Bias-variance, denoising, and the value of simulations

Kendrick Kay

http://cvnlab.net

Center for Magnetic Resonance Research (CMRR)
University of Minnesota, Twin Cities

UNIVERSITY OF MINNESOTA

Overview

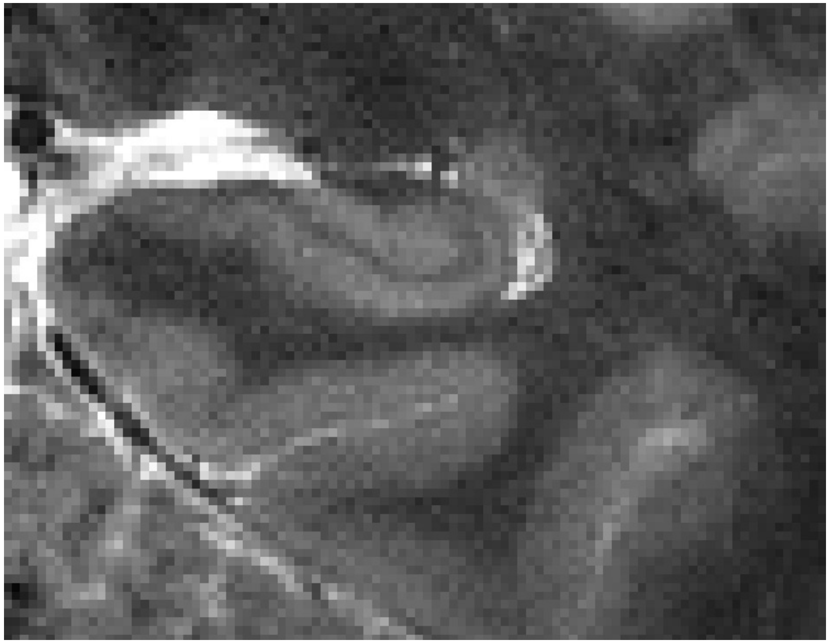
- Part 1: “The risk of bias in denoising methods” (Kay, *PLoS ONE*, 2022)

- Part 2: The value of simulations

Part 1: The risk of bias in denoising methods

Introduction

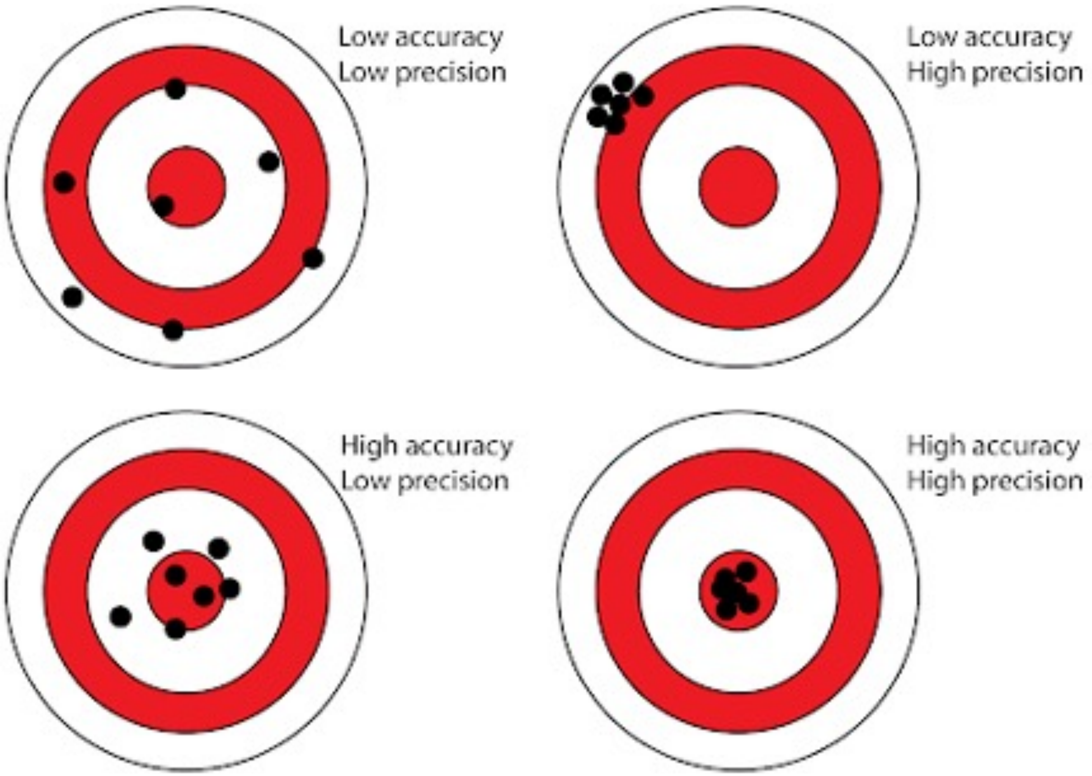
- MRI data can be noisy...



- Can’t we just ‘denoise’ the data?
  - There are many clever, interesting, and diverse proposals...

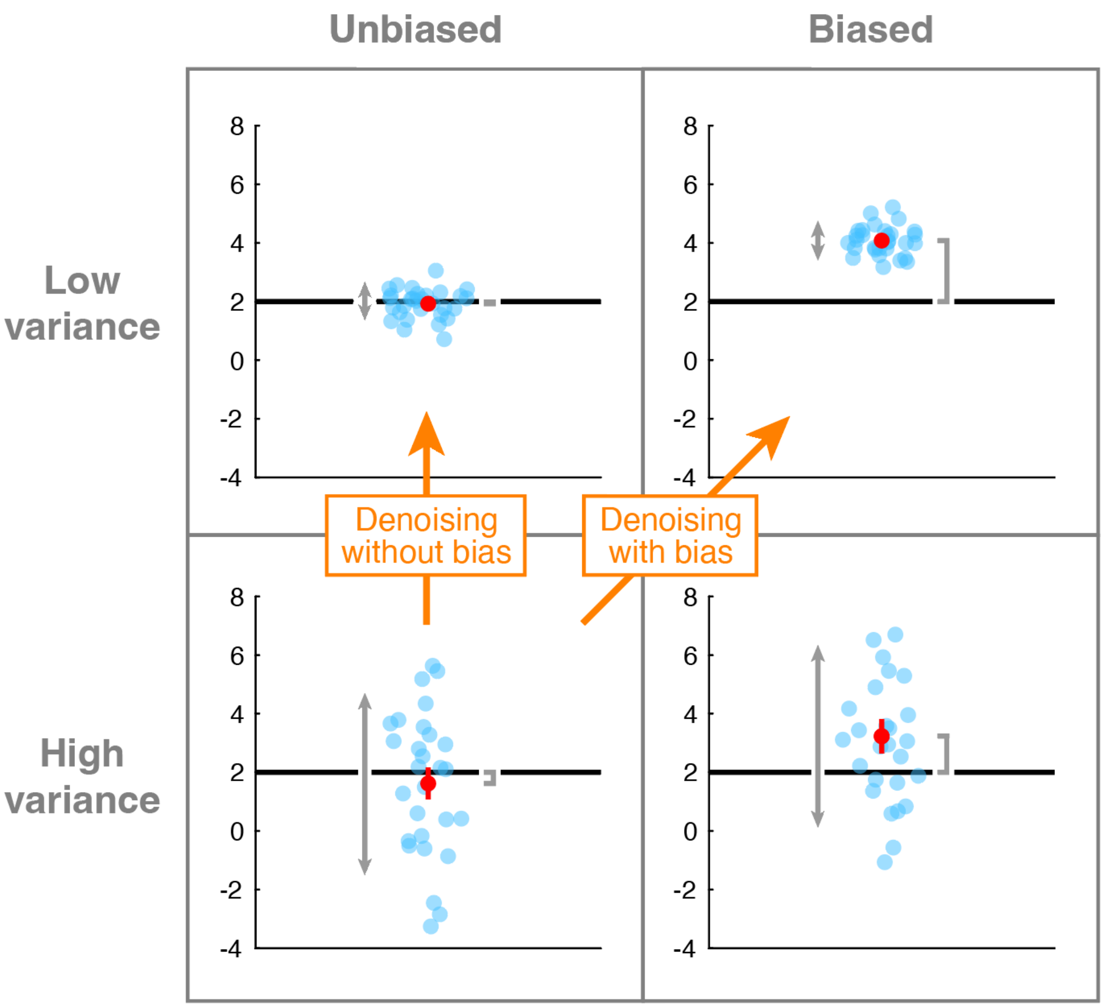
- But there is a major risk of bias.

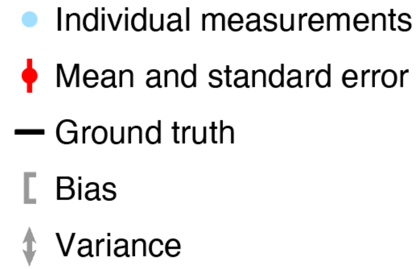
Bias and variance



Courtesy of O. Gulban

Bias and variance





$$\mathrm{BIAS} = \mathbb{E}[\hat{y}] - y$$

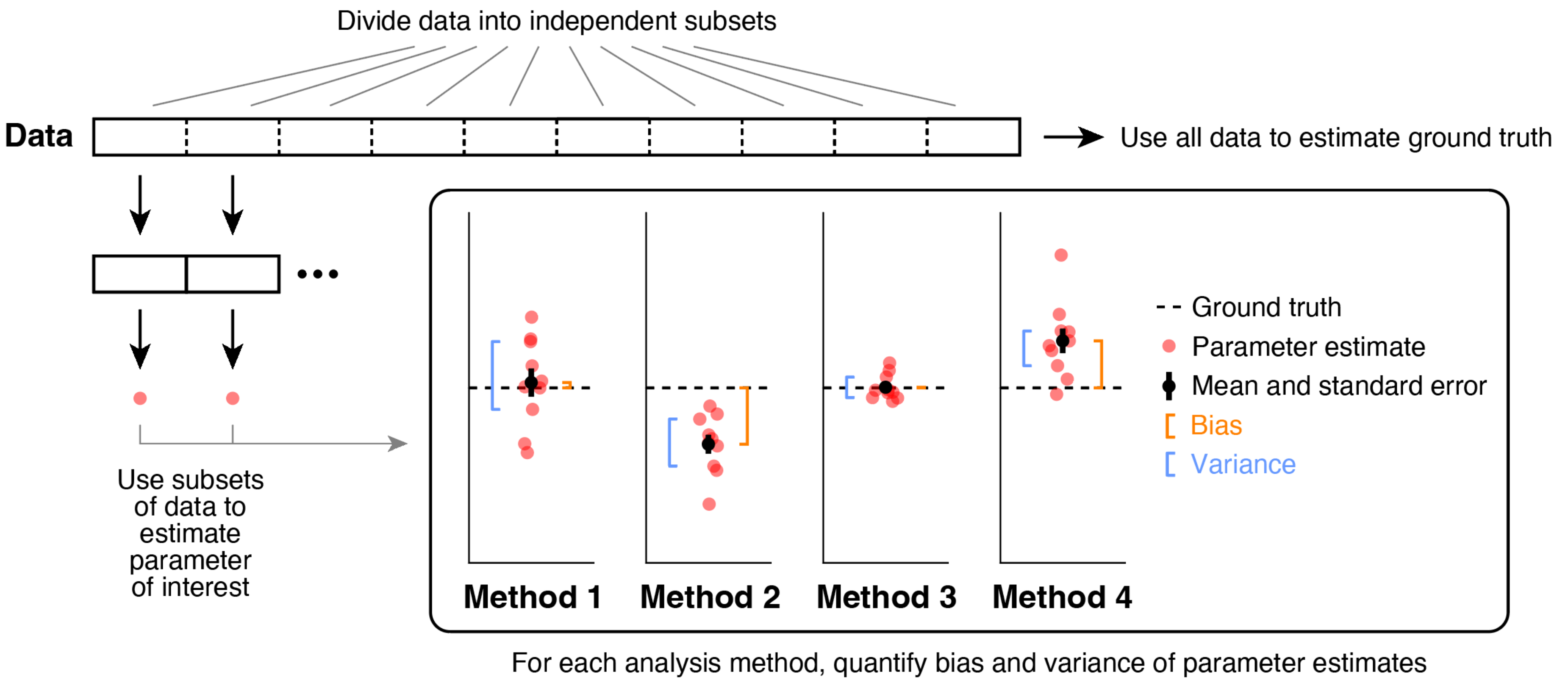
$$\mathrm{VARIANCE} = \mathbb{E}[(\hat{y} - \mathbb{E}[\hat{y}])^2]$$

Kendrick Kay, CMRR, University of Minnesota, http://cvnlab.net

Bias and variance

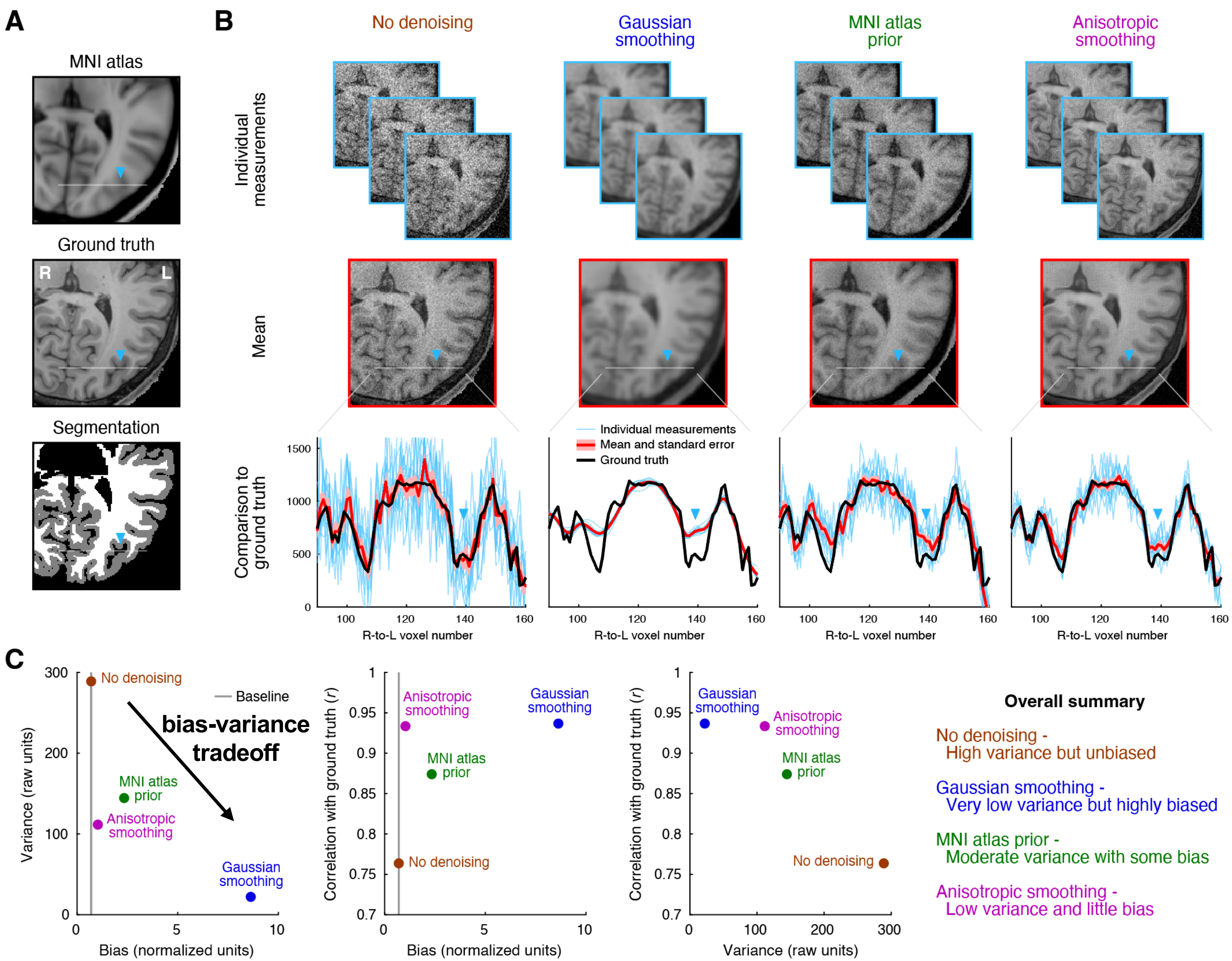
- In the paper:
  - Bias is quantified by computing, for each data point, the absolute deviation between the mean across analysis results and the ground truth, normalized by the standard error across analysis results.
    - Informally: “Across repeated simulations, does the mean converge to ground truth?”
  - Variance is quantified by computing, for each data point, the standard error across analysis results.
    - Informally: “How different are the results across repeated simulations?”
  - Error is quantified by computing Pearson’s correlation between each analysis result and the ground truth.
    - Informally: “How close do the simulations recover ground truth?”
- Important:
  - Error is a mixture of bias and variance.
  - Error does not directly tell us about bias.

How might we quantify bias/variance?



Kendrick Kay, CMRR, University of Minnesota, http://cvnlab.net

Example


A
MNI atlas
Ground truth
R
L
Segmentation
B
No denoising
Gaussian smoothing
MNI atlas prior
Anisotropic smoothing
Individual measurements
Mean
Comparison to ground truth
Individual measurements
Mean and standard error
Ground truth
1500
1000
500
0
100
120
140
160
R-to-L voxel number
C
300
200
100
0
0
5
10
Variance (raw units)
Bias (normalized units)
No denoising
Baseline
bias-variance tradeoff
MNI atlas prior
Anisotropic smoothing
Gaussian smoothing
Correlation with ground truth (r)
1
0.95
0.9
0.85
0.8
0.75
0.7
Variance (raw units)
Overall summary
No denoising -
High variance but unbiased
Gaussian smoothing -
Very low variance but highly biased
MNI atlas prior -
Moderate variance with some bias
Anisotropic smoothing -
Low variance and little bias

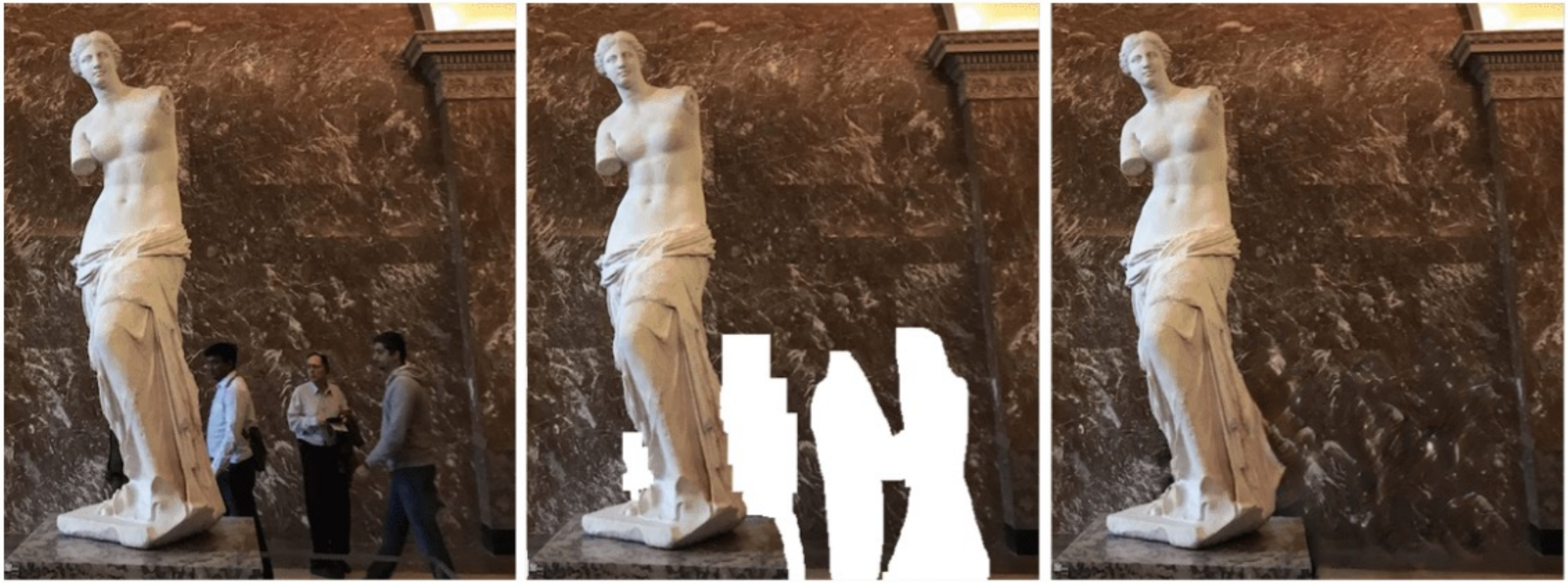
Bias-variance tradeoff

Bias-variance tradeoff:
The idea that often (but not always), a method can help reduce variance, but will do so at the expense of introducing bias.

JPEG analogy:
File size vs. compression artifact.
There is a fundamental tradeoff here!

Denoising methods as priors

- The weaker your data, the more tempting it is to try to fix it.
- If you make bad measurements, methods might invent them (bring priors).



Analogy of image inpainting
(from https://towardsdatascience.com)

For “image processing”, this might be fine and useful.
For rigorous science, we care about measurement and validity.

Take-home points

A careful stance:

- We should acknowledge bias
  [even speculation is useful]
- We should study bias
  [this can take a lot of work!]
- We should consider the risk of bias to one's goals
  [sometimes bias is OK; sometimes it isn’t]

Part 2: The value of simulations

What is a (ground-truth) simulation?

- This is a very general phrase...
  - Perhaps: code that implements analyses of highly controlled, possibly highly simplified data?

- Where does the data (ground truth) come from?
  - Empirical: find a high-quality dataset with lots of trials
  - Synthetic: generate data according to simple controlled functions
  - Random: generate data following simple statistical distributions

What are simulations good for?

Simulations can be useful for any of the following:

- Checking your knowledge (can you predict the outcome before running?)
- Checking your coding skills
- Checking your mathematical/statistical knowledge
- Checking your ability to explain concepts
- Checking your figure plotting abilities
- Demonstrating a principle clearly (to yourself and/or others)
- Comparing different methods in a controlled setting
- Exploring (cheaply) parameter space
- Finding edge cases
- Demonstrating a surprising effect that follows from simple principles

General steps in a simulation

1. Declare constants
2. Load or generate data (ground truth)
3. Repeat:
   1. Add noise to the ground truth?
   2. Analyze the resulting data
   3. Record results
4. Calculate metrics that summarize the results
5. Plot and interpret these metrics

Possibly embed the entire procedure in a loop that:

- *Systematically varies constants and parameters?*
- *Systematically varies the noise level?*
- *Explores what happens with random or shuffled data?*

Designing simulations

- Making a good simulation is hard work
- Design choices:
  - Make the code as modular as possible?
  - Who is the audience (you or others)?
  - What assumptions does your simulation involve?
  - What type of noise are you assuming?
  - Is the goal to demonstrate something new?
  - Is the goal to demonstrate a known concept?
  - Are the data intended to be realistic data, or highly controlled data?
- Think about:
  - Did you design your simulation well? Did you explore the relevant parameter ranges?
  - Are the math and code implementation correct?
  - Did you plot the results effectively?
  - Did you explain and reason about the underlying concepts well?