

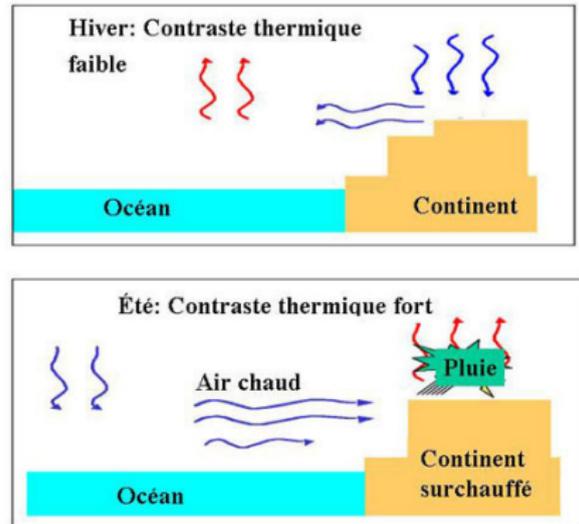
- 1** Modèle Physique
- 2** Analyse de sensibilité
- 3** Déploiement sur grille

Contexte Applicatif

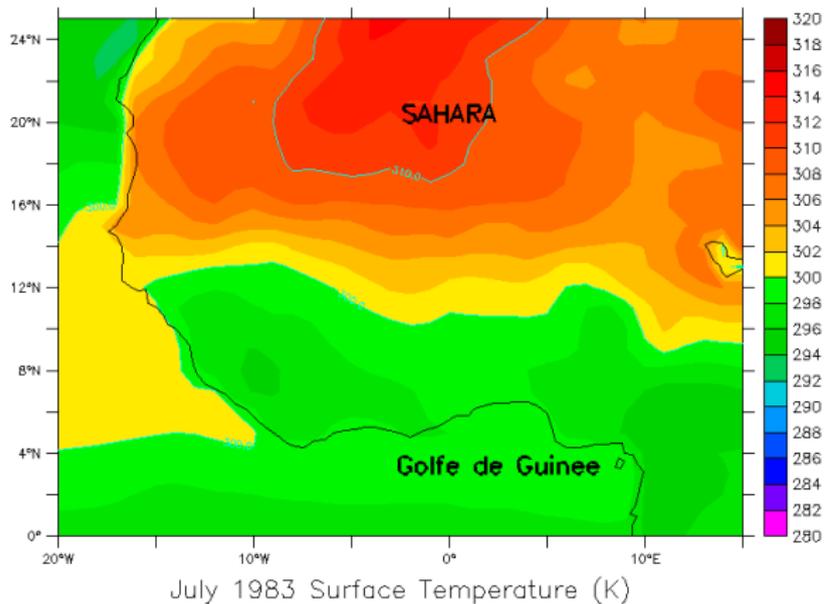
- Le climat de l'Afrique de l'ouest est piloté par un système de mousson avec une activité dynamique de mai à Septembre
- Le cumul pluviométrique d'une année se fait uniquement pendant la période du flux de mousson vers le Nord jusqu'au Sahel
- Ce cumul étant faible en zone Sahélienne (500 à 600mm), une faible variabilité à La baisse de ce cumul peut entraîner un stress hydrique irréversible pour les cultures et la végétation
- Si ce cumul perdure d'une année sur l'autre, les conséquences deviennent durables sur les écosystèmes qui tendent à passer à des écosystèmes de type Saharien
- Ce cycle saisonnier des précipitations est une composante climatique cruciale pour l'agriculture et la survie des populations locales dans les zones Sahéliennes et Soudano-Guinéennes.

Contrastes thermiques

- Le système de mousson d'Afrique de l'Ouest apparaît comme **un système couplé océan-atmosphère-continent**.
- Les anomalies de températures de surface de mer à l'échelle globale ont un impact important sur la **variabilité interannuelle et décennale des précipitations Sahéliennes**.
- Les **processus de surface continentale** (interactions avec la végétation et l'humidité du sol) jouent un rôle non négligeable dans la **dynamique de la circulation de mousson et des systèmes convectifs**



Température de surface Afrique de l'Ouest (Juillet 83)



Objectifs du projet

- Quelle est la réponse du climat Africain à la température de l'eau en surface du Golfe de Guinée?
- Quels sont les principaux processus responsables de la variabilité de la mousson? Prise en compte d'autres variables en entrée du système (albedo, végétation, . . .)
- Quelle est la dynamique spatio-temporelle du système?
- Utilisation d'un modèle régional atmosphérique: MAR
- Comment réagit le modèle MAR aux variations de ses entrées, variations dues à la physique et/ou variations dues aux incertitudes?
- . . .

Contexte de l'analyse de sensibilité

Phénomènes physiques \Rightarrow Modèles Mathématiques \Rightarrow Codes de Simulation

Modèle: Le modèle mathématique étudié est décrit par une fonction déterministe η qui fait correspondre à Chaque entrée (x_1, x_2, \dots, x_n) , une sortie scalaire ou vectorielle $y = (y_1, y_2, \dots, y_p)$

Enjeux: Les codes sont de plus en plus complexes, ils prennent en comptes de nombreuses variables en entrées et peuvent délivrer un grand nombre de résultats

Objectifs: Développer des méthodes performantes, même en grande dimension, permettant de mesurer la variabilité des sorties et de déterminer les facteurs en entrées les plus importants

Modèle Mathématique

Atmosphère

- **equations primitives** = Navier-Stokes
- **Cycle hydrologique**: equations de conservation (vapeur d'eau, gouttelettes, cristaux nuageux, flocons de neige, ...)
- **Paramétrisation des processus dynamiques sous-maille** (turbulence, ...)
- **Paramétrisation des nuages sous-maillles** (convection)

Surface: conservation de la chaleur et de l'eau des sols

Code de Simulation : code séquentielle - 20H pour simuler 15 journées - 8 mois simulés par année - 1.5-2Go en Entrée/Sorties

MAR - Implémentation

- Code séquentielle (fortran77 + un peu fortran90)
- Ressources nécessaires
 - 20H pour simuler 15 jours avec 1.5Go à 2Go d'E/S
 - Nécessité de simuler au moins un an avant d'avoir des sorties interprétables. Il faudra simuler deux ans par année observée
- En cours de "modernisation" nécessaire pour les contraintes de couplage et de performances :
 - passage au fortran90 (modules, interfaces génériques,...)
 - parallélisation
 - utilisation d'outils de déploiement
- Difficultés pour le déploiement sur grille
 - Volumes des données en E/S
 - La simulation de longues périodes se fait par intervalles de temps successifs \implies fortes dépendances entre les entrées sorties

DIET (Distributed Interactive Engineering Toolbox) va permettre de gérer les interdépendances entre les E/S ce que ne sait pas faire CIGRI

Bilan de masse (en surface)

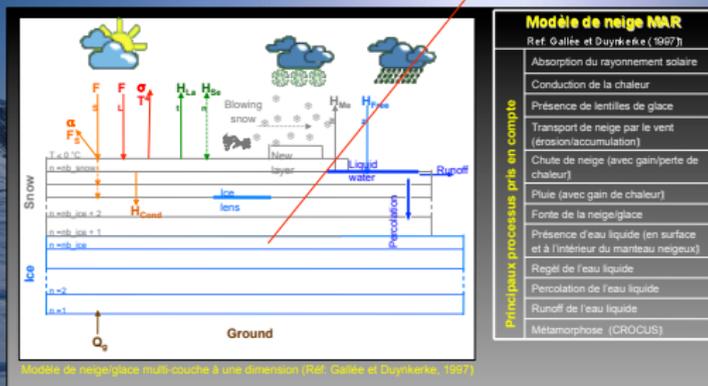
Objectif: étudier les processus climatiques

Outils:

- Modèle Atmosphérique Régional (MAR)
- Modèle Océan – Atmosphère couplé: MAR – NEMO (TANGO)
- Imbrication de MAR dans ECMWF

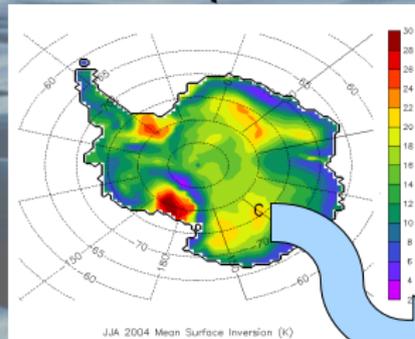
Modèle de neige du MAR

- Température
- Densité
- Teneur en eau liquide
- Taille des grains
- Dendricité
- Sphéricité
- Âges

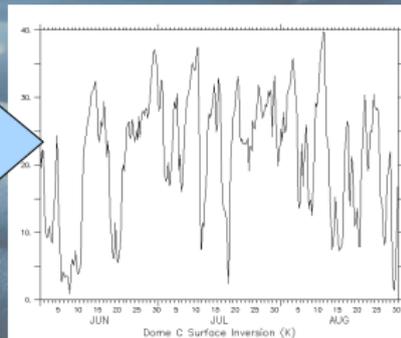


Fonte dans MAR: LWC > 1%

Inversion durant l'hiver antarctique

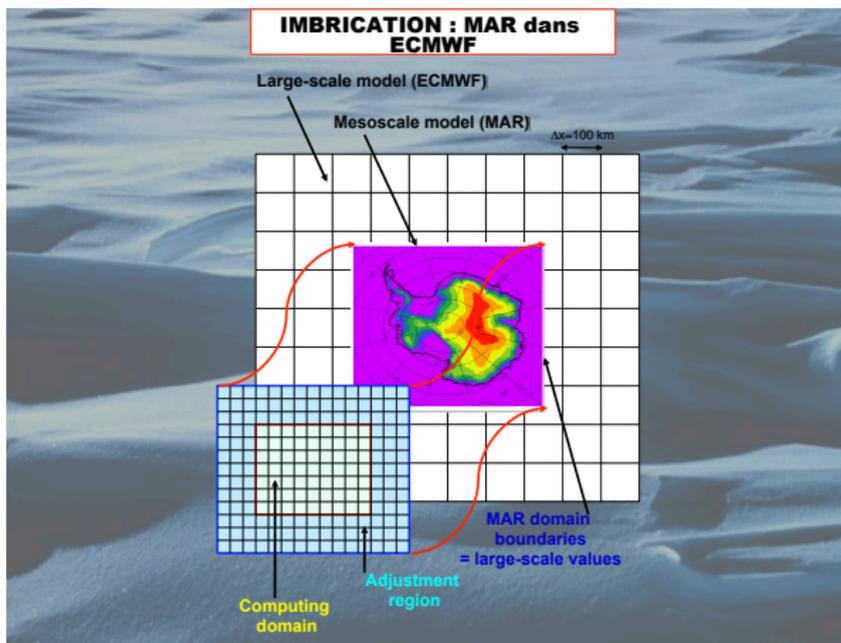


JJA 2004 Mean Surface Inversion (K)



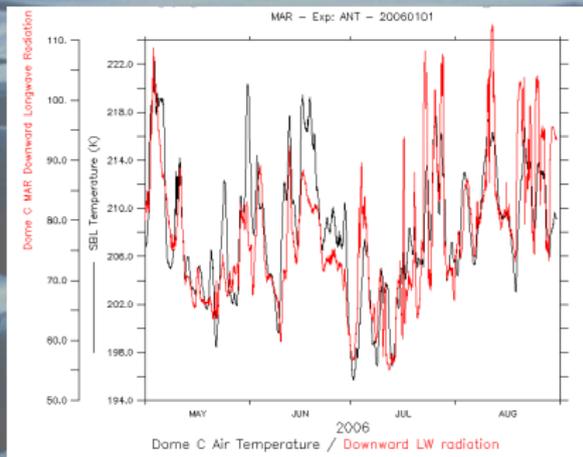
Dome C Surface Inversion (K)

ECMWF: European Centre for Medium-Range Weather Forecasts



L'hiver à Dome C

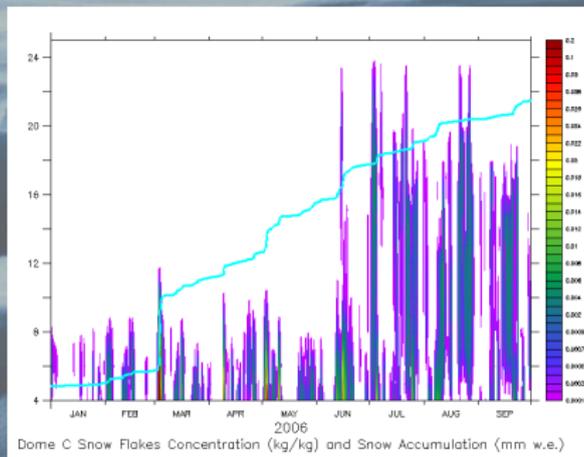
Simulation Temperature et **Flux infra-rouge descendant**



IR ↓ élevé
correspond à
une t° élevée

Nuages à Dome C

Particules de neige



Analyse d'incertitude - Analyse de sensibilité

Analyse d'incertitude: étude de **la variabilité des sorties** à partir de l'ensemble des évaluations du modèle.

Analyse de sensibilité: étude plus détaillée qui a pour objectif de **quantifier l'impact de la variation de chaque facteur** (variables, paramètres) **en entrée du modèle sur la variation des sorties**. En les classant éventuellement par degré d'influence.

- **Approche déterministe:** indicateurs locaux
- certaines **approches stochastiques** fournissent des indicateurs globaux.

Analyse de sensibilité - Difficultés

- P1 Les modèles physiques mettent en jeu des phénomènes **complexes** (non linéarité, effets de seuils,...) avec de fortes interactions entre les variables physiques
- P2 Les codes de calcul peuvent être **très couteux en ressources de calcul** (temps CPU, espace de stockage,...)
- P3 les modèles numériques peuvent avoir **un grand nombre de variables en entrées**
- P4 les modèles numériques fournissent **en sorties une grande nombre de variables intéressantes** qui peuvent dépendre du temps et de l'espace

Approche déterministe

- Une approche déterministe souvent utilisée par les physiciens consiste à faire varier les facteurs dans un intervalle donné et d'en déduire une intervalle de variation des sorties \implies intervalles non optimaux aux sens probabilistes puisqu'on suppose que les facteurs varient uniformément sur leur intervalle de valeurs.
Ces méthodes génèrent beaucoup de calcul lorsque le nombre de facteurs croit.
- L'information locale de ces variations peuvent êtes obtenues à partir des dérivées partielles de la fonction η
 - par dérivation de η lorsque η est connu analytiquement (rarement)
 - par approximation numérique lorsque η dérive d'un système d'équations aux dérivées partielles. Plusieurs méthodes sont employées, les plus efficaces reposent sur l'utilisation de l'adjoint du modèle

Approche stochastique

Pour obtenir **des informations globales** sur tout **le domaine des facteurs** on devra utiliser **une approche stochastique**

Soit η le modèle, **les facteurs d'entrées** sont les réalisations d'une **variable aléatoire vectorielle** $X = (X_1, X_2, \dots, X_n)$ dont les composantes sont indépendantes.

On associe à X une loi de probabilité composante par composante.

La sortie du modèle est un variable aléatoire $Y = \eta(X)$

$$F_Y(t) = P(Y \leq t) = \int_{\eta(x) \leq t} dF(x)$$

$$E(Y^k) = \int (\eta(x))^k dF(x)$$

$$P(Y > M) = \int_{\eta(x) > M} dF(x)$$

Évaluation des sorties du modèle sur un échantillon des entrées sélectionné au hasard \Rightarrow plans d'expérience.

Utilisation des sorties pour l'analyse d'incertitude et l'analyse de sensibilité.

5 étapes:

- 1** Détermination de l'étendue et de la distribution des variables en entrée X_i .
- 2** Échantillonnage des variables X_i .
- 3** Évaluation du modèle sur chaque élément de l'échantillon.
- 4** Analyse d'incertitude
- 5** Analyse de sensibilité \Rightarrow Calcul d'indices de sensibilité,...

Les codes de calcul deviennent de plus en plus complexe et couteux en ressources informatiques (temps CPU, stockage,...).

Décomposition de Sobol

Soient X_1, \dots, X_n variables aléatoires uniformément distribuées sur $[0, 1]$, f (modèle) une fonction de carré intégrable

La variance de $Y = f(X_1, \dots, X_n)$ se décompose de manière unique

$$\text{Var}(Y) = \sum_{i=1}^n V_i + \sum_{1 \leq i < j \leq n} V_{i,j} + \dots + V_{1,\dots,n}$$

$$V_i = \text{Var}[E(Y/X_i)]$$

$$V_{i,j} = \text{Var}[E(Y/X_i, X_j)] - V_i - V_j$$

$$V_{i,j,k} = \text{Var}[E(Y/X_i, X_j, X_k)] - V_i - V_j - V_k$$

...

Cette décomposition mène naturellement à la définition **des indices de sensibilité de Sobol**

Indice de sensibilité

On appelle **indice de sensibilité d'ordre s**

$$S_{i_1 i_2 \dots i_n} = \frac{V_{i_1, i_2, \dots, i_n}}{\text{Var}(Y)} \quad \text{avec}$$

$$1 \leq i_1 \leq \dots \leq i_s \leq n \quad \text{et} \quad 1 \leq s \leq n$$

$$\text{on a} \quad \sum_{s=1}^n \sum_{i_1 < \dots < i_s} S_{i_1 \dots i_s} = 1$$

$$S_{i_1 i_2 \dots i_n} = \frac{V_{i_1, i_2, \dots, i_n}}{\text{Var}(Y)} \quad \text{avec} \quad 1 \leq i_1 \leq \dots \leq i_s \leq n \quad \text{avec} \quad 1 \leq s \leq n$$

On a

$$\sum_{s=1}^n \sum_{i_1 < \dots < i_s} S_{i_1 \dots i_s} = 1$$

Chaque groupe de variables est responsable d'un pourcentage de la variabilité de la sortie Y

Méthode basées sur la variance

Indices mesurant le poids de la variance de Y expliquée par chaque facteur en entrée du modèle et par les interactions entre ces facteurs.

- **Modèle additif** on a

$\sum_{i=1}^n S_i = 1$ toute la variance est expliquée par les indices d'ordre 1

- **Modèle fortement couplé** on a

$\sum_{i=1}^n S_i = 1 \ll 1$ il faudra calculer les indices d'ordre supérieur

- Dans la pratique on ne calcule pas les indices au delà de l'ordre 2
- Les indices de Sobol s'adaptent aux phénomènes de non-linéarité et aux relations non monotones entre les sorties et les entrées (P1)
- Les méthodes de Monte-Carlo ne peuvent s'appliquer aux calculs des indices de sensibilité pour des codes couteux en CPU...

Vers la construction d'un méta-modèle

Difficultés Notre modèle physique:

- met en jeu des phénomènes complexes (non linéarité, effets de seuils,...) avec de fortes interactions entre les variables physiques
- le code de calcul associé est très coûteux en ressources de calcul (temps CPU, espace de stockage,...)
- a un grand nombre de variables en entrées qui peuvent dépendre du temps et de l'espace
- fournit en sorties un grand nombre de variables intéressantes qui peuvent dépendre du temps et de l'espace

Solution Construction d'un méta-modèle adapté

- Le **méta-modèle** devra simuler le comportement du code dans le domaine de variation des facteurs influents, à partir d'un nombre acceptable de runs.

Régression - Krigage

$$Y_{obs} = F_{obs}^T \beta + \epsilon$$

Y_{obs} : vecteur contenant les sorties observées réalisation d'un processus Gaussien

F_{obs} : base de régression évaluées sur les entrées observée

β, ϵ : vecteurs des paramètres à estimer, bruit

Régression linéaire: ϵ processus centré, ses composantes sont non corrélées $F_{obs}^T \beta$ est un modèle déterministe qui aurait récupéré toute l'information de Y_{obs}

- Krigeage:**
- la partie déterministe $F_{obs}^T \beta$ n'a pas récupérée toute l'information
 - ϵ possède une structure de covariance on devra en tenir compte lors de l'estimation des paramètres
 - Utilisation récente pour modéliser les résultats de code de calcul (Sacks, 1989; Santner, 2003)

Krigeage

Les **sorties du modèle** se décomposent en:

- une partie **déterministe** (tendance)
- une partie **stochastique**

Π : domaine de définition des facteurs en entrée du modèle

$$\pi \in \Pi \quad Y(\pi) = F(\pi)^T \beta + Z(\pi)$$

Z est un **processus aléatoire Gaussien**:

- de **moyenne nulle** et de **covariance** de la forme

$$E(Z(\pi_1), Z(\pi_2)) = \sigma^2 R(\theta, \pi_2 - \pi_1)$$

σ et θ sont **deux paramètres à estimer**, R est une fonction de corrélation **imposée à priori**.

Le processus Z **satisfait les hypothèses de stationnarités**, hypothèses qui simplifient le modèle, utiles pour l'estimation des paramètres.

Krigeage et notre application

Objectifs: On conserve dans un premier temps comme facteur influent, la température de surface de l'Océan dans le Golfe de Guinée (TSO), le modèle nous permet d'évaluer les précipitations sur les zones critiques (Sahel,...), on veut observer la dynamique spatio-temporelle du système sur une année

Modèle: Les entrées/sorties du modèle:

Entrées: On observe la température de surface de l'Océan (TSO) sur les points du maillage couvrant le Golfe de Guinée sur 117 années

Sorties: On évalue les précipitations grâce au modèle sur les points du maillage recouvrant les domaines critiques (Sahel,..) sur les 17 années

Les zones d'observations des entrées comme des sorties sont définies par les physiciens, elles sont issues de leurs connaissances à priori.

Krigeage et notre application(suite)

- la dimension des vecteurs aléatoire X et Y augmentent avec le nombre de points du maillage \Rightarrow **la taille de la matrice associée à la fonction de covariance** à inverser pour faire les estimations devient vite très grande
 - Les fonctions de covariance usuelles utilisées dans le krigeage imposent **l'hypothèse de stationnarité des dépendances spatiales** \Rightarrow incompatible avec notre problème
 - **Pour traiter la non stationnarité**, trois approches:
 - 1 Utiliser "the intrinsic random process"
 - 2 Modéliser la structure spatiale non stationnaire par une déformation d'espace
 - 3 Partitionner l'espace en région sur lesquelles l'hypothèse de stationnarité est acceptable
- \Rightarrow **on s'oriente vers une autre approche**

Méthodologie envisagée

On est dans le cas d'un système complexe où on doit considérer la variabilité spatio-temporelle des variables en entrée comme en sortie et récupérer la dynamique spatio-temporelle du système.

- Les entrées comme les sorties du modèle sont fonctionnelles, elles seront modélisées comme réalisations de deux éléments Gaussiens à valeurs dans un espace fonctionnel (Banach ou Hilbert), décomposition de Karuhnen-Loève
- La régression entre les entrées et les sorties sera modélisée par une régression fonctionnelle
- L'opérateur de régression sera décomposé sur une base fonctionnelle appartenant au dual topologique de l'espace des entrées
- Cette méthode va permettre de définir par dualité les indices de sensibilité et de les estimer

Décomposition de Karuhnen-Loève

Analyse en composantes principales fonctionnelle

X_1, \dots, X_m m fonctions aléatoires de domaine τ interval fermé borné

$$E(X(t)) = \mu(t) \quad t \in \tau$$

$$G(s, t) = \text{cov}(X(s), X(t)) \quad \text{fonction de covariance}$$

On suppose qu'il existe une expansion orthogonale (au sens de L^2) de fonctions propres ϕ_k et de valeurs propres λ_k

$$G(s, t) = \sum_k \lambda_k \phi_k(t) \phi_k(s) \quad t, s \in \tau$$

En analyse en composantes principales fonctionnelle on a

$$X_i(t) = \mu(t) + \sum_k \xi_{ik} \phi_k(t) \quad t \in \tau$$

$$\xi_{ik} \text{ v.a. } \perp / E(\xi_{ik}) = 0 \quad \text{et} \quad E(\xi_{ik}^2) = \lambda_k$$

$$\sum \lambda_k < \infty \quad \lambda_1 \leq \lambda_2 \leq \dots$$

Données :

\mathcal{R} zone du golf de Guinée sur laquelle on observe la TSO

\mathcal{R}' zone (Sub)-Saharienne sur laquelle on observe les précipitations

\mathcal{T} période de temps (grosso modo de mai à septembre)

17 années

$\mathbf{X}^i := (X_i(x, t))_{x \in \mathcal{R}, t \in \mathcal{T}}$ TSO année i

$\mathbf{Y}^i := (Y_i(x', t))_{x' \in \mathcal{R}', t \in \mathcal{T}}$ précipitations année i

Méthode :

A $x \in \mathcal{R}$ et $x' \in \mathcal{R}'$ fixés, régression linéaire fonctionnelle de

$\mathbf{Y}_{x'} := (Y(x', t))_{t \in \mathcal{T}}$ sur $\mathbf{X}_x := (X(x, t))_{t \in \mathcal{T}}$.

On suppose les entrées et les sorties centrées (prétraitement sinon)

étape 1

On effectue une **décomposition de Karuhnen-Loève** pour \mathbf{X}_{x_0} pour X_0 **fixé**

⇒ fonctions propres $e_k(t)$, $t \in \mathcal{T}$, valeurs propres $\lambda_k(x_0)$

En estime les fonctions e_k à l'aide des observation $\mathbf{X}_{x_0}^i$, $i = 1, \dots, 17$
(par la suite, on oublie leur dépendance en x_0)

On suppose les entrées et les sorties centrées (prétraitement sinon)

étape 1

On effectue une **décomposition de Karuhnen-Loève** pour \mathbf{X}_{x_0} pour X_0 **fixé**

\Rightarrow fonctions propres $e_k(t)$, $t \in \mathcal{T}$, valeurs propres $\lambda_k(x_0)$

En estime les fonctions e_k à l'aide des observation $\mathbf{X}_{x_0}^i$, $i = 1, \dots, 17$
(par la suite, on oublie leur dépendance en x_0)

étape 2

en p points x_j de \mathcal{R} on décompose \mathbf{X}_{x_j} sur la base e_k

$$\mathbf{X}_{x_j} = \left(\sum_k \lambda_k(x_j) e_k(t) \right)_{t \in \mathcal{T}}$$

à l'aide des observations $\mathbf{X}_{x_j}^i$, $i = 1, \dots, 17$ on estime et les $\lambda_k(x_j)$,
 $j = 1, \dots, p$

étape 3

On interpole (splines, noyau, ...) pour avoir $\lambda_k(x) \forall x \in \mathcal{R}$

étape 3

On interpole (splines, noyau, ...) pour avoir $\lambda_k(x) \forall x \in \mathcal{R}$

Hypothèse : A l'opérateur de régression

$$\mathbb{E}(\mathbf{Y}_{x'} | \mathbf{X}_x) = A(\mathbf{X}_x) = \sum_k \lambda_k(x) A e_k$$

Pour tout $x' \in \mathcal{R}'$, $\mathbf{Y}_{x'}$ s'écrit $\sum_k \beta_k(x') f_k$

étape 3

On interpole (splines, noyau, ...) pour avoir $\lambda_k(x) \forall x \in \mathcal{R}$

Hypothèse : A l'opérateur de régression

$$\mathbb{E}(\mathbf{Y}_{x'} | \mathbf{X}_x) = A(\mathbf{X}_x) = \sum_k \lambda_k(x) A e_k$$

Pour tout $x' \in \mathcal{R}'$, $\mathbf{Y}_{x'}$ s'écrit $\sum_k \beta_k(x') f_k$

étape 4

on choisit r points x'_l de \mathcal{R}' , à l'aide des observations $\mathbf{Y}_{x'_l}^i$,
 $i = 1, \dots, 17$ on estime les $\beta_k(x'_l)$, $l = 1, \dots, r$

étape 3

On interpole (splines, noyau, ...) pour avoir $\lambda_k(x) \forall x \in \mathcal{R}$

Hypothèse : A l'opérateur de régression

$$\mathbb{E}(\mathbf{Y}_{x'} | \mathbf{X}_x) = A(\mathbf{X}_x) = \sum_k \lambda_k(x) A e_k$$

Pour tout $x' \in \mathcal{R}'$, $\mathbf{Y}_{x'}$ s'écrit $\sum_k \beta_k(x') f_k$

étape 4

on choisit r points x'_l de \mathcal{R}' , à l'aide des observations $\mathbf{Y}_{x'_l}^i$,
 $i = 1, \dots, 17$ on estime les $\beta_k(x'_l)$, $l = 1, \dots, r$

étape 5

pour chaque $l = 1, \dots, r$, régression non paramétrique de $\beta_k(x'_l)$ sur les $\lambda_k(x_j), j = 1, \dots, p$

étape 5

pour chaque $l = 1, \dots, r$, régression non paramétrique de $\beta_k(x'_l)$ sur les $\lambda_k(x_j), j = 1, \dots, p$

étape 6

on fait un lissage (régression, noyau,...) pour avoir $\beta_k(x') \forall x' \in \mathcal{R}'$

étape 5

pour chaque $l = 1, \dots, r$, régression non paramétrique de $\beta_k(x'_l)$ sur les $\lambda_k(x_j), j = 1, \dots, p$

étape 6

on fait un lissage (régression, noyau,...) pour avoir $\beta_k(x') \forall x' \in \mathcal{R}'$

Bilan

estimation **off line** des $e_k, \lambda_k(x_j), j = 1, \dots, p$ et des $E(\beta_k(x'_l) | \mathbf{X}_{x_j}), l = 1, \dots, r, j = 1, \dots, p$

puis **on line** on estime $E(\mathbf{Y}_{x'} | \mathbf{X}_x)$ pour n'importe quel couple $(x, x') \in \mathcal{R} \times \mathcal{R}'$

étape 5

pour chaque $l = 1, \dots, r$, régression non paramétrique de $\beta_k(x'_l)$ sur les $\lambda_k(x_j), j = 1, \dots, p$

étape 6

on fait un lissage (régression, noyau,...) pour avoir $\beta_k(x') \forall x' \in \mathcal{R}'$

Bilan

estimation **off line** des $e_k, \lambda_k(x_j), j = 1, \dots, p$ et des $E(\beta_k(x'_l) | \mathbf{X}_{x_j}), l = 1, \dots, r, j = 1, \dots, p$

puis **on line** on estime $E(\mathbf{Y}_{x'} | \mathbf{X}_x)$ pour n'importe quel couple $(x, x') \in \mathcal{R} \times \mathcal{R}'$

au passage on récupère l'estimation de $\mathbb{E}(\mathbf{Y}_{x'} | \mathbf{X}_x)$ pour l'analyse de sensibilité (calcul des indices)

On considère dans un premier temps, une variable d'entrée, la TSO et la variable "précipitation" en variable de sortie du modèle

- **Modélisation fonctionnelle de la TSO en entrée du modèle:** on travaille sur un échantillon de 17 années observées sur un maillage de 516 points dans le golf de Guinée.
- **Modélisation fonctionnelle des précipitations en sortie du modèle:** on utilise les précipitations fournies par le modèle pendant les 17 années. On se limitera à la zone Sub-Saharienne.
- Pour chaque année observée, on simule 2 années
⇒ 17X2 années simulées ⇒ 816 runs de 15 jours (20h)

On adopte une méthode de déploiement sur grille permettant la gestion du "Workflow"

DIET: Distributed Interactive Engineering Toolbox

Permet de déployer de grosses applications sur des grilles de calculateurs.

- Une **boîte à outils** dont le but est de rendre l'utilisation d'une grille informatique aussi transparente que possible pour l'utilisateur
- C'est un **'intergiciel'** qui se situe donc entre la partie applicative d'intérêt et les ressources informatiques à disposition.
- **Se charge de:**

scheduling: quelle machine est la mieux à même d'effectuer une tâche à un moment donné.

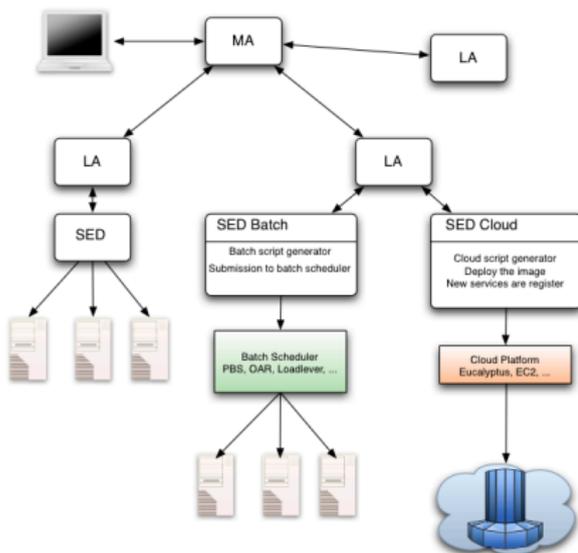
Data managment: persistance des données, (re)distribution des données, workflow, ?

interopérabilité: une grille pouvant être très hétérogène

Architecture de DIET

- Un **client** va interroger un **'Master Agent'**, point d'entrée vers un ensemble de ressources.
- Chaque **MA** peut éventuellement être **responsable de plusieurs 'Local Agents'**.
- Les **LA** communiquent directement avec des **Servers Deamons**, partie applicative sur chaque serveur.
- Les **SeD** informent leur agent sur **leur disponibilité**.
- **L'information à remonter au client** qui saura alors à quel **SeD** s'adresser pour concrètement effectuer sa requête.
- Nécessite **l'écriture d'un client et d'un serveur** spécifiques pour chaque application à porter sur la grille.
- **Le client s'occupe d'envoyer les données nécessaires** à l'accomplissement d'une tâche.

Le serveur attend ces données, effectue la tâche puis renvoie le résultat au client



Pourquoi DIET ?

Contexte au sein de la grille Rhône-Alpes:

- Un interfaçage éprouvé avec les outils OAR (allocation de ressources) et KaDeploy (déploiement à grande échelle), **tous deux développés à Grenoble.**

<http://graal.ens-lyon.fr/GRUDU> (Grid5000 Reservation Utility for Deployment Usage)

- DIET est disponible pour l'ensemble des architectures présentes (**hétérogénéité**)

Contexte propre au projet:

- Utilise les ressources **en mode “best-effort”**, DIET facilite la relance automatique d'une tâche qui aurait été interrompue.
- **Gestion avancée des flux de données** (workflow)
- **Conception dynamique** de DIET, adapté à des calculs longs multiples (découverte automatique de nouvelles ressources disponibles ou capacité à gérer la disparition d'un SeD).