# Statistics and Data Analysis in MATLAB

Lecture 6:
Model reliability

Kendrick Kay
Washington University in St. Louis
March 28, 2014

Error bars on model parameters via bootstrapping

Model bootstraps

Distribution of parameter estimates

True model
y = 0.5x + 2

Model to be fitted
y = ax + b

n = 10

Model bootstraps

Distribution of parameter estimates

True model
y = 0.5x + 2

Model to be fitted
y = ax + b

n = 200

Kendrick Kay, Washington Univ. in St. Louis
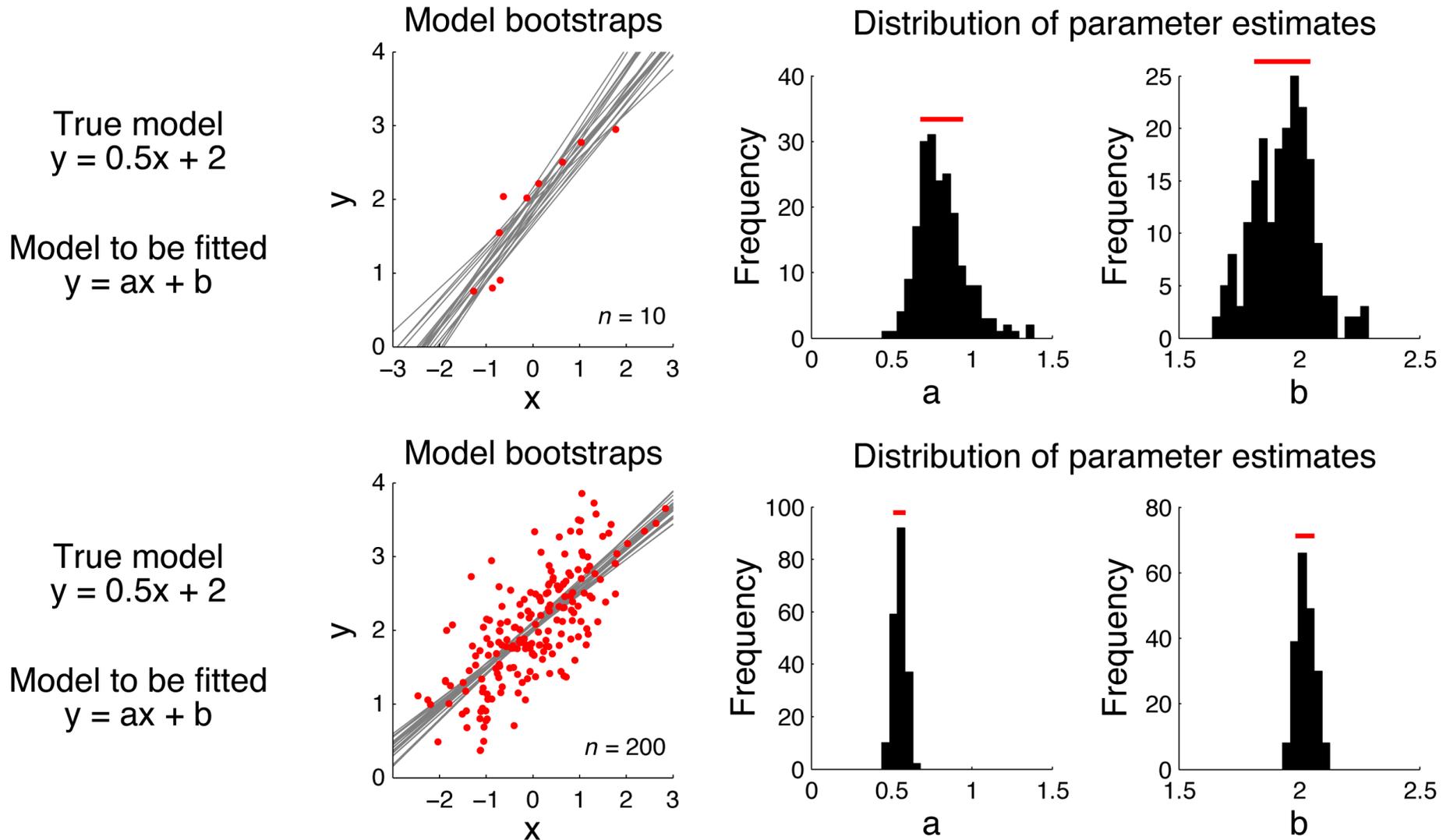
Correlated regressors yield correlated errors

Error surface

Distribution of parameter estimates

Model
$y = w_1 x_1 + w_2 x_2$

x_1 and x_2
uncorrelated

Error surface

Distribution of parameter estimates
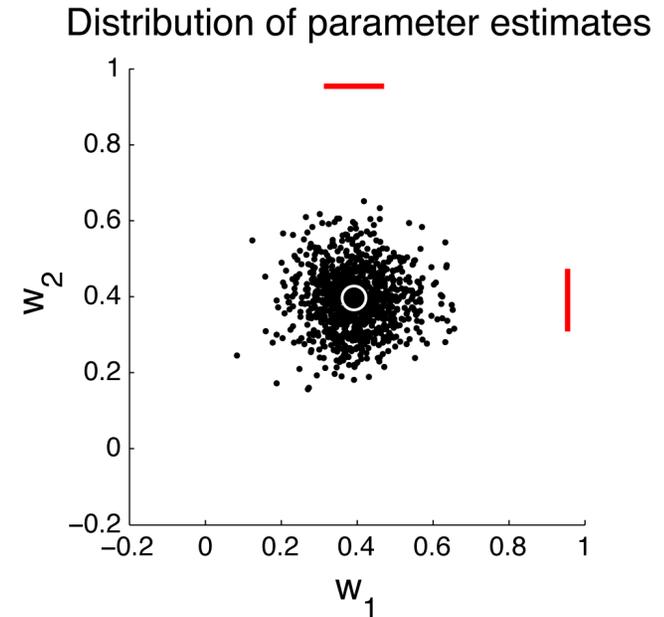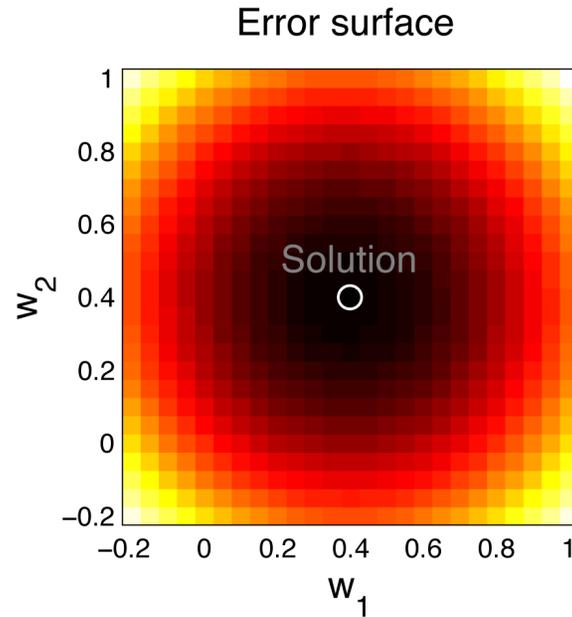
Model
$y = w_1 x_1 + w_2 x_2$

x_1 and x_2
correlated

Kendrick Kay, Washington Univ. in St. Louis
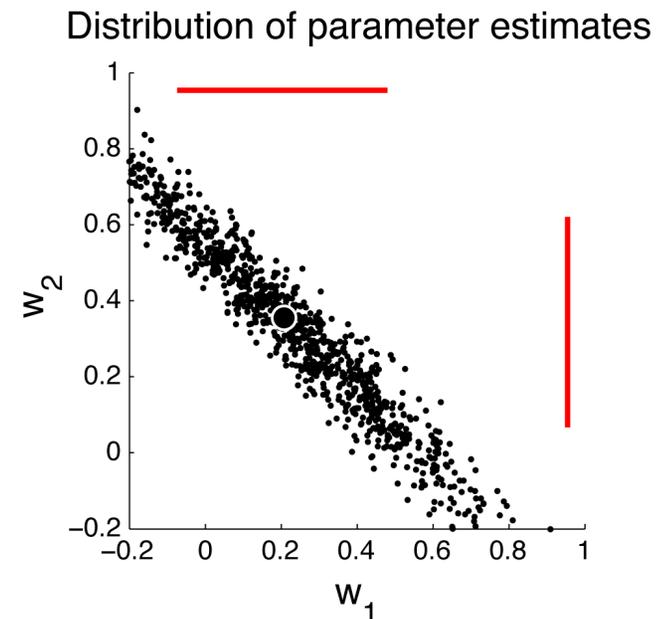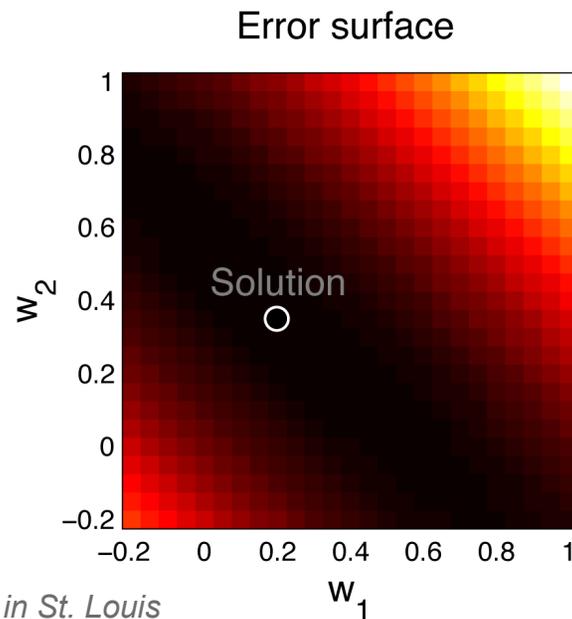
Distinction between accuracy and reliability

Kendrick Kay, Washington Univ. in St. Louis
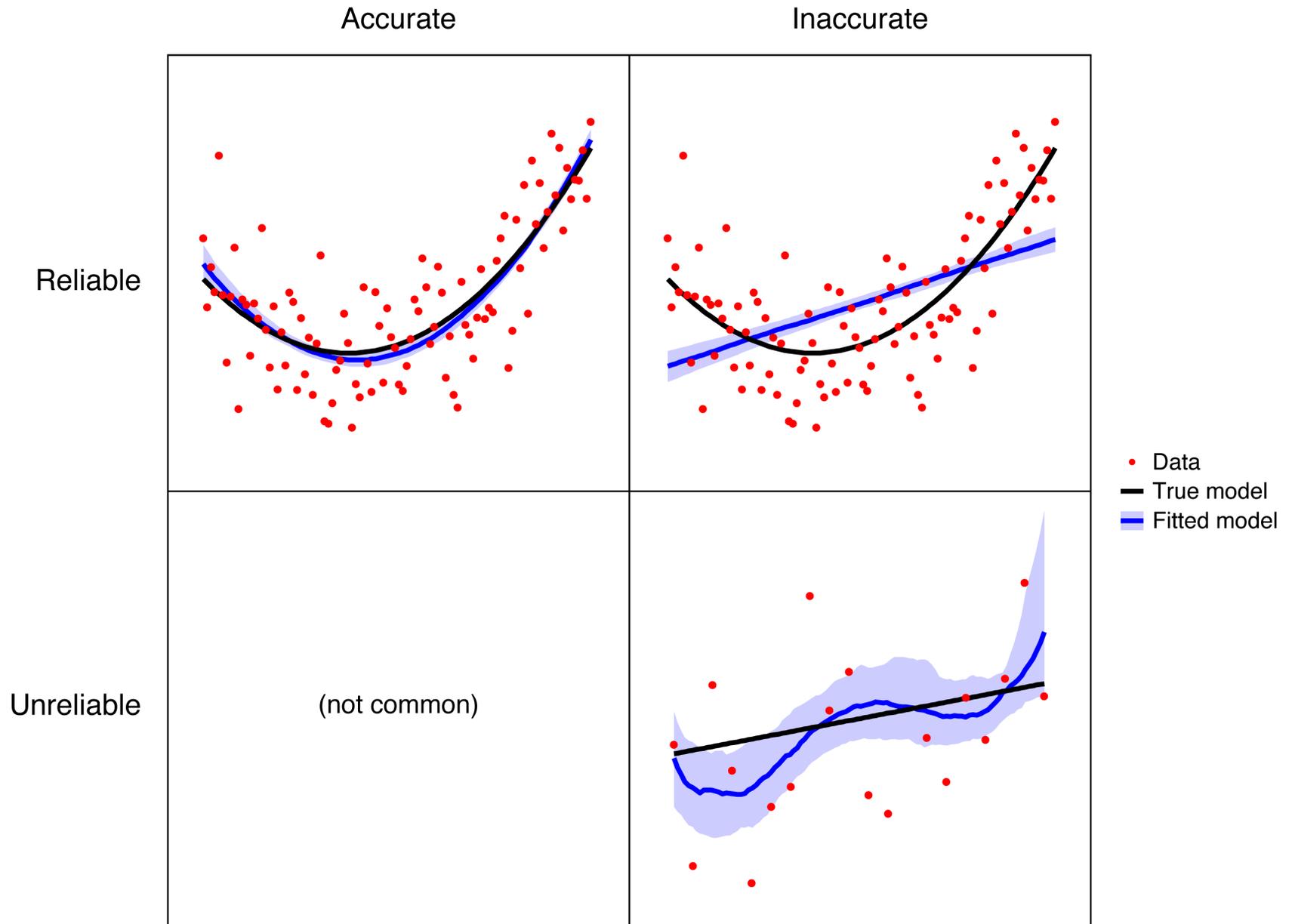
Bootstrapping and cross-validation applies to all models

Linear
(y = ax + b)

Linearized
(y = ax² + bx + c)

Parametric nonlinear
(y = axⁿ)

Nonparametric nonlinear
(Nearest-neighbor)
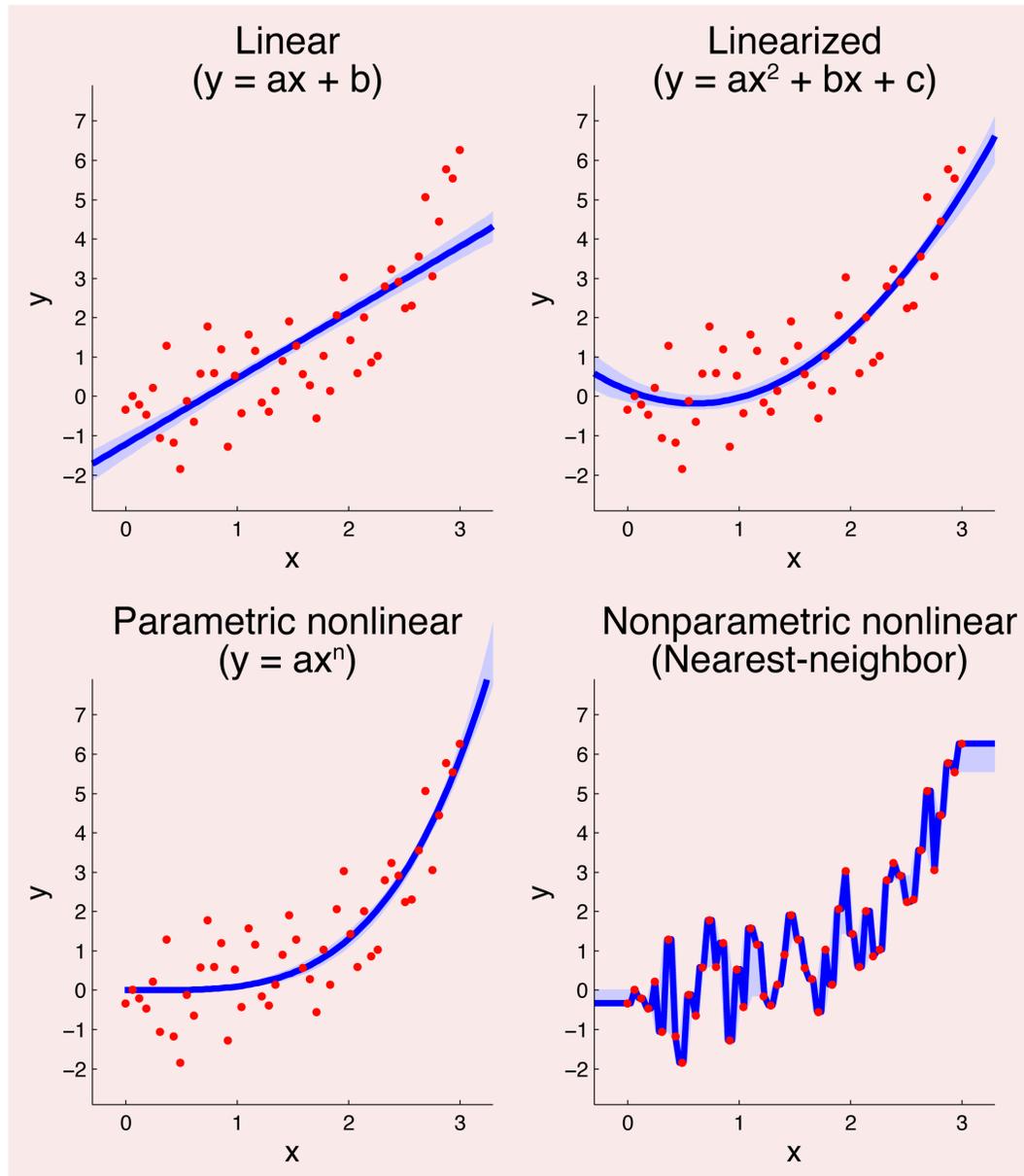
Use bootstrapping to estimate model reliability

Use cross-validation to estimate model accuracy

Cross–validated R^2

Model

y = ax + b

y = ax² + bx + c

y = axⁿ

Nearest-neighbor

Kendrick Kay, Washington Univ. in St. Louis