

Réunion PAN

Décembre 2005

Stéphane Gazut et Jean-Marc Martinez

**Plans d'expériences numériques, une comparaison entre la
D-Optimalité et une méthode utilisant le Bagging.**



Introduction

- Toute modélisation d'un processus inconnu se base sur un certain nombre d'expérimentations.



Introduction

- Toute modélisation d'un processus inconnu se base sur un certain nombre d'expérimentations.
- Elles conditionnent fortement la qualité du modèle construit et le choix est d'autant plus judicieux qu'elles sont couteuses (temps ou argent)



Introduction

- Toute modélisation d'un processus inconnu se base sur un certain nombre d'expérimentations.
- Elles conditionnent fortement la qualité du modèle construit et le choix est d'autant plus judicieux qu'elles sont coûteuses (temps ou argent)
- La théorie des plans d'expériences répond à cette problématique. *Obtenir le maximum d'information avec le minimum d'expérimentations.*

- Toute modélisation d'un processus inconnu se base sur un certain nombre d'expérimentations.
- Elles conditionnent fortement la qualité du modèle construit et le choix est d'autant plus judicieux qu'elles sont coûteuses (temps ou argent)
- La théorie des plans d'expériences répond à cette problématique. *Obtenir le maximum d'information avec le minimum d'expérimentations.*
- Nous proposons une méthode itérative de construction de plans d'expériences **numériques** grâce à une étude statistique *a posteriori* sur les sorties prédites par une population de modèles.



Introduction

Cette étude s'inscrit dans le cadre de l' "Active Learning" .

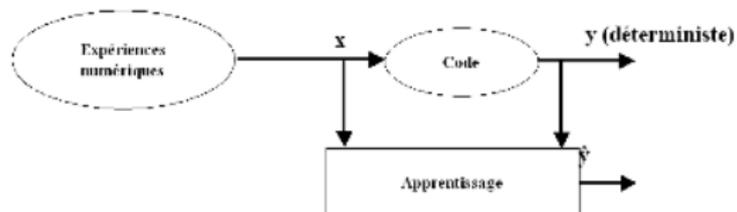
- Le but de l'apprentissage "*classique*" est d'estimer une fonction en utilisant un plan d'expériences donné.

Cette étude s'inscrit dans le cadre de l' "Active Learning" .

- Le but de l'apprentissage "*classique*" est d'estimer une fonction en utilisant un plan d'expériences donné.
- Le but de l'apprentissage actif est d'estimer une fonction en utilisant un plan d'expériences réduit et en spécifiant de manière itérative les points à ajouter.

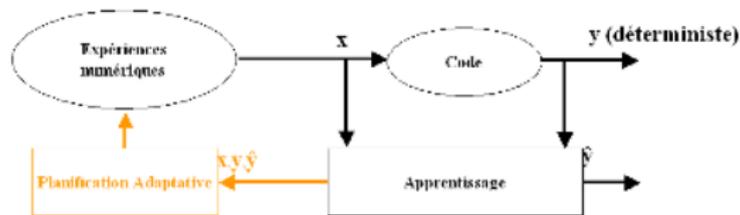
Cette étude s'inscrit dans le cadre de l' "Active Learning".

- Le but de l'apprentissage "*classique*" est d'estimer une fonction en utilisant un plan d'expériences donné.
- Le but de l'apprentissage actif est d'estimer une fonction en utilisant un plan d'expériences réduit et en spécifiant de manière itérative les points à ajouter.



Cette étude s'inscrit dans le cadre de l' "Active Learning".

- Le but de l'apprentissage "classique" est d'estimer une fonction en utilisant un plan d'expériences donné.
- Le but de l'apprentissage actif est d'estimer une fonction en utilisant un plan d'expériences réduit et en spécifiant de manière itérative les points à ajouter.



- 1 Principe des deux méthodes
 - La replanification D-Optimale
 - La replanification par Bagging

- 1 Principe des deux méthodes
 - La replanification D-Optimale
 - La replanification par Bagging
- 2 Les Cas test
 - Friedman
 - Homma et Saltelli
 - Sélection du modèle pour Homma et Saltelli

- 1 Principe des deux méthodes
 - La replanification D-Optimale
 - La replanification par Bagging
- 2 Les Cas test
 - Friedman
 - Homma et Saltelli
 - Sélection du modèle pour Homma et Saltelli
- 3 Conclusion

- 1 Principe des deux méthodes
 - La replanification D-Optimale
 - La replanification par Bagging
- 2 Les Cas test
 - Friedman
 - Homma et Saltelli
 - Sélection du modèle pour Homma et Saltelli
- 3 Conclusion



La replanification D-optimale

- Longtemps utilisée dans le cadre linéaire (modèles linéaires en les paramètres comme les polynômes).



La replanification D-optimale

- Longtemps utilisée dans le cadre linéaire (modèles linéaires en les paramètres comme les polynômes).
- **Le But** : Trouver les points dans l'espace des entrées qui minimisent la région de confiance des paramètres θ .



La replanification D-optimale

- Longtemps utilisée dans le cadre linéaire (modèles linéaires en les paramètres comme les polynômes).
- **Le But** : Trouver les points dans l'espace des entrées qui minimisent la région de confiance des paramètres θ .
- $\hat{\theta} = (X^T X)^{-1} X^T y$ donc
 $V(\hat{\theta}) = (X^T X)^{-1} X^T V(y) X (X^T X)^{-1}$



La replanification D-optimale

- Longtemps utilisée dans le cadre linéaire (modèles linéaires en les paramètres comme les polynômes).
- **Le But** : Trouver les points dans l'espace des entrées qui minimisent la région de confiance des paramètres θ .
- $\hat{\theta} = (X^T X)^{-1} X^T y$ donc
 $V(\hat{\theta}) = (X^T X)^{-1} X^T V(y) X (X^T X)^{-1}$
- Sous les hypothèses d'un bruit gaussien en sortie
 $\hat{\theta} = N(\theta, \sigma^2 (X^T X)^{-1})$



La replanification D-optimale

- Longtemps utilisée dans le cadre linéaire (modèles linéaires en les paramètres comme les polynômes).
- **Le But** : Trouver les points dans l'espace des entrées qui minimisent la région de confiance des paramètres θ .
- $\hat{\theta} = (X^T X)^{-1} X^T y$ donc
 $V(\hat{\theta}) = (X^T X)^{-1} X^T V(y) X (X^T X)^{-1}$
- Sous les hypothèses d'un bruit gaussien en sortie
 $\hat{\theta} = N(\theta, \sigma^2 (X^T X)^{-1})$
- La région de confiance des paramètres estimés $\hat{\theta}$ est une ellipse, dans l'espace des paramètres, dépendant de la variance σ^2 .

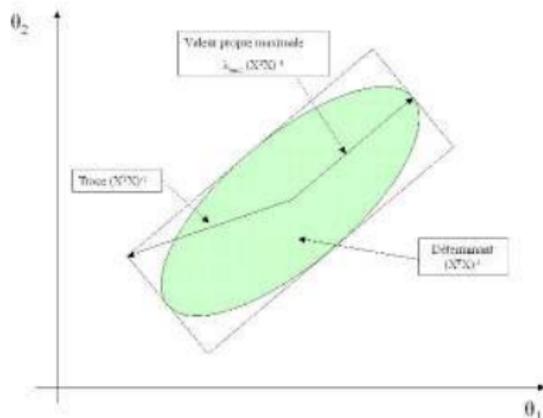


La replanification D-optimale

- $(\theta - \hat{\theta})^T (X^T X)^{-1} (\theta - \hat{\theta}) \leq \sigma^2 (1 - \alpha)$

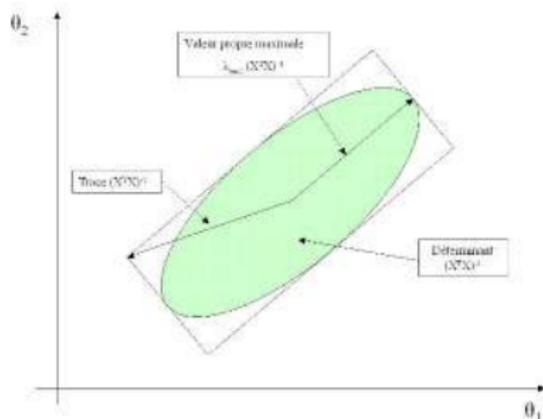
La replanification D-optimale

- $(\theta - \hat{\theta})^T (X^T X)^{-1} (\theta - \hat{\theta}) \leq \sigma^2 (1 - \alpha)$



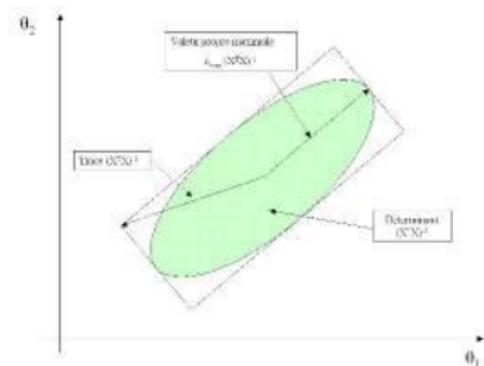
La replanification D-optimale

- $(\theta - \hat{\theta})^T (X^T X)^{-1} (\theta - \hat{\theta}) \leq \sigma^2 (1 - \alpha)$



- Le volume de l'ellipsoïde dépend de σ^2 et de la matrice X qui correspond au placement des points dans l'espace des entrées. Il apparait alors naturel de chercher un plan X qui minimise le volume de l'ellipsoïde. (*méthode a priori*)

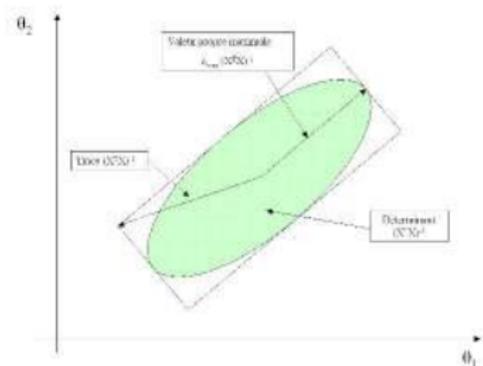
La replanification D-optimale



On peut réduire :

- le grand axe de l'ellipse (valeur propre maximale de $(X^T X)^{-1}$) : E-optimalité

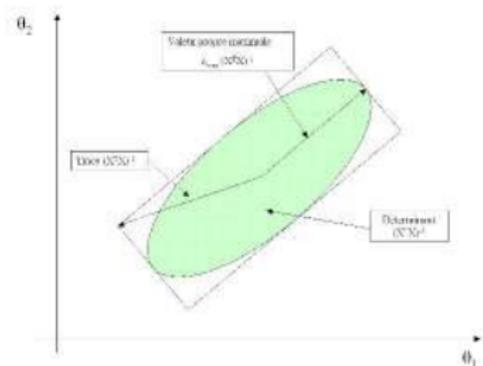
La replanification D-optimale



On peut réduire :

- le grand axe de l'ellipse (valeur propre maximale de $(X^T X)^{-1}$) : E-optimalité
- la diagonale du pavé exinscrit à l'ellipse (la trace de $(X^T X)^{-1}$) : A-optimalité

La replanification D-optimale



On peut réduire :

- le grand axe de l'ellipse (valeur propre maximale de $(X^T X)^{-1}$) : E-optimalité
- la diagonale du pavé exinscrit à l'ellipse (la trace de $(X^T X)^{-1}$) : A-optimalité
- le volume de l'ellipse (le déterminant de $(X^T X)^{-1}$) : D-optimalité



La replanification $D(\theta^0)$ -optimale

En non linéaire, on fait une approximation linéaire par développement limité de Taylor au premier ordre. La matrice jacobienne du modèle joue alors un rôle central.

$$Z(x_k, \theta) = \left[\frac{\partial f(x_k, \theta)}{\partial \theta_{ij}} \right]_{ij}$$



La replanification $D(\theta^0)$ -optimale

En non linéaire, on fait une approximation linéaire par développement limité de Taylor au premier ordre. La matrice jacobienne du modèle joue alors un rôle central.

$$Z(x_k, \theta) = \left[\frac{\partial f(x_k, \theta)}{\partial \theta_{ij}} \right]_{ij}$$

Il faut un jeu d'expériences initiales de taille n_0 pour estimer les paramètres θ afin de calculer les éléments de la matrice Jacobienne. On utilise généralement un plan LHS. On peut ainsi linéariser le modèle : $Y = Z(X, \Theta) \cdot \Theta$



La replanification $D(\theta^0)$ -optimale

En non linéaire, on fait une approximation linéaire par développement limité de Taylor au premier ordre. La matrice jacobienne du modèle joue alors un rôle central.

$$Z(x_k, \theta) = \left[\frac{\partial f(x_k, \theta)}{\partial \theta_{ij}} \right]_{ij}$$

Il faut un jeu d'expériences initiales de taille n_0 pour estimer les paramètres θ afin de calculer les éléments de la matrice Jacobienne. On utilise généralement un plan LHS. On peut ainsi linéariser le modèle : $Y = Z(X, \Theta) \cdot \Theta$

On utilise la matrice Z pour spécifier N expériences D-optimales en minimisant $\det((Z^T Z)^{-1})$.

Le plan d'expériences final est bien souvent constitué des $n_0 + N$ expériences.

- 1 Principe des deux méthodes
 - La replanification D-Optimale
 - La replanification par Bagging
- 2 Les Cas test
 - Friedman
 - Homma et Saltelli
 - Sélection du modèle pour Homma et Saltelli
- 3 Conclusion



Le Bagging

- Pour faire de la replanification par Bootstrap, nous utilisons le Bagging dans une version un peu modifiée.



Le Bagging

- Pour faire de la replanification par Bootstrap, nous utilisons le Bagging dans une version un peu modifiée.
- Le Bagging consiste à entraîner B modèles sur B répliques Bootstrap de la base initiale et de considérer la moyenne des réponses des B modèles comme étant la sortie estimée de la modélisation.

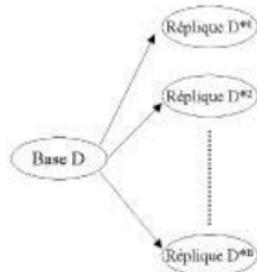
Le Bagging

- Pour faire de la replanification par Bootstrap, nous utilisons le Bagging dans une version un peu modifiée.
- Le Bagging consiste à entraîner B modèles sur B répliques Bootstrap de la base initiale et de considérer la moyenne des réponses des B modèles comme étant la sortie estimée de la modélisation.

Base D

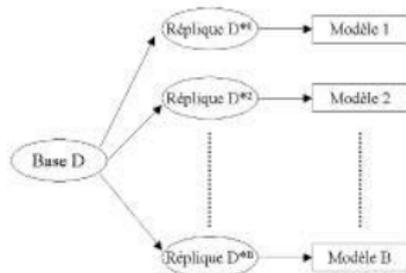
Le Bagging

- Pour faire de la replanification par Bootstrap, nous utilisons le Bagging dans une version un peu modifiée.
- Le Bagging consiste à entraîner B modèles sur B répliques Bootstrap de la base initiale et de considérer la moyenne des réponses des B modèles comme étant la sortie estimée de la modélisation.



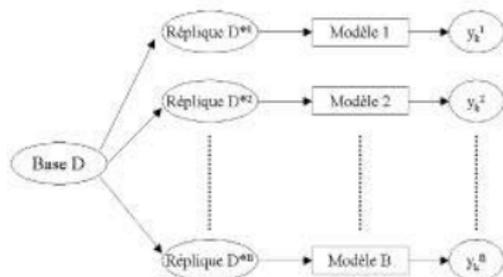
Le Bagging

- Pour faire de la replanification par Bootstrap, nous utilisons le Bagging dans une version un peu modifiée.
- Le Bagging consiste à entraîner B modèles sur B répliques Bootstrap de la base initiale et de considérer la moyenne des réponses des B modèles comme étant la sortie estimée de la modélisation.



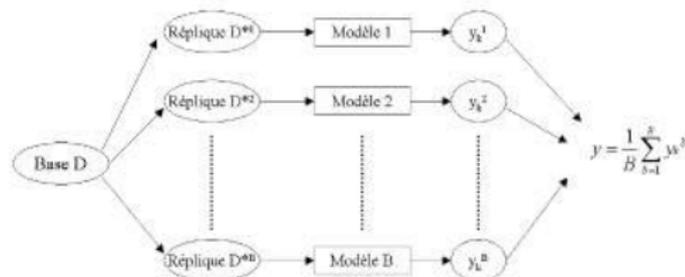
Le Bagging

- Pour faire de la replanification par Bootstrap, nous utilisons le Bagging dans une version un peu modifiée.
- Le Bagging consiste à entraîner B modèles sur B répliques Bootstrap de la base initiale et de considérer la moyenne des réponses des B modèles comme étant la sortie estimée de la modélisation.

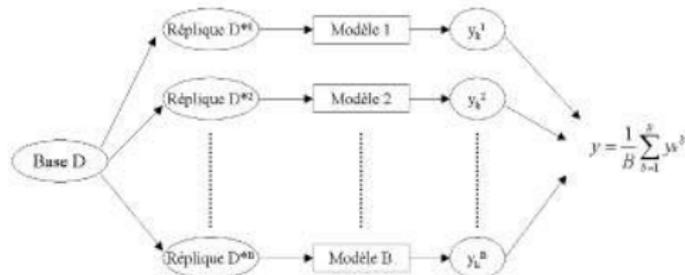


Le Bagging

- Pour faire de la replanification par Bootstrap, nous utilisons le Bagging dans une version un peu modifiée.
- Le Bagging consiste à entraîner B modèles sur B répliques Bootstrap de la base initiale et de considérer la moyenne des réponses des B modèles comme étant la sortie estimée de la modélisation.



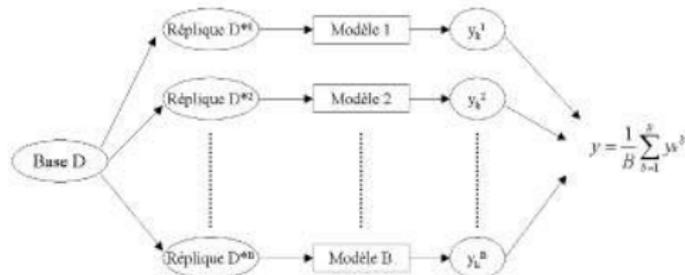
Le Bagging



- $x \in \mathbb{R}^n$ et $y \in \mathbb{R}$

Soit $\mathcal{D}_m = \{(x_i, y_i); i = 1 \dots m\}$ le plan initial de m expériences avec $y_i = P(x_i)$, P étant le processus inconnu.

Le Bagging



- $x \in \mathbb{R}^n$ et $y \in \mathbb{R}$
 Soit $\mathcal{D}_m = \{(x_i, y_i); i = 1 \dots m\}$ le plan initial de m expériences avec $y_i = P(x_i)$, P étant le processus inconnu.
- Soit $\mathcal{D}_m^{*k} = \{(x_i^*, y_i^*); i = 1 \dots m\}$ la réplique bootstrap k de \mathcal{D}_m .
 $\mathcal{D}_m^{*k} \rightsquigarrow f_k$



Le Bagging

- $x \in \mathbb{R}^n$ et $y \in \mathbb{R}$
Soit $\mathcal{D}_m = \{(x_i, y_i); i = 1 \dots m\}$ le plan initial de m expériences avec $y_i = P(x_i)$, P étant le processus inconnu.
- Soit $\mathcal{D}_m^{*k} = \{(x_i^*, y_i^*); i = 1 \dots m\}$ la réplique bootstrap k de \mathcal{D}_m .
 $\mathcal{D}_m^{*k} \rightsquigarrow f_k$
- Le modèle obtenu par Bagging est :

$$\bar{f}(x) = \frac{1}{B} \sum_{i=1}^B f_i(x)$$



La replanification par Bagging

- Pour la replanification Bagging nous n'utilisons pas la moyenne des B modèles issus de répliques bootstrap mais la variance. On replanifie le point de variance prédictive maximale.



La replanification par Bagging

- Pour la replanification Bagging nous n'utilisons pas la moyenne des B modèles issus de répliques bootstrap mais la variance. On replanifie le point de variance prédictive maximale.

$$V_m(x) = \frac{1}{B-1} \sum_{i=1}^B (f_i(x) - \bar{f}(x))^2$$



La replanification par Bagging

- Pour la replanification Bagging nous n'utilisons pas la moyenne des B modèles issus de répliques bootstrap mais la variance. On replanifie le point de variance prédictive maximale.

$$V_m(x) = \frac{1}{B-1} \sum_{i=1}^B (f_i(x) - \bar{f}(x))^2$$

- Replanification :

$$\mathcal{D}_{m+1} = \mathcal{D}_m + \{(x_{m+1}, y_{m+1})\}$$

avec $x_{m+1} = \text{Argmax}(V_m)$ et $y_{m+1} = P(x_{m+1})$.



La replanification par Bagging

- Pour la replanification Bagging nous n'utilisons pas la moyenne des B modèles issus de répliques bootstrap mais la variance. On replanifie le point de variance prédictive maximale.

$$V_m(x) = \frac{1}{B-1} \sum_{i=1}^B (f_i(x) - \bar{f}(x))^2$$

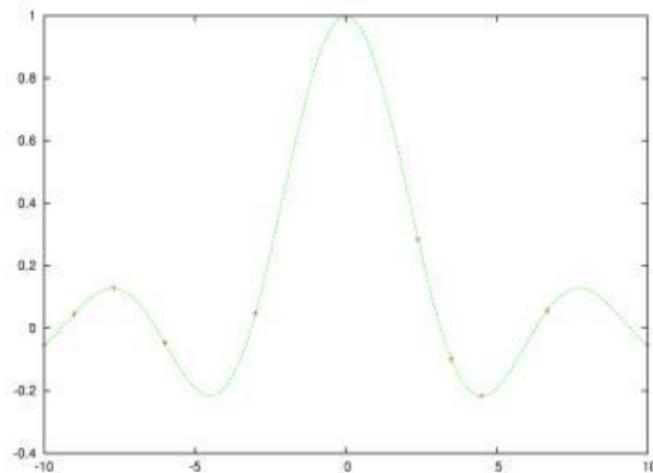
- Replanification :

$$\mathcal{D}_{m+1} = \mathcal{D}_m + \{(x_{m+1}, y_{m+1})\}$$

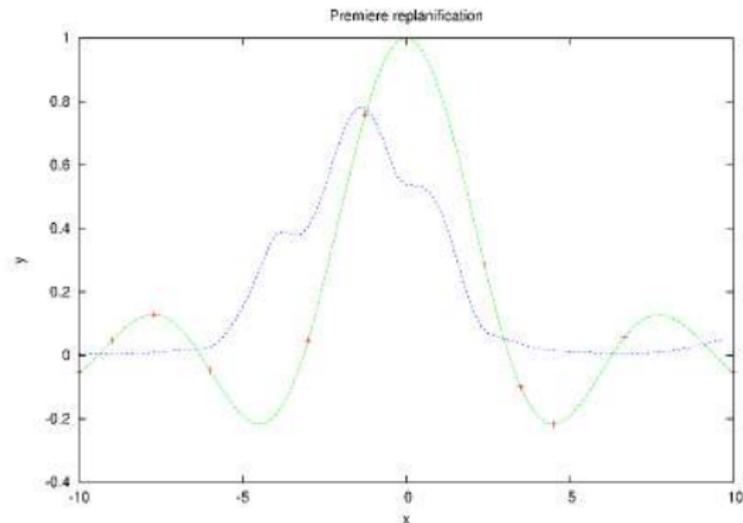
$$\text{avec } x_{m+1} = \text{Argmax}(V_m) \text{ et } y_{m+1} = P(x_{m+1}).$$

Apprentissage *classique* : **Relation Maître-Elève**

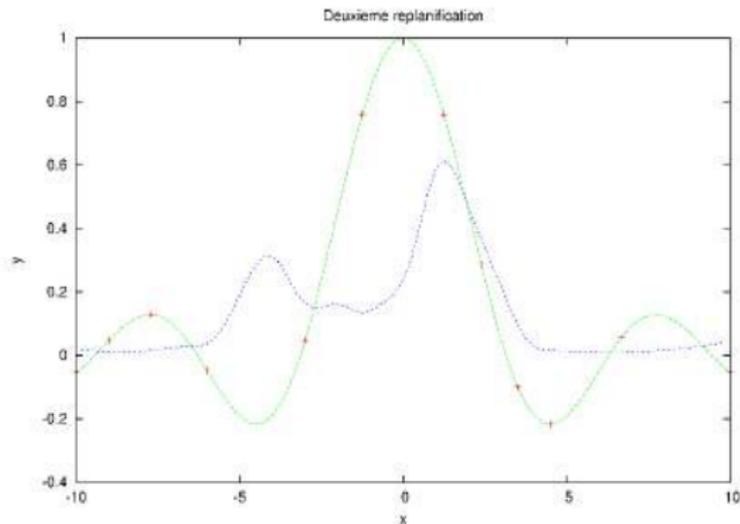
Apprentissage Actif par Bagging : **Relation Maître-Classe**



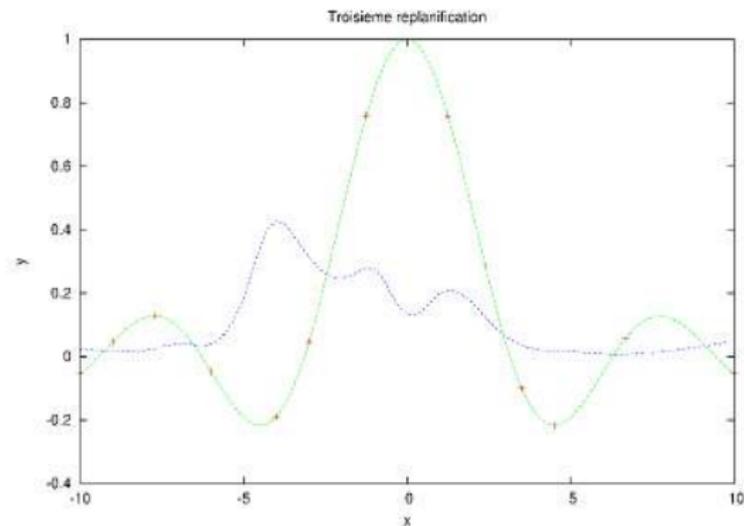
On calcul la variance prédictive sur ce plan d'expériences.

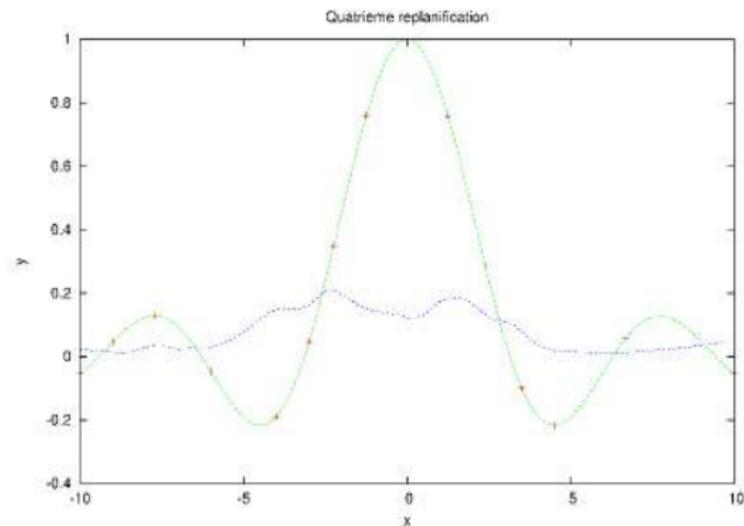


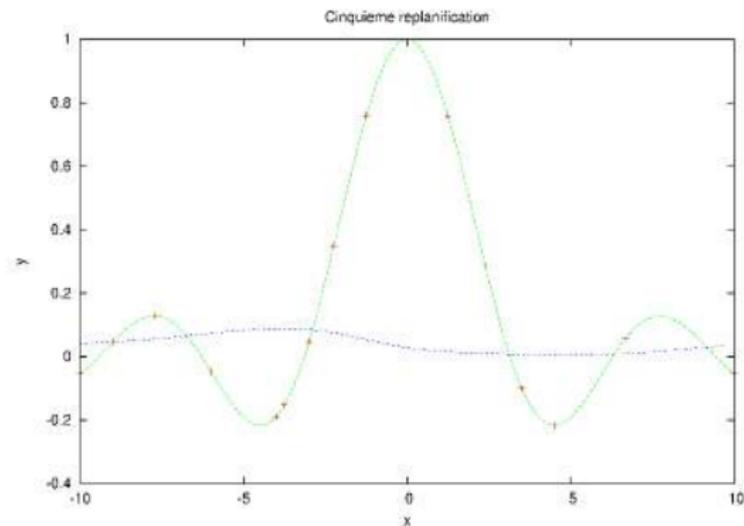
On replanifie le point de variance prédictive maximale.

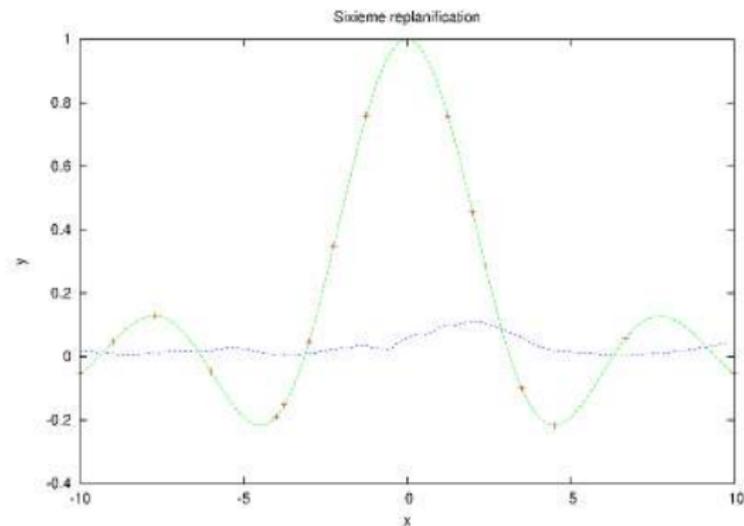


On recommence la procédure.









- 1 Principe des deux méthodes
 - La replanification D-Optimale
 - La replanification par Bagging
- 2 Les Cas test
 - **Friedman**
 - Homma et Saltelli
 - Sélection du modèle pour Homma et Saltelli
- 3 Conclusion



Le modèle de Friedman

$$f(x) = \theta_1 \sin(\pi x_1 x_2) + \theta_2 (x_3 - \theta_3)^2 + \theta_4 x_4 + \theta_5 x_5$$



Le modèle de Friedman

$$f(x) = \theta_1 \sin(\pi x_1 x_2) + \theta_2 (x_3 - \theta_3)^2 + \theta_4 x_4 + \theta_5 x_5$$

avec

$$\theta_1 = 0.4$$

$$\theta_2 = 0.8$$

$$\theta_3 = 0.5$$

$$\theta_4 = 0.4$$

$$\theta_5 = 0.2$$

$$f(x) = \theta_1 \sin(\pi x_1 x_2) + \theta_2 (x_3 - \theta_3)^2 + \theta_4 x_4 + \theta_5 x_5$$

avec

$$\theta_1 = 0.4$$

$$\theta_2 = 0.8$$

$$\theta_3 = 0.5$$

$$\theta_4 = 0.4$$

$$\theta_5 = 0.2$$

$$x_i \in [0; 1]$$

$$\forall i = 1, \dots, 5$$



Le modèle de Friedman

$$f(x) = \theta_1 \sin(\pi x_1 x_2) + \theta_2 (x_3 - \theta_3)^2 + \theta_4 x_4 + \theta_5 x_5$$

avec

$$\begin{array}{ccccc} \theta_1 = 0.4 & \theta_2 = 0.8 & \theta_3 = 0.5 & \theta_4 = 0.4 & \theta_5 = 0.2 \\ x_i \in [0; 1] & \forall i = 1, \dots, 5 & & & \end{array}$$

- Bonne approximation de la fonction f pour un modèle neuronal comportant 5 ou 6 neurones en couche cachée.



Le modèle de Friedman

$$f(x) = \theta_1 \sin(\pi x_1 x_2) + \theta_2 (x_3 - \theta_3)^2 + \theta_4 x_4 + \theta_5 x_5$$

avec

$$\begin{array}{ccccc} \theta_1 = 0.4 & \theta_2 = 0.8 & \theta_3 = 0.5 & \theta_4 = 0.4 & \theta_5 = 0.2 \\ x_i \in [0; 1] & \forall i = 1, \dots, 5 & & & \end{array}$$

- Bonne approximation de la fonction f pour un modèle neuronal comportant 5 ou 6 neurones en couche cachée.
- Nous utiliserons des réseaux de neurones 5_6_1



La méthode de comparaison

- On utilise un plan LHS initial de 100 points commun aux deux stratégies. On décide de replanifier 30 points dans les deux cas.



La méthode de comparaison

- On utilise un plan LHS initial de 100 points commun aux deux stratégies. On décide de replanifier 30 points dans les deux cas.
- On initialise 30 réseaux de neurones 5_6_1 et on choisit le meilleur en apprentissage, pour initialiser la replanification Bagging et l'algorithme de la D-optimalité (*Calcul de la jacobienne*).



La méthode de comparaison

- On utilise un plan LHS initial de 100 points commun aux deux stratégies. On décide de replanifier 30 points dans les deux cas.
- On initialise 30 réseaux de neurones 5_6_1 et on choisit le meilleur en apprentissage, pour initialiser la replanification Bagging et l'algorithme de la D-optimalité (*Calcul de la jacobienne*).
- On estime l'erreur de prédiction des modèles ayant appris sur une base replanifiée par tirage Monte-Carlo. C'est cette estimation que l'on utilisera pour déterminer la qualité de la base replanifiée.

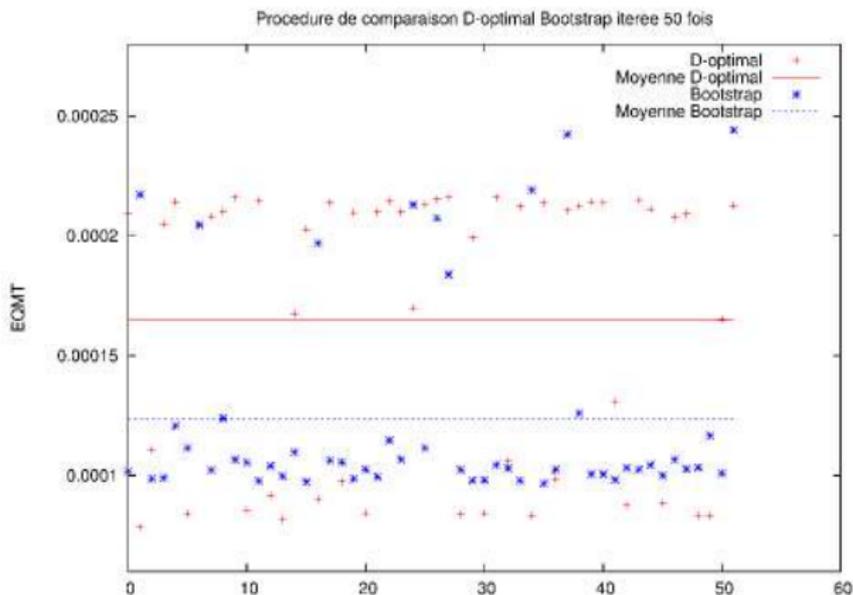


La méthode de comparaison

- On utilise un plan LHS initial de 100 points commun aux deux stratégies. On décide de replanifier 30 points dans les deux cas.
- On initialise 30 réseaux de neurones 5_6_1 et on choisit le meilleur en apprentissage, pour initialiser la replanification Bagging et l'algorithme de la D-optimalité (*Calcul de la jacobienne*).
- On estime l'erreur de prédiction des modèles ayant appris sur une base replanifiée par tirage Monte-Carlo. C'est cette estimation que l'on utilisera pour déterminer la qualité de la base replanifiée.
- Pour éviter de comparer avec un *bon* ou un *mauvais* aléatoire, nous comparerons les deux méthodes à 50 bases différentes replanifiées aléatoirement.

Résultats itérés 50 fois

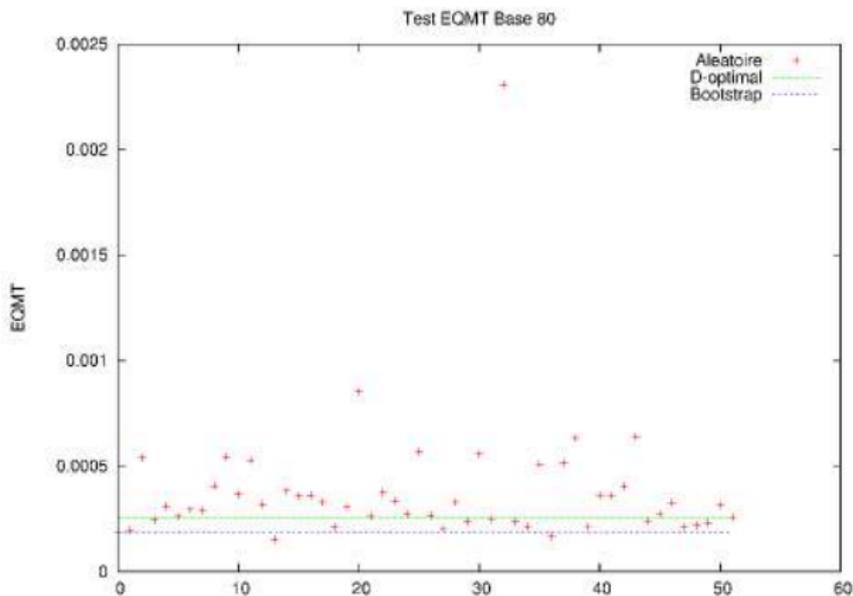
Pour s'affranchir de l'influence des initialisations sur les résultats, on a réitéré toute la procédure 50 fois.





Résultats avec la base initiale de 50 points

Même démarche avec une base contenant moins d'information (50 points)



cea Plan

- 1 Principe des deux méthodes
 - La replanification D-Optimale
 - La replanification par Bagging
- 2 Les Cas test**
 - Friedman
 - Homma et Saltelli**
 - Sélection du modèle pour Homma et Saltelli
- 3 Conclusion



Le modèle de Homma et Saltelli

$$f(x) = \sin(x_1) + \theta_1 \sin^2(x_2) + \theta_2 x_3^4 \sin(x_1)$$

avec $\theta_1 = 7$ et $\theta_2 = 0.1$

$x_i \in [-\pi; \pi] \forall i = 1, \dots, 3$



Le modèle de Homma et Saltelli

$$f(x) = \sin(x_1) + \theta_1 \sin^2(x_2) + \theta_2 x_3^4 \sin(x_1)$$

avec $\theta_1 = 7$ et $\theta_2 = 0.1$

$x_i \in [-\pi; \pi] \forall i = 1, \dots, 3$

- La fonction de Homma et Saltelli est difficile à modéliser. Pour obtenir de bonnes performances en généralisation il faut un nombre élevé de neurones en couche cachée, donc un nombre conséquent d'exemples.



Le modèle de Homma et Saltelli

$$f(x) = \sin(x_1) + \theta_1 \sin^2(x_2) + \theta_2 x_3^4 \sin(x_1)$$

avec $\theta_1 = 7$ et $\theta_2 = 0.1$

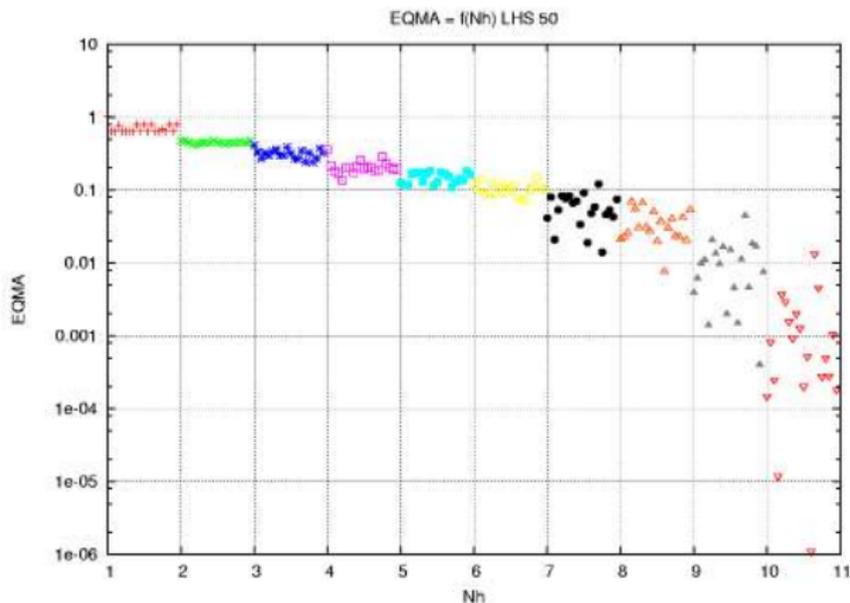
$x_i \in [-\pi; \pi] \forall i = 1, \dots, 3$

- La fonction de Homma et Saltelli est difficile à modéliser. Pour obtenir de bonnes performances en généralisation il faut un nombre élevé de neurones en couche cachée, donc un nombre conséquent d'exemples.
- Nous utiliserons des réseaux de neurones 3_12_1



Choix de l'architecture pour Homma et Saltelli

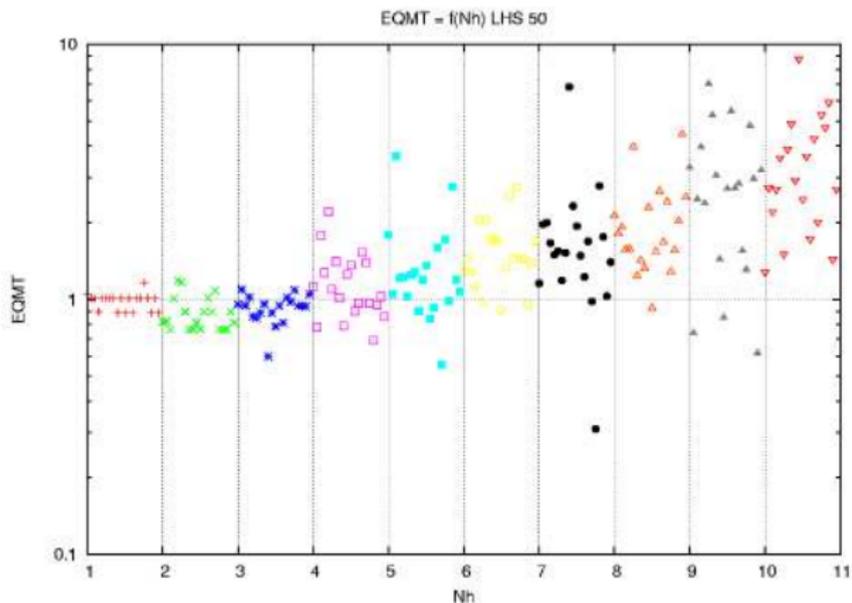
Graphiques reportant l'EQMA et l'EQMT pour différentes architectures et bases d'apprentissage.





Sélection du modèle pour Homma et Saltelli

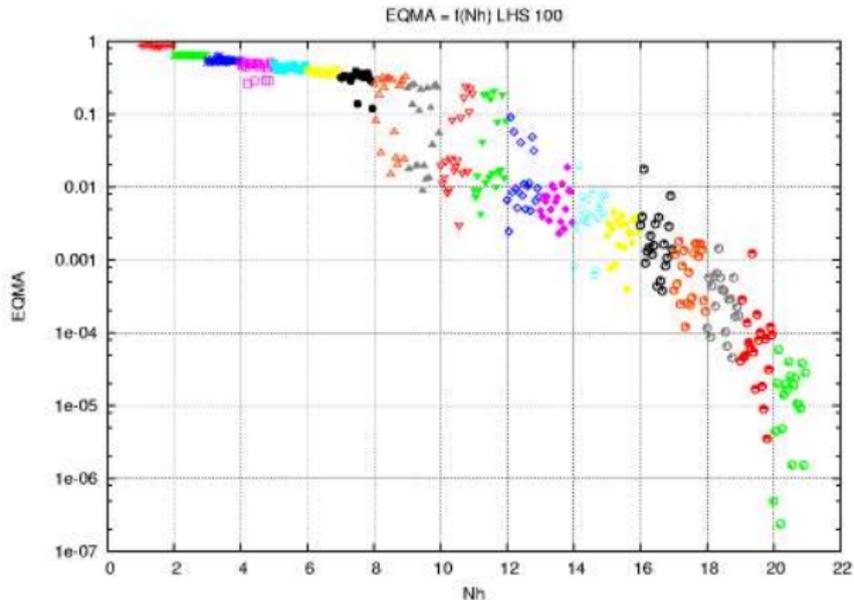
Graphiques reportant l'EQMA et l'EQMT pour différentes architectures et bases d'apprentissage.





Sélection du modèle pour Homma et Saltelli

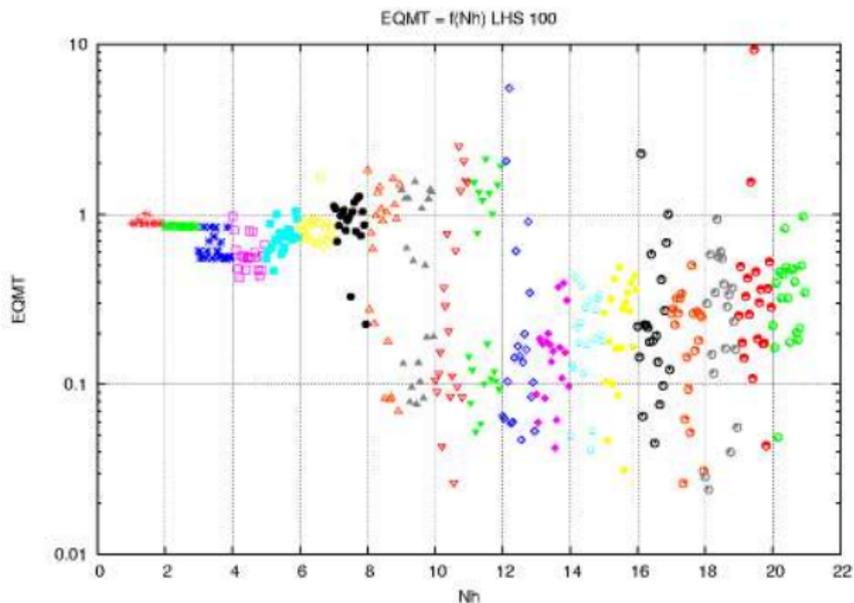
Graphiques reportant l'EQMA et l'EQMT pour différentes architectures et bases d'apprentissage.





Sélection du modèle pour Homma et Saltelli

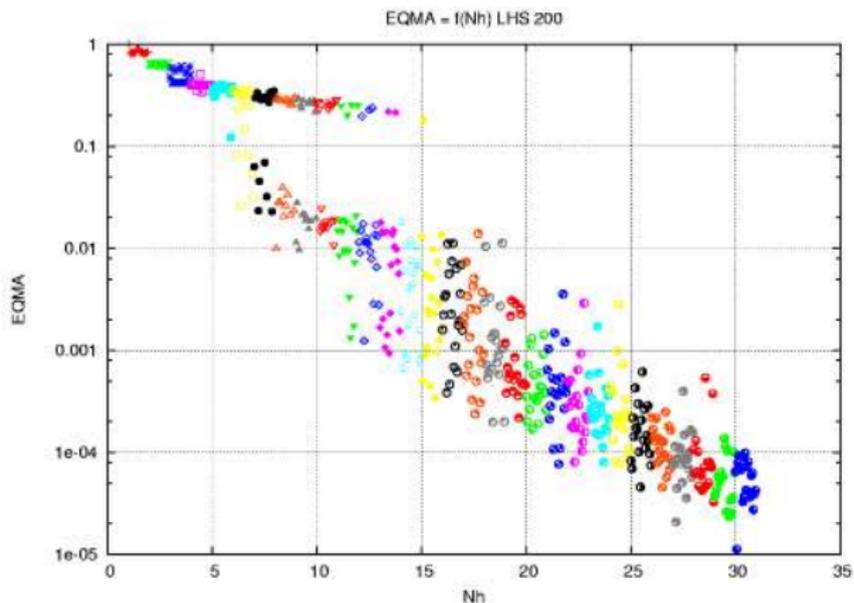
Graphiques reportant l'EQMA et l'EQMT pour différentes architectures et bases d'apprentissage.





Sélection du modèle pour Homma et Saltelli

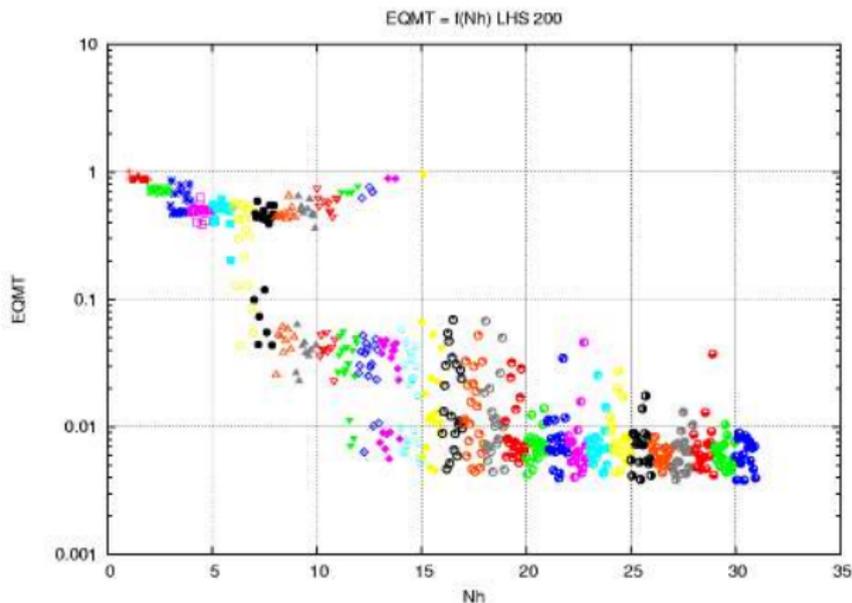
Graphiques reportant l'EQMA et l'EQMT pour différentes architectures et bases d'apprentissage.



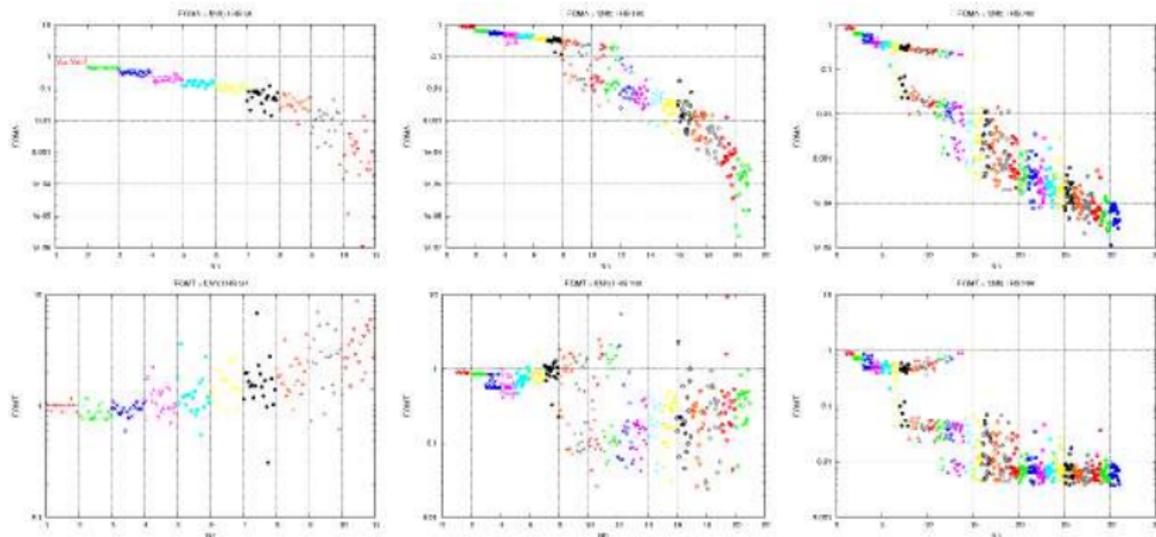


Sélection du modèle pour Homma et Saltelli

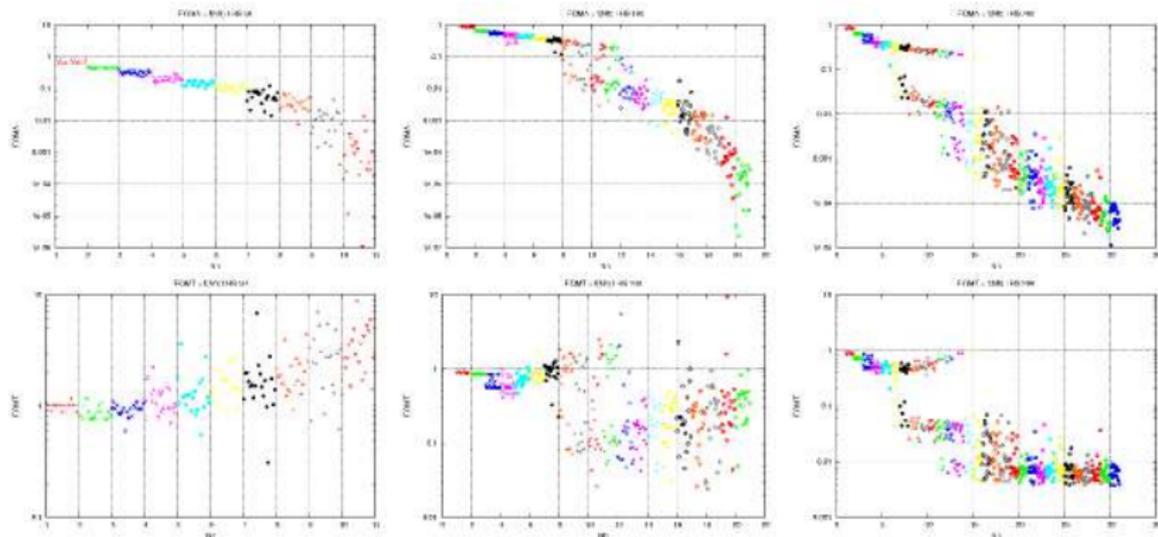
Graphiques reportant l'EQMA et l'EQMT pour différentes architectures et bases d'apprentissage.



Sélection du modèle pour Homma et Saltelli



Sélection du modèle pour Homma et Saltelli



On choisit d'étudier le cas test Homma et Saltelli sur un plan d'expériences initial LHS de 100 points et des réseaux 3_12_1



La méthode de comparaison

- Plan LHS initial \rightarrow 100 points (commun aux deux stratégies).
On décide de replanifier 2 fois 30 points.



La méthode de comparaison

- Plan LHS initial \rightarrow 100 points (commun aux deux stratégies).
On décide de replanifier 2 fois 30 points.
- On initialise 30 réseaux de neurones 3_12_1 et on choisit le meilleur en apprentissage pour l'initialisation des deux méthodes.



La méthode de comparaison

- Plan LHS initial \rightarrow 100 points (commun aux deux stratégies).
On décide de replanifier 2 fois 30 points.
- On initialise 30 réseaux de neurones 3_12_1 et on choisit le meilleur en apprentissage pour l'initialisation des deux méthodes.
- On effectue un tirage Monte-Carlo pour estimer l'erreur de prédiction des modèles obtenus à partir des bases replanifiées.

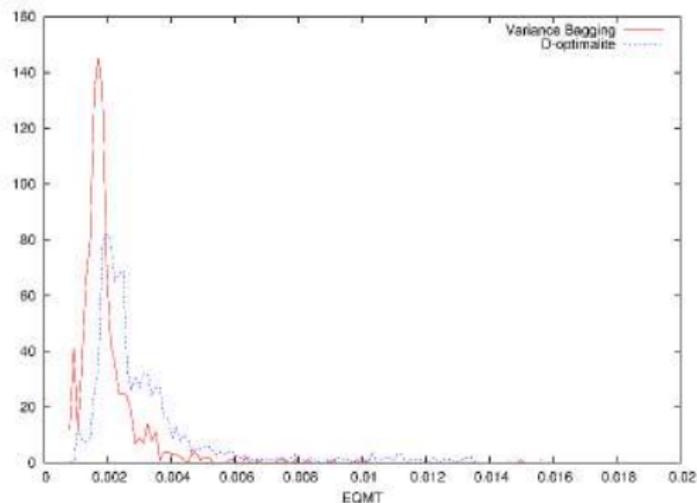


La méthode de comparaison

- Plan LHS initial \rightarrow 100 points (commun aux deux stratégies).
On décide de replanifier 2 fois 30 points.
- On initialise 30 réseaux de neurones 3_12_1 et on choisit le meilleur en apprentissage pour l'initialisation des deux méthodes.
- On effectue un tirage Monte-Carlo pour estimer l'erreur de prédiction des modèles obtenus à partir des bases replanifiées.
- Pour éviter de comparer avec un *bon* ou un *mauvais* aléatoire, nous comparerons les deux méthodes à 50 bases différentes replanifiées aléatoirement.

Procédure itérée 1000 fois

Afin de s'affranchir de l'influence de l'initialisation des modèles



$EQMT \leq 2.10^{-3}$ pour 73 % des modèles "Bagging".
 $EQMT \leq 2.10^{-3}$ pour 26 % des modèles D-optimaux.

Conclusion

- La méthode de réplanification utilisant la variance prédictive Bagging offre de bonnes performances.

Conclusion

- La méthode de réplanification utilisant la variance prédictive Bagging offre de bonnes performances.
- Notre présentation contribue à la passerelle qui pourrait être faite entre le monde des plans d'expériences et celui de l'apprentissage statistique avec l'Active Learning.



Conclusion

- La méthode de réplanification utilisant la variance prédictive Bagging offre de bonnes performances.
- Notre présentation contribue à la passerelle qui pourrait être faite entre le monde des plans d'expériences et celui de l'apprentissage statistique avec l'Active Learning.
- Cependant, il faut être conscient des limites de la méthode, notamment sur le sens à donner à la variance prédictive en présence d'*outlier*.

Conclusion

- La méthode de réplanification utilisant la variance prédictive Bagging offre de bonnes performances.
- Notre présentation contribue à la passerelle qui pourrait être faite entre le monde des plans d'expériences et celui de l'apprentissage statistique avec l'Active Learning.
- Cependant, il faut être conscient des limites de la méthode, notamment sur le sens à donner à la variance prédictive en présence d'*outlier*.
- Nous envisageons alors, en perspectives, de calculer la variance prédictive sur des fractiles de la population, ou d'utiliser la médiane, au lieu de la moyenne, et de calculer la variance prédictive sur une population réduite de modèles centrée sur la médiane.