect predictions, all observations would coincide with the unity line
- In other words: there is no residual

Residuals vs time



- The observations should be equally distributed on either side of x-axis
- Throughout the simulation

Residual distribution

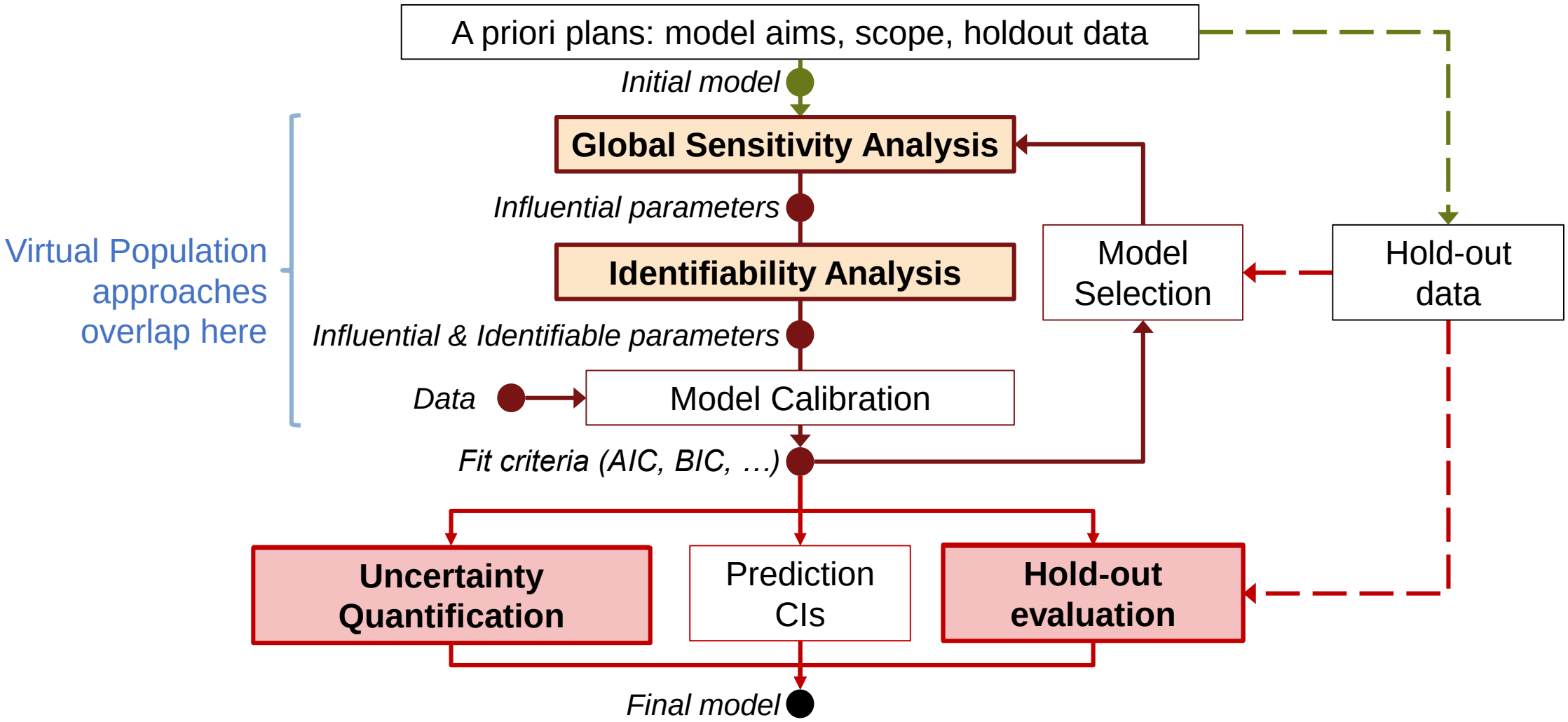


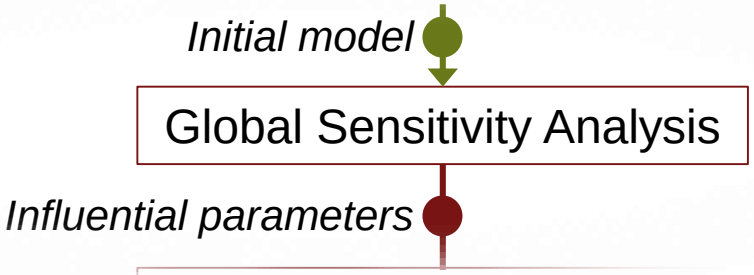
- Quantile-Quantile plot
- Compare residual distribution with normal distribution
- If the points lie mostly along the line, you can assume the residuals are approximately normally distributed

Agenda: Parameter Estimation

- Central concepts
  - Objective function
  - Optimization
- Mechanics of parameter estimation
  - Setting up an optimization
  - Pitfalls
- Workflow to guide estimation process
  - Sensitivity analysis
  - Identifiability analysis
  - Confidence intervals

General workflow for parameter estimation





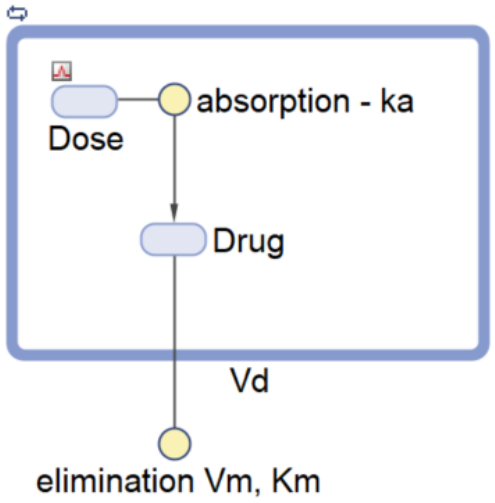
Sensitivity Analysis

Concepts in Sensitivity Analysis

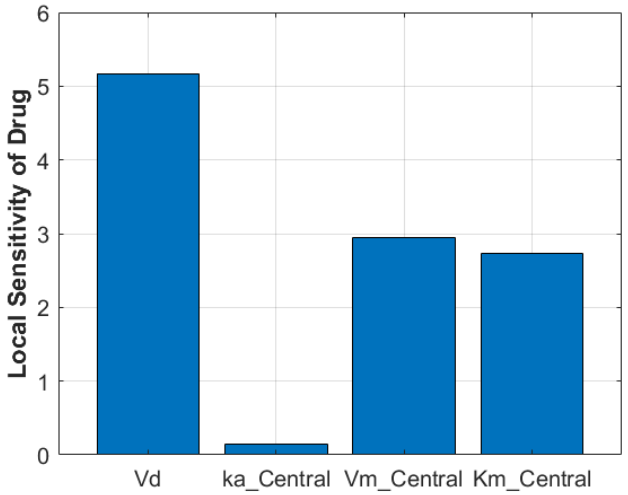
- Local or Global
  - Local: Analysis at a single point in parameter space; assumes correct calibration
  - Global: Analysis across a user-specified domain in parameter space
- Sampling
  - One-at-a-time: change one parameter at a time to calculate sensitivity index
    - Unable to observe interactions between parameters from single analysis
  - All-at-a-time: take random samples from parameter space to calculate sensitivity index
    - Observe interactions between parameters
    - Multiple methods to sample parameter space – Sobol, Latin hypercube, uniform
- Derivative, variance, or correlation based

Why use Global Sensitivity Analysis?

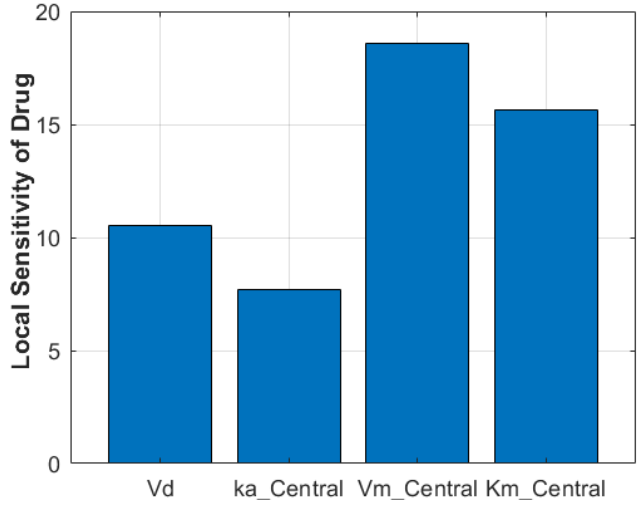
Example: 1-compartment model, oral dosing, enzymatic clearance



Local Sensitivity Analysis 1



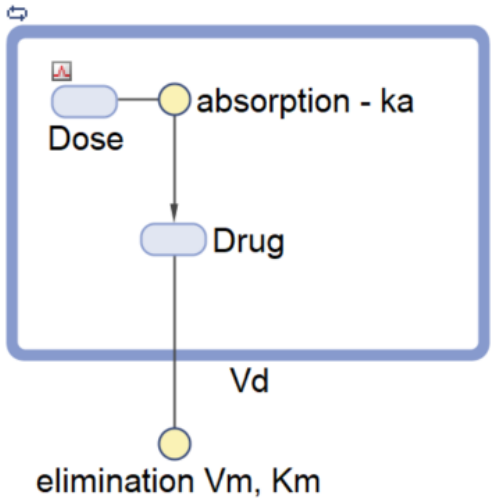
Local Sensitivity Analysis 2



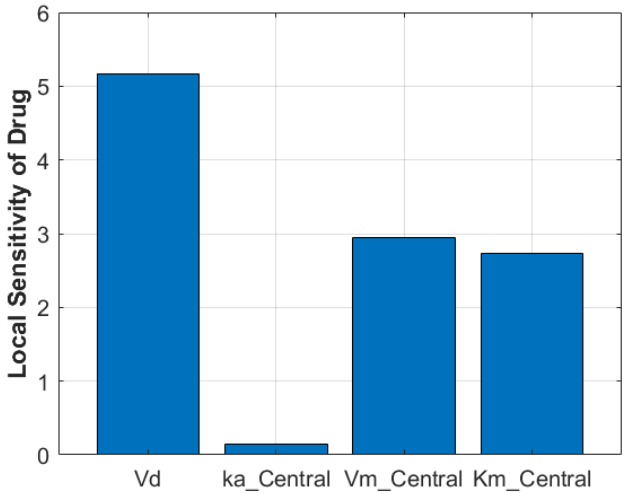
Which analysis will you base your modeling decisions on?

Why use Global Sensitivity Analysis?

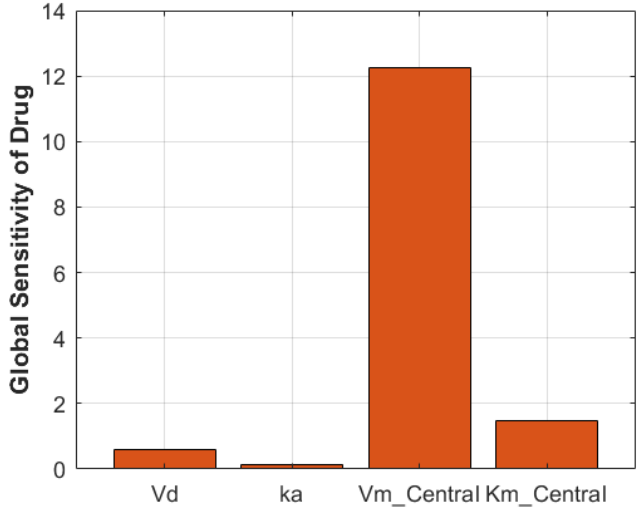
Example: 1-compartment model, oral dosing, enzymatic clearance



Local Sensitivity Analysis 1



Global Sensitivity Analysis (Sobol)



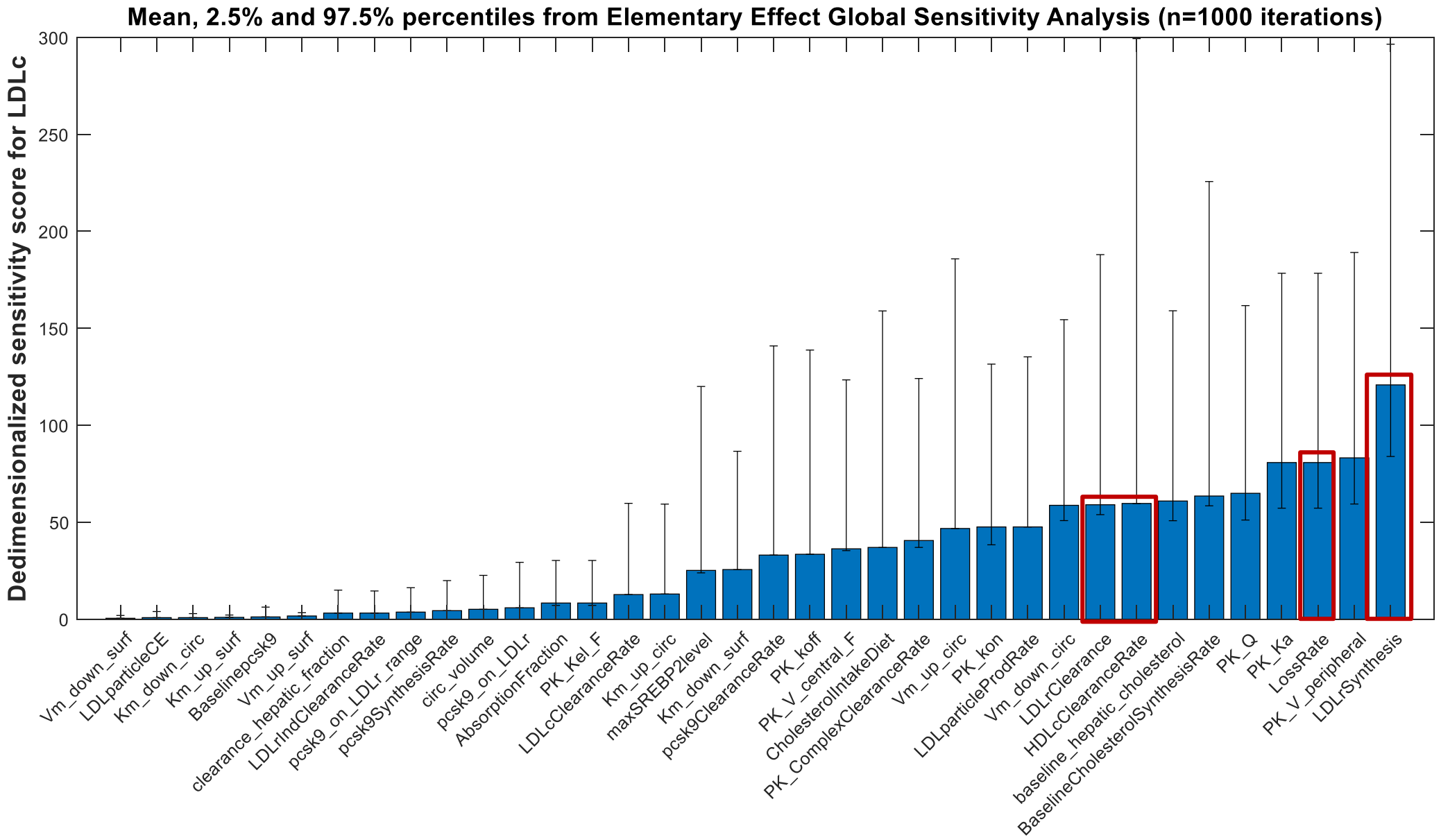
Use GSA: fewer assumptions, explores full input parameter domain

Global Sensitivity Analysis methods

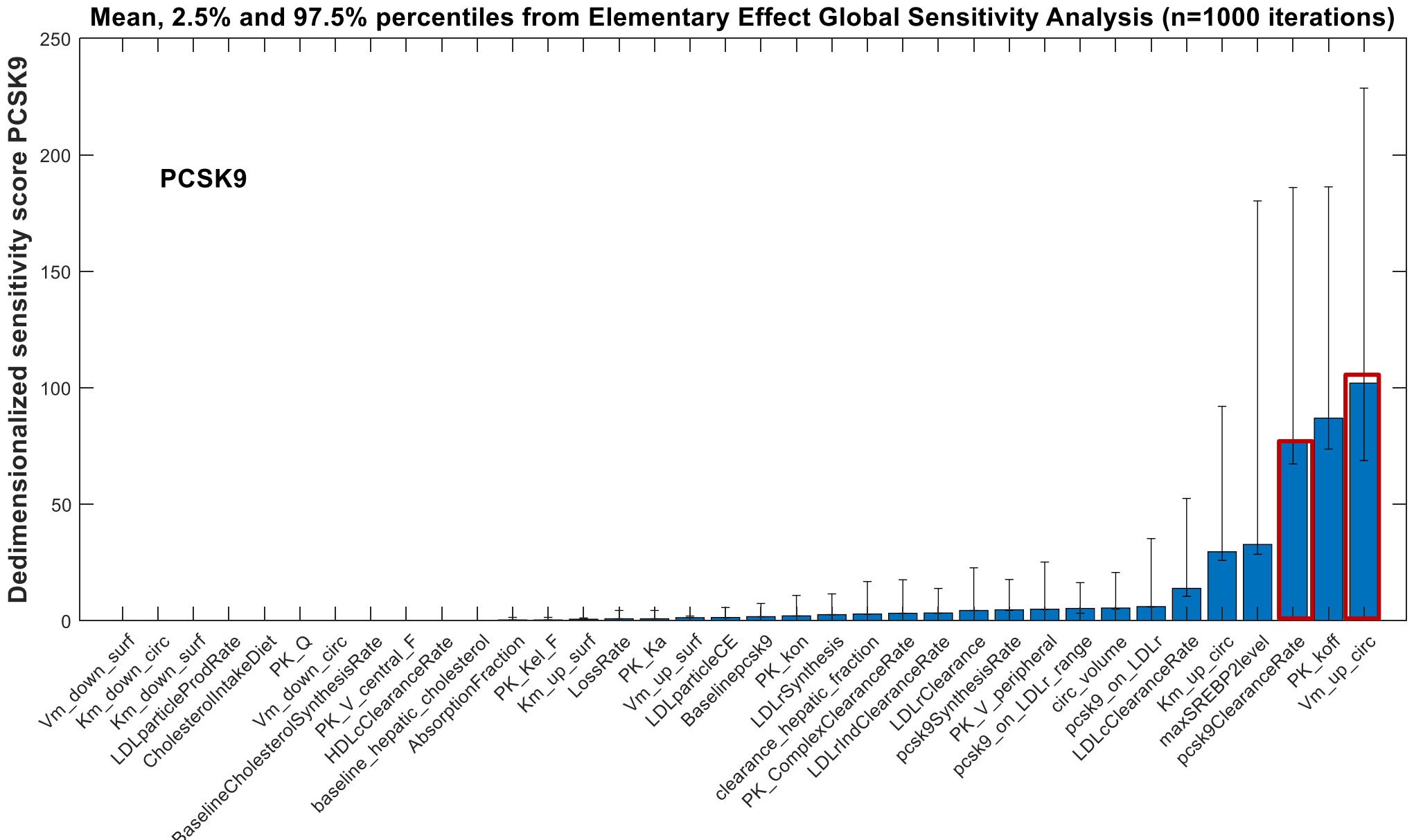
Method	Assumptions/Limitations	Approximate computational expense
Morris/Elementary Effect/ Weighted Average of Local Sensitivities Methods	Monotonicity, linearity, one-at-a-time	>M*10 M: number of parameters under investigation
Partial Rank Correlation Coefficient (PRCC)	Monotonicity in parameters, limited correlation between parameters, robust for nonlinear models	>M*100
Variance-based: Sobol indices, (extended) Fourier Amplitude Sensitivity Test	Very few. Variance is a good statistic to represent model output distribution; not appropriate when output distribution is highly skewed or multimodal	>M*1000

Pianosi et al. (2016) Env Modelling and Software
Zhang et al. (2015) CPT:PSP
Tiemann et al. (2013) PLOS Comp Bio
Zi (2011) IET Sys Bio
Marino et al. 2008) Journal of Theoretical Biology

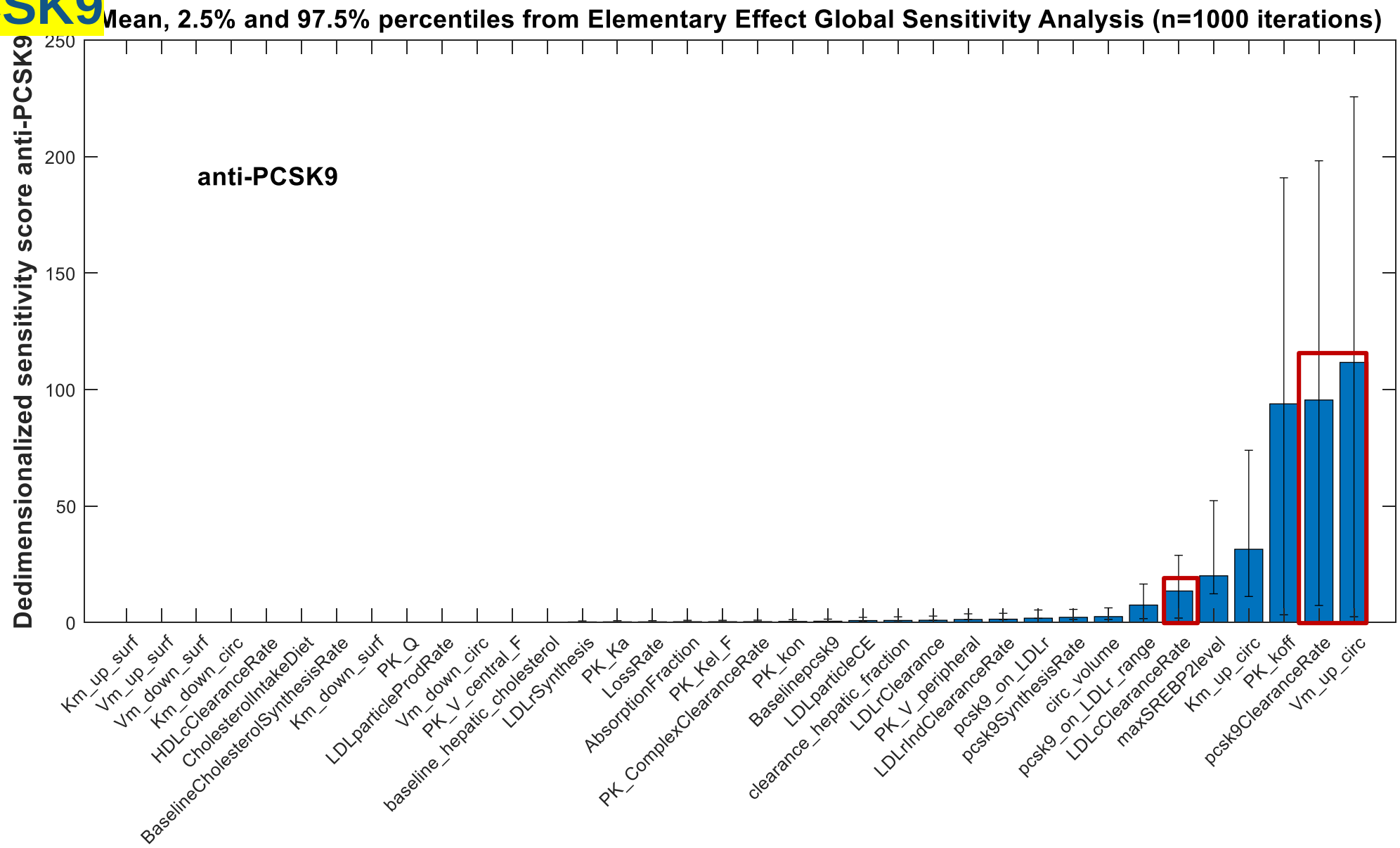
Rank-ordered sensitivities from Elementary Effects Method for LDLc



Rank-ordered sensitivities from Elementary Effects Method for **PCSK9**



Rank-ordered sensitivities from Elementary Effects Method for anti-PCSK9



Sobol indices

- Apportion variance in model output (Y) to model inputs (X)
- First-order sensitivity index: individual contributions

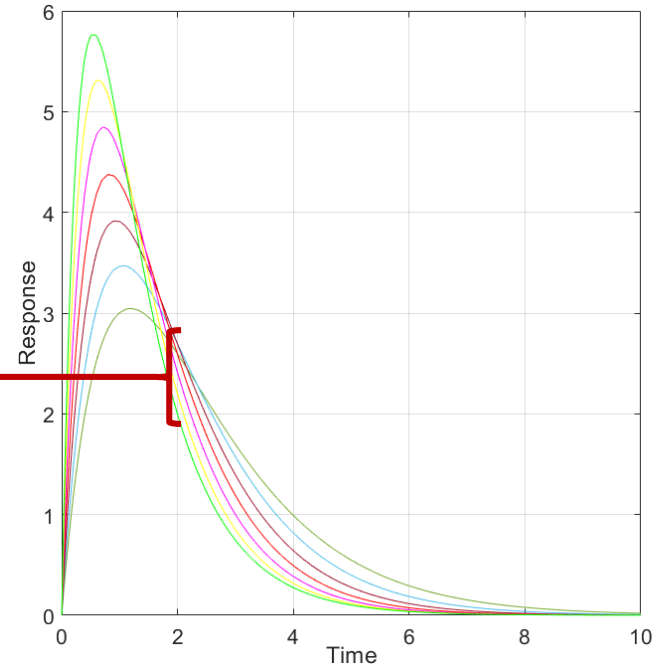
$$S_i = 1 - \frac{E_{X_i}\left(Var_{X_{\sim i}}(Y|X_i)\right)}{Var(Y)}$$

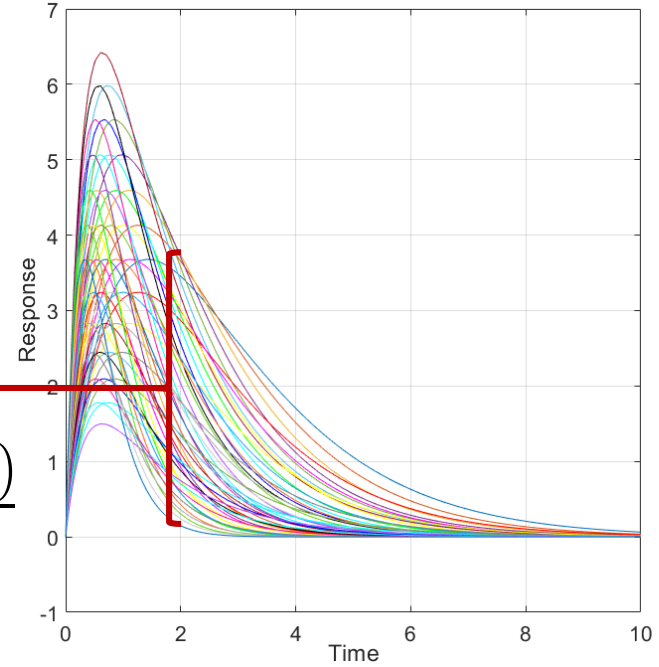
Variance not due to X_i

Total variance

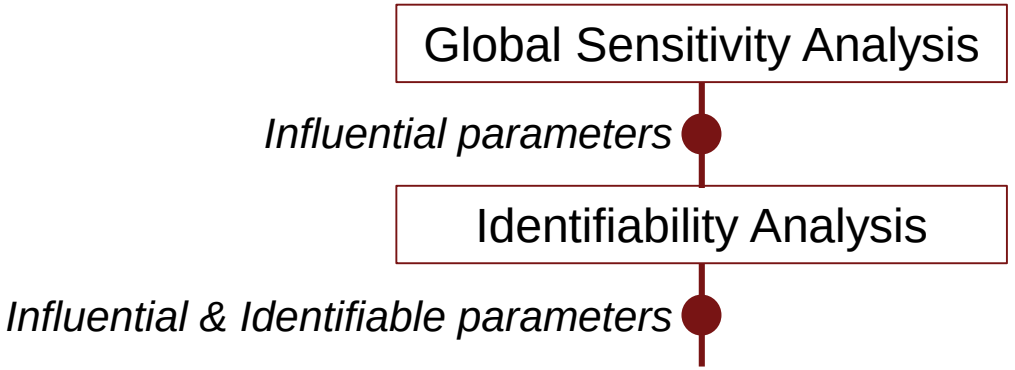
- Total effect: shows interactions

$$S_{T,1} = S_1 + S_{12} \cdots S_{1,2,\cdots,n} = 1 - \frac{Var_{X_{\sim 1}}\left(E_{X_1}(Y|X_{\sim 1})\right)}{Var(Y)} = \frac{E_{X_{\sim 1}}\left(Var_{X_1}(Y|X_{\sim 1})\right)}{Var(Y)}$$



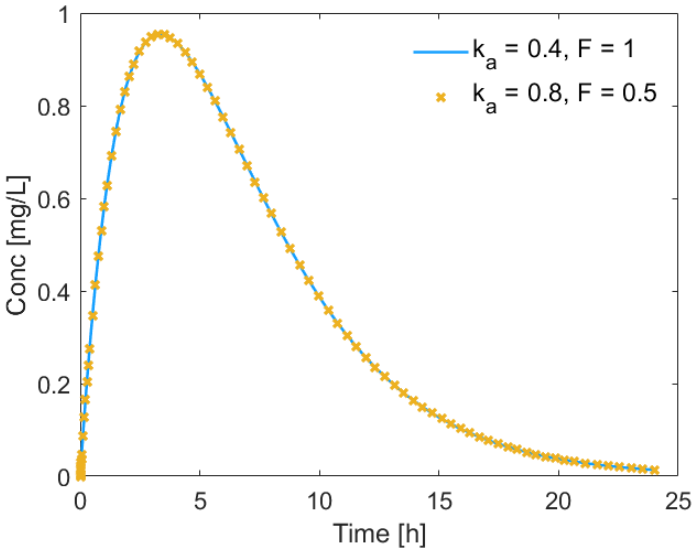


Saltelli et al. (2008) Global Sensitivity Analysis – The Primer, Wiley



Identifiability Analysis

Concepts in Identifiability



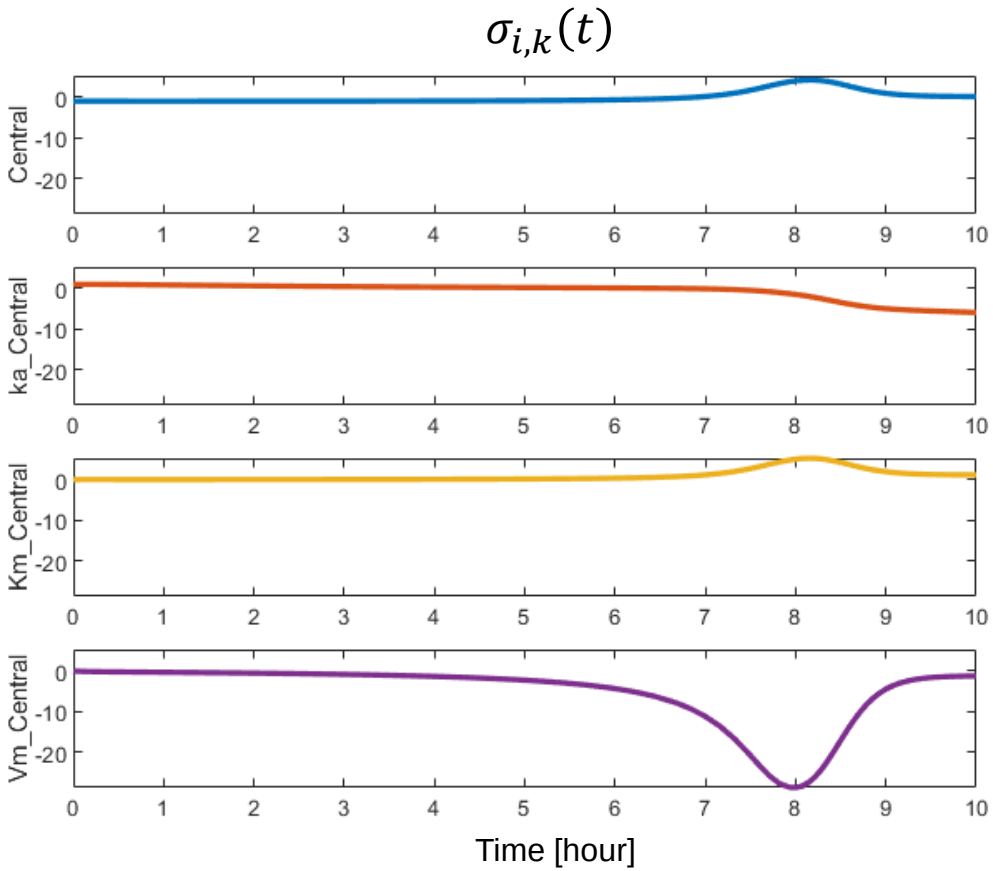
- Types
  - Structural identifiability, e.g., F·ka or parallel pathways
  - Practical identifiability, e.g., inadequate data to constrain
  - Note: identifiability is NOT “my runs converged”
- Structural identifiability computations
  - GenSSI: Generating Series for Structural Identifiability
  - COMBOS: identifiable parameter combinations using Groebner bases
  - DAISY: Differential Algebra for Identifiability of Systems
- Practical identifiability computations
  - Mostly sampling-based (e.g., profile likelihood, MCMC)

Aliasing score (1)

- Definition:

1. Start with local sensitivities:

   $\sigma_{i,k}(t) = \frac{\partial R_i(t)}{\partial p_i}$, where p_i is the i[th] parameter and R_k is the k[th] response in the model



Aliasing score (2)

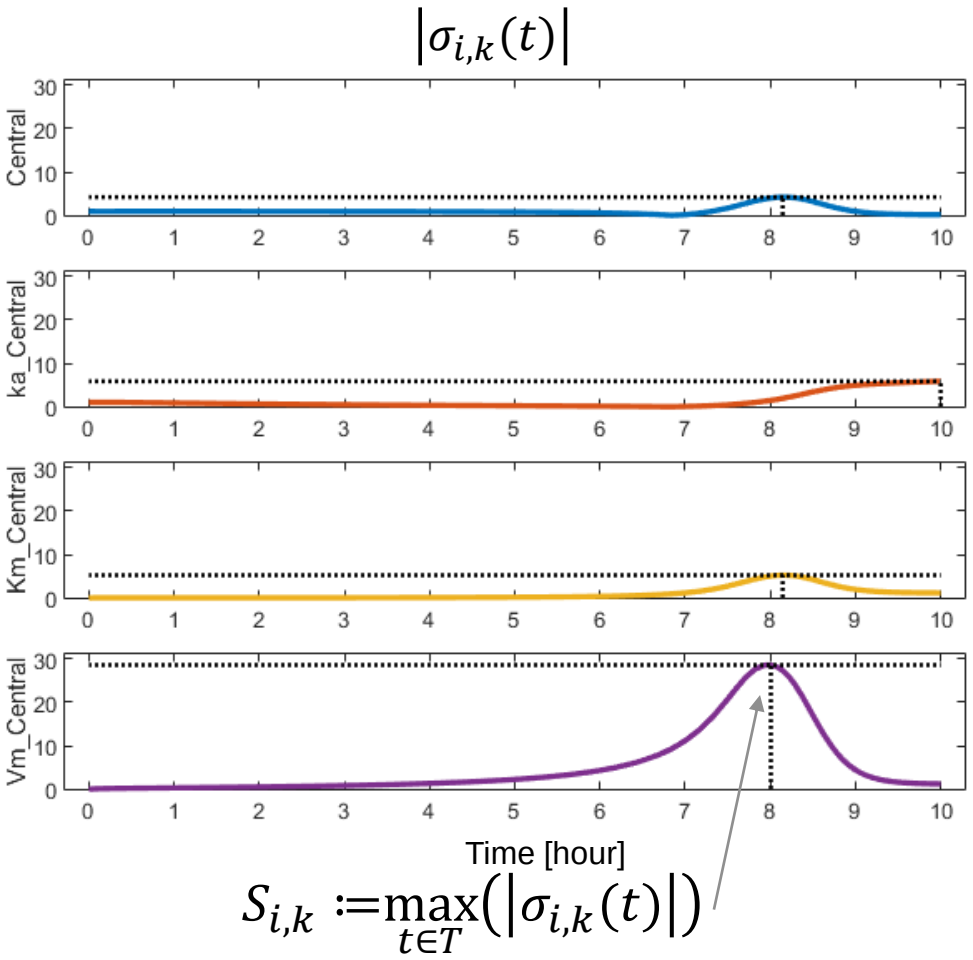
- Definition:

1. Start with local sensitivities:

   $\sigma_{i,k}(t) = \frac{\partial R_i(t)}{\partial p_i}$, where p_i is the i[th] parameter and R_k is the k[th] response in the model

2. Calculate the maximum value of $\left|\sigma_{i,k}(t)\right|$

   $S_{i,k} := \max_{t \in T}\left(\left|\sigma_{i,k}(t)\right|\right)$



Aliasing score (3)

- Definition:

3. Normalize local sensitivities

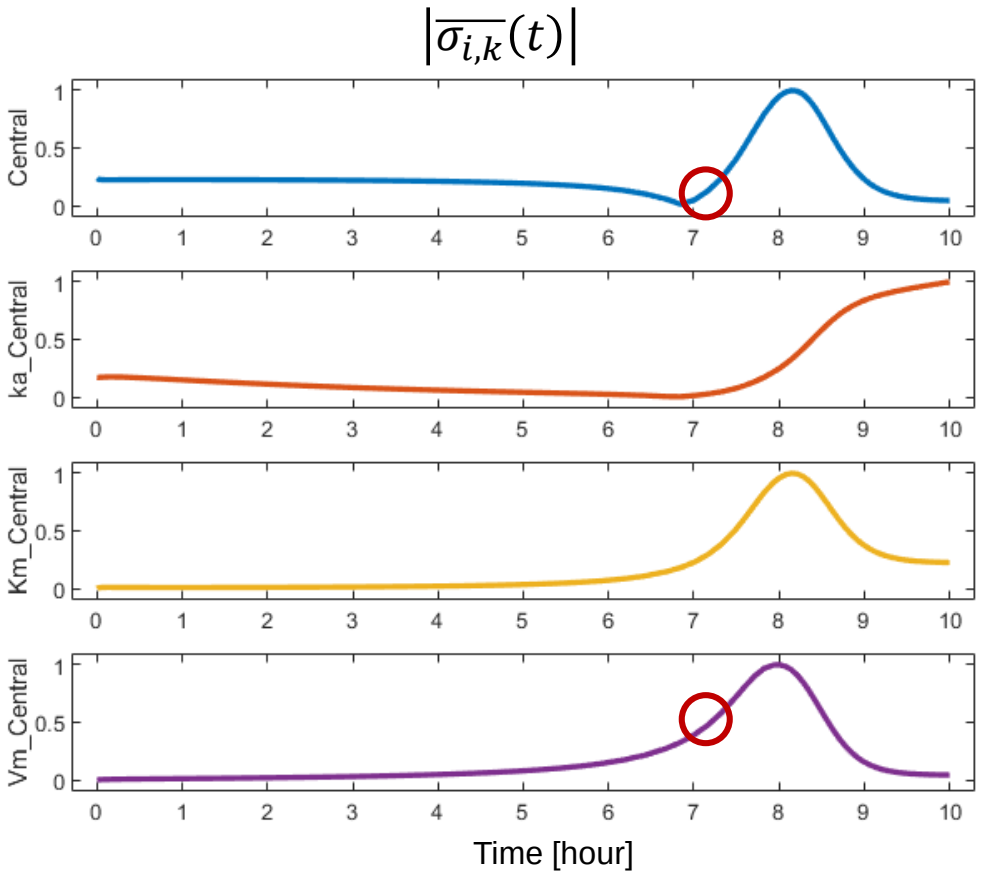
$$\overline{\sigma_{i,k}}(t) := \frac{\sigma_{i,k}(t)}{S_{i,k}}$$

4. Calculate pair-wise aliasing metric

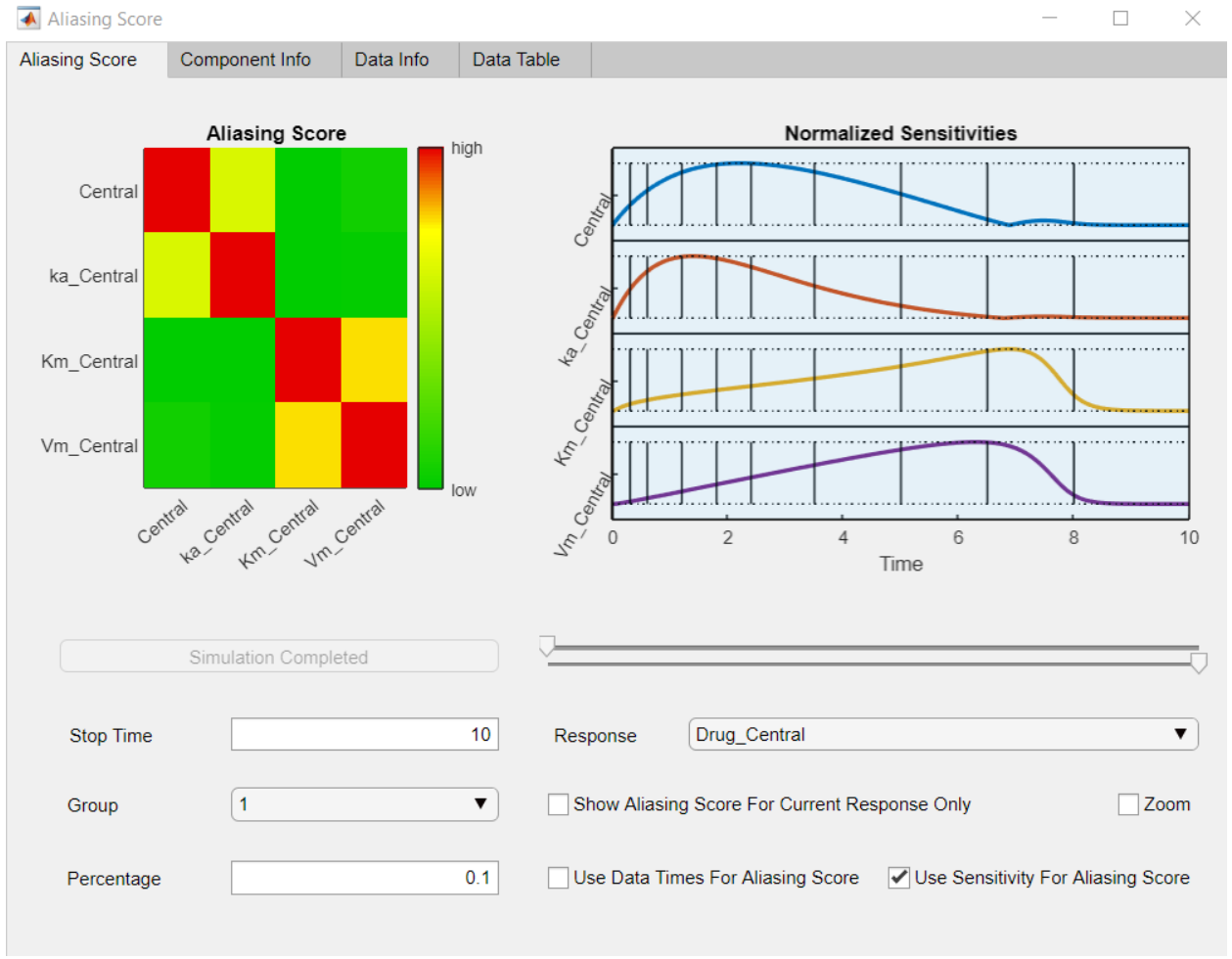
$$\alpha_{i,j}^{k} := \max_{t \in T} \left| \left| \overline{\sigma_{i,k}}(t) \right| - \left| \overline{\sigma_{j,k}}(t) \right| \right|$$

5. Transform to aliasing score

$$A_{i,j}^{k} := 100 \cdot \max\{1 - \alpha_{i,j}^{k}, 0\}$$

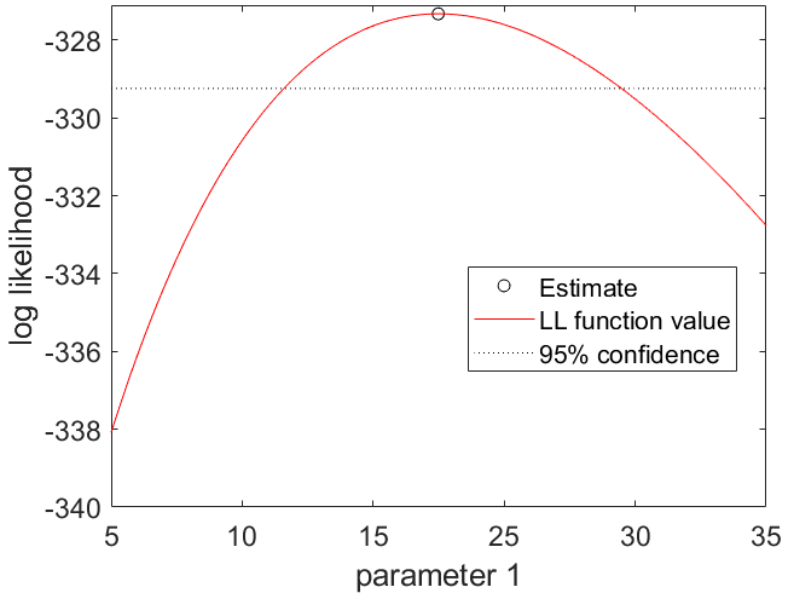


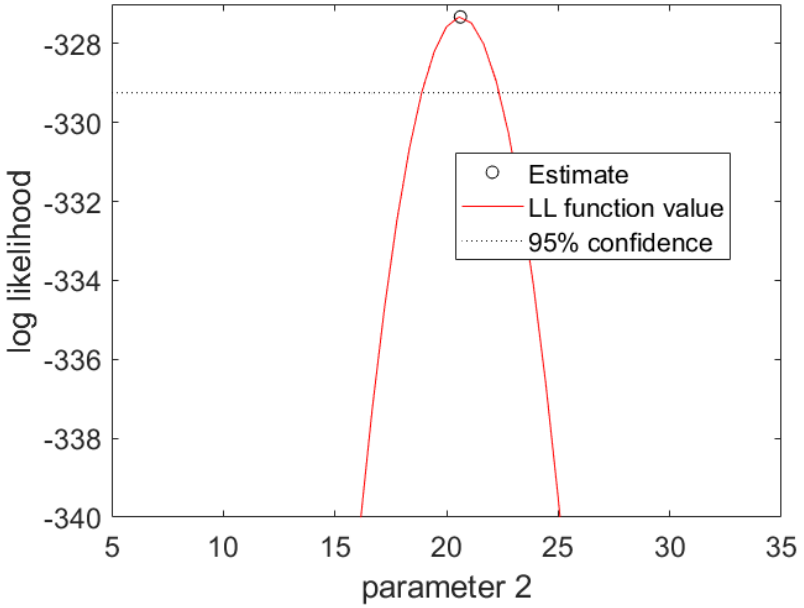
Working with the Aliasing Score



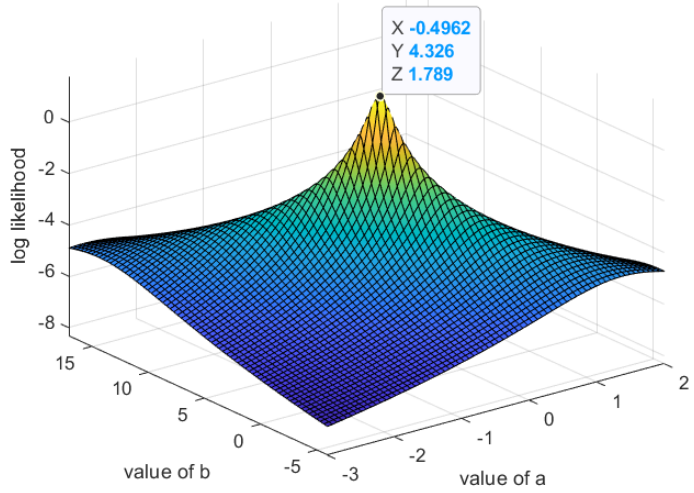
https://github.com/mathworks-SimBiology/AliasingScoreApp

Practical identifiability – profile-likelihood

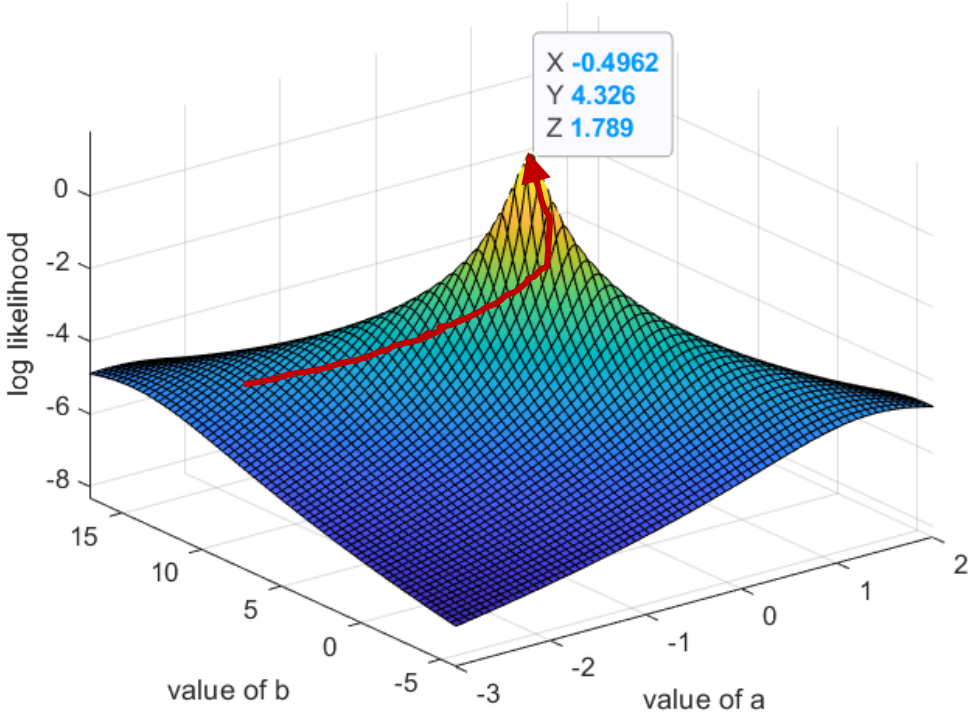


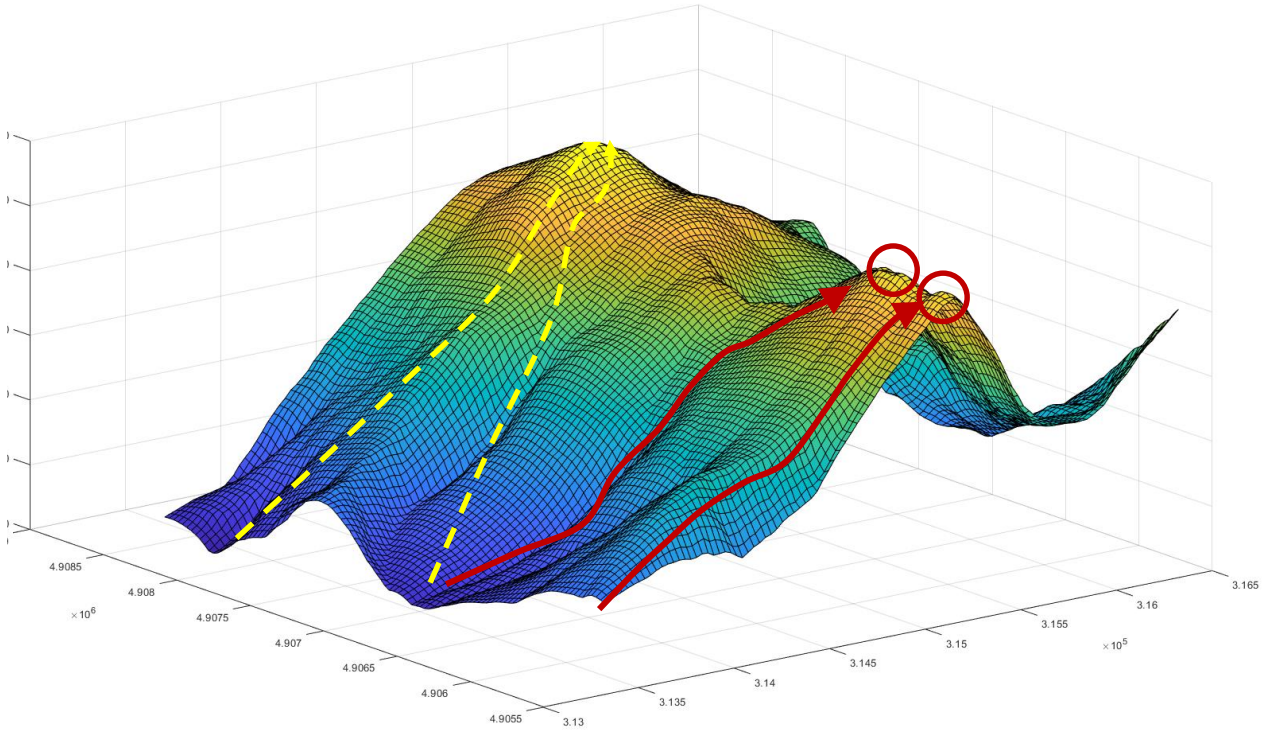


- One-at-a-time sampling of each parameter across its domain to calculate log-likelihood
- If log-likelihood is very flat and wide, the parameter is practically non-identifiable
- If log-likelihood function is sharp and narrow, the parameter is practically identifiable



Local vs Global Optimization

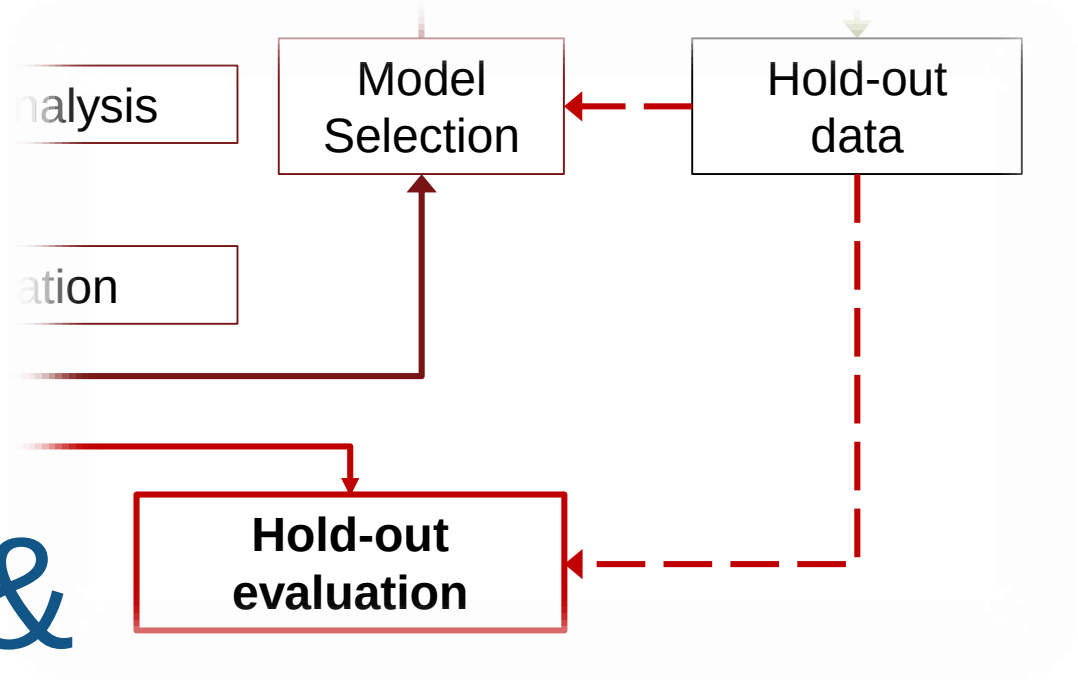




There is only one way to the top …
… because there is only one top

But what if there were local maxima (red)?
This is where global optimization can help

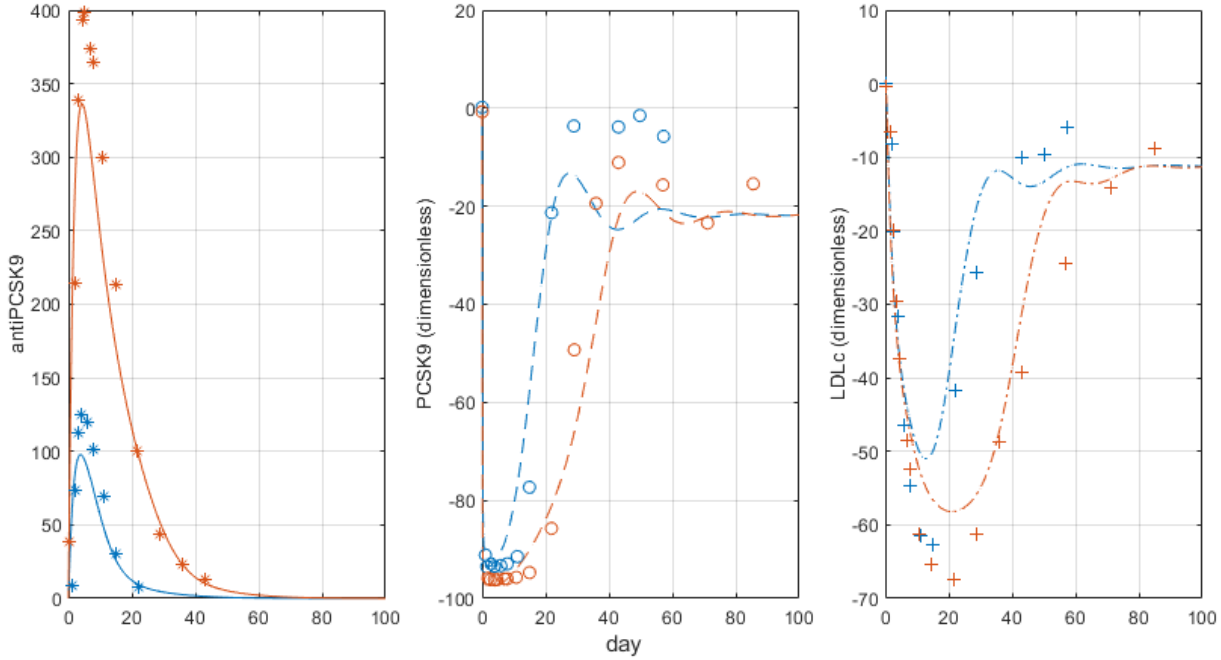
Note: global optimization will not necessarily solve identifiability problems!



Hold-out data & Uncertainty Quantification

Hold-out data can help gain confidence in model predictions

- Hold-out data: data the was never used to calibrate the model but that the model should be able to predict
  - In this case: data from Kasichayanula et al.
- Comparing predictions from the calibrated model to hold-out data
- Consider:
  - Re-estimating non-mechanistic PK parameters for hold-out data to account for differences between studies
  - Then test whether mechanistic PD model is able to predict hold-out results



Uncertainty Quantification

- Prediction confidence intervals = “This set of parameter estimates is a good approximation of the true parameter values that generated the data”
  - Assuming the model is able to represent the data
  - Post-hoc calculation: Gaussian approximation assumes model is linear in parameters (different from a linear ODE)
  - …or use bootstrapping – sampling approximation of confidence intervals (expensive)
- Monte Carlo approach (overlaps with Virtual Population approach)
  - Sample parameter space
  - Simulate model for each sample
  - Investigate statistic of model response, e.g. 95% quantile region in simulated AUC

Agenda: Parameter Estimation

- Central concepts
  - Objective function
  - Optimization
- Mechanics of parameter estimation
  - Setting up an optimization
  - Pitfalls
- Workflow to guide estimation process
  - Sensitivity analysis
  - Identifiability analysis
  - Confidence intervals
- Why did we do this?

Why did we do this again?

With my calibrated PCSK9 model, I can

- … go to the project team with this model and my colleagues can ask me to simulate hypothetical scenarios
- … have an appropriate level of confidence in model predictions
- … quantify the uncertainty
  - Parameter estimates
  - Predictions