

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 2 · ≈ 16 MIN

The Best Line

126 measured x values, one straight-line model, and infinitely many candidate lines. Before any software: decide what “best” means. Then let the computer find it.

2

Problem

›

Model

›

Compute

›

Analyze

›

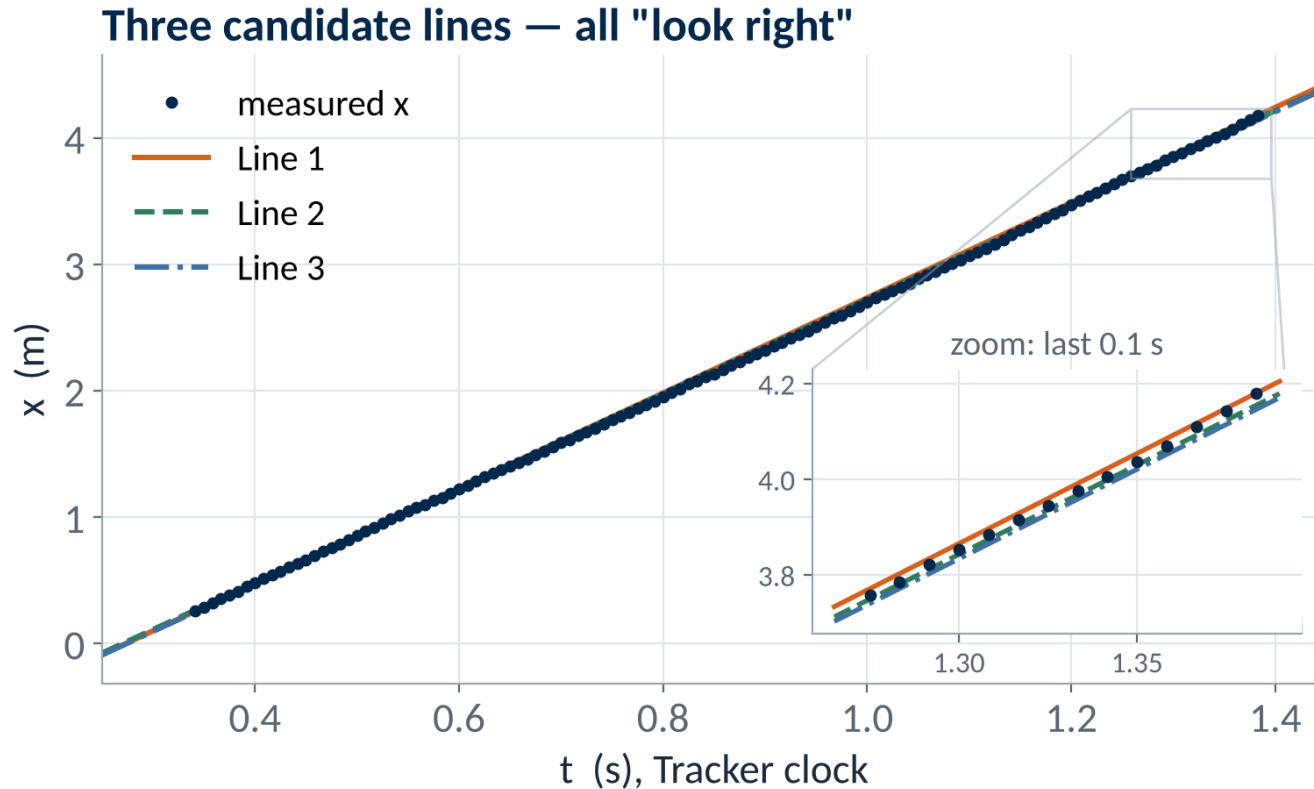
Validate

›

Improve



Many lines look right. Which one fits best?



THE MODEL

$$x(t) = A_x t + B_x$$

Two unknown coefficients. Every choice of A_x and B_x is a different line.

ASK THE ROOM

Which line is best? How would you **prove** it to a teammate who picked a different one?

Our eyes cannot rank these lines. We need a NUMBER that says how badly a line misses ALL the data.

Problem

**Model**

Compute



Analyze



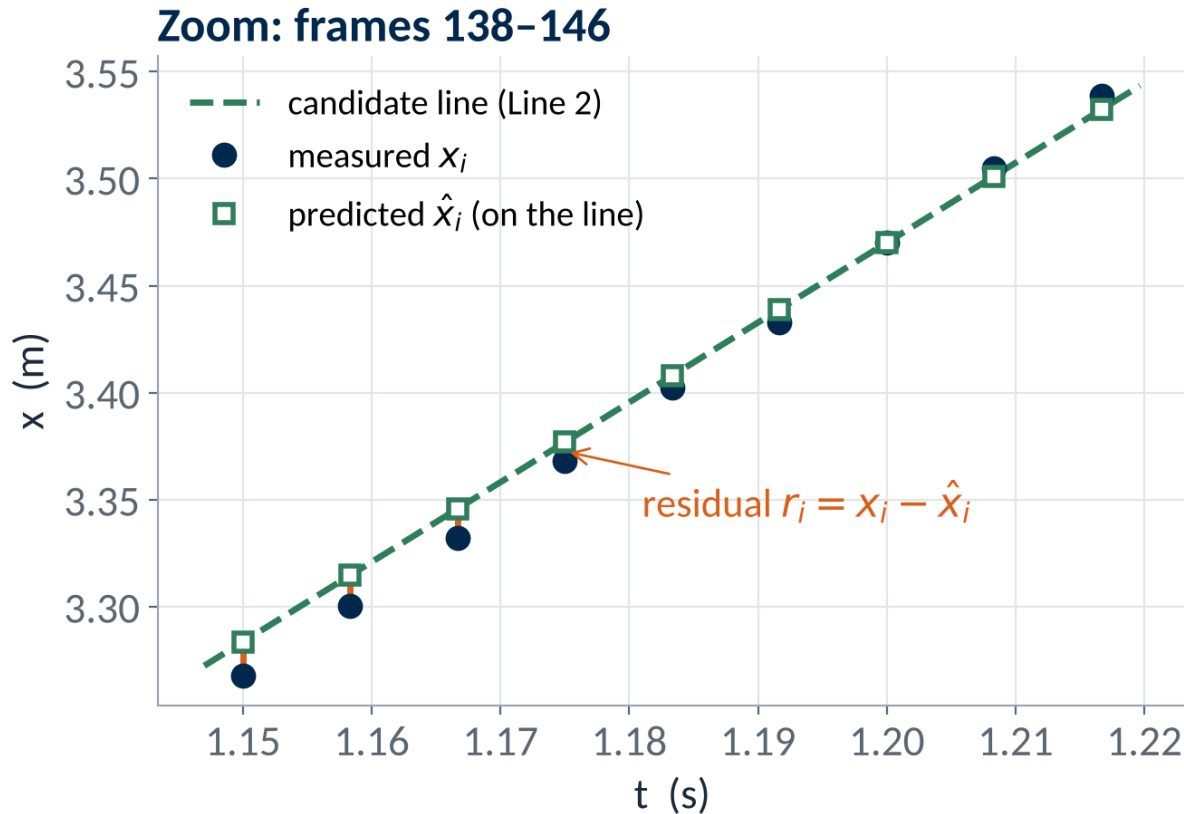
Validate



Improve



For every measured point: what did the line predict, and how far off was it?



PREDICTION

$$\hat{x}_i = A_x t_i + B_x$$

What the line says at the measured time t_i .

RESIDUAL

$$r_i = x_i - \hat{x}_i$$

Measured minus predicted, in metres. Positive: the point lies **above** the line. Negative: below.

One residual per measurement. Every candidate line produces its own 126 residuals.

Problem



Model



Compute



Analyze



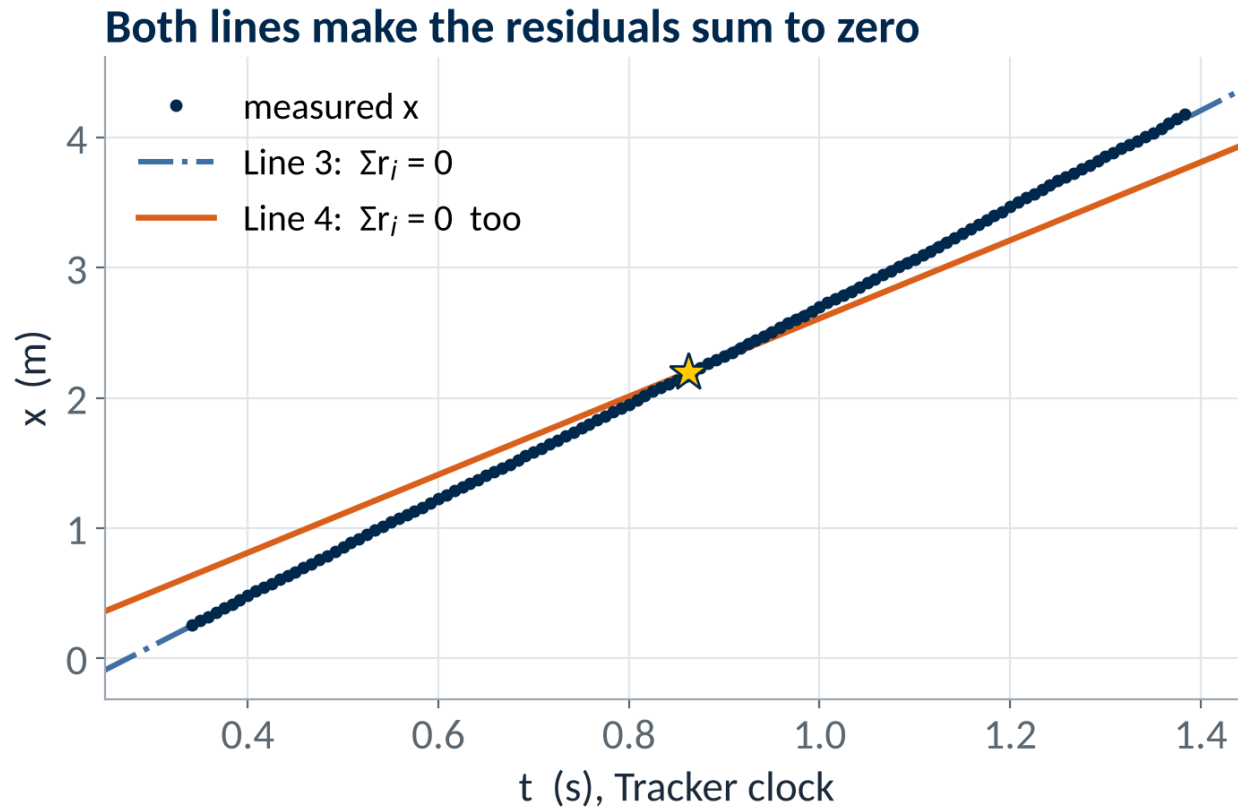
Validate



Improve



Proposal: the best line makes the residuals add up to zero



VOTE FIRST

Is $\sum r_i = 0$ a good rule for “best”? Yes · no · it depends.

THE CATCH

Line 4 is obviously wrong — and its residuals **also** sum to zero. So does every line through the star (the average point). Big positive misses cancel big negative ones.

$\sum r_i = 0$ cannot choose a line. Residuals of opposite sign cancel — a terrible line can score perfectly.

Problem



Model



Compute



Analyze



Validate



Improve



Square every residual, then add

$$SSE = \sum r_i^2 = \sum (x_i - \hat{x}_i)^2$$

EVERY MISS COUNTS

r^2 is never negative — nothing cancels.

BIG MISSES COUNT MORE

A 2-cm miss costs four times a 1-cm miss.

UNITS: m²

Smaller SSE = closer to all the data.

Candidate	$\sum r_i$ (m)	SSE (m ²)
Line 1 · first & last points	−2.48	0.0818
Line 2 · two interior points	−1.46	0.0417
Line 3	0.00	0.0244
Line 4 · slope 3.0 m/s	0.00	6.32

READ THE TABLE

Line 4 ties Line 3 on $\sum r$ — and loses on SSE by a factor of 260.

LEAST SQUARES

the best-fit line is the pair A_x, B_x that makes SSE as small as possible.

Problem



Model



Compute



Analyze



Validate



Improve



Predict, change one cell, watch SSE

$A_x =$ **3.70** m/s $B_x =$ **-1.00** m ← the two cells you change

t (s)	x measured (m)	$\hat{x} = A_x t + B_x$ (m)	$r = x - \hat{x}$ (m)	r^2 (m ²)
0.3417	0.2577	0.2642	-0.0065	0.000042
0.3500	0.2898	0.2950	-0.0052	0.000027
0.3583	0.3221	0.3258	-0.0037	0.000014
0.3667	0.3558	0.3567	-0.0008	0.000001
⋮ 126 rows	⋮	⋮	⋮	⋮
			SSE = $\sum r^2$	0.0482

YOUR TURN

1. Start: $A_x = 3.70$, $B_x = -1.00 \rightarrow \text{SSE} = 0.0482 \text{ m}^2$.
2. **PREDICT:** should A_x go up or down?
3. Change ONE yellow cell. Did SSE fall?
4. Keep going until SSE stops falling. Watch the residual chart as you go.

Understand the objective first. Then let software solve it.

Problem



Model



Compute



Analyze



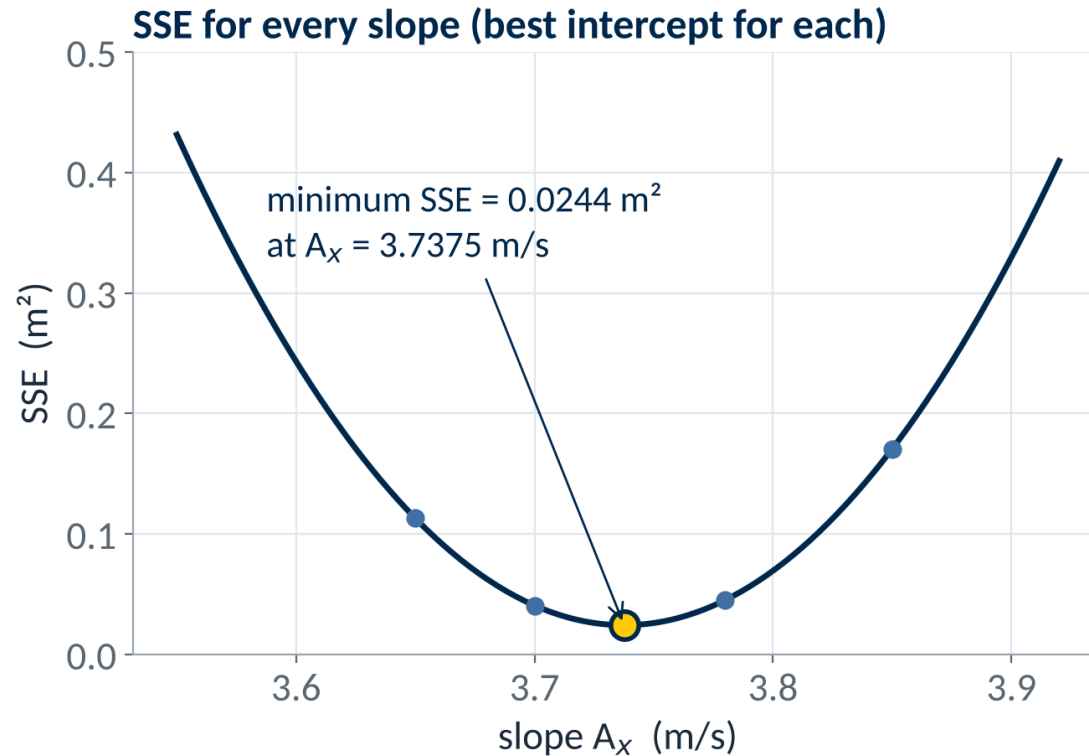
Validate



Improve



Regression: the computer minimizes the SSE you defined



WHAT SOFTWARE RETURNS

$$A_x = 3.7375 \text{ m/s}$$

$$B_x = -1.0247 \text{ m}$$

SSE = 0.0244 m² — the bottom of the bowl. Trendline, SLOPE/INTERCEPT and LINEST all agree. Line 3 was this line.

WHAT IT DID — AND DID NOT DO

It found the coefficients that minimize the quantity **we** chose, for the model **we** chose. It did not find “the truth.”

LINEAR REGRESSION least squares applied to a straight-line model. The algorithm is the definition; the software is the tool.

Problem ›

Model ›

Compute ›

Analyze ›

Validate ›

Improve



$$x(t) \approx 3.74 t - 1.02 \quad (x \text{ in m, } t \text{ in s})$$

$$A_x \approx 3.74 \text{ m/s}$$

Units m/s: a velocity.

Under the linear model, this is the ball's horizontal velocity in the camera's calibrated plane: about 3.7 m every second toward the hoop.

A fitted, average value for the whole flight — not a measurement at any single instant.

$$B_x \approx -1.02 \text{ m}$$

The line's x at Tracker t = 0.

That is 0.34 s BEFORE release — outside the measured flight. The ball was still in the shooter's hands.

Not the release position: at release (frame 41) the ball was measured at x = 0.26 m.

Use the line inside the measured range only: t = 0.342 to 1.383 s. Outside it, the line is extrapolating — the Chapter 6 guard still applies.

A coefficient means what the model and its time origin say it means. Hold that thought: Part 3 needs it.

Problem ›

Model ›

Compute ›

Analyze ›

Validate ›

Improve



One number for the fit — and what it means

161.7 m²

SST

how much x varies around its average

0.0244 m²

SSE

what the best line leaves unexplained

0.99985

$R^2 = 1 - \text{SSE}/\text{SST}$

fraction of the variation accounted for

READ IT AS

The line accounts for 99.985 % of how much x changes over the flight.

R^2 IS EVIDENCE — NOT A CERTIFICATE

It says how well the chosen model fits these data. It does not say the model is physically right.

High R^2 answers one question: "Does the line follow the data's overall trend?" Nothing more.

Problem



Model



Compute



Analyze



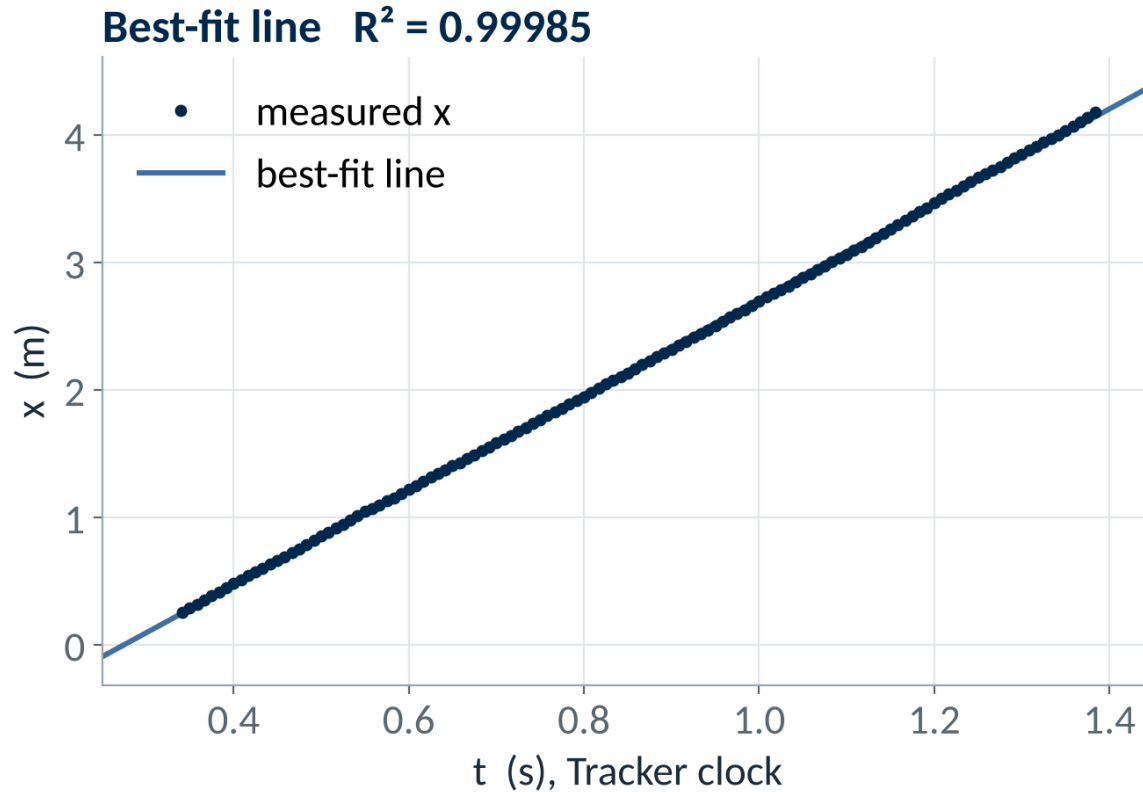
Validate



Improve



Is this conclusion justified by the evidence?



AI ANALYST

“ $R^2 = 0.99985$, so the linear model is essentially perfect. No further analysis of $x(t)$ is needed.”

VOTE: ACCEPT · MODIFY · REJECT

Then name ONE piece of evidence you would want to see before signing off on that sentence.

The plot looks perfect. R^2 looks perfect. What have we NOT looked at yet?

Problem



Model



Compute



Analyze



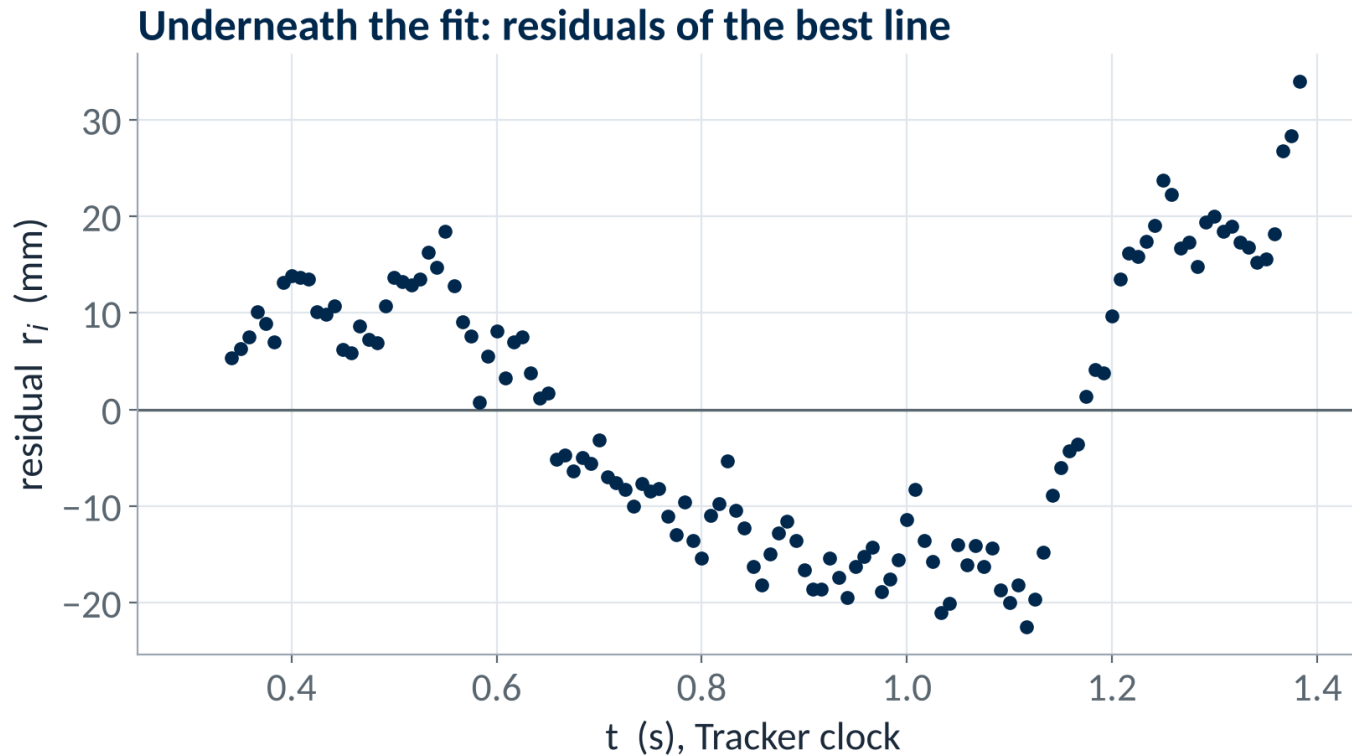
Validate



Improve



The fit looks excellent. Now look at what it left over.



WHAT DO YOU NOTICE?

Say it before anyone explains it.

WHAT WE CAN SAY

Not random scatter: positive $\rightarrow$ negative $\rightarrow$ positive. Only 3 sign changes in 126 points. RMS 13.9 mm; largest 34 mm.

Something systematic is unexplained. Its cause is not decided here.

$R^2 \approx 1$ and structured residuals can coexist. The fit is excellent. The investigation is not finished.

Problem



Model



Compute



Analyze

**Validate**

Improve



The best line is defined, found and inspected

MODEL ›

PREDICT $\hat{x}$ ›

RESIDUALS ›

SSE ›

MINIMIZE ›

 A_x, B_x ›**RESIDUAL PLOT**

LEAST SQUARES

chooses the coefficients. YOU chose the model and the definition of "best."

$$R^2 = 0.99985$$

is strong evidence that the line follows the overall trend.

THE RESIDUAL PLOT

shows what R^2 cannot: systematic structure, a few centimetres in size.

NEXT — PART 3 · THE VERTICAL MOTION

A straight line will not do for $y(t)$. What happens when the model needs more than a line — and do the fitted coefficients carry physics?

And why did the best line put the ball at $x = -1.02$ m before it was even thrown?

FITS ≠ VALIDATED — first evidence. An excellent fit and an unexplained pattern, in the same 126 points.

Problem ›

Model ›

Compute ›

Analyze ›

Validate ›

Improve