# Sensitivity and Robustness of Financial Models

Giray Ökten
Joint work with David Mandel

Florida State University

# Sobol' global sensitivity analysis: Notation

- Model: $f = f(x_1, ..., x_s)$
- Scale the input variables to $(0, 1)$
- Index set $D = \{1, 2, ..., s\}$
- $u \subseteq D$, $-u$ is the complement of $u$
- $f_u(x^u)$ is a function that only depends on $x^u$

# Anova Decomposition of Functions

$$f(x) = \sum_{u \subseteq D} f_u(x^u)$$

- Component functions $f_u$ are constructed recursively:

$$f_u(x) = \int f(x) dx^{-u} - \sum_{v \subset u} f_v(x^v)$$

- Decomposition is orthogonal, and, as a consequence

$$\sigma^2 = \sum_{u \subseteq D} \sigma_u^2$$

  where

$$\sigma^2 = \int f^2(x) dx - (\int f(x) dx)^2$$

$$\sigma_u^2 = \int f_u^2(x) dx^u$$

# Sensitivity indices

- Sobol' calls $\sigma_u^2/\sigma^2$ the global sensitivity index
- Different ways to measure importance of a set of variables:

$$\underline{S}_u = \frac{1}{\sigma^2}\sum_{v\subseteq u}\sigma_v^2 = \frac{\underline{\tau}_u}{\sigma^2}$$

$$\overline{S}_u = \frac{1}{\sigma^2}\sum_{v\cap u\neq\emptyset}\sigma_v^2 = \frac{\overline{\tau}_u}{\sigma^2}$$

# Freezing Variables

Use the sensitivity indices to identify the "unimportant" parameters of the model, and thus reduce the dimension of the model:

- Compute $\overline{S}_{\{i\}} = \frac{1}{\sigma^2} \sum_{v \cap \{i\} \neq \emptyset} \sigma_v^2$
- Freeze variable $x_i$ at its constant value if $\overline{S}_{\{i\}}$ is small

# Example: Rothermel's Fire Spread Model

- Output:
  rate of spread (ros)
  direction of max spread (sdr)
  effective wind speed (efw)
  reaction intensity (ri)

## Example: Rothermel's Fire Spread Model

- Output:
  rate of spread (ros)
  direction of max spread (sdr)
  effective wind speed (efw)
  reaction intensity (ri)
- Input: There are a lot of them...

| Parameter | Symbol | Value | Units |
|---|---|---|---|
| fuel bed depth | $d$ | 1.83 | m |
| low heat content | $heat$ | 18622.0 | kJ/kg |
| 1-h fuel moisture | $m_{d1}$ | 8.0 | % |
| 10-h fuel moisture | $m_{d2}$ | 8.0 | % |
| 100-h fuel moisture | $m_{d3}$ | 8.0 | % |
| live herbaceous fuel moisture | $m_{lh}$ | 150.0 | % |
| live woody fuel moisture | $m_{lw}$ | 150.0 | % |
| moisture of extinction | $mx$ | 20 | % |
| particle density | $\rho_p$ | 512.5 | $kg/m^3$ |
| effective mineral content | $s_e$ | 1.0 | % |
| slope | $slp$ | 14.04 | $^\circ$ |
| total mineral content | $s_t$ | 5.55 | % |
| 1-h surface area/vol ratio | $sv_{d1}$ | 6562.0 | $m^2/m^3$ |
| 10-h surface area/vol ratio | $sv_{d2}$ | 358.0 | $m^2/m^3$ |
| 100-h surface area/vol ratio | $sv_{d3}$ | 98.0 | $m^2/m^3$ |
| live herb surface area/vol ratio | $sv_{lh}$ | 4921.0 | $m^2/m^3$ |
| live woody surface area/vol ratio | $sv_{lw}$ | 4921.0 | $m^2/m^3$ |
| direction of wind vector (from upslope) | $\theta$ | 45 | $^\circ$ |
| 1-h fuel load | $w0_{d1}$ | 1.12 | $kg/m^2$ |
| 10-h fuel load | $w0_{d2}$ | 0.90 | $kg/m^2$ |
| 100-h fuel load | $w0_{d3}$ | 0.45 | $kg/m^2$ |
| live herbaceous fuel load | $w0_{lh}$ | 0 | $kg/m^2$ |
| live woody fuel load | $w0_{lw}$ | 1.12 | $kg/m^2$ |
| midflame wind speed | $wsp$ | 2.3 | m/s |

# Rothermel's Fire Spread Model

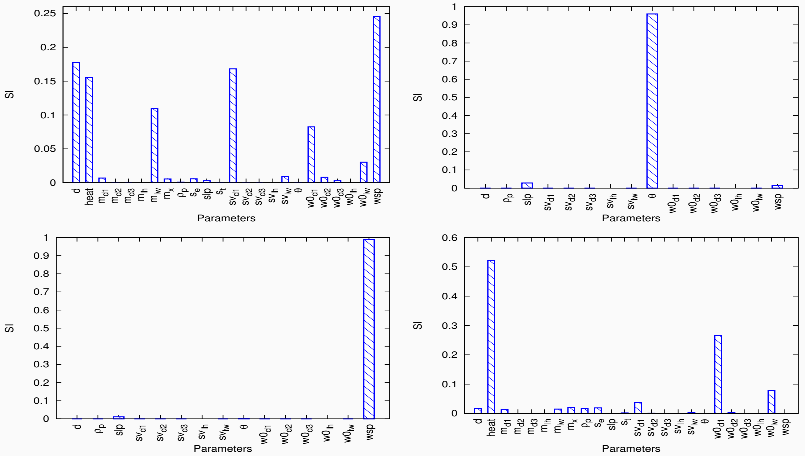



**Figure 1:** Sobol sensitivity analysis, left-right, up-down: *ros*, *sdr*, *efw*, *ri*

# Rothermel's Fire Spread Model

- Most input parameters had a normal distribution - measurement error - and a few had empirical distributions obtained from field data
- Only 7 input variables out of 24 had a significant effect on the model variance
- Used RQMC and variance reduction techniques for uncertainty quantification for the reduced model

# Application: Interest rate models

$$\text{Vasicek:} dr(t) = a(b - r(t))dt + \sigma dW(t)$$
$$\text{CIR:} dr(t) = a(b - r(t))dt + \sigma \sqrt{r(t)} dW(t)$$

Calibration:

- Use one year of interest rate data (yields on one-year T-bills)
- Use MLE to estimate $a, b, \sigma$

# Numerical results

| Year | Parameter | Vasicek | CIR |
|---|---|---|---|
| 1974 | $a$ | 4.05(2.62) | 3.36 (0.96) |
| | $b$ | 7.61(0.46) | 2.21 (0.54) |
| | $\sigma$ | 1.06(4.53) | 1.07 (0.47) |
| 1987 | $a$ | 4.73(2.54) | 2.72 (0.92) |
| | $b$ | 7.02(0.33) | 1.49 (0.60) |
| | $\sigma$ | 1.05(3.65) | 1.08 (0.37) |
| 2006 | $a$ | 6.26(2.14) | 4.32 (1.07) |
| | $b$ | 2.72(0.08) | 0.90 (0.12) |
| | $\sigma$ | 1.02(0.94) | 1.02 (0.21) |

Parameter estimates and standard errors for three calibrations

# Sobol' sensitivity indices

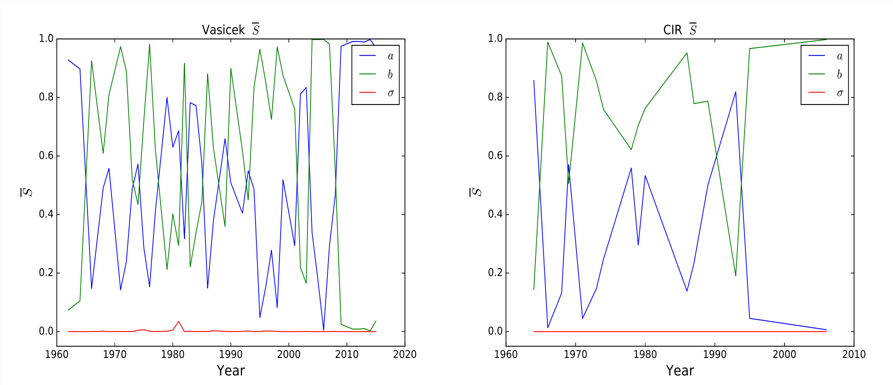
| Year | Sensitivity | Vasicek | CIR |
|---|---|---|---|
| 1974 | $\overline{S}_a$ | 0.57 | 0.25 |
| | $\overline{S}_b$ | 0.43 | 0.76 |
| | $\overline{S}_\sigma$ | 0.004 | 0.0 |
| 1987 | $\overline{S}_a$ | 0.38 | 0.23 |
| | $\overline{S}_b$ | 0.63 | 0.78 |
| | $\overline{S}_\sigma$ | 0.002 | 0.0 |
| 2006 | $\overline{S}_a$ | 0.005 | 0.007 |
| | $\overline{S}_b$ | 0.99 | 0.99 |
| | $\overline{S}_\sigma$ | 0.0002 | 0.0 |

Sensitivity indices for three calibrations

# What's important, what's not: an interesting behavior

- For each year (1962-2015) estimate the model parameters using maximum likelihood method
- The sampling distribution of the parameters is asymptotically normal
- Compute the Sobol' upper sensitivity index for each parameter, using its sampling distribution

# What's important, what's not: an interesting behavior

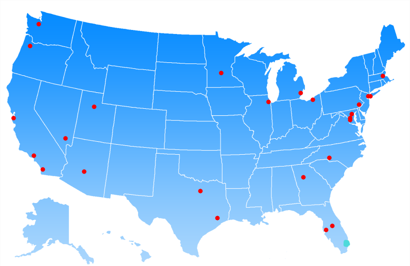



Upper Sobol' sensitivity indices for Vasicek and CIR models across years

**Observe:** What is important seems to change from year to year!

# Application: Temperature Models and Weather Derivatives

- Three models for daily average temperature: [Alaton 2002], [Benth 2007] and [Brody 2002]
- Number of parameters: 17 in Alaton, 14 in Benth, 18 in Brody
- Parameters estimated for twenty-five locations in the US



# Weather Derivatives

- Heating degree-day (HDD) on day $t \in \mathbb{N}$:

$$HDD(t) := \max\{65 - T(t), 0\}$$

- Total HDDs in contract period $[t_0, t_N]$:

$$H(t_N) := \sum_{n=0}^{N} HDD(t_n)$$

- European call option on total HDDs in $[0, T]$:

$$C = e^{-rT} E^P \left(\max\{H(t_N) - K, 0\}\right)$$

where $K$ is the strike

# Temperature Model

Alaton and Benth models assume the average daily temperature follows

$$dT(t) = ds(t) + a(s(t) - T(t))dt + \sigma(t)dW(t),$$

whereas Brody assumes

$$dT(t) = ds(t) + a(s(t) - T(t))dt + \sigma(t)dW^H(t),$$

where

$$s(t) = A + Bt + C\sin(\omega t) + D\cos(\omega t)$$

and $\omega = \frac{2\pi}{365}$.

- Alaton and Brody assume $\sigma(t)$ piecewise constant in each month:

$$\sigma(t) \in \{\sigma_{Jan}, \sigma_{Feb}, \ldots, \sigma_{Dec}\}$$

- Benth assumes a truncated Fourier series:

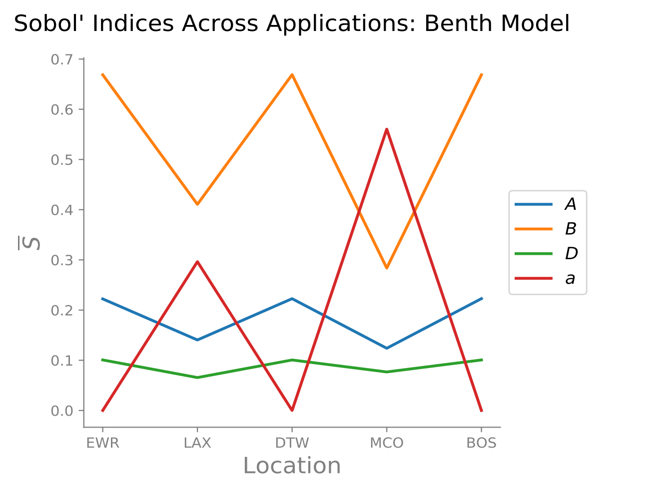
$$\sigma(t) = c_0 + \sum_{i=1}^{4} c_i \sin(i\omega t) + \sum_{j=1}^{4} d_j \cos(j\omega t)$$

- $H \in (0, 1)$ is the Hurst parameter in Brody

# What's important, what's not

- For each location, we use data from that location to estimate the model parameters using maximum likelihood method.
- The sampling distribution of the parameters is asymptotically normal.
- We compute the Sobol' upper sensitivity index for each parameter, using its sampling distribution. Here is how the index changes across locations, for the Benth model.

# What's important, what's not



## Observations and Remedies

Observations:

- Sensitivity indices are not fixed, but random
- Ranking of parameters in terms of sensitivity may not be stable for a given model

# Observations and Remedies

Observations:

- Sensitivity indices are not fixed, but random
- Ranking of parameters in terms of sensitivity may not be stable for a given model

Remedy:

- Randomized Sobol' indices and Anova decomposition
- Robustness of a model in terms of the stability of the sensitivity of its parameters

# Randomized Sobol' indices

Before:

$$\underline{S}_u = \frac{1}{\sigma^2} \sum_{v \subseteq u} \sigma_v^2$$

$$\overline{S}_u = \frac{1}{\sigma^2} \sum_{v \cap u \neq \emptyset} \sigma_v^2$$

After:

$$\underline{S}_u^{Y} = \frac{1}{(\sigma^{Y})^2} \sum_{v \subseteq u} (\sigma_v^{Y})^2$$

$$\overline{S}_u^{Y} = \frac{1}{(\sigma^{Y})^2} \sum_{v \cap u \neq \emptyset} (\sigma_v^{Y})^2$$

# Randomized Anova decomposition

Before:

$$f(x) = \sum_{u \subseteq \mathcal{D}} f_u(x_u)$$

$$\int_{[0,1]} f_u(x_i, x_{u \setminus \{i\}}) dx_i = 0$$

After:

$$f^{Y}(X) = \sum_{u \subseteq \mathcal{D}} f_u^{Y}(X_u)$$

$$\int_{R} f_u^{Y}(x_i, X_{u \setminus \{i\}}) \, \Lambda(dx_i \mid Y) = 0$$

# Robustness of a model

*"Robustness is the insensitivity to small deviations from the assumptions"*

# Robustness of a model

*"Robustness is the insensitivity to small deviations from the assumptions"*

- A model with lower average variance across applications is more robust than a model with higher variance

# Robustness of a model

*"Robustness is the insensitivity to small deviations from the assumptions"*

- A model with lower average variance across applications is more robust than a model with higher variance
- A model with a sensitivity pattern (*ordering of Sobol' sensitivity indices*) that stays fixed with a higher probability across applications is more robust than one with a lower probability

**Notation**

*If* $d = 3$ *then*

$$D! = \{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}$$

*If* $\ell = (2,1,3)$ *then*

$$P(\overline{S}_\ell) = P(\overline{S}_2 < \overline{S}_1 < \overline{S}_3)$$

**Definition**

Let $f : \mathbb{R}^d \to \mathbb{R}$ and let $D!$ denote the set of all permutations of the index set $D = \{1, \dots, d\}$.

For each $\ell \in D!$, let $\overline{S}_\ell$ denote the event $(\overline{S}_{\ell_1} < \overline{S}_{\ell_2} < \dots, \overline{S}_{\ell_d})$. The robustness of the mathematical model $f$ is

$$R = \frac{\max_{\ell \in D!} P(\overline{S}_\ell^Y)}{\mathrm{CoV}}$$

where $\mathrm{CoV} = \frac{\sqrt{E(Var(f(X)|Y))}}{E(f(X))}$.

# Robustness

**Definition**

Let $f : \mathbb{R}^d \to \mathbb{R}$ and let $\tilde{D}!$ denote the set of all permutations of the potentially important index set $\tilde{D} = \{i_1, \ldots, i_p\}$, $p \leq d$.

For each $\ell \in D!$, let $\overline{S}_\ell$ denote the event $(\overline{S}_{\ell_1} < \overline{S}_{\ell_2} < \ldots, \overline{S}_{\ell_d})$. The robustness of the mathematical model $f$ is

$$R = \frac{\max_{\ell \in \tilde{D}!} P(\overline{S}_\ell^Y)}{\mathrm{CoV}}$$

where $\mathrm{CoV} = \frac{\sqrt{E(Var(f(X)|Y))}}{E(f(X))}$.

# Robustness of interest rate models

| Event | Prob (Vasicek) | Prob (CIR) |
|---|---|---|
| $\overline{S}_a < \overline{S}_b < \overline{S}_\sigma$ | 0.0 | 0.0 |
| $\overline{S}_a < \overline{S}_\sigma < \overline{S}_b$ | 0.001 | 0.0 |
| $\overline{S}_b < \overline{S}_a < \overline{S}_\sigma$ | 0.0 | 0.0 |
| $\overline{S}_b < \overline{S}_\sigma < \overline{S}_a$ | 0.003 | 0.0 |
| $\overline{S}_\sigma < \overline{S}_a < \overline{S}_b$ | 0.53 | 0.69 |
| $\overline{S}_\sigma < \overline{S}_b < \overline{S}_a$ | 0.47 | 0.31 |

Parameter estimates and standard errors for three calibrations

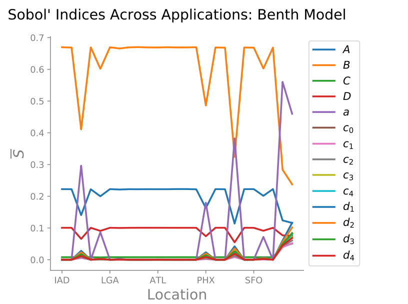
# Robustness of interest rate models

Results:

$$R_{Vasicek} = \frac{0.53}{\sqrt{0.00226}/0.97} = 10.86$$
$$R_{CIR} = \frac{0.69}{\sqrt{0.00183}/0.97} = 15.7$$

Therefore the CIR model is approximately $15.7/10.86 = 1.45$ times more robust than the Vasicek model.

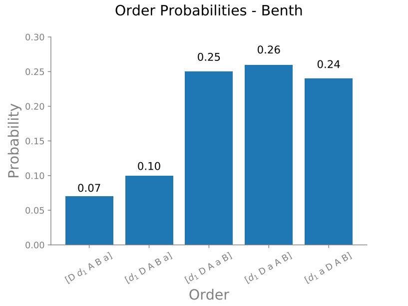
# Varying Sobol' Indices & Potentially Important Parameters




| | $P(\overline{S}_i > 0.1)$ | | |
|---|---|---|---|
| Parameter | Alaton | Benth | Brody |
| $A$ | 0.88 | 0.92 | 0.80 |
| $B$ | 0.98 | 0.99 | 0.94 |
| $C$ | $< 0.05$ | $< 0.05$ | $< 0.05$ |
| $D$ | 0.06 | 0.12 | 0.15 |
| $a$ | 0.71 | 0.73 | 0.77 |
| $H$ | - | - | 0.16 |
| $\sigma_1$ | 0.48 | - | 0.56 |
| $\sigma_2, \ldots, \sigma_{12}$ | $< 0.05$ | - | $< 0.05$ |
| $c_0, \ldots, c_4$ | - | $< 0.05$ | - |
| $d_1$ | - | 0.07 | - |
| $d_2, d_3, d_4$ | - | $< 0.05$ | - |

Estimated $P(\overline{S}_i > 0.1)$

# Robustness of the temperature models




$$R_{Alaton} = \frac{0.18}{\sqrt{73.67}/64.57} = 1.32$$

$$R_{Benth} = \frac{0.26}{\sqrt{73.86}/64.72} = 1.92$$

$$R_{Brody} = \frac{0.22}{\sqrt{74.69}/66.44} = 1.67$$

$$\boxed{R_{Alaton} < R_{Brody} < R_{Benth}}$$

# Thanks for listening!

Questions?

Contact: okten@math.fsu.edu