

Couplage efficace entre Optimisation et Simulation stochastique

Application à la maintenance optimale d'une constellation de satellites

Benoît Beghin

Pierre Baqué

André Cabarbaye

Centre National d'Etudes Spatiales (CNES)

18 avenue Edouard Belin

31401 Toulouse Cedex 4

Tél. 05 61 28 20 44

Mél : benoit.beghin@cnes.fr

CAB INNOVATION

3, rue de la Coquille

31500 Toulouse

Tél. 05 61 54 68 08

Mél : Andre.Cabarbaye@cabinnovation.fr

Web : www.cabinnovation.fr

Résumé

Dans le cadre de recherches sur l'optimisation des systèmes vis-à-vis de la Sûreté de Fonctionnement, cet article présente une technique originale de couplage entre optimisation (Algorithmes Génétiques et simplexe non linéaire) et simulation de Monte-Carlo qui tente de limiter la durée globale des traitements. Fondés sur une application réelle du domaine spatial relative au déploiement et à la maintenance optimale d'une constellation de satellites d'observation de la terre, les résultats obtenus sont encourageants (diminution des temps de calcul dans un rapport 5 environ).

Summary

Within the framework of research on the optimisation of the systems with respect to the Reliability, this article presents an original technique of coupling between optimisation (Genetic Algorithms and non-linear simplex) and simulation of Monte-Carlo which tries to limit the total duration of the treatments. Based on a real application of the space field relating to the optimal deployment and maintenance of a constellation of satellites, the results obtained are encouraging (reduction in the computing times in a ratio of approximately 5).

1. INTRODUCTION

L'évolution rapide des performances d'ordinateur a permis d'étendre sensiblement le champ couvert par les techniques de simulation de type Monte-Carlo et de les appliquer à des problématiques complexes, associant phénomènes déterministes et aléatoires, telles que celles relatives à la fiabilisation et la maintenance des systèmes.

Toutefois, ces techniques présentent des limitations, parfois rédhibitoires, concernant la durée de calcul et la précision atteinte. Définissant les bornes de l'intervalle de confiance en fonction du nombre de simulations effectuées, le théorème Central Limite se révèle en effet incontournable. Un rapport d'un millier environ diffère, et séparera toujours, les durées de traitement d'une même évaluation de fiabilité ou de disponibilité (d'au moins trois décimales de précision) réalisée par simulation de Monte-Carlo et par traitement analytique ou markovien. Pour ce type de problématique, les techniques de réduction de variance, utilisées notamment en fiabilité mécanique, sont en général d'un faible apport car elles supposent la connaissance a priori d'un domaine d'intérêt privilégié.

Par ailleurs, les techniques d'optimisation paramétrique de type stochastique (Algorithmes Génétiques, Recuit simulé, méthode Tabout...) offrent des possibilités intéressantes en recherche optimale car elles ne se fondent pas sur une expression mathématique de la fonction à traiter mais seulement sur son résultat. Cependant, ces techniques se révèlent particulièrement gourmandes en nombre d'évaluations à réaliser pour obtenir un résultat souvent pertinent mais dont l'optimalité ne peut être démontrée.

Bien que très prometteur, le couplage entre optimisation et simulation stochastique, qui consiste à rechercher une configuration optimale de paramètres d'un système à partir des résultats d'une fonction d'évaluation traitée par simulation de Monte-Carlo, s'avère donc très pénalisant en termes de durée de traitement. En première approximation, le nombre de cas de simulation à réaliser est égal au nombre d'évaluations nécessaires à l'optimisation multiplié par le nombre de cas de simulation requise par la précision recherchée.

Mais cette durée peut être sensiblement diminuée par le choix d'une stratégie pertinente consistant à faire varier

judicieusement, au cours du traitement, le nombre de simulations de chaque évaluation.

L'objet de cette communication est de montrer l'apport d'un tel couplage sur un cas d'application réel relatif au déploiement et au renouvellement d'une constellation de satellites d'observation de la terre.

Pour évaluer cet apport, les outils génériques de simulation SIMCAB [9] (Monte-Carlo) et d'optimisation GENCAB [7] (Algorithmes Génétiques couplés avec un simplexe non linéaire) de la société CAB INNOVATION ont été utilisés.

2. Le cas d'application

Pour étudier au mieux le couplage entre optimisation et simulation, l'application choisie doit être suffisamment riche pour aboutir à des résultats probants mais cependant pas trop complexe pour ne pas se tromper de problématique en s'approchant des limites propres aux techniques d'optimisation (on ne cherche pas ici à tester les Algorithmes Génétiques).

Aussi l'application retenue ne comporte qu'une dizaine de paramètres influents, de nature discrète ou continue, et concerne la définition d'un nouveau système de satellites d'observation de la terre basé sur un ensemble de mini ou micro satellites (acquisition d'images de type SPOT).

Illustrée par la figure 1, l'évaluation porte sur la performance globale obtenue par cette constellation tout au long d'une mission opérationnelle de longue durée (16 ans), en termes de nombre moyen journalier de prises de vues ou de taux de réalisation du carnet de commande, et sur les coûts associés ramenés à la date du début de développement du premier satellite par un taux d'intérêt (5 % environ).

En jouant sur un certain nombre de paramètres influents, tels que le nombre de satellites utilisés simultanément, leur performance propre, leur durée de vie, leur fiabilité, leur durée de fabrication, la stratégie de lancement et de renouvellement..., ayant chacun un impact sur les coûts, l'optimisation consiste à rechercher la configuration optimale de la constellation et de son exploitation selon différents critères tels que le coût minimum en respectant un objectif de performance donné, ou celui de la performance maximale dans une enveloppe de coût.

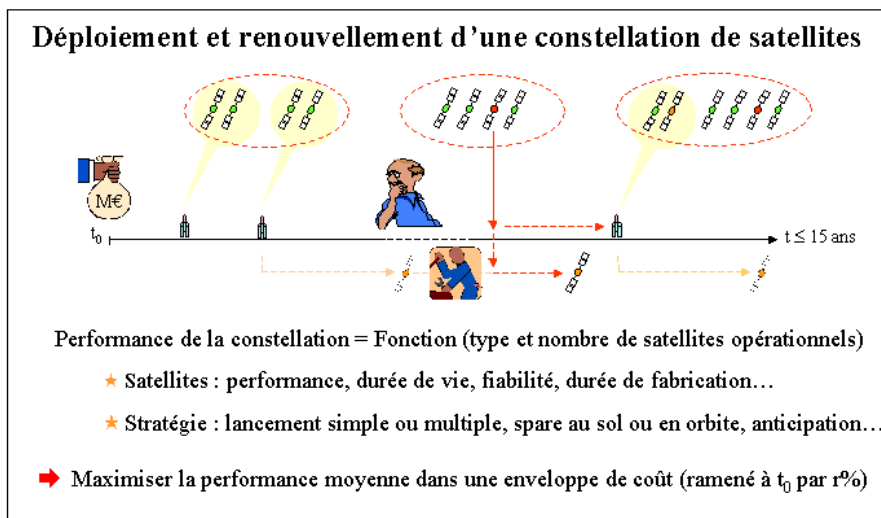


Figure 1. Cas d'application

Les hypothèses considérées dans ce cas d'application sont détaillées ci-après (bien que le problème soit réel, les données diffusées avec lesquelles cette évaluation a été réalisée sont fictives pour des raisons de confidentialité).

Constituée nominale de N satellites simultanément opérationnels ($1 \leq N \leq 4$), la constellation présente des performances qui varient avec le type (classe) et le nombre (P) de satellites fonctionnant réellement à l'instant courant, comme l'indique la table 1.

P/N	Classe 1	Classe 2	Classe 3	Classe 4
1	5,06	6,56	7,3	7,8
2	6,6	7,76	8,32	8,56
3	7,6	8,52	8,52	8,52
4	7,8	8,7	8,8	8,9

Table 1 : Performance globale de la constellation

A partir de sa date de lancement, chaque satellite est disponible après avoir éventuellement séjourné un certain temps en orbite en tant que rechange. Son activation opérationnelle est déclenchée par le sol mais n'est effective qu'après un délai de reconfiguration d'un mois.

La perte d'un satellite a lieu au moment du tir en cas de perte du lanceur, à l'occurrence d'une panne (loi exponentielle correspondant à une valeur de fiabilité à 5 ans) ou à la fin de sa durée de vie par épuisement des ergols (4 à 6 ans).

Les lancements peuvent être simples ou multiples, éventuellement différenciés entre le premier tir et les suivants, selon la capacité du lanceur.

Les dates de tir dépendent de la stratégie de renouvellement choisie et des contraintes de fabrication des satellites. Un lancement est décidé quand un nombre minimum de satellites opérationnels n'est plus assuré (1 à 4), en anticipant la fin de vie des satellites présents en orbite.

Le lancement effectif n'a lieu qu'après une durée de réservation du lanceur (6 mois) et de fabrication des satellites si des rechanges ne sont pas disponibles au sol. La fabrication de telles rechanges peut être programmée ou non et leur lancement en production peut être décalé par rapport à la date de tir des satellites précédents.

Les coûts déterministes et aléatoires du développement et du maintien opérationnel de la constellation sont ramenés à T_0 en considérant un taux d'intérêt de 5% ($C_0 = C / [1+r]t$).

Le coût récurrent et le coût de développement des satellites dépendent de leur durée de vie et de leur fiabilité, comme l'indique la table 2. D'autres caractéristiques propres aux diverses classes de satellites sont fournies en table 3.

		Durée de vie		
		4 ans	5 ans	6 ans
Fiabilité à 5 ans	0,4	90 / 15	94 / 16	100 / 19
	0,5	92 / 16	96 / 17	102 / 20
	0,6	94 / 17	98 / 18	104 / 21
Classe1				
	0,4	100 / 32	104 / 34	108 / 42
	0,5	102 / 34	106 / 36	110 / 44
	0,6	104 / 36	108 / 38	112 / 46
Classe2				
	0,4	134 / 58	140 / 60	150 / 70
	0,5	138 / 62	144 / 64	154 / 74
	0,6	142 / 66	148 / 68	158 / 78
Classe3				
	0,4	114 / 78	120 / 80	130 / 90
	0,5	118 / 82	124 / 84	134 / 94
	0,6	122 / 86	128 / 88	138 / 98
Classe4				

Table 2 : Coût développement / récurrent des satellites (M.Euros)

Classe de satellite	1	2	3	4
Coût du lanceur (M.Euros)	14	14	30	14
Nb max de satellite/lanceur	3	3	1	3
Taux de réussite lanceur	0,97	0,97	0,97	0,97
Temps minimum de réservation lanceur	0,5	0,5	0,5	0,5
Durée de développement et de fabrication du premier satellite (année)	3	3	3,5	3
Durée supplémentaire pour la fabrication des autres satellites du premier tir (temps masqué)	0,5	0,5		0,5
Durée de fabrication des satellites récurrents (année)	1	1	2	1

Table 3 : Autres caractéristiques

La durée de fabrication des satellites récurrents peut être diminuée d'un facteur K, mais cette diminution engendre un surcoût proportionnel défini par la formule ci-dessous :

$$D = D_{\text{nominal}} / K \quad C = C_{\text{nominal}} [0,8 + 0,2 K] \quad 1 \leq K \leq 2$$

Les paramètres sur lesquels porte l'optimisation sont les suivants :

- Classe de satellite (1 à 4)
- Fiabilité à 5 ans (0,4 0,5 ou 0,6)
- Durée de vie (4, 5 ou 6 ans)
- Effort sur la durée de fabrication (facteur K de 1 à 2)
- Nombre de satellites du premier lanceur
- Nombre de satellites des lanceurs suivants
- Nombre N de satellites de la constellation nominale
- Nombre de satellite conduisant à une décision de renouvellement (< N)
- Fabrication de satellites de rechange au sol (vrai ou faux)

- Décalage entre une date de tir et la fabrication de satellites de rechange (0 à 5 ans).

Traité par l'outil SIMCAB, ce cas d'application est évalué en 1 minute environ pour 5000 cas de simulation avec un Pentium 4 (figure 2). La recherche d'une configuration optimale nécessitant de l'ordre de 5000 évaluations (nombre nécessaire à l'outil GENCAB pour résoudre des problématiques de complexité similaire évaluées par traitements markoviens), la durée globale du traitement devrait avoisiner une centaine d'heures environ. Cette durée n'étant pas réellement compatible avec une activité industrielle, d'autant que l'application choisie est de complexité moyenne, une amélioration du couplage entre optimisation et évaluation stochastiques a été recherchée afin de rendre les durées de traitement non rédhibitoires.

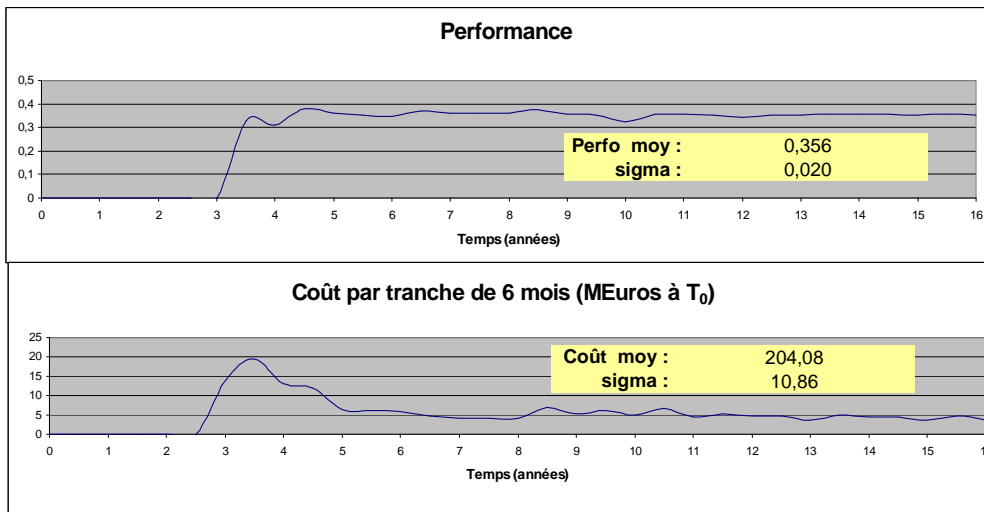


Figure 2. Résultats de l'évaluation d'une configuration

3. Technique de couplage mise en oeuvre

Cette technique originale de couplage consiste à faire varier judicieusement, tout au long du traitement, le nombre N de simulations de chaque évaluation, en exploitant la moyenne et la variance des résultats obtenus à partir d'un nombre limité N₀ de simulations.

Le principe de base de cette technique est d'accorder à chaque solution candidate une même probabilité de rejet inopportun, ce qui se traduit par une condition à respecter entre les valeurs respectives N_i et N_j du nombre de simulations à réaliser pour évaluer deux candidats i et j :

$$N_i/N_j = [(M-m_j)*\sigma_j / (M-m_i)*\sigma_i]^2$$

Cette condition résulte directement de l'application du théorème central limite comme l'illustre la figure 3 :

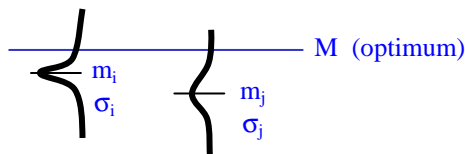


Figure 3. Comparaison de deux solutions après N₀ simulations

Trouvée dans la littérature scientifique [6], une technique récente de couplage dite optimale, l'algorithme OCBA (Optimal Computing Budget Allocation), applique ce même principe pour rechercher une valeur optimale parmi un nombre fini p de candidats. A chaque itération k, celui-ci autorise n nouvelles simulations distribuées selon les ratios indiqués en figure 4 avec i la meilleure solution courante trouvée durant l'itération k-1.

$$\frac{N_i}{N_j} = \sigma_i \sqrt{\sum_{j=1, j \neq i}^p \frac{1}{\sigma_j^2} \rho_{ij}^2}, \quad i \neq \hat{i},$$

$$\rho_{ij} = \left(\frac{\sigma_j / \Delta_j}{\sigma_i / \Delta_i} \right)^2, \quad i, j \in 1, 2, \dots, p, \quad i, j \neq \hat{i},$$

$$\Delta_i = J_i - J_i.$$

Figure 4. Algorithme OCBA

Dans le cas d'un nombre infini de candidats, ce principe peut être également appliqué sous réserve de certaines adaptations. C'est ce qui a été fait avec l'outil GENCAB basé sur une méthode hybride associant Algorithmes Génétiques et Simplexe non linéaire (algorithme de Nelder Mead) dont le fonctionnement est illustré par la figure 5 (le simplexe cherchant à améliorer localement les meilleures solutions trouvées par les Algorithmes Génétiques).

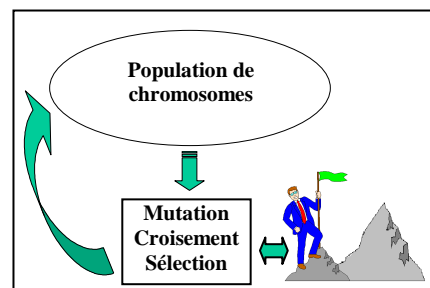


Figure 5. Couplage Algorithmes Génétiques & Simplexe

Le nombre limité de simulations N_0 et celui nécessaire à la précision requise N sont définis a priori par l'utilisateur. La population initiale de chromosomes (solutions potentielles) est alors évaluée à N_0 , puis la meilleure solution parmi celle-ci (en valeur moyenne) est réévaluée à N (par ajout de $N-N_0$ simulations). Au cours du traitement des Algorithmes Génétiques, chaque candidat i résultant d'une mutation, d'un croisement ou d'une recherche locale (simplexe) est évalué à N_0 puis réévalué soit à la valeur N_i obtenu par application de l'algorithme OCBA (uniquement si $N_i > N_0$) soit à la valeur N s'il s'avère meilleur que la solution optimale courante. La sommation utilisée par l'algorithme OCBA dans le calcul des ratios est mise à jour à chaque évaluation (afin de ne pas devoir réévaluer ultérieurement les anciens candidats) et est réinitialisée à l'émergence de toute meilleure solution.

4. Résultats

Pour évaluer l'apport du couplage mis en oeuvre, l'optimisation a été réalisée en cherchant à maximiser la performance moyenne de la constellation au cours d'une mission dans une enveloppe de coût (640 M€ sur 16 ans) exprimée à 1 sigma, soit avec un risque de 15 % environ de non satisfaction de cette contrainte. La configuration suivante a ainsi été obtenue au terme d'un traitement de 50 boucles, de 180 évaluations par boucle environ, avec couplage réglé à $N = 5000$ et $N_0 = 50$ (les Algorithmes Génétiques ayant conservé leur réglage d'usine à l'exception du simplexe réglé à 25 pas par boucle) :

- Classe de satellite = 2
- Fiabilité à 5 ans = 0,4
- Durée de vie = 5 ans)
- Facteur K sur la durée de fabrication = 1,74
- Nombre de satellites du premier lanceur = 2
- Nombre de satellites des lanceurs suivants = 1
- Nombre N de satellites de la constellation nominale = 3
- Nombre de satellite conduisant à une décision de renouvellement = 2
- Fabrication de satellites de rechange au sol = Faux
- Décalage entre une date de tir et la fabrication de satellites de rechange = N/A
- Performance moyenne : 6,76 ($\sigma = 0,14$)
- Coût : 572,8 M.Euros ($\sigma = 43,66$)

Si cette configuration semble intéressante, bien que son optimalité ne puisse être garantie, le gain apporté par le couplage au cours de ce premier traitement est apparu décevant en regard du gain escompté. En effet, le nombre moyen de simulations réalisées par évaluation était de 2980 soit seulement une diminution dans un rapport 1,7 de la durée globale de traitement.

Il est à noter cependant que la convergence semblait acquise bien avant la fin de la 50ème boucle.

Aussi des tests complémentaires ont-ils été réalisés afin de mieux appréhender les phénomènes mis en oeuvre. Selon les cas traités, l'instant considéré durant leur traitement, et la précision recherchée, une grande disparité du gain est apparue au cours de ces différents tests.

Cette observation s'explique par le principe même du couplage adopté qui est d'autant plus efficace que la solution optimale courante est meilleure que les candidats évalués.

Ainsi une fonction d'évaluation présentant de nombreux pics sera traitée beaucoup plus efficacement qu'une fonction aux sommets très arrondis. De même, le caractère plus ou moins stochastique de cette fonction ainsi que le nombre N de simulations requises influencent directement la variance des résultats obtenus et par la même l'efficacité du couplage, d'une part, et celle de l'ensemble du processus d'optimisation, d'autre part.

Comme l'illustre la figure 6, la précision demandée influence l'efficacité du couplage puisque celle-ci conduit à un rapprochement de la plupart des résultats obtenus par rapport à la solution optimale courante (l'intervalle de confiance au risque α de la performance moyenne d'une même configuration est par exemple 7 fois plus étroit si celle-ci est évaluée à 5000 simulations plutôt qu'à 100).

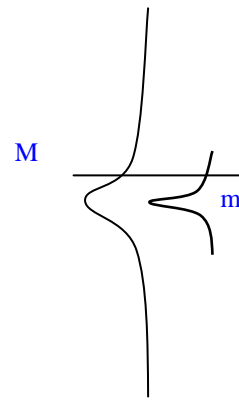


Figure 6. Distribution de la performance d'une même configuration évaluée à 100 et 5000 simulations

Le gain apporté par le couplage évalué précédemment passe ainsi de 1,7 pour 5000 simulations à 1,9 pour 1000 et 2,7 pour 100.

Ce gain croît au début du traitement, dès que les premières bonnes solutions émergent, puis diminue par la suite quand toute la population s'améliore. Le gain moyen était ainsi de 2,3 environ au cours des premières boucles du traitement à 5000 simulations avant de progressivement décroître jusqu'à atteindre la valeur de 1,7 au terme des 50 boucles.

De plus les Algorithmes Génétiques sont programmés (avec possibilité de réglage) de manière à ce que l'amplitude des sauts aléatoires dans l'espace des solutions diminue au cours des générations successives de population, ce qui devrait favoriser la convergence mais a tendance à uniformiser les résultats obtenus. En revanche, la mise à l'échelle qui est une transformation des résultats qui amplifie les écarts au fur et à mesure de la recherche pour accélérer la convergence (effet de zoom) est rendue en partie inopérante par la variance des résultats.

Une 51ème boucle réalisée à partir de la population finale obtenue précédemment présentait ainsi une performance de couplage améliorée de 1,85.

Par ailleurs, les évaluations exécutées par le simplexe sont pénalisantes car réalisées à proximité de l'optimum courant, ce qui a conduit à limiter le nombre de pas de simplexe effectué à chaque boucle (25 plutôt que 50 proposés par défaut).

Ces tests ont enfin mis en évidence une difficulté propre au couplage entre optimisation et simulation stochastique relative au traitement des contraintes quand celles-ci concernent des résultats de simulation (limitation du coût à 1 sigma dans notre exemple). Le dépassement des bornes définissant la contrainte se traduit par la soustraction d'une très forte pénalité au résultat, proportionnelle à l'amplitude du dépassement. Compte tenu de la variance des résultats, plusieurs évaluations d'une même solution à proximité des bornes peuvent fortement varier et conduire éventuellement à la perte inopinée d'une solution jugée précédemment optimale comme l'illustre la figure 7.

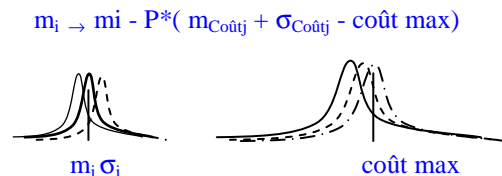


Figure 7 - Traitement des contraintes

Réglable mais prise par défaut par l'outil à la valeur 10^{50} , le facteur de pénalité P doit donc être correctement ajusté.

5. Améliorations et perspective

Les premiers tests effectués ont permis de mettre en évidence une faiblesse inhérente au principe de couplage : Le nombre N_i de simulations réalisées pour chaque évaluation est calculé à partir d'une estimation de l'écart entre la performance de celle-ci et celle de la solution optimale courante. Or cette dernière s'améliore lentement et, de par leur principe même, les Algorithmes Génétiques (et le simplexe a fortiori) privilégient l'évaluation de solutions candidates à proximité des meilleures solutions et donc avec des performances voisines. C'est pourquoi le gain apporté par ce couplage reste limité.

Il apparut toutefois judicieux de ne pas requérir la précision maximale dès les premiers calculs mais de faire croître la précision demandée tout au long du traitement parallèlement à l'amélioration progressive de la population de solutions ; un profil linéaire d'évolution du nombre de simulations allant de N_0 à N de la première à la dernière boucle a ainsi été implanté.

Une seconde amélioration a été apportée au couplage en n'évaluant plus systématiquement avec la précision maximale (N simulations) les candidats estimés meilleurs, après N_0 simulations, que la solution optimale courante, mais seulement avec la précision proposée par l'algorithme OCBA (N_i) dans la mesure où celle-ci n'excède pas la précision maximale ($N_i \leq N$).

Avec la modification concernant l'évolution progressive de la précision demandée, le gain apporté par le couplage pour 5000 simulations passe de 1,7 à 2,8. Ce gain passe de 1,7 à 5,3 (soit 950 simulations en moyenne au lieu de 5000 par évaluation) quand les deux modifications sont simultanément mises en oeuvre ; la convergence étant observée dès la 35^{ème} boucle avec des solutions très voisines de celles précédemment obtenues.

REFERENCES :

- [1] A. Pages & M. Gondran - Fiabilité des systèmes - Edition Eyrolles, Paris 1980
- [2] A. Villemeur - Sûreté de Fonctionnement des systèmes industriels - Edition Eyrolles, Paris 1987
- [3] David E. Goldberg, Algorithmes Génétiques, Exploration optimisation et apprentissage automatique, Addison-Wesley, 1994.
- [4] J-M. Renders, Algorithmes génétiques et réseaux de neurone, Hermes, 1995
- [5] R. S. Sutton , A. G. Barto, Reinforcement Learning : An Introduction, MIT Press, Cambridge, MA, 1998.
- [6] C. H. Chen, J. Lin, E. Yucesan, and S. E. Chick. Simulation budget allocation for further enhancing the efficiency of ordinal optimization. Journal of Discrete Event Dynamic Systems : Theory and Application, 10(3) :251-270; 2000
- [7] A. Cabarbaye – Outil générique d'optimisation par Algorithmes Génétiques et Simplexe - 8^{èmes} Journées Nationales du groupe

Plus conforme à notre attente, cet apport du couplage pourrait probablement être encore augmenté par le choix d'un profil mieux adapté d'amélioration de la précision. Le traitement d'un autre cas réel d'application relativement plus complexe que celui présenté dans cette communication devrait prochainement nous fournir des résultats complémentaires que nous espérons pouvoir présenter au cours du colloque $\lambda\mu$.

CONCLUSION

Le couplage entre optimisation et simulation stochastique, qui consiste à rechercher une configuration optimale de paramètres d'un système à partir des résultats d'une fonction d'évaluation traitée par simulation de Monte-Carlo, est très pénalisant en terme de durée de traitement.

Aussi, l'apport du couplage proposé dans cette communication est significatif (diminution de la durée globale des traitements dans un rapport 5 environ selon les problèmes à traiter), et devrait pouvoir être encore amélioré.

Parallèlement à l'augmentation continue des performances d'ordinateurs, ce couplage devrait faciliter la résolution des problématiques d'optimisation qui ne peuvent pas être traitées autrement que par simulation. Certaines de celles-ci concernent notamment la conception et l'exploitation de systèmes complexes vis-à-vis de la Sûreté de Fonctionnement ; Les enjeux sous-jacents étant parfois considérables tant notre activité, et notamment la maintenance, cache encore des gisements de productivité.

Mode (Mathématique de l'Optimisation et de la Décision) de la SMAI, Toulouse 23 - 25 mars 2000.

[8] Frédéric Garcia, Linda Tomasini, Julien Séroï, André Cabarbaye - Optimisation de la maintenance d'une constellation de satellites - 12^e Colloque National de Sûreté de Fonctionnement (Lambdamu 12), Montpellier 28 - 30 mars 2000.

[9] A. Cabarbaye - SIMCAB : Un outil générique de Simulation sous Microsoft Excel® - 3^e Conférence Francophone de Modélisation et Simulation (MOSIM'01), Troyes 25 au 27 avril 2001.

[8] A. Cabarbaye - SUPERCAB PRO : Un atelier d'Ingénierie Système sous Microsoft Excel® - 2^{ème} Conférence Annuelle d'Ingénierie Système, organisée par l'AFIS, TOULOUSE, 26-28 juin 2001.

[10] Roland Lautheret, Bruno Lacosta, André Cabarbaye - Minimisation du coût d'un système contraint par un objectif de disponibilité - QUALITA 2003 - Nancy 18-20 Mars 2003.