

Une approche différence-de-convexes pour un OPF sous contrainte probabiliste

Ksenia Syrtseva *

CMA, Mines Paris–PSL et EDF R&D
ksenia.syrtseva@minesparis.psl.eu

Mots-clés : *contraintes probabilistes, optimal power flow, différence de convexes.*

1 Introduction

Aujourd’hui, les réseaux de distribution électriques intègrent de nouveaux profils de consommateurs et de la production décentralisée, notamment des unités de génération s’appuyant sur des sources renouvelables variables et intermittentes, comme l’éolien ou le photovoltaïque. Cela engendre davantage d’incertitudes dans les prévisions de production et de demande, et donc un plus grand risque pour l’opérateur (Distribution System Operator ou DSO) de ne pas savoir anticiper un déséquilibre du réseau. Par contre, pour résoudre ces déséquilibres, le DSO possède maintenant de nouveaux leviers de flexibilité, notamment la modulation de puissance qu’il contractualise avec générateurs et producteurs individuellement.

La gestion opérationnelle court-terme du réseau de distribution se traduit alors par un problème d’Optimal Power Flow (OPF) sous incertitudes. Il s’agit pour le DSO de décider d’un niveau d’activation des leviers, dans leurs limites contractuelles et à moindre coût, permettant d’assurer l’équilibre du réseau pour des prévisions de volume de charge et de production stochastique. L’optimisation sous contraintes probabilistes est une approche de l’incertitude privilégiée dans la gestion des systèmes énergétiques [6]. Si les modèles probabilistes sont intuitifs, ils ne sont généralement ni convexes, ni différentiables, ce qui les rend difficiles à résoudre. Ses applications à l’OPF (voir par ex. [7, 1, 4]) reposent ainsi principalement sur des approximations linéaires ou convexes des contraintes probabilistes et des équations de flux de puissance en courant alternatif (AC-Load Flow ou LF). Par ailleurs, elles s’appuient sur l’indépendance supposée des variables aléatoires du modèle, permettant une reformulation déterministe ou robuste du problème.

L’objet de la présente étude est de prendre en considération la corrélation forte observée entre profils de charge et de production. Nous développons, pour ce faire, un modèle avec contraintes probabilistes jointes : étant donnée la réalisation d’un scénario x^ξ de charge et de production, la décision x de modulation doit assurer l’existence d’un équilibre réseau (i.e. une solution de LF) avec une probabilité minimale : $\mathbb{P}[x^\xi - x \in LF] \geq 1 - \alpha$, $\alpha > 0$. Résoudre ce modèle par une méthode numérique nécessite de disposer d’un *oracle* capable d’évaluer une fonction de probabilité multi-variée et d’en calculer un sous-gradient. Cette tâche complexe est souvent approchée par simulations Monte-Carlo qui, dans le cas convexe, conduit à une reformulation déterministe discrète. Dans le cas des équations non-convexes de flux, et pour éviter d’avoir à les approximer, nous proposons à la place une reformulation Différence-de-Convexes (DC) et développons l’oracle associé, naturellement parallélisable. Nous appliquons alors la méthode numérique proposée dans [5], un algorithme des faisceaux pour les problèmes DC à contraintes DC, et illustrons la performance de notre approche sur le test d’un réseau de 33 noeuds.

*Cet article est extrait et traduit de : Ksenia Syrtseva, Welington de Oliveira, Sophie Demassey, Hugo Morais, Paul Javal, Bhargav Swaminathan. *Difference-of-Convex Approach to Chance-Constrained Optimal Power Flow modelling the DSO Power Modulation Lever for Distribution Networks*, à paraître dans *Sustainable Energy, Grids and Networks* (2023)

2 AC-OPF avec contraintes probabilistes jointes

Le réseau de distribution est représenté par un graphe sur des ensembles $\mathcal{A}$ d'arcs (les lignes) et $\mathcal{N} \cup \{s\}$ de noeuds (les bus, incluant le noeud s de raccordement au réseau de transport). À chaque noeud $i \in \mathcal{N}$, peut être raccordé au plus un générateur ou consommateur dont il faut décider la réduction de puissance produite ou consommée x_i . On considère la puissance apparente dans $\mathbb{C}$, identifiée ici au couple $x_i \in \mathbb{R}^2$ des puissances active (partie réelle) et réactive (partie imaginaire). La prévision de puissance est un paramètre incertain noté x_i^ξ ($x_i^\xi \geq 0$ si produite ou $x_i^\xi \leq 0$ si consommée) et associé à la réalisation de scénario d'un vecteur aléatoire $\xi \in \Xi \in \mathbb{R}^d$. Les contraintes contractuelles s'expriment comme des bornes et des coûts sur x_i : l'ensemble des bornes définit un polytope $X \ni x = (x_i)_i$ et l'ensemble des coûts définit la fonction objectif $f(x)$ à minimiser, que l'on suppose quadratique convexe. En décomposant x_i comme une somme de variables avec des bornes et coûts propres, on peut alors modéliser différentes formes de flexibilité : modulation, écrêtage, effacement, à puissance garantie, etc.

Déterminer l'existence d'un équilibre réseau opérationnel pour un flux de puissance $p_i = x_i^\xi - x_i$ en chaque noeud $i \in \mathcal{N}$ c'est prouver l'existence d'une solution de Load Flow, c'est à dire des valeurs de tension V_i et de phase δ_i en chaque noeud et une puissance appelée $p_s \geq 0$ sur le réseau de transport, pris dans un ensemble convexe $(V, \delta, p_s) \in L$ (qui traduit les contraintes de bornes sur chaque variable relatives au fonctionnement du réseau et des contraintes quadratiques convexes de congestion sur les arcs), et vérifiant les équations de Load Flow de la forme $p_i = l_i(V, \delta)$ avec l_i une fonction non-linéaire (trigonométrique) en chaque noeud $i \in \mathcal{N} \cup \{s\}$. Nous formulons ainsi le problème de gestion opérationnel par le modèle probabiliste suivant :

$$(P) : \min_{x \in X} f(x) : \mathbb{P}[x^\xi - x \in LF] \geq 1 - \alpha$$

où

$$LF = \{p : \exists (V, \delta, p_s) \in L, p_i = l_i(V, \delta) \forall i \in \mathcal{N} \cup \{s\}\}$$

désigne l'ensemble des flux de puissance opérationnels et où le paramètre $\alpha > 0$ représente la tolérance au risque d'instabilité. Suivant l'usage, on suppose l'ensemble support Ξ du vecteur aléatoire ξ représenter un nombre fini de scénarios Monte-Carlo et donc être un ensemble discret fini. En associant une variable binaire à la réalisation de chacun des scénarios et en les liant aux contraintes de LF par la méthode big-M, on peut reformuler (P) en un modèle déterministe. Mais, dans ce cas précis où les contraintes LF sont non-linéaires, et à moins de les linéariser, cela produit un programme non-convexe en nombres entiers de très grande taille.

3 Programmation Différence-de-Convexes (DC)

Nous choisissons l'alternative de reformuler (P) en un programme continu non-convexe, plus précisément, en un programme continu sous contraintes linéaires (représentées ici par le polytope X) et une contrainte différence-de-convexes (DC) non nécessairement différentiable :

$$(P_{DC}) : \min_{x \in X} f(x) : c_1(x) - c_2(x) \leq 0,$$

avec c_1 et c_2 deux fonctions convexes (non-différentiables). La programmation DC est un sous-domaine important de la programmation non-convexe non-lisse et qui rencontre un intérêt croissant (voir [2]) ; d'une part, elle trouve de nombreuses applications pratiques, et d'autre part, elle offre des conditions d'optimalité et permet de développer des algorithmes heuristiques ou exactes au sens de ces conditions. Pour le cas le plus général qui s'applique ici, où non seulement la fonction objectif peut être DC, mais aussi des contraintes, il existe notamment un algorithme proximal de faisceaux [5] garantissant théoriquement l'obtention d'un point critique et avec de bonnes performances pratiques.

3.1 Algorithme proximal de faisceaux

L'algorithme de [5] intègre les contraintes DC dans l'objectif, non pas par pénalisation, mais via la *fonction améliorante* $H_\tau(x) = \max\{f(x) - \tau_f, c_1(x) - c_2(x) - \tau_c\}$ qui possède elle même une décomposition DC :

$$H_\tau = F_\tau - c_2 \text{ avec } F_\tau(x) = \max\{f(x) + c_2(x) - \tau_f, c_1(x) - \tau_c\} \text{ convexe.}$$

La méthode des faisceaux est appliquée à $(P_\tau) : \min_{x \in X} H_\tau(x)$, un modèle DC sous contraintes convexes, en l'adaptant de façon à ajuster le paramètre τ progressivement vers $\tau^* = (f^*, 0)$, où f^* dénote l'optimum (inconnu) de (P_{DC}) . On observe, en effet, que les solutions de (P_{τ^*}) résolvent (P_{DC}) .

La méthode proximale des faisceaux est une méthode numérique du premier ordre pour le cas non-lisse qui explore l'espace X de manière itérative, dirigée par les informations proximales (i.e. la distance au centre de stabilité $\hat{x}$ correspondant à la dernière solution calculée avec une amélioration sérieuse du coût) et par les sous-gradients (i.e. une généralisation de la dérivée). Pour ce faire, la méthode appelle à chaque itération $k \in \mathbb{N}$ pour le point candidat $x^k \in X$, une procédure **Oracle** pour calculer la valeur $c_i(x^k)$ et un sous-gradient arbitraire $s_i^k \in \partial c_i(x^k)$ pour $i \in \{1, 2\}$. On construit ainsi, progressivement, une approximation par défaut des fonctions convexes par un modèle de plans coupants :

$$\check{c}_i^k(x) = \max_{j=0 \dots k} \{c_i(x^j) + \langle s_i^j, x - x^j \rangle\} \leq c_i(x),$$

On en dérive une approximation de H_τ :

$$\check{H}_\tau^k(x) = \max\{\check{f}^k(x) + \check{c}_2^k(x) - \tau_f, \check{c}_1^k(x) - \tau_c\} - \langle s_2^k, x \rangle.$$

L'algorithme itère sur le candidat choisi comme solution du programme quadratique convexe :

$$x^{k+1} \in \arg \min_{x \in X} \check{H}_\tau^k(x^k) + \mu \|x - \hat{x}\|^2 \text{ pour un facteur proximal } \mu \text{ donné.}$$

Si le nouveau candidat x^{k+1} améliore suffisamment l'objectif H_τ , il devient alors le nouveau centre de stabilité et le paramètre τ est mis à jour :

$$\hat{x} \leftarrow x^{k+1}, \tau_f \leftarrow f(\hat{x}) + \rho(c_1(\hat{x}) - c_2(\hat{x}))^+, \tau_c \leftarrow \sigma(c_1(\hat{x}) - c_2(\hat{x}))^+,$$

où $(.)^+ = \max\{., 0\}$ et $\rho > 0$ et $\sigma \in [0, 1)$ sont librement choisis. Cette règle de mise à jour assure la convergence de l'algorithme vers un point critique du problème (P_{DC}) [5]. Par ailleurs, l'algorithme admet différentes améliorations comme la sélection du faisceau de plans coupants, ou la mise à jour du facteur proximal.

Pour appliquer cet algorithme à (P_{DC}) , il nous faut maintenant expliciter ce modèle DC et l'**Oracle** associé. Nous obtenons (P_{DC}) du modèle probabiliste (P) en deux étapes de reformulation DC : (1) la décomposition exacte de la condition d'appartenance $x^\xi - x \in LF$ et (2) la décomposition approchée de la mesure de probabilité $\mathbb{P}$. Nous présentons ci-après ces deux reformulations et les oracles correspondants, dans un contexte général, puis dans le contexte de l'OPF.

3.2 Modèle DC d'une condition d'appartenance $y \in Y$ et oracle

Dans l'espace euclidien $\mathbb{R}^n$, la condition d'appartenance $y \in Y$, avec Y fermé, s'écrit en terme de distance :

$$0 \geq d^2(y) := \min_{z \in Y} \frac{1}{2} \|z - y\|^2.$$

Si Y n'est pas convexe, la distance au carré d^2 ne l'est pas non plus, mais elle admet une décomposition DC. Par l'identité $\frac{1}{2} \|z - y\|^2 = \frac{1}{2} \|y\|^2 - \langle z, y \rangle + \frac{1}{2} \|z\|^2$, on décompose :

$$d^2(y) = g_1(y) - g_2(y) \text{ avec } g_1(y) = \frac{1}{2} \|y\|^2 \text{ et } g_2(y) = \max_{z \in Y} \{\langle z, y \rangle - \frac{1}{2} \|z\|^2\} \text{ convexes.}$$

L'oracle pour cette décomposition est une procédure qui prend $y \in \mathbb{R}^n$ en entrée et retourne $g_1(y)$, $s_1 \in \partial g_1(y)$, $g_2(y)$ et $s_2 \in \partial g_2(y)$.

g_1 est une fonction quadratique simple, différentiable et $\partial g_1(y) = \{y\}$. En revanche, g_2 est la fonction de valeur d'un problème d'optimisation non-convexe, et est donc éventuellement non-différentiable. Les sous-gradients de g_2 en $y \in \mathbb{R}^n$ sont les projetés de y sur Y , en effet :

$$\partial g_2(y) = \{s_2 \in \mathbb{R}^n : g_2(z) - g_2(y) \geq \langle s_2, z - y \rangle, \forall z \in \mathbb{R}^n\} = \arg \max_{z \in Y} \langle z, y \rangle - \frac{1}{2} \|z\|^2 = \arg \min_{z \in Y} \frac{1}{2} \|z - y\|^2.$$

Si Y est convexe (fermé) alors, il existe un unique projeté ou sous-gradient et donc g_2 est différentiable. Évaluer $g_2(y)$ et calculer un sous-gradient $s_2 \in \partial g_2(y)$ est donc direct si $y \in Y$ puisque g_2 coïncide avec g_1 sur Y , sinon il faut résoudre le problème d'optimisation $\min_{z \in Y} \frac{1}{2} \|z - y\|^2$.

L'oracle en un point $y \in \mathbb{R}^n$ est défini par la procédure suivante :

Oracle_Y($y; g_1, g_2$) : retourne ($g_1(y)$, s_1 , $g_2(y)$, s_2)

1. Définir $g_1(y) \leftarrow \frac{1}{2} \|y\|^2$, $s_1 \leftarrow y$
2. Si $y \in Y$ alors définir $g_2(y) \leftarrow \frac{1}{2} \|y\|^2$, $s_2 \leftarrow y$;
3. Sinon calculer $y^* \leftarrow \arg \min_{z \in Y} \frac{1}{2} \|z - y\|^2$ et définir $g_2(y) \leftarrow \langle y, y^* \rangle - \frac{1}{2} \|y^*\|^2$, $s_2 \leftarrow y^*$.

Dans le contexte AC-OPF, on identifie $y = x$, la décision de modulation de puissance, et $Y = Y^\xi$ pour toute réalisation ξ où :

$$Y^\xi := x^\xi - LF = \{x : x^\xi - x \in LF\}. \quad (1)$$

Le test $y \in Y$ à l'étape 2 de l'oracle consiste à résoudre un problème de Load Flow pour vérifier l'opérabilité des flux de puissance $x^\xi - x$. Si le Load Flow n'a pas de solution alors la décision de modulation x n'est pas réalisable dans la réalisation ξ . On peut alors recourir à un solveur OPF (déterministe à fonction quadratique convexe) de la littérature pour résoudre le problème à l'étape 3 de l'oracle. On notera que ce problème consiste à trouver la perturbation minimale sur y pour le rendre opérable. La réponse est $y^* - y$ soit bien une direction améliorante, i.e. celle opposée au sous-gradient $s_1 - s_2$.

3.3 Modèle DC de la loi de probabilité $\mathbb{P}$ et oracle

Nous cherchons maintenant à reformuler la contrainte probabiliste

$$\mathbb{P}[y \in Y^\xi] \geq 1 - \alpha$$

étant donnée la reformulation DC vue dans la section précédente :

$$y \in Y^\xi \iff g_1^\xi(y) - g_2^\xi(y) \leq 0 \text{ avec } g_1^\xi \text{ et } g_2^\xi \text{ convexes.}$$

Pour une décision $y \in \mathbb{R}^n$ donnée, on considère la variable aléatoire $v(\xi) = g_1^\xi(y) - g_2^\xi(y)$. Si $\mathbb{E}[\cdot]$ dénote l'espérance et χ la fonction caractéristique de $\mathbb{R}_*^+$ définie par $\chi(v) = 1$ si $v > 0$ et $\chi(v) = 0$ si $v \leq 0$, on a l'équivalence :

$$\mathbb{P}[v(\xi) \leq 0] \geq 1 - \alpha \iff \mathbb{P}[v(\xi) > 0] \leq \alpha \iff \mathbb{E}[\chi(v(\xi))] \leq \alpha.$$

A l'instar de [5], nous considérons une approximation paramétrique de la fonction χ pour lever la discontinuité en 0 :

$$\chi^t(v) = \begin{cases} 0 & \text{si } v \leq 0, \\ \frac{v}{t} & \text{si } 0 < v \leq t, \\ 1 & \text{si } t < v. \end{cases}$$

Elle est continue sur $\mathbb{R}$ pour tout $t > 0$ et tend vers χ quand t tend vers 0 : $\lim_{t \downarrow 0} \chi^t(v) = \chi(v)$. De plus, elle accepte la décomposition DC $\chi^t(v) = (\frac{v}{t})^+ - (\frac{v-t}{t})^+$, et sa composition avec une fonction DC reste DC :

$$\chi^t(v(\xi)) = \frac{\max\{g_1^\xi(y), g_2^\xi(y)\}}{t} + 1 - \frac{\max\{g_1^\xi(y), g_2^\xi(y) + t\}}{t}.$$

Comme d'usage (voir [3]), nous estimons l'espérance $\mathbb{E}[\chi(v(\xi))]$ en l'approchant par sa moyenne empirique $\frac{1}{N} \sum_{j=1}^N \chi(v(\xi_j))$ sur un échantillon de tirages Monte-Carlo $\Xi = \{\xi_1, \dots, \xi_N\}$.

On obtient ainsi une décomposition DC approchée $c_1(y) - c_2(y) \leq 0$ de la contrainte probabiliste $\mathbb{P}[y \in Y^\xi] \geq 1 - \alpha$, en définissant les fonctions convexes suivantes :

$$c_1(y) := \frac{1}{N} \sum_{j=1}^N \max\{g_1^{\xi_j}(y), g_2^{\xi_j}(y)\} + t(1 - \alpha) \text{ et } c_2(y) := \frac{1}{N} \sum_{j=1}^N \max\{g_1^{\xi_j}(y), g_2^{\xi_j}(y) + t\}. \quad (2)$$

Connaissant des sous-gradients $s_i^{\xi_j} \in \partial g_i^{\xi_j}(y)$ pour $j \in \{1, \dots, N\}$ et $i \in \{1, 2\}$, alors on peut calculer des sous-gradients $s_1 \in \partial c_1(y)$ et $s_2 \in \partial c_2(y)$ de manière analytique.

L'oracle en un point $y \in \mathbb{R}^n$ est défini par la procédure suivante, dont l'étape 1, la plus coûteuse, est naturellement parallélisable :

Oracle _{α} ($y; c_1, c_2$) : retourne ($c_1(y), s_1 \in \partial c_1(y), c_2(y), s_2 \in \partial c_2(y)$)

1. Calculer $(g_1^{\xi_j}(y), s_1^{\xi_j}, c_2^{\xi_j}(y), s_2^{\xi_j}) \leftarrow \mathbf{Oracle}_{Y_j^\xi}(y; g_1^{\xi_j}, g_2^{\xi_j})$ pour tout $j \in \{1, \dots, N\}$.
2. Définir $c_1(y)$ et $c_2(y)$ selon la formule (2) et calculer $s_1 \in \partial c_1(y)$ et $s_2 \in \partial c_2(y)$

Dans le contexte AC-OPF, on identifie de nouveau y à la modulation de puissance x et Y^ξ est l'ensemble défini par (1). On peut alors définir les fonctions c_1 et c_2 de la reformulation (P_{DC}) telles que dans l'équation (2). On applique alors l'algorithme proximal des faisceaux décrit dans la section 3.1 en appelant la procédure **Oracle _{α} ($y; c_1, c_2$)** à chaque itération.

4 Résultats numériques

Nous avons évalué le modèle probabiliste et l'algorithme des faisceaux sur une instance AC-OPF construite sur un réseau de 33 bus dont deux générateurs biomasse et un générateur éolien, dont la puissance peut être modulée selon deux types de contrat : FiT (Feed-in tariff) pour l'écrêtage ou SCP (smart connection point) pour la modulation jusqu'à un seuil garanti et l'écrêtage au-delà. Par ailleurs, tous les consommateurs supportent l'effacement FiT. L'objectif f est la somme de : (1) coûts linéaires de modulation suivant une hiérarchie entre effacement, FiT et SCP modulation et (2) coûts quadratiques uniformes privilégiant une répartition équilibrée des flexibilités. Nous générons, à partir de Enedis Open Data, $N = 1000$ scénarios de consommation et de génération, proches des limites techniques du réseau, et en attribuant une variance $\sigma^2 \in [10^{-5}, 10^{-2}]$ individuelle aux générateurs et consommateurs.

Nous évaluons la méthode sur un cas d'instabilité qui doit être corrigée par modulation, et faisons varier le paramètre de sécurité $1 - \alpha$ entre 0.75 et 1 (le cas parfaitement robuste). L'algorithme trouve une solution correctrice quelque soit le niveau de sécurité et converge en moyenne en moins de 30 minutes (sur un ordinateur portable standard), jusqu'à 38 minutes pour $1 - \alpha = 0.775$. On mesure le ratio de scénarios où la correction est effective. Le ratio effectif diffère du ratio attendu ($1 - \alpha$) à cause de l'approximation de la contrainte probabiliste (liée au paramètre t fixé à 10^{-5}) mais l'écart est faible et diminue avec le niveau de tolérance attendu α ou avec le paramètre t : le ratio effectif est par exemple de 0.876 pour $1 - \alpha = 0.9$ et de 0.893 si on approxime avec $t = 10^{-6}$. À noter que sans modulation, le ratio effectif tombe à 0.545, soit une chance sur deux que le réseau soit opérable. Par ailleurs, le coût de modulation augmente naturellement avec le niveau de sécurité visé, et plus fortement pour $1 - \alpha \geq 0.92$, quand il s'agit de couvrir les scénarios les plus extrêmes. L'activation de la modulation est donc primordiale pour la stabilité du réseau, mais une décision trop conservatrice vis-à-vis de l'incertitude peut engendrer un surcoût pour le DSO et une qualité de service moindre pour les utilisateurs connectés. L'approche probabiliste présentée ici permet de calculer une décision fiable et offre au DSO la possibilité d'évaluer différents niveaux de sécurité.

5 Conclusion et perspectives

Dans cette étude, nous présentons un nouveau modèle d'optimisation sous contraintes probabilistes jointes pour l'AC-OPF dans le contexte de la gestion opérationnelle des réseaux électriques de distribution intégrant toujours plus de génération distribuée, variable et incertaine. Nous proposons des techniques de décomposition DC pour traiter les contraintes probabilistes et de l'OPF, sans recourir à la linéarisation ou la convexification. Ces techniques sont générales et dépassent le cadre de l'OPF. Enfin, nous développons également les oracles associés permettant d'appliquer les méthodes numériques de l'optimisation non-lisse ; et mettons en œuvre un algorithme proximal dans un cas pratique pour illustrer la fiabilité de l'approche et son intérêt pour le DSO. Les perspectives futures portent sur l'intégration d'autre leviers de flexibilité, et notamment des leviers discrets.

Références

- [1] D. Bienstock, M. Chertkov, and S. Harnett. Chance-constrained opf : Risk-aware network control under uncertainty. *SIAM Review*, 56(3) :461–495, 2014.
- [2] W. de Oliveira. The ABC of DC programming. *Set-Valued and Variational Analysis*, 28(4) :679–706, 2020.
- [3] B. K. Pagnoncelli, S. Ahmed, and A. Shapiro. Sample average approximation method for chance constrained programming : Theory and applications. *Journal of Optimization Theory and Applications*, 142(2) :399–416, Aug 2009.
- [4] L. Roald and G. Andersson. Chance-constrained AC opf : Reformulations and efficient algorithms. *IEEE Transactions on Power Systems*, 33(3) :2906–2918, 2018.
- [5] W. van Ackooij, S. Demassey, P. Javal, H. Morais, W. de Oliveira, and B. Swaminathan. A bundle method for nonsmooth DC programming with application to chance-constrained problems. *Computational Optimization and Application*, 78(1) :451–490, 2021.
- [6] W. van Ackooij, R. Zorgati, R. Henrion, and A. Möller. Chance constrained programming and its applications to energy management. 02 2011.
- [7] H. Zhang and P. Li. Chance constrained programming for optimal power flow under uncertainty. *IEEE Transactions on Power Systems*, 26(4) :2417–2424, 2011.