# Reification: One Way to (formally) Incorporate Expert Judgement (I.e., Humans) in Uncertainty Quantification

**Leanna House**



# Consider a computer model describing...









Who is the expert? ... Any human with in the analytic process.

# What is “Human-in-the-Loop”?

# What is “Human-in-the-Loop”?

**Traditional Analytic Pipeline**

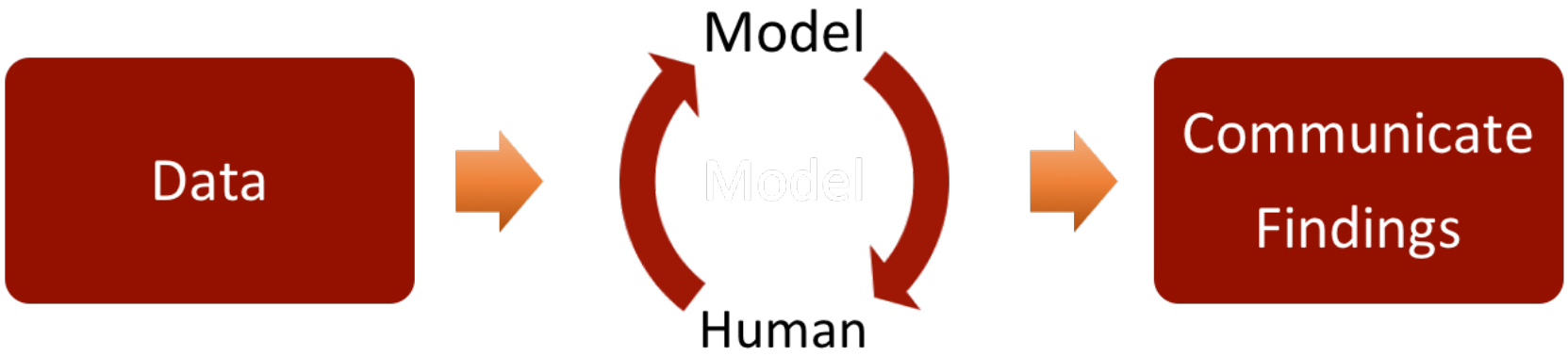


# What is “Human-in-the-Loop”?

**Traditional Analytic Pipeline**




**Human-in-the-Loop Analytic Pipeline**

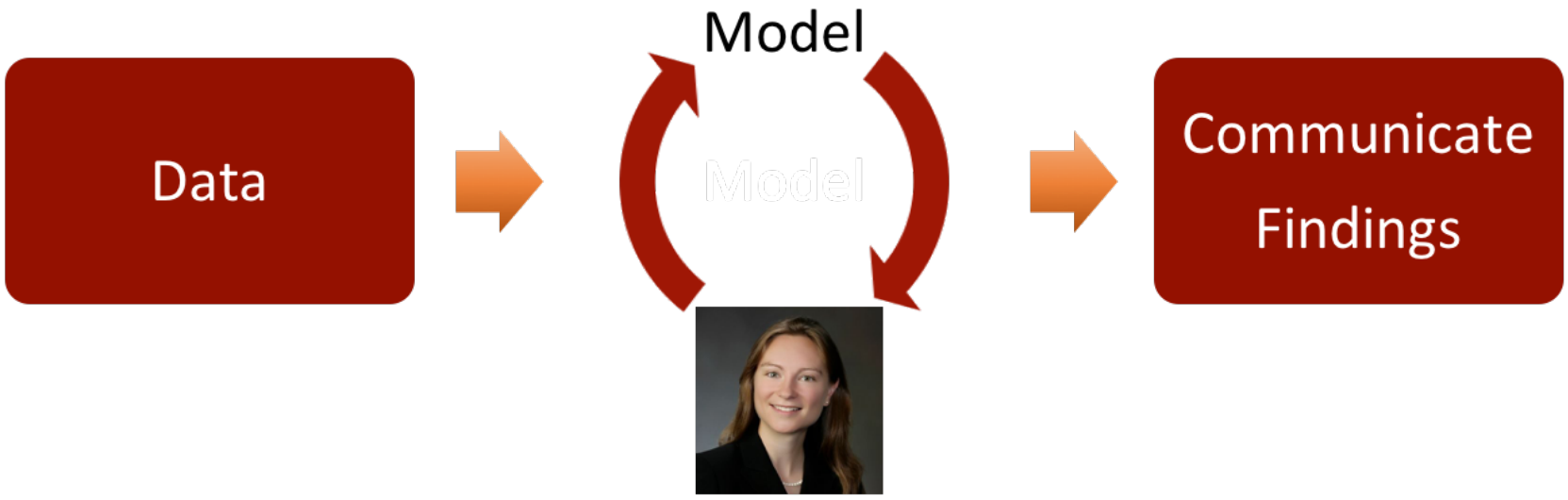


# What is “Human-in-the-Loop”?

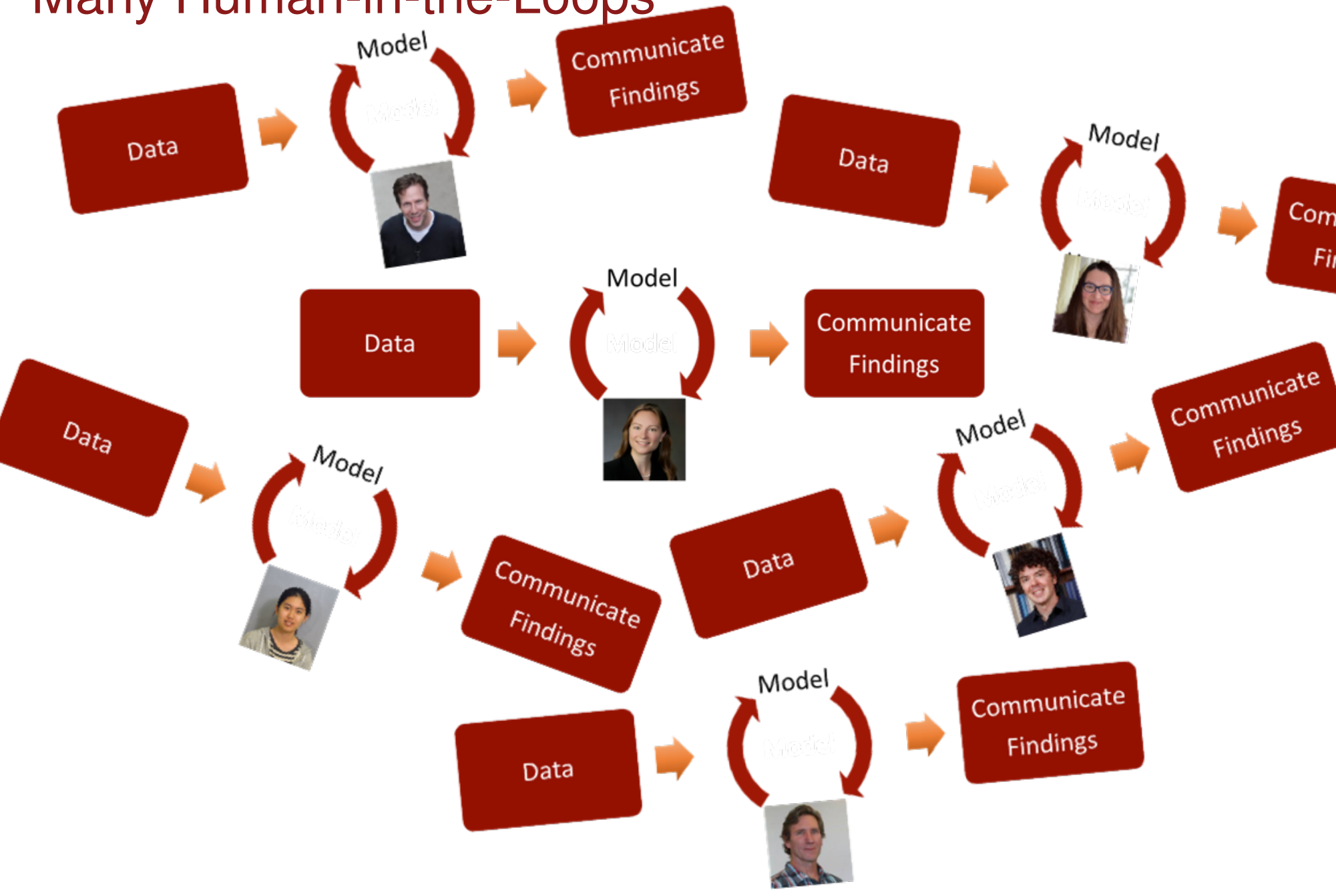
**Traditional Analytic Pipeline**




**Human-in-the-Loop Analytic Pipeline**



# Many Human-in-the-Loops



# Is it ok to have many loops?

- For some, this is a very old topic of conversation

# Is it ok to have many loops?

- For some, this is a very old topic of conversation
- Relevant again, in the era of big data

# Is it ok to have many loops?

- For some, this is a very old topic of conversation
- Relevant again, in the era of big data
- No: Implicit bias

# Is it ok to have many loops?

- For some, this is a very old topic of conversation
- Relevant again, in the era of big data
- No: Implicit bias
- Yes:
  - Of course, use all available information

# Is it ok to have many loops?

- For some, this is a very old topic of conversation
- Relevant again, in the era of big data
- No: Implicit bias
- Yes:
  - Of course, use all available information
  - ALWAYS cases when humans out perform machines

# Is it ok to have many loops?

- For some, this is a very old topic of conversation
- Relevant again, in the era of big data
- No: Implicit bias
- Yes:
  - Of course, use all available information
  - ALWAYS cases when humans out perform machines
    → By including humans, we identify cases and capitalize

# Is it ok to have many loops?

- For some, this is a very old topic of conversation
- Relevant again, in the era of big data
- No: Implicit bias
- Yes:
  - Of course, use all available information
  - ALWAYS cases when humans out perform machines
    → By including humans, we identify cases and capitalize
  - Human behavior and taking action

# Is it ok to have many loops?

- For some, this is a very old topic of conversation
- Relevant again, in the era of big data
- No: Implicit bias
- Yes:
  - Of course, use all available information
  - ALWAYS cases when humans out perform machines
    → By including humans, we identify cases and capitalize
  - Human behavior and taking action
    → Human involvement may alleviate uncertainty anxiety

# Is it ok to have many loops?

- For some, this is a very old topic of conversation
- Relevant again, in the era of big data
- No: Implicit bias
- Yes:
  - Of course, use all available information
  - ALWAYS cases when humans out perform machines
    → By including humans, we identify cases and capitalize
  - Human behavior and taking action
    → Human involvement may alleviate uncertainty anxiety
    → Enable action from findings

# Is it ok to have many loops?

- For some, this is a very old topic of conversation
- Relevant again, in the era of big data
- No: Implicit bias
- Yes:
  - Of course, use all available information
  - ALWAYS cases when humans out perform machines
    → By including humans, we identify cases and capitalize
  - Human behavior and taking action
    → Human involvement may alleviate uncertainty anxiety
    → Enable action from findings
  - ...
- Yes/No: Inference vs Data exploration

# Outline: Application of Reification

- Define
- Context
- Approach
- Results
- Discussion
- How?
- Wrap up

# What is Reification? First, notation

- $z$: Real Data
- $y$: Features of a true process
- $f$: Simulator
- $t$: Inputs to the simulator
- $D$: $\{F; T\}$, a set of simulator evaluations from design $T$

# What is Reification? First, notation

- $z$: Real Data
- $y$: Features of a true process
- $f$: Simulator
- $t$: Inputs to the simulator
- $D$: $\{F; T\}$, a set of simulator evaluations from design $T$
- *Very* simple model:

$$\begin{aligned} z|y,\theta &\sim \text{No}(y(\theta), {\phi_z}^{-1}) \\ y(t)|f,t &\sim \text{No}(f(t), {\phi_y}^{-1}) \\ f(t)|D,t &\sim \text{No}(m(t), p(t)^{-1}) \end{aligned}$$

# What is Reification? First, notation

- $z$: Real Data
- $y$: Features of a true process
- $f$: Simulator
- $t$: Inputs to the simulator
- $D$: $\{F; T\}$, a set of simulator evaluations from design $T$
- *Very* simple model:

$$\begin{aligned} z|y,\theta &\sim \text{No}(y(\theta), {\phi_z}^{-1}) \\ y(t)|f,t &\sim \text{No}(f(t), {\phi_y}^{-1}) \\ f(t)|D,t &\sim \text{No}(m(t), p(t)^{-1}) \end{aligned}$$

- $\theta$: Ideal value for $t$
- $\phi_z$,$\phi_y$: Precisions
- $m(t)$, $p(t)$: Posterior predictive parameters for the emulator.

# What is Reification? Answer

- Goal: Assess the uncertainty in our predictions from $f$ for $y$

# What is Reification? Answer

- Goal: Assess the uncertainty in our predictions from $f$ for $y$
- Alas, we know that $\mathrm{E}[y(t)|f,t] \neq f(t)$

# What is Reification? Answer

- Goal: Assess the uncertainty in our predictions from $f$ for $y$
- Alas, we know that $\mathrm{E}[y(t)|f,t] \neq f(t)$
- Introduce discrepancy terms: $\mathrm{E}[y(t)|f,t,\rho,\delta] = \rho f(t) + \delta$

# What is Reification? Answer

- Goal: Assess the uncertainty in our predictions from $f$ for $y$
- Alas, we know that $\mathrm{E}[y(t)|f,t] \neq f(t)$
- Introduce discrepancy terms: $\mathrm{E}[y(t)|f,t,\rho,\delta] = \rho f(t) + \delta$
- Consider a hypothetical, expert informed $f^*$

$$
\begin{aligned}
y(t)|f^*,t &\sim \mathrm{No}(f^*(t), \phi_y^{-1}) \\
f^*|f,t &\sim \mathrm{No}(m^*(t), p^*(t)^{-1}) \\
f(t)|D,t &\sim \mathrm{No}(m(t), p(t)^{-1})
\end{aligned}
$$

# What is Reification? Answer

- Goal: Assess the uncertainty in our predictions from $f$ for $y$
- Alas, we know that $\mathrm{E}[y(t)|f,t] \neq f(t)$
- Introduce discrepancy terms: $\mathrm{E}[y(t)|f,t,\rho,\delta] = \rho f(t) + \delta$
- Consider a hypothetical, expert informed $f^*$

$$\begin{aligned} y(t)|f^*,t &\sim \mathrm{No}(f^*(t),\phi_y^{-1}) \\ f^*|f,t &\sim \mathrm{No}(m^*(t),p^*(t)^{-1}) \\ f(t)|D,t &\sim \mathrm{No}(m(t),p(t)^{-1}) \end{aligned}$$

- "$f^*$ is a reified $f$"

# What is Reification? Answer

- Goal: Assess the uncertainty in our predictions from $f$ for $y$
- Alas, we know that $\mathrm{E}[y(t)|f,t] \neq f(t)$
- Introduce discrepancy terms: $\mathrm{E}[y(t)|f,t,\rho,\delta] = \rho f(t) + \delta$
- Consider a hypothetical, expert informed $f^*$

$$\begin{aligned} y(t)|f^*,t &\sim \mathrm{No}(f^*(t),\phi_y^{-1}) \\ f^*|f,t &\sim \mathrm{No}(m^*(t),p^*(t)^{-1}) \\ f(t)|D,t &\sim \mathrm{No}(m(t),p(t)^{-1}) \end{aligned}$$

- "$f^*$ is a reified $f$"

Reification (Goldstein and Rougier, 2009):
The process of eliciting $f^*$ from experts.

# What is Reification? Answer

- Goal: Assess the uncertainty in our predictions from $f$ for $y$
- Alas, we know that $\mathrm{E}[y(t)|f,t] \neq f(t)$
- Introduce discrepancy terms: $\mathrm{E}[y(t)|f,t,\rho,\delta] = \rho f(t) + \delta$
- Consider a hypothetical, expert informed $f^*$

$$\begin{aligned} y(t)|f^*,t &\sim \mathrm{No}(f^*(t),\phi_y^{-1}) \\ f^*|f,t &\sim \mathrm{No}(m^*(t),p^*(t)^{-1}) \\ f(t)|D,t &\sim \mathrm{No}(m(t),p(t)^{-1}) \end{aligned}$$

- "$f^*$ is a reified $f$"

Reification (Goldstein and Rougier, 2009):
The process of eliciting $f^*$ from experts.

- Def: To reify is to make an abstract concept more concrete.

# Advantages of $f^*$

Advantages:

# Advantages of $f^*$

Advantages:

- $f^*$ breaks down the discripency: $y - f = y - f^* + f^* - f$

# Advantages of $f^*$

Advantages:

- $f^*$ breaks down the discripency: $y - f = y - f^* + f^* - f$
  $\rightarrow$ Discrepancy $y - f$ is hard to conceptualize directly

# Advantages of $f^*$

Advantages:

- $f^*$ breaks down the discripency: $y - f = y - f^* + f^* - f$
  $\rightarrow$ Discrepancy $y - f$ is hard to conceptualize directly
- Often experts have an understanding $f^* - f$

# Advantages of $f^*$

Advantages:

- $f^*$ breaks down the discripency: $y - f = y - f^* + f^* - f$
  $\rightarrow$ Discrepancy $y - f$ is hard to conceptualize directly
- Often experts have an understanding $f^* - f$
- Consider general reification, structural reification, and both

# Advantages of $f^*$

Advantages:

- $f^*$ breaks down the discripency: $y - f = y - f^* + f^* - f$
  $\rightarrow$ Discrepancy $y - f$ is hard to conceptualize directly
- Often experts have an understanding $f^* - f$
- Consider general reification, structural reification, and both
- Thoughtful specifications of $\{m^*(t), p^*(t)\}$ enable

  structural and random inaccuracies within $f$ to be quantified for inference about $y$

# Advantages of $f^*$

Advantages:

- $f^*$ breaks down the discripency: $y - f = y - f^* + f^* - f$
  $\rightarrow$ Discrepancy $y - f$ is hard to conceptualize directly
- Often experts have an understanding $f^* - f$
- Consider general reification, structural reification, and both
- Thoughtful specifications of $\{m^*(t), p^*(t)\}$ enable

  structural and random inaccuracies within $f$ to be quantified for inference about $y$

Comments:

# Advantages of $f^*$

Advantages:

- $f^*$ breaks down the discripency: $y - f = y - f^* + f^* - f$
  $\rightarrow$ Discrepancy $y - f$ is hard to conceptualize directly
- Often experts have an understanding $f^* - f$
- Consider general reification, structural reification, and both
- Thoughtful specifications of $\{m^*(t), p^*(t)\}$ enable

  structural and random inaccuracies within $f$ to be quantified for inference about $y$

Comments:

- Subjectivity in analyses always face push-back

# Advantages of $f^*$

Advantages:

- $f^*$ breaks down the discripency: $y - f = y - f^* + f^* - f$
  $\rightarrow$ Discrepancy $y - f$ is hard to conceptualize directly
- Often experts have an understanding $f^* - f$
- Consider general reification, structural reification, and both
- Thoughtful specifications of $\{m^*(t), p^*(t)\}$ enable

  structural and random inaccuracies within $f$ to be quantified for inference about $y$

Comments:

- Subjectivity in analyses always face push-back
- In my work, I offered a process to justify/validate reification when using compartmental simulators

# Advantages of $f^*$

Advantages:

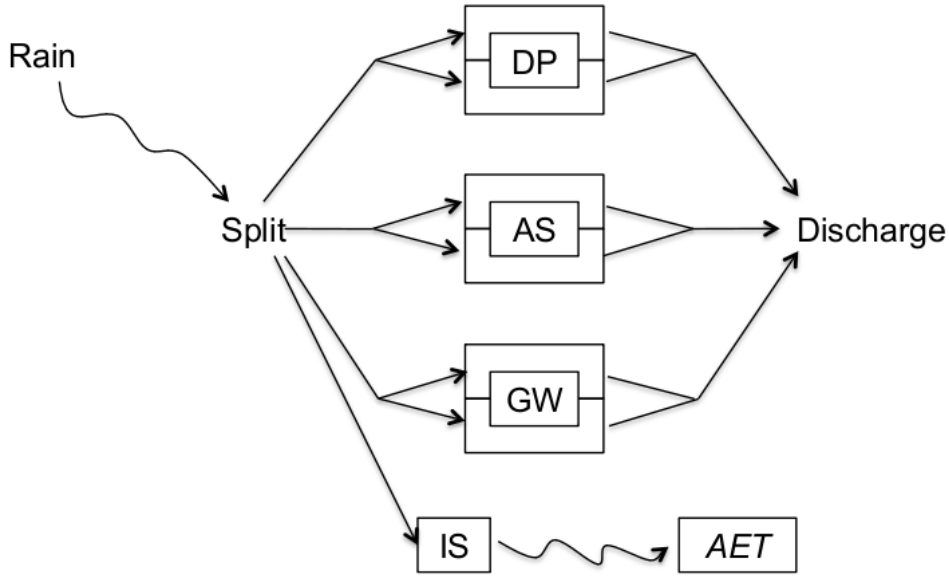
- $f^*$ breaks down the discripency: $y - f = y - f^* + f^* - f$
  $\rightarrow$ Discrepancy $y - f$ is hard to conceptualize directly
- Often experts have an understanding $f^* - f$
- Consider general reification, structural reification, and both
- Thoughtful specifications of $\{m^*(t), p^*(t)\}$ enable

  structural and random inaccuracies within $f$ to be quantified for inference about $y$

Comments:

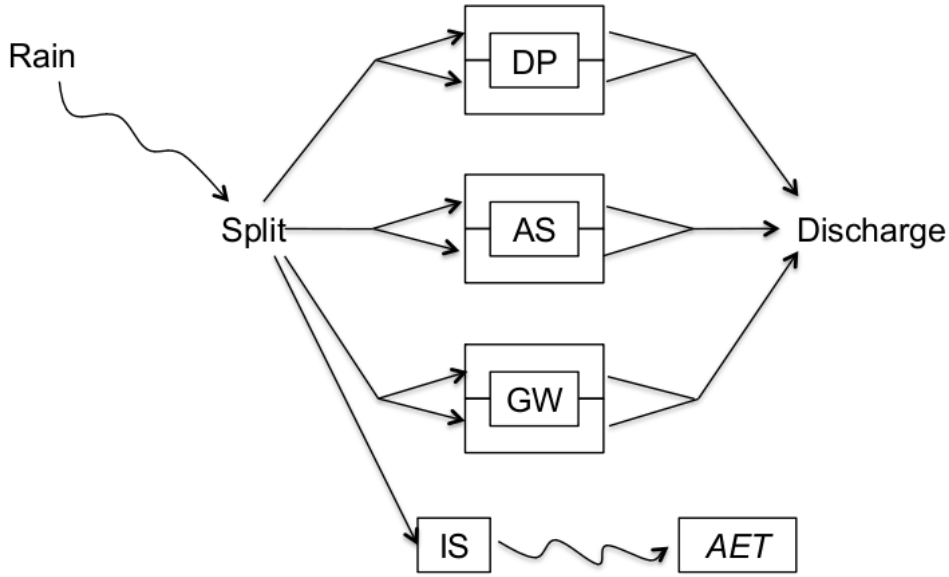
- Subjectivity in analyses always face push-back
- In my work, I offered a process to justify/validate reification when using compartmental simulators
- Today, show reification and learn some foundations to apply it.

# Rainfall-Runoff Model (Iorgulescu et al., 2005)



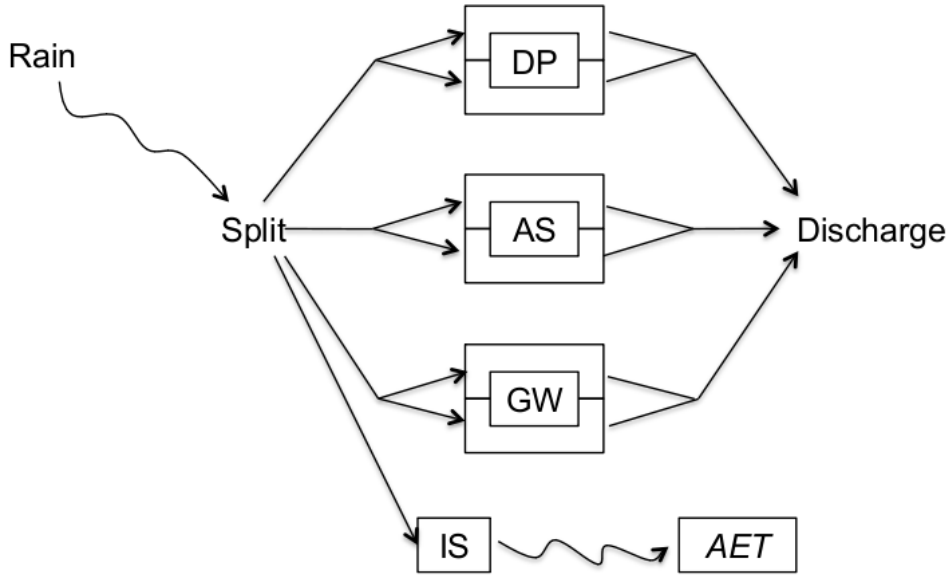

- 3 compartment (GW, AS, DP), conceptual model

# Rainfall-Runoff Model (Iorgulescu et al., 2005)



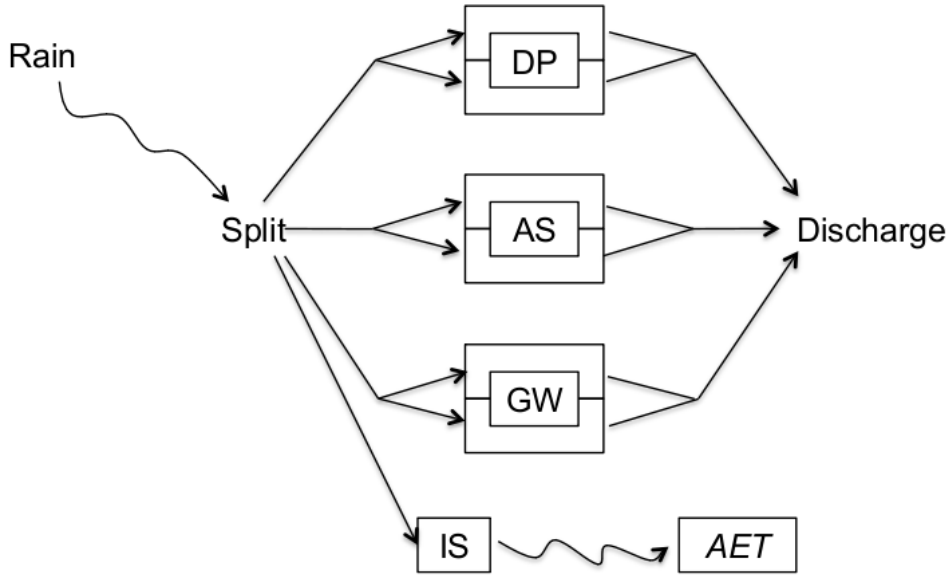

- 3 compartment (GW, AS, DP), conceptual model
- Conditions: Rain over time (e.g., 839 hours) and AET

# Rainfall-Runoff Model (Iorgulescu et al., 2005)



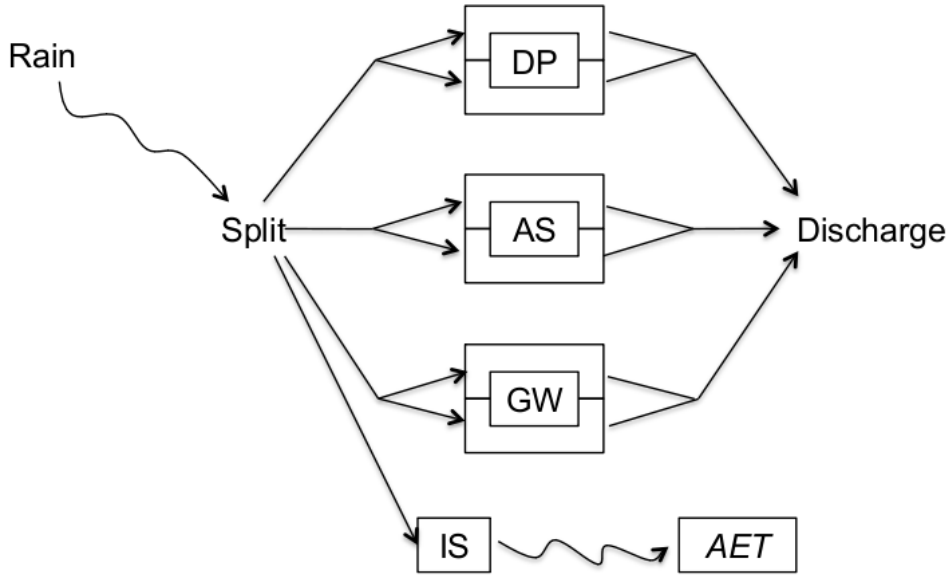

- 3 compartment (GW, AS, DP), conceptual model
- Conditions: Rain over time (e.g., 839 hours) and AET
- 17 inputs: how waters splits, travels, and exits the system

# Rainfall-Runoff Model (Iorgulescu et al., 2005)




- 3 compartment (GW, AS, DP), conceptual model
- Conditions: Rain over time (e.g., 839 hours) and AET
- 17 inputs: how waters splits, travels, and exits the system
- 3 outputs over time: 2 tracer concentrations and discharge

# Rainfall-Runoff Model (Iorgulescu et al., 2005)




- 3 compartment (GW, AS, DP), conceptual model
- Conditions: Rain over time (e.g., 839 hours) and AET
- 17 inputs: how waters splits, travels, and exits the system
- 3 outputs over time: 2 tracer concentrations and discharge
- Fast

# General Approach

- Goal: Reify *f*

# General Approach

- Goal: Reify $f$
- Use what we learn from a low fidelity simulator relative to $f$

# General Approach

- Goal: Reify $f$
- Use what we learn from a low fidelity simulator relative to $f$
- Remove one compartment to create

$f^{/}$: an incomplete simulator

# General Approach

- Goal: Reify $f$
- Use what we learn from a low fidelity simulator relative to $f$
- Remove one compartment to create

$f^{/}$: an incomplete simulator

- Design an experiment so that runs from $f^{/}$ and $f$ align

# General Approach

- Goal: Reify $f$
- Use what we learn from a low fidelity simulator relative to $f$
- Remove one compartment to create

$f^{l}$: an incomplete simulator

- Design an experiment so that runs from $f^{l}$ and $f$ align
- Emulate and *honestly* reify the $f^{l}$

# General Approach

- Goal: Reify $f$
- Use what we learn from a low fidelity simulator relative to $f$
- Remove one compartment to create

$f^I$: an incomplete simulator

- Design an experiment so that runs from $f^I$ and $f$ align
- Emulate and *honestly* reify the $f^I$
- Compare reified predictions from $f^I$ to output $f$

# General Approach

- Goal: Reify $f$
- Use what we learn from a low fidelity simulator relative to $f$
- Remove one compartment to create

$f^l$: an incomplete simulator

- Design an experiment so that runs from $f^l$ and $f$ align
- Emulate and *honestly* reify the $f^l$
- Compare reified predictions from $f^l$ to output $f$
- Consider the use of $f^l - f$ to inform $f^*$, a reified $f$

# Specifics

- Consider univariate responses from the model

# Specifics

- Consider univariate responses from the model
- E.g., focus on water discharge at t=620

# Specifics

- Consider univariate responses from the model
- E.g., focus on water discharge at t=620
- Rely on high-dimensional multivariable linear regression (MLR) to emulate $f^l$ and $f$, e.g.,

$$f^l(t^l)|\beta_m, \beta_e, t^l = g_m(t^l)\beta_m + g_e(t^l)\beta_e + e, \quad e \sim \text{No}(0, {\phi_f}^{-1}),$$

# Specifics

- Consider univariate responses from the model
- E.g., focus on water discharge at t=620
- Rely on high-dimensional multivariable linear regression (MLR) to emulate $f^l$ and $f$, e.g.,

$$f^l(t^l)|\beta_m, \beta_e, t^l = g_m(t^l)\beta_m + g_e(t^l)\beta_e + e, \quad e \sim \text{No}(0, {\phi_f}^{-1}),$$

- Parameterize expert judgements $f^*$ in a form that is compatible with MLR

# Justification

- This was a fast model, yet I still emulate
  → enables including expert judgments probabilistically.

# Justification

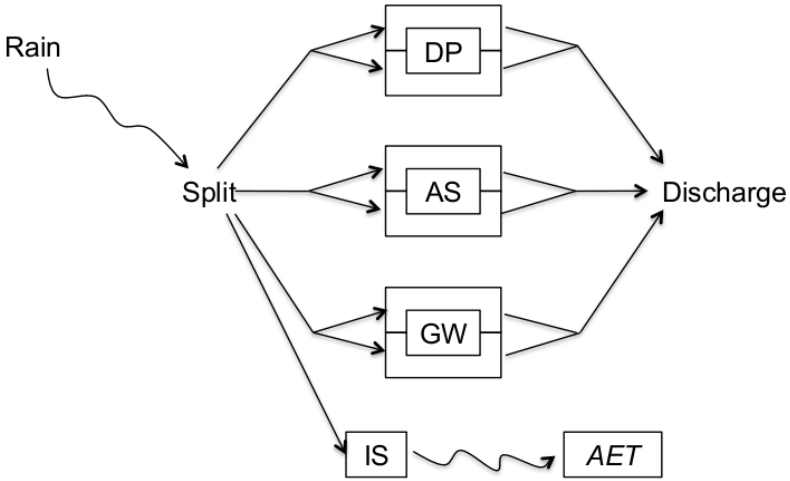
- This was a fast model, yet I still emulate
  → enables including expert judgments probabilistically.
- I used MLR, not Gaussian Proceses
  → being a fast model, I do not need to extrapolate.
  → so many models terms, residuals are negligable.

# Justification

- This was a fast model, yet I still emulate
  → enables including expert judgments probabilistically.
- I used MLR, not Gaussian Proceses
  → being a fast model, I do not need to extrapolate.
  → so many models terms, residuals are negligable.
- Modeled univariate responses, rather than multivariate
  → with so many runs, I didn't need to pool information

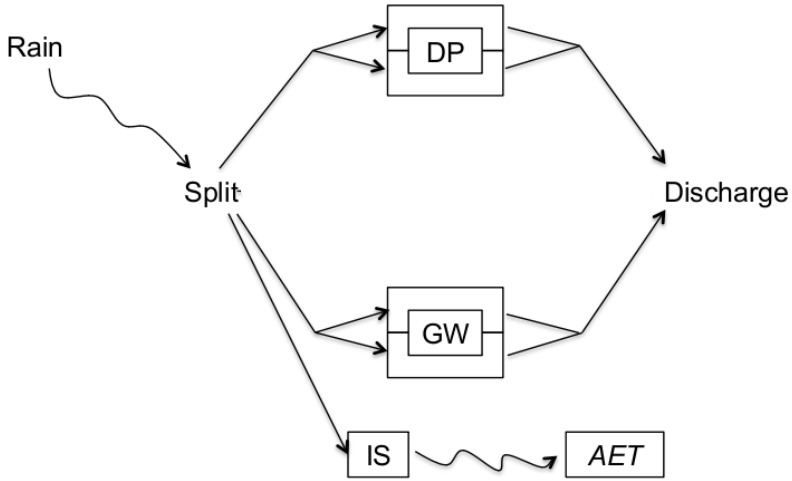
# Full vs Incomplete Simulators

## Complete (3 Compartments)




17 inputs

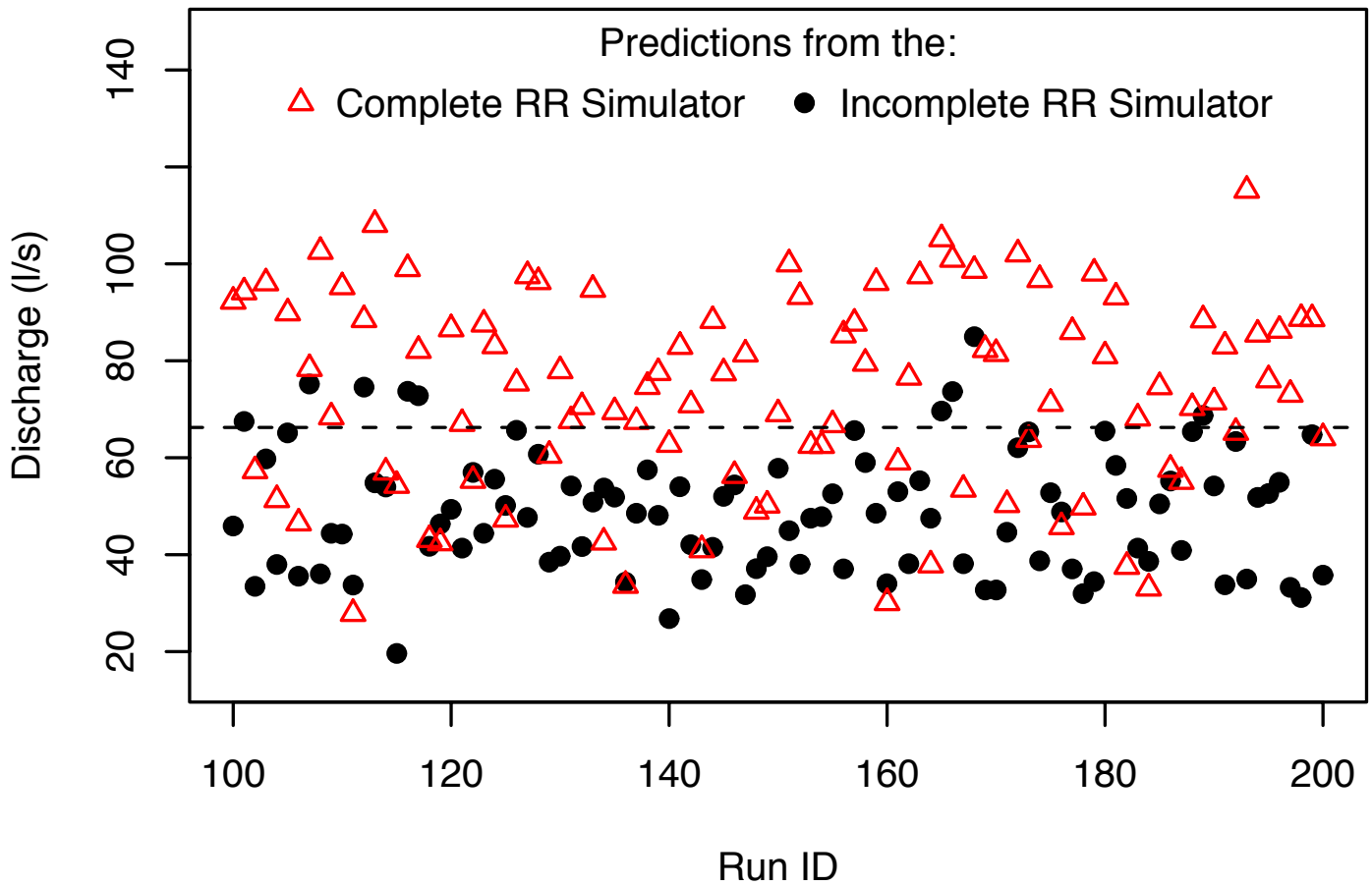
## Incomplete (2 Compartments)




11 inputs

# Visualize Predictions From Reified $f^I$

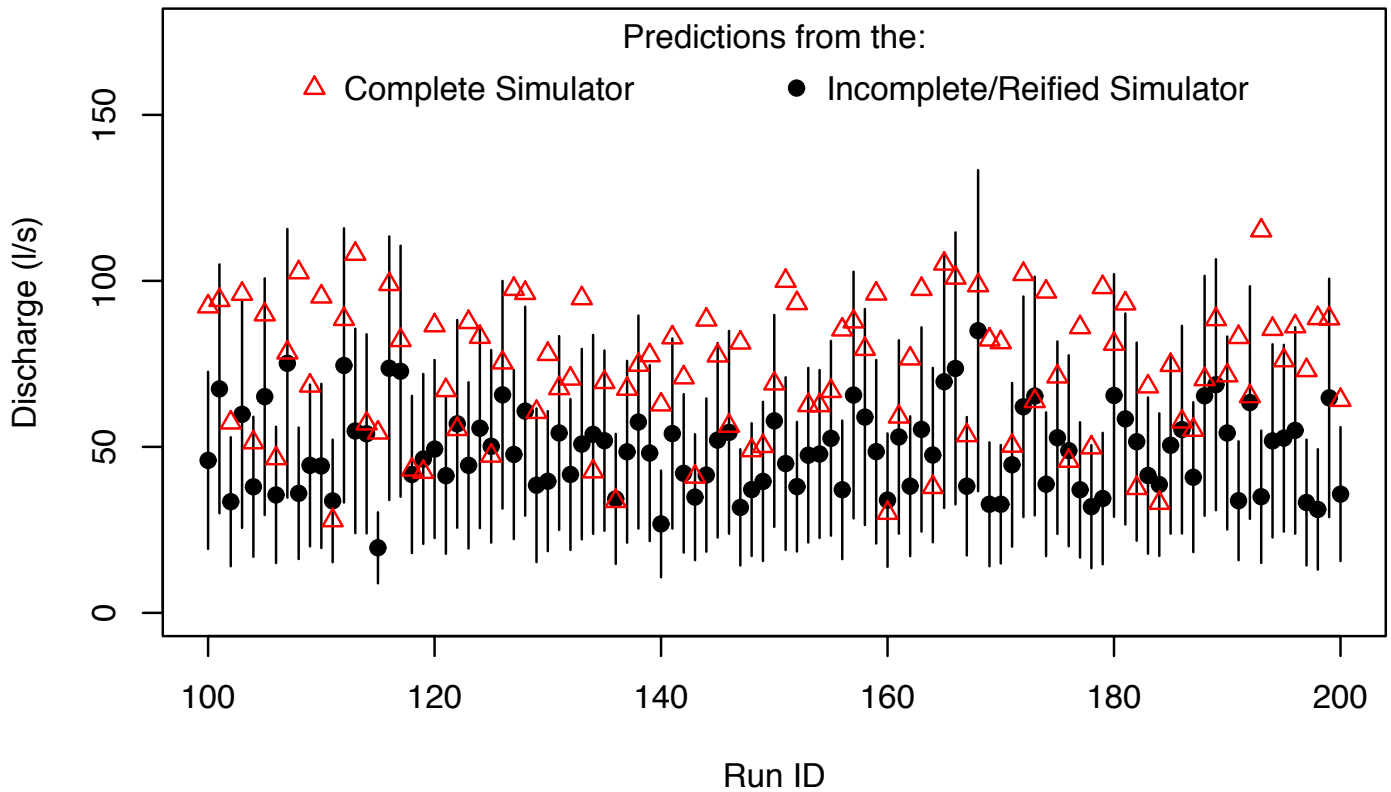
**Goal:** *Reify $f^I$ honestly, and compare it $f$.*



# Visualize Predictions From Reified $f^I$

General Reification

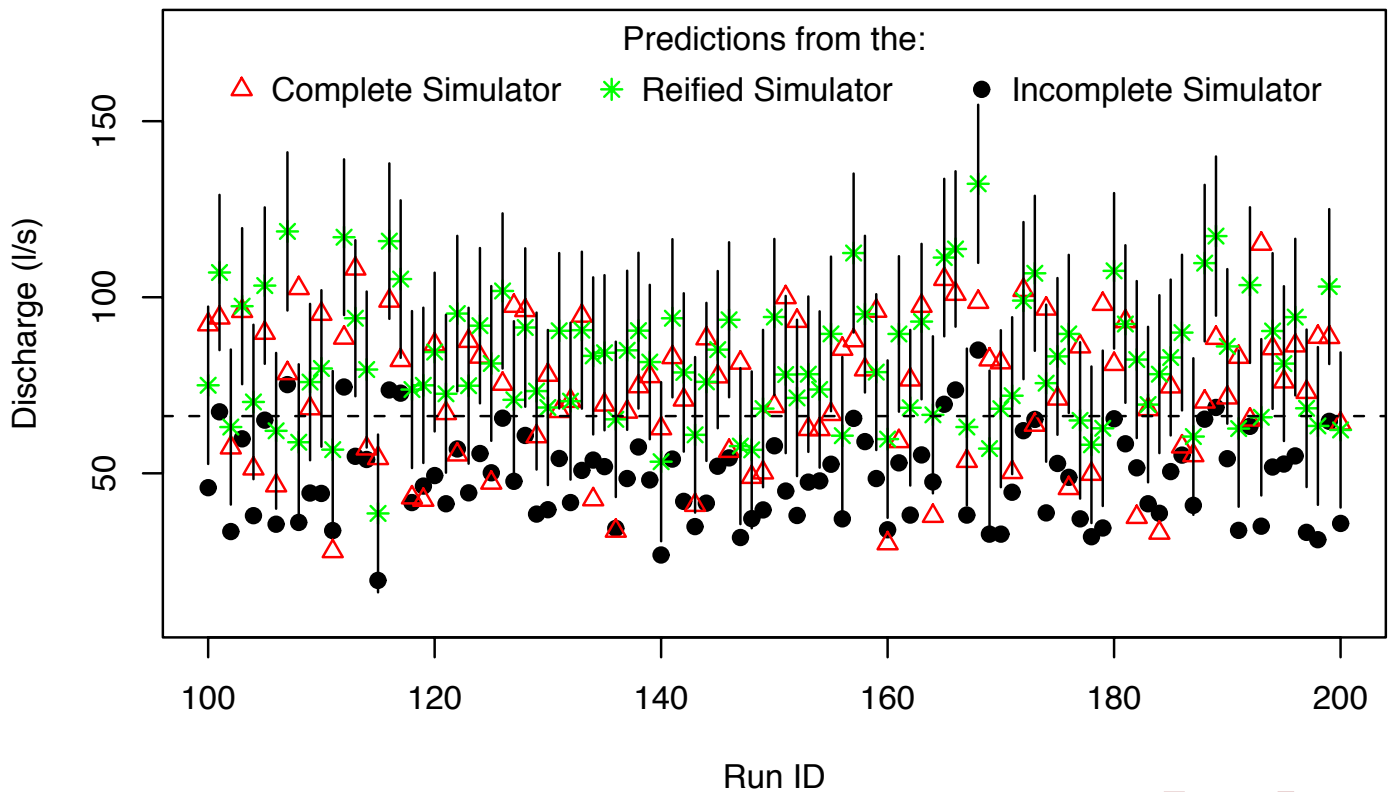
**Expert Judgement:** *Conditional on rain (water input) and parameter specifications, simulators without the AS compartment will likely underestimate discharge by 50%.*



# Visualize Predictions From Reified $f^I$

## Structural Reification

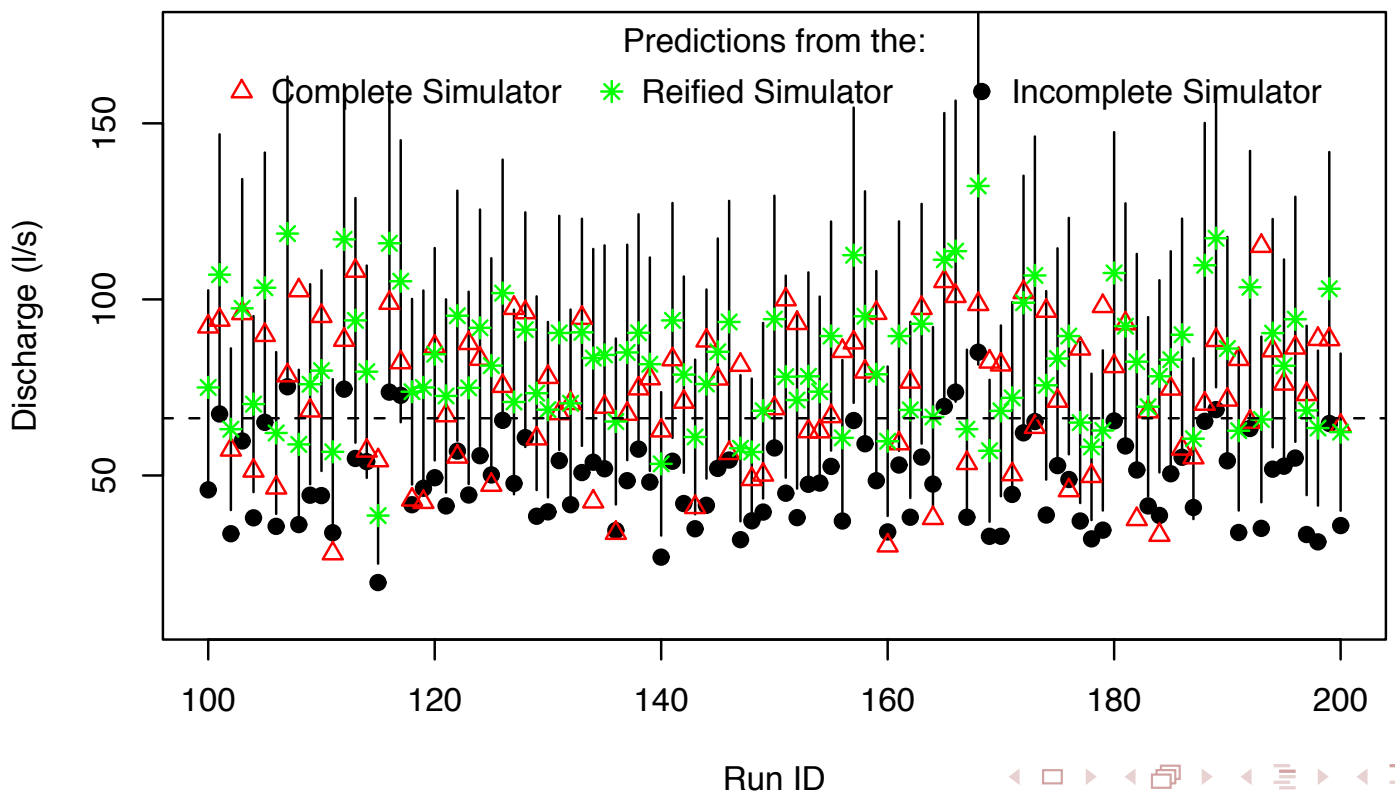
**Expert Judgement:** *Compartments DP and AS should present similar patterns of discharge. Thus, conditional on rain (water input) and parameter specifications, a reified simulator would approximately double the output from DP.*



# Visualize Predictions From Reified $f^I$

General and Structural Reification

**Expert Judgement:** *Compartments DP and AS should present similar, but not exact, patterns of discharge. Thus, conditional on equal rainfalls (water input) and parameter specifications, a reified simulator may approximately double the output from DP*



# Reified $f$: General

**Expert Judgement:** *The relative change in predictions between the incomplete and complete simulators is indicative of a relative change in predictions between the complete simulator and a reified version.*

# Reified $f$: General

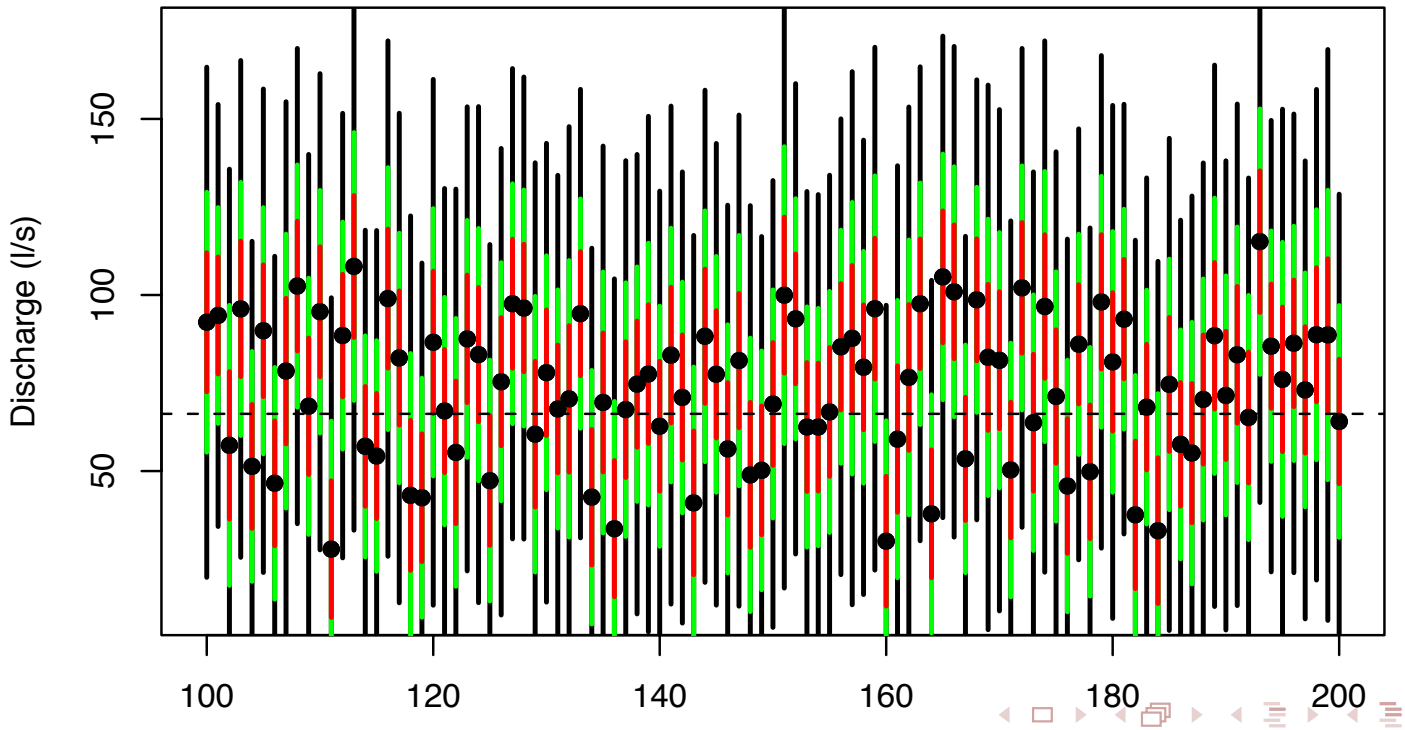
**Expert Judgement:** *The relative change in predictions between the incomplete and complete simulators is indicative of a relative change in predictions between the complete simulator and a reified version.*

- Model $d(t) = f(t) - f^I(t^I)$ with MLR to assess get degree to which additional components may influence outcomes
- Include hyperparameters to assess sensitivity

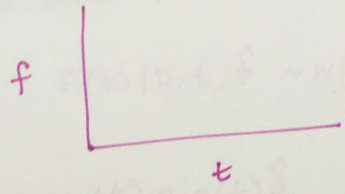
# Reified $f$: General

**Expert Judgement:** *The relative change in predictions between the incomplete and complete simulators is indicative of a relative change in predictions between the complete simulator and a reified version.*

- Model $d(t) = f(t) - f^I(t^I)$ with MLR to assess get degree to which additional components may influence outcomes
- Include hyperparameters to assess sensitivity



Use doc-cam to show general reification.




Notation:

$f$ = model/simulator output

$t$ = model input

$F$ = evaluated/observed

$T$ = evaluated inputs

$\} D$

① Model the simulator

$$f(t) \mid \beta, \phi \sim N(g_m(t)\beta_m + g_e(t)\beta_e, \phi_f^{-1})$$

$$= N(g(t)\beta, \phi_f^{-1})$$

$$\pi(\beta) \propto 1$$

$$\pi(\phi) \propto \phi^{-1}$$

$$E[\beta \mid \phi D] = (g(T)'g(T))^{-1} g(T)'F = \hat{\beta}$$

$$V[\beta \mid \phi D] = (g(T)'g(T))^{-1} \phi^{-1}$$

$$E[\phi \mid \beta D] = \frac{1}{MSE} = \hat{\phi}$$

(Plug-in)

② Derive the Posterior Predictive Distribution

$$f(t) \mid D, t, \hat{\phi} \sim N(m(t), p(t)^{-1})$$

$$m(t) = g(t)\hat{\beta}$$

$$p(t) = \left(g(t)\left(g(T)'g(T)\right)^{-1}g(t) + 1\right)^{-1}$$

③ Reify:

Ex ante var may differ from β

$$f^*(t, v) \mid \beta^*, t, v = g(t)\beta^* + h(t, v)\theta^* + e^*(t, v)$$

$h(t, v)\theta^*$: structural shift

$e^*(t, v)$: new residual

General: $f^*(t) = f(t)$ predictions

$$\mathbb{V}[f(t) \mid D, t] < \mathbb{V}[f^*(t) \mid D, t]$$

so that credible interval of $f^*(t)$ contains $c \cdot f(t)$

Solution: ① inflate $p(t)^{-1}$ or ② integrate out parameters

Choose ②

$$f^*(t) \mid D, \beta^+ \sim N\left(g(t)(\hat{\beta} + \beta^+), p(t)^{-1}\right)$$

$$\beta^+ \sim N(0, \Sigma^+)$$

$$\Rightarrow f^*(t) \mid D \sim N\left(g(t)\hat{\beta},\ g(t)\Sigma^+ g(t)' + p(t)^{-1}\right)$$

Now ... choose $\Sigma^+$

In my paper: $\Sigma^+ = c_1 V(\beta) + c_2$

↑ scaler ↑ matrix: $p \times p \propto (g(\Pi)'g(\Pi))^{-1}$

# Discussion

- Human role is essential for UQ
- Reification formalizes how to include humans
- The process for reifying $f^I$ could be iterative
- The intention for reifying is to enable inference

# Thank You!

Generally reified version

$$
\begin{aligned}
f^{l*}(t^l)|D^l, \beta^{+*} &\sim \mathrm{No}(g(t^l)(\hat{\beta} + \beta^{+*}), p(t^l)^{-1}) \\
f^{l*}(t^l)|D^l &\sim \mathrm{No}(g(t^l)\hat{\beta}, g(t^l)\Sigma^{+*}g(t^l)' + p(t^l)^{-1}).
\end{aligned}
$$

Structurally reified version

$$
\begin{aligned}
f^{l\circledast}(t^l, \nu)|f, t^l, \beta^{+\circledast} &= f(t^l) + g^{+\circledast}(t^l, \nu_m)\beta_m^{+\circledast} + e^{+\circledast}(t^l, \nu) \\
\mathrm{E}[f^{l\circledast}(t^l, \nu)|D, \nu] &= \Big(g_m(t^l) + \nu_1 g_m(t^l)\Big)\hat{\beta}_m + \Big(g_e(t^l) + \\
&\qquad \nu_1\nu_3 g_e(t^l)\Big)\hat{\beta}_e + \nu_1\nu_2(1 - k^{(\mathrm{DP})})\hat{\beta}_{k^{(\mathrm{DP})}} \\
\mathrm{Var}[f^{l\circledast}(t^l, \nu)|D, \nu] &= p(t^l)^{-1} + \nu_5 g_m^{+\circledast}(t^l)g_m^{+\circledast}(t^l)' + \\
&\qquad \nu_1^2\nu_3^2 g_e(t^l)g_e(t^l)' + \nu_1^2\nu_4.
\end{aligned}
$$