

# Gestion des incertitudes paramétriques en simulation de systèmes chimiques complexes.

## Application aux plasmas photochimiques N<sub>2</sub>/CH<sub>4</sub>

**P. Pernot<sup>1</sup>, Z. Peng<sup>1</sup>, S. Plessis<sup>1</sup> and N. Carrasco<sup>2</sup>**

<sup>1</sup>LCP, UMR8000 CNRS/Univ. Paris-Sud, Orsay

<sup>2</sup>LATMOS, UMR8190 CNRS/UVSQ, UPMC, Guyancourt

J. Simul 2013/11/14

- 1 Incertitudes en modélisation
  - Gestion des incertitudes paramétriques
- 2 MCUP en modélisation photochimique
  - Spécification des incertitudes
- 3 Application aux plasmas N<sub>2</sub>/CH<sub>4</sub>
- 4 Conclusion

- 1 **Formulation du modèle :**  
approximations, incertitudes structurales (incomplétude)
- 2 **Implémentation numérique du modèle :**  
précision des algorithmes, seuils de convergence,  
modèles stochastiques
- 3 **Incertitudes paramétriques :**  
précision des paramètres définissant le modèle

- 1 **Formulation du modèle** :  
approximations, incertitudes structurales (incomplétude)
- 2 **Implémentation numérique du modèle** :  
précision des algorithmes, seuils de convergence,  
modèles stochastiques
- 3 **Incertitudes paramétriques** :  
précision des paramètres définissant le modèle

• Les incertitudes paramétriques conditionnent fortement :

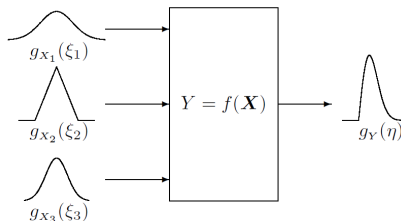
- la validation d'un modèle (1)
- la qualité de l'implémentation numérique (2)

- 1 **Formulation du modèle** :  
approximations, incertitudes structurales (incomplétude)
- 2 **Implémentation numérique du modèle** :  
précision des algorithmes, seuils de convergence,  
modèles stochastiques
- 3 **Incertitudes paramétriques** :  
précision des paramètres définissant le modèle
  - Les incertitudes paramétriques conditionnent fortement :
    - la validation d'un modèle (1)
    - la qualité de l'implémentation numérique (2)

- 1 **Formulation du modèle** :  
approximations, incertitudes structurales (incomplétude)
- 2 **Implémentation numérique du modèle** :  
précision des algorithmes, seuils de convergence,  
modèles stochastiques
- 3 **Incertitudes paramétriques** :  
précision des paramètres définissant le modèle
  - Les incertitudes paramétriques conditionnent fortement :
    - la validation d'un modèle (1)
    - la qualité de l'implémentation numérique (2)

## Équation de Markov

$$g_Y(\eta) = \int d\xi_1 \dots d\xi_k \delta(\eta - f(\xi_1, \dots, \xi_k)) g_{X_1, \dots, X_k}(\xi_1, \dots, \xi_k)$$



**Réf. :** *Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement".* JCGM 101 :2008.

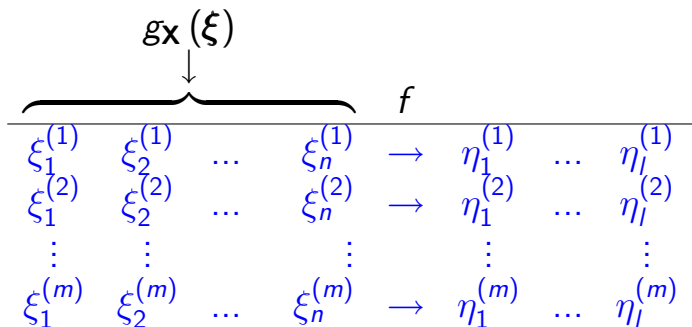
[http://www.bipm.org/utis/common/documents/jcgm/JCGM\\_101\\_2008\\_E.pdf](http://www.bipm.org/utis/common/documents/jcgm/JCGM_101_2008_E.pdf)





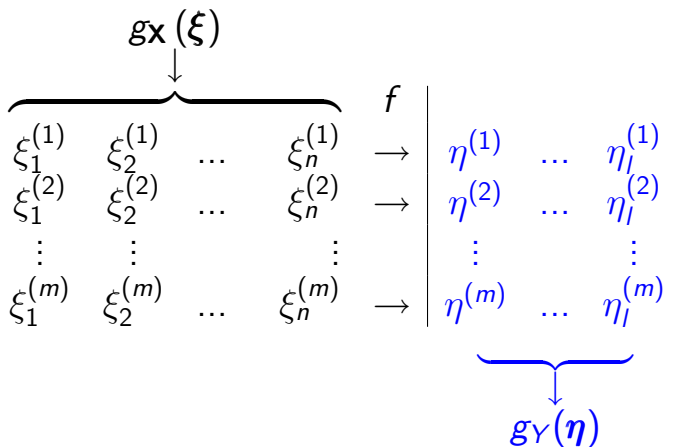
# Propagation des distributions par Monte Carlo

Étape 2 : application du modèle à chacun des points de l'échantillon



# Propagation des distributions par Monte Carlo

Étape 3 : analyse statistique de l'échantillon des sorties du modèle



## ① Identification et caractérisation des paramètres incertains

- requiert une bonne expertise dans la détermination des paramètres considérés
- peu de bases de données avec des données fiables à ce sujet : gros travail à prévoir

## ② Propagation des incertitudes par Monte Carlo

- temps de calcul éventuellement [très] long, mais parallélisme idéal (cloud computing)...

## ③ Analyse de sensibilité

- analyse de sensibilité **globale**, basée les échantillons aléatoires :
  - la plus simple est basée sur l'estimation de  $\text{corr}(Y, X_i)$  à partir des échantillons

## ① Identification et caractérisation des paramètres incertains

- requiert une bonne expertise dans la détermination des paramètres considérés
- peu de bases de données avec des données fiables à ce sujet : gros travail à prévoir

## ② Propagation des incertitudes par Monte Carlo

- temps de calcul éventuellement [très] long, mais parallélisme idéal (cloud computing)...

## ③ Analyse de sensibilité

- analyse de sensibilité **globale**, basée les échantillons aléatoires :
  - la plus simple est basée sur l'estimation de  $\text{corr}(Y, X_i)$  à partir des échantillons

## ① Identification et caractérisation des paramètres incertains

- requiert une bonne expertise dans la détermination des paramètres considérés
- peu de bases de données avec des données fiables à ce sujet : gros travail à prévoir

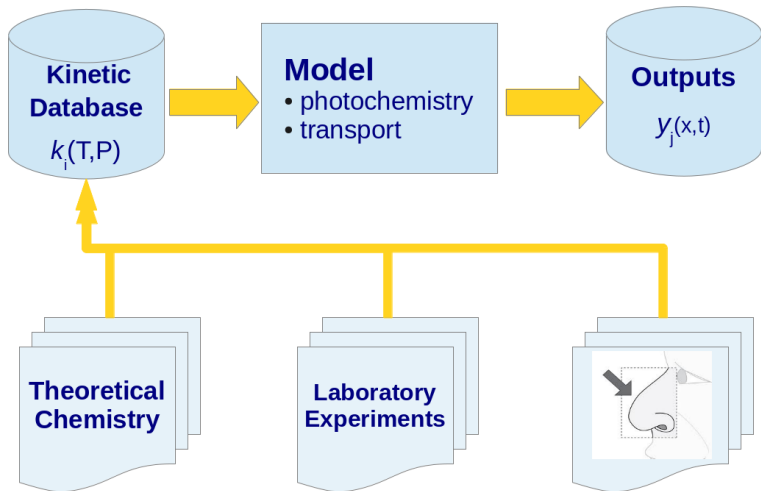
## ② Propagation des incertitudes par Monte Carlo

- temps de calcul éventuellement [très] long, mais parallélisme idéal (cloud computing)...

## ③ Analyse de sensibilité

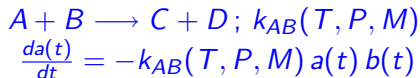
- analyse de sensibilité **globale**, basée les échantillons aléatoires :
  - la plus simple est basée sur l'estimation de  $\text{corr}(Y, X_i)$  à partir des échantillons

# Modélisation de photochimie détaillée



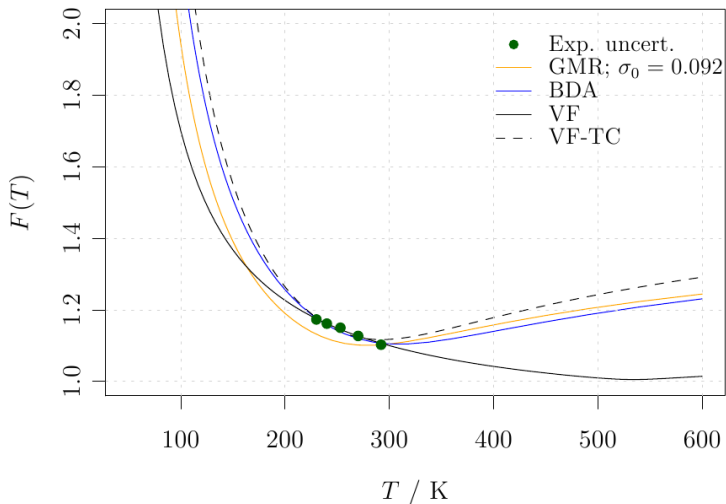
Ch.Th. : Burke et al. (2013) *Proc. Comb. Inst.* **34** :547

- **les équations chimiques sont basées sur des paramètres empiriques**



- **les constantes de vitesses sont tirées de mesures et/ou d'estimations** (peu de calculs)
  - toujours évaluées avec une [ [très] grande ] incertitude
  - dans certains modèles, il y a de nombreux paramètres estimés (par analogie)
  - *dans les modèles de Titan, moins de 10 % des constantes de vitesse de réaction ont été mesurées dans des conditions (T,P,M) représentatives (170 K, 10<sup>-6</sup> mb, N<sub>2</sub>/CH<sub>4</sub>)...*

# Incertitude d'extrapolation

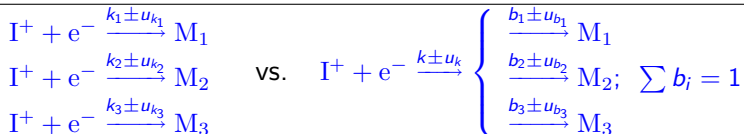


Nagy et al. (2011) *Int. J. Chem. Kinet.* **43** :359





# Le cas des réactions multi-voies



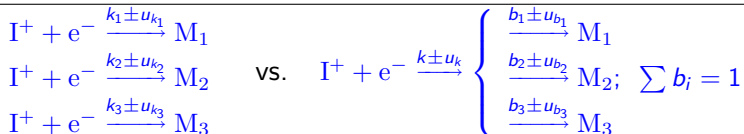
- les constantes de vitesse ( $k$ ) et les rapports de branchement ( $b_i$ ) (RB) sont typiquement mesurés dans des expériences différentes
- Règle d'Or de la Propagation des Incertitudes :

**“Une séparation des sources d'incertitude tu garderas !”**

- essentiel pour l'implémentation de la somme à 1 des  $b_i$
- fort impact sur les incertitudes et l'analyse de sensibilité

Carrasco et al. (2007) *J. Phys. Chem. A* **111** :350 

# Le cas des réactions multi-voies



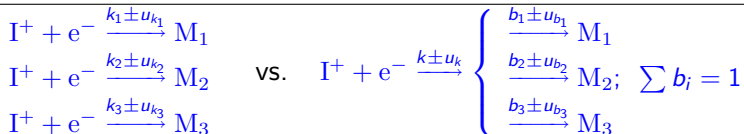
- les constantes de vitesse ( $k$ ) et les rapports de branchement ( $b_i$ ) (RB) sont typiquement mesurés dans des expériences différentes
- Règle d'Or de la Propagation des Incertitudes :

**“Une séparation des sources d'incertitude tu garderas !”**

- essentiel pour l'implémentation de la somme à 1 des  $b_i$
- fort impact sur les incertitudes et l'analyse de sensibilité

Carrasco et al. (2007) *J. Phys. Chem. A* **111** :350 

# Le cas des réactions multi-voies



- les constantes de vitesse ( $k$ ) et les rapports de branchement ( $b_i$ ) (RB) sont typiquement mesurés dans des expériences différentes
- Règle d'Or de la Propagation des Incertitudes :

**“Une séparation des sources d'incertitude tu garderas !”**

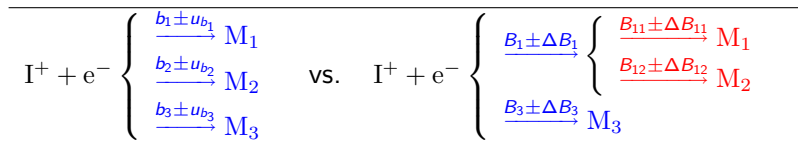
- essentiel pour l'implémentation de la somme à 1 des  $b_i$
- **fort impact sur les incertitudes et l'analyse de sensibilité**

Carrasco et al. (2007) *J. Phys. Chem. A* **111** :350 

## Exemple

Réaction à 3 voies de sortie avec des probas ( $b_1, b_2, b_3$ )

- dans une expérience on a mesuré  $\{M_1, M_2\}$  ( $B_1$ ) et  $M_3$  ( $B_3 = 1 - B_1$ );
- et dans une autre  $M_1$  ( $B_{11}$ ) et  $M_2$  ( $B_{12} = 1 - B_{11}$ ) (ou bien ils sont indéterminés...)

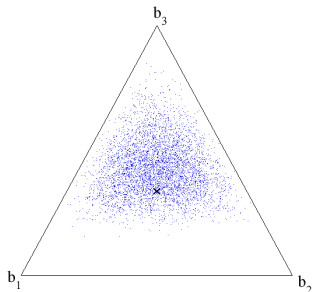


Plessis *et al.* (2010) *J. Chem. Phys.* **133** :134110



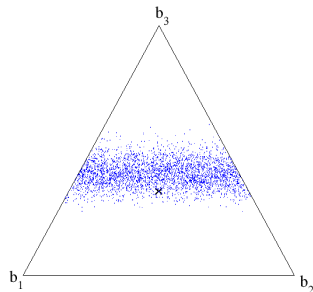
Ex. :  $B_1 = 0.6 \pm 0.1$ ,  $B_3 = 0.40 \pm 0.05$ ,  $B_{11} \in [0, 1]$ , and  $B_{12} \in [0, 1]$

Représentation 1 niveau



vs.

Représentation en arbre



# Le cas des recombinaisons dissociatives

Pour la majorité des systèmes mesurés les données sont partielles

p.ex. :  $\text{CH}_2\text{CHCNH}^+ + \text{e}^-$  [Vigren *et al.* (2009) *ApJ* **695** :317-324]

Produits	Probabilité	Voies exoergiques
$\text{C}_3\text{NH}_x + y \text{H} + w\text{H}_2$	$0.50 \pm 0.04$	6
$\text{C}_2\text{H}_x + \text{CNH}_y + z\text{H} + w\text{H}_2$	$0.49 \pm 0.04$	7
$\text{C}_3\text{H}_x + \text{NH}_y$ $\text{C}_2\text{NH}_x + \text{CH}_y + z\text{H}$	$0.01 \pm 0.01$	9
$x + y + z + 2w = 4$		22

# Le cas des recombinaisons dissociatives

Pour la majorité des systèmes mesurés les données sont partielles

p.ex. :  $\text{CH}_2\text{CHCNH}^+ + \text{e}^-$  [Vigren *et al.* (2009) *ApJ* **695** :317-324]

Produits	Probabilité	Voies exoergiques
$\text{C}_3\text{NH}_x + y \text{H} + w\text{H}_2$	$0.50 \pm 0.04$	6
$\text{C}_2\text{H}_x + \text{CNH}_y + z\text{H} + w\text{H}_2$	$0.49 \pm 0.04$	7
$\text{C}_3\text{H}_x + \text{NH}_y$ $\text{C}_2\text{NH}_x + \text{CH}_y + z\text{H}$	$0.01 \pm 0.01$	9
$x + y + z + 2w = 4$		22



Qu'est-ce que je vais bien pouvoir faire de ça ???



# Nouveau paradigme : les arbres probabilistes

**Table 1**

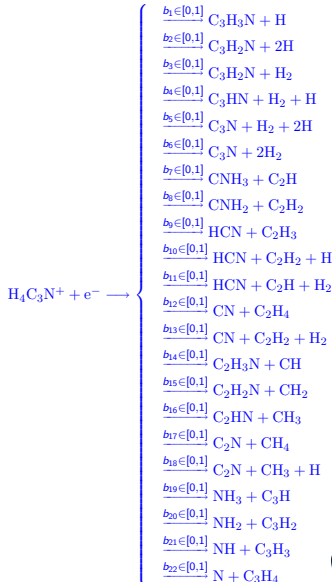
The Exoergic Channels in the Reaction  $\text{CH}_2\text{CHCNH}^+ + \text{Electron}$

Channel	$\Delta H$ (eV)	Channel	$\Delta H$ (eV)
(6a) $\text{C}_3\text{H}_3\text{N} + \text{H}$	-5.4	(6l) $\text{CNH}_3 + \text{C}_2\text{H}$	-3.2
(6b) $\text{C}_3\text{H}_2\text{N} + \text{H}_2$	-5.2	(6m) $\text{CNH}_2 + \text{C}_2\text{H}_2$	-4.7
(6c) $\text{C}_3\text{H}_2\text{N} + 2\text{H}$	-0.7	(6n) $\text{HCN} + \text{C}_2\text{H}_3$	-5.0
(6d) $\text{C}_3\text{HN} + \text{H} + \text{H}_2$	-3.6	(6o) $\text{HCN} + \text{C}_2\text{H}_2 + \text{H}$	-3.4
(6e) $\text{C}_3\text{N} + 2\text{H}_2$	-5.2	(6p) $\text{HCN} + \text{C}_2\text{H} + \text{H}_2$	-3.2
(6f) $\text{C}_3\text{N} + 2\text{H} + \text{H}_2$	-0.7	(6q) $\text{CN} + \text{C}_2\text{H}_4$	-4.5
(6g) $\text{C}_2\text{H}_3\text{N} + \text{CH}$	-2.6	(6r) $\text{CN} + \text{C}_2\text{H}_2 + \text{H}_2$	-2.6
(6h) $\text{C}_2\text{H}_2\text{N} + \text{CH}_2$	-2.9	(6s) $\text{NH}_3 + \text{C}_3\text{H}$	-2.3
(6i) $\text{C}_2\text{HN} + \text{CH}_3$	-3.0	(6t) $\text{NH}_2 + \text{C}_3\text{H}_2$	-2.3
(6j) $\text{C}_2\text{N} + \text{CH}_4$	-4.5	(6u) $\text{NH} + \text{C}_3\text{H}_3$	-2.1
(6k) $\text{C}_2\text{N} + \text{CH}_3 + \text{H}$	0.0	(6v) $\text{N} + \text{C}_3\text{H}_4$	-2.7

$$\sum_{i=1,22} b_i = 1$$

Carrasco *et al.* (2007) *J. Phys. Chem. A* **111** :3507

Plessis *et al.* (2010) *J. Chem. Phys.* **133**:134110





# Nouveau paradigme : les arbres probabilistes

**Table 2**

The Branching Fractions in the Dissociative Recombination of  $\text{CH}_2\text{CHCNH}^+$

Fragmentation		Percentage
(6a--6f)	$\text{C}_3\text{NH}_x + y\text{H} + w\text{H}_2$	50% $\pm$ 4%
(6l--6r)	$\text{C}_2\text{H}_x + \text{CNH}_y + z\text{H} + w\text{H}_2$	49% $\pm$ 4%
(6g--6k) and (6s-6v)	$\text{C}_3\text{H}_x + \text{NH}_y; \text{C}_2\text{NH}_x + \text{CH}_y + z\text{H}$	1% $\pm$ 1%

**Note.** For each line  $x + y + z + 2w = 4$ .

$$\sum_{i=1,3} B_i = 1$$

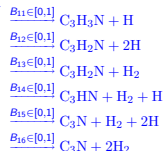
$$\sum_{j=1, n_j} B_{ij} = 1; i = 1, 3$$

Carrasco *et al.* (2007) *J. Phys. Chem. A* **111** :3507

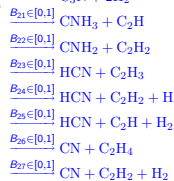
Plessis *et al.* (2010) *J. Chem. Phys.* **133**:134110



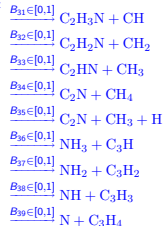
$B_1 = 0.50 \pm 0.04$



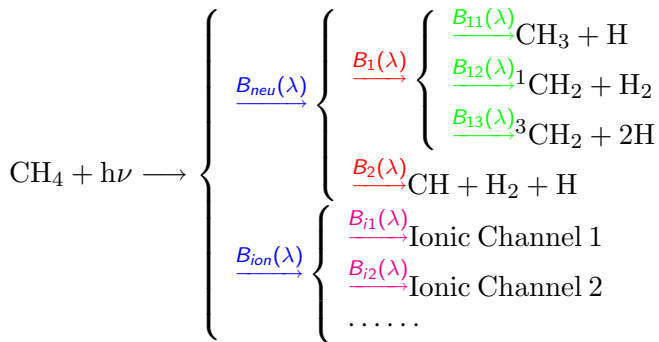
$B_2 = 0.49 \pm 0.04$



$B_3 = 0.01 \pm 0.01$



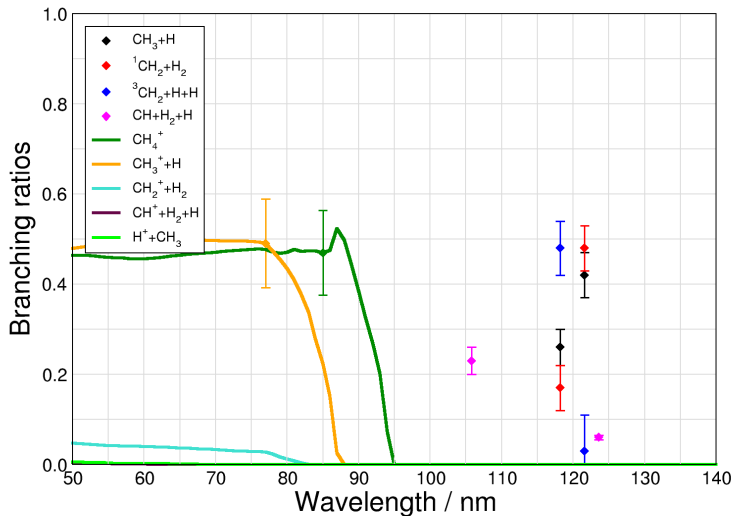
# Extension à des quantités continues



Gans *et al.* (2013) *Icarus* **223** :330-342



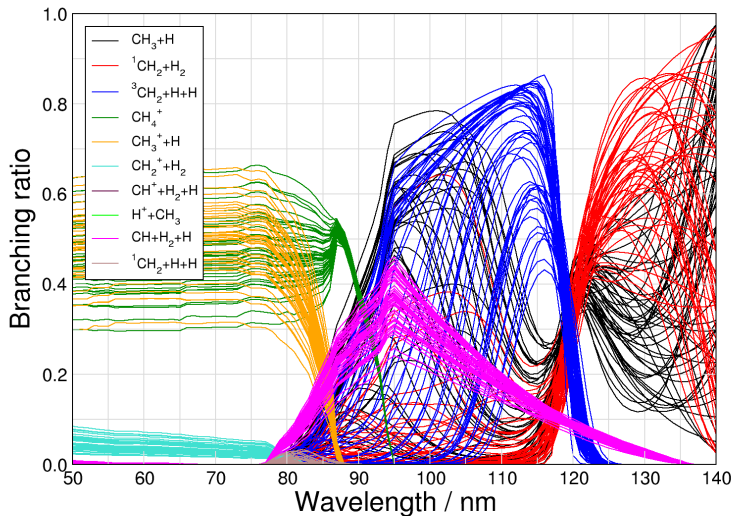
# Extension à des quantités continues



Gans et al. (2013) *Icarus* **223** :330-343

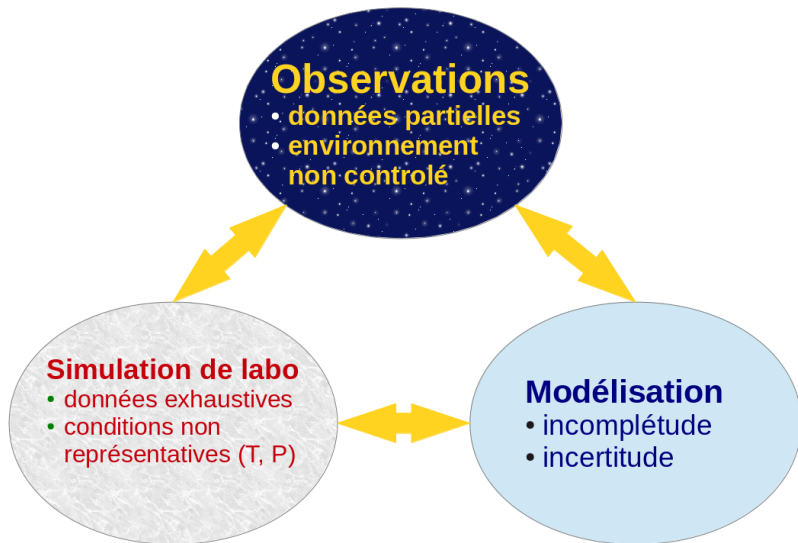


# Extension à des quantités continues

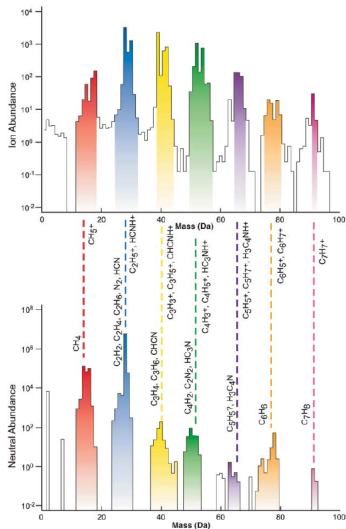


Gans et al. (2013) *Icarus* 223 :330-343

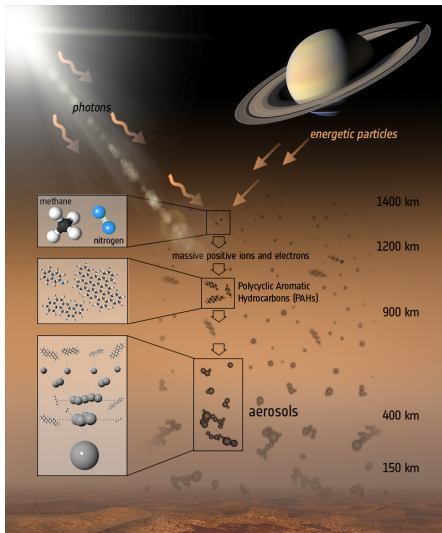




# Titan : une usine à grosses molécules



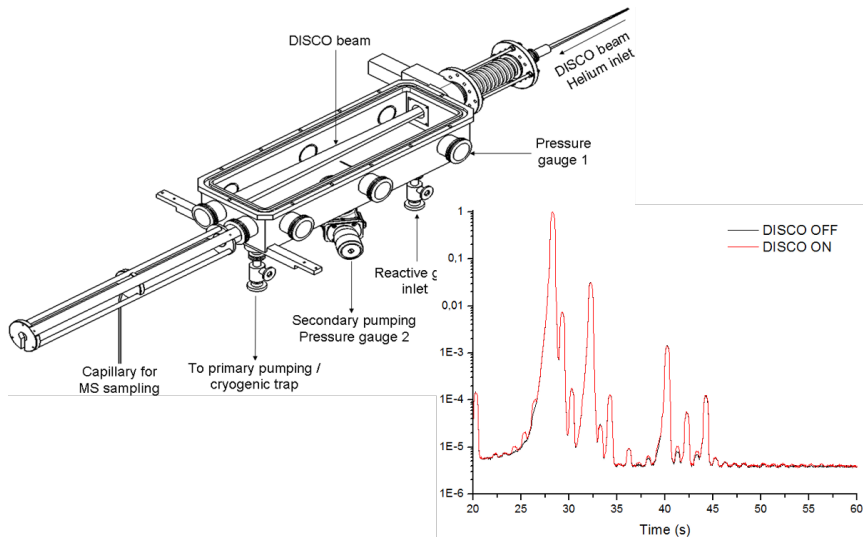
Waite *et al.* (2007) *Science*



Lopez-Puertas *et al.* (2013) *ApJ*

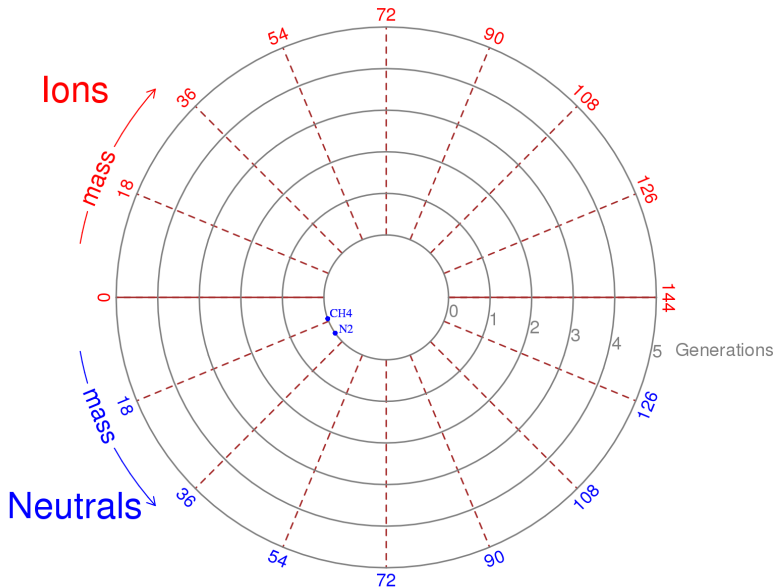


# APSYS : Atmospheric Photochem. Simul. by Synchrotron



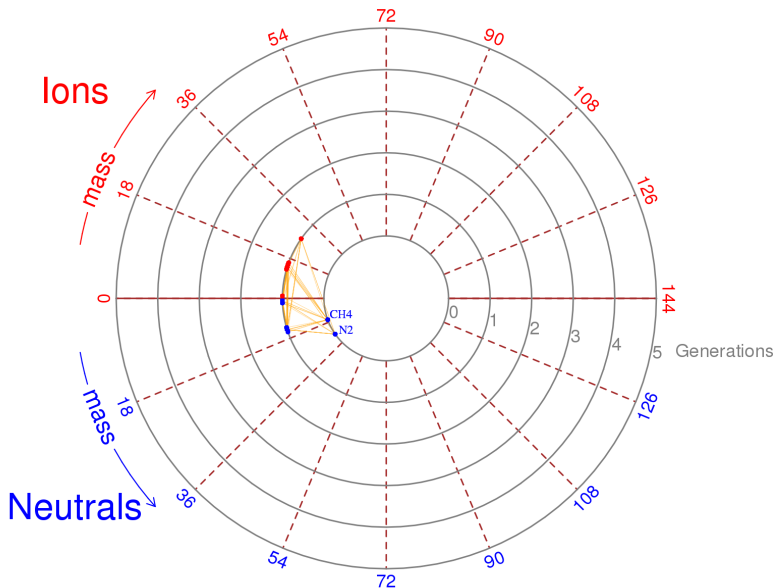
Peng *et al.* (2013) *JGR Planets* **118** :778-788

# Croissance moléculaire dans les plasmas VUV N<sub>2</sub>/CH<sub>4</sub>

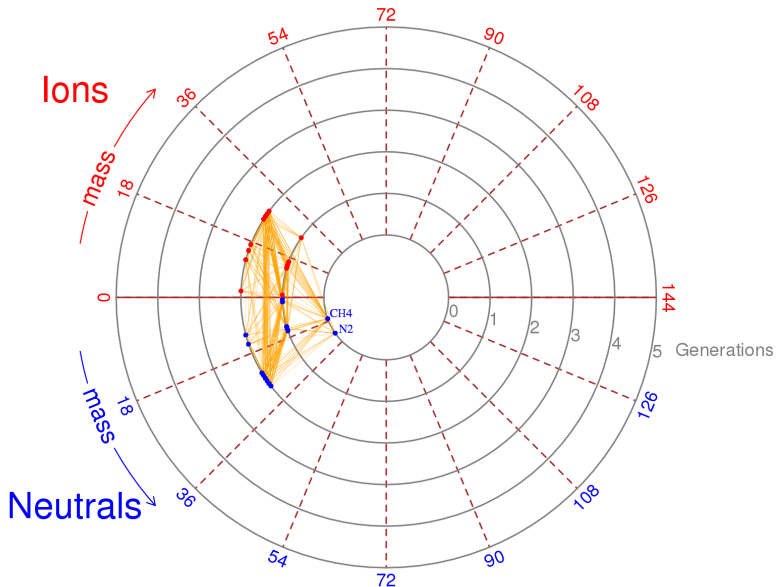




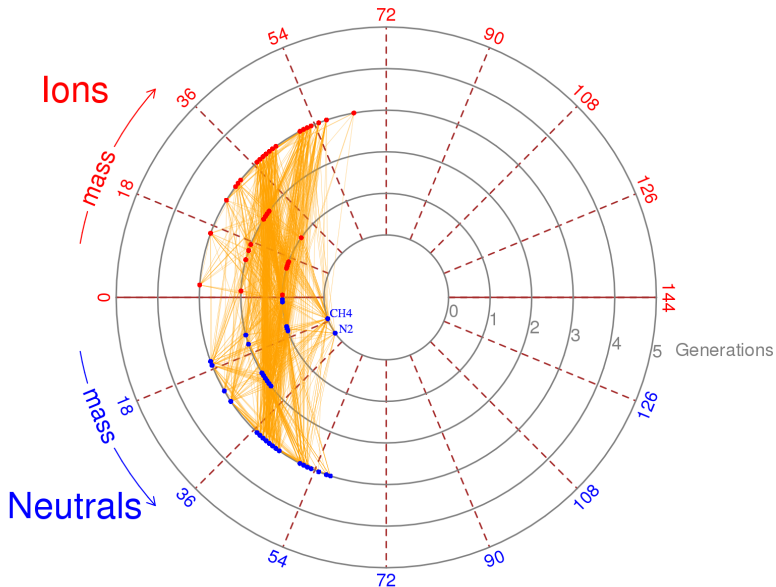
# Croissance moléculaire dans les plasmas VUV N<sub>2</sub>/CH<sub>4</sub>



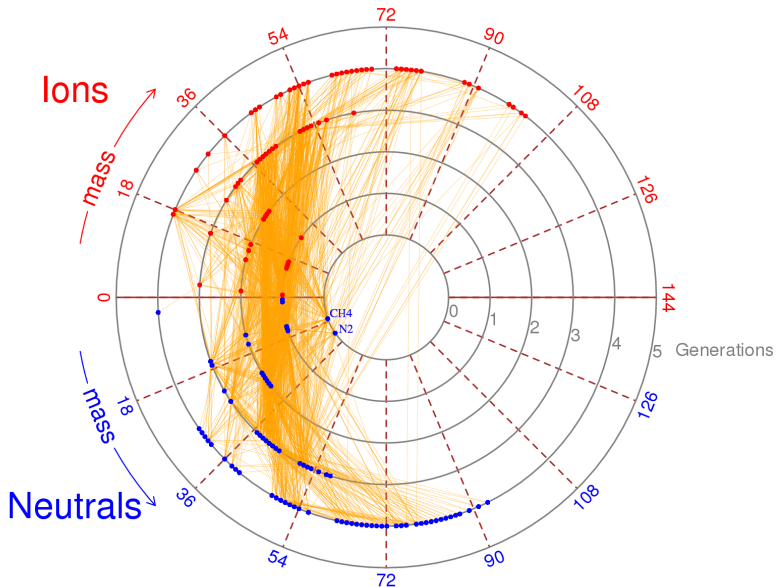
# Croissance moléculaire dans les plasmas VUV N<sub>2</sub>/CH<sub>4</sub>



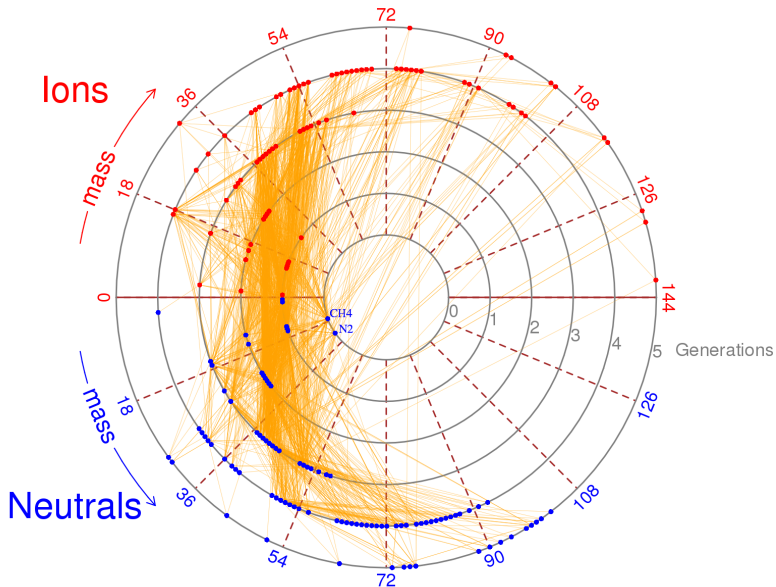
# Croissance moléculaire dans les plasmas VUV N<sub>2</sub>/CH<sub>4</sub>



# Croissance moléculaire dans les plasmas VUV N<sub>2</sub>/CH<sub>4</sub>



# Croissance moléculaire dans les plasmas VUV N<sub>2</sub>/CH<sub>4</sub>



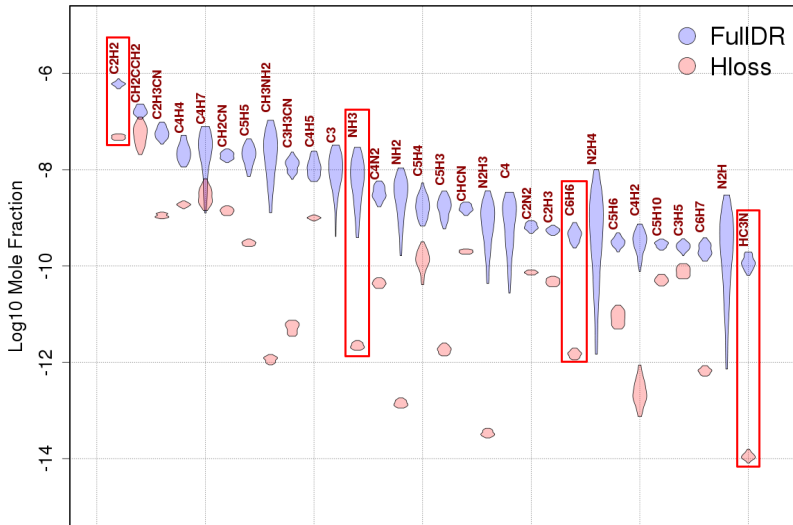
Notre modèle de photochimie VUV de  $N_2/CH_4$  contient  
**1700 réactions** (photolyse, bi- et ter-moléculaire, ion-molécule et recomb. dissoci.) et  
**237 species** (neutres et ions positifs).

- ① notre implémentation de toutes les données RD : **FullDR**  
arbres probabilistes pour 58 ions (sur 116)  
448 voies  
62 produits neutres
- ② la version "standard" basée sur la perte de H : **Hloss**  
63 voies  
48 produits neutres

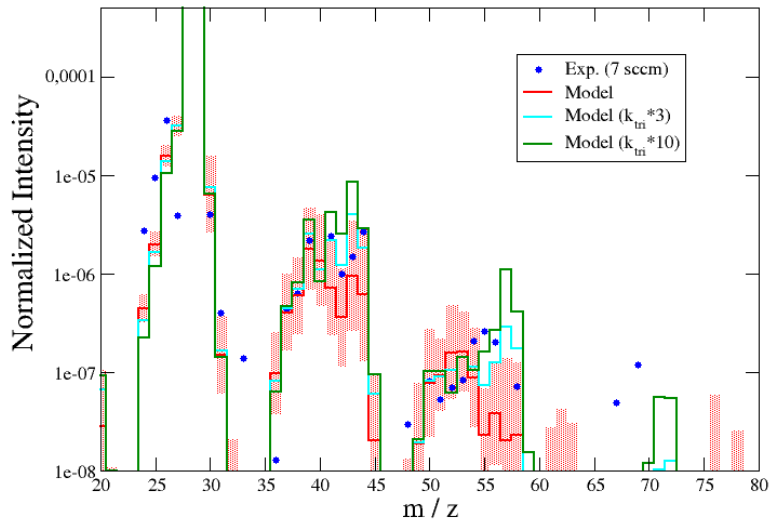
Plessis *et al.*, *JCP* (2010)



# H-loss vs. Full model



# Simulation du SM/IE des neutres APSIS

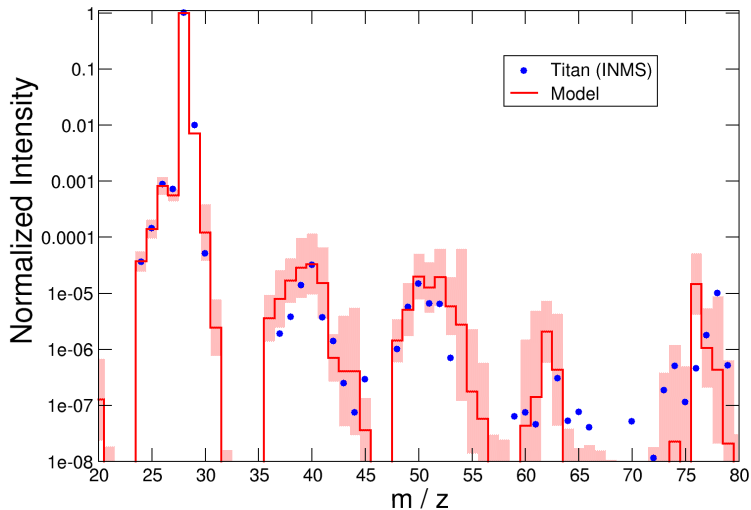


Z. Peng PhD Thesis (Sept 2013)





# Simulation du SM/IE des neutres "Titan"



Z. Peng PhD Thesis (Sept 2013)



- la gestion des incertitudes paramétriques est une étape incontournable pour la [in-]validation d'un modèle
- c'est un outil précieux pour réduire l'incertitude de prédiction
- la représentation probabiliste des paramètres incertains permet de valoriser des données inexploitable autrement

- la gestion des incertitudes paramétriques est une étape incontournable pour la [in-]validation d'un modèle
- c'est un outil précieux pour réduire l'incertitude de prédiction
- la représentation probabiliste des paramètres incertains permet de valoriser des données inexploitable autrement

- la gestion des incertitudes paramétriques est une étape incontournable pour la [in-]validation d'un modèle
- c'est un outil précieux pour réduire l'incertitude de prédiction
- **la représentation probabiliste des paramètres incertains permet de valoriser des données inexploitable autrement**

**Depuis plusieurs années, nous organisons une formation d'initiation à la propagation et à la gestion des incertitudes.**

Cette formation est offerte au 1er semestre 2014 dans deux contextes :

- 1 Form. CNRS (Ch ; Ens-Ch ; ITA/IATOS)  
<http://metrologie.cnrs.fr>
- 2 Form. doctorale (ED Chimie, MIPEGE...)  
<http://www.ed-chimie.u-psud.fr/spip.php?breve154>