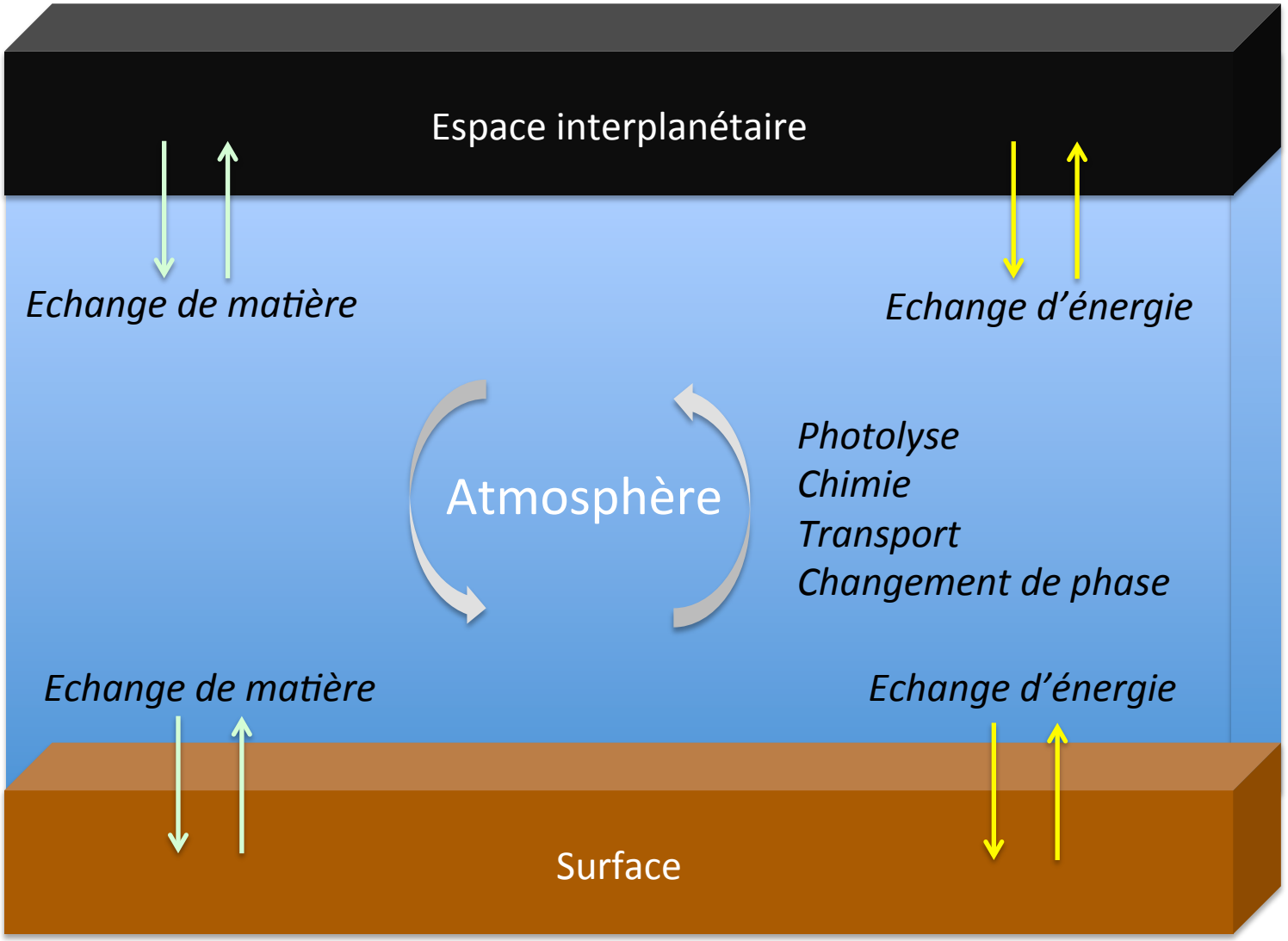
Propagation d'incertitudes et étude de sensibilité dans les modèles d'atmosphères planétaires

Michel Dobrijevic

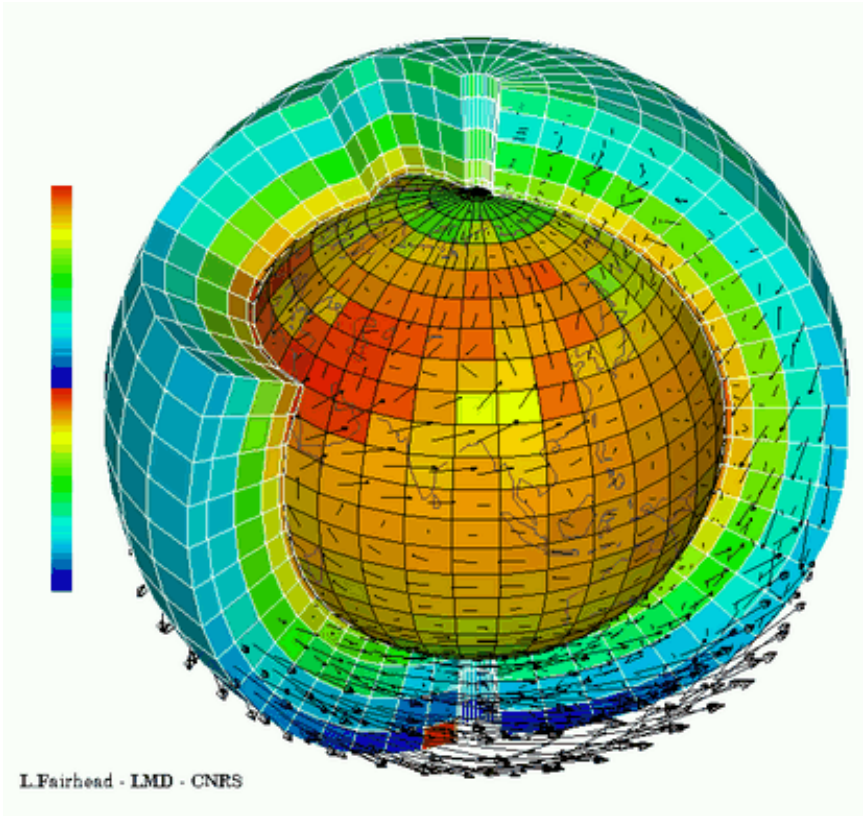
Laboratoire d'Astrophysique de Bordeaux

L’atmosphère, un système complexe

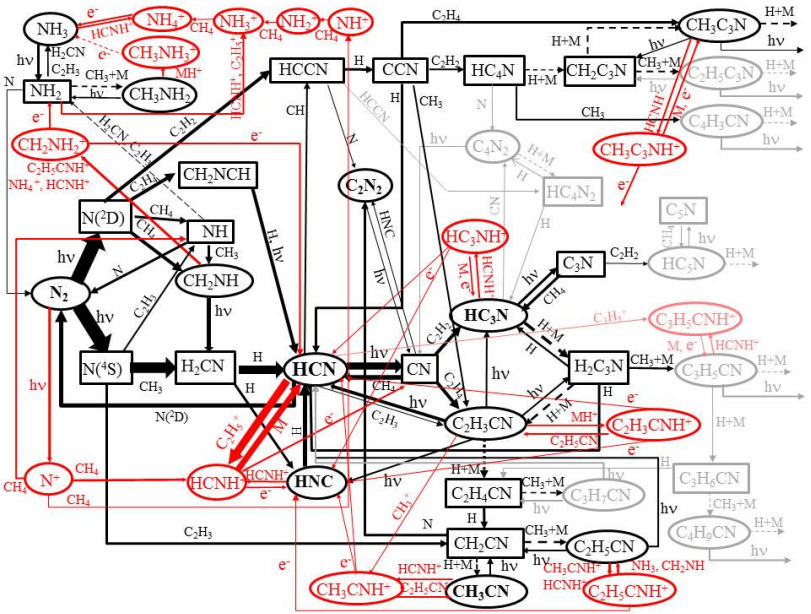


Modélisation : différents types de modèles

Modèle météorologique ou climatique



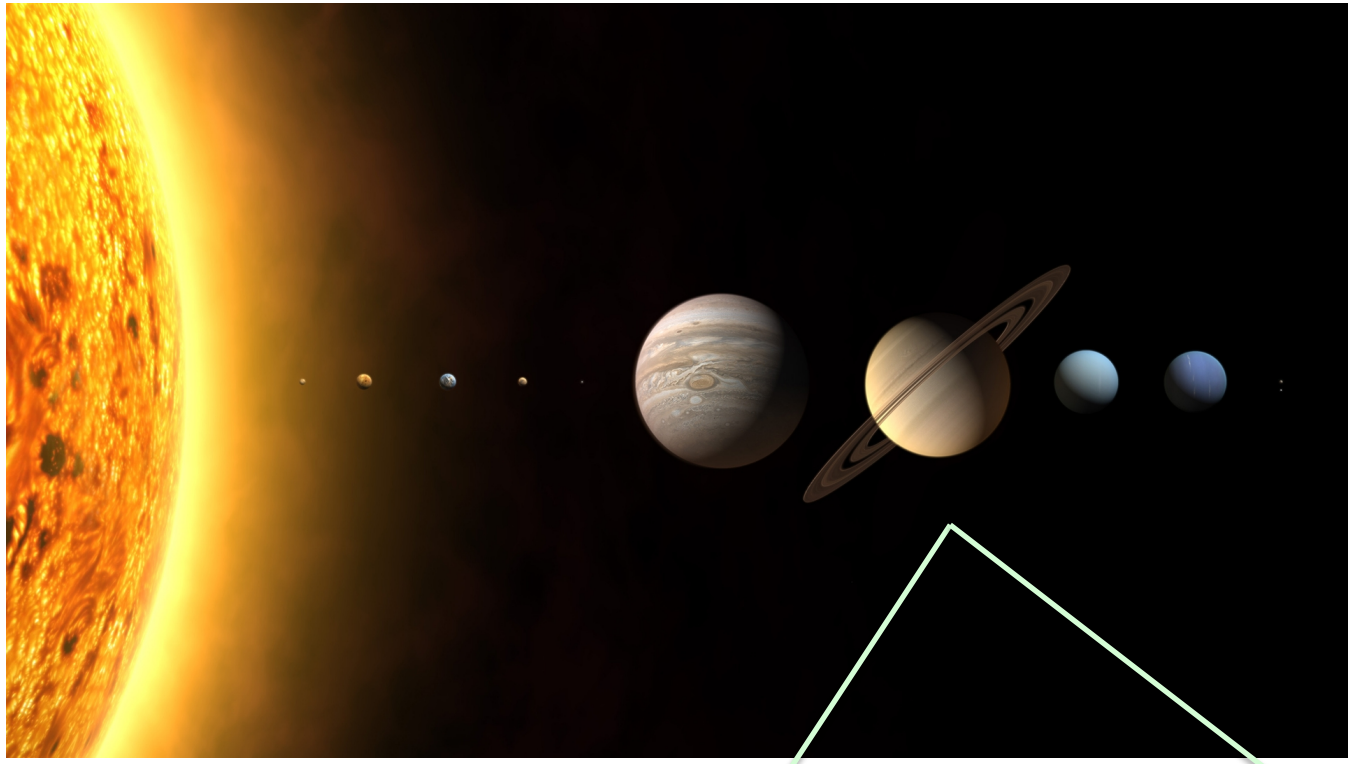
Modèle photochimique

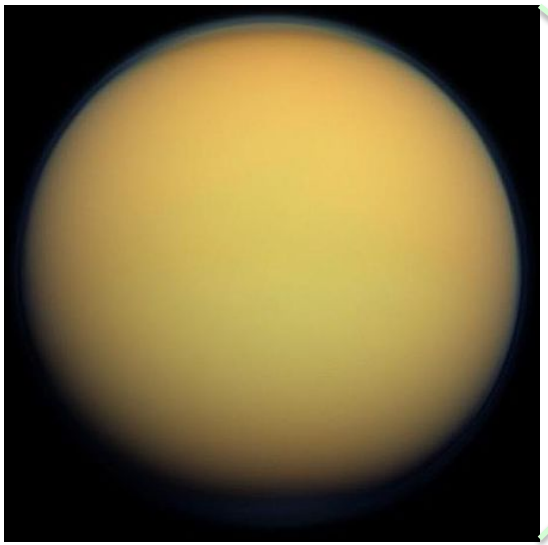


Plus de physique

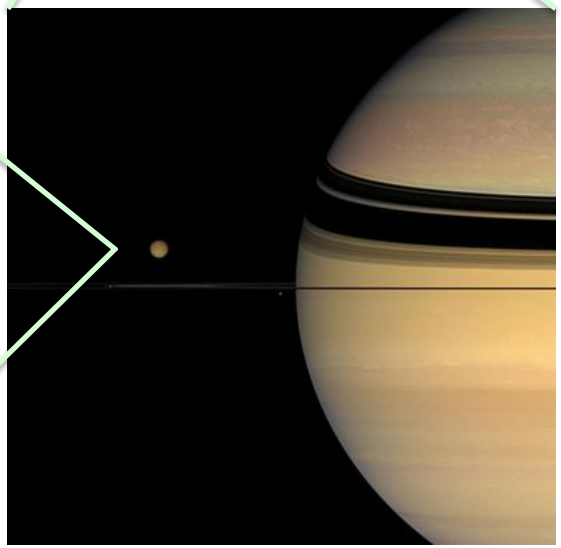
Plus de chimie

Titan, un satellite remarquable !



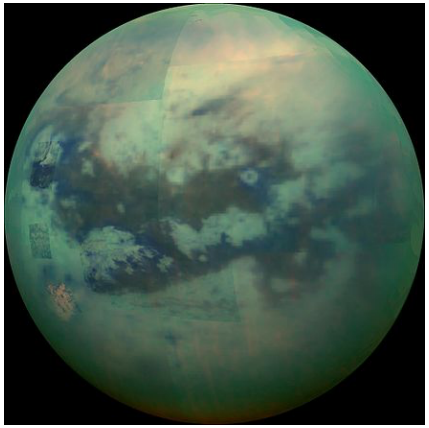


Titan dans le visible (Cassini, NASA)

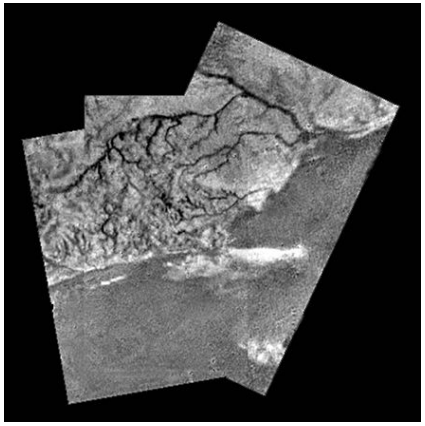


Titan : un satellite de Saturne (Cassini, NASA)

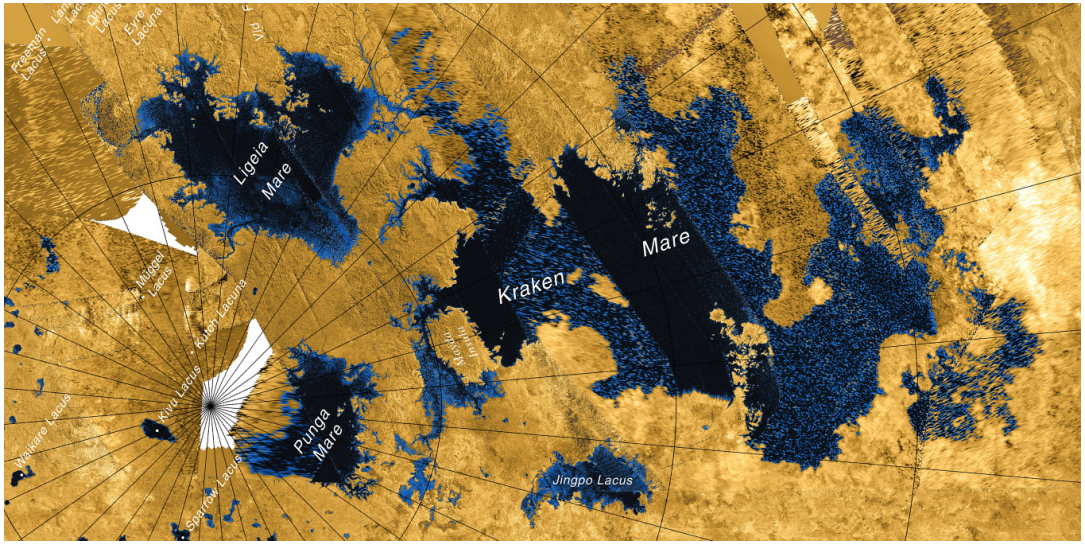
Des lacs d'hydrocarbures à la surface



Titan en Infra-rouge (Cassini, NASA)

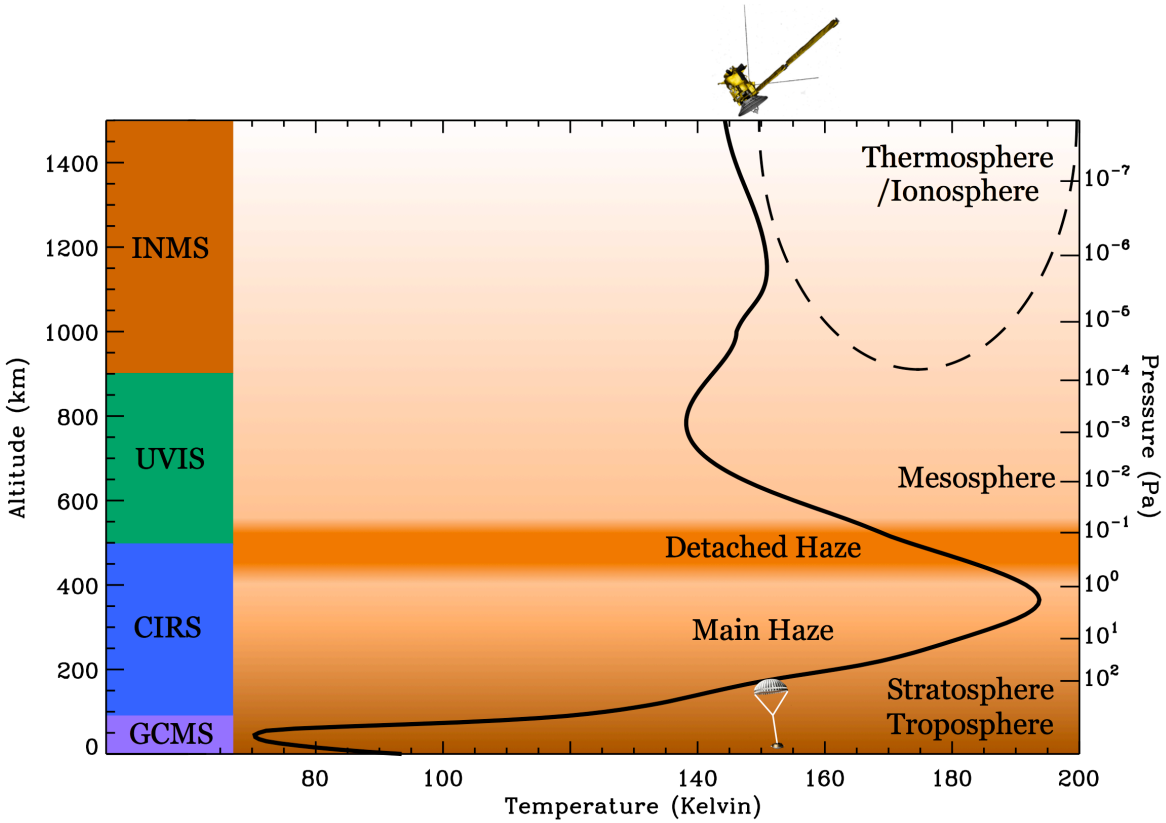


ESA Huygens DISR camera



Mers et lacs d'hydrocarbures au niveau des pôles
(Radar images from the NASA/ESA Cassini spacecraft)

L'atmosphère de Titan



Sarah M. Hörst

Atmosphère

<u>Conditions à la surface</u>

$P = 1.5$ bar

$T = 93$ K

<u>Composition</u>

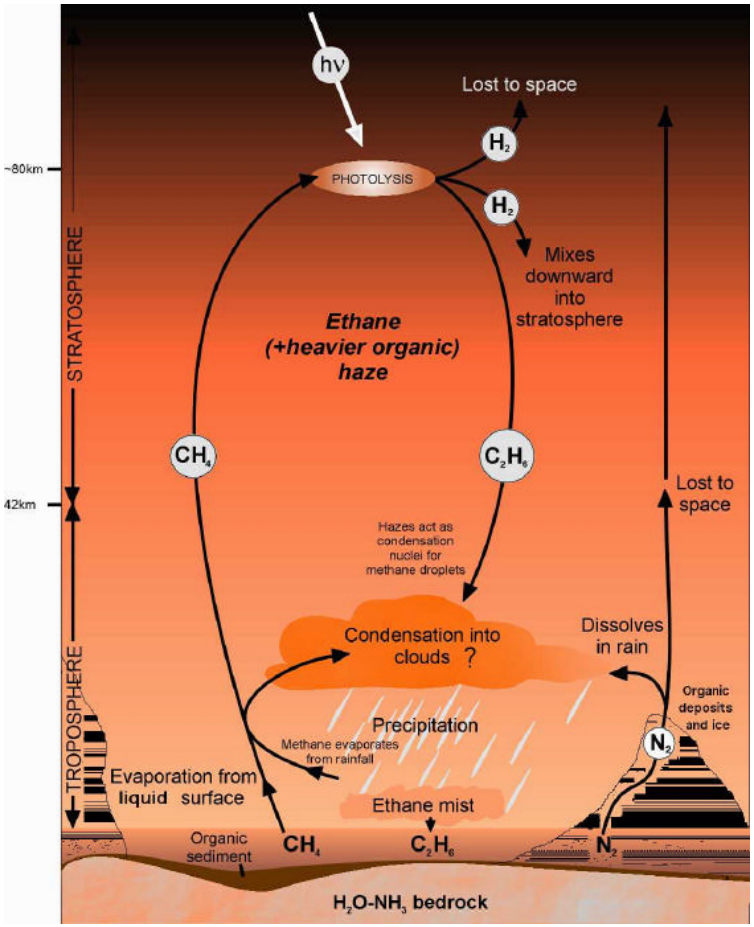
Azote N_2 : 95%

Méthane CH_4 : 4.9%

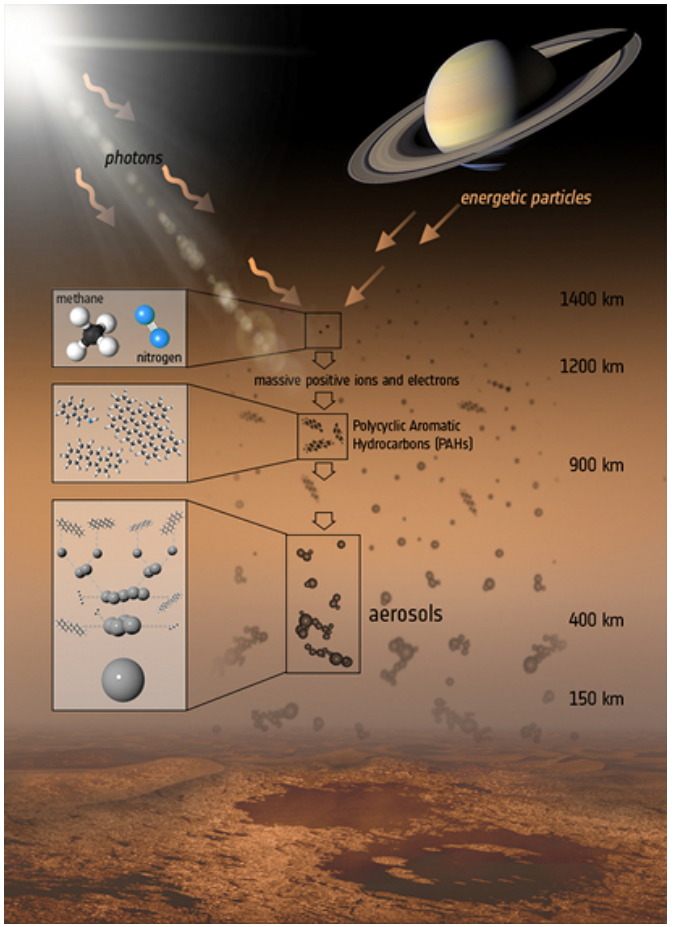
Hydrogène H_2 : 0.1%

Gravité : 1.352 $m.s^{-2}$

Titan : un réacteur chimique à l'échelle d'une planète



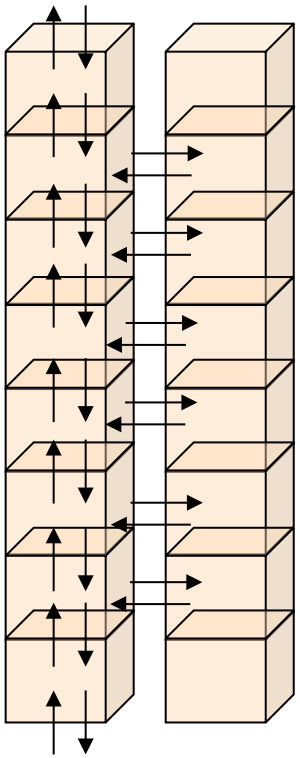
De nombreux composés condensent dans la basse atmosphère



Des aérosols sont produits chimiquement dans la haute atmosphère

Plusieurs mètres de HCN, C_2H_2, C_6H_6, etc. se sont accumulés à la surface !

Modélisation physique et chimique de l'atmosphère de Titan



Pour chaque espèce i, à chaque niveau d'altitude :

Abondance

$$\frac{dy_i}{dt} = \frac{P_i}{n} - y_i L_i - \frac{div(\overrightarrow{\Phi_i})}{n}$$

+ autres processus (condensation, apports et pertes aux limites,…)

Chimie

Transport vertical (et horizontal)

180 espèces et 140 niveaux d'altitude (1200 réactions chimiques)

→ Environ **25 000** équations différentielles non-linéaires et couplées

La photo-chimie

Des centaines (voire des milliers) de réactions :

- Photodissociation : $A + hn \rightarrow B + C$
- Photoionisation : $A + hn \rightarrow A^+ + e^-$
- Réactions neutres : $A + B \rightarrow C + D$ (*)
- Réactions ions-neutres : $A^+ + B \rightarrow C^+ + D$
- Recombinaison dissociatives : $A^+ + e^- \rightarrow B + C$
- etc.

Les constantes de réaction sont mesurées en laboratoire ou calculées via des modèles théoriques

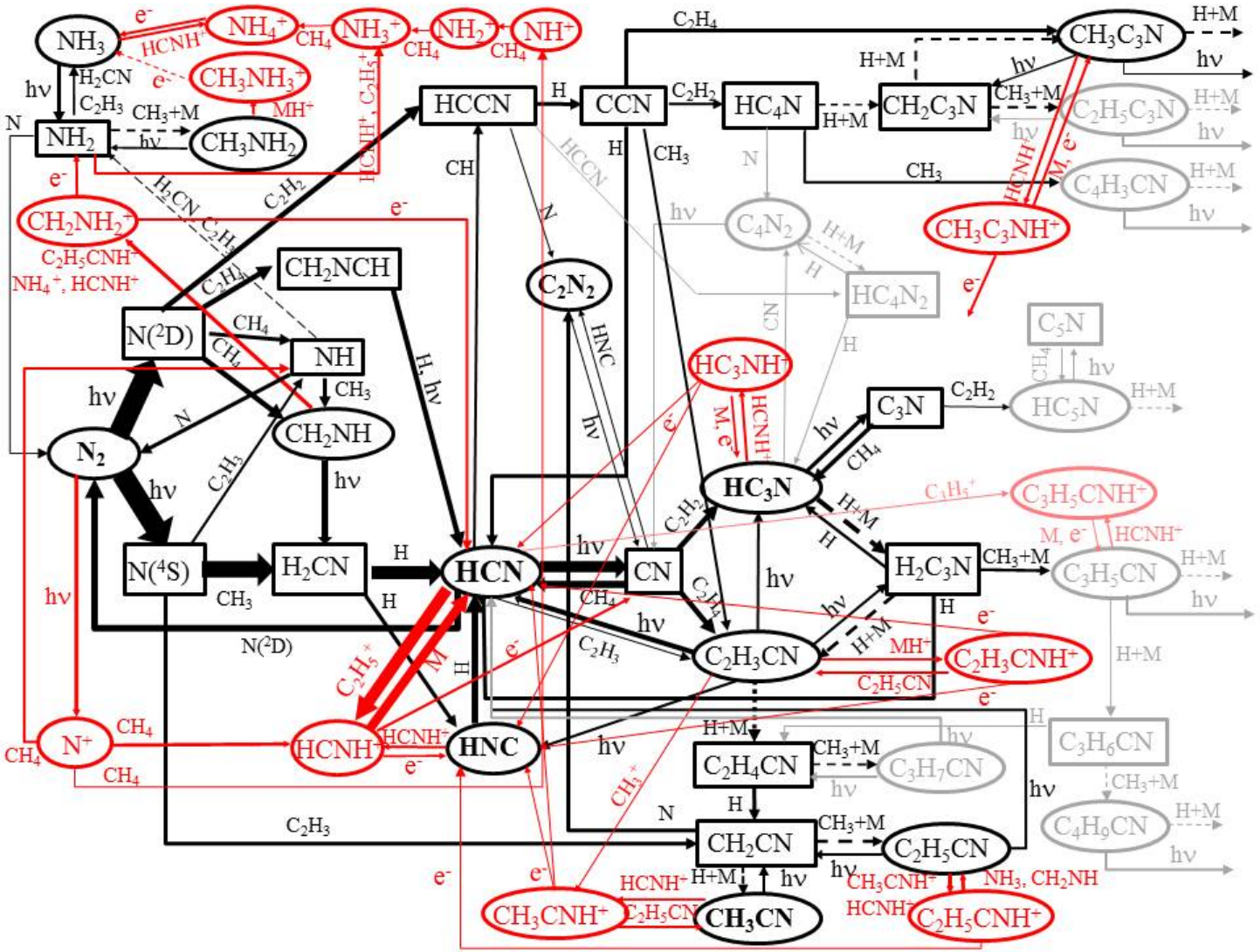
Constante de réaction

$$^{(*)}\, P_C = k \times [A] \times [B] + \ldots$$

Taux de production du produit C

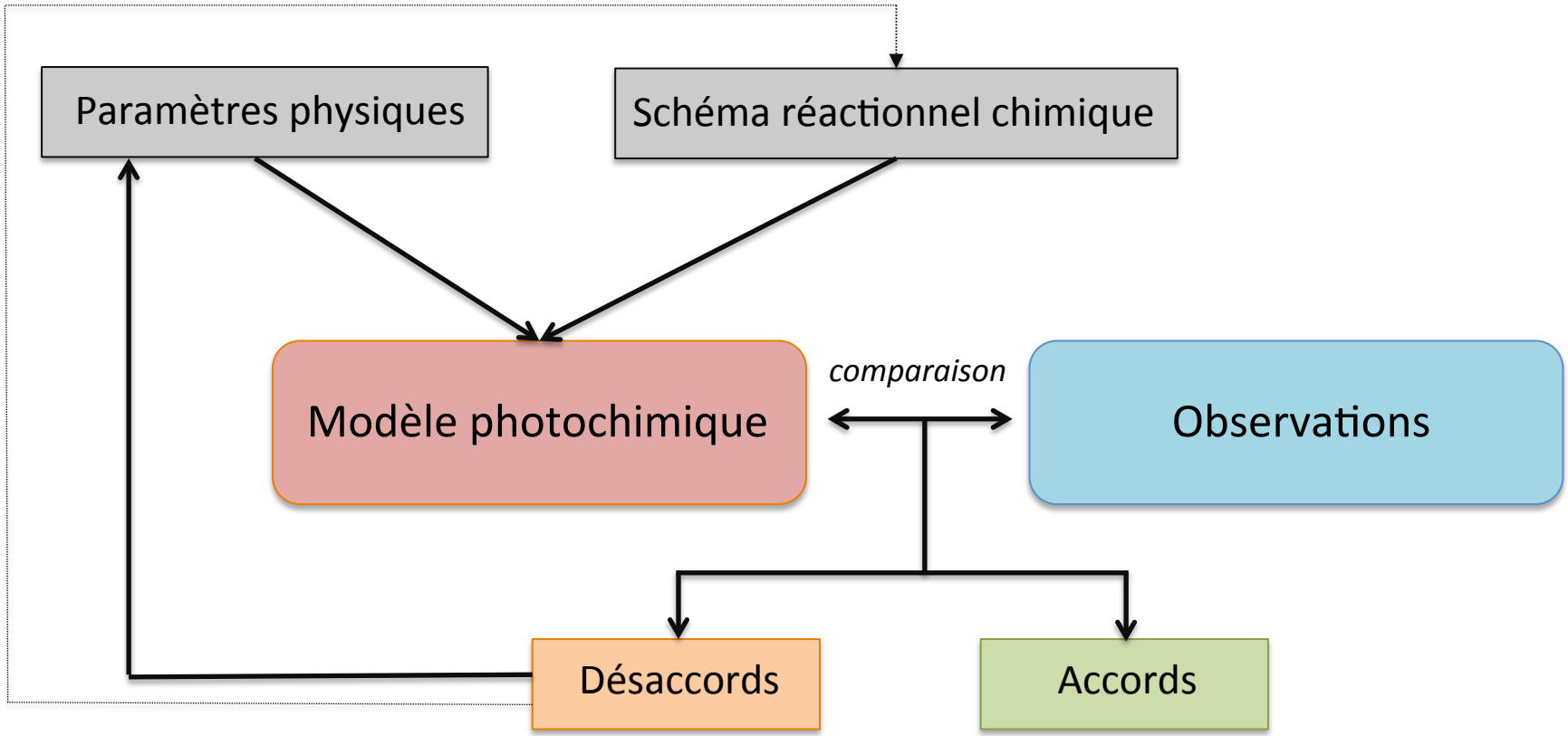
Concentrations des réactants

Chimie des composés azotés



Dobrijevic et al. 2016

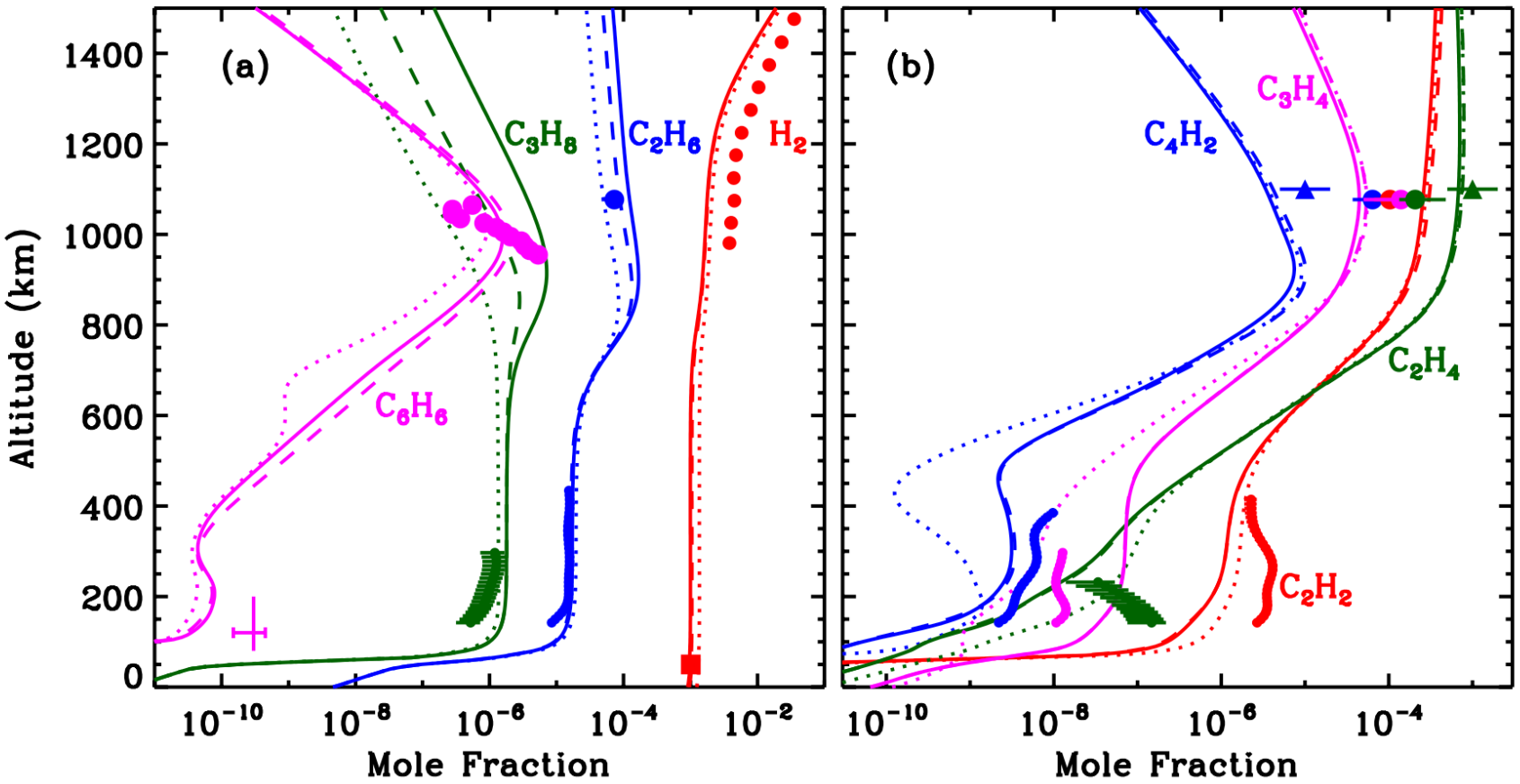
Méthodologie répandue



Implicitement, la chimie est considérée connue,
mais pas tout le temps !?

Exemple

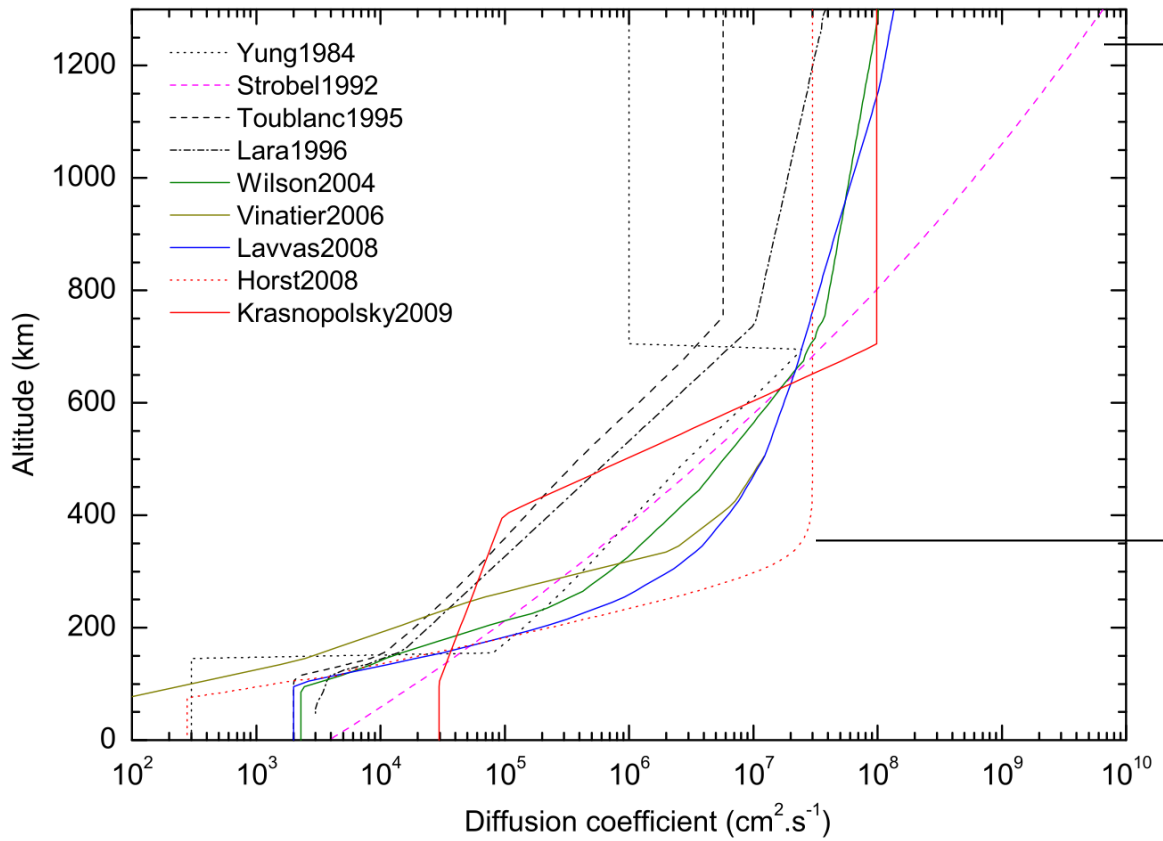
Abondances et étude de sensibilité locale



→ Comparaison entre les résultats du modèle et les observations

Vuitton et al. 2012

Des années de tentative de détermination du paramètre physique $K(z)$



$K_{\text{Strob}} = 6000 \times K_{\text{Yung}}$

$K_{\text{Horst}} = 300 \times K_{\text{Krasno}}$

A chaque nouvelle étude, on a une nouvelle détermination de ce paramètre physique…

Une chimie imprécise et mal connue

Incertitudes sur les données

→ Des mesures très dispersées (expériences difficiles)

→ Des mesures obtenues dans des conditions inadaptées (température, gaz porteur)

Informations manquantes

→ Un grand nombre de réactions non étudiées

→ Beaucoup d'espèces non étudiées (isomères, espèces lourdes, radicaux très réactifs)

Exemple 1 : $CH_3 + CH_3 + M \rightarrow C_2H_6 + M$

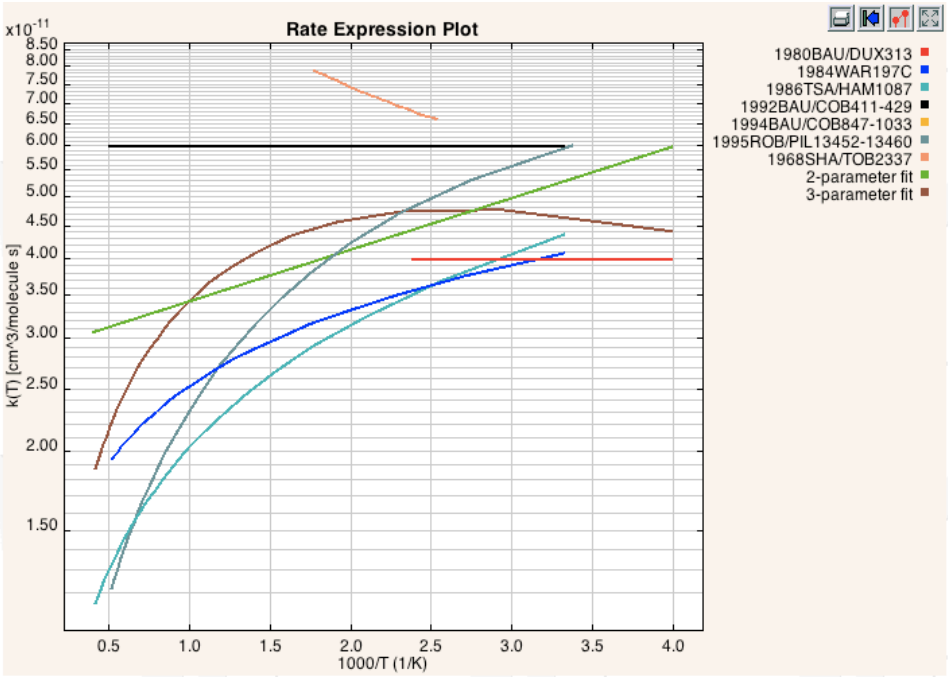
Search Results

Kinetics Database

Click on a link in the table below to see detail on the selected reaction.

Records	Reaction
1 records matched	$\cdot CH_3 + CH_3CO \rightarrow C_2H_6 + CO$
119 records matched	$\cdot CH_3 + \cdot CH_3 \rightarrow C_2H_6$
1 records matched	$(CH_3)3Ga + \cdot CH_3 \rightarrow C_2H_6 + \cdot CH_3 + CH_3Ga$
2 records matched	$(CH_3)4Sn + \cdot CH_3 \rightarrow C_2H_6 + (CH_3)3Sn$
1 records matched	$(CH_3)2Hg + \cdot CH_3 \rightarrow C_2H_6 + \cdot CH_3 + Hg$
1 records matched	$C_2H_6 + \cdot CH_3 \rightarrow C_2H_6 + \cdot CH_3$
3 records matched	$CH_4 + \cdot CH_3 \rightarrow C_2H_6 + H\cdot$
1 records matched	$GaMe_2NH_2 + \cdot CH_3 \rightarrow GaNH_2 + C_2H_6 + \cdot CH_3$

Search returned 129 records.



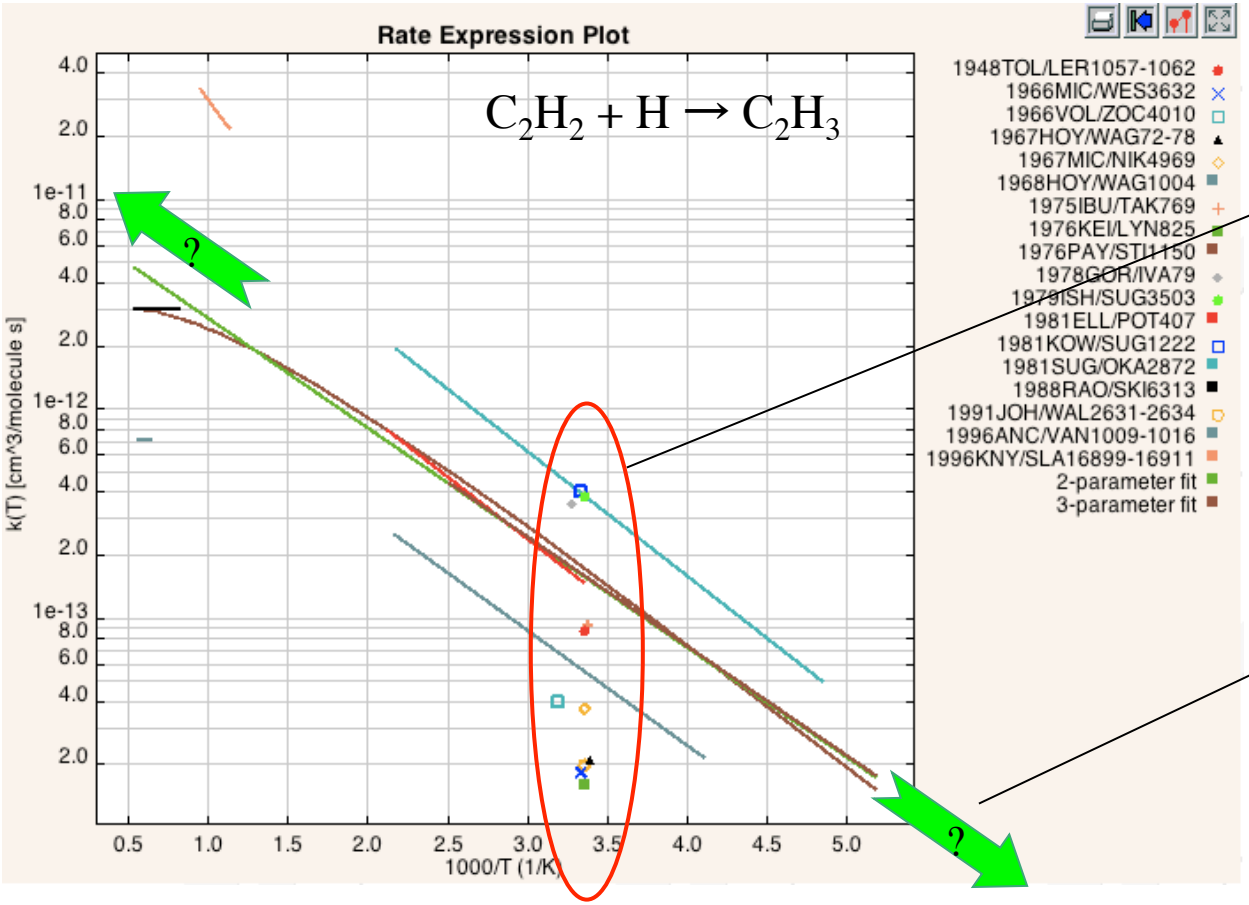
⇒ Quelle constante choisir et comment ?

<u>Plusieurs facteurs pouvant expliquer les différences</u>

- Temperature
- Gaz porteur
- Technique utilisée
- Date des expériences (?)
- Expertise de l'équipe (?)
- etc.

Exemple 2 : $H + C_2H_2 + M \rightarrow C_2H_3 + M$

La dispersion est plus grande que les incertitudes !



Facteur 20 entre les mesures…
Quelle est la valeur nominale ?
Quel est le facteur d'incertitude ?

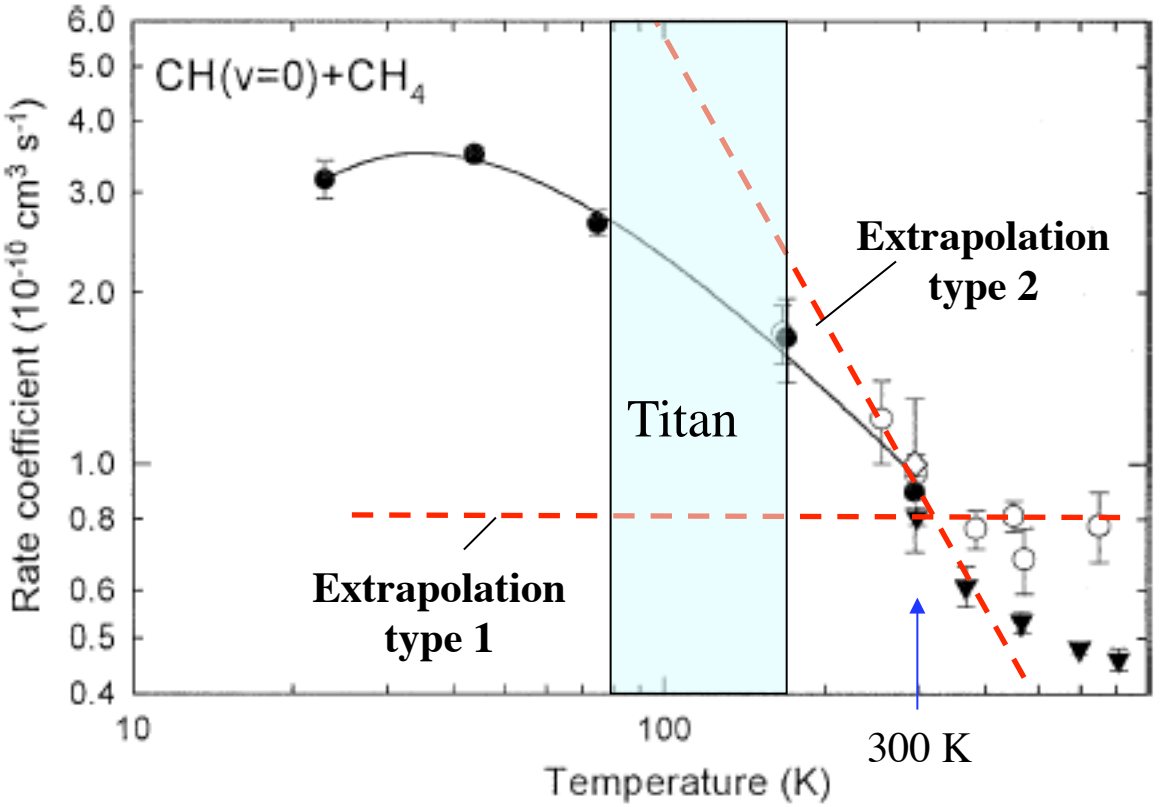
Comment extrapoler à basse température ?

Kinetics Database

Constantes de réaction à basse température

- La plupart des réactions sont mesurées à 300 K.
- L'extrapolation à basse température n'est pas triviale.

Exemple:
Mesure de CH + CH_4
à basse
température



Canosa et al. 1997

Quels effets sur les modèles ?

Choix des réactions

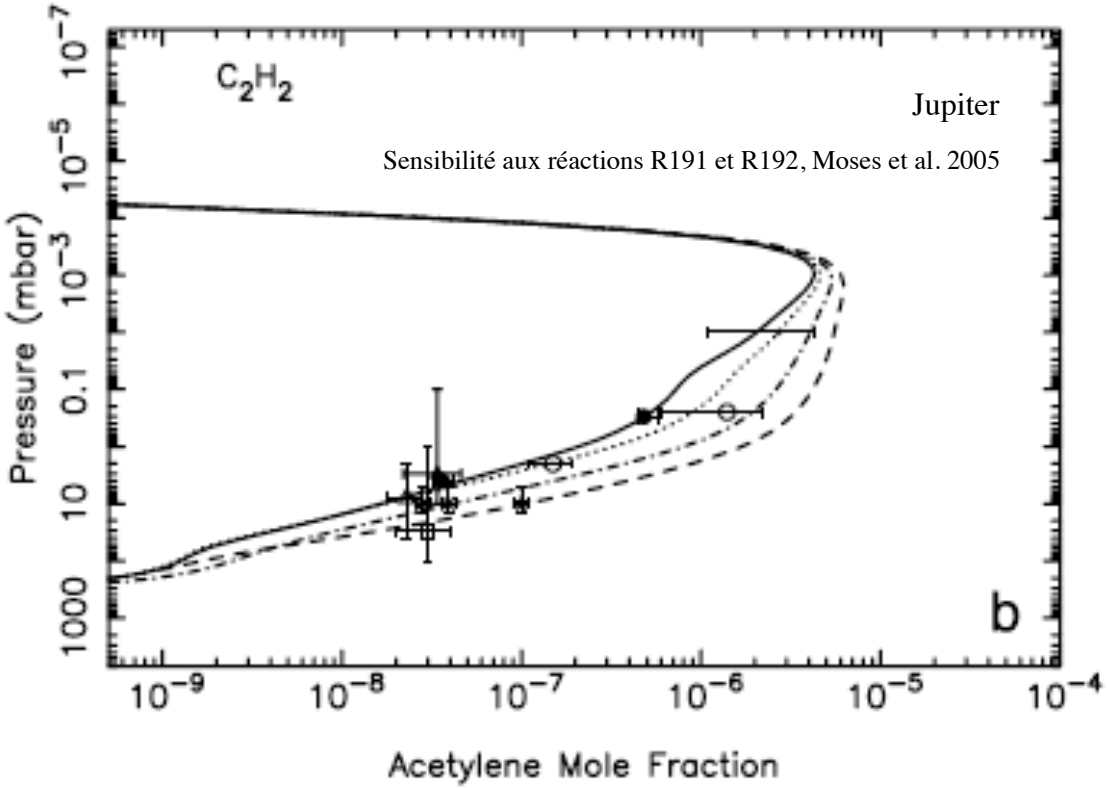
→ On calcule les taux de production relatifs (et les taux de perte relatifs) de chaque espèce du modèle

Méthodologie

→ On fait varier chaque paramètre individuellement d'un certain facteur et on regarde son impact sur les résultats

Méthode "One-factor-at-a-time"

Etude de sensibilité locale



➜ Cette méthode ne donne qu'une information partielle…
Elle ne prend pas en compte la non-linéarité et les couplages du système !

Comment tenir compte du couplage des espèces et de la non-linéarité et du système ?

$$\frac{d[X_1]}{dt} = f_1\left([X_1](t), [X_2](t), ..., [X_n](t)\right)$$

$$\frac{d[X_2]}{dt} = f_2\left([X_1](t), [X_2](t), ..., [X_n](t)\right)$$

$$...$$

$$\frac{d[X_n]}{dt} = f_n\left([X_1](t), [X_2](t), ..., [X_n](t)\right)$$

→ En faisant varier tous les facteurs en même-temps !

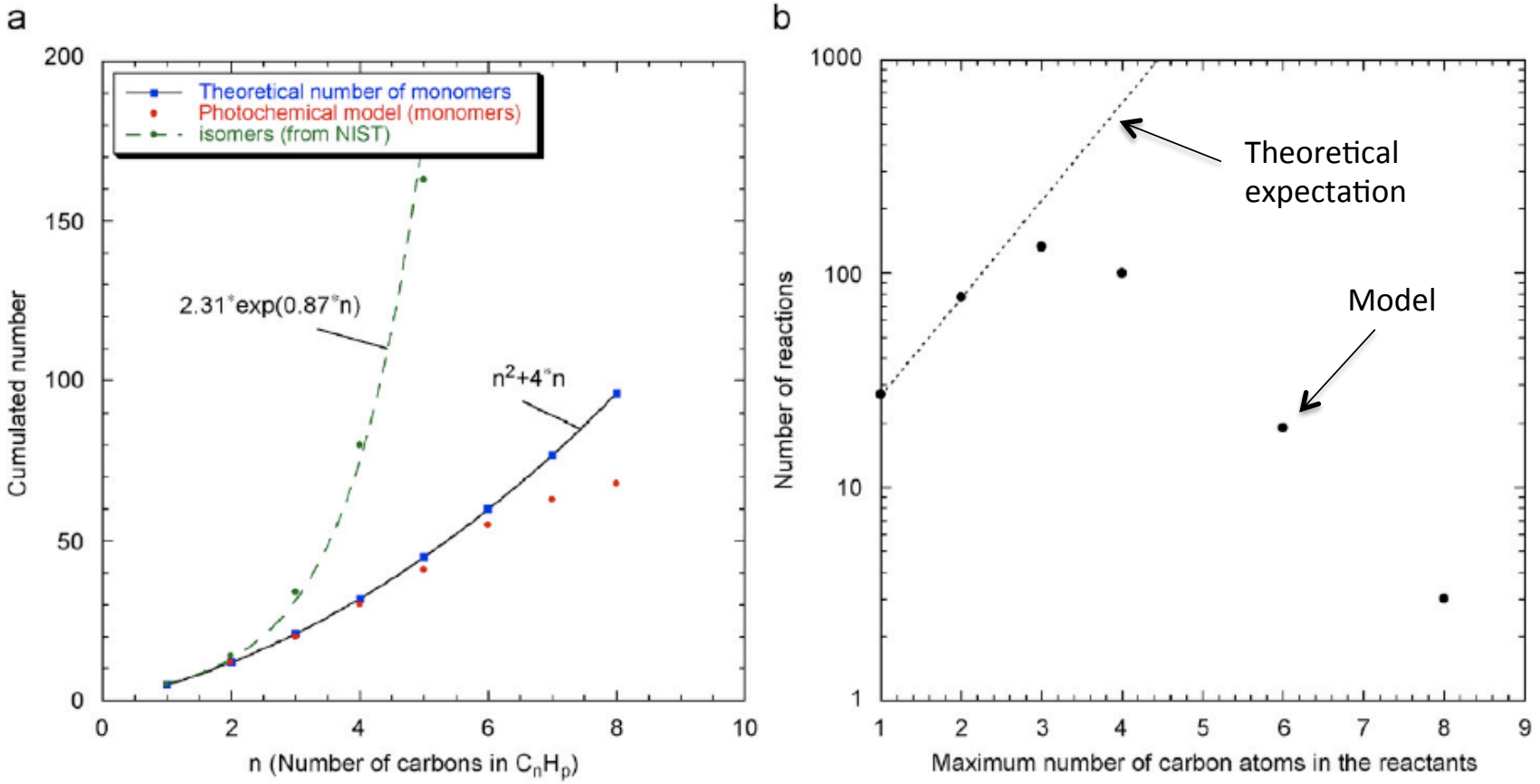
Propagation des incertitudes sur les paramètres du modèle (facteurs d'entrée)

- Ce que l'on souhaite faire :
  - Tenir compte du caractère non linéaire et des couplages du système
  - Tenir compte des incertitudes sur les paramètres chimiques
  - Ajouter des processus pour lesquels nous n'avons que des estimations grossières

- Ce qu'il nous faut :
  - Une revue sérieuse de la chimie (par des spécialistes) avec une estimation des incertitudes pour toutes les constantes…
  - Des outils pour générer les entrées et analyser les sorties

Problème de l'incomplétude des schémas réactionnels

De nombreuses réactions et de nombreuses espèces ne sont pas introduits dans les modèles par manque d'informations !



Comment compléter les schémas réactionnels ?

1. On ajoute des réactions dont on ne sait pas grand chose (pas d'étude expérimentale, pas d'étude théorique poussée)

2. On regarde l'effet sur les résultats du modèle.
   - S'il y a un effet, il faut l'ajouter au système.
   - Sinon, on ne s'en occupe pas (?)

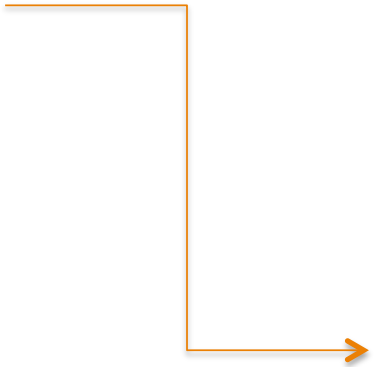
Attention :
- il faut tenir compte de l'incertitude de sa constante !
- Il faut tenir compte des couplages non-linéaires du système !

→ Une réaction peut être importante de manière subtile…

Mon expérience sur la détermination des facteurs clefs

Les principales méthodes utilisées :

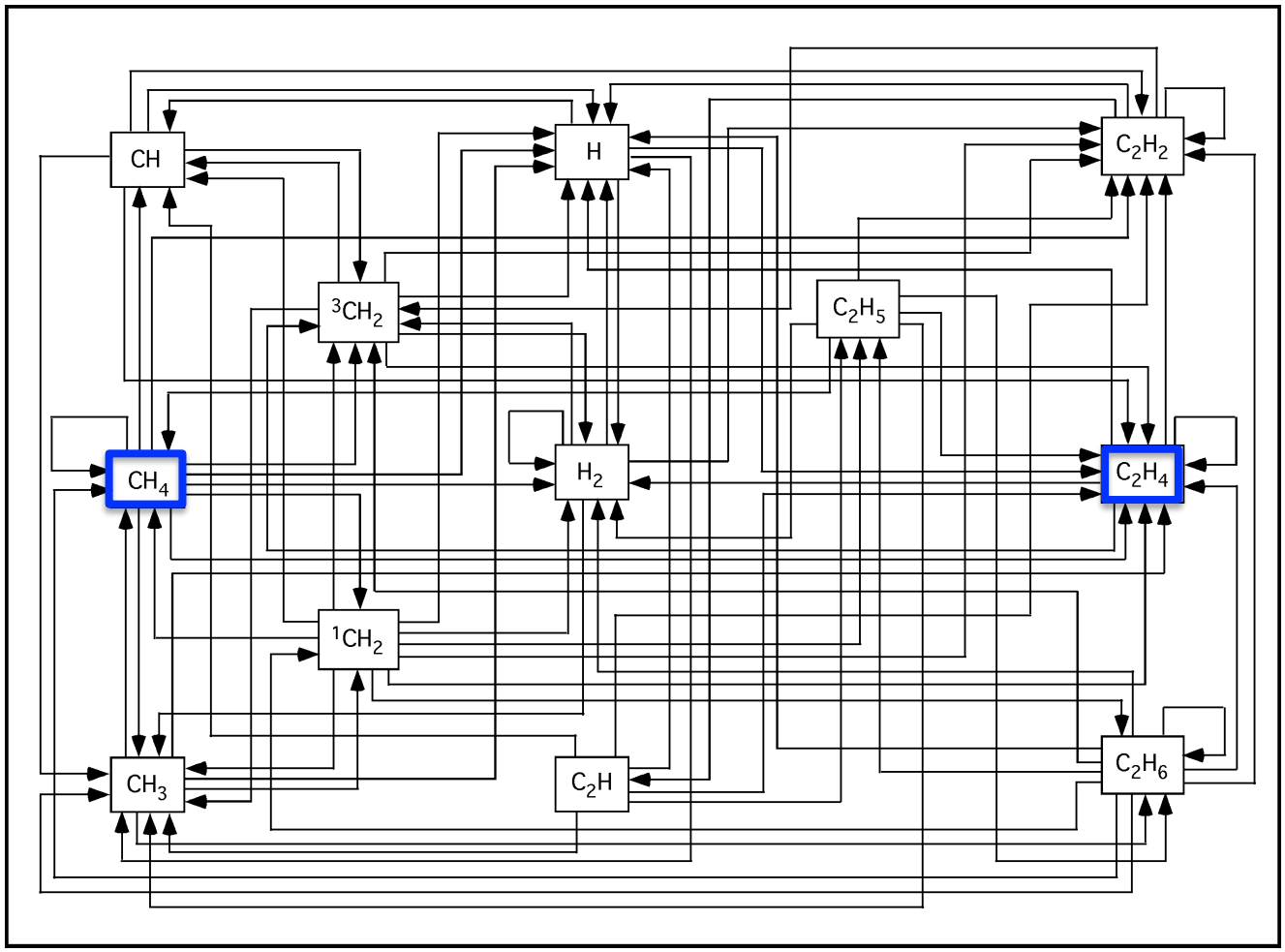
- Analyse locale
- Graphes à transmission de signaux
- **Monte-Carlo**
- Plans d'expériences
- Latin Hyper-cube



Collaboration
avec l'équipe M^2

Les graphes à transmission de signaux

Schéma réactionnel de 12 espèces et 48 réactions



12 nœuds et 75 branches (dont des boucles)

Il y a 11 357 chemins entre CH_4 et C_2H_4 !

→ Permet une étude quantitative exhaustive de tous les chemins !

→ Impossible à utiliser pour des systèmes plus gros !!

(Dobrijevic et al. 2010)

La méthode de sensibilité locale

Abondance de l'espèce *i* après modification de la réaction *j*.

Abondance de l'espèce *i* nominale

$$S_j^i = \frac{\log X_j^i - \log X_0^i}{\log F_j}$$

Coefficient de sensibilité de la réaction *j* sur l'espèce *i*.

Facteur d'incertitude de la réaction *j*.

→ A priori, toutes les réactions doivent être testées !

→ Les incertitudes en entrée doivent être faibles…

→ Echoue à donner les réactions clefs dans le bon ordre dans un cas général !

La méthode de propagation locale d'incertitudes

Coefficient de sensibilité de la réaction *j* sur l'espèce *i*.

$$U_j^i = \frac{\left(\log F_j\right)^2 \left(S_j^i\right)^2}{\sum_j \left(\log F_j\right)^2 \left(S_j^i\right)^2}$$

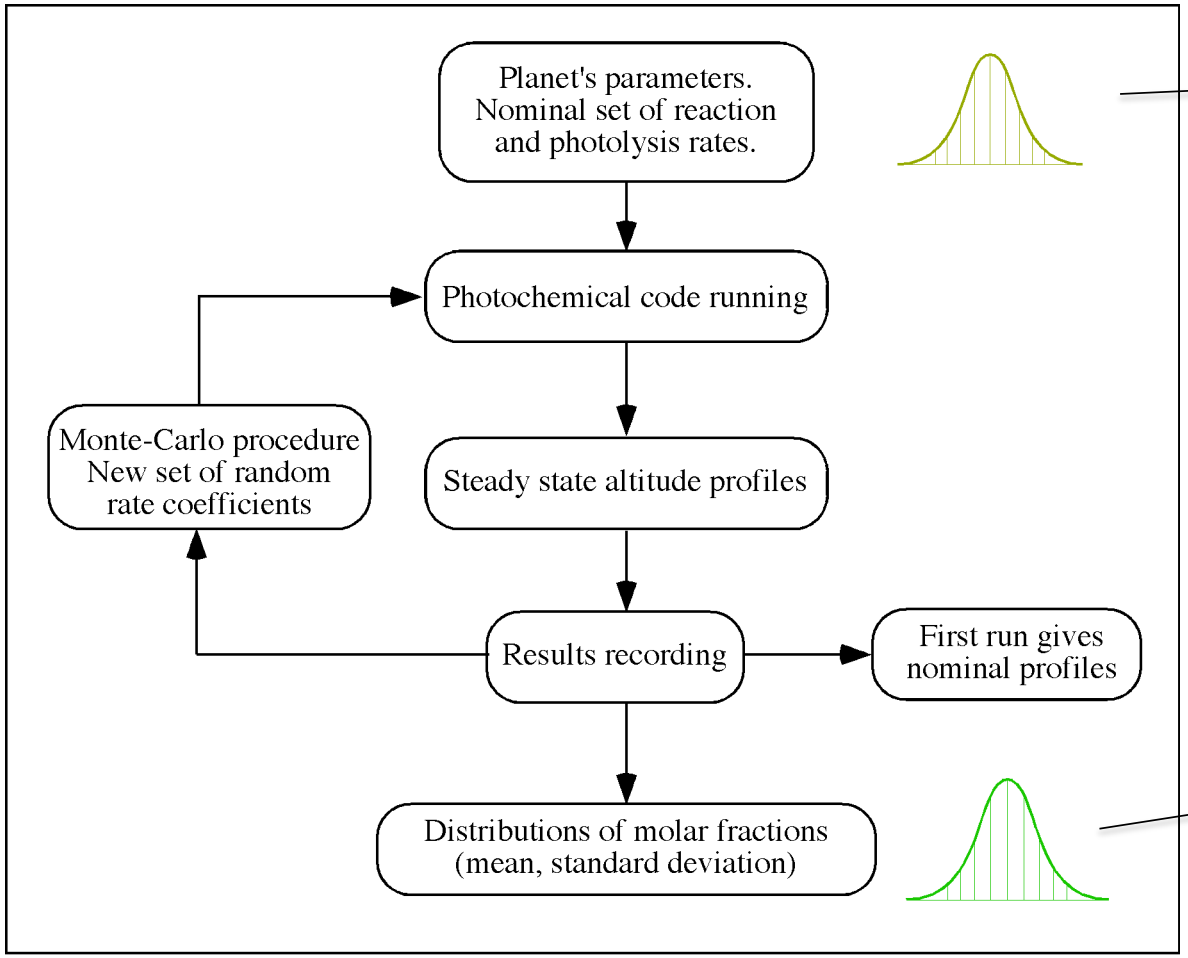
Importance relative de la réaction *j* pour l'espèce *i*.

Facteur d'incertitude de la réaction *j*.

→ Echoue à donner les réactions clefs dans le bon ordre pour les « gros » systèmes !

Une méthode globale : Monte-Carlo

Le principe est simple…



<u>Entrée</u> :

Le schéma réactionnel
- constantes de réaction recommandées sur des considérations expérimentales ou théoriques,
- estimations des facteurs d'incertitude de chaque réaction.

<u>Sortie</u> :

Distributions des abondances de chaque composé

…mais les temps de calculs sont importants

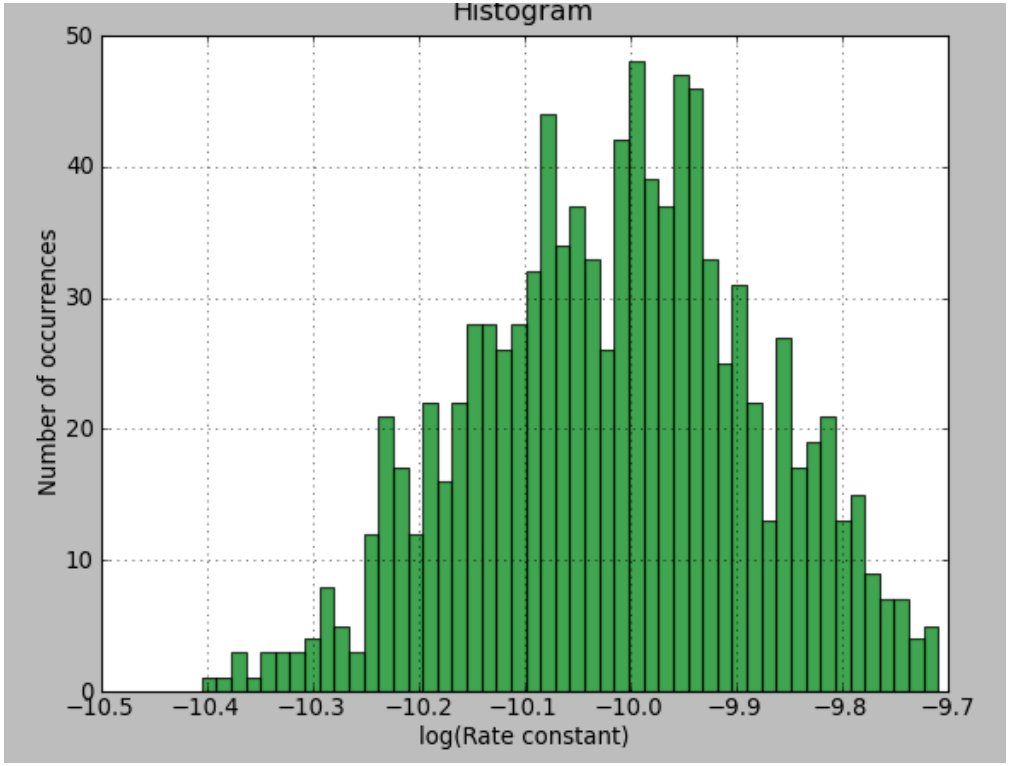
Générer des constantes de réaction aléatoirement

Variable aléatoire qui suit une distribution normale

$$\log(k_j) = \log(k_j^0) \pm \varepsilon \log(F_j)$$

Constante de la réaction *j*

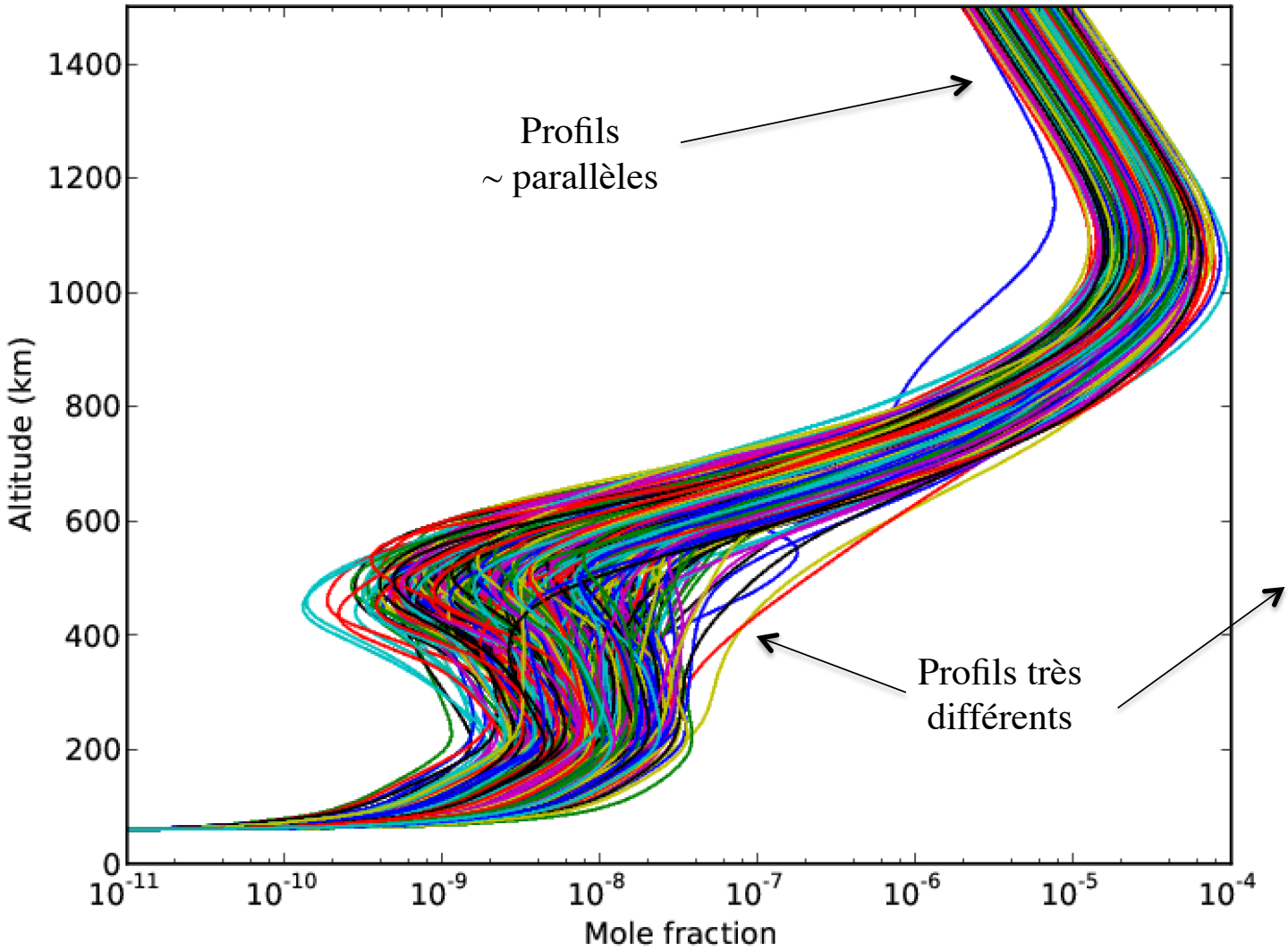
Facteur d'incertitude de la constante de réaction *j*



Rejet des valeurs > à 2σ

Exemple de résultat (1)

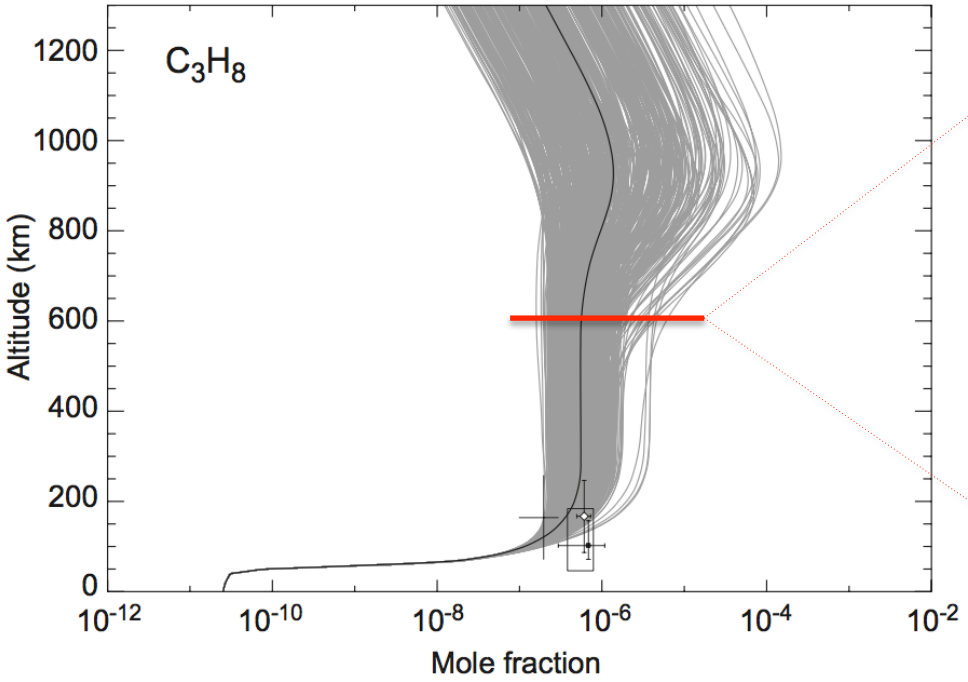
Profils des abondances en fonction de l'altitude



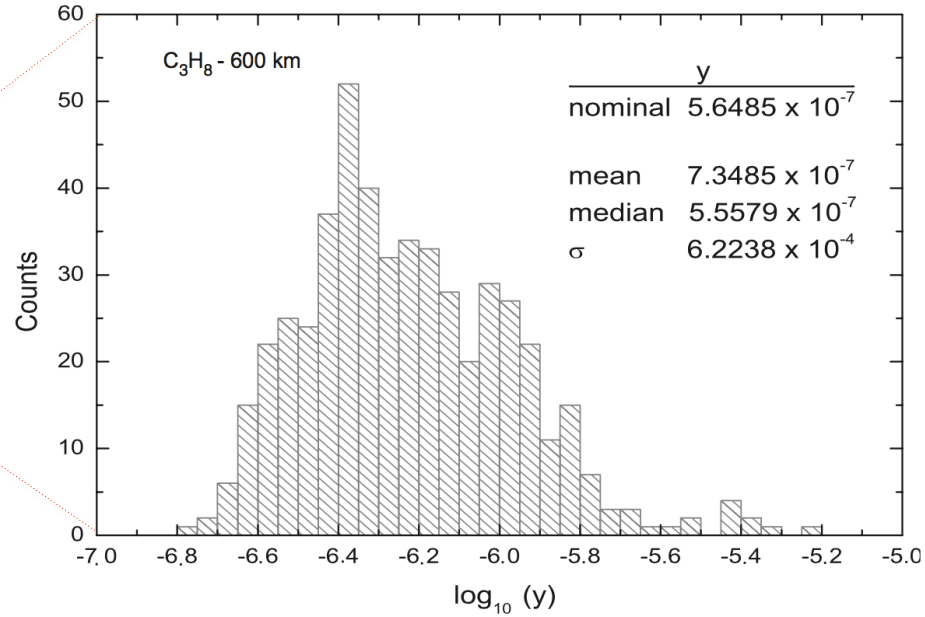
Couplage
non-linéaire
des espèces

Exemple de résultat (2)

Profils des abondances en fonction de l'altitude



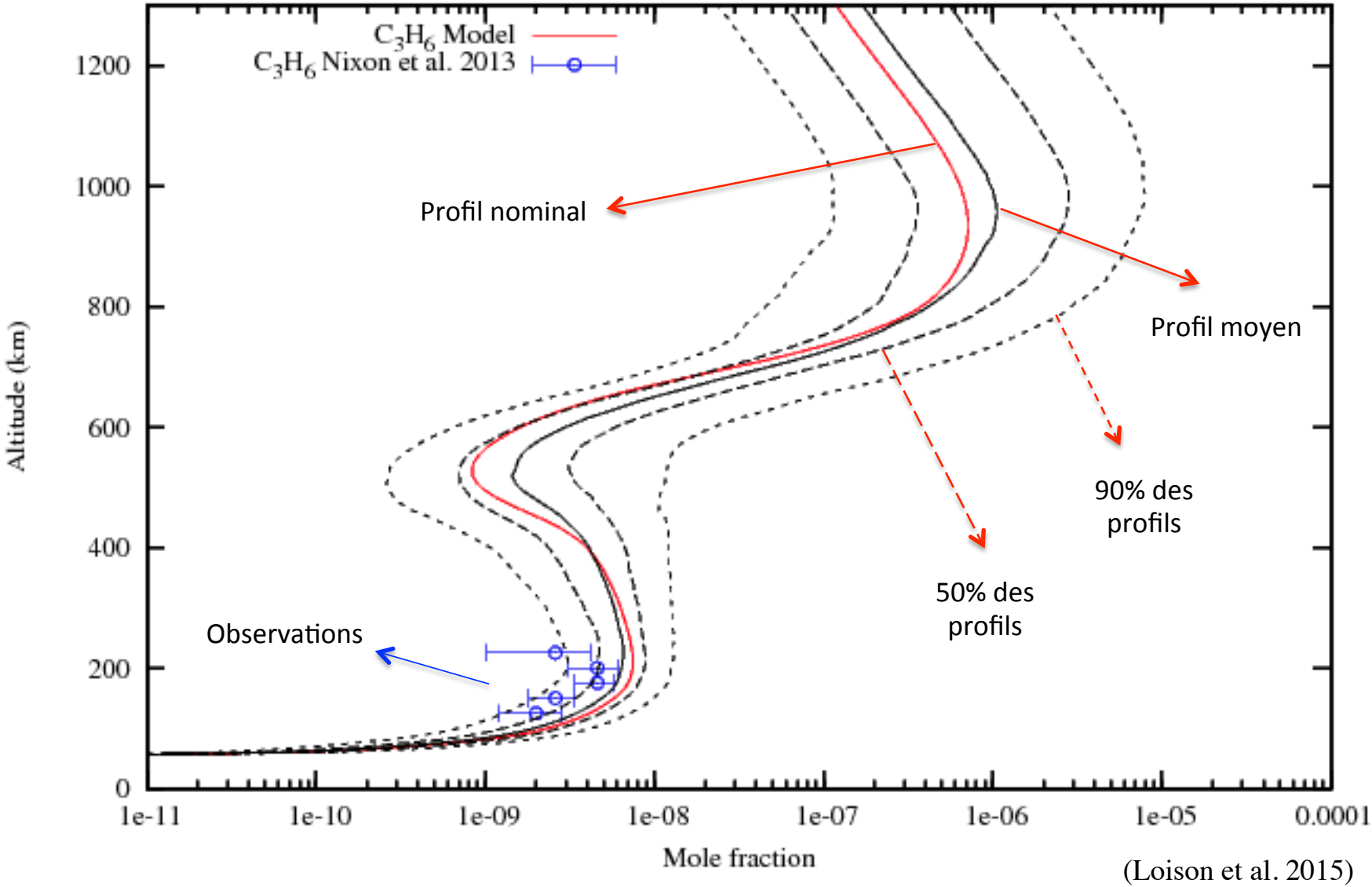
Distribution des abondances pour une altitude donnée



Calcul du facteur d'incertitude pour chaque altitude

(Hébrard et al. 2007)

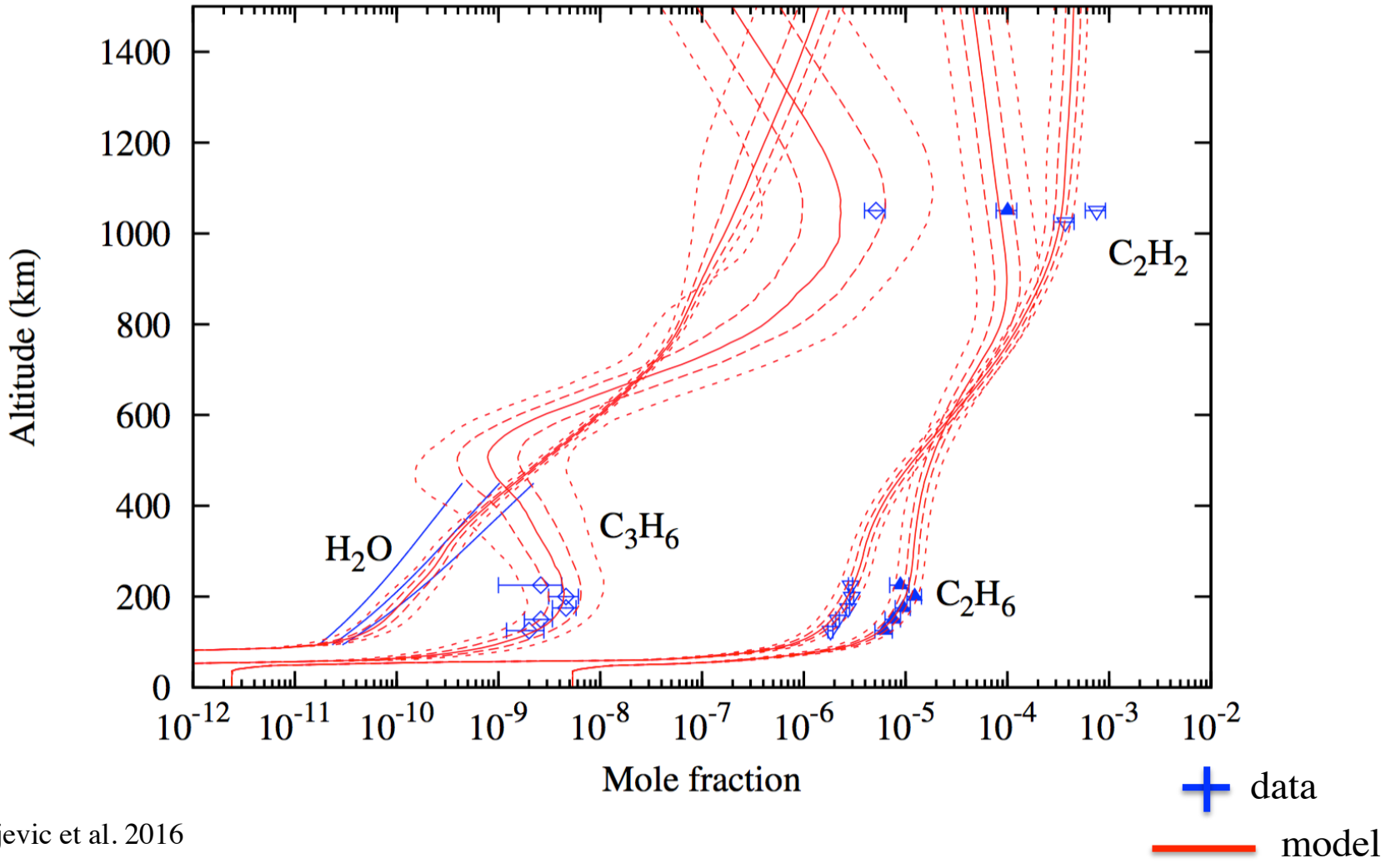
Exemple de résultat (3)



(Loison et al. 2015)

Interprétation des résultats

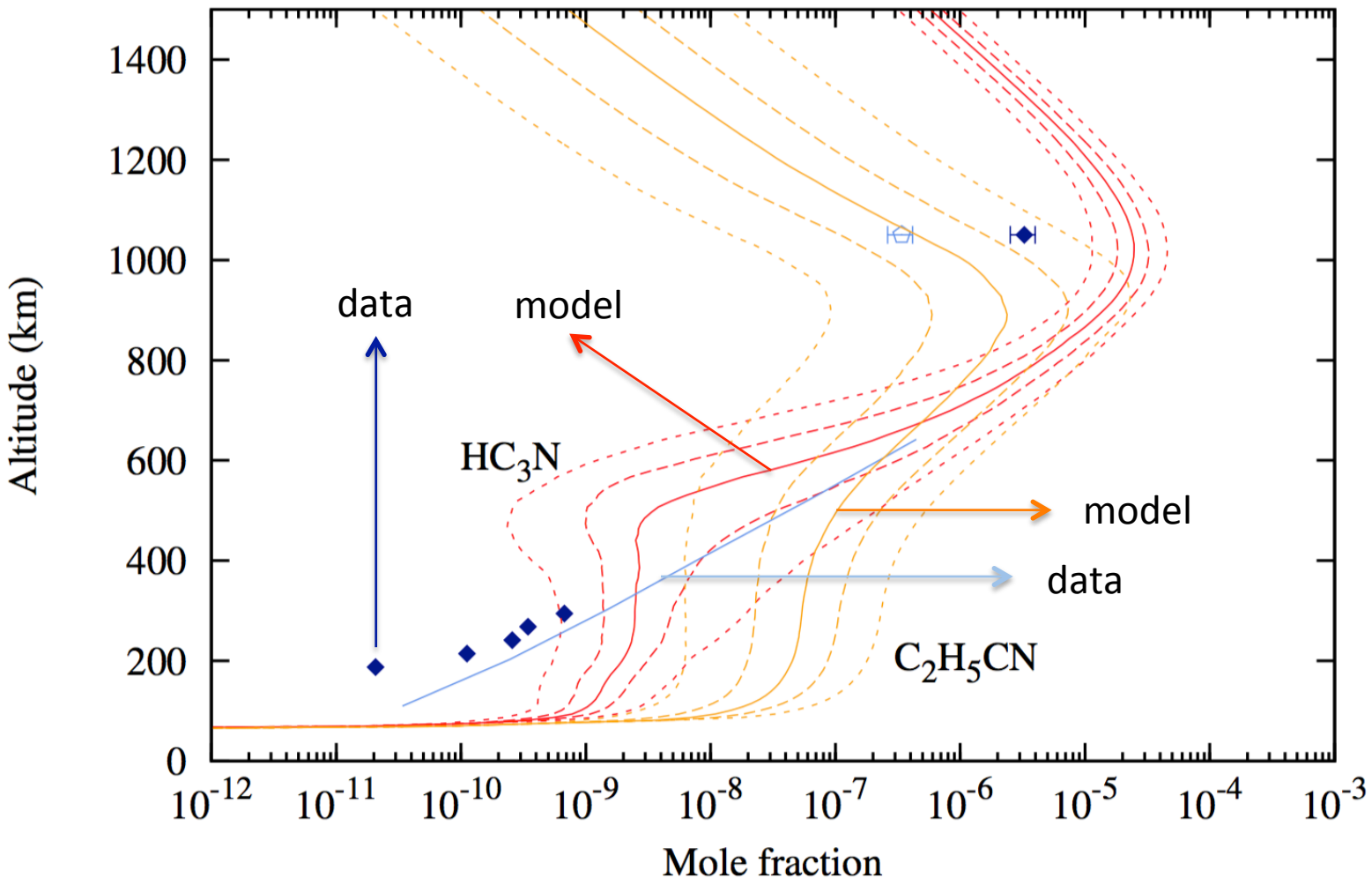
Bons accords entre le modèle et les observations



Dobrijevic et al. 2016

Interprétation des résultats (2)

Il y a aussi des désaccords…



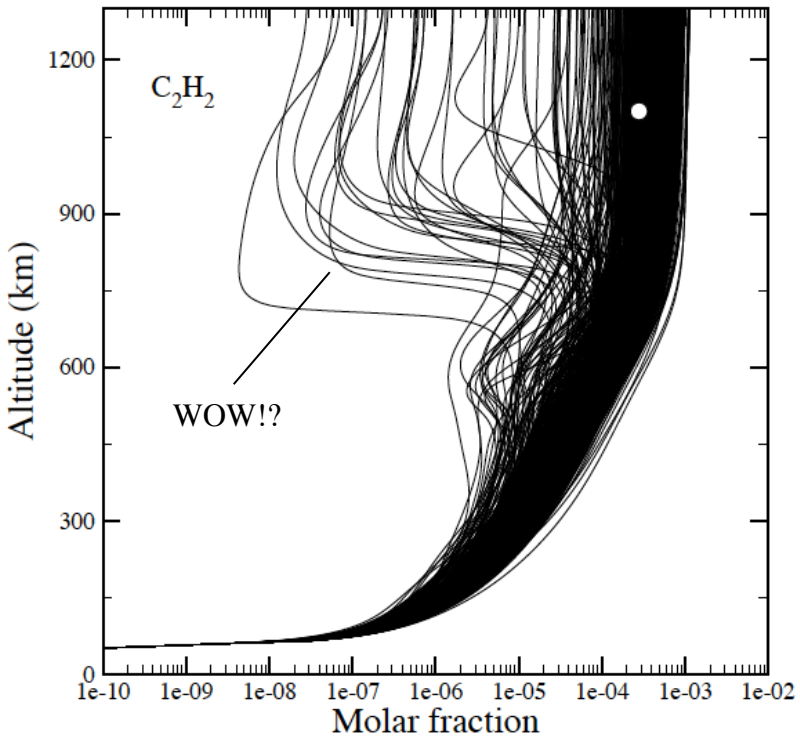
Dobrijevic et al. 2016

Bimodalités épistémiques

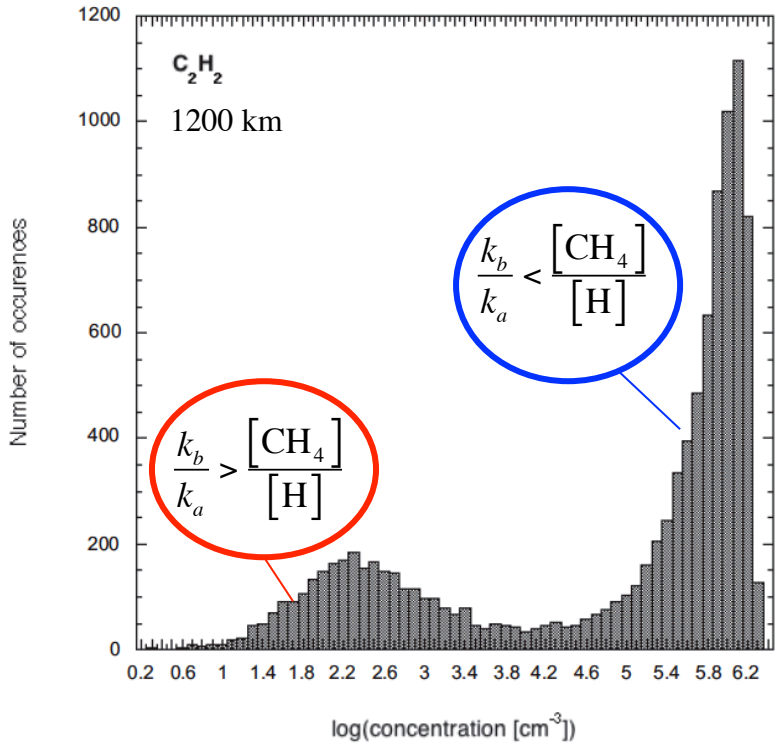
Des bimodalités apparaissent à cause de la mauvaise connaissance de certaines constantes réactionnelles !

(Dobrijevic et al. 2008)

1D model: 600 runs



0D model: 10 000 runs



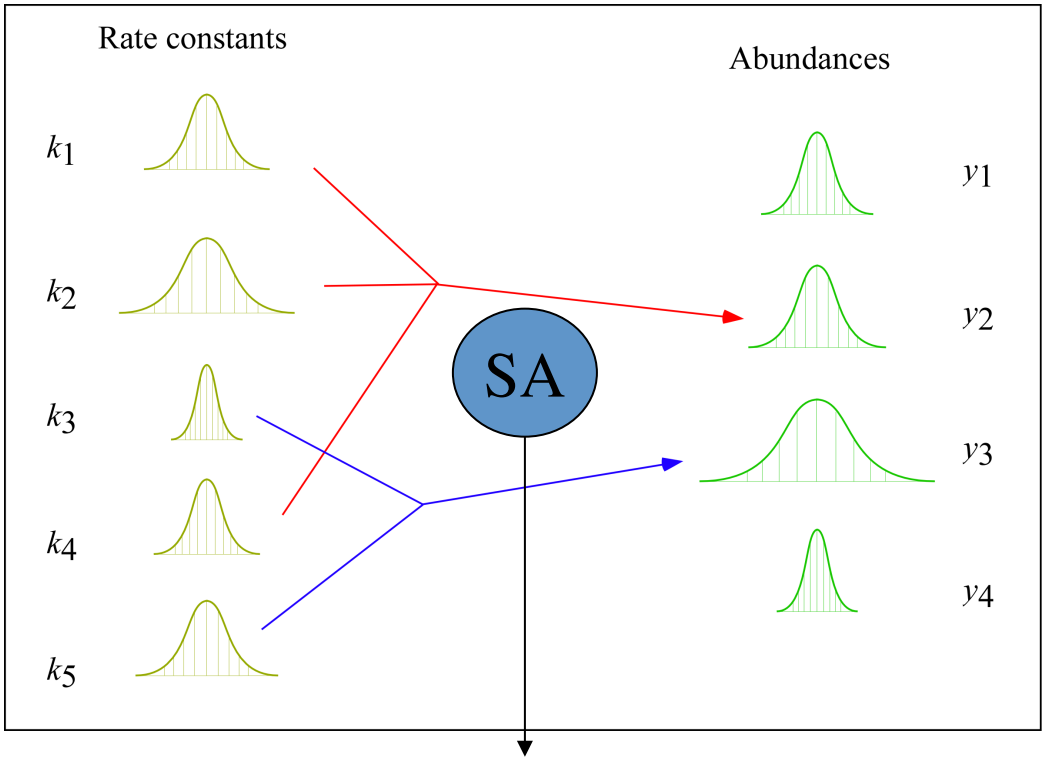
Sensitivity Analysis highlights 2 key reactions:

$CH + H \rightarrow C + H_2$ (k_b)

$CH + CH_4 \rightarrow C_2H_4 + H$ (k_a)

Analyse de sensibilité globale

$$X_i = \overline{X_i} \pm \Delta X_i$$
$$X_i = \log(k_i)$$



$$Y_j = \overline{Y_j} \pm \Delta Y_j$$
$$Y_j = \log(y_j)$$

Détermination des **facteurs clefs** (constantes de réaction) responsables des larges incertitudes sur les sorties (abondances)

Coefficients de corrélation de rang de Spearman

Utilisé pour des variables dont la relation est non linéaire mais monotone

Pour chaque composé i, chaque réaction j et N tirages Monte-Carlo

$$\{X_i, Y_j\} \rightarrow \{r_{X_i}, r_{Y_j}\}$$

Abondances (sorties)

Constantes de réaction (entrées)

Rang des variables ($N \times N$ variables)

Coefficient de rang de i avec j :

$$\rho_{i,j} = \frac{\mathrm{cov}(r_{X_i}, r_{Y_j})}{\sigma_{X_i}\sigma_{Y_j}}$$

Coefficients de corrélation de rang de Spearman

Reaction	RCC at 300 km	RCC at 200 km	RCC at 100 km
acetylene C_2H_2			
$C_2H + C_2H_2 \rightarrow C_4H_2 + H$	-0.23	-0.24	-0.24
$H + C_2H_2 + M \rightarrow C_2H_3 + M$	-0.23		-0.30
$C_2H_3 + CH_4 \rightarrow C_2H_4 + CH_3$			0.21
ethylene C_2H_4			
$H + C_2H_2 + M \rightarrow C_2H_3 + M$	0.32		0.27
$H + C_2H_4 + M \rightarrow C_2H_5 + M$	-0.25	-0.23	-0.37
$C_2H_3 + CH_4 \rightarrow C_2H_4 + CH_3$		0.49	0.59
ethane C_2H_6			
$^1CH_2 + CH_4 \rightarrow CH_3 + CH_3$	0.33	0.33	0.32
$H + C_2H_2 + M \rightarrow C_2H_3 + M$	-0.27		
$CH_3 + CH_3 + M \rightarrow C_2H_6 + M$	0.21	0.21	
$CH_3 + C_2H_5 + M \rightarrow C_3H_8 + M$	-0.31	-0.32	-0.32
$C_2H_3 + CH_4 \rightarrow C_2H_4 + CH_3$		0.22	0.39
propyne CH_3C_2H			
$H + C_2H_2 + M \rightarrow C_2H_3 + M$	0.42		0.25
$H + C_3H_3 + M \rightarrow CH_3C_2H + M$	0.48	0.57	0.45
$H + CH_3C_2H + M \rightarrow C_3H_5 + M$		-0.27	
$C_2H_3 + CH_4 \rightarrow C_2H_4 + CH_3$			-0.37
propene C_3H_6			
$H + C_3H_6 + M \rightarrow C_3H_7 + M$	-0.50	-0.35	
$H + C_2H_2 + M \rightarrow C_2H_3 + M$	0.31		0.32
$H + C_3H_5 + M \rightarrow C_3H_6 + M$	0.33	0.25	
$CH_3 + C_2H_3 + M \rightarrow C_3H_6 + M$		0.51	0.35
$CH_3 + C_2H_3 \rightarrow C_3H_5 + H$		-0.21	
$C_2H_3 + CH_4 \rightarrow C_2H_4 + CH_3$			-0.43
propane C_3H_8			
$C_2H + C_3H_8 \rightarrow C_3H_7 + C_2H_2$	-0.20	-0.20	
$CH_3 + C_2H_5 + M \rightarrow C_3H_8 + M$	0.72	0.71	0.72
$C_2H_3 + CH_4 \rightarrow C_2H_4 + CH_3$		0.24	0.45

RCC

Rank Correlation Coefficient

⇒ Les réactions qui ont un RCC (en valeur absolue) proche de 1 sont des réactions clefs

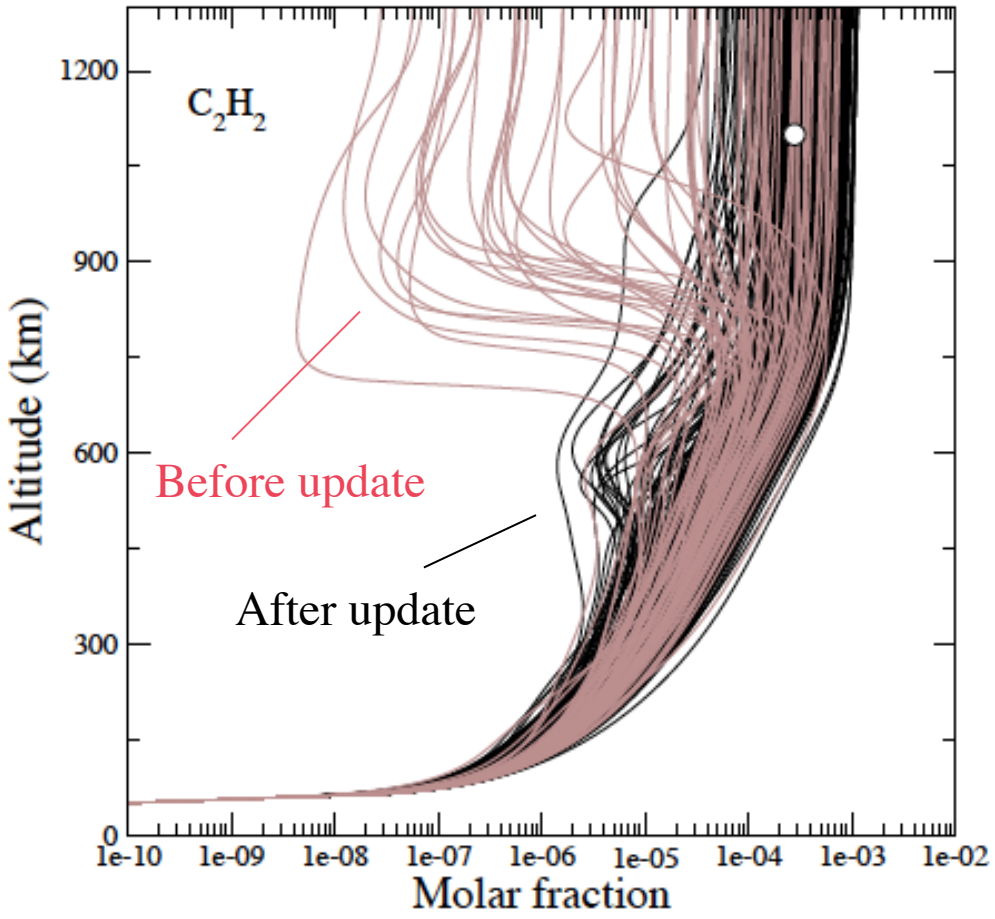
De nouvelles mesures des constantes clefs améliorent le modèle !

Exemple: nouvelle évaluation de 38 constantes à basse température

$CH + Réactants \rightarrow Produits$
$C_2H + Réactants \rightarrow Produits$
$C_4H + Réactants \rightarrow Produits$
$CN + Réactants \rightarrow Produits$

⇒ La diminution des facteurs d'incertitude des constantes clefs permettent :

- de réduire les incertitudes sur les sorties
- de supprimer les bimodalités

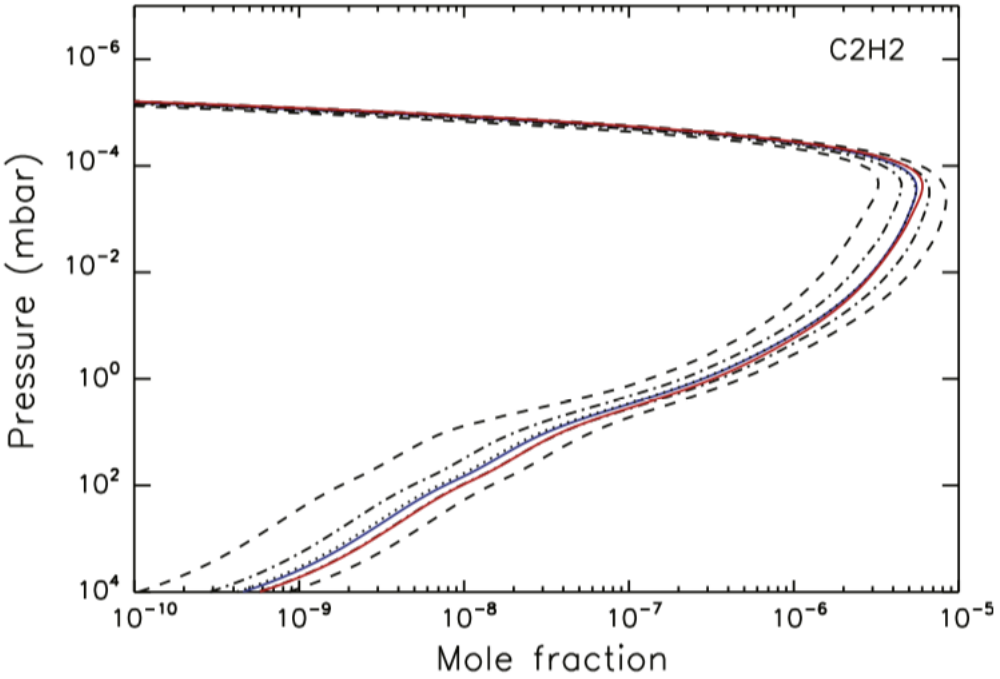


(Dobrijevic et al. 2008a)

Détermination de schémas réduits

Méthodologie : pour un ensemble d'espèces données, on ne conserve que les réactions dont les RCCs sont importants !

Schéma réactionnel initial :	→	Schéma réactionnel final :
- 124 espèces - 1313 réactions		- 22 espèces - 57 réactions



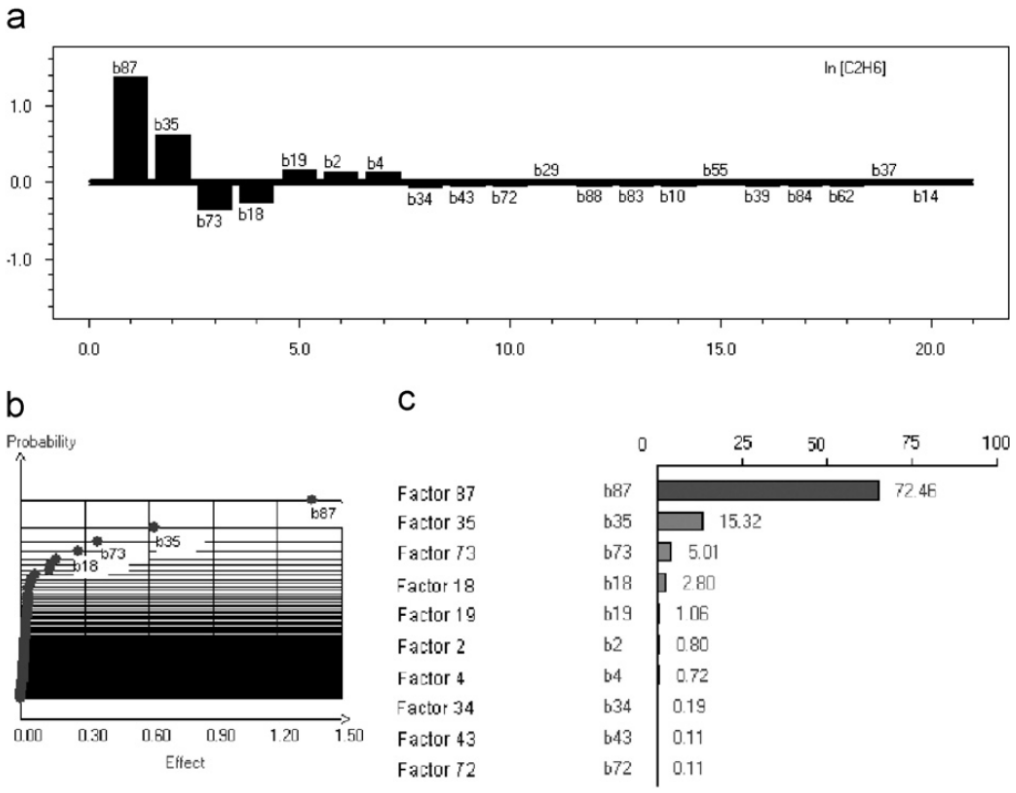
⇒ Le modèle réduit est en accord avec le modèle "complet"

(Hue et al. 2015)

Utilisation des plans d'expérience

Méthodologie : Construction d'un méta-modèle et recherche des facteurs clefs

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \cdots$$



(Dobrijevic et al. 2008b)

Utilisation des plans d'expérience

Sur des modèles simples

96 simulations

Key reactions determined by the experimental design

Compound	Key reactions (left: predominant; right: important)	
1: H_2	R18; R19; R4	R31; R6; R17; R43; R73; R29
2: H	R35; R18; R2; R17; R6; R73; R29	R37; R4; R31; R34
3: CH_4	R35; R4; R6; R73; R2; R19	R18; R17; R31
4: CH_3	R35; R18; R73	R19; R2; R4
5: $^3CH_2$	R19	R43; R73
6: $^1CH_2$	R18; R73	R6; R31; R19; R4
7: CH	R29; R17	R19; R18; R3
8: C_2H_2	R18; R73; R29; R35; R17	R2; R6; R4; R10
9: C_2H	R58	R18; R7; R73; R29; R17; R35
10: C_2	R56	R8; R18; R29; R17
11: C_2H_3	R80; R22; R51	
12: C_2H_4	R17; R29; R18	R2; R6; R26
13: C_2H_6	R87; R35	R73; R18; R19; R2; R4
14: C	R17; R29	R19; R2; R26; R18
15: C_2H_5	R87; R23; R39	R18; R35; R73; R19

Rank correlation coefficients for main reactions (5000 runs)

Reaction number	Rank correlation coefficients for C_2H_2 (absolute value >0.1)
R18	0.40
R35	−0.38
R6	0.34
R73	0.32
R17	−0.27
R10	0.27
R4	0.25
R29	0.20
R11	0.19
R2	0.18
R31	−0.13

⇒ Les plans d'expériences donnent les bonnes réactions clefs (mais pas forcément dans le bon ordre)

(Dobrijevic et al. 2008b)

Conclusion

Ce dont j'ai besoin pour mes modélisations dépend de mes objectifs :

1. Avoir une bonne estimation des incertitudes sur les sorties du modèle
2. Déterminer les facteurs clefs
3. Rechercher des multi-modalités

- Parcourir tout l'espace des paramètres d'entrée du modèle ?
- Utiliser une méthode "statistique" de type Monte-Carlo ?

Contrainte forte : le temps de calcul !

⇒ *Comment réduire le nombre de simulations ?*

⇒ *Comment déterminer la pertinence des facteurs clefs ?*

Une seule méthode pour tour faire ?