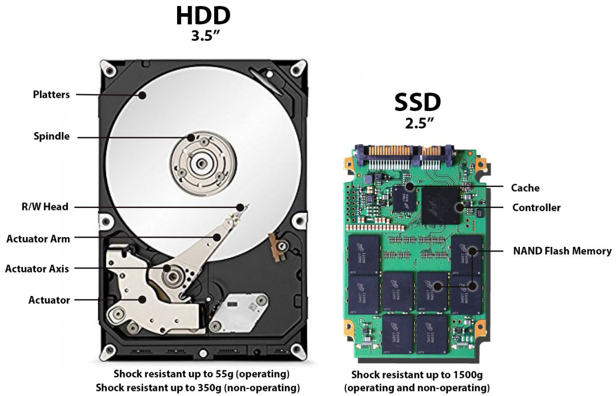
SSD Failures in the Field: Symptoms, Causes, and Prediction Models

Jacob Alter
Ji Xue
Alma Dimnaku
Evgenia Smirni

Solid State Drives (SSDs)

- Storage medium of choice for datacenter and enterprise applications
- High throughput, low energy footprint
- SSD sales are expected to surpass HDD sales by 2021*



*Statista. "Shipments of Hard and Solid State Disk (HDD/SSD) Drives Worldwide from 2015 to 2021 (in Millions)."

Reliability: SSDs vs HDDs

- SSD reliability characteristics are different
- Extreme temperatures, shocks, magnetic fields: SSDs are more resi
- Higher bi
- Laborato

Workload and reliability: complex relationship

Long-term wear? Smaller but realistic workloads?

Existing SSD Field Studies

- Facebook data (Meza et al., SIGMETRICS 2015)
  - Looks at memory layout effects, effects of temperature, aging e
- Google d ani et al., ATC 2017
  - Raw bit error rates and uncorrectable error rates
  - Prediction of uncorrectable errors and bad blocks

Our goal: study the life and death of SSDs

Research Questions

- What are the longitudinal effects of aging and wear?
- How do the effects of write wear compare to the effects of drive aging?
- How predictable are drive failures? How far in advance can they be predicted?

Outline

1. Data description
2. Failure characterization: failures/repairs
   a. Timeline of repairs process
   b. Symptoms and causes
3. Failure prediction
   a. Model interpretation

Field Data

Collected “in the wild” at a Google data center

Daily activity reports

- Span a period of 6 years
- 3 MLC drive models
- 10,000 unique drives
- 40,000,000 reports, total

Drive Attributes

Age
Model
Factory bad block count

Status Flags

Status dead
Status read-only

Workload

P/E cycle count
Read count
Write count
Erase count

Error Counts

Bad block count
Correctable error
Erase error
Final read error
Final write error
Meta error
Read error
Response error
Timeout error
Uncorrectable error
Write error
Drive failure

Drive Attributes

Age
Model
Factory bad block count

Status Flags

Status dead
Status read-only

Workload

P/E cycle count
Read count
Write count
Erase count

Error Counts

Bad block count
Correctable error
Erase error
Final read error
Final write error
Meta error
Read error
Response error
Timeout error
Uncorrectable error
Write error
Drive failure

Program/Erase (P/E) cycle

One write operation + one erase operation

Uncorrectable Error (UE)

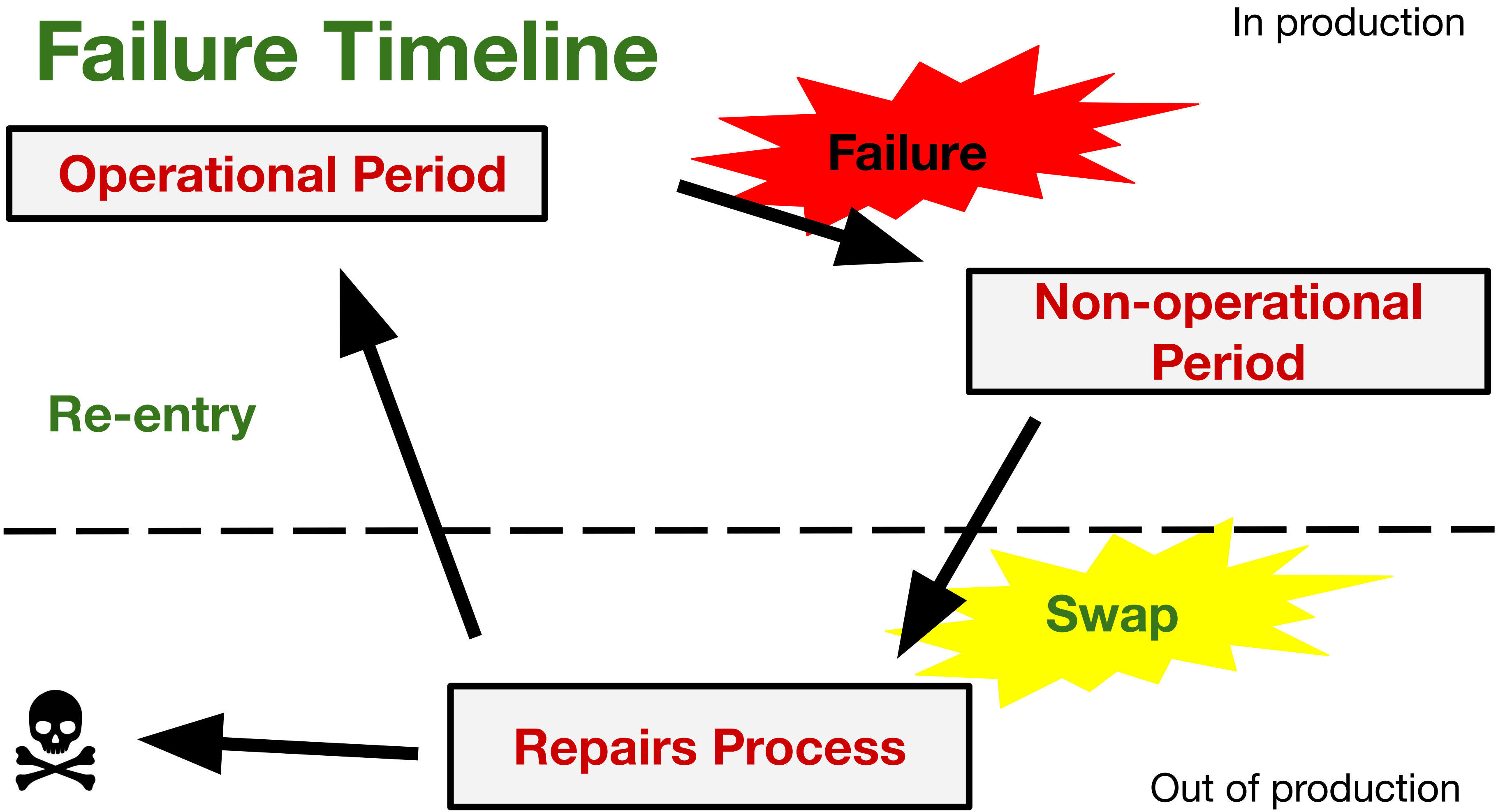
Read operation that fails to be corrected by ECC

Drive Failure

Catastrophic events requiring intervention from data center maintenance

Failures & the Repairs Process

Failure Timeline



Failure Incidence Rates

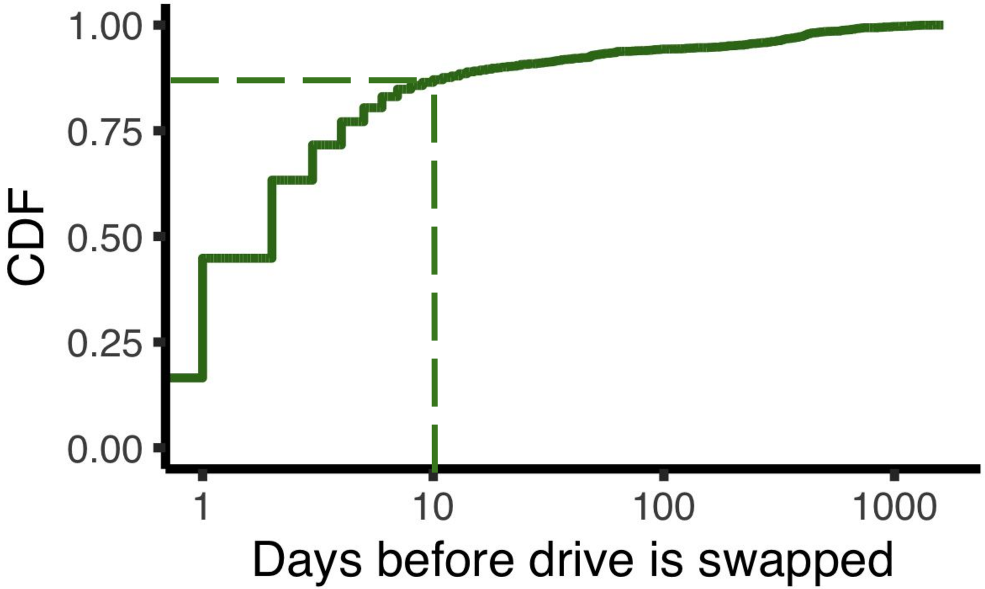
Model	# Failures	% Failed
MLC-A	734	6.95%
MLC-B	1565	14.3%
MLC-D	1580	12.5%
All	3879	11.3%

# Failures	% of drives
0	88.7%
1	10.1%
2	1.04%
3	0.133%
4	0.001%

Observation:

Failures are common! Drives may fail many times

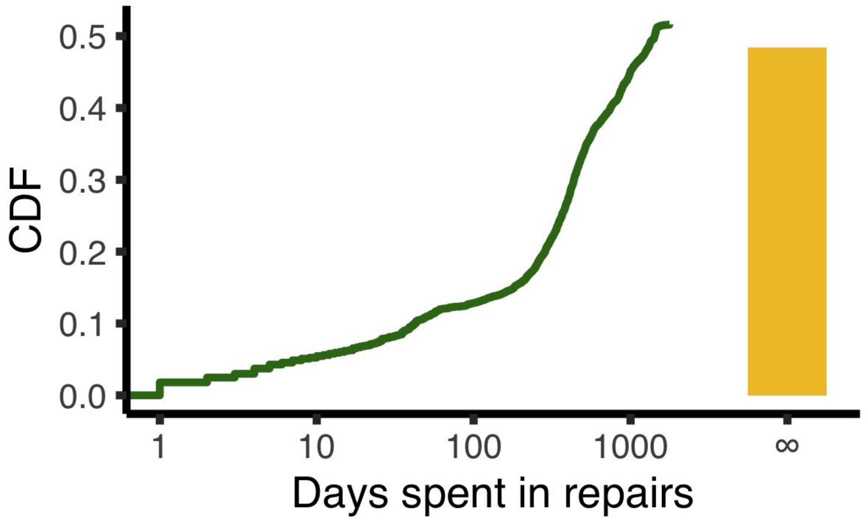
The Repairs Process



Observation:

Failed drives may remain in the system for upwards of a year

The Repairs Process



Observation:

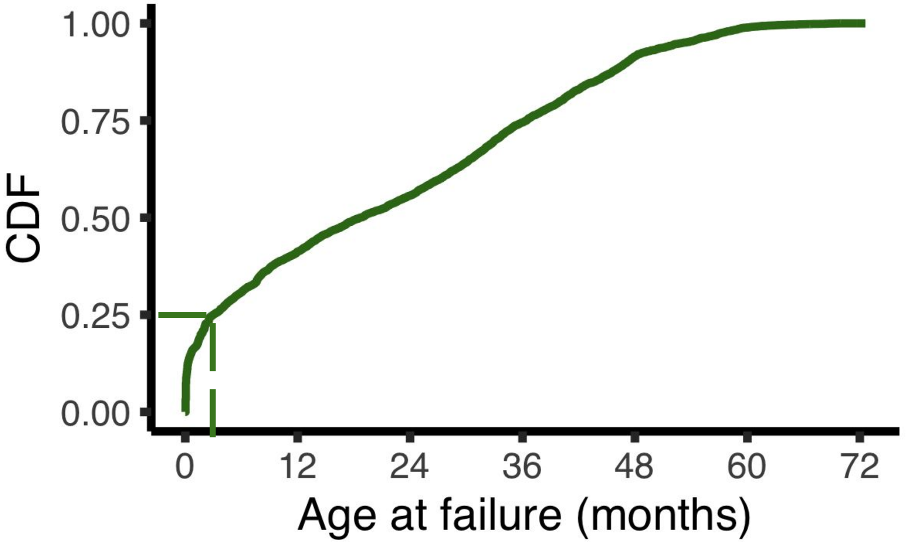
Only half of failed drives are observed to re-enter the field

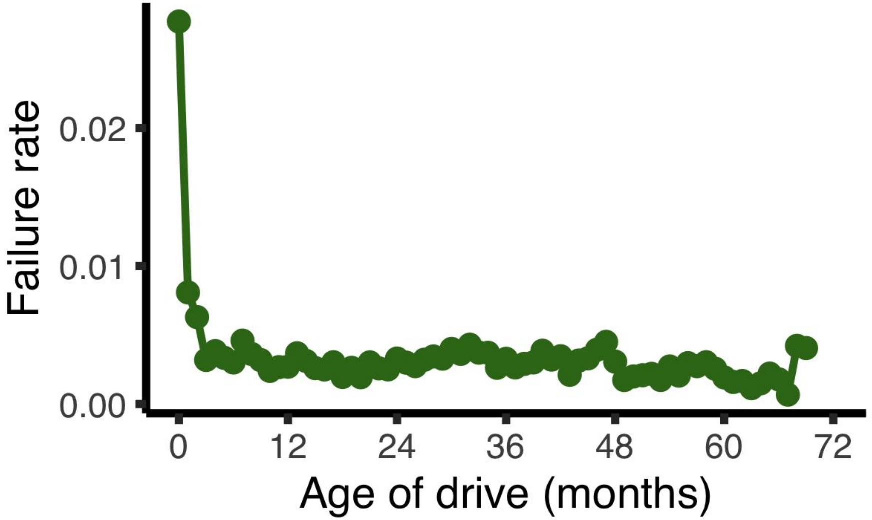
Observation:

Drive repairs take upwards of a year, on average

Symptoms & Causes of Drive Failure

Failures and Drive Age





Observation:

25% of failures occur within the first 3 months of the drive's life

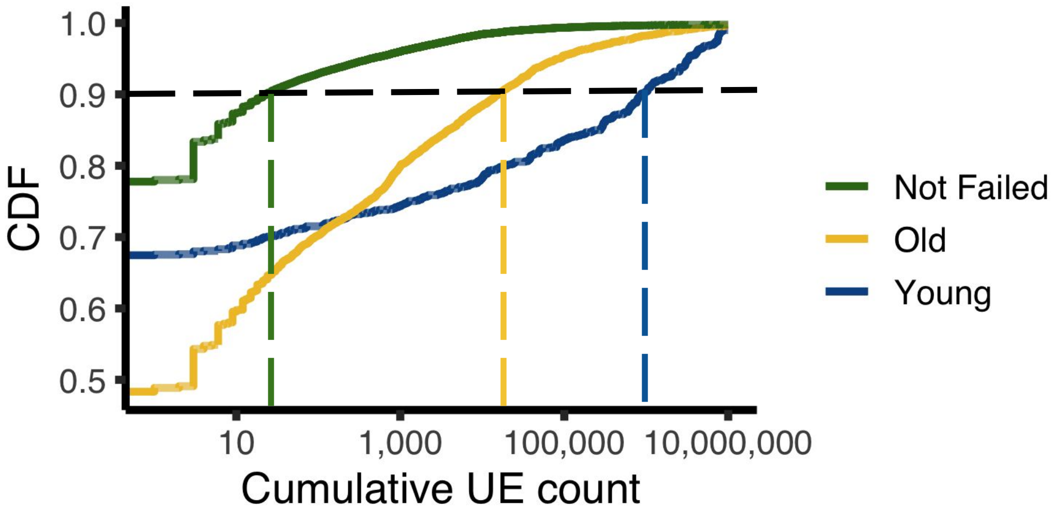
Observation:

Failure rate is roughly constant among older drives

Infant Failure / Young Failure

A failure occurring within the first three months of a drive's life

Failures and Errors



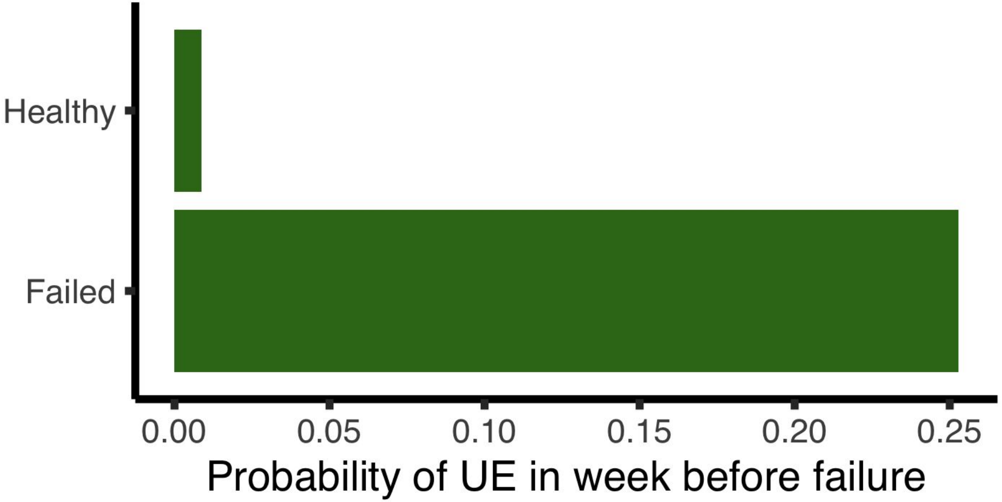
Observation:

Failed drives have higher error incidence, but over half of them experience no errors before failure.

Observation:

Young failed drive failures are more likely to have very high error counts

Failures and Errors



Observation:

75% of failed drives do not experience any UEs in the 7 days prior to their failure

Takeaways

- Failures occur spuriously and without warning
- Young drives fail more catastrophically than old drives
- We must turn to advanced, nonlinear techniques to detect drive failure

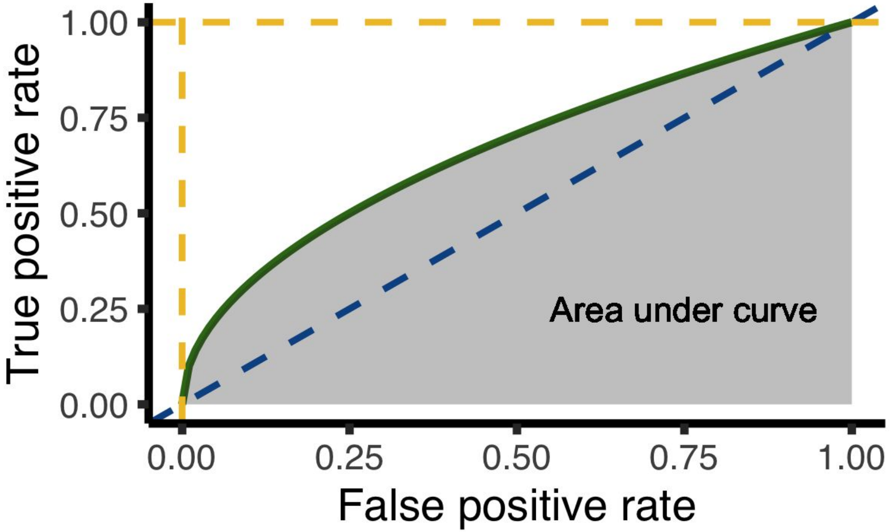
Failure Prediction

Failure Prediction Model

- Machine learning to predict upcoming failures
  - Gain insights into the failure process
- Model input: activity summary statistics
  - Daily activity & error counts
  - **Cumulative** activity & error counts
- Model output: probability that the drive will fail within the next N days

Evaluation

- Problem: high imbalance (1 failure case for each 10,000 failure cases)
- Evaluation metric: ROC AUC score
  - True positives vs false positives as prediction threshold is varied
  - Ranges from 0.5 to 1.0



Prediction Results

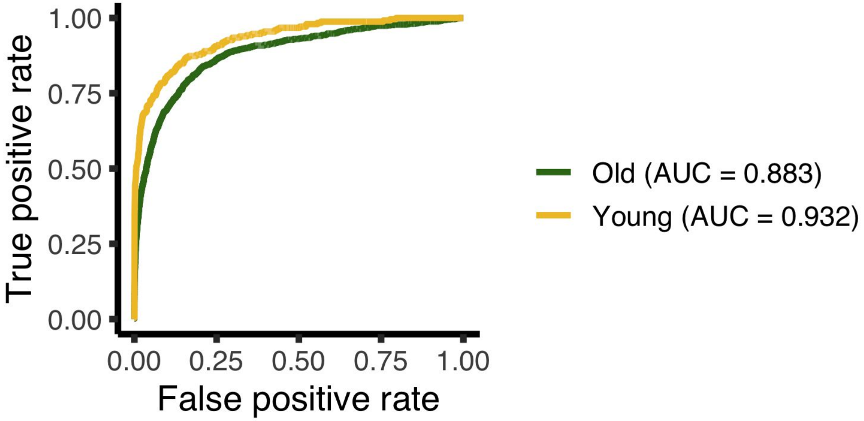
	N = 1	*N* = 2	*N* = 7
Logistic Regression	0.796	0.765	0.713
kNN	0.816	0.791	0.716
SVM	0.821	0.795	0.728
Neural Network	0.857	0.828	0.770
Decision Tree	0.872	0.840	0.780
Random Forest	**0.905**	**0.859**	**0.803**

Model Interpretation

Observation:

Model performance differs greatly between old and young drives

- Splitting the **training data by age** improves model performance!
  - Young AUC = 0.970
  - Old AUC = 0.890
  - Indicates that the indicators of failure differ between the two sets



Model Interpretation

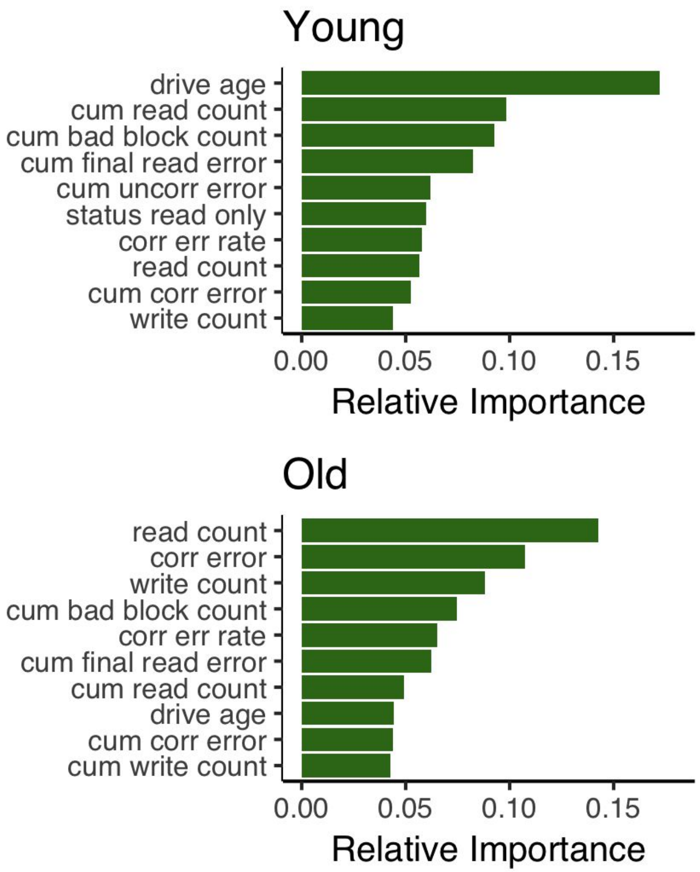
- Random forests produce rankings of most important features

Observation:

P/E cycles are nowhere to be found!

Observation:

Young drives: age & error counts
Old drives: drive wear, age irrelevant



Take Home Message

- We find **no association** in our data between P/E cycles and drive failure incidence
- The **age** of a drive is crucial to understanding its failure characteristics
- We are able to achieve **high quality** short-term forecasts of drive failure using machine learning techniques

Thank you!

Future Directions

- Confirmation of effects of drive age vs. P/E cycle wear
  - More fine-grained workload information required
- What are the mechanical/operational reasons for heightened early failure rates?
  - How can they be mitigated?
  - Can we detect that a drive will fail early?