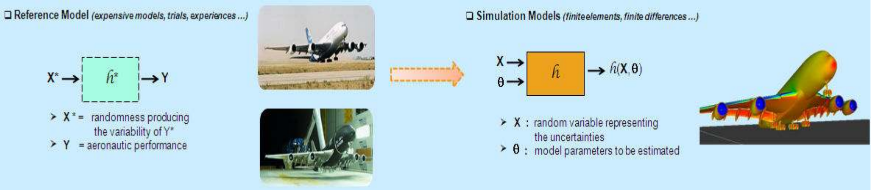
Statistical Learning for Computer Experiments - Application to Aeronautics

— Ateliers GdR MASCOT-NUM, 17 of May 2011 —

Nabil Rachdi, PhD student at Institut de Mathématiques de Toulouse
Advised by: JC Fort (Paris V), Thierry Klein (Toulouse III)
and Fabien Mangeant (EADS IW)

Context

- From Real life to Simulated Life...



- Y = Variable of Interest (uncertain !)
- ρ^* = Quantity of Interest (*quantile, pdf, exceed. probability ...*)
- **Challenge :**
  From ref. data $Y_1, ..., Y_n$ or $(\mathbf{X}_1^*, Y_1), ..., (\mathbf{X}_n^*, Y_n)$ (n limited !)

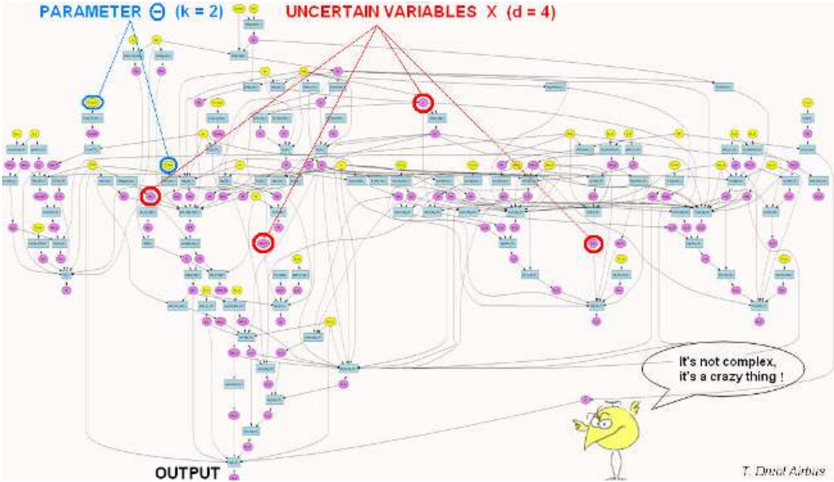
  $\longrightarrow$ Choose h and $\boldsymbol{\theta}$ to predict ρ^* with simulation model(s) h

Examples

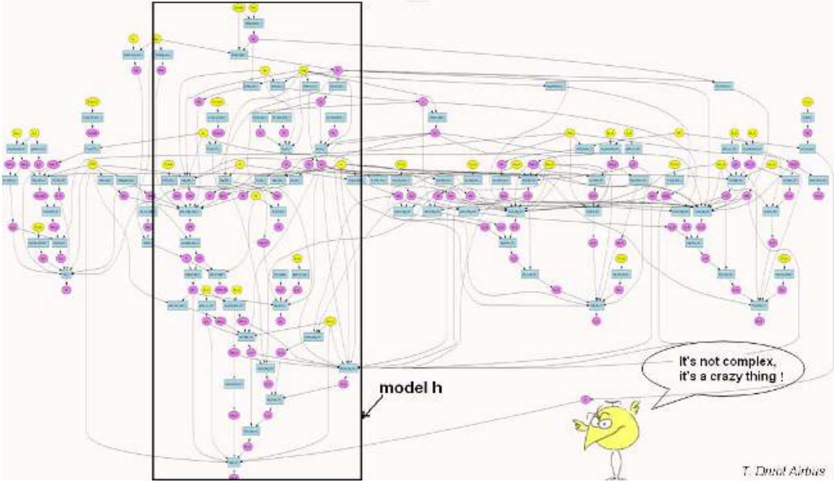
- $Y = TOFL$ (= *Take-Off Field Length*)
  - Quantity of Interest : $\mathbb{P}(TOFL > tofl_{req})$
  - Ref. data $TOFL_1, ..., TOFL_n$ providing from tests, former aircrafts etc...

    $\Rightarrow$ n too small for evaluating $\mathbb{P}(TOFL > tofl_{req})$
  - $h =$ aeronautic model with parameters & uncertainties
- $Y = Range$ (= *distance an aircraft can travel*)
  - Quantity of Interest : $\mathbb{P}(Range < range_{req})$
  - Ref. data $Range_1, ..., Range_n$ providing from tests, former aircrafts etc...
  - $h =$ aeronautic model with parameters & uncertainties

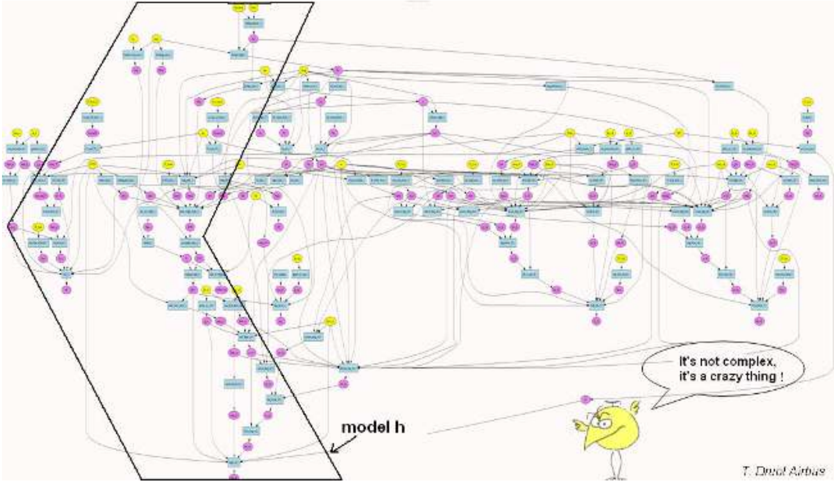
Numerical Simulations under Uncertainties



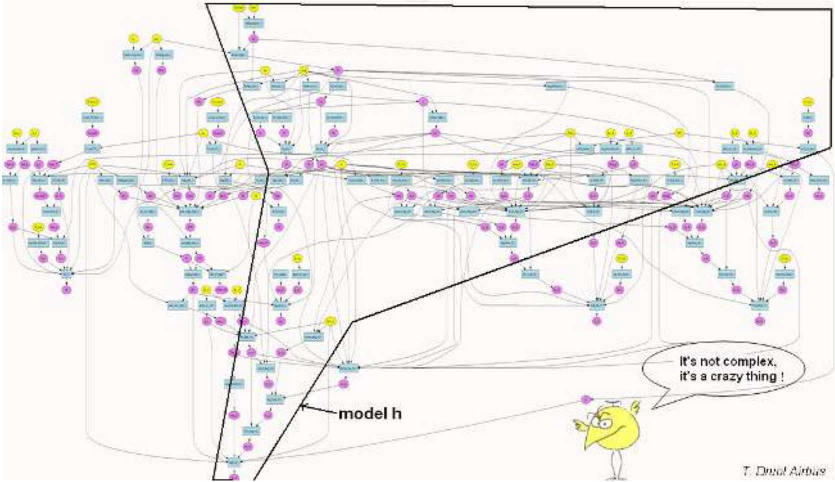
Numerical Simulations under Uncertainties



Numerical Simulations under Uncertainties



Numerical Simulations under Uncertainties



General Framework

- Reference data : $Z_1 = (\mathbf{X}_1, Y_1), ..., Z_n = (\mathbf{X}_n, Y_n)$ with (unknown) dist. $Q^{\mathbf{z}}$ and denote by Q the marginal dist. of Y
  - $\mathbf{X}_1, ..., \mathbf{X}_n$ may be unobserved (too complex, $\neq$ input codes etc...)
- Models : $\{\mathbf{x} \in \mathcal{X} \mapsto h(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{Y}, \quad \boldsymbol{\theta} \in \Theta\}$
  - mathematical models : $h(\mathbf{x}, \boldsymbol{\theta}) = \sum_{i=1}^{l=q} \phi(\mathbf{x})\, \boldsymbol{\theta}$ etc ...
  - physical/simulation models : $h(\mathbf{x}, \boldsymbol{\theta})$ is the result of a computer code
- Uncertainty : equip $\mathcal{X}$ with a prob. measure $P^{\mathbf{x}}$ - $\mathbf{x} \to \mathbf{X} \in (\mathcal{X}, P^{\mathbf{x}})$
  - stochastic codes, Monte-Carlo codes, uncertain variables etc...

General Framework

- Reference data : $Z_1 = (\mathbf{X}_1, Y_1), ..., Z_n = (\mathbf{X}_n, Y_n)$ with (unknown) dist. $Q^{\mathbf{z}}$ and denote by Q the marginal dist. of Y
  - $\mathbf{X}_1, ..., \mathbf{X}_n$ may be unobserved (too complex, $\neq$ input codes etc...)
- Models : $\{\mathbf{x} \in \mathcal{X} \mapsto h(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{Y}, \quad \boldsymbol{\theta} \in \Theta\}$
  - mathematical models : $h(\mathbf{x}, \boldsymbol{\theta}) = \sum_{i=1}^{l=q} \phi(\mathbf{x})\,\boldsymbol{\theta}$ etc ...
  - physical/simulation models : $h(\mathbf{x}, \boldsymbol{\theta})$ is the result of a computer code
- Uncertainty : equip $\mathcal{X}$ with a prob. measure $P^{\mathbf{x}}$ - $\mathbf{x} \rightarrow \mathbf{X} \in (\mathcal{X}, P^{\mathbf{x}})$
  - stochastic codes, Monte-Carlo codes, uncertain variables etc...
- **Goal : Predict a feature of the output Y**

General Framework

- Reference data : $Z_1 = (\mathbf{X}_1, Y_1), ..., Z_n = (\mathbf{X}_n, Y_n)$ with (unknown) dist. $Q^{\mathbf{z}}$ and denote by Q the marginal dist. of Y
  - $\mathbf{X}_1, ..., \mathbf{X}_n$ may be unobserved (too complex, $\neq$ input codes etc...)
- Models : $\{\mathbf{x} \in \mathcal{X} \mapsto h(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{Y}, \quad \boldsymbol{\theta} \in \Theta\}$
  - mathematical models : $h(\mathbf{x}, \boldsymbol{\theta}) = \sum_{i=1}^{l=q} \phi(\mathbf{x})\, \boldsymbol{\theta}$ etc ...
  - physical/simulation models : $h(\mathbf{x}, \boldsymbol{\theta})$ is the result of a computer code
- Uncertainty : equip $\mathcal{X}$ with a prob. measure $P^{\mathbf{x}}$ - $\mathbf{x} \rightarrow \mathbf{X} \in (\mathcal{X}, P^{\mathbf{x}})$
  - stochastic codes, Monte-Carlo codes, uncertain variables etc...
- **Goal : Predict a feature of the output** Y
  2 kind of predictions ...
  - Given a new input $\mathbf{X}$, predict the output Y

EADS

General Framework

- Reference data : $Z_1 = (\mathbf{X}_1, Y_1), ..., Z_n = (\mathbf{X}_n, Y_n)$ with (unknown) dist. $Q^{\mathbf{z}}$ and denote by Q the marginal dist. of Y
  - $\mathbf{X}_1, ..., \mathbf{X}_n$ may be unobserved (too complex, $\neq$ input codes etc...)
- Models : $\{\mathbf{x} \in \mathcal{X} \mapsto h(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{Y}, \quad \boldsymbol{\theta} \in \Theta\}$
  - mathematical models : $h(\mathbf{x}, \boldsymbol{\theta}) = \sum_{i=1}^{l=q} \phi(\mathbf{x})\, \boldsymbol{\theta}$ etc ...
  - physical/simulation models : $h(\mathbf{x}, \boldsymbol{\theta})$ is the result of a computer code
- Uncertainty : equip $\mathcal{X}$ with a prob. measure $P^{\mathbf{x}}$ - $\mathbf{x} \to \mathbf{X} \in (\mathcal{X}, P^{\mathbf{x}})$
  - stochastic codes, Monte-Carlo codes, uncertain variables etc...
- **Goal : Predict a feature of the output** Y
  2 kind of predictions ...
  - Given a new input $\mathbf{X}$, predict the output Y
  - probabilistic feature on Y (mean, quantile, pdf, exceedance prob., etc...)

Remark :
prediction = param. estimation + computation under the param.

Motivations

- Consider the classical modeling
  - $Y_i = h(\mathbf{X}_i, \boldsymbol{\theta}^\star) + \varepsilon_i\ , \quad i = 1, ..., n$
  - $\varepsilon_i \sim \mathcal{N}(0, 1)$ independent of $\mathbf{X}_i$,
- Suppose $\mathbf{X}_1, ..., \mathbf{X}_n$ observed

  $\rightarrow$ classical statistical learning (regression etc...)

Motivations

- Consider the classical modeling
  - $Y_i = h(\mathbf{X}_i, \boldsymbol{\theta}^\star) + \varepsilon_i \ , \quad i = 1, \dots, n$
  - $\varepsilon_i \sim \mathcal{N}(0,1)$ independent of $\mathbf{X}_i$,
- Suppose $\mathbf{X}_1, \dots, \mathbf{X}_n$ observed

  $\rightarrow$ classical statistical learning (regression etc...)

– Questions –

- If the $\mathbf{X}_i$'s are not observed ? How to calibrate ? (e.g Monte-Carlo codes, input code $\neq$ experimental conditions etc...)
- even if they are observed, should we always use regression parameters for prediction?
- meaning of ...

for example $\mathbb{P}(h(\mathbf{X}, \widehat{\theta}_{reg}) > s)$, $pdf_{h(\mathbf{X}, \widehat{\theta}_{reg})}$

... *duality* between estimation procedure and target prediction

EADS

Motivations

- Consider the classical modeling
  - $Y_i = h(\mathbf{X}_i, \boldsymbol{\theta}^\star) + \varepsilon_i$, $i = 1, ..., n$
  - $\varepsilon_i \sim \mathcal{N}(0,1)$ independent of $\mathbf{X}_i$,
- Suppose $\mathbf{X}_1, ..., \mathbf{X}_n$ observed

  $\rightarrow$ classical statistical learning (regression etc...)

– Questions –

- If the $\mathbf{X}_i$'s are not observed ? How to calibrate ? (e.g Monte-Carlo codes, input code $\neq$ experimental conditions etc...)
- even if they are observed, should we always use regression parameters for prediction?
- meaning of ...

  for example $\mathbb{P}(h(\mathbf{X}, \widehat{\theta}_{reg}) > s)$, $pdf_{h(\mathbf{X}, \widehat{\theta}_{reg})}$

  ... *duality* between estimation procedure and target prediction
- **Robust Prediction**

EADS

Calibration - Parameter estimation

- Feature Space $\mathcal{F}$ associated to a feature of $Q^{\mathbf{z}}$:
  - $(\mathbf{x} \mapsto \mathbb{E}(Y/\mathbf{X} = \mathbf{x})) \in \mathcal{F} \subset \{\rho : \mathcal{X} \to \mathcal{Y}\}$
  - $\mathbb{E}(Y) \in \mathcal{F} \subset \mathbb{R}$
  - (pdf of Y)$\in \mathcal{F} \subset \{\rho : \mathcal{Y} \to \mathbb{R}\}$
  - etc ...
- $\mathcal{F}$-Contrast function Ψ :

$$\begin{array}{rcl} \Psi \,:\, \mathcal{F} & \longrightarrow & L_1(Q^{\mathbf{z}}) \\ \rho & \longmapsto & \Psi(\rho, \cdot) \,:\, (\mathbf{x}, y) \in \mathcal{X} \times \mathcal{Y} \longmapsto \Psi(\rho, (\mathbf{x}, y)) \end{array}$$

- Ψ-Risk (to be minimized !) :

$$\mathcal{R}_\Psi(\rho) := \mathbb{E}_{Q^{\mathbf{z}}} \Psi\left(\rho\,,\, Z\right) \underset{\rho=\rho(\boldsymbol{\theta})}{\rightarrow} \mathcal{R}_\Psi(\boldsymbol{\theta}) := \mathbb{E}_{Q^{\mathbf{z}}} \Psi\left(\rho(\boldsymbol{\theta})\,,\, Z\right)$$

Rmk : think $\mathcal{R}_\Psi$ as a "distance" between model and true features

$$\mathcal{R}_\Psi(\boldsymbol{\theta}) \approx \mathcal{D}_\Psi(\rho(\boldsymbol{\theta}), \rho^*)$$

ex. $\mathcal{D}(\mathbb{E}(h(\mathbf{X}, \boldsymbol{\theta})), \mathbb{E}(Y))$, $\mathcal{D}(pdf_{h(\mathbf{X},\boldsymbol{\theta})}, pdf_Y)$, $\mathcal{D}(h(\cdot, \boldsymbol{\theta}), \mathbb{E}(Y/\mathbf{X} = \cdot))$ ...

Example of contrasts (*"the way of minimizing"*)

- $\mathcal{F} \subset \{\rho : \mathcal{X} \to \mathcal{Y}\}$

  regression contrast

$$\Psi\left(\rho, (\mathbf{x}, y)\right) = \left(y - \rho(\mathbf{x})\right)^2$$

- $\mathcal{F} \subset \mathbb{R}$: $\rho = \mathbb{E}(Y)$, $\mathbb{P}(Y > s)$ etc...

  mean contrast

$$\Psi\left(\rho, (\mathbf{x}, y)\right) = \Psi\left(\rho, y\right) = (y - \rho)^2$$

- $\mathcal{F} \subset$ {density functions on $\mathcal{Y}$}

  log-**contrast**

$$\Psi\left(\rho, (\mathbf{x}, y)\right) = \Psi\left(\rho, y\right) = -\log \rho(y)$$

  L_2-**contrast**

$$\Psi\left(\rho, (\mathbf{x}, y)\right) = \Psi\left(\rho, y\right) = ||\rho||_2^2 - 2\,\rho(y)$$

- etc...

Application to computer experiments

Recall: $("\mathbf{X}_i", Y_i)_{1..n}$, $\{\mathbf{x} \in \mathcal{X} \mapsto h(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{Y}, \boldsymbol{\theta} \in \Theta\}$, $\mathbf{X} \sim P^{\mathbf{x}} \rightarrow \boldsymbol{\theta}?$

Application to computer experiments

Recall: $(\text{"}\mathbf{X}_i\text{"}, Y_i)_{1..n}$, $\{\mathbf{x} \in \mathcal{X} \mapsto h(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{Y}, \boldsymbol{\theta} \in \Theta\}$, $\mathbf{X} \sim P^{\mathbf{x}} \rightarrow \boldsymbol{\theta}?$

- Learning Procedures
  - Regression: $\boldsymbol{\theta}_{reg} = \text{Argmin}_{\boldsymbol{\theta} \in \Theta} \mathbb{E}_{Q^{\mathbf{z}}} \left(Y - h(\mathbf{X}, \boldsymbol{\theta})\right)^2$

Application to computer experiments

Recall: $("\mathbf{X}_i", Y_i)_{1..n}$, $\{\mathbf{x} \in \mathcal{X} \mapsto h(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{Y}, \boldsymbol{\theta} \in \Theta\}$, $\mathbf{X} \sim P^{\mathbf{x}} \rightarrow \boldsymbol{\theta}$?

- Learning Procedures
  - Regression: $\boldsymbol{\theta}_{reg} = \text{Argmin}_{\boldsymbol{\theta} \in \Theta} \mathbb{E}_{Q^z} (Y - h(\mathbf{X}, \boldsymbol{\theta}))^2$
  - Density (log-)contrast: $\boldsymbol{\theta}_{log} = \text{Argmin}_{\boldsymbol{\theta} \in \Theta} \underbrace{-\mathbb{E}_Q(\log(\rho_{\boldsymbol{\theta}}(Y)))}_{"\mathcal{D}(pdf_{h(\mathrm{X},\boldsymbol{\theta})}, pdf_Y)"}$

    (where $Y \sim Q$ and $\rho_{\boldsymbol{\theta}}$ = pdf of $h(\mathrm{X}, \boldsymbol{\theta})$ both unknown !)

Application to computer experiments

Recall: $(\text{"}\mathbf{X}_i\text{"}, Y_i)_{1..n}$, $\{\mathbf{x} \in \mathcal{X} \mapsto h(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{Y}, \boldsymbol{\theta} \in \Theta\}$, $\mathbf{X} \sim P^{\mathbf{x}} \rightarrow \boldsymbol{\theta}$?

- Learning Procedures
  - Regression: $\boldsymbol{\theta}_{reg} = \text{Argmin}_{\boldsymbol{\theta} \in \Theta} \mathbb{E}_{Q^{\mathbf{z}}} (Y - h(\mathbf{X}, \boldsymbol{\theta}))^2$
  - Density (log-)contrast: $\boldsymbol{\theta}_{log} = \text{Argmin}_{\boldsymbol{\theta} \in \Theta} \underbrace{-\mathbb{E}_Q(\log(\rho_{\boldsymbol{\theta}}(Y)))}_{\text{"}\mathcal{D}(pdf_{h(\mathbf{X}, \boldsymbol{\theta})}, pdf_Y)\text{"}}$

    (where $Y \sim Q$ and $\rho_{\boldsymbol{\theta}}$ = pdf of $h(\mathbf{X}, \boldsymbol{\theta})$ both unknown !)
- Learning Algorithms (depend on database: $(\mathbf{X}_i, Y_i)_{1..n}$ or $Y_1, ..., Y_n$)
  - Regression (if $\mathbf{X}_i$ observed !):

    $\widehat{\boldsymbol{\theta}}_{reg} = \text{Argmin}_{\boldsymbol{\theta} \in \Theta} \frac{1}{n} \sum_{i=1}^{n} (Y_i - h(\mathbf{X}_i, \boldsymbol{\theta}))^2$ (well studied)

Application to computer experiments

Recall: $("\mathbf{X}_i", Y_i)_{1..n}$, $\{\mathbf{x} \in \mathcal{X} \mapsto h(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{Y},\ \boldsymbol{\theta} \in \Theta\}$, $\mathbf{X} \sim P^{\mathbf{x}} \rightarrow \boldsymbol{\theta}$?

- Learning Procedures
  - Regression: $\boldsymbol{\theta}_{reg} = \text{Argmin}_{\boldsymbol{\theta} \in \Theta} \mathbb{E}_{Q^z} \left(Y - h(\mathbf{X}, \boldsymbol{\theta})\right)^2$
  - Density (log-)contrast: $\boldsymbol{\theta}_{log} = \text{Argmin}_{\boldsymbol{\theta} \in \Theta} \underbrace{-\mathbb{E}_Q(\log(\rho_{\boldsymbol{\theta}}(Y)))}_{"\mathcal{D}(pdf_{h(\mathbf{X},\boldsymbol{\theta})},\, pdf_{Y})"}$

    (where $Y \sim Q$ and $\rho_{\boldsymbol{\theta}}$ = pdf of $h(\mathbf{X}, \boldsymbol{\theta})$ both unknown !)
- Learning Algorithms (depend on database: $(\mathbf{X}_i, Y_i)_{1..n}$ or $Y_1, ..., Y_n$)
  - Regression (if $\mathbf{X}_i$ observed !):
    $\widehat{\boldsymbol{\theta}}_{reg} = \text{Argmin}_{\boldsymbol{\theta} \in \Theta} \frac{1}{n} \sum_{i=1}^{n} (Y_i - h(\mathbf{X}_i, \boldsymbol{\theta}))^2$ (well studied)
  - Density contrast for computer experiments (N. Rachdi et al. 2010):

$$\widehat{\boldsymbol{\theta}}_{\log} = \underset{\boldsymbol{\theta} \in \Theta}{\text{Argmin}} - \sum_{i=1}^{n} \log \left(\sum_{j=1}^{m} K_b(Y_i - h(\mathbf{X}_j, \boldsymbol{\theta})) \right)$$

where $\mathbf{X}_1, ..., \mathbf{X}_m$ i.i.d from $P^{\mathbf{x}}$, $K()$ is a kernel, b bandwidth
($\rho_{\boldsymbol{\theta}}$ was estimated by a kernel smoothing)

EADS

General Result

Consider the general procedure

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}} \sum_{i=1}^{n} \Psi(\rho^m(\boldsymbol{\theta}), Y_i) \left(\approx \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}}\, \mathcal{D}_\Psi(\rho^m(\boldsymbol{\theta}), \rho^n) \right)$$

- $\rho^m(\boldsymbol{\theta})$ = emp. feature of $h(\mathbf{X}, \boldsymbol{\theta})$ based on $\mathbf{X}_{1..m}$
- and ρ^n = emp. feature of Y based on $Y_{1..n}$

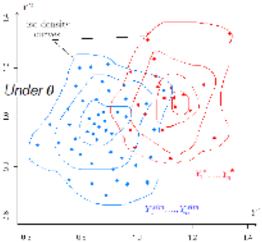
General Result

Consider the general procedure

$$\widehat{\theta} = \underset{\theta \in \Theta}{\text{Argmin}} \sum_{i=1}^{n} \Psi(\rho^m(\theta), Y_i) \left(\approx \underset{\theta \in \Theta}{\text{Argmin}}\, \mathcal{D}_\Psi(\rho^m(\theta), \rho^n) \right)$$

- $\rho^m(\theta)$ = emp. feature of $h(\mathbf{X}, \theta)$ based on $\mathbf{X}_{1..m}$
- and ρ^n = emp. feature of Y based on $Y_{1..n}$

- In blue: Simulated data
- In red: Reference data



EADS

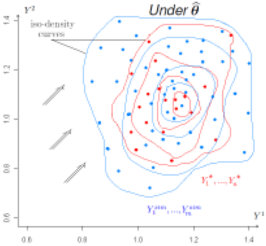
General Result

Consider the general procedure

$$\widehat{\theta} = \underset{\theta \in \Theta}{\text{Argmin}} \sum_{i=1}^{n} \Psi(\rho^m(\theta), Y_i) \left(\approx \underset{\theta \in \Theta}{\text{Argmin}}\, \mathcal{D}_\Psi(\rho^m(\theta), \rho^n) \right)$$

- $\rho^m(\theta)$ = emp. feature of $h(\mathbf{X}, \theta)$ based on $\mathbf{X}_{1..m}$
- and ρ^n = emp. feature of Y based on $Y_{1..n}$

- In blue: Simulated data
- In red: Reference data



EADS

General Result

Consider the general procedure

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\Theta}{\mathrm{Argmin}} \sum_{i=1}^{n} \Psi(\rho^m(\boldsymbol{\theta}), Y_i) \left(\approx \underset{\boldsymbol{\theta}\in\Theta}{\mathrm{Argmin}}\, \mathcal{D}_\Psi(\rho^m(\boldsymbol{\theta}), \rho^n) \right)$$

- $\rho^m(\boldsymbol{\theta})$ = emp. feature of $h(\mathbf{X}, \boldsymbol{\theta})$ based on $\mathbf{X}_{1..m}$
- and ρ^n = emp. feature of Y based on $Y_{1..n}$

Theorem: Oracle Inequality (N. Rachdi *et al* 2010)

Under some conditions on the contrast Ψ and under tightness conditions, for all $\varepsilon > 0$, with probability at least $1 - 2\varepsilon$ it holds

$$\mathcal{R}_\Psi(\widehat{\boldsymbol{\theta}}) \leq \inf_{\boldsymbol{\theta}\in\Theta} (\mathcal{R}_\Psi(\boldsymbol{\theta})) + \frac{K^{\varepsilon}_{(\widetilde{\rho},\Psi)}}{\sqrt{n}} \left(1 + \sqrt{\frac{n}{m}} (K^{\varepsilon}_{(\widetilde{\rho},h)} + B_m) \right)$$

where $K^{\varepsilon}_{(\widetilde{\rho},\Psi)}$, $K^{\varepsilon}_{(\widetilde{\rho},h)}$ some *concentration constants* and B_m a bias factor

Resume

- From **experimental data** $("\mathbf{X}_i", Y_i)_{i=1..n}$ and **simulated data** $h(\mathbf{X}_1, \theta), ..., h(\mathbf{X}_m, \theta)$, we propose others estimation procedures adapted to the *quantity of interest* we want to predict.
- In practice, regression parameters $(\widehat{\theta}_{reg})$ may be used to predict a lot of quantities:
  - quantile
  - exceedance probability
  - density function
  - etc ...

EADS

Resume

- From **experimental data** $("\mathbf{X}_i", Y_i)_{i=1..n}$ and **simulated data** $h(\mathbf{X}_1, \theta), \ldots, h(\mathbf{X}_m, \theta)$, we propose others estimation procedures adapted to the *quantity of interest* we want to predict.
- In practice, regression parameters $(\widehat{\theta}_{reg})$ may be used to predict a lot of quantities:
  - quantile
  - exceedance probability
  - density function
  - etc ...

  Question : How to quantify the estimation procedure error ?

EADS

Resume

- From **experimental data** ("$\mathbf{X}_i$", $Y_i)_{i=1..n}$ and **simulated data** $h(\mathbf{X}_1, \theta), ..., h(\mathbf{X}_m, \theta)$, we propose others estimation procedures adapted to the *quantity of interest* we want to predict.
- In practice, regression parameters $(\widehat{\theta}_{reg})$ may be used to predict a lot of quantities:
  - quantile
  - exceedance probability
  - density function
  - etc ...

  Question : How to quantify the estimation procedure error ?
- It can be investigated in terms of "*distance between contrasts*"

  (N. Rachdi *"A note about predicting with computer experiments"* (In preparation))

  $\Rightarrow$ key point: each quantity of interest is viewed as an *Argmin* of some Ψ-Risk $\mathcal{R}_\Psi$

Academic example

- Let consider

$$Y_i = \sin(X_i) + 0.01\,\varepsilon_i \quad i = 1, \ldots, n$$

  - $X_i \sim \mathcal{N}(0,1)$
  - $\varepsilon_i \sim \mathcal{N}(0,1)$ independent of X_i

- Model

$$h(\mathbf{X}, \boldsymbol{\theta}) = \theta_0 + \theta_1\, X + \theta_2\, X^3\,, \quad \mathbf{X} \sim \mathcal{N}(0,1)$$

- Goal: Predict the pdf of Y

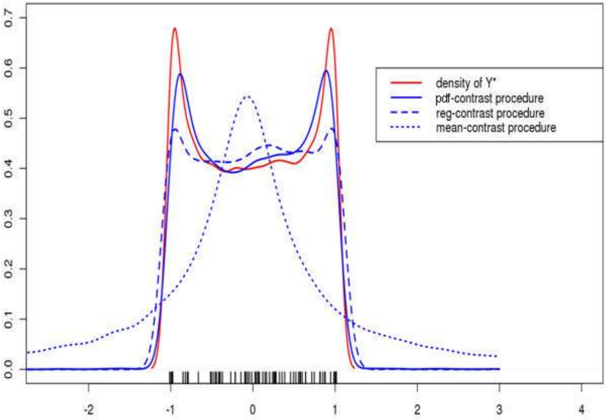
$\Rightarrow$ for this

- Compute some $\widehat{\boldsymbol{\theta}}$
- Compute the prediction propagating Uncertainties of $\mathbf{X}$ through the model $\mathbf{x} \mapsto h(\mathbf{x}, \widehat{\boldsymbol{\theta}})$

Predictions

- We compute density predictions by propagating uncertainties through models

$h(\mathbf{x}, \widehat{\theta}_{pdf})$ (solid line), $h(\mathbf{x}, \widehat{\theta}_{reg})$ (dashed line), $h(\mathbf{x}, \widehat{\theta}_{mean})$ (dotted line)



Perspectives

- Academic
  - Constants improvement in inequalities
  - Central Limit Theorems for the calibration parameter $\widehat{\boldsymbol{\theta}}_{\Psi}$
  - Functional analysis of *contrast functions*
- Industrial Applications
  - Run the learning algorithms with *real* computer codes...
  - EADS Applications

    $\rightarrow$ **Non Destructive Testing (*Prediction*)**
    combine reference data and simulated data for POD estimation

    $\rightarrow$ **Electromagnetism (*Inverse problem*)**
    Characterize slot parameters from sensors data and uncertain numerical models

Thank you for your attention !