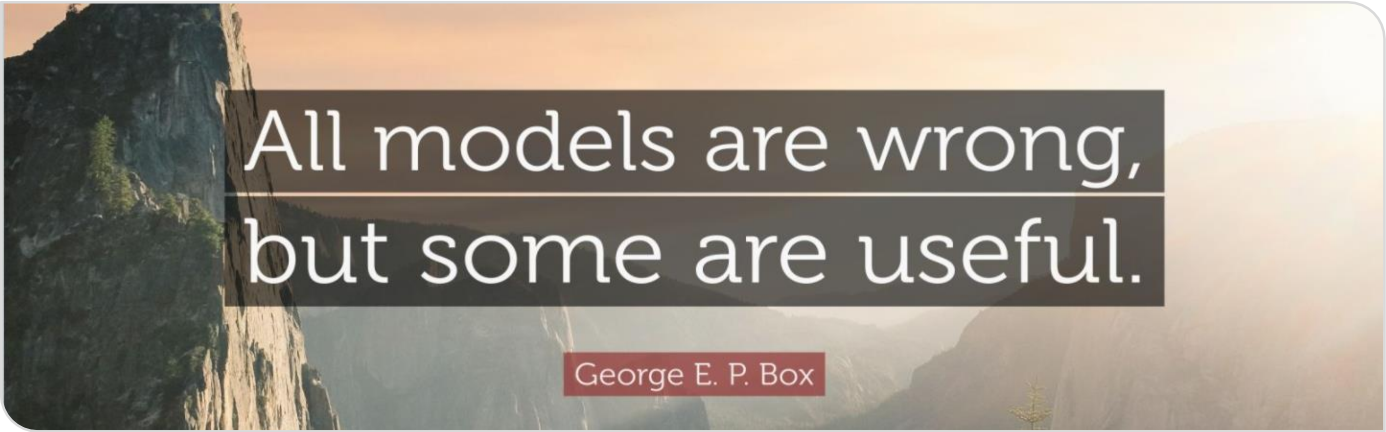
Karlsruher Institut für Technologie

Geodatenanalyse I: Fortgeschrittene Sensitivitätsanalyse

Kathrin Menberg



Stundenplan

	08:30 – 12:30 Uhr	13:30 – 17:30 Uhr
Montag	Tag 1 / Block 1	Tag 1 / Block 2
Dienstag	Tag 2 / Block 1	Tag 2 / Block 2
Mittwoch	Tag 3 / Block 1	**Tag 3 / Block 2**
Donnerstag	Tag 4 / Block 1	Tag 4 / Block 2
Freitag	Tag 5 / Block 1	Tag 5 / Block 2

- 2.7 Monte Carlo Methoden
- 2.8 Grundlagen der Sensitivitätsanalyse
- **2.9 Fortgeschrittene Sensitivitätsanalyse**

Lernziele Block 2.9

Am Ende der Stunde werden die Teilnehmer:

- … mit den mathematischen Grundlagen der Varianzdekomposition vertraut sein.
- … verschiedene quantitative Methoden zur Sensitivitätsanalyse und deren Einsatzgebiete kennen.
- … in Python Sobol Indizes bestimmen und interpretieren können.

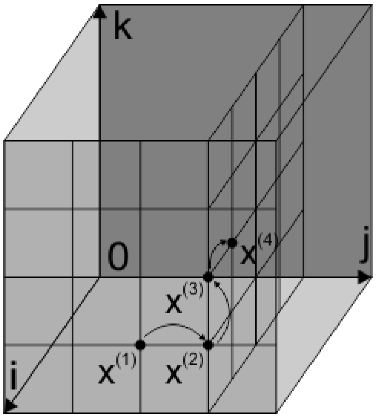
Parameter Screening

- Ziel: Identifikation von nicht-einflussreichen Parameter
- Semi-quantitative Methoden zum Parameter Ranking
- Effiziente Methode um große Anzahl an Parameter zu testen

- Morris Method
- Factorial Sampling Strategie
  - Multidimensionalen Parameterraum durch wenige Stichproben abdecken
  - Anzahl an benötigten Modellauswertungen gering halten

Factorial Sampling

- Parameter-Kombinationen entlang von „Pfaden“ t (engl. trajectory) durch den Parameterraum
  - Unterteilung des Parameterraums in ein regelmäßiges Gitter mit Anzahl an Leveln (oft p)
  - Zwischen jedem Schritt entlang des Pfades ändert sich genau ein Wert



X_1	X_2	X_3
1	1	1
-1	1	1
1	-1	1
-1	-1	1
1	1	-1
-1	1	-1
1	-1	-1
-1	-1	-1

trajectory

Saltelli et al. (2008)

Elementary Effects

- Änderung im Modell-Output zwischen konsekutivem Schritten

$$EE_i = \frac{Y\,(X + \ e_i\Delta_i) - Y(X)}{\Delta_i}$$

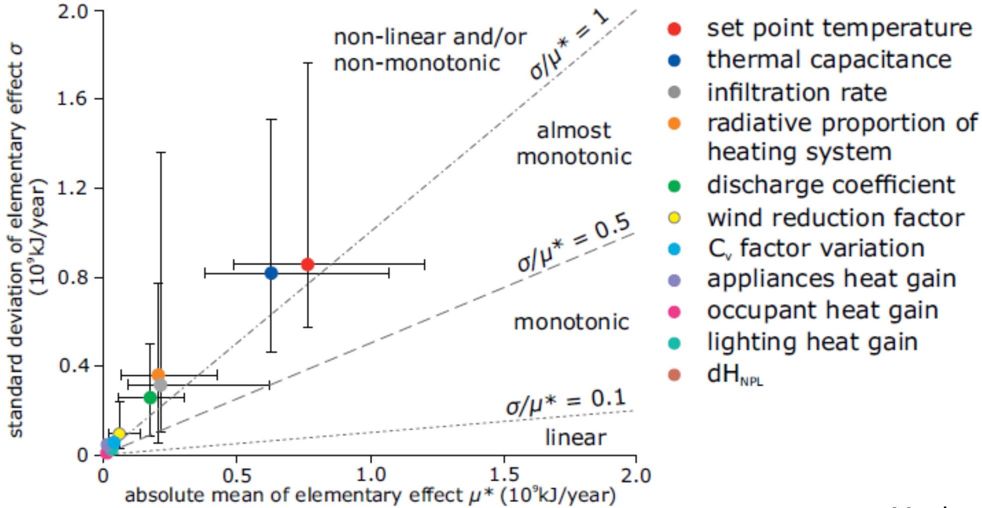
- Statistische Auswertung der Elementary Effects sortiert nach Parametern über alle Trajectories
  - z.B. bei 10 Trajectories, 10 mal Änderungen von Wert von X_i

$$\mu^*{}_i = 0.5 \sum_{t=1}^{k} |EE_{it}| \qquad \sigma = \sqrt{\frac{1}{(\mathrm{k}-1)} \sum_{t=1}^{k} (EE_{it} \ - \ \mu_i)^2}$$

- Absoluten Mittelwert: Einfluss von X_i auf Modell-Output
- Standardabweichung: Maß für die Varianz in Modell-Output aufgrund von X_i

Auswertung Morris Method

- Parameter Ranking nach Mittelwerten der Elementary Effects
- Analyse des Verhältnisses $\sigma/_{\mu}$ der einzelnen Parameter



Menberg et al. (2016)

Varianzdekomposition

- Generisches Modell $Y = f(X_1, X_2, \ldots, X_k)$

$$Y = f_0 + \sum_{i=1}^{d} f_i(X_i) + \sum_{j<i}^{d} f_{ij}(X_i, X_j) + \ \ldots$$

$$V(Y) = \sum_{i} V_i + \sum_{i} \sum_{j<i} V_{ij} + \ \ldots$$

- Ableitung von verschiedenen Indizes zur Sensitivitätsanalyse für nicht-lineare Systeme, bzw. Modelle
- Sobol Indizes (nach Ilya Sobol):
  - Sobol Effekte erster Ordnung, Effekte höherer Ordnung, Totale Effekte

Sobol Effekte erster Ordnung

Bedingte Varianzen (1. Option)

- Generisches Modell $Y = f(X_1, X_2, \dots, X_k)$
- Jedes X variiert innerhalb eines bestimmten Wertebereichs
  - Quantifizieren über Varianz V_{X_i}
- Sensitivität definieren als Effekt den das Fixieren von X_i auf die Varianz in Y hat
  - $E_{\mathbf{X}_{\sim i}}(Y|X_i)$
- Für gesamten Wertebereich von X_i und normiert auf die Gesamtvarianz

$$S_i = \frac{V_{\mathbf{X}_i}(E_{\mathbf{X}_{\sim i}}(Y|X_i))}{V(Y)}$$

Sobol Effekte erster Ordnung

- Sampling Strategie über Monte Carlo Simulation mit Zufallswerten für Modell-Inputs
- Referenzmatrix und eine Vergleichs-Matrix für jedes X_i

$$X^{ref} = \begin{bmatrix} X_{11}^{ref} & \cdots & X_{1i}^{ref} & \cdots & X_{1k}^{ref} \\ X_{21}^{ref} & \cdots & X_{2i}^{ref} & \cdots & X_{2k}^{ref} \\ \vdots & & \cdots & & \vdots \\ X_{N1}^{ref} & \cdots & X_{Ni}^{ref} & \cdots & X_{Nk}^{ref} \end{bmatrix} \qquad X^{(i)} = \begin{bmatrix} \boxed{\begin{matrix} X_{11}^{(i)} \\ X_{21}^{(i)} \\ \vdots \\ X_{N1}^{(i)} \end{matrix}} & \begin{matrix} \cdots \\ \cdots \\ \\ \cdots \end{matrix} & \begin{matrix} X_{1i}^{ref} \\ X_{2i}^{ref} \\ \cdots \\ X_{Ni}^{ref} \end{matrix} & \begin{matrix} \cdots \\ \cdots \\ \\ \cdots \end{matrix} & \boxed{\begin{matrix} X_{1k}^{(i)} \\ X_{2k}^{(i)} \\ \vdots \\ X_{Nk}^{(i)} \end{matrix}} \end{bmatrix}$$

→ die Werte von X_i bleiben gleich, alle anderen werden variiert

- Benötigte Modelldurchläufe: $(1 + \ Anzahl\ Paratemeter) * N$

Sobol Totale Effekte

Bedingte Varianzen (2. Option)

- Generisches Modell $Y = f(X_1, X_2, \ldots, X_n)$
- Jedes X variiert innerhalb eines bestimmten Wertebereichs
  - Quantifizieren über Varianz V_{X_i}
- Sensitivität als Effekt den Variieren von X_i auf die Varianz in Y hat, wenn alle $X_{\sim i}$ fixiert sind
  - $V_{X_i}(Y|\boldsymbol{X}_{\sim i})$
- Für gesamten Wertebereich von X_i und normiert auf die Gesamtvarianz

$$S_{T_i} = \frac{E_{\boldsymbol{X}_{\sim i}}(V_{X_i}(Y|\boldsymbol{X}_{\sim i}))}{V(Y)} \quad = S_i + S_{ij} + \; S_{ij\ldots k}$$

Sobol totale Effekte

- Sampling Strategie über Monte Carlo Simulation mit Zufallswerten für Modell-Inputs
- Referenzmatrix und eine Vergleichs-Matrix für jedes X_i

$$X^{ref} = \begin{bmatrix} X_{11}^{ref} & \cdots & X_{1i}^{ref} & \cdots & X_{1k}^{ref} \\ X_{21}^{ref} & \cdots & X_{2i}^{ref} & \cdots & X_{2k}^{ref} \\ \vdots & & \cdots & & \vdots \\ X_{N1}^{ref} & \cdots & X_{Ni}^{ref} & \cdots & X_{Nk}^{ref} \end{bmatrix} \qquad X^{(i)} = \begin{bmatrix} X_{11}^{ref} & \cdots & \boxed{X_{1i}^{i}} & \cdots & X_{1k}^{ref} \\ X_{21}^{ref} & \cdots & X_{2i}^{i} & \cdots & X_{2k}^{ref} \\ \vdots & & \cdots & & \vdots \\ X_{N1}^{ref} & \cdots & X_{Ni}^{i} & \cdots & X_{Nk}^{ref} \end{bmatrix}$$

→ Nur die Werte von X_i werden variiert, alle anderen bleiben gleich

- Benötigte Modelldurchläufe: $(1 + \ Anzahl\ Paratemeter) * N$

Effekte erster Ordnung vs. Totale Effekte

- Definition der „Wichtigkeit“ von Parametern

- Ziel: Reduzierung von Unsicherheiten
  - Parameter der bei Fixierung Varianz in Y verringert
  - Effekt erster Ordnung

- Ziel: Modellvereinfachung
  - Parameter, der bei Variation möglichst viel der Varianz in Y erhält
  - Totale Effekte

- für additive Modellen ergibt die Summe aller Effekte erster Ordnung 1

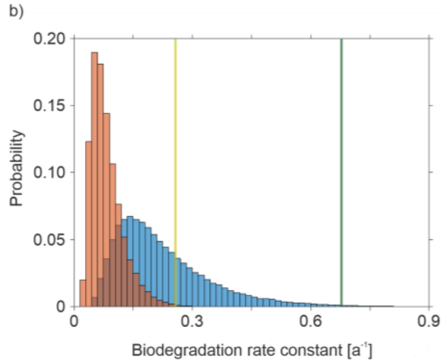
Fazit Sensitivitätsanalyse

- Methoden mit unterschiedlichem Aufwand für verschiedene Zwecke
- Sobol Indizes: <u>**die**</u> Methode für quantitative und globale Sensitivitätsanalyse
- Rechenkosten für Effekte erster Ordnung und totale Effekte: $(1 + \ Anzahl\ Parameter) * N\ * 2$
  - … schnell auf mehrere 10.000 iterative Simulationen
- Parameter Reihenfolge mit Screening Methoden (z.B. Morris Method) verlässlich und effizient zu bestimmen

Übung 2.9: Sensitivitätsanalyse II

- Basierend auf MC Simulation aus Übung 2.7 fortgeschrittene Methoden zur Sensitivitätsanalyse
  - Factorial sampling
  - Morris Method
  - Sobol Indizes
  - Visualisierung

$$\lambda = \frac{\Delta\delta^{13}C \cdot \ k_f \cdot i}{\varepsilon \cdot s \cdot n_e}$$



- Aufgaben in Jupyter Notebook: geodatenanalyse_1-2-9

Aufgabenbesprechung

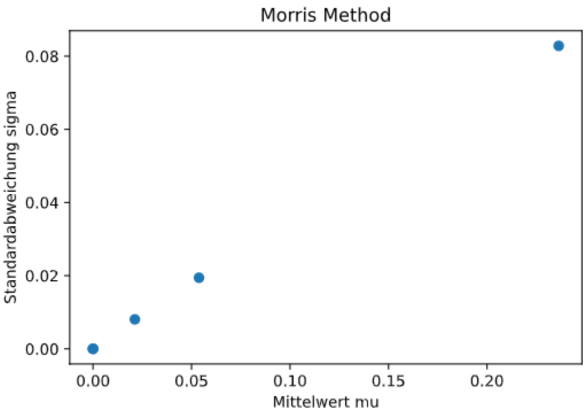
- Factorial Sampling

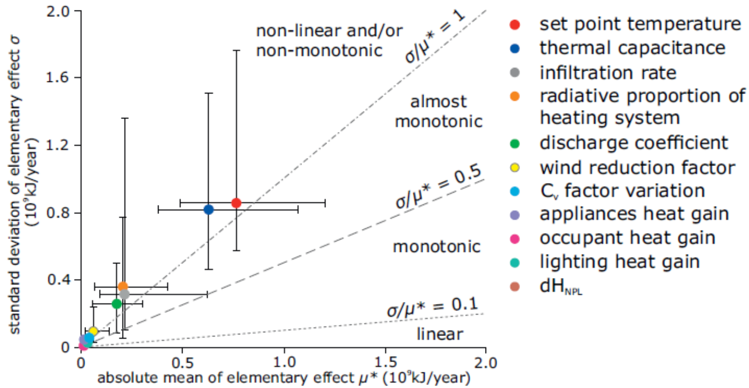
index ▲	0	1	2	3	4	5
0	4.085714…	-2.6	416.1428…	0.171428…	31285714…	0.001385…
1	4.657142…	-2.6	416.1428…	0.171428…	31285714…	0.001385…
2	4.657142…	-2.6	420.7142…	0.171428…	31285714…	0.001385…
3	4.657142…	-2.6	420.7142…	0.171428…	330000000	0.001385…
4	4.657142…	-1.4	420.7142…	0.171428…	330000000	0.001385…
5	4.657142…	-1.4	420.7142…	0.171428…	330000000	0.001557…
6	4.657142…	-1.4	420.7142…	0.274285…	330000000	0.001557…
7	4.514285…	-2.6	415	0.145714…	32571428…	0.001514…
8	4.514285…	-2.6	415	0.145714…	32571428…	0.001342…
9	4.514285…	-2.6	419.5714…	0.145714…	32571428…	0.001342…
10	3.942857…	-2.6	419.5714…	0.145714…	32571428…	0.001342…
11	3.942857…	-2.6	419.5714…	0.145714…	30857142…	0.001342…
12	3.942857…	-2.6	419.5714…	0.248571…	30857142…	0.001342…
13	3.942857…	-1.4	419.5714…	0.248571…	30857142…	0.001342…
14	4.657142…	-3.2	423	0.171428…	31714285…	0.001428…
15	4.657142…	-3.2	418.4285…	0.171428…	31714285…	0.001428…
16	4.657142…	-3.2	418.4285…	0.274285…	31714285…	0.001428…
17	4.657142…	-2	418.4285…	0.274285…	31714285…	0.001428…
18	4.085714…	-2	418.4285…	0.274285…	31714285…	0.001428…
19	4.085714…	-2	418.4285…	0.274285…	31714285…	0.0016
20	4.085714…	-2	418.4285…	0.274285…	300000000	0.0016
21	4.371428…	-2	421.8571…	0.248571…	32571428…	0.001385…
22	4.371428…	-2	421.8571…	0.248571…	30857142…	0.001385…
23	4.371428…	-2	421.8571…	0.248571…	30857142…	0.001557…
24	3.8	-2	421.8571…	0.248571…	30857142…	0.001557…
25	3.8	-2	417.2857…	0.248571…	30857142…	0.001557…

Aufgabenbesprechung

- Morris Method

Parameter	Mu_Star	Mu	Mu_Star_Conf	Sigma
deltaC	0.054	0.054	0.002	0.019
epsilon	0.236	0.236	0.008	0.083
s	0.000	0.000	0.000	0.000
neff	0.000	-0.000	0.000	0.000
t	0.021	-0.021	0.001	0.008
grad	0.000	0.000	0.000	0.000





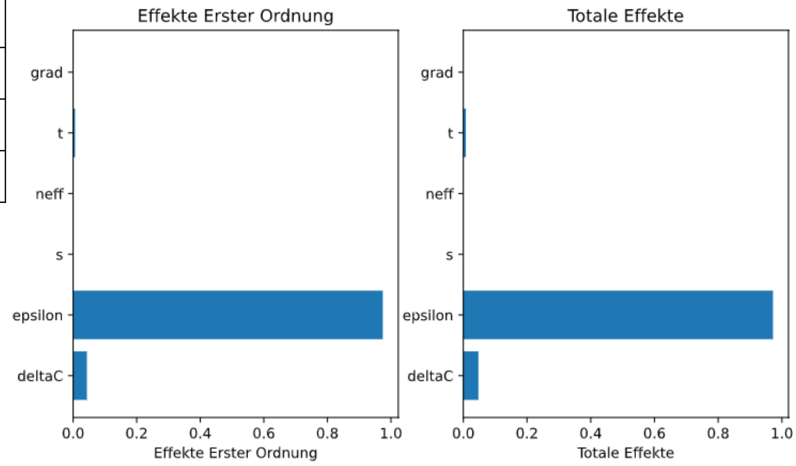
Menberg et al. (2016)

Aufgabenbesprechung

Karlsruher Institut für Technologie

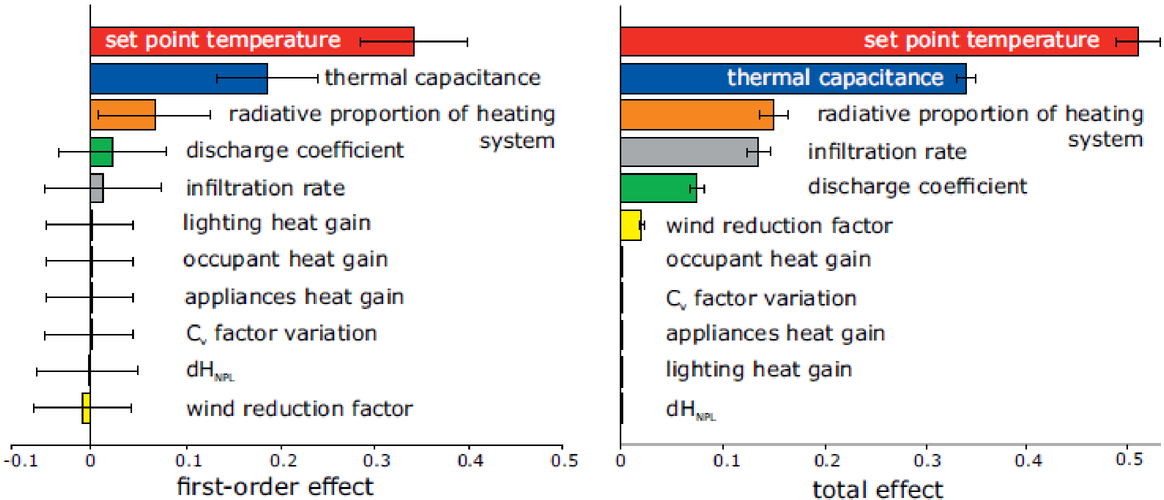
- Sobol Indizes

	1st order effects	total effects
deltaC	0.043	0.047
Epsilon	0.974	0.972
Distance	0.000	0.000
n_eff	0.000	0.000
time	0.006	0.007
Gradient	0.000	0.00



Aufgabenbesprechung

- Beispiel für unterschiedliche Sobol Indizes



Menberg et al. (2016)

Literatur

Karlsruher Institut für Technologie

- Saltelli et al. (2008): Global Sensitivity Analysis. The Primer, John Wiley & Sons.
- Menberg et al. (2016): Sensitivity analysis methods for building energy models: Comparing computational costs and extractable information, Energy and Buildings 133, 433-445.
- Würth et al. (2021): Quantifying biodegradation rate constants of o-xylene by combining compound-specific isotope analysis and groundwater dating. Journal of Contaminant Hydrology, 238, 103757


thank you
danke
謝謝
teşekkür ederim
gracias
спасибо
dank je
tapadh leat
mochchakkeram
go raibh maith agat
dziękuję
bedankt
obrigado
sukriya
kop khun krap
terima kasih
grazie
arigatō
dakujem
takk
merci
мерси
감사합니다
ありがとう
hvala
sagolun
Баярлалаа
blagodaram
хвала