Welcome to...

Sensing and Observation Theory

part 2

Modelling, Uncertainty, and Data for Engineers

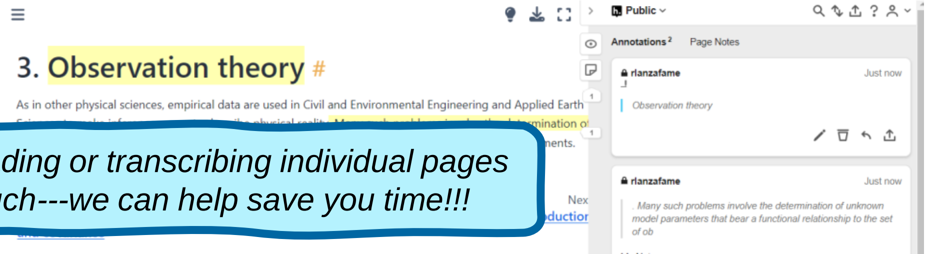
WEEK 4 !

Week 1.4 Announcements

- Assignment Portfolio: Deadlines and Points (reminder)
  - Group Assignments: submit each Friday at 12:30; feedback returned middle of following week
  - BuddyCheck (finish by 11:00 Monday – they are now worth points!). No late submissions allowed.
  - PA: best to finish before Friday each week; ultimate deadline for points is Week 1.9, Monday (Oct 28)
- Solutions from last week online; also, a widget to explore model parameters, see14
- `.../files/GA_1_3/`
- Programming Tutorials Continue: Monday at 10:45, Room 1.9**8** (focus is on basics programming skills)
- PA1.4: access link on MUDE Files page: `.../files/Week_1_4/README.html`

Miss taking notes in a textbook? Make the book "yours" with Hypothesis Browser Extension:

https://web.hypothes.is/start/ (→ later this year we will provide you a way to "save" your notes/book)



Note: if you are downloading or transcribing individual pages of the book, get in touch---we can help save you time!!!

Review

Estimators - overview

Functional model: $\mathbb{E}(Y) = \mathrm{A} \cdot \mathrm{x}$

Stochastic model: $\mathbb{D}(Y) = \Sigma_Y$

- Weighted Least-Squares estimation : minimizing weighted sum of squared errors

  allows to give different weights to observations

  $$\hat{X} = (\mathrm{A}^T \mathrm{W} \mathrm{A})^{-1} \mathrm{A}^T \mathrm{W} \cdot Y$$

- Best Linear Unbiased estimation : $\min \left(trace(\Sigma_{\hat{X}}) \right)$ (best), $\hat{X} = \mathrm{L}^T \cdot Y$ (linear), $\mathbb{E}(\hat{X}) = \mathrm{x}$ (unbiased)

  $$\hat{X} = \left(\mathrm{A}^T \Sigma_Y^{-1} \mathrm{A}\right)^{-1} \mathrm{A}^T \Sigma_Y^{-1} Y$$

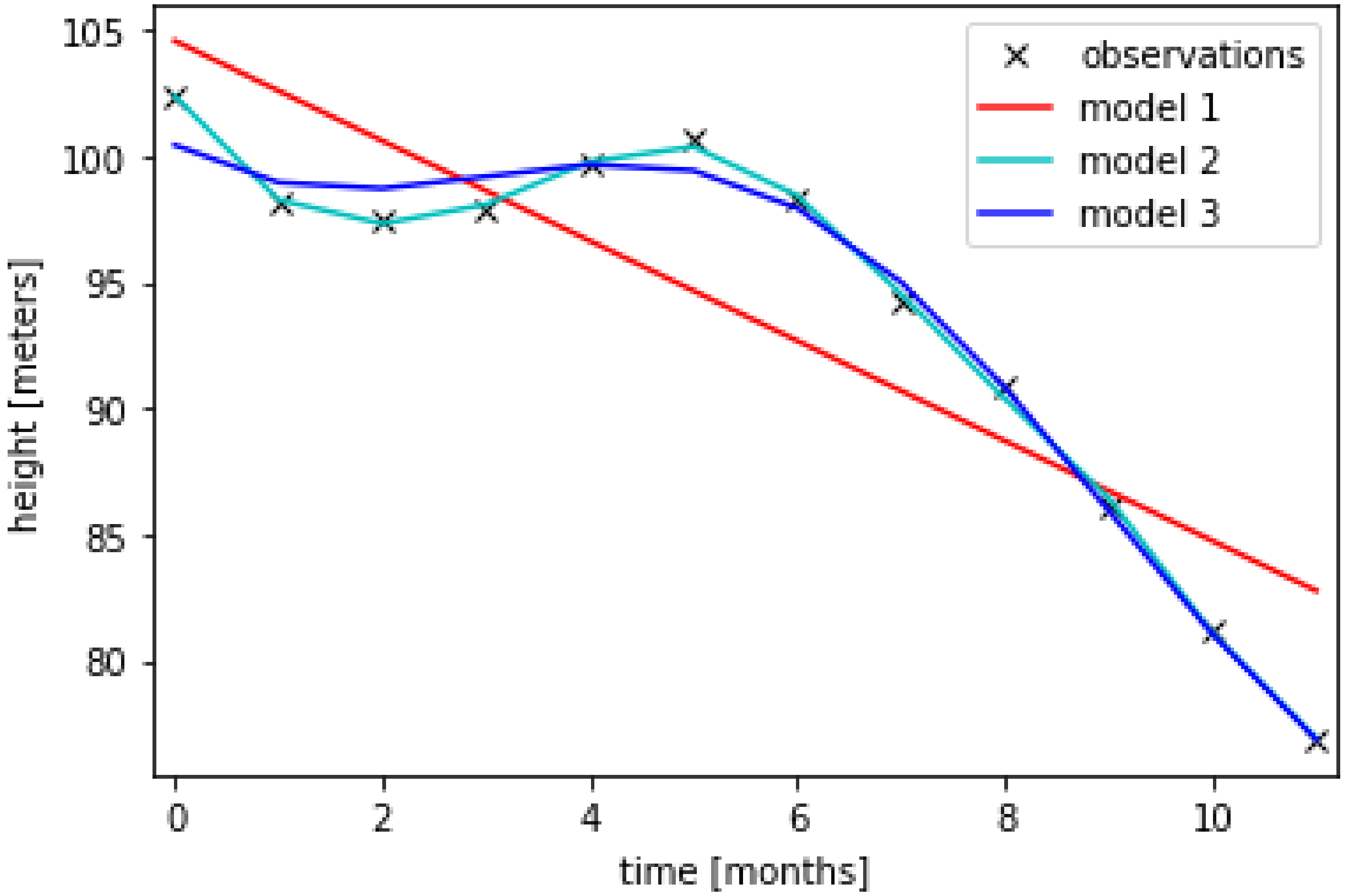
- Maximum Likelihood estimation : most likely x for given y,

  for normally distributed data same as BLUE

TU Delft

Results from notebooks

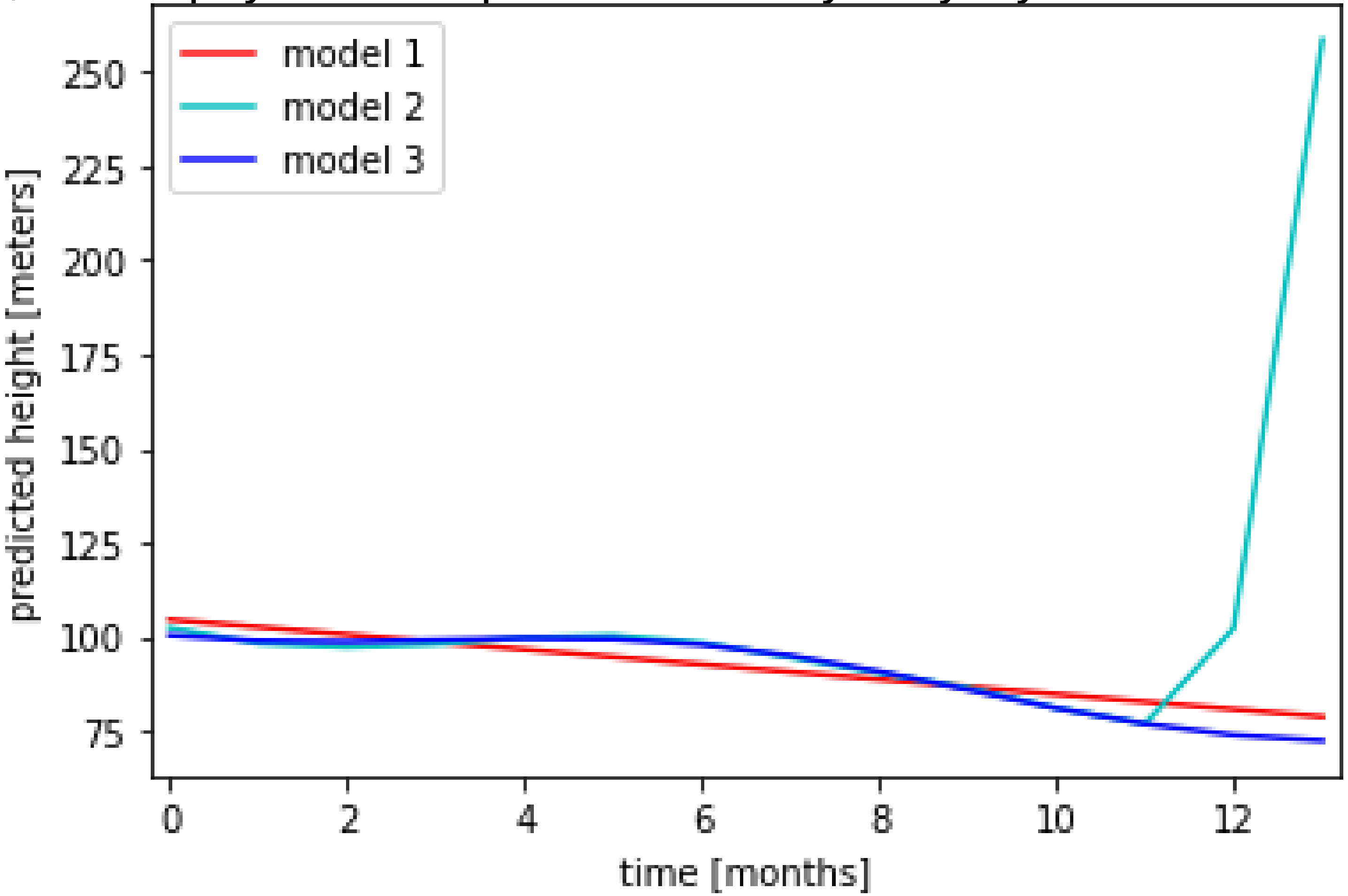
- Underfitting: model too simplistic, does not capture the real signal
- Overfitting: nearly perfect fit, but no physical interpretation



TU Delft

Results from notebooks

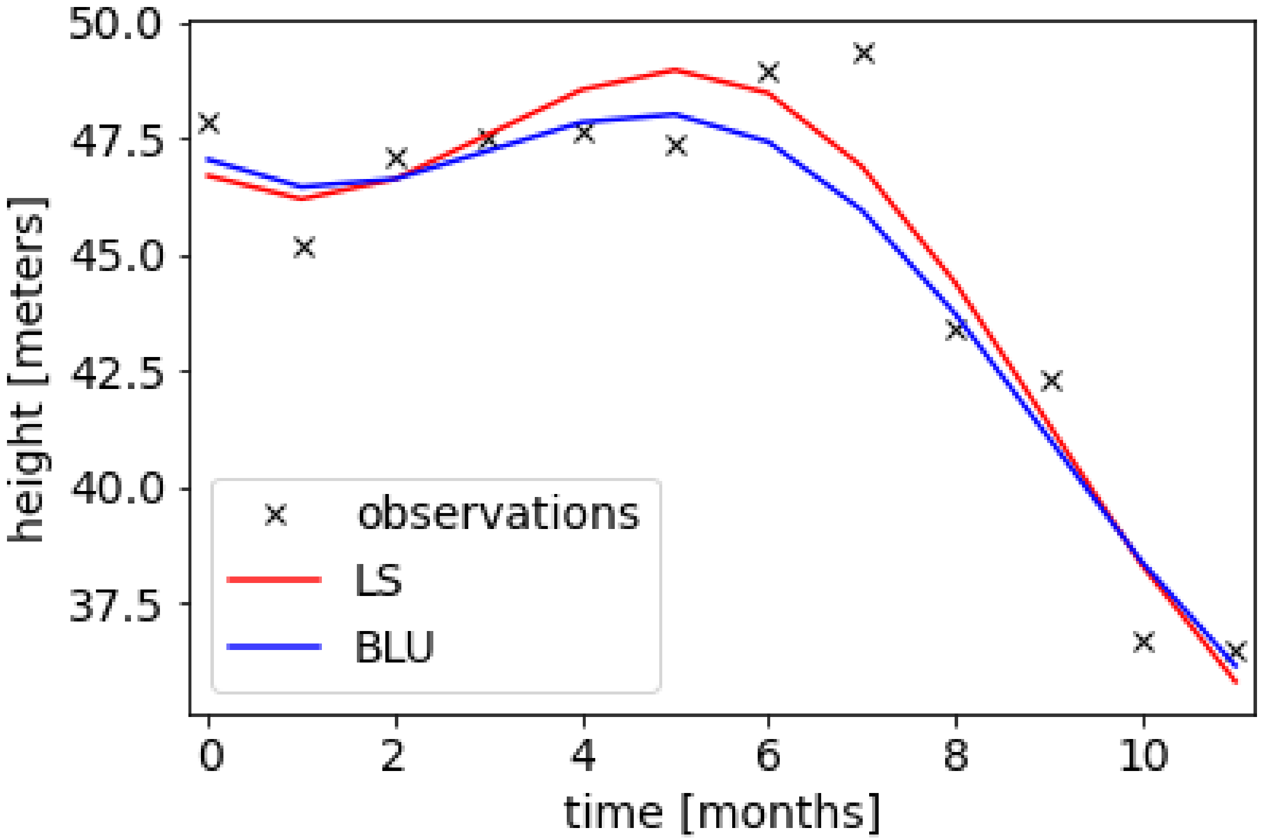
- Underfitting: model too simplistic, does not capture the real signal
- Overfitting: nearly perfect fit, but no physical interpretation → very risky if you use model for prediction



Best linear unbiased estimator = best weighted least squares estimator

Weight matrix is inverse covariance matrix:

Makes sense: high precision → small variance → large weight



$$Y = \mathrm{A} \cdot \mathrm{x} + \epsilon$$
$$\mathbb{D}(Y) = \Sigma_Y = \Sigma_\epsilon$$

Question that may have popped up:

Where does Σ_Y come from?

→ Calibration:

- Repeated measurements
- Calculate standard deviation

Usually observables are assumed to be independent, since the random errors are independent (error of observation Y_i does not depend on the error of observation Y_j

When would observable be dependent?

- due to signal processing in sensor (often when sampling rate is too high)
- if we use *differential* observables
- if we apply a common correction to our observations which is *stochastic*

TU Delft

Open questions

- (How to come up with a model?)
- What if my model is non-linear?
- Does my model really fit?
- Which models fits best?

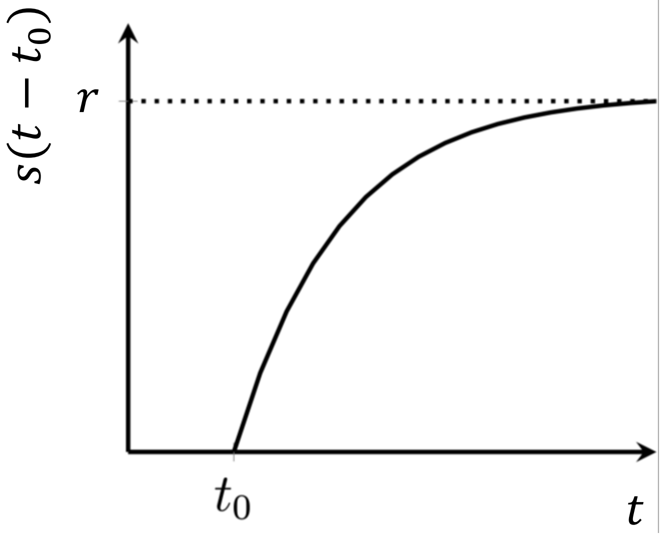
TU Delft

What if my observation equations are non-linear?

Observed: ground water level rise due to rainfal

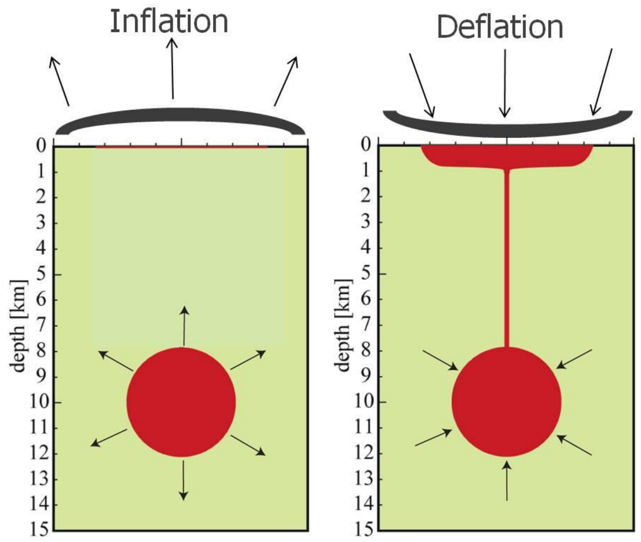
$$E(Y_i) = p \cdot \underbrace{r\left(1 - \exp\left(-\frac{t-t_0}{a}\right)\right)}_{s(t-t_0)}$$

- Known parameter:
  - p [m]: constant water inflow during rain event
- Unknown parameters:
  - **scaling parameter a [days] (memory of system),**
  - **response r [m/m] of the aquifer depending on the amount of rainfall**

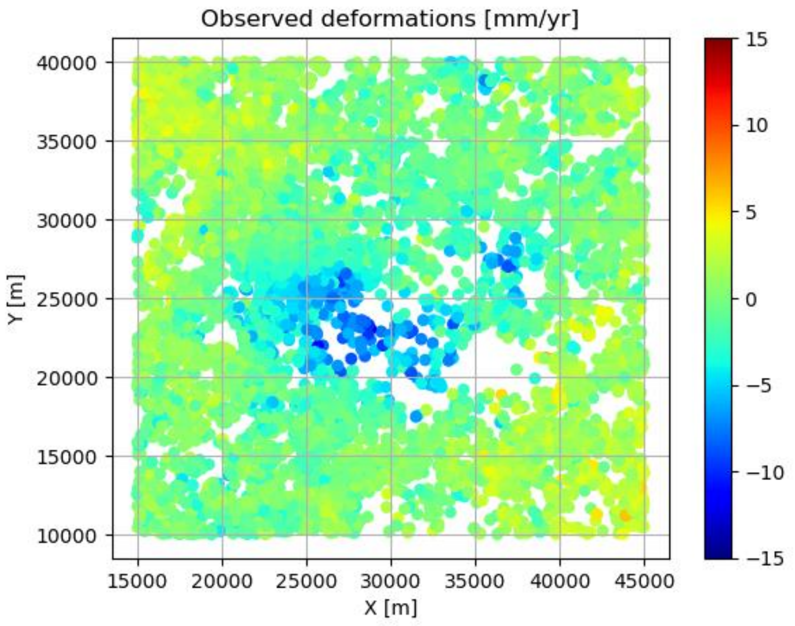


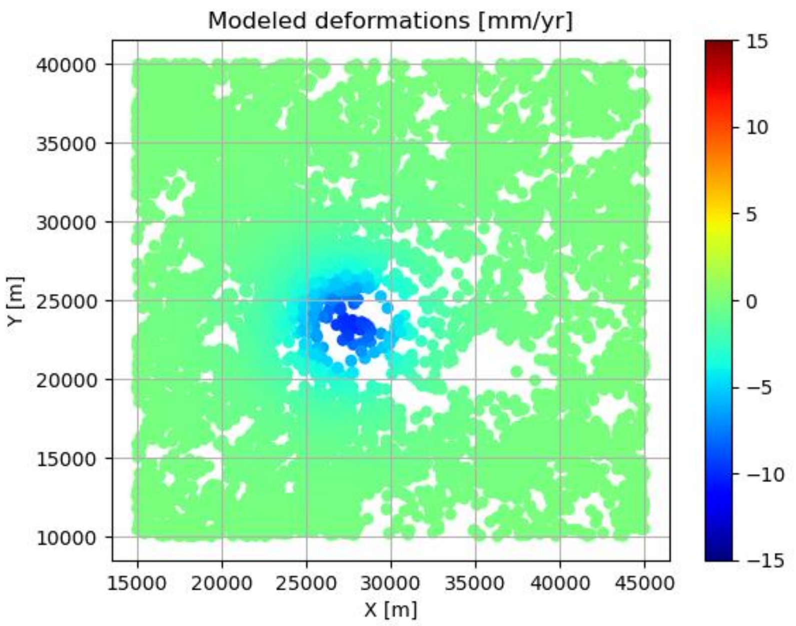
Volcano deformation rates at known locations (x_i, y_i)

$$\mathbb{E}(Y_i) = \frac{0.73\Delta V}{\pi d^2} \cdot \left(1 + \frac{1}{d^2}((x_i - x_s)^2 + (y_i - y_s)^2)\right)^{-\frac{3}{2}}$$



Unknown parameters: **volume change ΔV, depth of magma chamber d, (x_s, y_s) horizontal coordinates of centre**





Linearized observation equation using 1st order Taylor approximation
1 observation 1 unknown

$$y = q(x) + \epsilon \approx q(x_{[0]}) + \partial_x q(x_{[0]})(x - x_{[0]}) + \epsilon$$

initial guess

for now: omit ϵ from equations

$$\Delta y = y - q(x_{[0]}) \approx \partial_x q(x_{[0]}) \underbrace{(x - x_{[0]})}_{\Delta x}$$

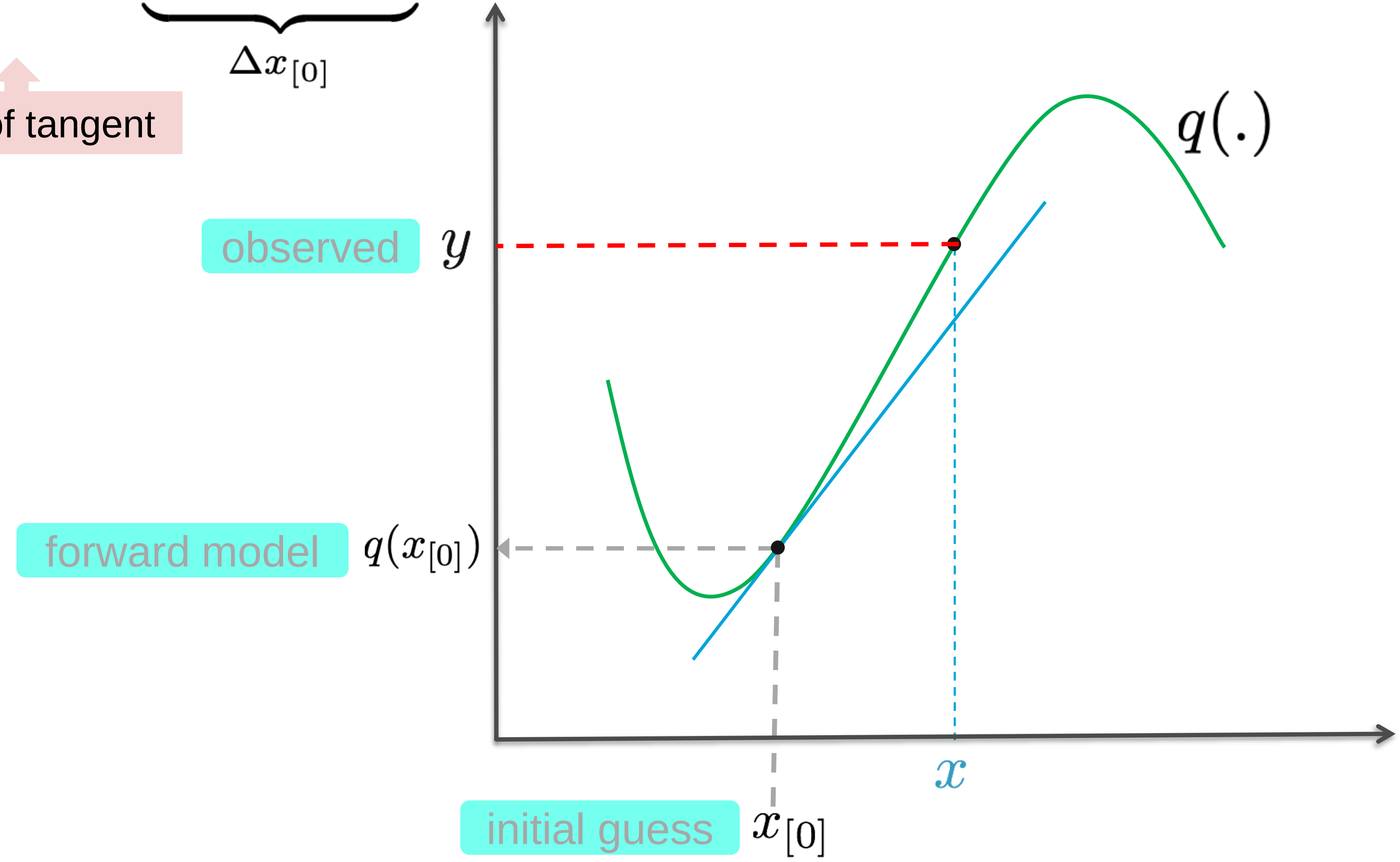
observed-minus-computed

Input:

- observation y
- initial guess $x_{[0]}$

$$\Delta y_{[0]} = y - q(x_{[0]}) \approx \partial_x q(x_{[0]}) \underbrace{(x_{[1]} - x_{[0]})}_{\Delta x_{[0]}}$$

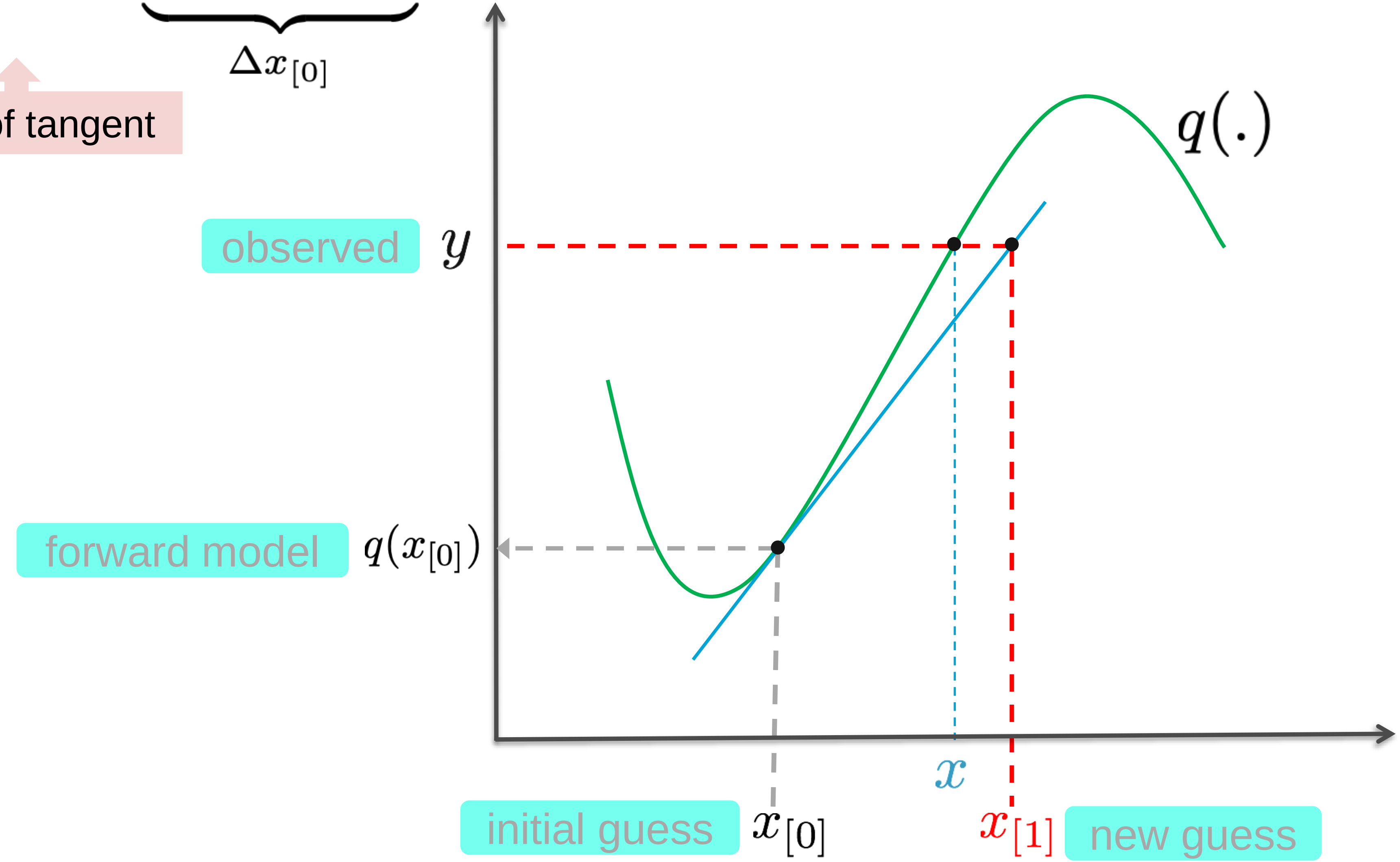
slope of tangent



Input:
- observation y
- initial guess $x_{[0]}$

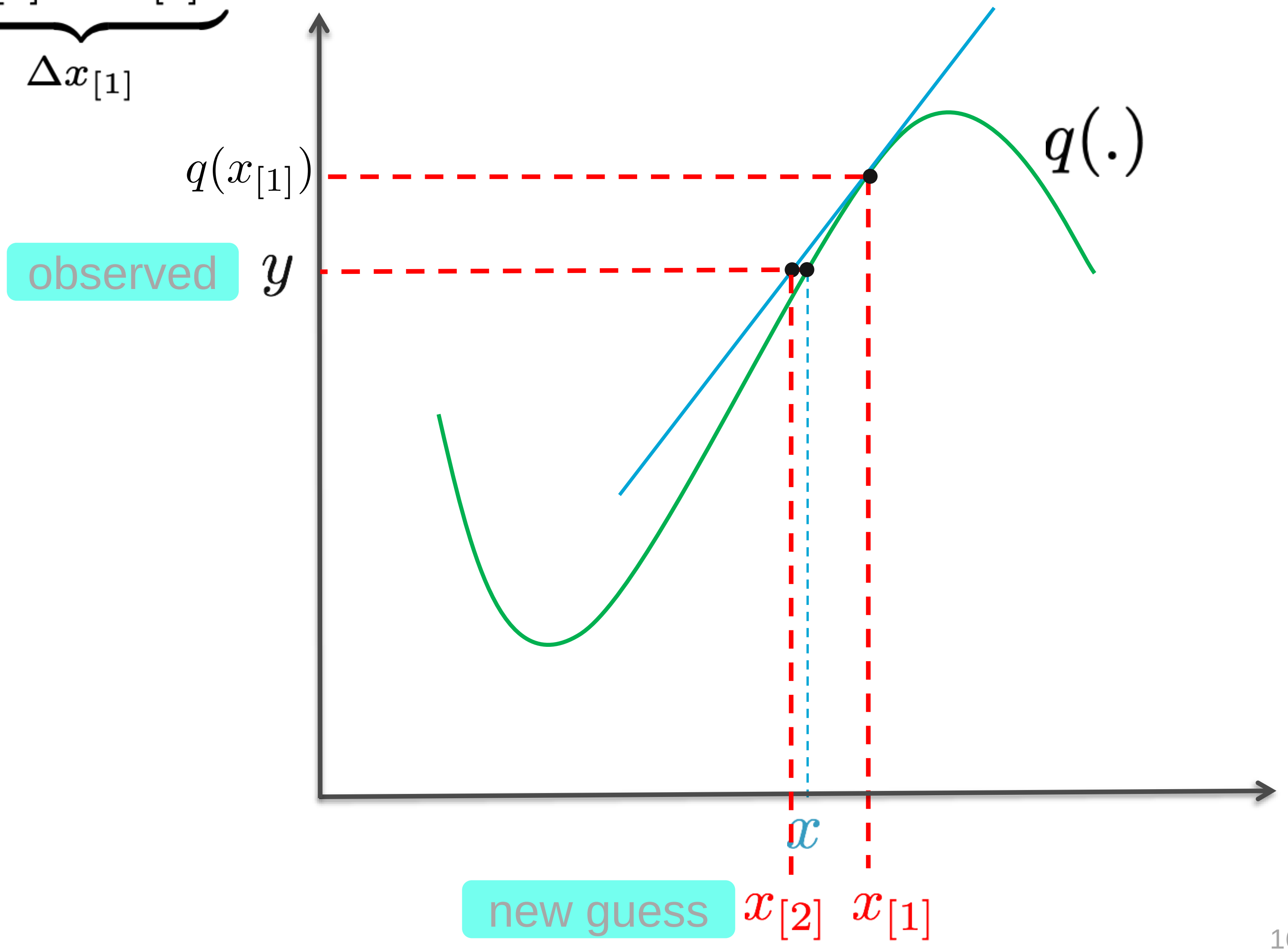
$$\Delta y_{[0]} = y - q(x_{[0]}) \approx \partial_x q(x_{[0]}) \underbrace{(x_{[1]} - x_{[0]})}_{\Delta x_{[0]}}$$

slope of tangent



Input:
- observation y
- new guess $x_{[1]}$

$$\Delta y_{[1]} = y - q(x_{[1]}) \approx \partial_x q(x_{[1]}) \underbrace{(x_{[2]} - x_{[1]})}_{\Delta x_{[1]}}$$



→ Gauss-Newton iteration

Continue until $\Delta x_{[i]}$ is very small

TU Delft

Linearized observation equation using 1st order Taylor approximation
1 observation n unknowns

$$\Delta y_{[i]} = y - \underset{n \times 1}{q(\mathrm{x}_{[i]})} \approx \partial_{\mathrm{x}} q(\mathrm{x}_{[i]}) \underbrace{(\mathrm{x} - \mathrm{x}_{[i]})}_{\Delta \mathrm{x}_{[}i]}$$

$$= \begin{bmatrix} \partial_{x_1} q(\mathrm{x}_{[i]}) & \partial_{x_2} q(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q(\mathrm{x}_{[i]}) \end{bmatrix} \begin{bmatrix} x_1 - x_{1,[i]} \\ x_2 - x_{2,[i]} \\ \vdots \\ x_n - x_{n,[i]} \end{bmatrix}$$

i is the iteration index

TU Delft

Non-linear functional model

$$\mathbb{E}\left(\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix}\right) = \begin{bmatrix} q_1(\mathrm{x}) \\ q_2(\mathrm{x}) \\ \vdots \\ q_m(\mathrm{x}) \end{bmatrix}$$

Linearized functional model

$$\mathbb{E}\left(\begin{bmatrix} \Delta Y_1 \\ \Delta Y_2 \\ \vdots \\ \Delta Y_m \end{bmatrix}_{[i]}\right) = \begin{bmatrix} \partial_{x_1} q_1(\mathrm{x}_{[i]}) & \partial_{x_2} q_1(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_1(\mathrm{x}_{[i]}) \\ \partial_{x_1} q_2(\mathrm{x}_{[i]}) & \partial_{x_2} q_2(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_2(\mathrm{x}_{[i]}) \\ \vdots & \vdots & & \vdots \\ \partial_{x_1} q_m(\mathrm{x}_{[i]}) & \partial_{x_2} q_m(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_m(\mathrm{x}_{[i]}) \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{bmatrix}_{[i]}$$

Jacobian ← takes the role of design matrix A

Gauss-Newton iteration

Start with initial guess $\mathrm{x}_{[0]}$, and start iteration with $i = 0$

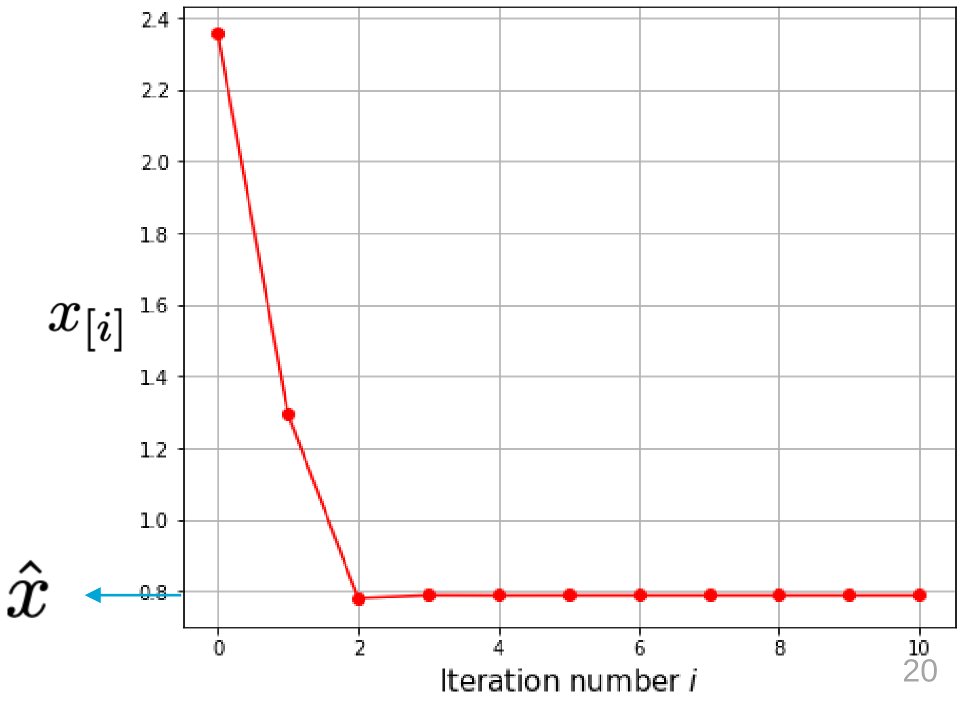
1. Calculate observed-minus-computed $\Delta \mathrm{y}_{[i]}$
2. Determine the Jacobian
3. Estimate $\Delta\hat{\mathrm{x}}_{[i]}$ by applying BLUE
4. New guess $\mathrm{x}_{[i+1]} = \Delta\hat{\mathrm{x}}_{[i]} + \mathrm{x}_{[i]}$

WHEN TO STOP?

$$\mathbb{E}\left(\begin{bmatrix}\Delta Y_1\\ \Delta Y_2\\ \vdots\\ \Delta Y_m\end{bmatrix}_{[i]}\right) = \begin{bmatrix}\partial_{x_1} q_1(\mathrm{x}_{[i]}) & \partial_{x_2} q_1(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_1(\mathrm{x}_{[i]})\\ \partial_{x_1} q_2(\mathrm{x}_{[i]}) & \partial_{x_2} q_2(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_2(\mathrm{x}_{[i]})\\ \vdots & \vdots & & \vdots\\ \partial_{x_1} q_m(\mathrm{x}_{[i]}) & \partial_{x_2} q_m(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_m(\mathrm{x}_{[i]})\end{bmatrix}\begin{bmatrix}\Delta x_1\\ \Delta x_2\\ \vdots\\ \Delta x_n\end{bmatrix}_{[i]}$$

Jacobian ← takes the role of design matrix A

Convergence



TU Delft

Gauss-Newton iteration

Start with initial guess $\mathrm{x}_{[0]}$, and start iteration with $i = 0$

1. Calculate observed-minus-computed $\Delta\mathrm{y}_{[i]}$
2. Determine the Jacobian
3. Estimate $\Delta\hat{\mathrm{x}}_{[i]}$ by applying BLUE
4. New guess $\mathrm{x}_{[i+1]} = \Delta\hat{\mathrm{x}}_{[i]} + \mathrm{x}_{[i]}$

WHEN TO STOP?

$$\mathbb{E}\left(\begin{bmatrix}\Delta Y_1\\ \Delta Y_2\\ \vdots\\ \Delta Y_m\end{bmatrix}_{[i]}\right) = \begin{bmatrix}\partial_{x_1} q_1(\mathrm{x}_{[i]}) & \partial_{x_2} q_1(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_1(\mathrm{x}_{[i]})\\ \partial_{x_1} q_2(\mathrm{x}_{[i]}) & \partial_{x_2} q_2(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_2(\mathrm{x}_{[i]})\\ \vdots & \vdots & & \vdots\\ \partial_{x_1} q_m(\mathrm{x}_{[i]}) & \partial_{x_2} q_m(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_m(\mathrm{x}_{[i]})\end{bmatrix}\begin{bmatrix}\Delta x_1\\ \Delta x_2\\ \vdots\\ \Delta x_n\end{bmatrix}_{[i]}$$

Jacobian ← takes the role of design matrix A

Gauss-Newton iteration

Stop criterion

$$\Delta\hat{\mathrm{x}}_{[i]}^T \cdot \Sigma_{\hat{\mathrm{x}}}^{-1} \cdot \Delta\hat{\mathrm{x}}_{[i]} < \text{small value}$$

an estimated parameter with small variance should have a relatively small deviation compared to a parameter with large variance

Start with initial guess $\mathrm{x}_{[0]}$, and start iteration with $i = 0$

1. Calculate observed-minus-computed $\Delta\mathrm{y}_{[i]}$
2. Determine the Jacobian
3. Estimate $\Delta\hat{\mathrm{x}}_{[i]}$ by applying BLUE
4. New guess $\mathrm{x}_{[i+1]} = \Delta\hat{\mathrm{x}}_{[i]} + \mathrm{x}_{[i]}$
5. If **stop criterion** is met: set $\hat{\mathrm{x}} = \mathrm{x}_{[i+1]}$ and break, otherwise set $i := i + 1$ and go to step 1

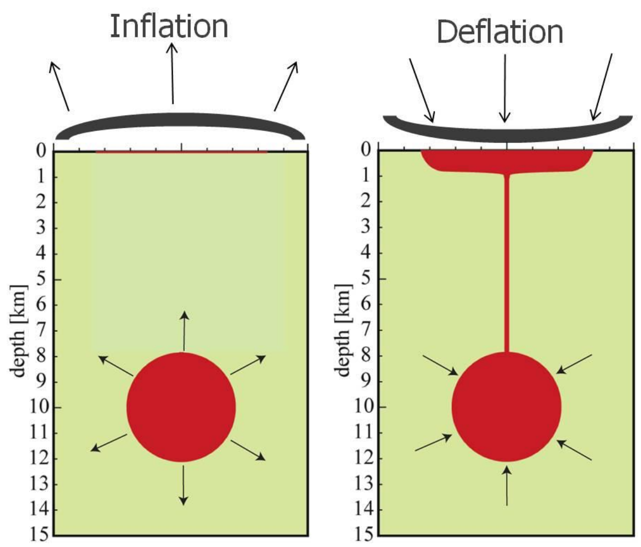
$$\mathbb{E}\left(\begin{bmatrix}\Delta Y_1 \\ \Delta Y_2 \\ \vdots \\ \Delta Y_m\end{bmatrix}_{[i]}\right) = \begin{bmatrix}\partial_{x_1} q_1(\mathrm{x}_{[i]}) & \partial_{x_2} q_1(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_1(\mathrm{x}_{[i]}) \\ \partial_{x_1} q_2(\mathrm{x}_{[i]}) & \partial_{x_2} q_2(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_2(\mathrm{x}_{[i]}) \\ \vdots & \vdots & & \vdots \\ \partial_{x_1} q_m(\mathrm{x}_{[i]}) & \partial_{x_2} q_m(\mathrm{x}_{[i]}) & \cdots & \partial_{x_n} q_m(\mathrm{x}_{[i]})\end{bmatrix}\begin{bmatrix}\Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n\end{bmatrix}_{[i]}$$

Jacobian ← takes the role of design matrix A

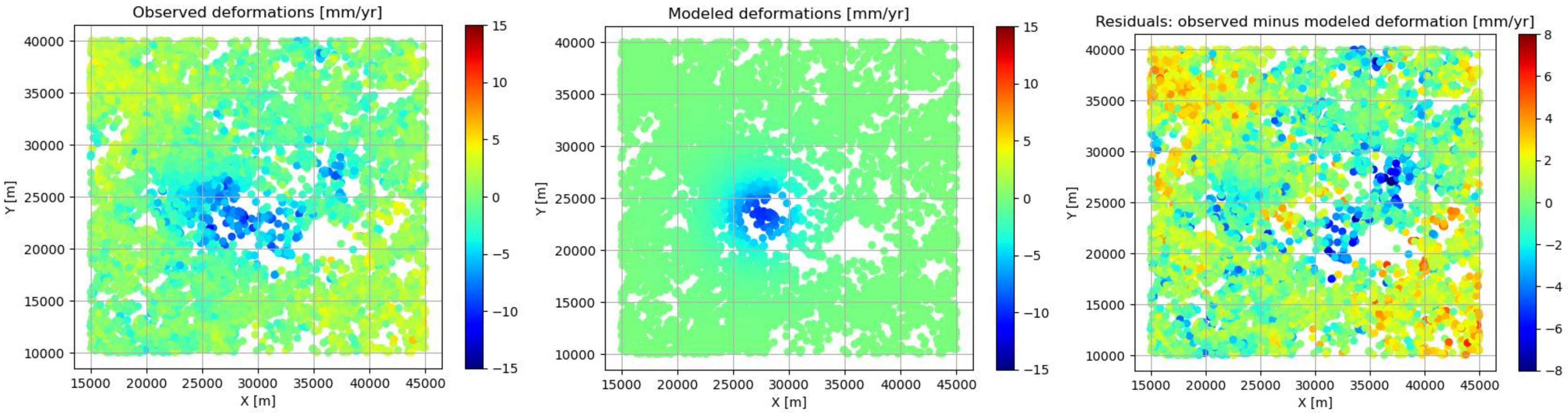
TU Delft

Volcano deformation rates at known locations (x_i, y_i)

$$\mathbb{E}(Y_i) = \frac{0.73\Delta V}{\pi d^2} \cdot \left(1 + \frac{1}{d^2}\left((x_i - x_s)^2 + (y_i - y_s)^2\right)\right)^{-\frac{3}{2}}$$



Unknown parameters: **volume change ΔV, depth of magma chamber d, (x_s, y_s) horizontal coordinates of centre**



Volcano deformation – precision of estimated parameters

observed deformations at (x_i, y_i) as function of volume change, depth, horiz. position of centre

$$\mathbb{E}(Y_i) = \frac{0.73\Delta V}{\pi d^2} \cdot (1 + \frac{1}{d^2}((x_i - x_s)^2 + (y_i - y_s)^2))^{-\frac{3}{2}}$$

$$\begin{bmatrix} \hat{\Delta V} \\ \hat{d} \\ \hat{x}_s \\ \hat{y}_s \end{bmatrix} = \begin{bmatrix} -552352.169\ m^3 \\ 3562.319\ m \\ 27528.535\ m \\ 23540.619\ m \end{bmatrix}$$

$$\begin{bmatrix} \sigma_{\hat{\Delta V}} \\ \sigma_{\hat{d}} \\ \sigma_{\hat{x}_s} \\ \sigma_{\hat{y}_s} \end{bmatrix} = \begin{bmatrix} 1582.769\ m^3 \\ 8.986\ m \\ 8.238\ m \\ 7.239\ m \end{bmatrix}$$

seems large, but look at units, and look at size compared to estimate !

TU Delft

Is it a good fit?

TU Delft

Sensing and observation theory - why

Needed for monitoring and prediction
e.g., natural processes, human-induced deformations, structural health, climate & environment, geo-energy and geo-resources, …

- Process measurements (= observations) to estimate parameters of interest
- In order to use estimation results for further analysis and interpretations (eventually to make decisions)

= uncertainty quantification

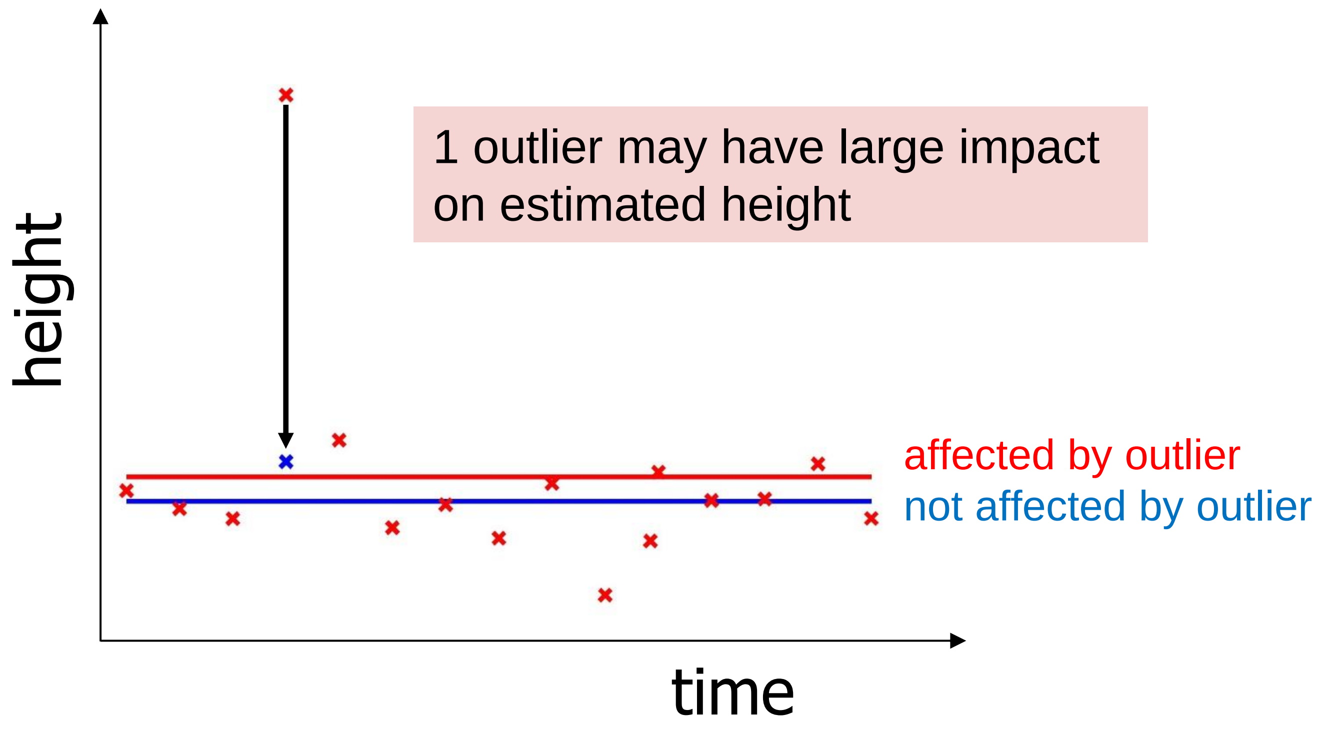
= detection of errors in data (outliers, systematic biases)
+ correction / adaption for these errors

= model validation

- detect model misspecifications
- multiple candidate models → decide which one is best

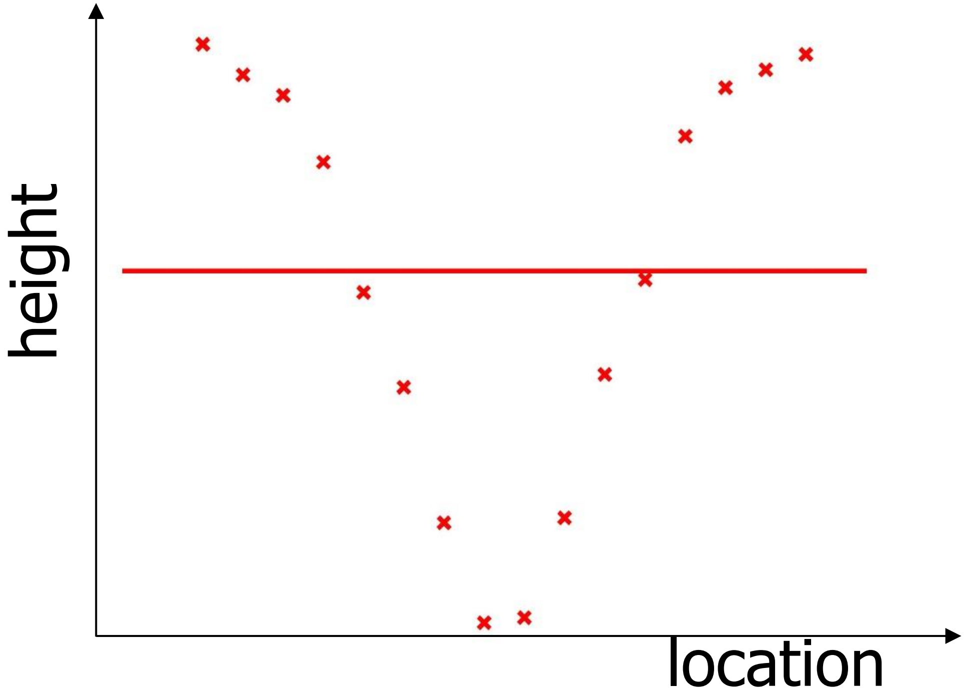
TU Delft

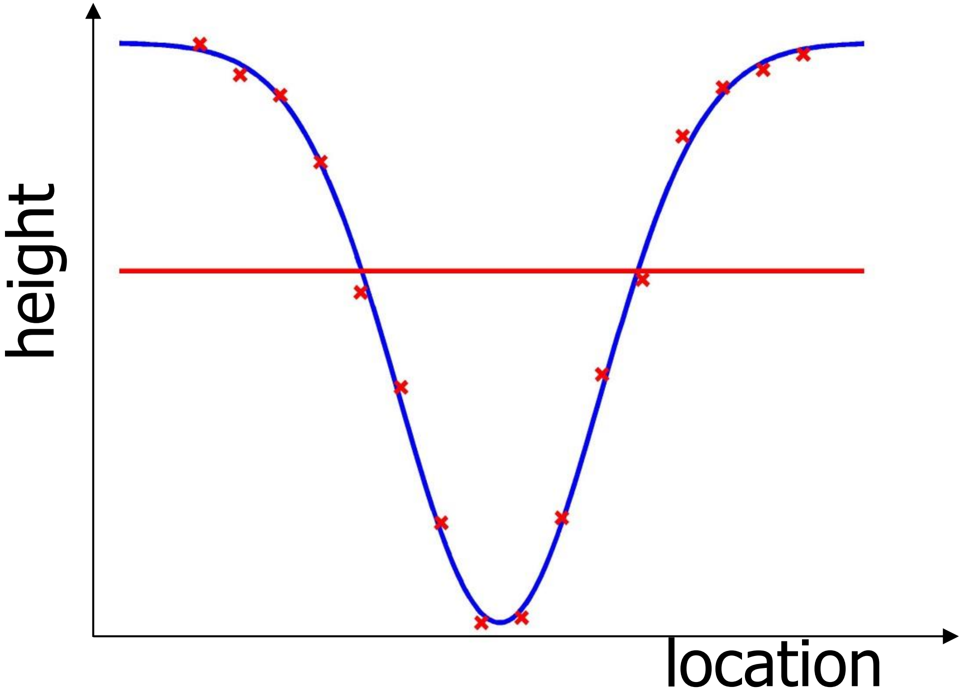
Example: outlier



Example: model misspecification

Wrong model → large residuals
(difference observations and fitted model)





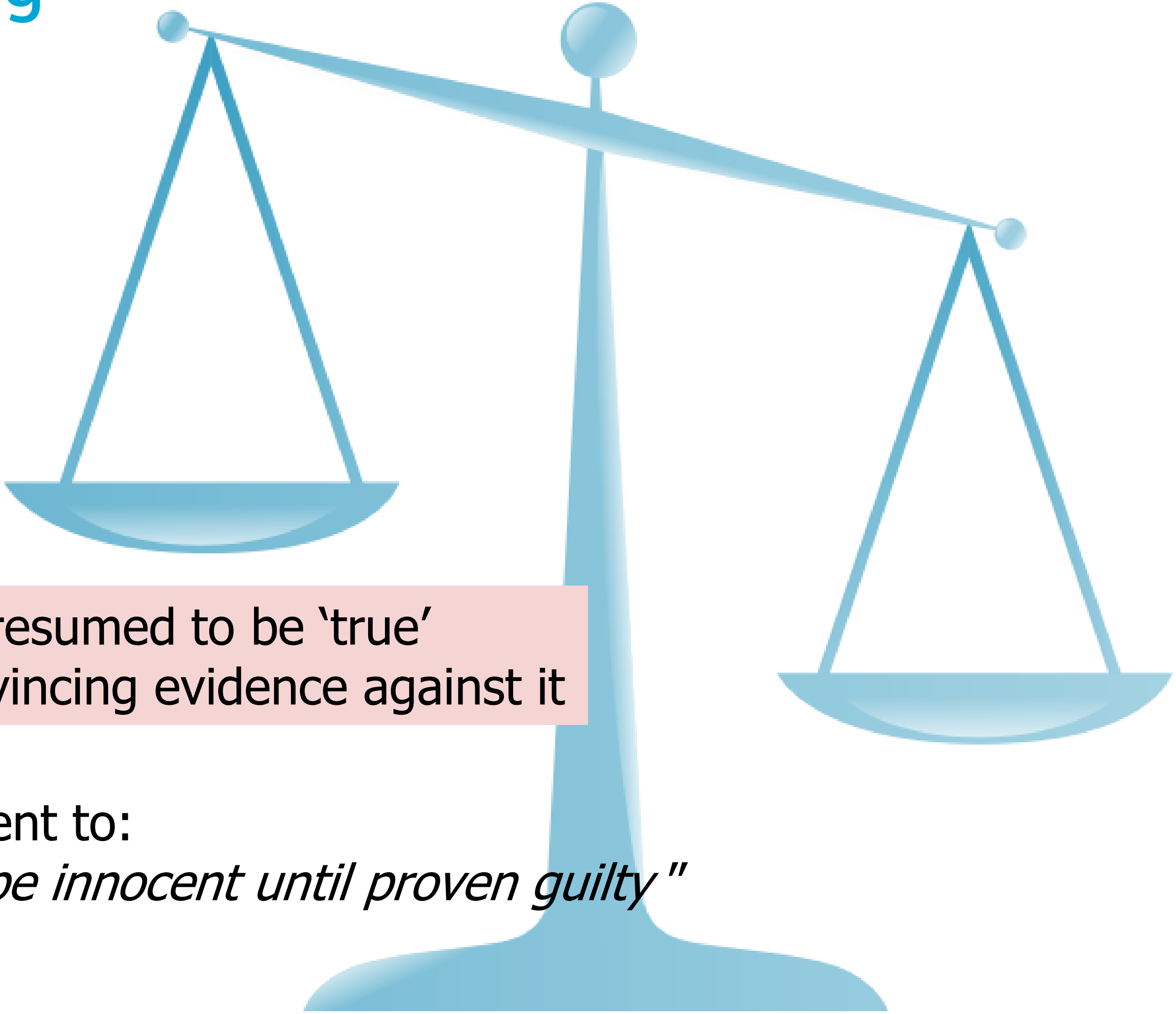
TU Delft

Statistical hypothesis testing

→test for compliance of model and data

Two competing hypothesis:

- Null hypothesis (nominal model): $\mathcal{H}_0$
- Alternative hypothesis: $\mathcal{H}_a$



Null hypothesis presumed to be 'true'
until data provide convincing evidence against it

equivalent to:
" *the defendant is presumed to be innocent until proven guilty* "

TU Delft

Project: Road deformation

$$\mathbb{E}(Y_i) = d_0 + vt_i + k \text{ GW}$$

$$\mathbb{E}(Y_i) = d_0 + R \left(1 - \exp \frac{-t_i}{a}\right) + k \text{ GW}$$

Apply non-linear least-squares

How to decide between the two models?

TU Delft

Enjoy…

Modelling, Uncertainty, and Data for Engineers

WEEK 4 !

TU Delft