

Towards more general sensitivity estimates

Applications considering model structural uncertainties, grouping of parameters, and large-scale analyses

Juliane Mai

SAMO 2022 - Florida State University
March 14-16, 2022



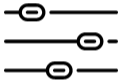
Motivation

– Challenges in Sensitivity Analyses –

Motivation

– Challenges in Sensitivity Analyses –

Model parameters
only

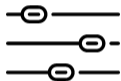


Source Icon

Motivation

– Challenges in Sensitivity Analyses –

Model parameters
only



Source Icon

Challenge #1:

Limits process
understanding



Source Icon

Motivation

– Challenges in Sensitivity Analyses –

Model parameters
only



Source Icon

Dependence on
model structure



Source Icon

Challenge #1:

Limits process
understanding



Source Icon

Motivation

– Challenges in Sensitivity Analyses –

Model parameters only



Source Icon

Dependence on model structure



Source Icon

Challenge #1:

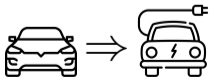
Limits process understanding



Source Icon

Challenge #2:

Transferability to other models



Source Icon

Source Icon

Motivation

– Challenges in Sensitivity Analyses –

Model parameters
only



Source Icon

Dependence on
model structure



Source Icon

Dependence on
location



Source Icon

Challenge #1:

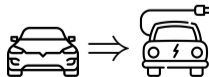
Limits process
understanding



Source Icon

Challenge #2:

Transferability to
other models



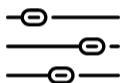
Source Icon

Source Icon

Motivation

– Challenges in Sensitivity Analyses –

Model parameters
only



Source Icon

Dependence on
model structure



Source Icon

Dependence on
location



Source Icon

Challenge #1:

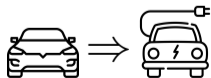
Limits process
understanding



Source Icon

Challenge #2:

Transferability to
other models

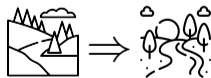


Source Icon

Source Icon

Challenge #3:

Sensitivities at new
location



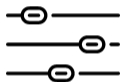
Source Icon

Source Icon

Motivation

– Challenges in Sensitivity Analyses –

Model parameters only



Source Icon

Dependence on model structure



Source Icon

Dependence on location



Source Icon

Data sharing & re-usability



Source Icon

Challenge #1:

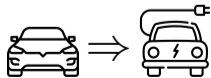
Limits process understanding



Source Icon

Challenge #2:

Transferability to other models

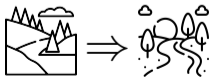


Source Icon

Source Icon

Challenge #3:

Sensitivities at new location



Source Icon

Source Icon

Motivation

– Challenges in Sensitivity Analyses –

Model parameters only



Source Icon

Dependence on model structure



Source Icon

Dependence on location



Source Icon

Data sharing & re-usability



Source Icon

Challenge #1:

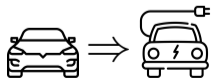
Limits process understanding



Source Icon

Challenge #2:

Transferability to other models

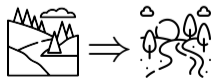


Source Icon

Source Icon

Challenge #3:

Sensitivities at new location



Source Icon

Source Icon

Challenge #4:

Extracting data from previous studies



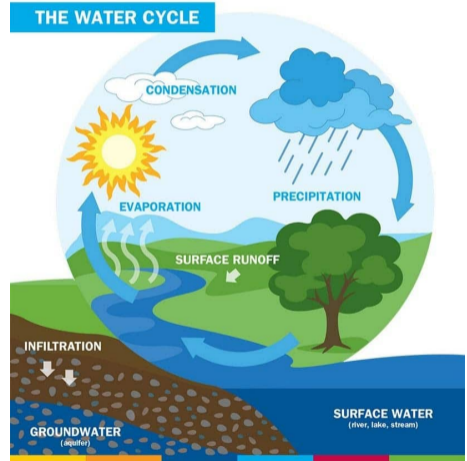
Source Icon

Disclaimer

- Challenges and **methods presented are general** but presented for applications in hydrologic modeling

Disclaimer

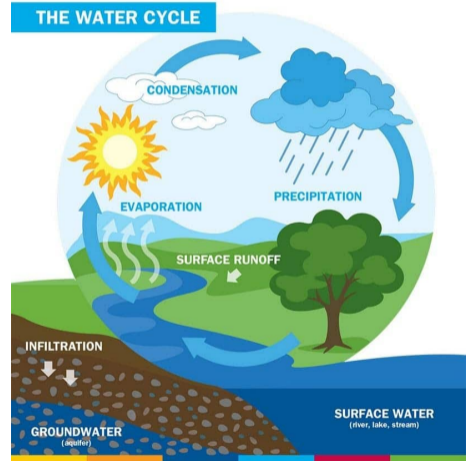
- Challenges and **methods presented are general** but presented for applications in hydrologic modeling
- Hydrologic models describe **water cycle to determine streamflow**, i.e., amount of water passing through a river at a certain location (e.g., Danube in Vienna)



Source Image

Disclaimer

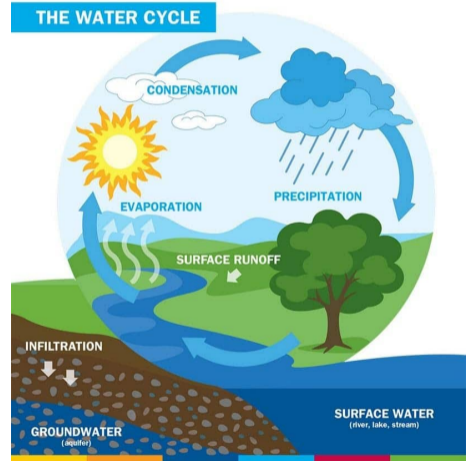
- Challenges and **methods presented are general** but presented for applications in hydrologic modeling
- Hydrologic models describe **water cycle to determine streamflow**, i.e., amount of water passing through a river at a certain location (e.g., Danube in Vienna)
- Output: streamflow



Source Image

Disclaimer

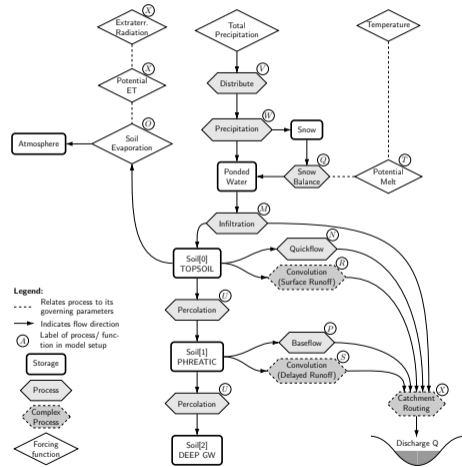
- Challenges and **methods presented are general** but presented for applications in hydrologic modeling
- Hydrologic models describe **water cycle to determine streamflow**, i.e., amount of water passing through a river at a certain location (e.g., Danube in Vienna)
- Output: streamflow
- Inputs:
 - dynamic: radiation, precipitation, temperature, ...
 - static: elevation, slope, latitude, longitude, ...



Source Image

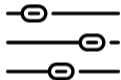
Disclaimer

- Challenges and **methods presented are general** but presented for applications in hydrologic modeling
- Hydrologic models describe **water cycle to determine streamflow**, i.e., amount of water passing through a river at a certain location (e.g., Danube in Vienna)
- Output: streamflow
- Inputs:
 - dynamic: radiation, precipitation, temperature, ...
 - static: elevation, slope, latitude, longitude, ...
 - parameters** describing evaporation, infiltration, etc.



Source Image

Model parameters
only



Source Icon



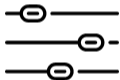
Challenge #1:

Limits process
understanding



Source Icon

Model parameters
only



Source Icon



Challenge #1:

Limits process
understanding



Source Icon



Define **SA** for
processes (groups
of parameters)
rather than
parameters only

Challenge 1: Define SA for processes

– Sobol' method for parameters –

Main effect:

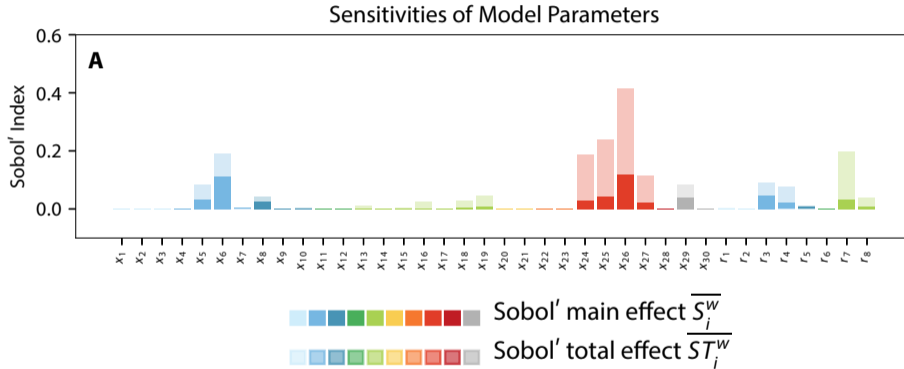
$$S_i = \frac{\text{Variance of model, if **one** parameter is variable}}{\text{Variance of model, if **all** parameters are variable}}$$

Total effect:

$$S_{T_i} = \frac{\text{Variance of model, if **all incl. one** parameters are variable}}{\text{Variance of model, if **all** parameters are variable}}$$

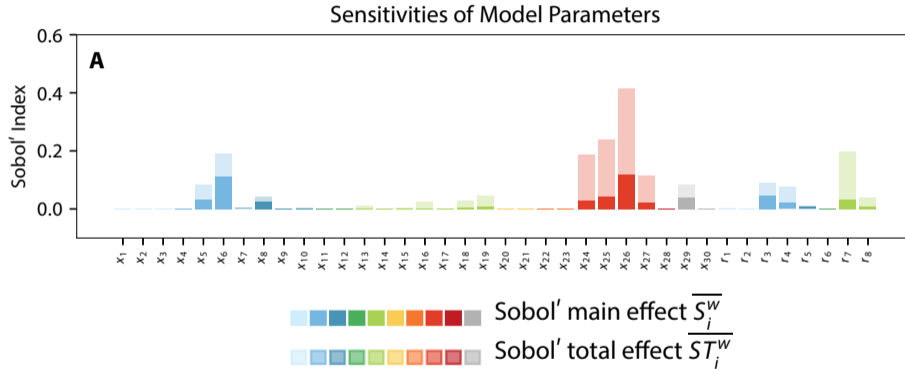
Challenge 1: Define SA for processes

– Sobol' method for parameters –



Challenge 1: Define SA for processes

– Sobol' method for parameters –



Estimates sensitivity of model regarding choice of model parameters.
(40,000 model runs)

Challenge 1: Define SA for processes

– Sobol' method for grouped parameters –

Main effect:

$$S_G = \frac{\text{Variance of model, if **only parameters in group G** are variable}}{\text{Variance of model, if **all** parameters are variable}}$$

Total effect:

$$S_{TG} = \frac{\text{Variance of model, if **all parameters that incl. group G** are variable}}{\text{Variance of model, if **all** parameters are variable}}$$

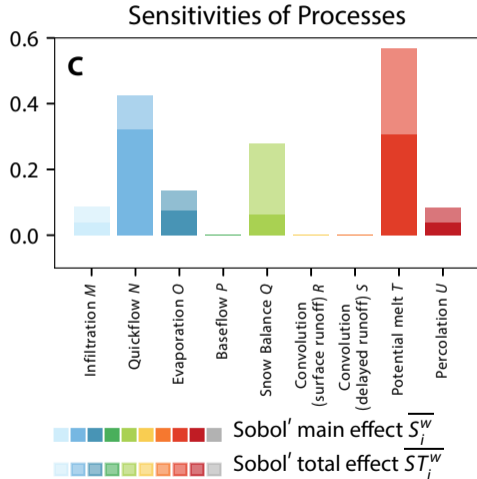
Challenge 1: Define SA for processes

– Sobol' method for grouped parameters –

ID	Process	Parameters active
1	Infiltration	$\{x_1, x_2, x_3, x_{29}\}$
2	Quickflow	$\{x_4, x_5, x_6, x_7, x_{29}\}$
3	Soil evaporation	$\{x_8, x_9, x_{10}, x_{29}\}$
4	Baseflow	$\{x_{11}, x_{12}\}$
5	Snow balance	$\{x_{13}, x_{14}, x_{15}, x_{16}, x_{17}, x_{18}, x_{19}\}$
...

Challenge 1: Define SA for processes

– Sobol' method for grouped parameters –



Estimates sensitivity of model regarding choice of model parameters.

↪ 11,000 model runs for 9 processes instead of 40,000 model runs for 38 parameters

Dependence on
model structure

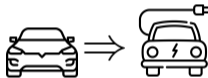


Source Icon



Challenge #2:

Transferability to
other models



Source Icon

Source Icon

Dependence on
model structure

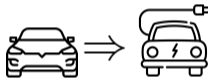


Source Icon



Challenge #2:

Transferability to
other models



Source Icon

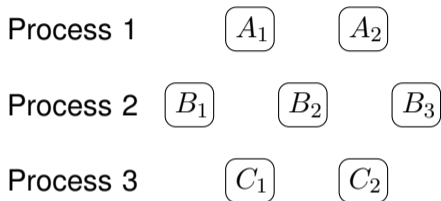
Source Icon



Define model
**blending multiple model
structures** into
one hyper-model

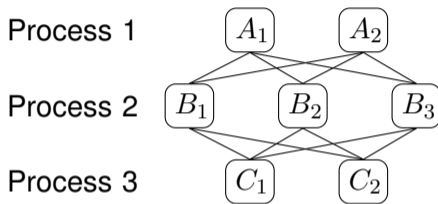
Challenge 2: Define blended models

– Sensitivity Analysis regarding Model Structure –



Challenge 2: Define blended models

– Sensitivity Analysis regarding Model Structure –



Model: $f(\vec{x}) = A_i \cdot B_j + C_k$ with $i \in \{1, 2\}$, $j \in \{1, 2, 3\}$, $k \in \{1, 2\}$

$2 \times 3 \times 2 = 12$ models with varying "active" parameters
 $\Rightarrow 12$ sensitivity analyses

Challenge 2: Define blended models

– Sensitivity Analysis regarding Model Structure –

n	Model			$S_{x_1}^n$	$S_{x_2}^n$	$S_{x_3}^n$	$S_{x_4}^n$	$S_{x_5}^n$	$S_{x_6}^n$	$S_{x_7}^n$
1	A_1	B_1	C_1	0.383	0.000	–	–	–	0.001	–
2	A_1	B_1	C_2	0.171	0.000	0.007	–	–	–	0.548
3	A_1	B_2	C_1	0.656	–	0.000	–	–	0.036	–
4	A_1	B_2	C_2	0.013	–	0.012	–	–	–	0.968
5	A_1	B_3	C_1	0.000	–	–	0.000	0.000	0.132	–
6	A_1	B_3	C_2	0.000	–	0.013	0.000	0.000	–	0.983
7	A_2	B_1	C_1	0.024	0.937	–	–	–	0.000	–
8	A_2	B_1	C_2	0.024	0.926	0.000	–	–	–	0.012
9	A_2	B_2	C_1	0.145	0.382	0.224	–	–	0.001	–
10	A_2	B_2	C_2	0.052	0.138	0.138	–	–	–	0.583
11	A_2	B_3	C_1	0.000	0.000	–	0.237	0.237	0.003	–
12	A_2	B_3	C_2	0.000	0.000	0.010	0.043	0.043	–	0.810

What is the overall most sensitive parameter?

Challenge 2: Define blended models

– Sensitivity Analysis regarding Model Structure –

$$f^n(\vec{x}) = A_i \cdot B_j + C_k \quad \text{with } i \in \{1, 2\}, j \in \{1, 2, 3\}, k \in \{1, 2\}$$

Challenge 2: Define blended models

– Sensitivity Analysis regarding Model Structure –

$$f^n(\vec{x}) = A_i \cdot B_j + C_k \quad \text{with } i \in \{1, 2\}, j \in \{1, 2, 3\}, k \in \{1, 2\}$$

↓

$$g(\vec{x}) = (w_1 A_1 + w_2 A_2) \cdot (w_3 B_1 + w_4 B_2 + w_5 B_3) + (w_6 C_1 + w_7 C_2)$$

with

$$w_1 + w_2 = 1$$

$$w_3 + w_4 + w_5 = 1$$

$$w_6 + w_7 = 1$$

Challenge 2: Define blended models

– Sensitivity Analysis regarding Model Structure –

$$f^n(\vec{x}) = A_i \cdot B_j + C_k \quad \text{with } i \in \{1, 2\}, j \in \{1, 2, 3\}, k \in \{1, 2\}$$

↓

$$g(\vec{x}) = (w_1 A_1 + w_2 A_2) \cdot (w_3 B_1 + w_4 B_2 + w_5 B_3) + (w_6 C_1 + w_7 C_2)$$

with

$$w_1 + w_2 = 1$$

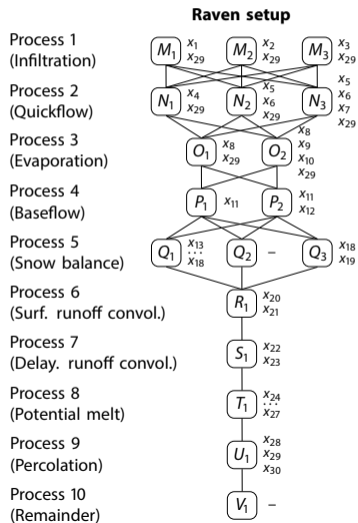
$$w_3 + w_4 + w_5 = 1$$

$$w_6 + w_7 = 1$$

Sampling of weights is another story... See [4].

Challenge 2: Define blended models

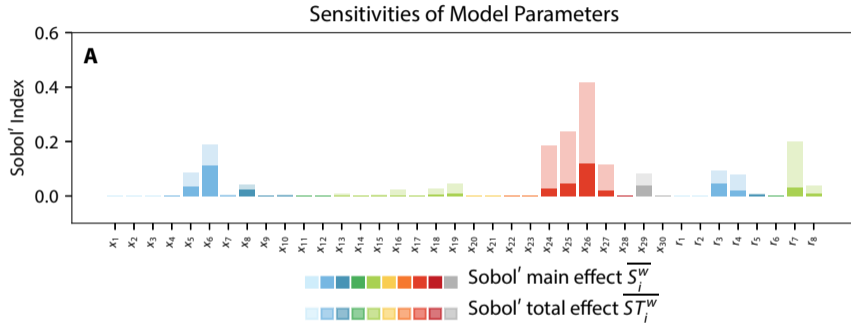
– Sensitivity Analysis regarding Model Structure –



108 models with 30 model parameters and 8 additional parameters to generate weights

Challenge 2: Define blended models

– Sensitivity Analysis regarding Model Structure –

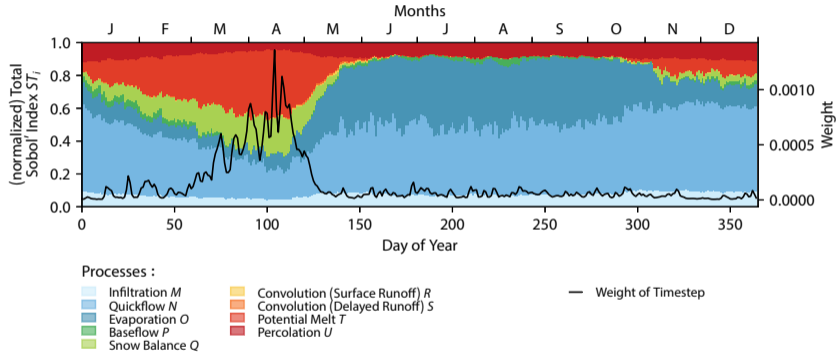


↪ 40,000 model runs for blended model with 38 parameters instead of

2,718,000 model runs to analyze 108 models individually each with 18 to 29 parameters

Challenge 2: Define blended models

– Sensitivity Analysis regarding Model Structure –



↪ 40,000 model runs for blended model with 38 parameters instead of

2,718,000 model runs to analyze 108 models individually each with 18 to 29 parameters

Dependence on
location

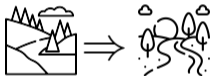


Source Icon



Challenge #3:

Sensitivities at new
location



Source Icon

Source Icon

Dependence on
location

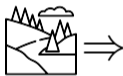


Source Icon



Challenge #3:

Sensitivities at new
location



Source Icon



Source Icon

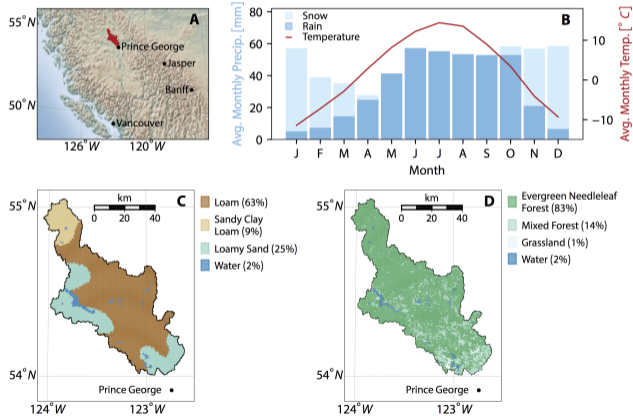


**Large-sample
experiment** to
obtain functional
relationships to
approx. sensitivities

Challenge 3: Large-sample experiments

– Inferring functional relationships for sensitivities –

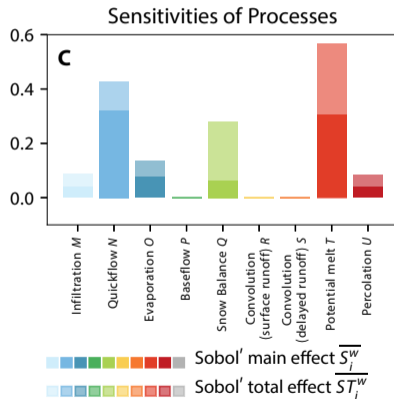
Example model setup for Salmon River catchment (4230km²)



Challenge 3: Large-sample experiments

– Inferring functional relationships for sensitivities –

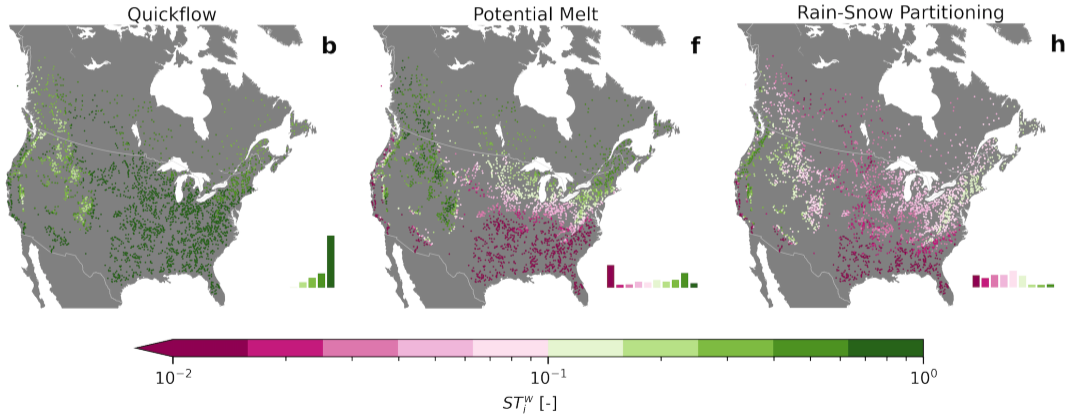
Example model setup for Salmon River catchment (4230km²)



Challenge 3: Large-sample experiments

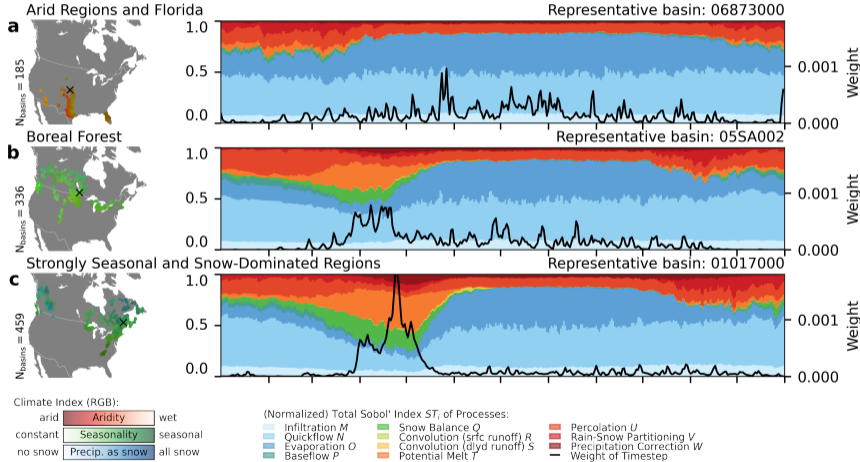
– Inferring functional relationships for sensitivities –

Model setups for 3316 catchments in North America



Challenge 3: Large-sample experiments

– Inferring functional relationships for sensitivities –



Challenge 3: Large-sample experiments

– Inferring functional relationships for sensitivities –

multi-variate regression using location characteristics to approximate sensitivities

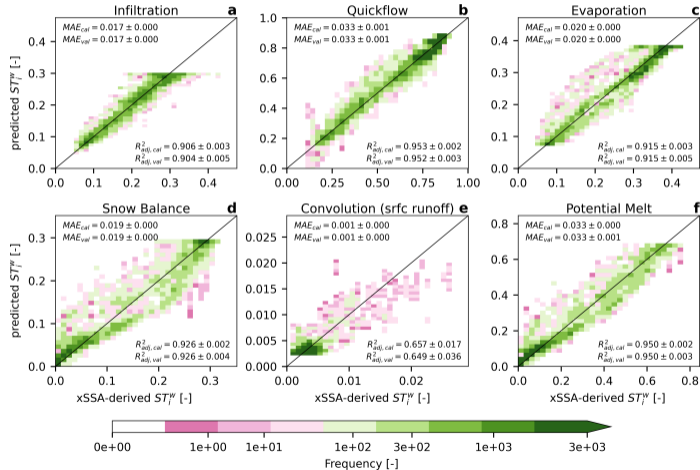
Table 1 Deduced functional relationships using basin attributes to estimate process sensitivities.

Process	Pred. x	Pred. y	Estimated functional relationship $\widehat{ST}_i^w = f(x, y)$	Adj. Coeff. of Determ. $R_{adj}^2(\widehat{ST}_i^w, ST_i^w)$
Infiltration	f_s	-	$0.2985 - 0.8113x + 1.0443x^2 - 0.5203x^3$	0.905
Quickflow	f_s	ΣP	$0.9044 - 3.4122x^2 + 2.6596x^3 - 4.771 \times 10^{-5}y - 0.001153xy + 0.001870x^2y$	0.953
Evaporation	f_s	-	$0.3835 - 0.7713x + 0.1945x^2 + 0.3964x^3$	0.915
Snow Balance	f_s	-	$-0.009051 + 0.4161x + 1.3796x^2 - 2.0661x^3$	0.926
Convolution (srfc runoff)	\bar{T}	f_{cold}	$0.06531 - 0.00639x + 1.645 \times 10^{-4}x^2 - 0.0008165y + 4.104 \times 10^{-5}xy + 2.6508 \times 10^{-6}y^2$	0.657
Potential Melt	f_s	-	$-0.01494 + 0.8236x + 2.4265x^2 - 3.1523x^3$	0.950
Percolation	f_s	-	$0.2829 - 0.7275x + 0.8276x^2 - 0.3582x^3$	0.909
Rain-Snow Partitioning	f_s	f_{cold}	$-0.001715 + 3.7710x - 0.002150y - 0.03638xy + 1.302 \times 10^{-5}y^2 + 8.753 \times 10^{-5}xy^2$	0.815
Precipitation Correction	f_s	$I_{m,r}$	$0.07343 + 0.5829x^2 - 1.0106x^3 + 0.01122y^3 - 0.1677xy^2 + 0.5264x^2y$	0.809

Note. Results of the regression analysis identifying the functional relationships between predictors and the variance-weighted total Sobol' sensitivities ST_i^w based on the results of the 3316 basins analyzed by xSSA. The regression is performed using polynomials with one predictor (up to degree three with at most six coefficients) and polynomials with two predictors (up to degree three with at most five coefficients). The one-predictor polynomial is used unless the use of a pair of predictors led to an improvement of the adjusted coefficients of determination R_{adj}^2 by at least 0.05. The basin characteristics used as predictors are derived from geophysical attributes and the meteorology available for each basin (1950 to 2010). The adjusted coefficients of determination R_{adj}^2 between the xSSA-derived sensitivities ST_i^w and the predicted sensitivities \widehat{ST}_i^w are given in the last column. Description of predictors: f_s is the fraction of precipitation that is snow [mm/mm], ΣP is the annual sum of precipitation [mm], \bar{T} is the annual average temperature in [°C], f_{cold} is the average annual number of days below 0 °C, $I_{m,r}$ is the measure of seasonality ranging between [0,2].

Challenge 3: Large-sample experiments

– Inferring functional relationships for sensitivities –



Data sharing &
re-usability



Source Icon



Challenge #4:

Extracting data from
previous studies



Source Icon

Data sharing &
re-usability



Source Icon



Challenge #4:

Extracting data from
previous studies



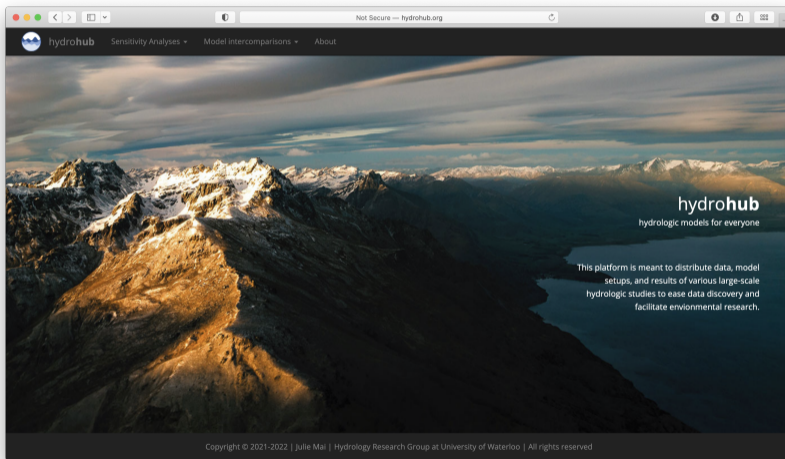
Source Icon



Web-based
interactive
presentation of
inputs, setups, and
results

Challenge 4: Sharing of large datasets

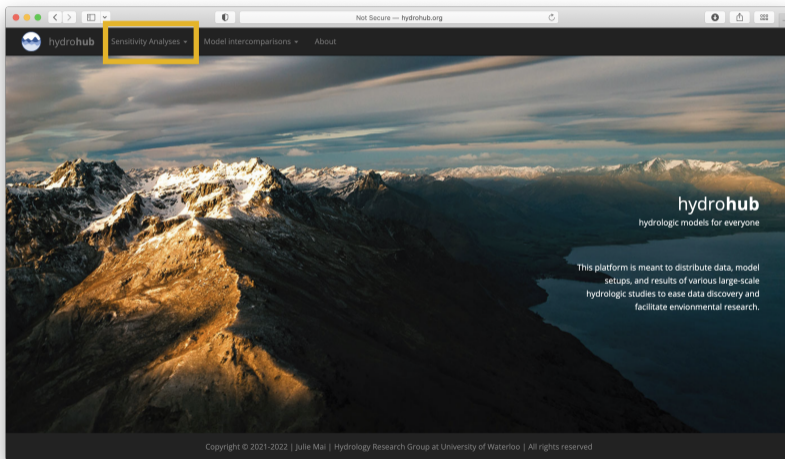
– Interactive websites –



Portal for interactive
websites of
large-scale projects:
hydrohub.org

Challenge 4: Sharing of large datasets

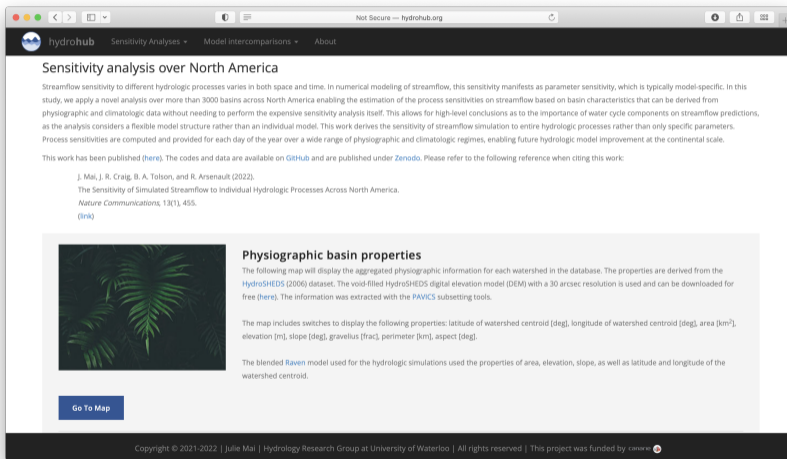
– Interactive websites –



Portal for interactive
websites of
large-scale projects:
hydrohub.org

Challenge 4: Sharing of large datasets

– Interactive websites –

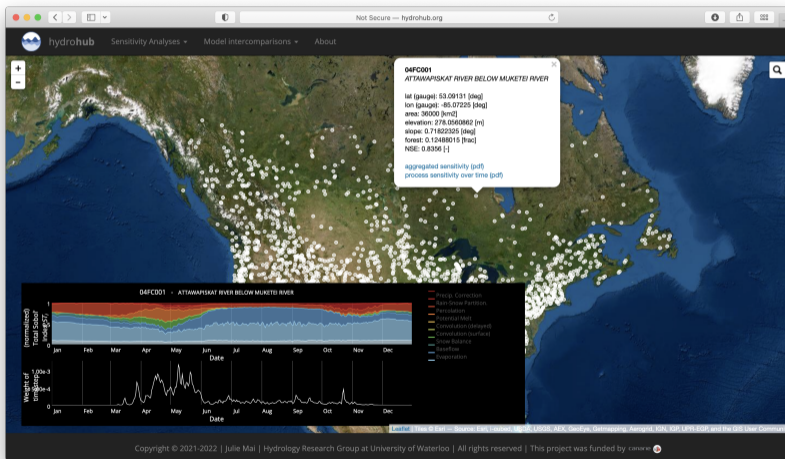


The screenshot shows a web browser window with the URL `hydrohub.org`. The page title is "Sensitivity analysis over North America". The main text describes a study on streamflow sensitivity to hydrologic processes across North America, mentioning the use of 3000 basins and the HydroSHEDS dataset. A citation is provided: "J. Mai, J. R. Craig, B. A. Tolson, and R. Arsenault (2022). The Sensitivity of Simulated Streamflow to Individual Hydrologic Processes Across North America. Nature Communications, 13(1), 455. (link)". Below this is a section titled "Physiographic basin properties" with a small image of green ferns. The text explains that the map displays aggregated physiographic information for each watershed, derived from the HydroSHEDS (2006) dataset. It lists properties like latitude, longitude, area, elevation, slope, gravel, perimeter, and aspect. A "Go To Map" button is visible at the bottom left of the section. The footer contains copyright information: "Copyright © 2021-2022 | Julie Mai | Hydrology Research Group at University of Waterloo | All rights reserved | This project was funded by CONNSRC".

Details about SA over
North America
&
List of available
interactive maps

Challenge 4: Sharing of large datasets

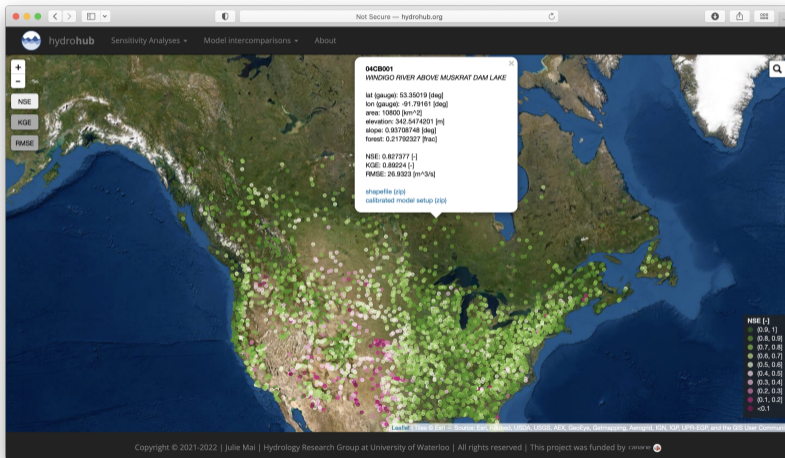
– Interactive websites –



For example,
time-dependent
**sensitivity of
processes** (incl. plots
for download)

Challenge 4: Sharing of large datasets

– Interactive websites –



For example,
calibration results
(incl. **model setups**
for download)

Summary

– Challenges and new approaches in Sensitivity Analyses –

Define **SA for processes** (groups of parameters) rather than parameters only

Define model **blending multiple model structures** into one hyper-model

Large-sample experiment to obtain functional relationships to approx. sensitivities

Web-based **interactive presentation** of inputs, setups, and results

Challenge #1:

Limits process understanding



Source Icon

Challenge #2:

Transferability to other models



Source Icon

Source Icon

Challenge #3:

Sensitivities at new location



Source Icon

Source Icon

Challenge #4:

Extracting data from previous studies



Source Icon



Thanks for your
attention!

Looking forward to
your questions.