

Apprentissage machine pour l'amélioration des modèles piloté par la représentation physique des processus

Frédéric Hourdin¹, Fleur Couvreur², Catherine Rio², Arnaud Jam¹, Ionela Musat¹, Romain Roehrig², Daniel Williamson³, Victoria Volodina³

Projet ANR HighTune / projet Dephy

1 : LMD/IPSL/SU/CNRS

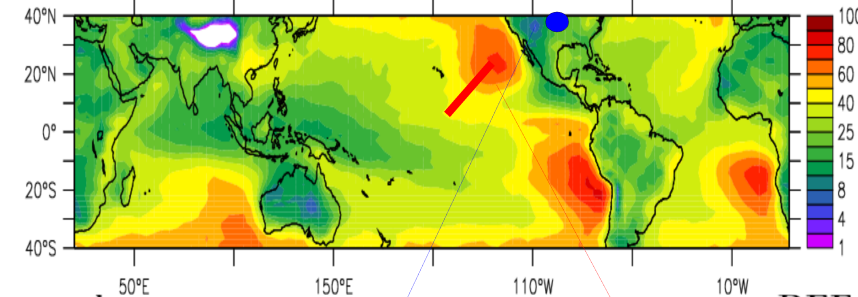
2 : CNRM/MeteoFrance

3 : Extete University

Réunion Pedallons « constantes de temps longues », 4 juin 2020

I. Exemple d'amélioration de LMDZ basée sur une amélioration de la représentation des processus nuageux.

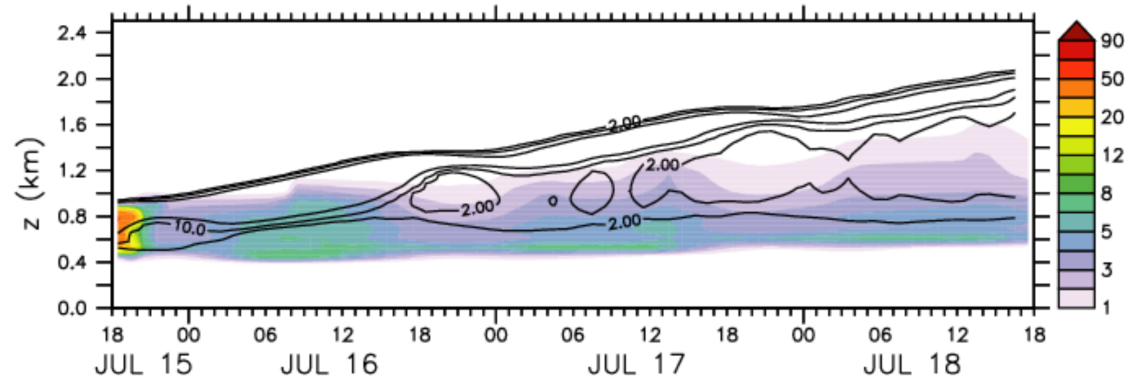
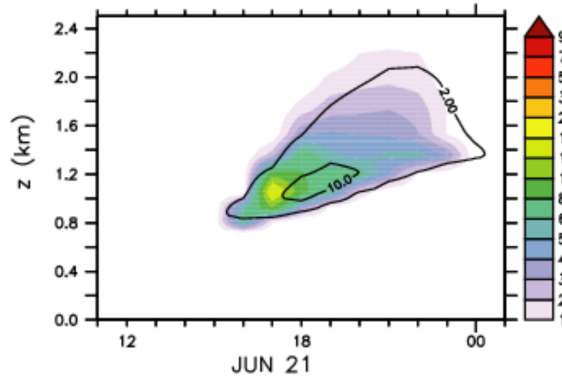
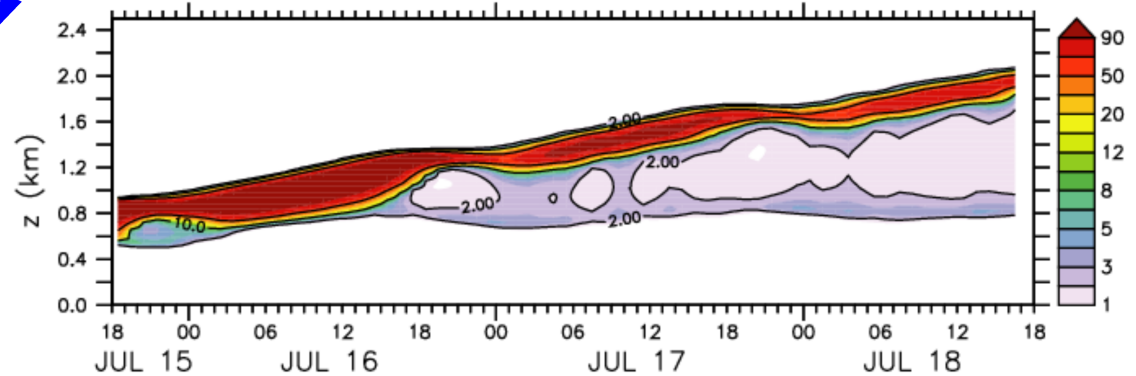
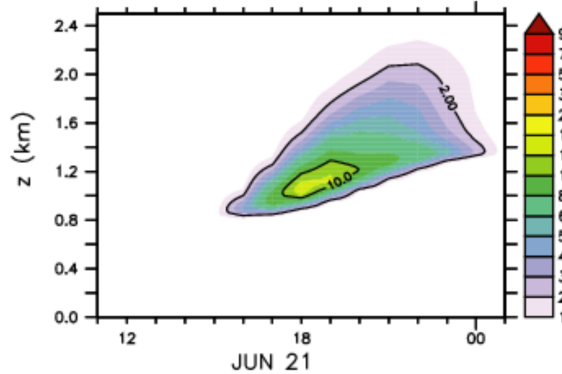
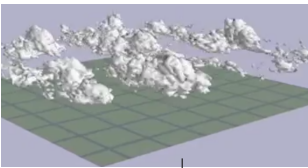
II. Revisite à l'aide d'émulateur (processus gaussien) et de refocalisation itérative (ou history matching) du tuning de LMDZ6A à l'échelle des processus.



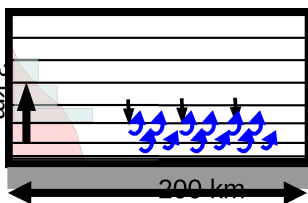
ARM cumulus case

REF transition case

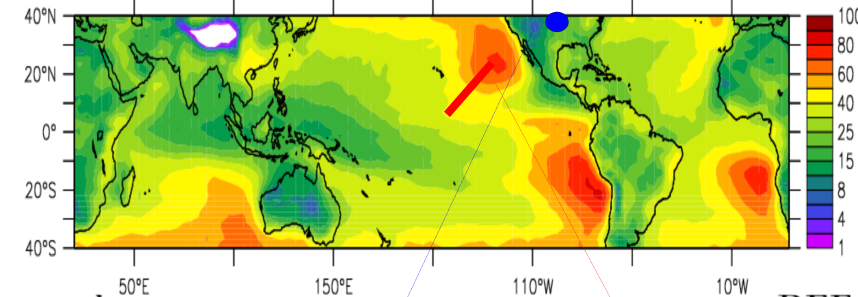
LES



LMDZ/L130



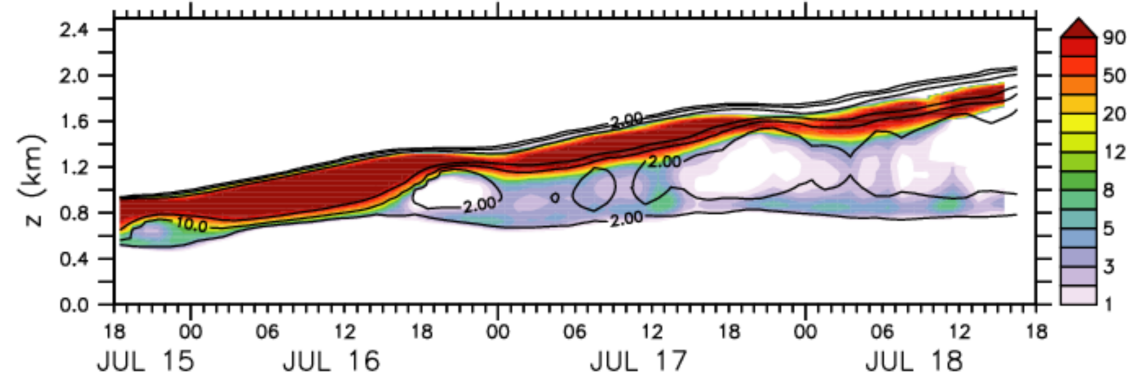
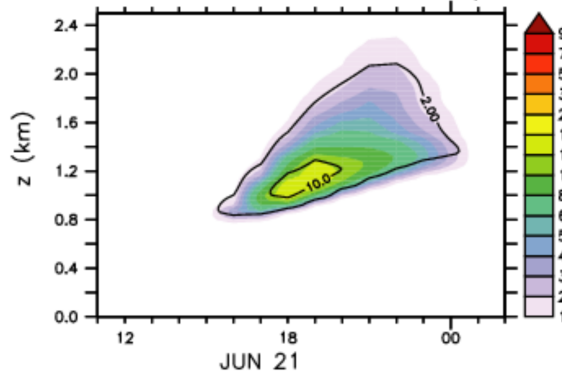
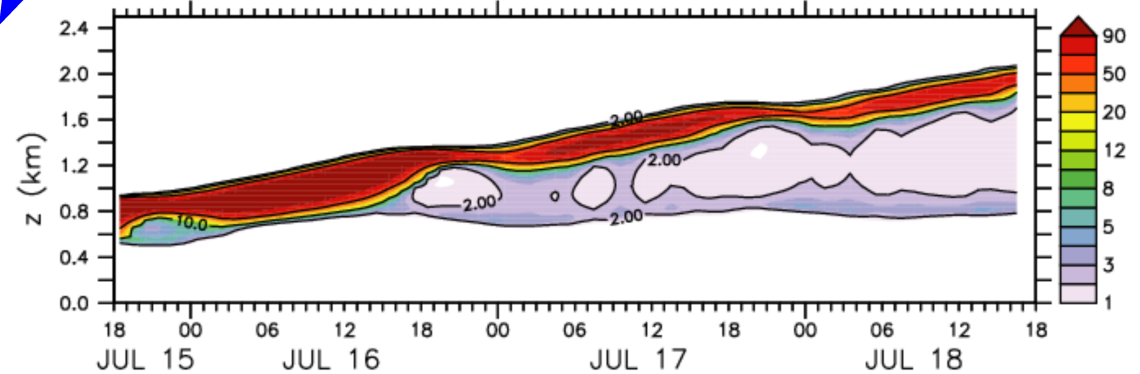
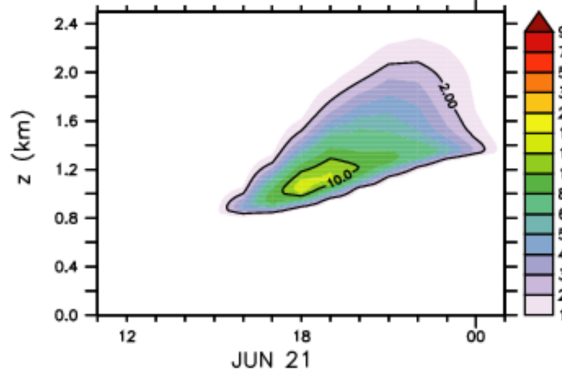
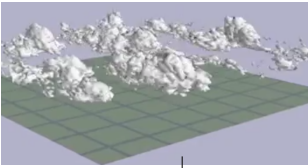
- Comparaison « Single Column Model » (SCM) à des « Large Eddy simulations » (LES)
- Avant modification** du « modèle du thermique » (version 5B sans « bidouille stratocu »)
- **Représentation raisonnable des cumulus continentaux**
 - **Mais pas de stratocumulus ni de transition stratocumulus vers cumulus**



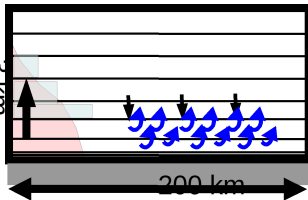
ARM cumulus case

REF transition case

LES



LMDZ/L130



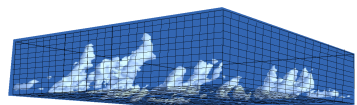
Comparaison « Single Column Model » (SCM) à des « Large Eddy simulations » (LES)

Après modification du « modèle du thermique » (version 6A)

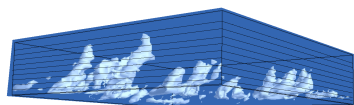
- Représentation raisonnable des cumulus continentaux
- Stratocumulus et transition stratocumulus vers cumulus raisonnables

1/ Développement de nouvelles paramétrisations

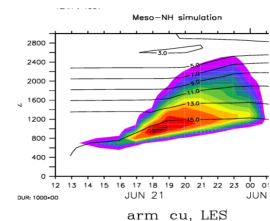
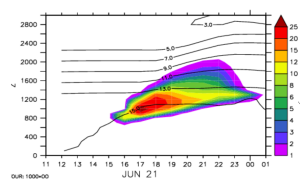
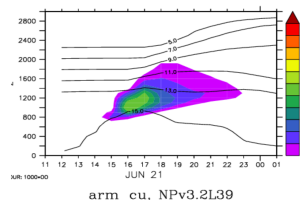
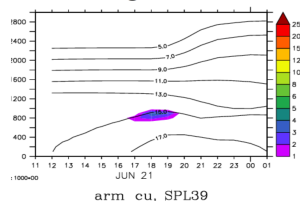
Développement et test en mode uni-colonne



Explicit simulations, $\Delta x \sim 20-100$ m

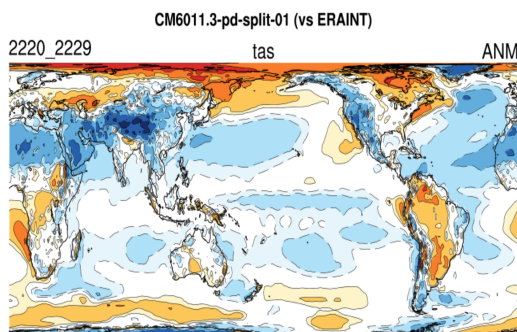
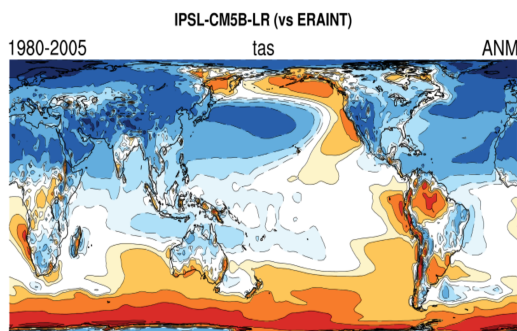
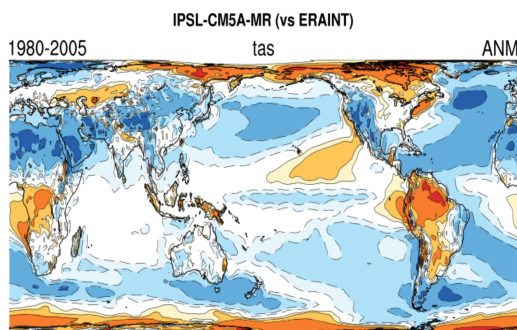


Climate model, parameterizations
« single-column » mode



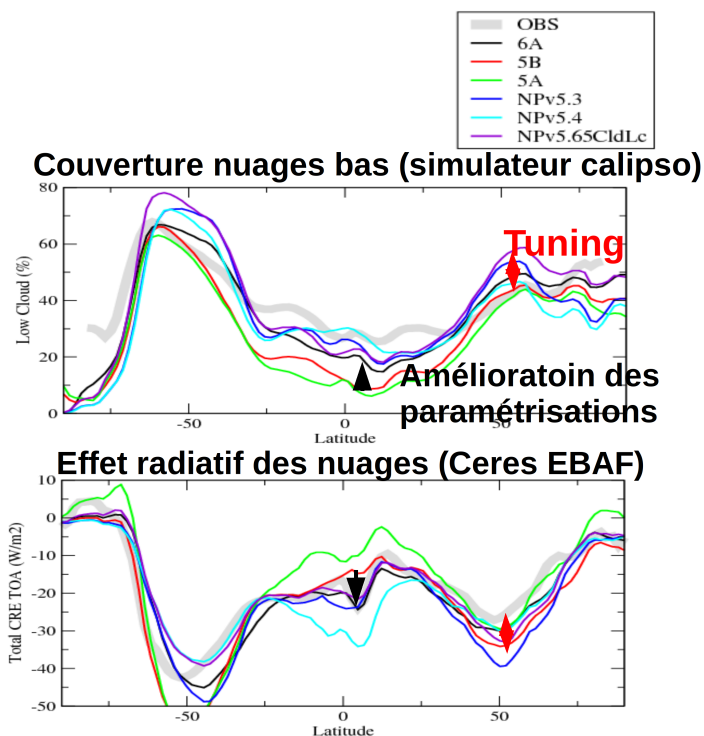
3/ Vérification des améliorations modèle couplé océan/atmosphère

Erreur moyenne sur la température
de surface



2/ Tuning énergétique du modèle 3D (températures océaniques prescrites)

Ajustement des paramètres nuageux pour
améliorer la représentation des bilans
radiatifs.



Tuning atmosphère pour CMIP6
- 738 Mutli-atlas (Ensemble de diagnostics mutualisés)
- 2600 simulations: “forcées” 2-30 ans, “couplées” 50-300 ans.
Beaucoup plus de personnes impliquées que pour CMIP5 = + exigeant et + de travail
2 ans de mobilisation pour l'équipe.

LES

Modèle du thermique modifié pour différentes valeurs du paramètre **DZ**

- Métrique « hauteur nuages » utilisé par la suite (la hauteur du max de la fraction nuageuse)

II : Revisite du tuning de LMDZ6A avec émulateur (processus gaussien) et méta modèle

Choix des paramètres (3, volontairement petit par soucis pédagogique) :

DZ : contrôle du détrainement au sommet des thermiques (thermals_ed_dz)

CLC : seuil d'auto conversion d'eau nuageuse en précipitation (cld_lc_lsc)

EVA : coefficient sur la réévaporation des pluies (coef_evap)

Choix des métriques (5, expertise et tests)

SANDU : (transition strato-cu cumulus) hauteur max des nuages, heure 55-60

ARMCU (cumulus continental): θ et q_v à 500 hPa et couverture nuageuse, heure: 7-9

RICO (cumulus d'alizés) : couverture nuageuse, heure 7-9

$$\int f^4 z dz / \int f^4 dz$$

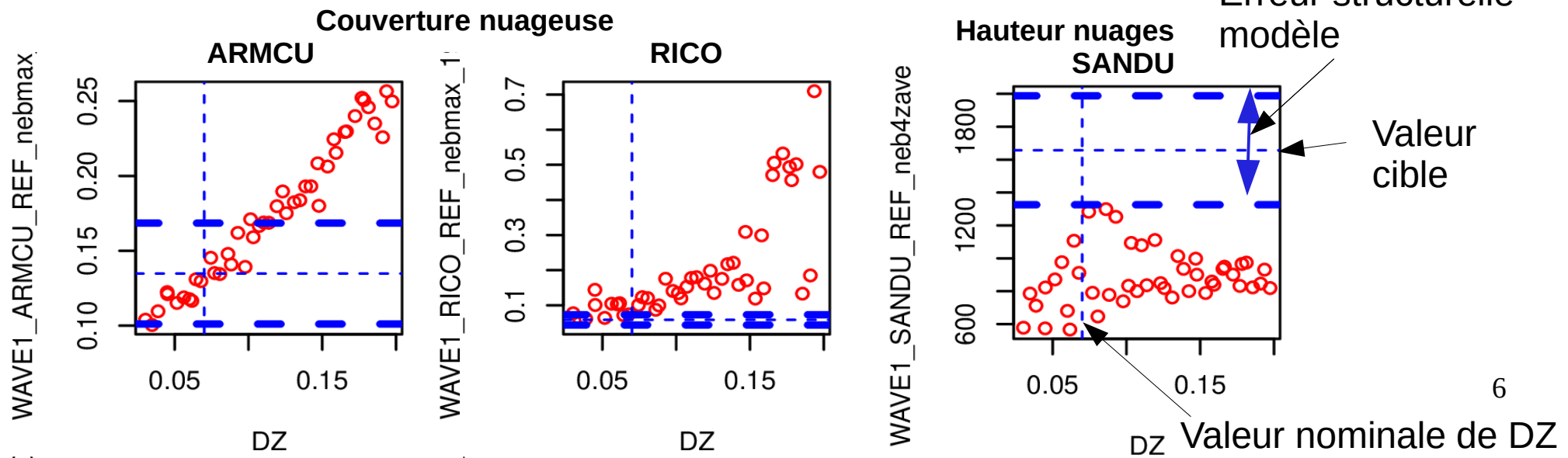
Réalisation d'une série d'expériences avec le modèle (ici 1D)

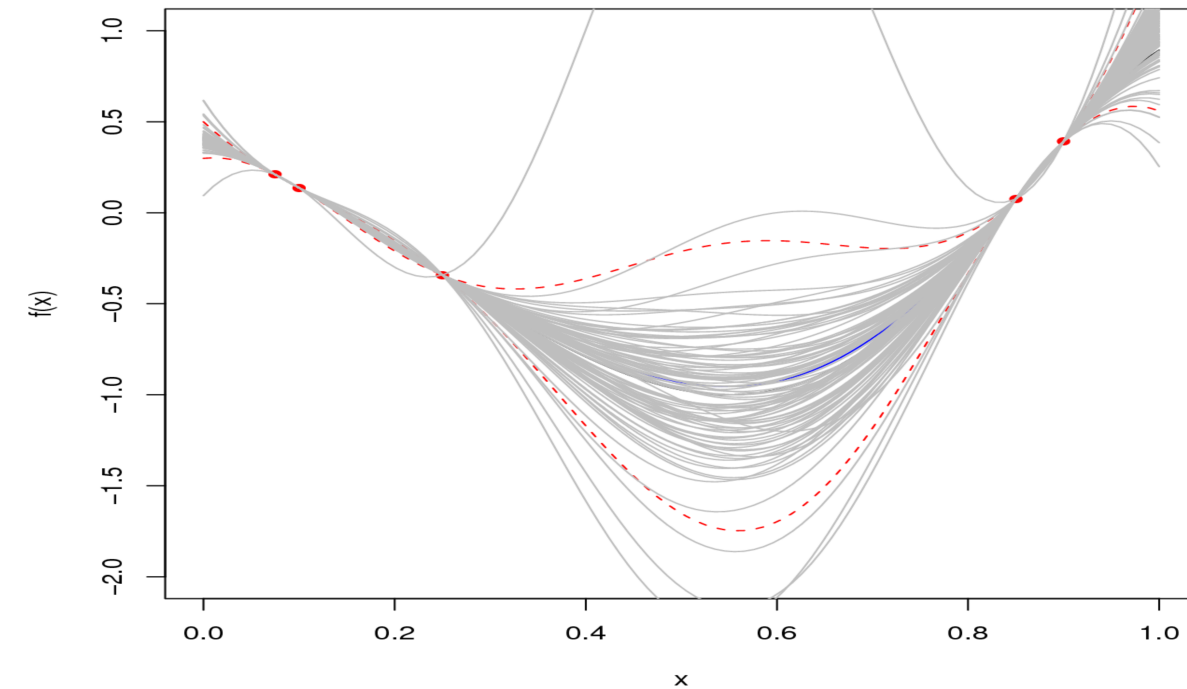
Dans le cube $[Dz_{\min}, Dz_{\max}] \times [CLC_{\min}, CLC_{\max}] \times [EVA_{\min}, EVA_{\max}]$

échantillonnage par hypercube latin (LHS) ; au moins $10 \times N$ où N est le nombre de paramètres.

Illustration :

3 métriques en fonction de la valeur de DZ pour 45 simulations





Construction d'un meta-modèle :

Modèle statistique ultra-rapide donnant la valeur d'une des 5 métriques en tout point de l'espace des paramètres comme l'espérance d'une variable aléatoire, dont les paramètres sont appris sur les 45 expériences réalisées.

Ici, meta-modèle=

Somme de fonctions polynomiales + sinusoïdales + **processus gaussien**

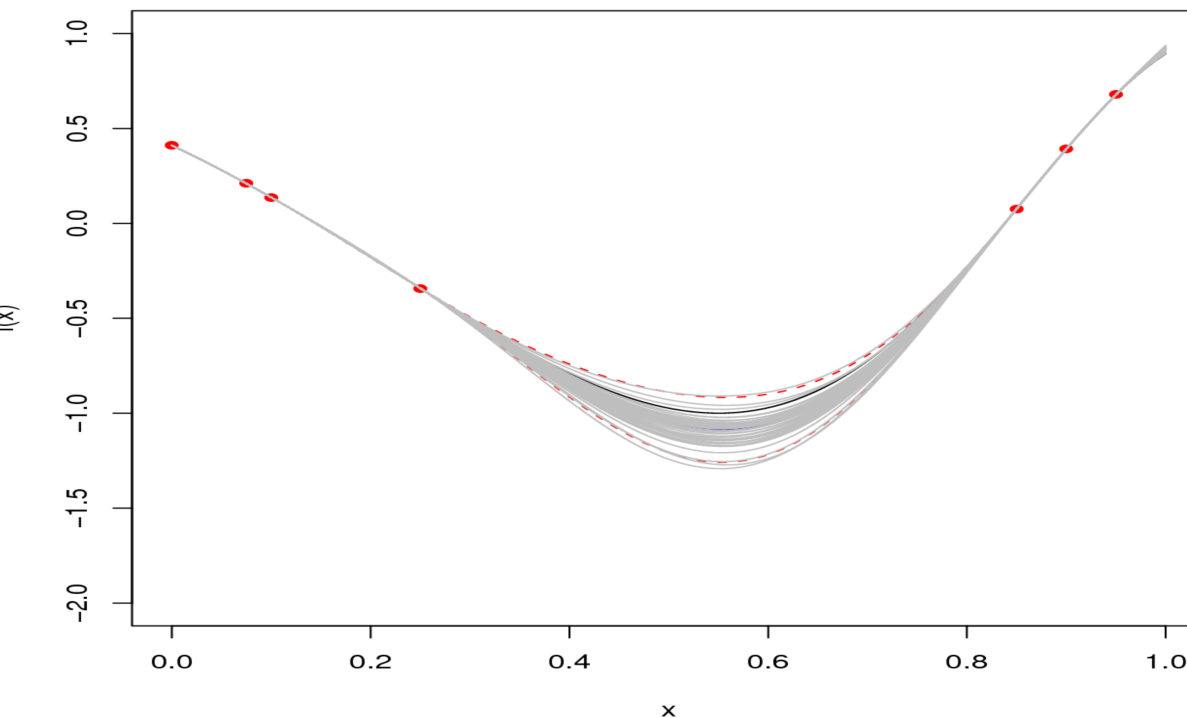
Le processus gaussien est basé sur une idée de longueur de corrélation.

Expert : on suppose ce processus stationnaire dans l'espace des paramètres, c'est à dire que les propriétés statistiques sont les mêmes partout.

Propriétés :

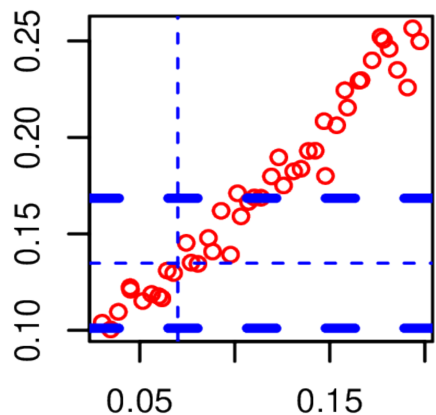
Plus on est loin des point sondés par les expériences, plus on est incertains.

Plus on ajoute d'expériences, plus l'incertitude décroît.

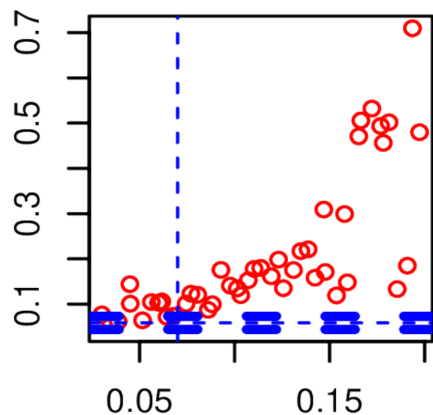


Couverture nuageuse

ARMCU

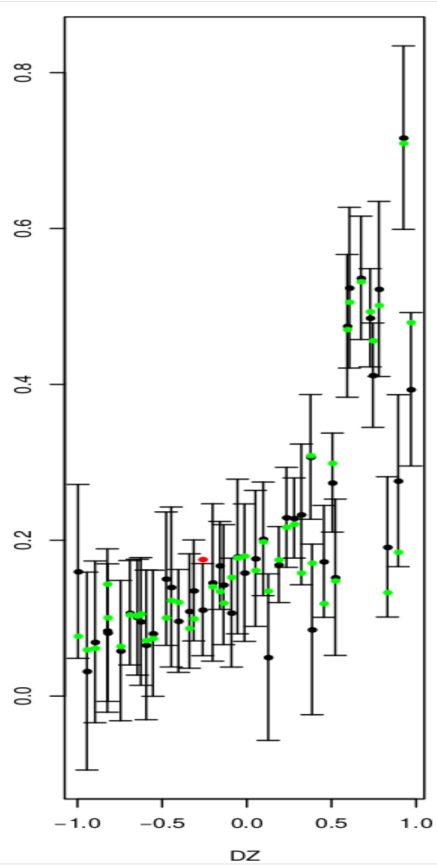
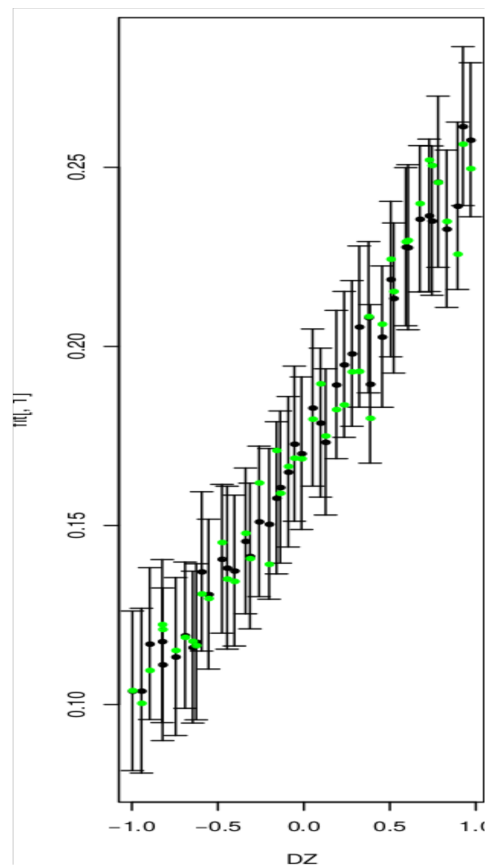


RICO

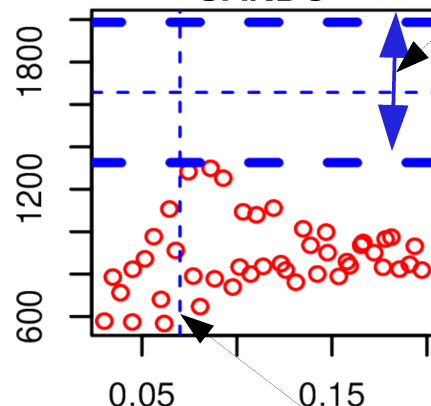


DZ

DZ



Hauteur nuages SANDU



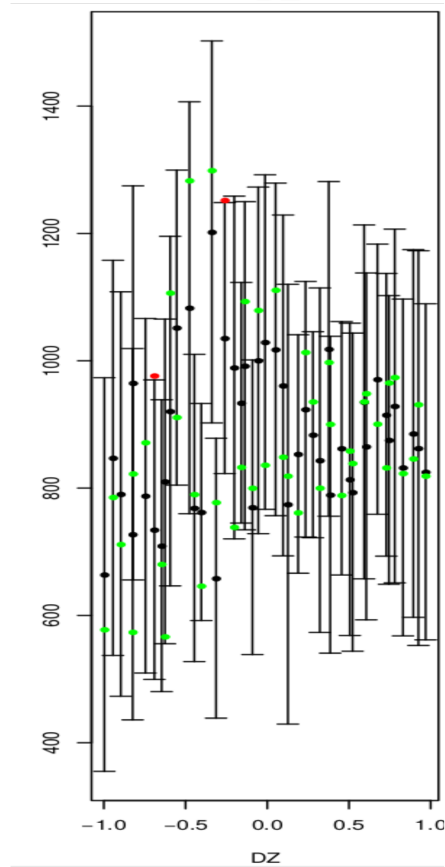
Tolérance à l'erreur ($\pm 2 \sigma$)
Erreur référence (LES) +
Erreur structurelle modèle

Valeur
cible

Expériences SCM

Valeur nominale de DZ

DZ



Graphiques LOO (Leave One Out)

Emulateur :
● Espérance $\pm 2 \sigma$
Supposant une
expérience
manquante

Expérience :
● intérieur $\pm 2 \sigma$
● extérieur $\pm 2 \sigma$

Ref ou valeur cible de la métrique

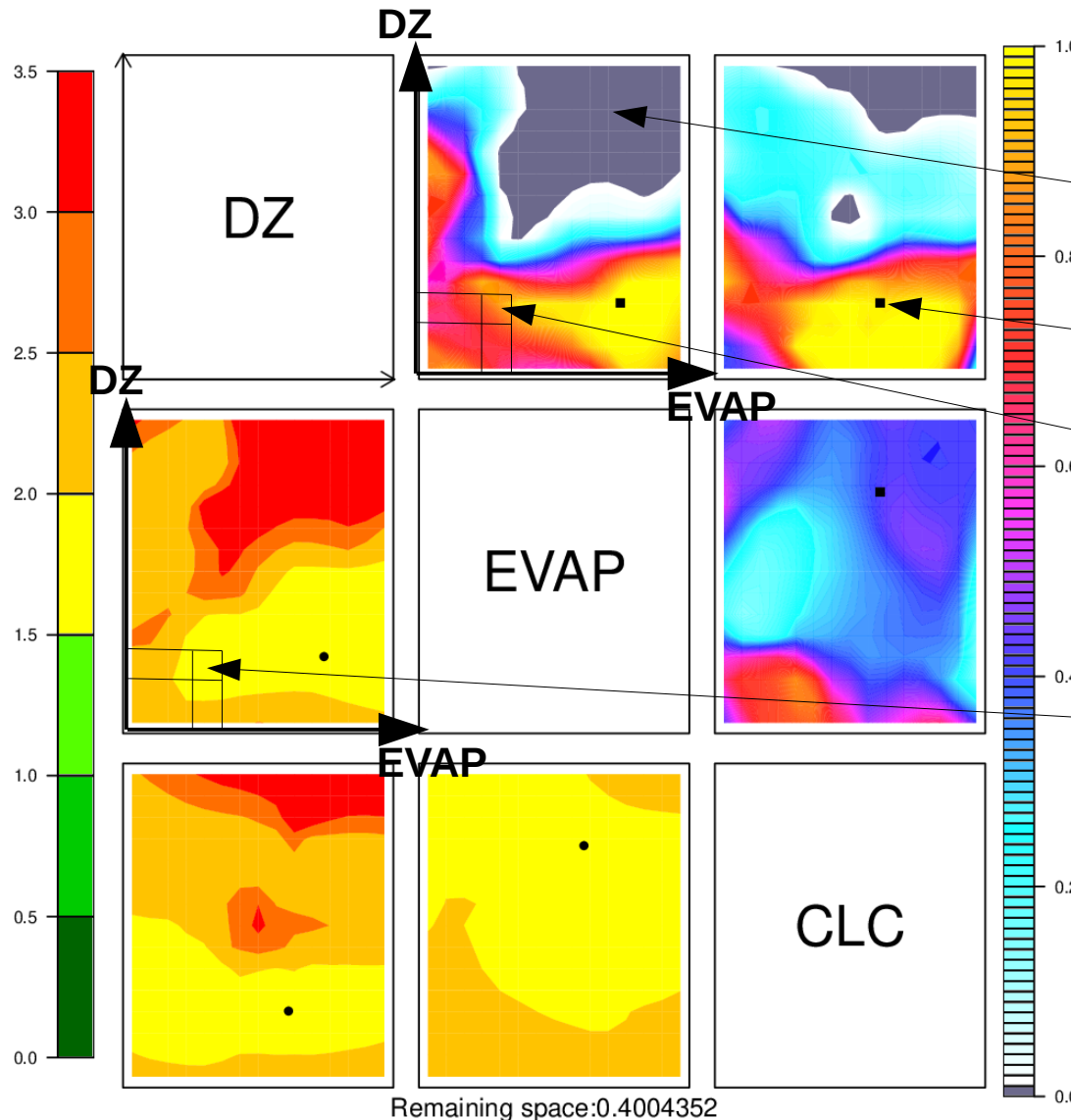
Espérance de l'émulateur

$$\text{Implausibilité : } I_f(\lambda) = \frac{|r_f - E[e(\lambda)]|}{\sqrt{\underbrace{\sigma_{r,f}^2 + \sigma_{d,f}^2}_{\text{Tolérance à l'erreur}} + \text{Var}[e(\lambda)]}}$$

Variance de l'émulateur

Vecteur des paramètres

Tolérance à l'erreur
Référence + structurelle



Remaining space:0.4004352

Valeur de DZ et EVAP rejeté
quelque soit la valeur de CLC

Valeur nominale des paramètres

Fraction de l'espace avec une
implausibilité < 3 pour les valeurs
de DZ et EVAP fixées

Implausibilité minimale
rencontrée pour les valeurs de
DZ et EVAP fixées

→ On ne retient dans l'espace « Not Ruled Out Yet » que les points avec Implausibilité < 3
Ici 40 % de l'espace retenu à la 1ère vague.

→ On refait des expérience pour ~10 N jeux de paramètres dans le NROY → vague 2
En ajoutant ou non des métriques

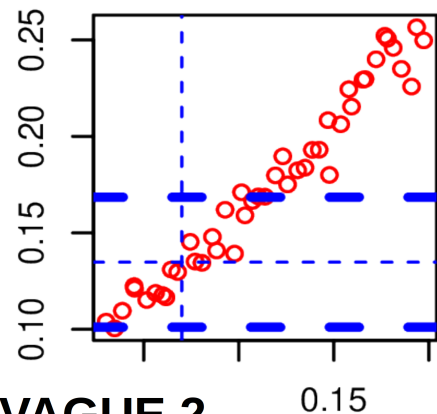
Tolérance à l'erreur ($\pm 2 \sigma$)
 Erreur référence (LES) +
 Erreur structurale modèle

Valeur
cible

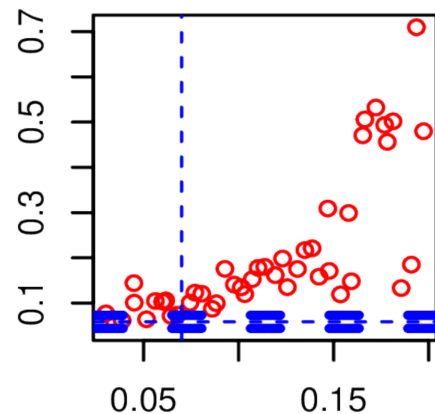
Expériences SCM

Valeur nominale de DZ

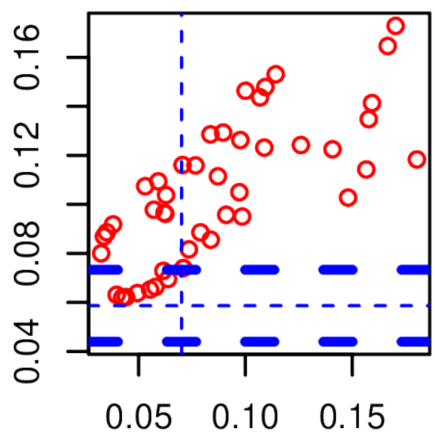
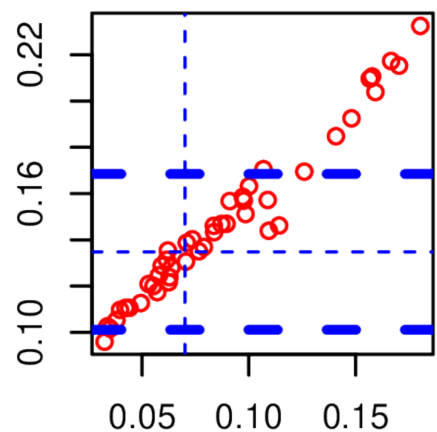
VAGUE 1 Couverture nuageuse
 ARM CU



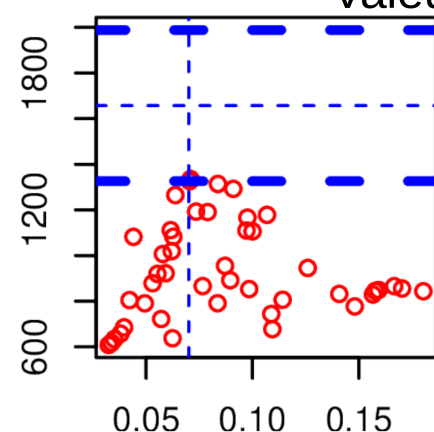
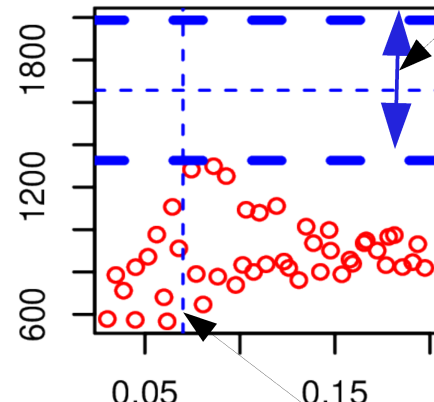
RICO



VAGUE 2



SANDU



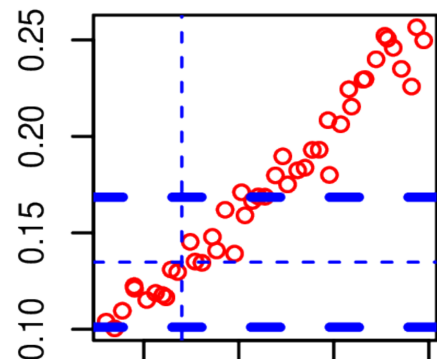
Tolérance à l'erreur ($\pm 2 \sigma$)
 Erreur référence (LES) +
 Erreur structurale modèle

Valeur
cible

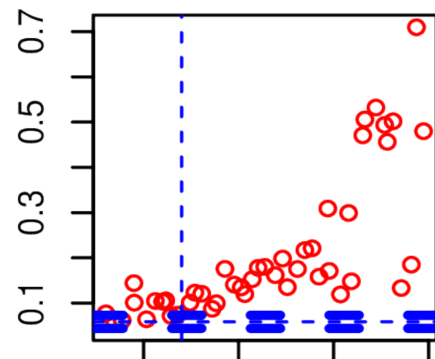
Expériences SCM

Valeur nominale de DZ

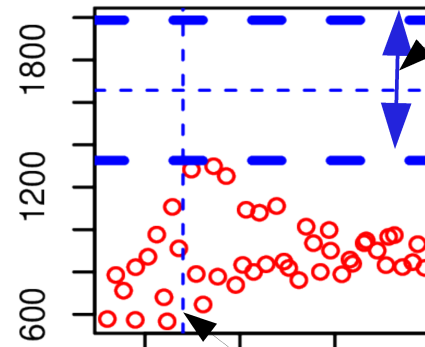
VAGUE 1 Couverture nuageuse
 ARM CU



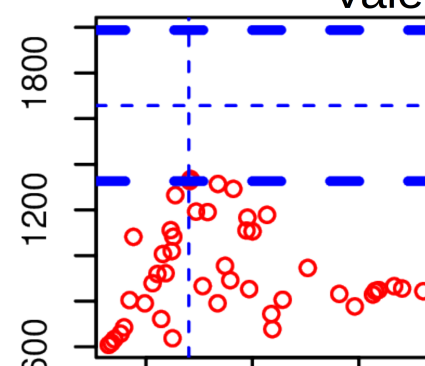
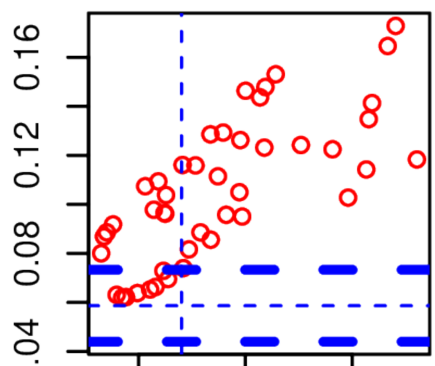
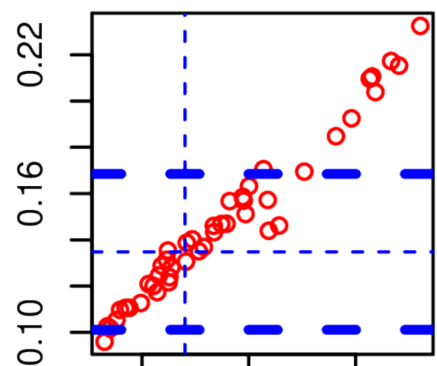
RICO



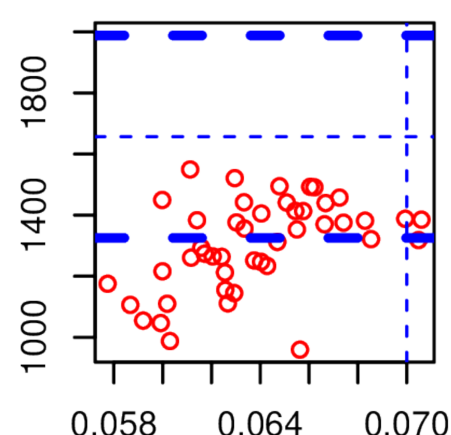
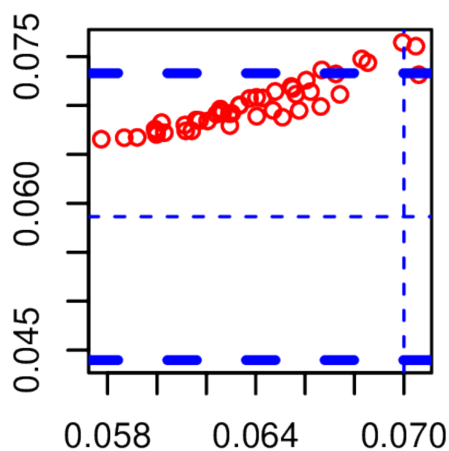
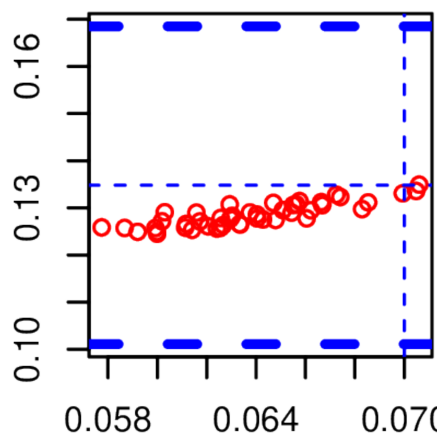
Hauteur nuages
 SANDU

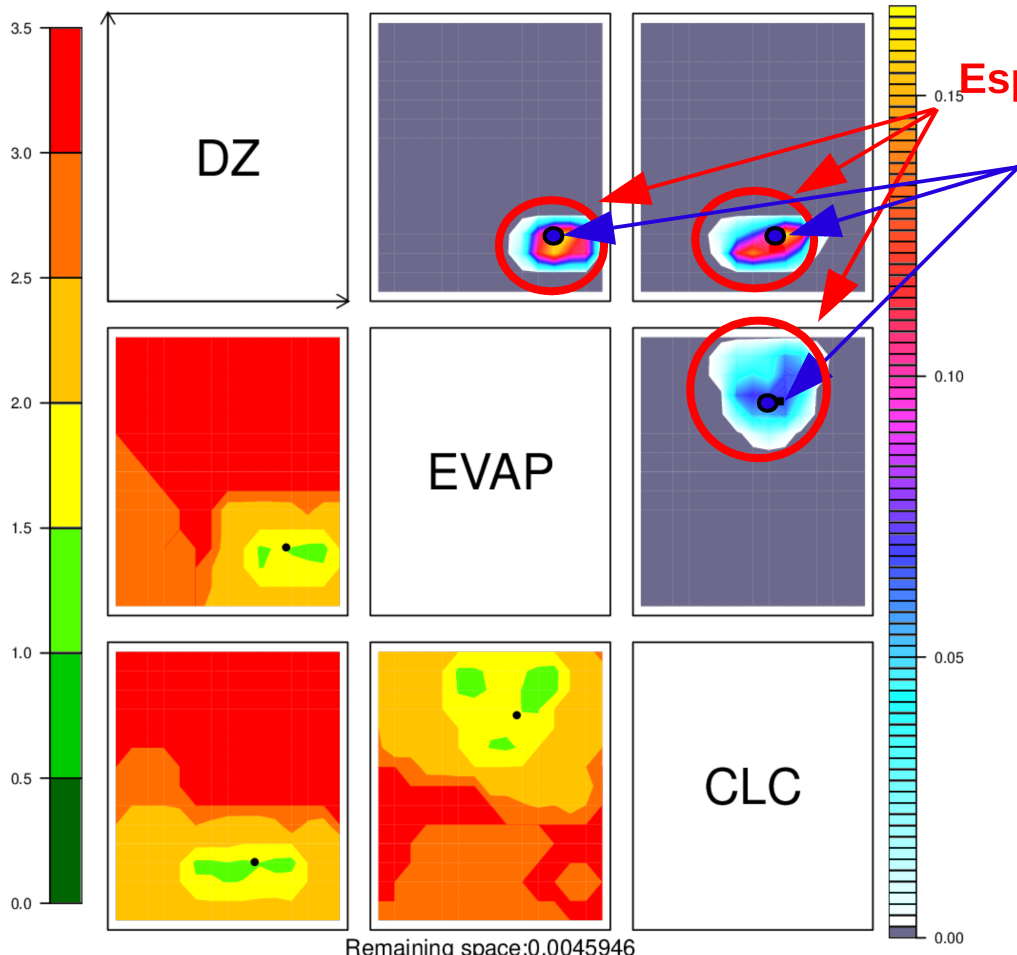


VAGUE 2 0.15



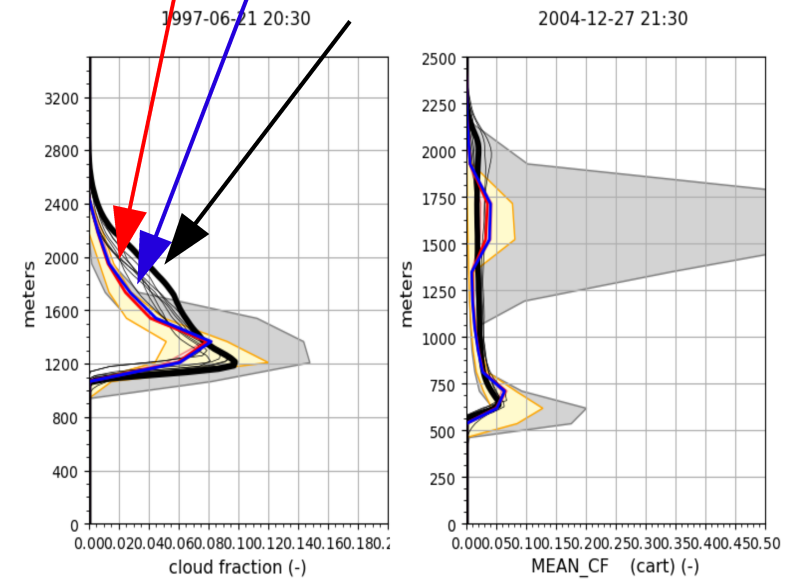
VAGUE 20 0 0.15





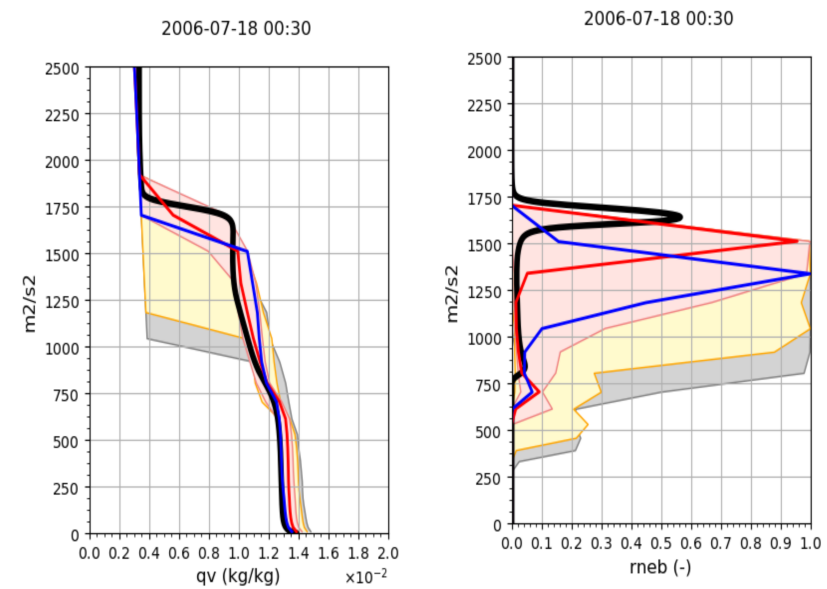
Espace non éliminée après 20 vagues successives
Valeur retenue pour la version LMDZ6A

Référence LES



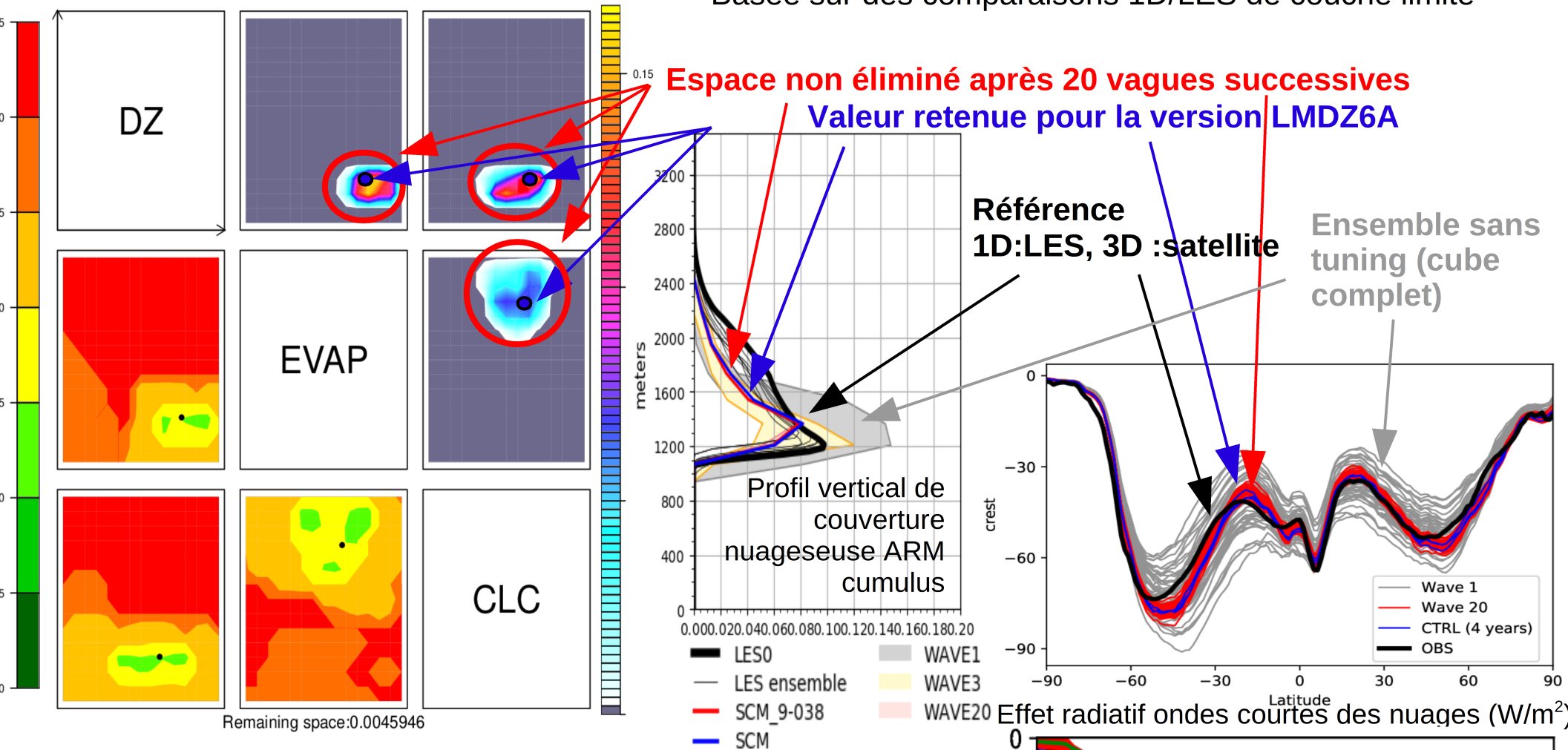
- LES0
- LES ensemble
- SCM_9-038
- SCM
- WAVE1
- WAVE3
- WAVE20

- LES0
- LES ensemble
- SCM_9-038
- SCM
- WAVE1
- WAVE3
- WAVE20



Revisite du tuning manuel de LMDZ6A :

Exemple de réduction de l'espace des paramètres
Basée sur des comparaisons 1D/LES de couche limite



Conclusion : Le tuning SCM LES permet

- 1/ de choisir des paramètres garantissant une bonne simulation des scènes nuageuses
- 2/ d'ajuster en partie les effets radiatifs des nuages.
- 3/ Peut être utilisé pour pré-conditionner un tuning 3D

→ Exposé Sébastien