

FUNDAMENTALS OF AI-AUGMENTED ENGINEERING: A CONVERSATIONAL APPROACH

Taner Onsay, Ph.D. & Sebnem Onsay
Authors
October 2026 · Revision 1.0

© 2026 AI First Engineering



The engineering workflow you will use throughout this book

CHAPTER - 7

PART 5 · ≈ 30 MIN · IN LAB

Put the Model on Trial

The model and the test do not agree perfectly. Three suspects — the MODEL, the DATA, the COMPARISON — and one basketball shot's worth of evidence. Examine it, weaken what it cannot support, and decide what it does justify.

5

Problem



Model



Compute



Analyze



Validate



Improve



The model and the test do not agree perfectly. What is the most defensible explanation?

Evidence	For the ideal model	Against — or unexplained
x(t), straight line	$R^2 = 0.99985$	U-shaped residuals, up to 34 mm
y(t), quadratic	$R^2 = 0.99982$	structured residuals, up to 14 mm
a _{y,fit}	-9.43 m/s ² , close to g	3.9 % from -9.81 m/s ²
apex	fit and camera agree within 2 mm	—
robustness	stable to small trims	-8.9 ... -10.0 across data windows
test blind spot	—	R ² cannot see a wrong ruler

MODEL 1 on trial: ideal projectile — gravity only, no drag, no spin. Do not answer yet. First: who are the suspects?

An investigation, not a verdict by R². Examine evidence. Weaken unsupported explanations. Decide what is justified.

Problem



Model



Compute



Analyze



Validate



Improve



Every explanation you proposed belongs to one of three suspects

1 · MODEL

Is important physics missing — or an assumption wrong?

e.g. drag · spin · non-planar motion

2 · DATA / TEST

Are the measurements accurate enough to support the conclusion?

e.g. clicking · pixels · blur · calibration

3 · COMPARISON

Are model and experiment compared consistently?

e.g. axes · scale plane · clock · same quantity

EXPERIMENTAL DATA ARE EVIDENCE — NOT UNQUESTIONABLE GROUND TRUTH. Suspect 2 is not secondary to Suspect 1.

Activity (3 min): each team places every item from the board list in one column — then defends one placement.

Problem



Model



Compute



Analyze



Validate



Improve



If the simple model is incomplete, what physics might matter?

MODEL 1

gravity only

Parts 1–4. VEL's own label: "no drag, spin, lateral error."

MODEL 2

gravity + aerodynamic drag

Drag acts OPPOSITE the ball's velocity — it slows the ball.

MODEL 3

gravity + drag + spin (Magnus)

A spinning ball pushes air aside unevenly: a force PERPENDICULAR to the velocity. Its direction depends on the spin — backspin tilts it upward, topspin downward.

Engineering hypotheses — not upgrades. Each must earn its place with evidence. Conceptual only: no equations of motion here.

Problem



Model



Compute



Analyze



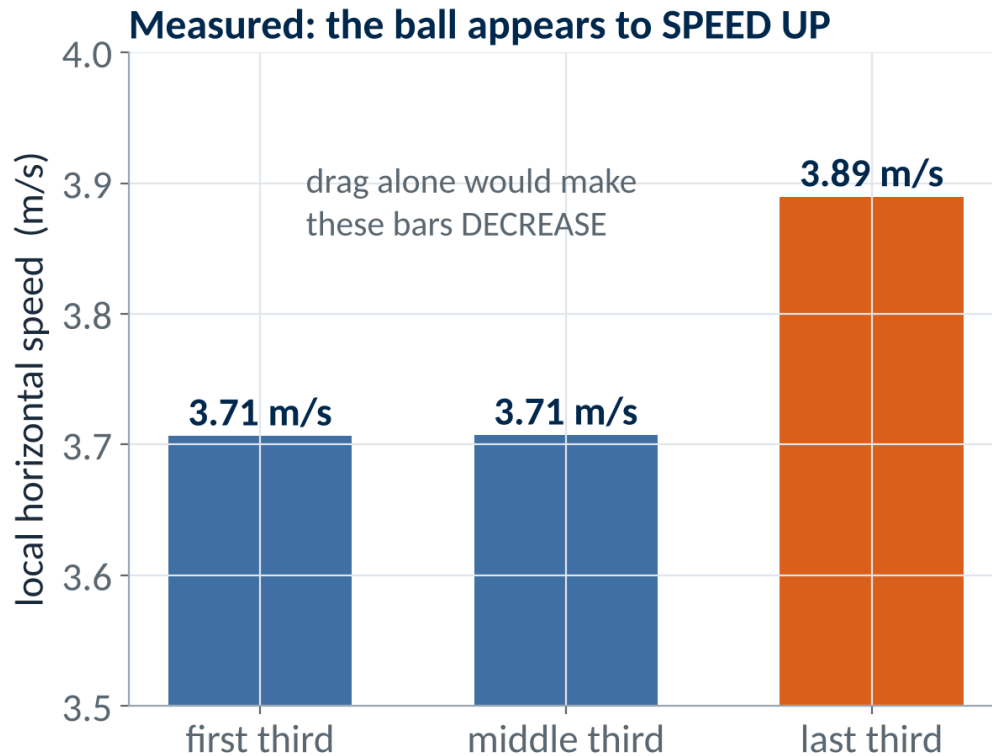
Validate



Improve



If drag were the story, which way should $x(t)$ bend?



PREDICT FIRST

Drag opposes the motion. Over the flight, should the horizontal speed go up, down, or stay the same?

WHAT THE DATA SHOW

Local horizontal speed: $3.71 \rightarrow 3.71 \rightarrow 3.89$ m/s. The ball **appears to speed up** — the U-shaped x residual from Part 2.

Drag alone would make it slow down.

Drag alone does not explain the x pattern. That does NOT prove drag is absent — only that drag is not this pattern.

Problem



Model



Compute



Analyze



Validate



Improve



ACCEPT, MODIFY or REJECT?

AI MODELING PARTNER

“The systematic residuals prove that the ideal projectile model is missing aerodynamic drag. Add drag and the model will become more accurate.”

RESIDUAL DIRECTION

Which way do the x residuals bend?

DRAG'S EFFECT

Which way would drag bend them?

TEST QUALITY

Are the effects larger than the measurement errors?

TEST GEOMETRY

Could the camera set-up mimic the pattern?

Team verdict + one sentence of evidence. Not "AI is wrong because drag doesn't exist" — but: does the evidence justify assigning the discrepancy to drag?

Students must demand evidence. The AI may propose physics. It does not get to decide what the residuals mean.

Problem



Model



Compute



Analyze



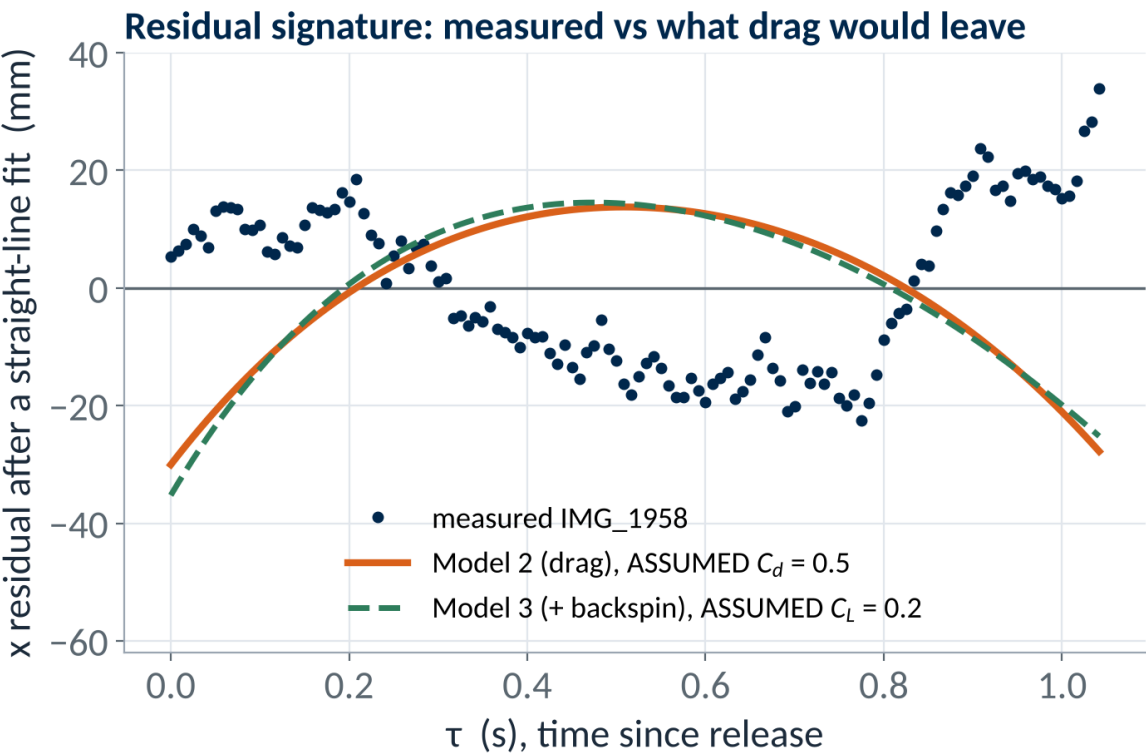
Validate



Improve



Compute the prediction — with the assumptions in plain sight



Fit Model 1 to...	$a_{y,fit}$	x bends
the measured shot	-9.43	speeds up
Model 2 · drag	-9.86	slows down
Model 3 · drag + backspin	-9.75	slows down

ASSUMED, not measured: $C_d = 0.5$, backspin lift coefficient 0.2, typical textbook values.

READ IT AS

With these assumptions, both models bend x the other way — and are only cm in size.

An assumed parameter is not experimental evidence. It tells us what to LOOK for — not what happened.





A human created these 126 numbers from video. How accurate is each one?



f041 · release, blur



f102 · ball at the top edge



f126 · still cut off



f166 · inside rim / net

BALL SIZE

about 55–60 px across $\approx$ 4 mm
per pixel (estimated)

CENTRE PICK

a few px off $\approx$ 1–2 cm, by hand,
frame by frame

MOTION

$\approx$ 15 px of travel per frame —
visible blur

BACKGROUND

wall, dark ceiling, scoreboard,
net

Could a few pixels of systematic centre-picking matter? Here, a few pixels is the size of every residual we have discussed.

Problem



Model



Compute



Analyze



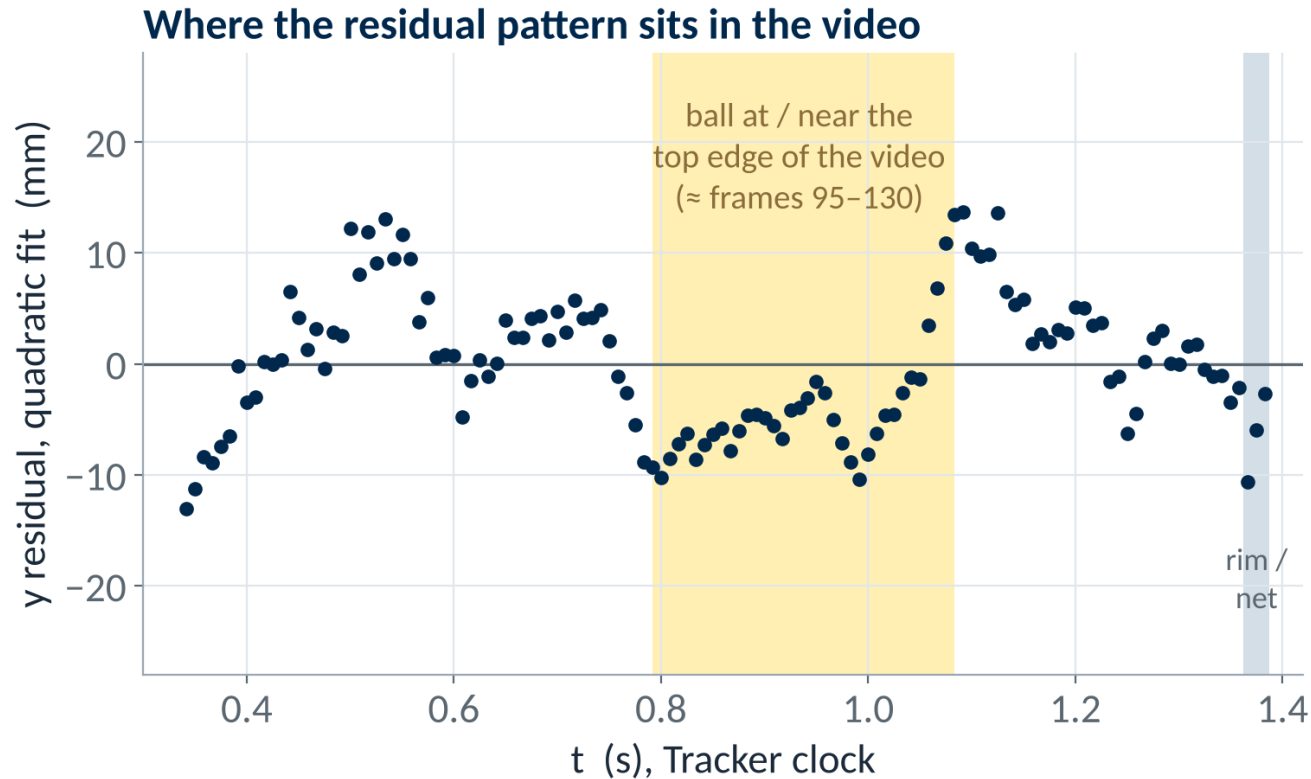
Validate



Improve



Does the residual pattern coincide with a known measurement problem?



OBSERVE

Across the top-edge frames the residuals run negative, then jump about +15 mm near $t \approx 1.05$ s — about when the whole ball is visible again.

CONCLUDE — CAREFULLY

Consistent with a **centre bias at the frame edge**.
Consistent is not proven.

Evidence that a pattern lines up with a known test problem raises that suspect — it does not convict it.

Problem



Model



Compute



Analyze



Validate



Improve



Does the conclusion change if suspect points are set aside — for a stated reason?

Analysis	Engineering reason	N	$a_{y,fit}$	x bend (m/s^2)	RMS y
PRIMARY · all points	nothing removed without a reason	126	-9.43	+0.28	6.1 mm
without rim / net	ball inside rim; contact possible	123	-9.42	+0.27	6.1 mm
without top-edge frames	ball cut off by the frame (f95–130)	90	-9.50	+0.23	4.9 mm
without both	both reasons	87	-9.49	+0.21	4.9 mm

IDENTIFY

STATE WHY

REPEAT

REPORT

Result: the residuals shrink, but the 3.2–3.9 % gap and the x bend remain. The conclusion does not collapse — and these points alone do not explain it.

Data removal requires an engineering reason. Never "delete until the fit improves." Report ALL data first.

Problem



Model



Compute

**Analyze**

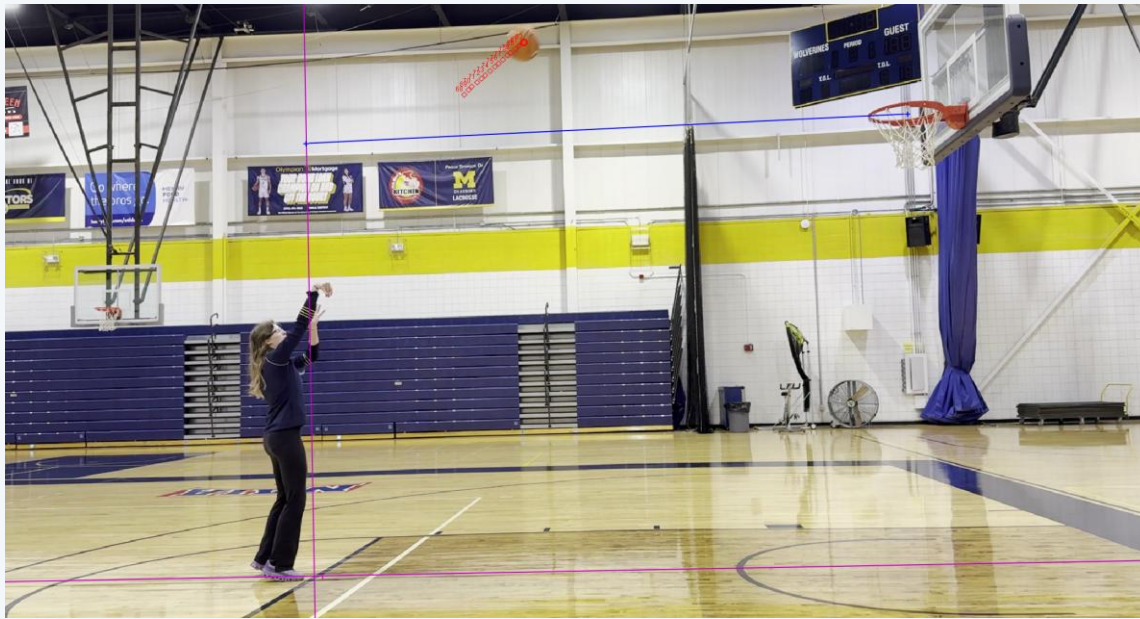
Validate



Improve



Model coordinates vs image coordinates: are we comparing like with like?



Frame 82: Tracker axes (magenta) and calibration stick (blue).

AXIS ORIENTATION

A $\approx 1.6^\circ$ tilt between the axes and true horizontal would create the observed x bend. The axes sit $\approx 1^\circ$ from the image edge — is the image level?

SCALE PLANE

Calibration fixes metres-per-pixel in ONE plane. A ball path 4 % of the camera distance nearer or farther changes every length by $\approx 4\%$.

CLOCK & ORIGIN

Tracker clock vs time since release (Part 3). Correct numbers, wrong reading.

A MODEL can be right and TEST DATA reasonable — while the COMPARISON between them is flawed.

Different suspects may explain different symptoms. An axis tilt could explain the x bend — but not the 3.9 % in y.

Problem



Model



Compute



Analyze



Validate



Improve



Can our experiment detect what Models 2 and 3 would add?

TO DEFEND MODEL 2

- drag coefficient
- ball diameter / area
- air density
- velocity dependence
- a numerical solution (ODE)

TO DEFEND MODEL 3, ALSO

- spin rate
- spin axis
- lift (Magnus) relationship
- more parameters to justify

Size of ...

about

frame-to-frame scatter

≈ 2 mm

one pixel

≈ 4 mm

systematic test effects seen

5 – 30 mm

drag / spin left after refitting

$\approx 4 - 13$ mm

THE KEY QUESTION

Is the improvement more physics would bring larger than what the test can resolve? **Here: no — they are the same size.**

Every added term needs assumptions, parameters and evidence.

An experiment that cannot tell the models apart is not a failure. It is a conclusion about the experiment's information content.

Problem



Model



Compute



Analyze



Validate



Improve



We are testing a very well-established physical model with one basketball shot. Is one test enough?

ONE TEST ›

REPEAT ›

COMPARE ›

VARIATION? ›

PATTERNS? ›

ACCUMULATE ›

JUDGE

ANALYZE EACH SHOT ON ITS OWN

Shot 1 $\rightarrow a_{y,fit,1}$ Shot 2 $\rightarrow a_{y,fit,2}$... Shot N $\rightarrow a_{y,fit,N}$

Then compare the fitted parameters, residual patterns and predictions. Never average $x(t)$ and $y(t)$ point by point across shots — that builds a trajectory no player threw.

AVERAGING HELPS — AND DOESN'T

Averaging independent tests reduces RANDOM variation. **It does not remove SYSTEMATIC error.** A 4 % calibration bias in every video survives 20 repeats.

REPEATABILITY $\neq$ ACCURACY

Hypothetical: -9.43, -9.44, -9.42, -9.43, -9.44 — very repeatable.

Repeating a biased experiment can produce a very **precise wrong answer.**

Before changing the physics model: run another test. Then another.

Problem ›

Model ›

Compute ›

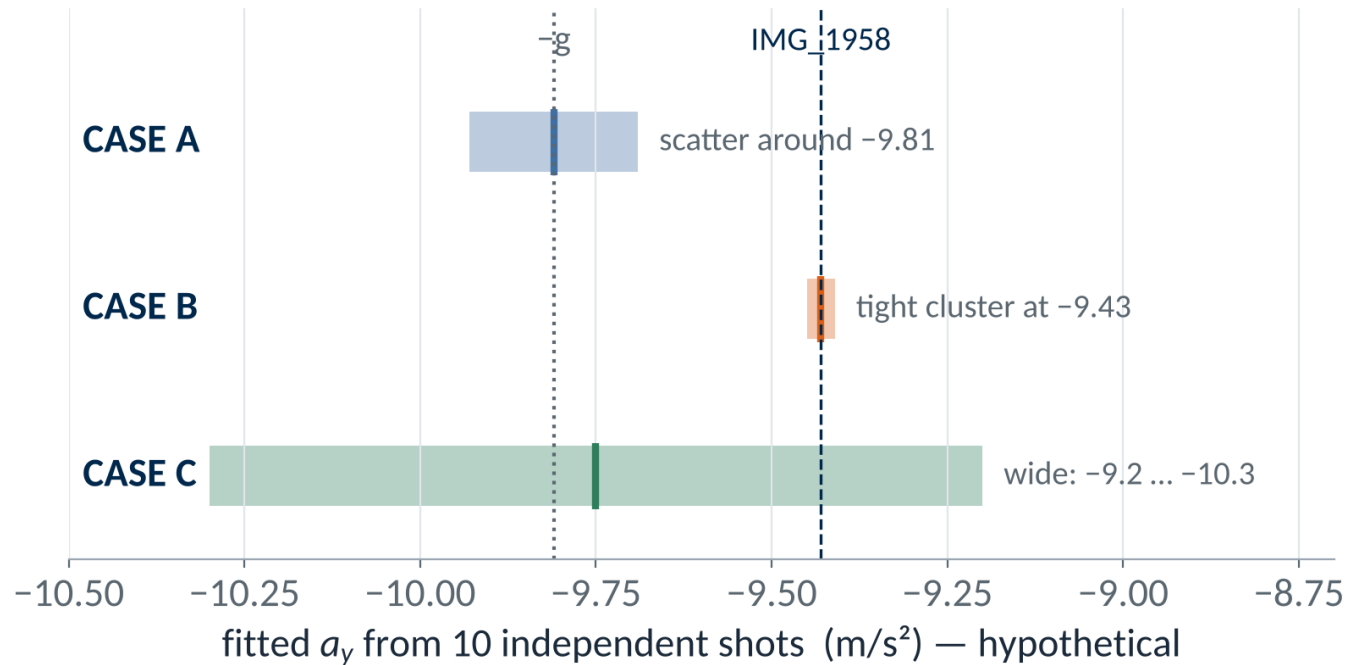
Analyze ›

Validate ›

Improve



Suppose we analyze 10 more shots the same way. What would each outcome suggest?



CASE A

Scatter around $-9.81 \rightarrow$ IMG_1958 may mostly reflect shot-to-shot variation.

CASE B

Tight cluster at $-9.43 \rightarrow$ something systematic. Still not proven: MODEL, DATA or COMPARISON?

CASE C

Wide spread $\rightarrow$ the method may lack the repeatability this claim needs.

Ranges are hypothetical — for discussion. No additional shots have been analyzed.

Accumulating evidence changes what we can claim — without a single formula.

Problem



Model



Compute



Analyze



Validate



Improve



Before you change the physics model, what test would you run next?

YOUR TEAM'S PROPOSAL (4 min)

1. Which suspect does it target — model, data or comparison?
2. What will you change between tests?
3. What result would point to each suspect?
4. What would it cost in time and equipment?

IDEAS TO START FROM

repeat shots, same player

several players

camera farther away, squarer to the plane

calibration object in the ball's plane

level the camera; check axes

a second analyst tracks the same video

higher resolution / less blur

keep the ball inside the frame

measure spin

drop a ball as a known-g check

Validation is not only "calculate a metric." It is "design the next piece of evidence."

Problem



Model



Compute



Analyze



Validate



Improve



IMPROVE does not only mean "improve the model"



...then repeat the cycle. Validation is not the end of the workflow — it tells you what to do next.

The course workflow, one level deeper: Problem › Model › Compute › Analyze › Validate › Improve — and back again.





What does the evidence actually justify?

OUR CONCLUSION

The ideal projectile model represents the measured trajectory extremely well for many practical purposes.

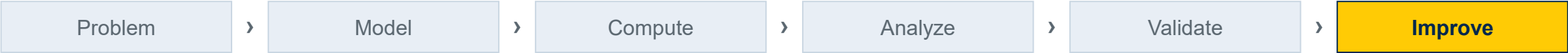
The remaining systematic discrepancies are comparable with plausible limitations of the video/Tracker experiment.

One test cannot uniquely attribute them to drag, spin, or a specific measurement or comparison error.

So the evidence does not yet justify adding model complexity. More controlled tests are needed.

Is Model 1 adequate for ...	Verdict
visualizing the trajectory	yes
estimating release speed	yes, with the ruler caveat
estimating apex height	yes
predicting hoop crossing	Project 1 will test it
measuring g accurately	not with this test
detecting drag / quantifying Magnus	not with this test

This is not indecision. This is engineering judgment: a claim sized to the evidence.





Six questions to trust a model — three suspects when it disagrees

1	FIT	Does the model agree numerically with the measurements?
2	RESIDUALS	Is systematic structure left unexplained?
3	PHYSICS	Are the coefficients and behaviour physically plausible?
4	STABILITY	Do reasonable analysis changes give similar conclusions?
5	TEST QUALITY	Are the measurements accurate, repeatable and sufficient for the claim?
6	PURPOSE	Is the model accurate enough for the decision?

IF SOMETHING DOES NOT AGREE

**MODEL
+ DATA / TEST
+ COMPARISON**

NEXT · PROJECT 1

Chapter 7: complete trajectory → fit → interpret → question → validate.

Project 1: partial trajectory → estimate → model → predict the unseen → reveal → validate.

The model now has to predict something it has not seen.

A reusable ENGR 200 framework — for basketballs, beams, sensors, circuits and anything else you will fit.



CHAPTER 7 · FROM DATA TO MODEL

Data do not choose their own model.
A good fit does not prove the physics.
Test data are evidence, not absolute truth.
More complexity requires more evidence.
One test may not be enough.
Repeatability is not the same as accuracy.
When evidence is ambiguous, design the next test.

FITS $\neq$ VALIDATED

**AI ASSISTS.
HUMAN VERIFIES.
ENGINEER DECIDES.**

Acknowledgment

Special thanks to **Jennifer Henegar**, a University of Michigan–Dearborn ME345 Engineering Dynamics student, for volunteering the basketball shot that made Chapter 7's experimental investigation possible.

Problem



Model



Compute



Analyze



Validate



Improve