



HAL
open science

Méthodes de réduction et de propagation d'incertitudes : application à un modèle de chimie-transport pour la modélisation et la simulation des impacts

Jaouad Boutahar

► To cite this version:

Jaouad Boutahar. Méthodes de réduction et de propagation d'incertitudes : application à un modèle de chimie-transport pour la modélisation et la simulation des impacts. Modélisation et simulation. Ecole des Ponts ParisTech, 2004. Français. ⟨NNT : ⟩. ⟨tel-00007557⟩

HAL Id: tel-00007557

<https://pastel.hal.science/tel-00007557v1>

Submitted on 30 Nov 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



THESE DE DOCTORAT DE L'ECOLE NATIONALE DES PONTS ET CHAUSSEES

Spécialité
Mathématiques et Informatique

Présentée par
Jaouad Boutahar

pour obtenir le titre de Docteur de l'Ecole Nationale des Ponts et Chaussées

**Méthodes de réduction et de propagation d'incertitudes :
Application à un modèle de Chimie-Transport pour
la modélisation et la simulation des impacts.**

Soutenue le 30 Septembre 2004 devant le jury composé de :

Jérôme Jaffré	(INRIA)	Rapporteur
Alexis Coppalle	(INSA Rouen)	Rapporteur
Laurence Rouil	(INERIS)	Examinatrice
Pierre-Marc Riboud	(EDF)	Examineur
Pierre Rouchon	(ENSMP)	Directeur de thèse
Bruno Sportisse	(ENPC)	Co-Directeur de thèse

A ma famille.

"As far as the laws of mathematics refer to reality, they are not certain, as far as they are certain, they do not refer to reality."

Albert Einstein.

Remerciements

Je tiens à commencer ce mémoire de thèse en remerciant chaleureusement Bruno Sportisse de m'avoir proposé ce projet de recherche, m'avoir accueilli dans son groupe et avoir dirigé mes travaux tout au long de cette thèse. Je lui suis reconnaissant pour sa disponibilité, son soutien, sa confiance et ses qualités humaines. Toute ma gratitude à toi Bruno !

Je tiens à remercier également Pierre Rouchon pour avoir assumé la direction de cette thèse ; bien que très pris par ses obligations il a toujours été là pour m'écouter et me conseiller.

Je remercie Electricité De France et l'Ecole Nationale des Ponts et Chaussées qui m'ont permis de réaliser ce projet.

Accepter de faire partie d'un jury de thèse revient à consacrer une partie importante de son temps au travail présenté et donc manifester un certain intérêt pour celui-ci. J'exprime donc ma profonde reconnaissance à Laurence Rouil, Pierre-Marc Riboud, Jérôme Jaffré et Alexis Coppalle qui m'ont fait l'honneur de participer à mon jury et de faire des commentaires très enrichissants.

J'adresse un grand merci à Luc Musson Genon et Marc Bocquet qui ont accepté de relire ma thèse.

Ma gratitude s'adresse aussi à ceux qui ont été tour à tour mes collègues de bureau, Alex Keiko, Bastien Albriet, Marc Bocquet et enfin Denis Quélo.

A tous les "CEREAns", chercheurs et futurs docteurs, je dis merci pour la bonne ambiance et tous mes encouragements aux doctorants : Monica Kryska, Yelva Roustan, Vivien Mallet, Maya Milliez, Edouard Debry et à tous ceux déjà cités.

A Jacques Daniel pour sa disponibilité et ses interventions toujours efficaces, je dis merci.

Mon estime à Catherine Charleux, Annick Piazza, Khadija Elouali pour leur bonne humeur, leur disponibilité et surtout leur gentillesse.

Je tiens à remercier Martine Ohanna pour son aide logistique quand j'ai eu un besoin urgent d'une impression en couleur.

Je n'oublie pas mes collègues de CER EVE que je remercie aussi ici, plus particulièrement les deux directeurs : Jean Claude Deutsch (l'ancien) et Jean Marie Mouchel (l'actuel), mais aussi Jérôme Dispan (un ancien), Bruno Barroca et Anne Belbeoch.

Ma reconnaissance s'adresse aussi à tous ceux qui, de près ou de loin, ont permis que ce travail se réalise.

Je remercie enfin "last but not least" pour leur soutien mes parents, ma famille et mes amis.

Contents

Partie I	Introduction	7
1	Introduction	9
1.1	Modélisation des impacts	10
1.2	Fil directeur de la thèse	13
1.2.1	Modélisation de Chimie-Transport	15
1.2.2	Réduction de modèles	16
1.2.3	Propagation d'incertitude et analyse de sensibilité	18
1.2.4	Matrices de transfert par approche adjointe	20
1.2.5	Modèle statistique moyen : Journées et année représentatives	21
1.3	Organisation de la thèse	22
Partie II	Modélisation de Chimie-Transport	25
2	Le modèle de Chimie-Transport POLAIR3D	27
2.1	Les diverses approches de la modélisation	28
2.1.1	Les modèles Gaussiens	29
2.1.2	Modèles de boîtes.	29
2.1.3	Modèles tridimensionnels.	30
2.1.4	Description de la structure des entrée/sorties et des données nécessaires à un modèle de Chimie-Transport	31
2.1.5	Sorties	32
2.2	Les diverses paramétrisations utilisées	32
2.2.1	Émissions et sources	33
2.2.2	Chimie	34
2.2.3	Cinétique chimique	34
2.2.4	SPACK : Preprocesseur chimique	37
2.2.5	Diffusion	37
2.2.6	Paramétrisation sous maille : effet convectif	40
2.2.7	Conditions initiales et aux limites	42
2.2.8	Processus de perte	45

2.3	Description de la plateforme numérique de Chimie-Transport, POLAIR3D	48
2.3.1	Géométrie du modèle	49
2.3.2	Méthodes de séparation d'opérateurs : Splitting	51
2.3.3	Intégration de l'advection L_{advec}	60
2.3.4	Discrétisation du terme de la diffusion	66
2.3.5	Intégration de la chimie (L_{chem}) et la diffusion (L_{diffz})	68
3	Evaluation du modèle Chimie-Transport POLAIR3D par comparaison aux données du réseau EMEP pour l'été 2001	73
3.1	La géométrie du modèle	74
3.2	Les données météorologiques	75
3.3	La chimie	75
3.4	Les conditions initiales et aux limites	76
3.5	Les émissions	76
3.5.1	Les émissions dans le modèle	76
3.5.2	Inventaire d'émissions et distribution spatiale	77
3.5.3	Distribution temporelle	77
3.6	Méthodologie d'analyse	80
3.6.1	Les mesures du réseau EMEP	80
3.6.2	Mesures statistiques de l'erreur	83
3.6.3	Résultats de l'évaluation	84
	Partie III Réduction de modèle	97
4	Tabulation réduite avec la méthode HDMR (High Dimensional Model Representation)	99
4.1	Présentation de la méthode HDMR	101
4.1.1	Formulation	101
4.1.2	Coût de calcul des tables HDMR	104
4.1.3	Détermination des fonctions composantes de HDMR	104
4.1.4	Méthode Cut-HDMR	106
4.1.5	Quelques propriétés des fonctions de décomposition	109
4.1.6	Méthode MPcut-HDMR	110
4.2	Implémentation	114
4.3	Modèles simples d'illustration	116
4.3.1	Premier exemple d'illustration	116
4.3.2	Deuxième exemple d'illustration	117
4.4	Application à la cinétique chimique	118
4.4.1	Comparaison des profils de concentrations	119

4.4.2	Filtrage des fonctions HDMR	122
4.4.3	Test d'efficacité et coût de calcul	124
4.5	Sensibilité de Cut-HDMR au point de référence et au nombre d'échantillonnage	127
4.6	Application de HDMR au modèle POLAIR3D	130
4.6.1	Le modèle FEOM (Fully Equivalent Operationel Model	130
4.6.2	Variables d'entrée/sortie	131
4.6.3	Résultats	132
4.6.4	Gain en temps de calcul	141
5	Application de la méthode POD (Proper Orthogonal Decomposition) à la cinétique chimique	143
5.1	Fondement théorique de la méthode POD	145
5.1.1	Méthode des Snapshots	145
5.1.2	Projection du modèle d'origine sur la nouvelle base réduite	147
5.2	Implémentation algorithmique	148
5.3	Application à des cinétiques chimiques	150
5.3.1	Application au mécanisme CBM IV	150
5.3.2	Application au mécanisme RACM	152
5.3.3	Etudes de sensibilité POD-RACM	154
Partie IV	Modélisation de l'incertitude	171
6	Propagation d'incertitude dans un modèle de Chimie-Transport	173
6.1	Contexte méthodologique	177
6.1.1	Méthodes classiques de propagation d'incertitudes	177
6.1.2	Représentation de l'incertitude	180
6.1.3	Méthode de Monte Carlo classique	181
6.2	Présentation de la méthode DEMM/SRSM	182
6.2.1	Représentation des paramètres d'entrée	182
6.2.2	Formulation	182
6.3	Mise en oeuvre de la méthode DEMM/SRSM	186
6.3.1	Algorithme	186
6.3.2	Points de collocation	188
6.3.3	Un modèle simple	189
6.4	Application à un modèle 0D	191
6.4.1	Modèle	191
6.4.2	Résultats	192
6.5	Application à un modèle 3D sur l'Europe dans le cadre d'une étude d'impact	194

6.6	Couplage de DEMM/SRSM et de la différentiation automatique . . .	206
7	Analyse de sensibilité par la méthode RS-HDMR (Random Sampling - High Dimensional Model Representation)	211
7.1	Méthode RS-HDMR	212
7.1.1	Formulation	212
7.1.2	Projection sur des bases de polynômes orthogonaux	215
7.1.3	Application : Modèle Simple	217
7.1.4	Réduction de variance	219
7.2	Application : cinétique chimique de POLAIR3D	222
7.3	Etude de sensibilité par RS-HDMR	222
7.3.1	Test : modèle simple	225
7.3.2	Cinétique chimique EMEP	225
Partie V	Matrices de Transfert et Modèles moyens	235
8	Calcul des matrices de transfert par modèle adjoint	237
8.1	Non linéarité du modèle	239
8.1.1	Non linéarité des effets inter-pays, mono-espèce	241
8.1.2	Non linéarité des effets inter-espèces	242
8.2	Définition de la matrice de transfert	245
8.2.1	Modèle RAINS	246
8.3	Calcul de la matrice de transfert par modèle adjoint	247
8.3.1	Modèles dérivés	247
8.3.2	Calcul des dérivées partielles	248
8.3.3	Adjointisation du modèle POLAIR	251
8.3.4	Application et résultats	256
8.3.5	Coût calcul de la matrice de transfert	259
9	Modèles statistiques moyens : Journées et années représentatives	275
9.1	Les simulations annuelles de POLAIR3D	277
9.2	Le choix d'une année type	277
9.3	Choix d'un jour "type" par mois	287
9.4	Stratégie proposée pour le calcul des matrices de transfert	296
Partie VI	Conclusions et perspectives générales	301
10	Conclusions générales et perspectives	303
10.1	Modèle de Chimie-Transport (POLAIR3D)	303
10.2	Réduction de modèle	304

10.2.1	Proper Orthogonal Decomposition (POD)	304
10.2.2	High Dimensional Model Representation (HDMR)	304
10.3	Propagation d'incertitude	306
10.3.1	Random Sampling HDMR (RSHDMR)	306
10.3.2	Deterministic Equivalent Modeling Method/Stochastic Response Surface Method (DEMM/SRSM)	306
10.4	Matrices de transfert	307
10.5	Modèle statistique moyen : journée et année représentatives	308
10.6	Perspectives et futurs travaux	308
10.6.1	Développement de la version parallèle du modèle de Chimie Transport POLAIR3D	309
10.6.2	Réduction de modèle	309
10.6.3	Propagation d'incertitude	311
10.6.4	Matrices de transfert	311
10.6.5	Modèle statistique moyen	312
10.6.6	Modélisation intégrée	312
Partie VII Annexes		321
A Mécanisme CBM-IV		323
B La chimie d'acidification : EMEP		329
B.1	Réactions de EMEP ; Partie linéaire	329
B.2	Réaction de EMEP : Partie non linéaire	329
C Mécanisme chimique RACM		333
D Stations de Mesures du réseau Européen EMEP		343

Partie I
Introduction

Chapitre 1

Introduction

Les progrès effectués ces dernières années en matière de modélisation et de simulation numérique permettent de traiter des systèmes de plus en plus complexes, notamment dans le domaine de la géophysique. Compte tenu de l'état de l'art des moyens de calcul, les modèles restent cependant encore très coûteux en temps de calcul et un défi permanent reste celui de la réduction du coût calcul à qualité des résultats conservée. Ce problème est en pratique résolu en négligeant des phénomènes physiques, en les paramétrisant ou en utilisant des techniques numériques appropriées.

La modélisation de la cinétique chimique est un exemple typique de cette problématique. Sa réduction est une préoccupation qui intéresse les modélisateurs depuis plusieurs années que ce soit dans le domaine de la combustion ([1]) et plus récemment dans celui de la chimie atmosphérique ([2]).

En effet, la cinétique chimique fait intervenir un grand nombre de réactions et d'espèces chimiques dont les échelles de temps ont des ordres de grandeur très différents les unes des autres. Cette dispersion des échelles en temps induit la raideur des équations différentielles qui décrivent la chimie. Ces équations sont, de plus, fortement couplées entre elles et sont non-linéaires. L'évolution vers des modèles multiphasiques (gaz, eau, aérosols) amplifie ces caractéristiques. Ce problème exige l'utilisation de techniques numériques appropriées, les meilleurs algorithmes numériques restant extrêmement chers pour résoudre de telles équations ([3]). Dans le cadre d'une application atmosphérique tridimensionnelle, ce problème devient encore plus important, à cause du domaine de calcul qui est très grand (selon le type d'application, régionale, continentale ou globale) et une résolution verticale qui peut aller jusqu'à des dizaines de niveaux verticaux.

Les modèles de Chimie Transport, qui décrivent l'évolution en temps et en espace

d'espèces réactives, sont utilisés pour différentes applications :

- dans un but de **prévision** de l'état chimique de l'atmosphère (photochimique notamment) à l'échelle urbaine, régionale ou continentale.
- ou bien dans le cadre d'une **modélisation intégrée des impacts** des émissions anthropiques en estimant l'effet sur la qualité de l'air d'activité industrielles, d'infrastructures de transport ou de la production énergétique.

1.1 Modélisation des impacts

Les pays européens, conscients de l'effet nocif des émissions d'espèces chimiques (soufre, NO_x , COV, mercure, métaux lourds,...) sur la dégradation des écosystèmes¹ et sur la santé humaine, ont imposé une réduction de ces émissions à travers plusieurs protocoles et directives européennes. Il est établi pour chaque pays et par polluant une charge critique⁽²⁾ à ne pas dépasser. Il est clair, que pour un pays européen donné, les espèces chimiques mesurées ne proviennent pas uniquement des émissions nationales mais sont aussi importées d'autres pays. A titre d'exemple, selon la Commission Européenne, en 1997, la France exportait 70% de ses émissions de SO_2 et 63% de ses émissions de NO_x à travers l'Europe [4]. Ceci correspond à ce que l'on appelle classiquement "la pollution transfrontalière".

Dans ces conditions, les pays européens doivent utiliser des réductions d'émissions plus efficaces et être capable de quantifier pour un taux de réduction d'émissions par polluant, son effet sur le dépassement des charges critiques sur l'Europe et le coût engendré par une telle réduction. Il est donc indispensable de disposer d'outils d'aide à la décision basés sur des fondements scientifiques capables de fournir les informations les plus fidèles possibles sur l'effet d'une stratégie de réduction d'une émission polluante sur l'écosystème et la santé publique.

Dans ce contexte, le développement d'un outil de calcul d'impact (Figure 1.1) devient primordial et il doit comporter et représenter les différents processus qui interviennent jusqu'à l'impact final :

- un inventaire d'émission,
- un module de dispersion qui fournit les concentrations des espèces chimiques,

¹définis comme un ensemble dynamique d'organismes vivants (plantes, animaux et micro-organismes) qui interagissent entre eux et avec le milieu (sol, air, eau, lumière) dans lequel ils vivent.

²charge la plus élevée qui ne causera pas de changements chimiques menant à des effets nocifs à long terme dans les systèmes écologiques les plus sensibles. Ces charges sont établies à base de critères géologiques, pédologiques, hydrologiques et écologiques voire démographiques.

- un module de calcul d'exposition,
- un module d'estimation des impacts, par exemple via une loi de type dose-effet pour les impacts sanitaires,
- un modèle macro-économique capable d'estimer les coûts engendrés par les scénarios de réductions des émissions. Un exemple de coût engendré par une stratégie de réduction est présenté dans le tableau 1.1.

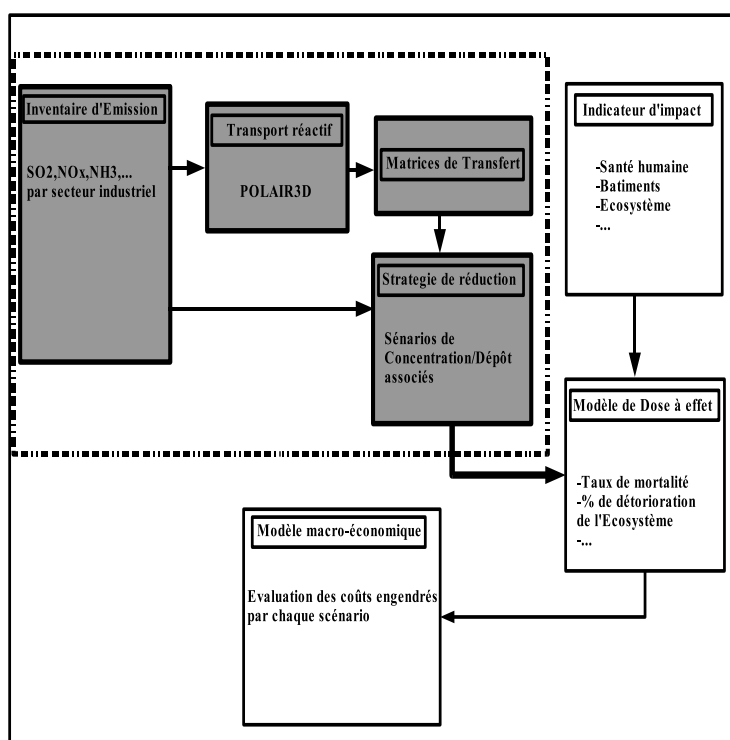


Figure 1.1: Schéma synthétique représentant le fonctionnement d'un outil intégré de modélisation d'impact

Un des outils d'impact les plus utilisés par la Commission Européenne est le modèle RAINS [5] développé par l'IISA (International Institute for Applied Systems Analysis, [6]) outil de modélisation intégré. Ce système a permis, par exemple, à la commission Européenne sur le problème de l'acidification (dépôts en NO_x et SO_2) de demander une réduction de 93% des émissions soufrées de la France pour les grosses installations de combustion [4].

Généralement, dans la chaîne d'évaluation d'impact, le maillon relatif à la dispersion et au calcul des concentrations, pour des raisons de temps de calcul,

	SO₂ 1000 tonnes	NO_x 1000 tonnes	COV 1000 tonnes	NH₃ 1000 tonnes	Coûts M-Euros/an
Emissions en 1998	837	1654	2341	828	
Protocole Goeteborg (Objectifs pour 2010)	400	860	1100	780	8745

Table 1.1: Stratégie de réduction des émissions pour 2010 demandée par le Protocole de Goeteborg (1999) pour la France (Source : CEE-ONU, Commission européenne, IIASA, [4]).

est souvent simplifié et peu fondé. Souvent pour les mêmes raisons, les réactions chimiques ne sont pas prises en compte dans la représentation des phénomènes. Ceci donne bien évidemment des résultats à faible précision pour des espèces réactives. Nous montrons, à titre d'exemple, dans la figure 1.2 l'apport de la chimie, avec le modèle POLAIR3D, en comparant la moyenne journalière sur l'Europe de NO_2 , pendant l'année 2001 par rapport au résultat obtenu avec la version passive ; le mécanisme chimique utilisé dans ce cas réactif est le schéma EMEP [7] (Annexe B).

Par conséquent et afin de s'approcher le plus possible de la réalité et produire des calculs d'impact fiables, le maillon relatif à la dispersion doit comporter un modèle de Chimie Transport capable de représenter les principaux processus physiques et chimiques, notamment :

- le transport horizontal par advection et la diffusion horizontale,
- le transport vertical par advection et la diffusion verticale,
- le dépôt sec/humide,
- les processus d'émissions surfaciques et volumiques,
- les réactions chimiques.

Les études d'impact nécessitent par ailleurs une utilisation spécifique des modèles :

- on souhaite tester plusieurs scénarii du modèle pour les différentes entrées du modèle (typiquement les champs météo et les émissions),
- ou on peut vouloir faire propager des incertitudes sur les entrées du modèle pour les différentes sorties. Dans les deux cas, un grand nombre d'évaluations du modèle est nécessaire.

Cela est bien sûr irréalisable avec des modèles de Chimie-Transport complexes.

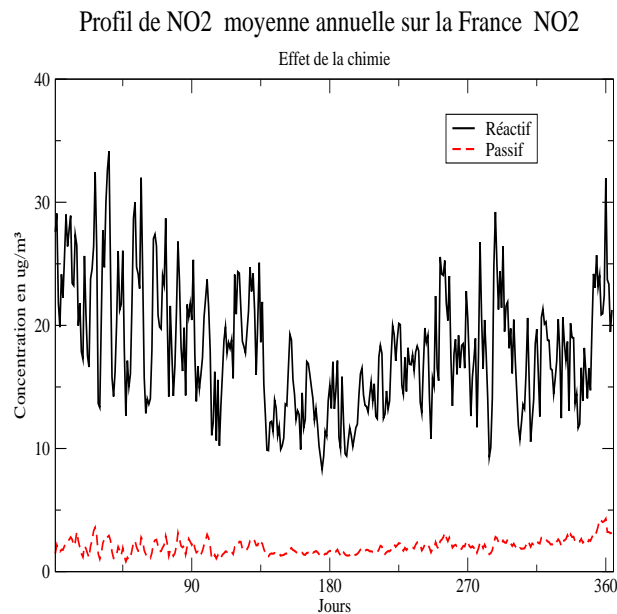


Figure 1.2: Apport de la chimie (Mécanisme EMEP [7]) année 2001

A titre d'exemple, une intégration du modèle de Chimie Transport POLAIR3D pendant une année sur l'Europe ($47 \times 32 \times 14 = 21056$ mailles), en utilisant une chimie simplifiée EMEP avec une dizaine d'espèces chimiques, prend environ 6h de temps CPU sur une machine Pentium 4 à 1 GHZ. Dans le cas multi scénarii qui nécessiterait plus de 1000 exécutions de modèle, cela demanderait plus d'une année de calcul. Dans le cas de la propagation d'incertitude où l'utilisation des méthodes de Monte Carlo nécessite un grand nombre d'échantillonnages (disons 5000 exécutions du modèle), ceci demanderait plus de 3 ans et demi de temps CPU.

1.2 Fil directeur de la thèse

Pour trouver une issue à cette difficulté, deux approches sont envisageables :

- la première est fondée sur une stratégie économe de propagation d'incertitude en comparaison de l'approche Monte Carlo, en ne faisant qu'un nombre restreint de simulations, en fonction de l'incertitude sur les entrées (méthode DEMM "Deterministic Equivalent Modeling Method" , [8]).
- la seconde revient à chercher à diminuer le coût d'évaluation du modèle CTM utilisé, en construisant un **modèle réduit**. Cette approche est appliquée à la cinétique chimique, module le plus consommateur en temps CPU dans un

CTM.

Pour atteindre l'objectif fixé, nous avons abordé cinq thèmes qui structurent cette thèse :

1. **La compréhension de la modélisation de chimie transport avec une analyse de ses capacités et son coût calcul.**

Un modèle de Chimie-Transport, performant numériquement, le modèle POLAIR3D [9], a été développé. POLAIR3D utilise les dernières techniques numériques de résolution des équations différentielles non linéaires, avec une modularité capable d'utiliser n'importe quel mécanisme chimique, gazeux ou de particules. Le développement d'un tel modèle a été aussi motivé par le souci de disposer d'un modèle adjoint pour effectuer aisément des études de sensibilité ou d'assimilation de données. Un tel développement permet une maîtrise des différents modules qui composent un modèle de chimie transport afin de mener à bien l'objectif de cette thèse qui est de construire des modèles réduits.

2. **L'utilisation des techniques de réduction de la cinétique chimique.**

Le second objectif est de réduire le coût calcul demandé par l'exécution de tels modèles dans un contexte tridimensionnel, en utilisant des méthodes récentes qui ont donné de bons résultats dans d'autres domaines d'application, par exemple, dans la turbulence ou dans la combustion.

3. **La réduction propagation d'incertitudes.**

Dans les études d'impact, plusieurs scénarii doivent être testés à base de plusieurs entrées, où on doit identifier l'effet des incertitudes des données d'entrée sur les sorties du modèle. Le troisième objectif est de fournir une alternative à la méthode de Monte Carlo dans le cas d'une approche multi-scénarii.

4. **Les stratégies capables de quantifier l'impact par le calcul des matrices de transfert.**

Un point clé dans une modélisation intégrée est de pouvoir quantifier d'une manière la plus rapide et la plus efficace possible le pourcentage de réduction dans la concentration dans l'air et/ou déposée d'espèces chimiques suite à des réductions d'émissions. Cette quantification est représentée classiquement

sous forme de matrices appelées matrices de transfert. Nous avons développé une méthode efficace de calcul de matrices de transfert en utilisant le modèle adjoint de POLAIR3D, et étudié les limites de cette approche (linéarisée).

5. La définition d'un modèle statistique moyen.

Dans une modélisation d'impact, le calcul des matrices de transfert se fait sur plusieurs années. Afin de réduire le nombre de calculs à effectuer, nous étudions la faisabilité de définir un modèle statistique moyen, à partir des différentes simulations du modèle POLAIR3D, qui permet de faire des calculs d'impact sur un faible nombre de jours.

1.2.1 Modélisation de Chimie-Transport

Dans cette partie nous développons un modèle de Chimie-Transport POLAIR3D. On cherche à décrire l'évolution en temps et en espace des composés chimiques dans l'atmosphère, dans un premier temps gazeux. Pour toute espèce c_i , son équation d'évolution peut s'écrire selon :

$$\frac{\partial c_i}{\partial t} = \underbrace{-div(V c_i)}_{Advection} + \underbrace{div(K \nabla c_i)}_{Diffusion} + \underbrace{\chi_i(c, T)}_{Chimie} + \underbrace{E_i}_{Emissions} - \underbrace{V_{dep}^i c_i}_{Deposition} + \underbrace{F_i}_{Autres} \quad (1.1)$$

où l'on prend en compte l'advection par le vent (V), la diffusion horizontale et verticale (K est le tenseur de diffusion), la cinétique chimique qui représente l'interaction entre les différents espèces considérées dans le mécanisme choisi, les émissions (E_i), le dépôt sec associé à la vitesse V_{dep}^i et d'autre paramétrisations notamment de lessivage (F_i).

En vue du calcul des impacts, et plus spécifiquement des dépôts au sol, une attention forte doit être accordée au traitement des conditions aux limites : le dépôt sec et humide à travers les précipitations au sol et les émissions qui comportent les termes surfaciques (rejets diffus, par exemple biogéniques, ou liés au trafic routier) et volumiques (rejets des cheminées industrielles).

D'autres processus tel que la convection par exemple, peuvent être pris en compte dans la modélisation de chimie transport.

Notons qu'il existe par ailleurs une large variété de modèles selon l'échelle couverte et l'application désirée. A titre d'exemple nous pouvons citer :

- EURAD [10] et EMEP [11] pour des applications à l'échelle continentale sur l'Europe,
- CMAQ Modèle Americain multi-échelle (urbaine et régionale) [12],
- MOZART [13] à l'échelle globale,
- CHIMERE conçu pour des applications régionales et continentales ([14]).

1.2.2 Réduction de modèles

Dans les modèles 3D de Chimie-Transport, la cinétique chimique peut consommer entre 50% et 90% du temps CPU total de la simulation. Cette situation devient encore plus critique quand des espèces et des réactions chimiques sont à rajouter dans les modèles pour les rendre plus complets et plus réalistes (par exemple : la prise en compte de plusieurs phases). Pour réduire le coût calcul, trois possibilités sont envisageables

- l'usage de méthodes numériques de plus en **plus rapides et efficaces**.
- l'exploitation du potentiel de développement des moyens de calcul en utilisant les techniques de **parallélisme**. Depuis l'apparition des machines à architecture parallèle des techniques de parallélisation sont utilisées dans les modèles afin de réduire leur coût calcul. Cette méthode nécessite des développements conséquents et des modifications importantes sur le modèle. Elle nécessite aussi l'accès à ces ressources de calcul intensif.
- l'utilisation de techniques de **réduction** :
Une première approche pour répondre à cette difficulté est liée à la réduction dynamique de tels systèmes. En se fondant sur le caractère lent/rapide des systèmes différentiels résultants, des modèles réduits de plus petite taille peuvent être construits en filtrant les composantes rapides [15].
Une seconde approche est liée au comportement statistique de ces systèmes et est fondée sur une décomposition en modes propres des solutions [16].
Une troisième approche, basée sur des développements en polynômes paramétrisés et sur des interpolations de tables pré-calculées (Look-Up Tables) qui peuvent être utilisées comme modèle de représentation pour remplacer les solveurs de chimie [17].

L'objectif de cette partie est de construire un modèle (ou des modèles) simplifiés capable d'être utilisé sur de longues périodes en remplacement du modèle d'origine. Notre travail a porté plus spécifiquement sur les approches suivantes :

1.2.2.1 Proper Orthogonal Decomposition (POD)

Proper Orthogonal Decomposition (POD) est une technique de réduction déjà utilisée dans plusieurs applications notamment dans le traitement de la turbulence [18, 19, 20]. Elle est basée sur la recherche d'une base optimale qui contient toute l'information de la solution exacte (le modèle d'origine). Le modèle initial est ensuite projeté sur cette base.

Si on considère un système de cinétique chimique (c est la concentration des espèces chimiques) décrit par une équation de la forme :

$$\frac{dc}{dt} = f(c), \quad c(0) = c_0 \in R^n \quad (1.2)$$

où n est la dimension du système (typiquement le nombre d'espèces chimiques), c est la concentration des espèces chimiques. On utilise en pratique la méthode dite de "snapshots" :

- Dans un premier temps, la méthode consiste à calculer une trajectoire exacte du modèle ($C(t)$) durant N itérations temporelles (t_1, t_2, \dots, t_N).
- L'ensemble de ces informations (C_1, C_2, \dots, C_N) génère un sous-espace vectoriel de base $\Phi = (\Phi_1, \Phi_2 \dots \Phi_d)$.
- L'objectif est alors de choisir d'une manière optimale une base orthonormale réduite $\Psi = (\Psi_1, \Psi_2 \dots \Psi_p)$ avec $p < d$.
- Le modèle initial est ensuite projeté sur la base réduite Ψ .

1.2.2.2 High Dimensional Model Representation (HDMR)

Pour un modèle ayant x_1, x_2, \dots, x_n variables d'entrée, pour une résolution en $1/s$ (s points d'échantillonnage par direction), la construction des traditionnelles tables pré-calculées d'interpolation nécessite s^n exécutions du modèle initial. En plus des difficultés liées aux interpolations à n dimensions, ces approches restent prohibitives pour des modèles chimiques qui comprennent un très grand nombre de degrés de liberté (une centaine d'espèces chimiques typiquement).

L'objectif de la méthode HDMR est de tabuler de manière efficace la relation entrée-sortie d'un modèle, ie en remplaçant le calcul d'un modèle $y = f(x)$ par la tabulation de cette relation fonctionnelle. Ceci nécessite bien sûr de tabuler l'espace X des valeurs possibles du paramètre d'entrée x , ce qui est en pratique irréalisable. La méthode HDMR permet de choisir les points "appropriés" de tabulation.

Pour un vecteur de paramètres d'entrée (x_1, x_2, \dots, x_n) , on va représenter la

fonction f par un développement fonctionnel sous la forme :

$$\begin{aligned}
 f(x) = & f_0 + \sum_{i=1}^{i=n} f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) \\
 & + \sum_{1 \leq i < j < k \leq n} f_{ijk}(x_i, x_j, x_k) + \dots + f_{1,2,\dots,n}(x)
 \end{aligned}
 \tag{1.3}$$

f_0 correspond à un effet moyen, la fonction f_i correspond à l'effet isolé de la variable x_i , la fonction f_{ij} aux effets couplés des variables x_i et x_j , etc. Le dernier terme $f_{1,2,\dots,n}$ correspond à un terme résiduel, une fois pris en compte tous les uplets de variables d'entrée.

La formulation HDMR est très efficace si on peut tronquer ce développement à un ordre $l < n$. En pratique, les tests utilisés sur des cinétiques chimiques issues de la chimie stratosphérique ([21, 22, 23, 24]) ont montré la très rapide convergence de ce développement (l'ordre 2 est suffisant pour les cas cités).

Cette méthode est utilisée dans le but de remplacer le module de chimie du modèle de CTM, par un modèle efficace et rapide basée sur de simples interpolations en utilisant les tables HDMR.

1.2.3 Propagation d'incertitude et analyse de sensibilité

Actuellement, les modèles de chimie transport permettent de faire des simulations de plus en plus réalistes du transport réactif des différentes espèces chimiques dans l'atmosphère. Cependant, la complexité des phénomènes étudiés, la variabilité spatiale et temporelle des espèces prise en compte engendrent un ensemble d'erreurs et d'incertitudes dans les données d'entrée comme par exemple :

- les émissions (il s'agit certainement de l'une des plus grandes sources d'incertitudes [25]),
- les champs météorologiques,
- la chimie à travers les réactions, leur constantes cinétiques et les simplifications effectuées dans les mécanismes pour les rendre moins complexes (regroupement d'espèces par exemple),
- les conditions aux limites.
- les algorithmes numériques choisis pour résoudre les équations du modèle
- ...

Lorsque ces données sont mal décrites ou présentent des erreurs, elles peuvent engendrer des erreurs importantes sur les résultats du modèle.

Pour analyser ces erreurs, nous devons procéder à une analyse d'incertitude

qui consiste à étudier l'importance de l'incertitude des entrées du modèle sur l'incertitude des sorties de modèle. Il existe schématiquement deux manières pour étudier l'incertitude dans un modèle :

- Par analyse de sensibilité locale qui calcule l'effet du changement dans les variables d'entrée de modèle sur les sorties de modèle ($\frac{\partial f}{\partial x}$). Nous proposons une méthode basée sur la représentation fonctionnelle de la variance de sortie par la méthode High Dimensional Model Representation (RS-HDMR).
- La seconde façon est de calculer l'incertitude des sorties de modèle à partir des incertitudes supposées sur les entrées de modèle. Les méthodes traditionnelles pour les études de propagation d'incertitudes sont les méthodes classiques de Monte Carlo, mais elles demandent un grand nombre d'exécution de modèles ce qui rend, dans le contexte des modèles de chimie transport, cette tâche très difficile. Nous présentons ici une autre méthode basée sur l'optimisation du nombre de calculs à effectuer (Deterministic Equivalent Modeling Method/DEMM [8] et Stochastic Response Surface Method/SRSM [26]).

1.2.3.1 Analyse de sensibilité par Random Sampling HDMR

La méthode HDMR peut être utilisée comme solveur pour remplacer un modèle mathématique très coûteux en temps de calcul en construisant un modèle efficace dit FEOM (Fully Equivalent Operationnel model), pour identifier les variables importantes du modèle ou pour une étude d'incertitude.

Dans le cas d'une utilisation pour analyse de sensibilité la méthode HDMR est utilisée pour décomposer la variance d'une sortie de modèle σ_g^2 en différentes contributions des entrées de modèle selon :

$$\sigma_g^2 = \sum_{i=1}^n \sigma_i^2 + \sum_{1 \leq i < j \leq n} \sigma_{ij}^2 + \sum_{1 \leq i < j < k \leq n} \sigma_{ijk}^2 + \dots \quad (1.4)$$

où σ_i^2 est la contribution de la variable d'entrée i sur la variance sortie, agissant seule (n est le nombre total des entrées de modèle). σ_{ij}^2 est la contribution couplée de la variable d'entrée i et j sur la variance sortie. Nous utiliserons la version RS-HDMR (Random Sampling) de la méthode HDMR, basée sur des calculs de moyenne sur le domaine Ω des variables d'entrée, pour calculer les différentes contributions σ_i^2 et σ_{ij}^2 .

1.2.3.2 DEMM/SRSM

L'objectif de cette partie est de mesurer d'une part l'influence des incertitudes sur la qualité des résultats et d'autre part d'être capable de présenter les résultats de modèles sous forme de probabilité de scénarios. Une approche de type Monte Carlo est malheureusement inutilisable telle quelle, vu le coût calcul requis. On utilisera des techniques de réductions de modèles développées au MIT par G. Mac Rae, pour l'étude de la propagation d'incertitude (DEMM). La méthode est basée sur la représentation d'une sortie de modèle supposée incertaine comme un développement en série de fonctions de variables d'entrée prises aléatoires, selon :

$$\begin{aligned}
Y &= Y_0 + \sum_{i=1}^n Y_i^{(1)} H_1(\xi_i) + \sum_{i=1}^n Y_i^{(2)} H_2(\xi_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(1)} H_1(\xi_i) H_1(\xi_j) \\
&+ \sum_{i=1}^n Y_i^{(3)} H_3(\xi_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(2)} H_1(\xi_i) H_2(\xi_j) \\
&+ \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(3)} H_2(\xi_i) H_2(\xi_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n Y_{ijk}^{(1)} H_1(\xi_i) H_1(\xi_j) H_1(\xi_k). \\
&+ \dots
\end{aligned} \tag{1.5}$$

où n est le nombre des paramètres d'entrée incertains, ξ_i est une variable aléatoire standard suivant une loi normale $N(0,1)$ correspondant à la valeur d'entrée incertaine x_i , Y est une sortie de modèle, $H_m(\xi_i)$ sont les séries de polynômes orthogonaux d'Hermite de degré m .

L'objectif est de pouvoir estimer les différents coefficients Y_i^1 , Y_i^2 , Y_i^3 , Y_{ij}^1 et Y_{ijk}^1 , en effectuant un nombre limité d'exécutions du modèle Y (dans notre cas le modèle de chimie transport POLAIR3D). Pour cela nous utiliserons la méthode SRSM (Stochastic Response Surface Method) [26].

1.2.4 Matrices de transfert par approche adjointe

Les matrices de transfert source/récepteur représentent le lien qui existe entre la concentration dans l'air ou déposée d'une espèce chimique sur plusieurs pays ou une région (ensemble de pays) pendant une période temporelle (journalière, mensuelle ou annuelle) et les émissions d'une autre espèce (ou la même espèce) issues d'un pays donné ou de quelques sources ponctuelles. On peut représenter cette relation par :

$$(D_X)_i = \sum_{j=1}^{j=N_{srce}} M_{ij} * (E_Y)_j \quad (1.6)$$

où X représente les espèces chimiques du modèle avec $(D_X)_i$ le terme de concentration déposée de l'espèce X sur un point (pays) récepteur i . Y symbolise l'espèce chimique émise avec $(E_Y)_j$ terme des émissions de l'espèce émise Y du point (pays) j (N_{srce} dans ce cas est le nombre de points (pays) émetteurs). M_{ij} est l'élément correspondant de la matrice de transfert.

Dans le système RAINS, les termes de ces matrices sont calculés d'une manière simplifiée, par des simples divisions des quantités déposées sur un pays et la quantité émise à partir d'un pays ou un point donné :

$$M_{ij} = (D_X)_i / (E_Y)_j \quad (1.7)$$

L'objectif ici est de calculer les éléments M_{ij} en calculant les dérivées partielles :

$$M_{ij} = \frac{\partial f_i}{\partial E_j} \quad (1.8)$$

f_i est le terme agrégé de sortie du modèle CTM (concentration dans l'air ou déposé sur un pays ou une région d'une espèce, après une intégration temporelle typiquement d'un an) et E_j est le terme d'émission. Ces gradient sont calculés d'une manière exacte par l'adjoint du modèle en utilisant le différentiateur automatique Odyssee [27].

1.2.5 Modèle statistique moyen : Journées et année représentatives

Traditionnellement, le calcul des matrices de transfert se fait sur des longues périodes de simulation (de plusieurs années), ce qui pourrait être très pénalisant dans un cadre opérationnel de modélisation d'impact. Afin de pouvoir réduire ces calculs, nous proposons une étude sur la variabilité interannuelle des champs de concentrations simulés par le modèle POLAIR3D à jeu d'émission annuel constant sur les différents années de calcul POLAIR3D. Dans le même objectif de réduction du temps de calcul, nous étudions la possibilité de représenter chaque mois de calcul POLAIR3D par un jour "type" capable de bien représenter le comportement moyen mensuel des champs de concentrations simulés par POLAIR3D.

1.3 Organisation de la thèse

Ce travail de thèse se compose de quatre parties. Dans la figure 1.3 est présenté un schéma synthétique de la structure et les différentes parties de cette thèse :

- **(I)** Après ce chapitre d'introduction (Chapitre 1), nous présentons dans **la première partie** le code POLAIR3D (Chapitre 2), ses performances par rapport aux mesures d'EMEP à l'échelle européenne sur une période d'été 2001 (Chapitre 3).
- **(II)** Dans **la seconde partie** nous étudions les principales méthodes de réduction envisagées :
 - Méthodes de tabulations HDMR : Cut-HDMR et une application 3D au modèle POLAIR3D avec l'utilisation de la chimie EMEP (Chapitre 4).
 - Méthodes de projections orthogonales (POD) avec des application 0D à la chimie de CBM-IV et RACM (Chapitre 5).
- **(III)** Dans la **troisième partie** nous présentons la méthode de propagation d'incertitude DEMM et son application au modèle de chimie transport POLAIR3D, dans le cadre d'une étude à l'échelle européenne (chapitre 6). Un autre chapitre porte sur l'analyse de sensibilité par la méthode RS-HDMR et une application au modèle POLAIR3D (Chapitre 7).
- **(IV)** **La quatrième partie** est composée de deux chapitres. Le premier est consacré à l'utilisation de la différentiation automatique pour le calcul des matrices de transfert avec une application à l'échelle européenne en utilisant POLAIR3D et le mécanisme chimique EMEP (Chapitre 8). Le deuxième chapitre de cette partie est consacré à la représentation par un modèle statistique moyen du modèle POLAIR3D (Chapitre 9).
- Les conclusions générales tirées à l'issue de ce travail, ainsi qu'un nombre de points ouverts et de perspectives sont enfin présentés pour clore ce manuscrit (Chapitre 10).

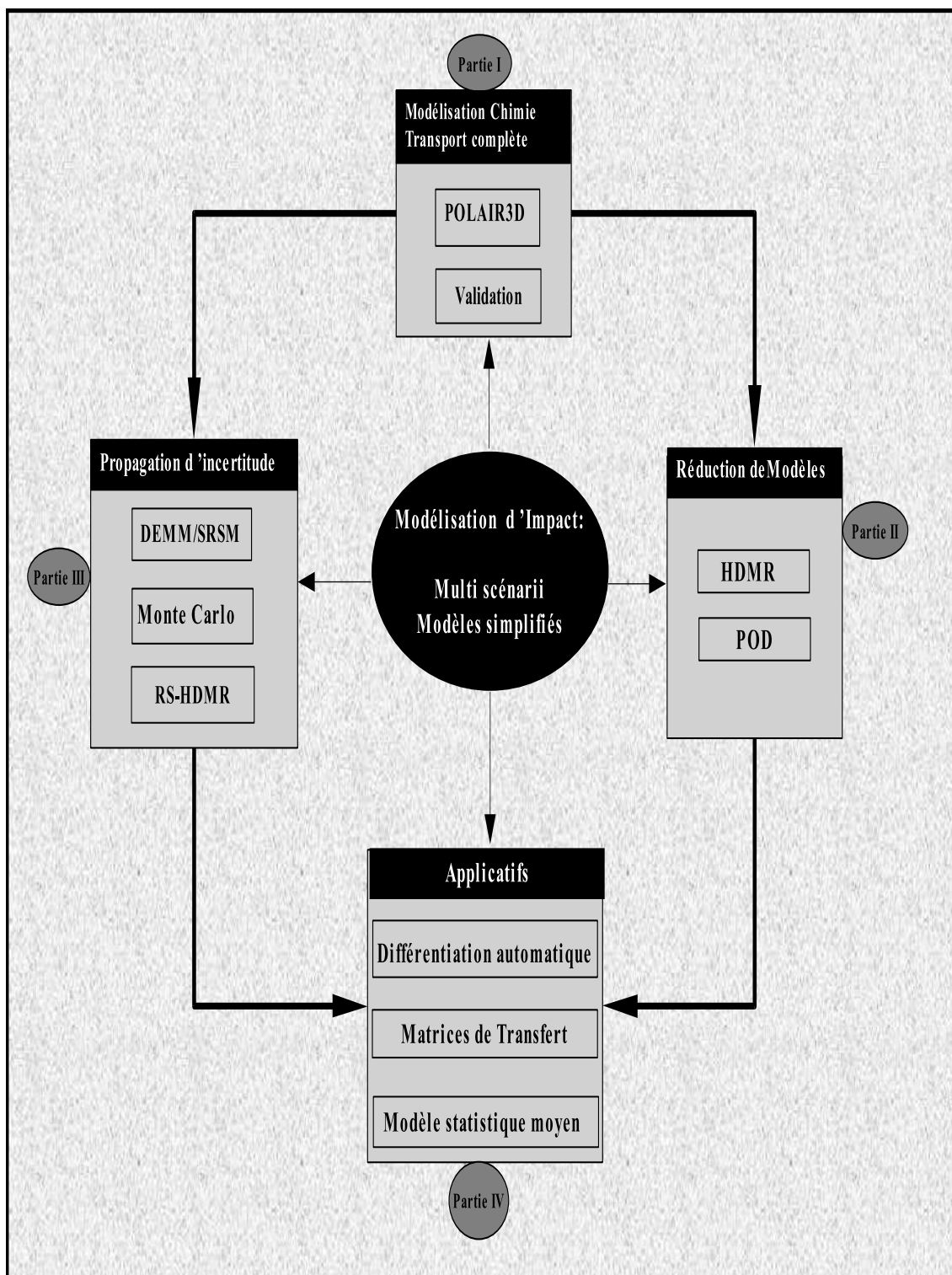


Figure 1.3: Schéma synthétisant la structure et les différentes parties de la thèse.

Partie II

Modélisation de Chimie-Transport

Chapitre 2

Le modèle de Chimie-Transport POLAIR3D

Résumé

Dans ce premier chapitre nous présentons les diverses approches de modélisation, (section 1), les différentes paramétrisations physiques et chimiques utilisées dans la dispersion atmosphérique (section 2) et la plateforme numérique de Chimie-Transport POLAIR3D que nous avons développée (section 3). Ce travail a fait l'objet d'une publication ([9])

2.1 Les diverses approches de la modélisation

Les modèles de qualité de l'air sont utilisés pour mieux comprendre la composition chimique de l'atmosphère et son effet sur la qualité de l'air. Ils simulent la dispersion des espèces chimiques en prenant en compte plusieurs processus : la chimie (en phase gazeuse, en phase aqueuse, particules), l'advection par le vent, la diffusion turbulente, les phénomènes convectifs, les émissions aux sol, les processus de dépôt sec et humide.

Pour décrire tous les phénomènes qui rentrent en jeu, les modèles doivent disposer de plusieurs données de terrain, des émissions et des données météorologiques. Ces dernières sont soit issues des modèles des services météorologiques (dans ce cas le modèle CTM est de type '**off-line**') ou soit simulées par le même modèle de chimie-transport qui sera de type '**on-line**'.

Il existe plusieurs techniques de modélisation et simulation de la qualité de l'air, qui sont regroupées principalement en deux catégories :

- les approches eulériennes,
- les approches lagrangiennes.

Chacune de ces méthodes présente des avantages et des inconvénients. La différence majeure entre les deux approches consiste en la description des phénomènes de dispersion :

- pour une approche eulérienne, le référentiel est fixe, on se placera en un point donné où l'on décrit le mouvement des particules qui passent par ce point,
- alors que dans la deuxième approche, le référentiel suit le mouvement des particules fluides.

Le choix d'une approche au lieu d'une autre dépend du type de phénomène et de l'échelle spatiale/temporelle considérés. Dans notre cas nous nous intéressons à l'approche eulérienne.

Pour effectuer des calculs de dispersion atmosphérique, on dispose généralement de plusieurs familles de modèles à savoir :

- les modèles de type Gaussien,
- les modèles de boîtes "Box models",
- les modèles tridimensionnels.

Nous présentons dans ce qui suit un aperçu sur les deux premiers types de modèles avant de nous intéresser plus en détail à la modélisation eulérienne tridimensionnelle.

2.1.1 Les modèles Gaussiens

Les premiers modèles de dispersion utilisés ont été de type Gaussien. Ces modèles, dits de dispersion passive, supposent que la répartition du polluant suit une loi Gaussienne dont les écarts-types dépendent de la distance à la source ou du temps écoulé depuis l'émission, appelé aussi temps de transfert, ainsi que des caractéristiques de la structure de l'atmosphère (ou classe de stabilité, [28]). Leur équation générale peut s'écrire de la manière suivante ([29]) :

$$C = \frac{Q}{2\pi\sigma_h\sigma_z\bar{u}} \exp\left[-\frac{1}{2}\left(\frac{y_r}{\sigma_y}\right)^2\right] \exp\left[-\frac{1}{2}\left(\frac{h_e - z_r}{\sigma_z}\right)^2\right] \quad (2.1)$$

où C est la concentration au point $r = (x_r, y_r, z_r)$ due à l'émission au point $s = (x_s, y_s, z_s)$; Q est le taux d'émissions, σ_h et σ_z sont les écarts types (horizontal et vertical) de la distribution spatiale de la concentration; h_e est la hauteur effective du rejet d'émissions; \bar{u} est le vent moyen à la hauteur des émissions.

Les modèles de type Gaussien s'appliquent à un polluant qui n'est soumis qu'à l'action du fluide porteur (air), c'est-à-dire qui, par sa densité, sa température ou sa concentration ne provoque pas de modification significative des caractéristiques de l'air, en particulier ne perturbe pas son écoulement. Ceci suppose qu'au niveau du rejet la vitesse du vent reste la même avec ou sans polluant.

Ces modèles sont assez bien adaptés :

- aux conditions météorologiques moyennes (turbulence homogène),
- loin de la source (plus de 100 mètres de la source),
- aux conditions d'absence du cisaillement vertical du vent qui n'éloignerait pas trop les polluants du sol,
- en absence d'obstacles (sans présence de relief),
- avec des vents de direction constante et de vitesse non nulle.

Les modèles Gaussiens, traitant de rejets continus sont généralement peu coûteux en ressources de calcul et couramment utilisés pour des applications opérationnelles liées notamment à la problématique du risque industriel. Cependant ce type de modèles reste une représentation très simplifiée de la réalité.

2.1.2 Modèles de boîtes.

Ces modèles se basent sur la conservation de la masse du polluant à l'intérieur d'une boîte qui peut généralement représenter un large domaine comme par exemple l'atmosphère au dessus d'une ville. Ils sont décrits par l'équation :

$$\frac{d\bar{c}_i}{dt} = \chi_i(\bar{c}, T, t) + S_i(t) + \frac{1}{H(t)}(E_i(t) - v_{dep}^i \bar{c}_i) \quad (2.2)$$

où \bar{c}_i représente la concentration de l'espèce chimique i ; χ_i est le terme de chimie; $S_i(t)$ est le terme source de l'espèce i ; $H(t)$ est la hauteur de mélange; $E_i(t)$ est le

terme d'émission surfacique ; v_{dep}^i est la vitesse de dépôt de l'espèce chimique i . Pour des modèles qui font intervenir plusieurs boîtes, chaque boîte est gouvernée par une équation de type (2.2).

Ces modèles sont souvent utilisés pour leur coût raisonnable de mise en oeuvre ainsi que le temps de calcul nécessaire à l'exécution du modèle. Dans [30] nous retrouvons une comparaison de cette modélisation avec les modèles eulériens.

2.1.3 Modèles tridimensionnels.

Les modèles tridimensionnels essayent de mieux représenter et simuler les différents processus physiques et chimiques qui interviennent de façon significative dans la dispersion des polluants dans l'atmosphère. La complexité de la chimie atmosphérique conduit à l'utilisation de mécanismes chimiques détaillés qui font interagir des centaines d'espèces chimiques, à travers des centaines de réactions chimiques, sur des domaines spatiaux qui vont de l'échelle d'une ville à l'échelle continentale et sur des périodes temporelles d'un jour à une année en fonction du type d'applications auxquelles nous nous intéressons. Ce type de modèle s'avère très coûteux en matière de temps de calcul.

L'équation fondamentale du modèle décrivant l'évolution en temps et en espace des espèces chimiques, basée sur la conservation de la masse, sur une grille eulérienne s'écrit :

$$\frac{\partial c_i}{\partial t} + \nabla(Uc_i) = \nabla(K.\nabla c_i) + \chi_i(c, T) + E_i + S_i - v_{dep}^i.c_i - \Lambda_i(t)c_i + F_i(c_i) \quad (2.3)$$

Elle tient compte des processus suivants :

- **Advection par le vent :** $-\nabla(Uc_i)$ qu'on notera $L_{adv}(c_i)$, où c_i est la concentration de l'espèce chimique i composante du vecteur global de concentrations c de dimension N_{esp} (en microgrammes m^{-3} , molécules cm^{-3} , ou en ppb,...). U est le champs du vent tridimensionnel de composantes (u,v,w) en (ms^{-1}). On distinguera l'advection selon les trois dimensions :

$$L_{adv} = L_{advx} + L_{advy} + L_{advz} \quad (2.4)$$

où :

$$L_{advx} = -\nabla(uc_i), \quad L_{advy} = -\nabla(vc_i), \quad L_{advz} = -\nabla(wc_i) \quad (2.5)$$

- **Diffusion turbulente** : $\nabla(K\nabla c_i)$ notée $L_{diff}(c_i)$, où K est le tenseur de diffusion turbulente. Il est classiquement pris diagonal avec des diffusions horizontales K_x , K_y constantes et K_z la diffusion verticale calculée selon une paramétrisation (par exemple [31]). Nous distinguerons la diffusion le long des directions x , y et z :

$$L_{diff} = L_{diffx} + L_{diffy} + L_{diffz} \quad (2.6)$$

où :

$$\begin{aligned} L_{diffx}(c) &= \nabla_x(K_x \nabla_x c_i), & L_{diffy}(c) &= \nabla_y(K_y \nabla_y c_i), \\ L_{diffz}(c) &= \nabla_z(K_z \nabla_z c_i) \end{aligned} \quad (2.7)$$

- **Production chimique** : $L_{chem}(c) = \chi(c, T, P, H, I)$ est le terme de production chimique pour l'espèce i , qui dépend notamment de la température T , de la pression P , de l'humidité H , de l'état "radiatif" I (photolyse), et d'autres paramètres tels que le contenu en eau liquide (pour le cas d'une chimie aqueuse).
- **Emissions** : E_i (resp. S_i) relative à l'espèce chimique i . Elle dépend du temps et de l'espace, les rejets peuvent être surfaciques (resp. volumiques) selon qu'ils proviennent de l'industrie, du transport routier ou autre.
- **Dépôt sec** : $v_{dep}^i \cdot c_i$ où v_{dep}^i est la vitesse de dépôt de l'espèce i , On notera L_{dep} le processus de dépôt sec.
- **Dépôt humide** : $\Lambda_i \cdot c_i$ où Λ_i est une paramétrisation du lessivage humide spécifique à l'espèce i . On notera L_{less} le processus de lessivage humide.
- **Le terme F** peut décrire le processus de diffusion verticale causé par les phénomènes convectifs caractérisés par des ascendances et descendances assez significatives dans certaines conditions météorologiques ([32]). On le notera $L_{conv}(c_i)$.

2.1.4 Description de la structure des entrée/sorties et des données nécessaires à un modèle de Chimie-Transport

2.1.4.1 Entrée

Les modèles de Chimie-Transport mettent en oeuvre plusieurs données d'entrée :

- **Les données météorologiques** : la vitesse du vent 3D U , l'humidité, la température, la pression, le rayonnement solaire, le contenu en eau dans le sol, la chaleur latente, la nébulosité et les précipitations.
- **Les concentrations initiales** sont éventuellement demandées, elles peuvent être prises à partir des données d'observation ou à partir des résultats de modèle.
- **Les concentrations au bord** sont nécessaires pour les processus de transport. Pour un modèle à aire limitée, un modèle d'échelle plus grande par rapport au domaine d'étude peut fournir ces conditions aux limites.
- **Les émissions** des espèces apparaissent dans l'équation (2.39), à travers le terme E pour les surfaciques et S pour les volumiques. Elles proviennent de l'inventaire des émissions, basée notamment sur le trafic routier et sur les émissions industrielles.
- Finalement nous devons rajouter des données utilisées par les différents processus pris en compte dans le modèle CTM par exemple les données caractérisant le type de terrain "**Land use Coverage data**" pour le processus de dépôt (données fournies par des instituts spécialisés tel que USGS (United States Geological Survey) "edcdaac.usgs.gov/glcc/glcc.html"), des **constantes photolytiques** pour la chimie d'ozone...

2.1.5 Sorties

Les sorties du modèle sont les concentrations des espèces chimiques sur tout le domaine et à des fréquences temporelles de sauvegarde définies auparavant. Naturellement, si le CTM traite des processus plus complexes, d'autres sorties peuvent être produites, par exemple des distributions d'aérosol.

2.2 Les diverses paramétrisations utilisées

Nous revenons à présent plus en détails sur les diverses paramétrisations physiques et chimiques utilisées.

La Figure 2.1 montre la structure typique d'un modèle de Chimie-Transport comme POLAIR3D.

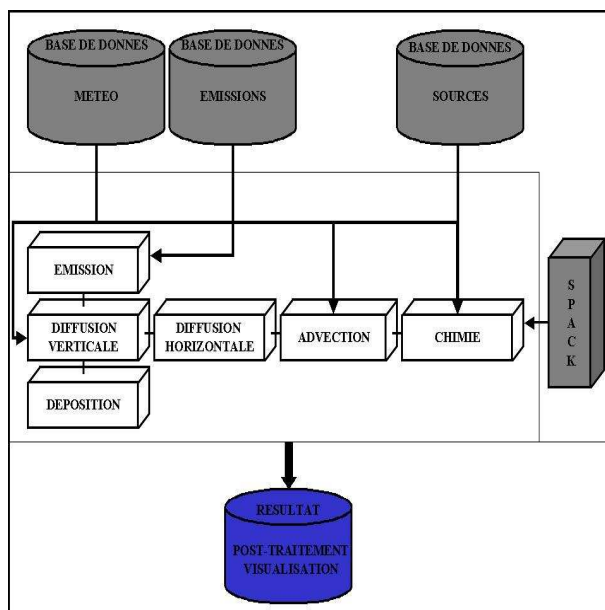


Figure 2.1: Structure du modèle de chimie-transport POLAIR3D

2.2.1 Émissions et sources

Les émissions sont parmi les plus importantes données d'entrée pour tous les modèles de Chimie-Transport. Elles sont données à partir d'inventaire d'émission sur la base de modèles ou de collecte de données. Elles sont particulièrement incertaines.

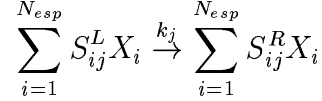
Les sources diffuses issues du trafic routier ou de l'activité biogénique sont prises classiquement comme conditions aux limites de la diffusion verticale (émissions surfaciques) tandis que les sources ponctuelles issues des rejets industriels sont traitées comme des émissions volumiques dans les premiers niveaux en fonction de la surhauteur de rejet estimée. Ces sources sont des termes de gain, rajoutés au terme de chimie L_{chem} de l'équation du transport (2.39).

L'ensemble de ces sources est généralement fourni à partir d'inventaires des agences environnementales, en général sous forme de totaux annuels. Pour les utiliser dans le modèle POLAIR3D un travail de prétraitement est primordial afin de pouvoir disposer des émissions et des sources à chaque point de grille et à chaque pas de temps. Nous reviendrons sur cette procédure dans le prochain chapitre.

2.2.2 Chimie

2.2.3 Cinétique chimique

Nous supposons que le schéma cinétique est décrit par un ensemble de N_r réactions et N_{esp} espèces chimiques. Les réactions sont symbolisées par :



où S_{ij}^L et S_{ij}^R sont des entiers qui composent la matrice dite stoechiométrique de dimension $(N_{esp} \times N_r)$ décrivant la perte et la production en nombre de molécules de l'espèce X_i dans la réaction j et k_j est la constante cinétique de la réaction qui dépend du temps t , la température T , et éventuellement d'autres variables selon le processus décrit.

La vitesse de la réaction j , ω_j , est exprimée à partir de la loi d'action de masse pour la cinétique chimique, en fonction de sa constante cinétique k_j , selon :

$$\omega_j = k_j \prod_{i=1}^{N_{esp}} X_i^{S_{ij}^L} \quad (2.8)$$

qui est le produit de toutes les concentrations de la partie gauche de la réaction. En prenant en compte le bilan de toutes les réactions qui font intervenir l'espèce X_i , son équation différentielle représentant le terme de production et de perte s'écrit :

$$\frac{\partial X_i}{\partial t} = \chi_i(X, T, L, I) = \sum_{j=1}^{N_r} (S_{ij}^R - S_{ij}^L) \omega_j(X_i, T, L, I) \quad (2.9)$$

Ce système peut être écrit pour $X = (X_1, X_2, \dots, X_{N_{esp}})^T$ sous la forme :

$$\frac{\partial X}{\partial t} = \chi(X, T, L, I) = S\omega(X, t) \quad (2.10)$$

où $S = S_{ij}^R - S_{ij}^L$ est la matrice stoechiométrique et $\omega(X, t) = \omega_j(X, t)$ le vecteur de vitesses de réactions de dimension N_R .

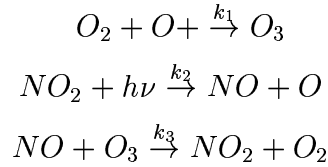
Le système 2.10 est souvent utilisé sous la forme :

$$\frac{\partial X_i}{\partial t} = P_i(t, X(t)) - L_i(t, X(t)) \cdot X_i(t) \quad (2.11)$$

où P est le vecteur de production, $P_i(t, X) = \sum_{j=1}^{N_r} S_{ij}^R \omega_j(t, X)$, et L le terme de perte, $L_i(t, X) = \sum_{j=1}^{N_r} S_{ij}^L \omega_j(t, X) / X_i$.

2.2.3.1 Un exemple : Cycle de Chapman troposphérique

Pour illustrer les différentes équations citées ci dessus, nous prenons comme exemple le cycle de Chapman de l'ozone troposphérique avec trois réactions faisant intervenir les espèces chimiques : $X_1 = [O_3]$, $X_2 = [O]$, $X_3 = [NO]$ et $X_4 = [NO_2]$



la matrice stoechiométrique S correspondante à ce mécanisme étant alors :

$$S = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

le vecteur de vitesses de réactions $\omega(t, X) = (k_1 X_2, k_2 X_4, k_3 X_1 X_3)^T$

La terme de production chimique pour l'ozone est alors : $\chi_{O_3} = \omega_1 - \omega_3$. le système d'équations différentielles du Cycle de Chapman s'écrit donc sous la forme :

$$\frac{d[O_3]}{dt} = k_1[O_2][O] - k_3[NO][O_3] = \omega_1 - \omega_3 \quad (2.12)$$

$$\frac{d[NO_2]}{dt} = k_3[NO][O_3] - k_2[NO_2] = \omega_3 - \omega_2 \quad (2.13)$$

$$\frac{d[NO]}{dt} = -k_3[NO][O_3] + k_2[NO_2] = \omega_3 - \omega_2 \quad (2.14)$$

$$\frac{d[O]}{dt} = k_1[O_2][O] + k_2[NO] = \omega_2 - \omega_1 \quad (2.15)$$

2.2.3.2 Les constantes cinétiques

Les constantes cinétiques sont calculées selon le type de réaction :

- **Les réactions thermiques**, pour lesquelles on utilise la loi de Arrhenius :

$$k_j(T) = A_j T^{B_j} \exp(-C_j/T) \quad (2.16)$$

A_j , B_j et C_j sont des constantes données pour chaque réaction du mécanisme.

– **Les réactions "fall-off"**

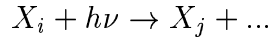
Elles font intervenir des espèces intermédiaires dites de désactivation collisionnelle : $A + B(+M) \xrightarrow{k_1} C(+M)$ où M est l'espèce intermédiaire, dans le cas atmosphérique généralement N_2 , O_2 et H_2O . Leurs constantes cinétiques dépendent de la température (T) et de la pression (P) et s'écrivent selon la paramétrisation de Troe [33] par exemple sous la forme :

$$k_j(P, T) = \frac{k_0(T)M}{(1 + k_0(T)M)k_1(T)} \cdot 0.6^{(1 + (\log_{10}(k_0(T)k_1(T)M))^2)^{-1}} \quad (2.17)$$

avec $k_0(T) = A \cdot 10^B (T/300)^{-n}$, $k_1(T) = D \cdot 10^E (T/300)^{-m}$, M est la concentration du tiers corps. Les paramètres A, B, D, E, n et m sont donnés pour chaque réaction du mécanisme de ce type.

– **Les réactions photolytiques**

Ce sont des réactions de photo-dissociation causées par le rayonnement solaire. Pour un photon de fréquence ν , on note classiquement les réactions selon :



La vitesse de réaction s'exprime par :

$$\omega_j(X, I) = J_j(I)X_i \quad (2.18)$$

où la constante photolytique J_j est calculée selon :

$$J_j(\theta) = \int_{\lambda_j^{min}}^{\lambda_j^{max}} I(\theta) q_\lambda(j) \sigma_\lambda^a(i) d\lambda \quad (2.19)$$

où θ est l'angle zénital, $I(\theta)$ est le flux actinique, $q_\lambda(j)$ est le rendement quantique pour la réaction j et $\sigma_\lambda^a(i)$ est la section efficace d'absorption pour l'espèce i , λ^{min} (Resp. λ^{max}) la limite inférieure (Resp. supérieure) des longueurs d'onde considérées pour la photodissociation. Sous certaines conditions de ciel clair, la constante photolytique est souvent paramétrisée en fonction de θ , par :

$$J_j(\theta) = A \cdot \exp(-B / \cos(\theta)) \quad (2.20)$$

où A et B sont des constantes qui dépendent de la réaction photochimique et des espèces prises en compte.

En pratique, ces constantes photolytiques dans POLAIR3D sont calculées en 'off-line' avec le modèle radiatif FAST-J ([34]) ou fournies par le module JPROC de CMAQ (<http://www.epa.gov/AMD/models3/doc/science/science.html>).

2.2.4 SPACK : Préprocesseur chimique

Le calcul des constantes cinétiques, des flux de réactions, du Jacobien associé à la cinétique chimiques (pour des résolutions numériques implicites), ..., sont des calculs préliminaires nécessaires pour résoudre numériquement l'équation de transport-chimie.

Ceci justifie l'utilité d'un préprocesseur indépendant du code de calcul. Nous utilisons en pratique le préprocesseur SPACK (Simplified Preprocessor for Atmospheric Chemical Kinetics) ([35]), ses fonctionnalités étant principalement :

- le passage de la forme réaction chimique à l'équation de cinétique chimique selon les équations (2.10) et (2.11) ;
- l'identification des espèces éventuelle de lumping ;
- la production d'une manière automatique de codes sources en Fortran pour un modèle de chimie transport (POLAIR3D dans notre cas) du terme $L_{chem}(c)$ et la matrice jacobienne associée $A_{chem} = \frac{\partial l_{chem}}{\partial c}$.

La fonction $L_{chem}(c)$ est donnée sous la forme ($\dot{c} = L_{chem}(c)$) ou séparée en terme de production et de consommation ($\dot{c} = P(c) - L(c)c$)

En entrée SPACK a besoin essentiellement d'un fichier décrivant de manière symbolique le mécanisme réactionnel. Le Tableau 2.1 montre un exemple des deux fichiers d'entrée de SPACK pour le mécanisme RACM.

2.2.5 Diffusion

2.2.5.1 Diffusion verticale K_z

Nous utilisons la paramétrisation classique de Louis pour la diffusion verticale due à la turbulence ([31]) :

$$K_z = l(z)^2 \left\| \frac{\partial V}{\partial z} \right\| F(R_i, z) \quad (2.21)$$

$$l(z) = \frac{K_a \cdot z}{1 + \frac{k_a \cdot z}{L}} \quad (2.22)$$

où $K_a \simeq 0.4$ est la constante de Von Karman et z_0 et L sont des facteurs d'échelles ($(z_0 \simeq 1 \text{ m et } L \sim 100 \text{ m})$).

<pre> File for chemical species CBM4(name and molar mass) # gaseous species # aqueous species 31 0 ---Gas-phase--- RXPAR 9.926E-06 OH 7.356E-05 PHO 1.292E-06 C2O3 6.776E-05 NO3 6.888E-05 HO2 4.888E-04 XO2 2.824E-04 HNO4 1.336E-03 N2O5 2.304E-03 </pre>	<pre> %% % File for gas-phase mechanism CBM4 % % Reference : % % Modifications : % %%% % UNIT GAS PPB %-----1----- NO2 > NO + O3P % KIN ARR1 0.5833D-2 % END </pre>
--	--

Table 2.1: fichiers d'entrée de SPACK

- R_i est le Nombre de Richardson :

$$R_i = g \frac{\partial \ln \theta}{\partial z} \frac{1}{\left\| \frac{\partial V}{\partial z} \right\|^2} \quad (2.23)$$

avec θ la température potentielle, g la constante de pesanteur et $\left\| \frac{\partial V}{\partial z} \right\|$ le module du gradient du vent horizontal.

- F est la fonction suivante :

$$\begin{cases} F(R_i, z) = \frac{1}{1 + 3B R_i \sqrt{1 + D R_i}} & \text{if } R_i \geq 0 \\ F(R_i, z) = 1 - \frac{3B R_i}{1 + 3B C \sqrt{\frac{|R_i|}{27}} \left(\frac{l(z)}{z + z_0} \right)^2} & \text{if } R_i \leq 0 \end{cases} \quad (2.24)$$

où habituellement $B = C = D = 5$.

Un exemple de profil vertical de K_z obtenu avec cette paramétrisation est illustré dans la Figure 2.2.

2.2.5.2 Diffusion horizontale K_x et K_y

La paramétrisation de la diffusion horizontale est un point faible de la modélisation, les processus de turbulence horizontale étant mal représentés. Il est classi-

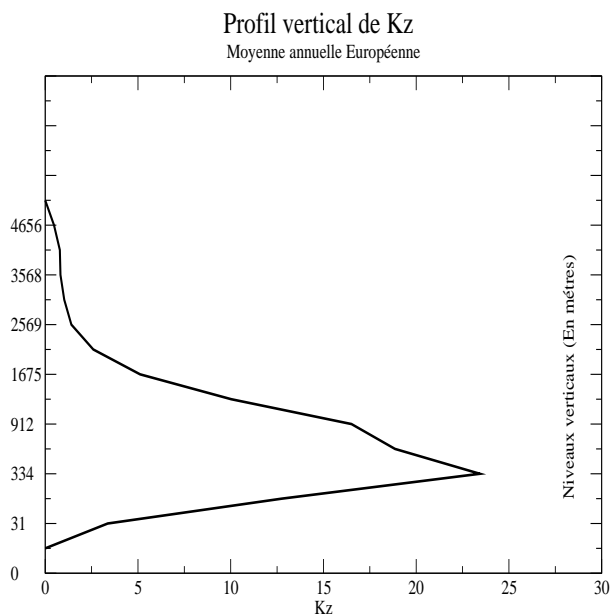


Figure 2.2: Profil de la diffusion verticale (moyenne annuelle européenne - année 2001)

quement recommandé de ne pas prendre de diffusion horizontale, de toute manière souvent noyée dans la diffusion numérique associée au schéma d'advection.

Dans les modèles de Chimie-Transport, une valeur constante est habituellement utilisée. Sa valeur dépend de l'échelle ; pour l'échelle régionale (typiquement $100\text{km} \times 100$ kilomètre), nous prenons typiquement :

$$K_x = K_y \simeq 500\text{m}^2/\text{s} \quad (2.25)$$

tandis que nous préconisons pour l'échelle continentale :

$$K_x = K_y \simeq 10^5\text{m}^2/\text{s} \quad (2.26)$$

Il est à noter que des paramétrisations de la diffusion horizontale plus sophistiquées existent dans la littérature [36].

Remarque : impact de la diffusion horizontale

Nous étudions l'impact de la diffusion horizontale à l'échelle continentale pour une chimie réactive utilisant le mécanisme RACM. Nous distinguons deux périodes l'été et l'hiver pour étudier l'effet de la diffusion horizontale sur la concentration des espèces chimiques. Nous calculons la moyenne spatiale sur la France (Figures 2.3

et 2.4), avec le modèle POLAIR3D. On remarque la faible incidence de la prise en compte à l'échelle européenne de K_H .

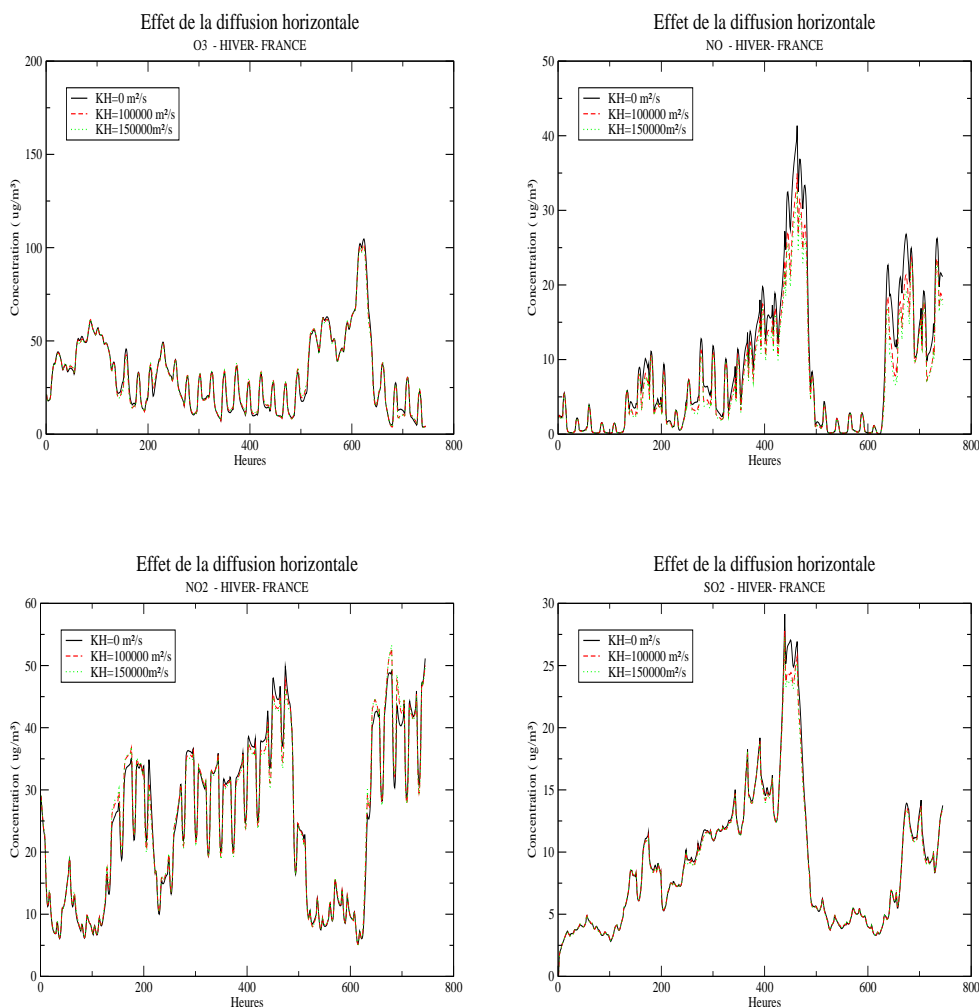


Figure 2.3: Comparaison des profils de concentration sur la France pendant l'hiver avec différentes valeurs de K_H

2.2.6 Paramétrisation sous maille : effet convectif

Un autre processus de transport vertical peut s'avérer important dans les modèles de chimie-transport, le transport convectif. La paramétrisation des processus convectifs est nécessaire pour des applications à grande échelle afin de tenir compte du transport sous-maille induit par les gradients d'humidité liés aux nuages de petite

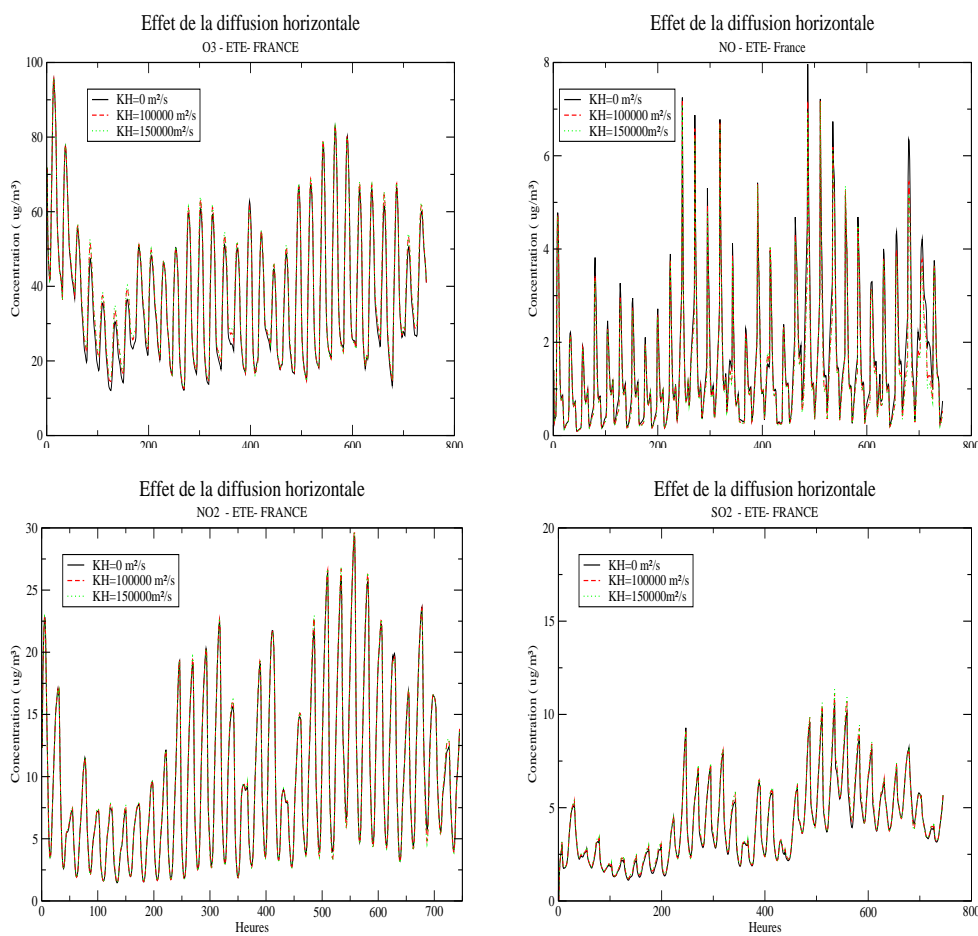


Figure 2.4: Comparaison des profils de concentration sur la France pendant l'été avec différentes valeurs de K_H

échelle (cumulus). Les flux de sous-maille sont définis de la façon suivante :

$$L_{conv}(c) = -\frac{\partial}{\partial z}(F_u(c) + F_d(c) + F_s(c)) \quad (2.27)$$

où F_u (resp. F_d) représente le flux ascendant (resp. descendant) et F_s est le flux de subsidence. Ces flux sont calculés sur la base des taux d'entraînement et de detraînement pendant l'ascendance et la descendance selon par exemple le schéma décrit par Tiedke ([32, 37]). Après discrétisation, le schéma proposé ([32]) peut s'écrire sous forme :

$$L_{conv}(c) = M_{conv}c \quad (2.28)$$

où M_{conv} exprimé en $kg.m^{-2}s^{-1}$ est déterminé par le modèle de nuage en fonction des données météorologiques ([32]).

Dans la version actuelle du modèle POLAIR3D le transport convectif n'est pas pris en compte.

2.2.7 Conditions initiales et aux limites

Afin d'initialiser les espèces chimiques du modèle nous avons besoin des conditions initiales qui peuvent être prises à partir des réseaux de mesures par le biais d'interpolation, des résultats antérieurs du modèle ou encore à partir d'autres modèles qui couvrent le domaine d'étude. Il est à noter, pour des applications sur de longues périodes en temps, la faible dépendance aux conditions initiales après une période transitoire de quelques jours (spin-up).

Les conditions aux limites (latérales, au sommet et au sol) sont nécessaires aux modèles à aire limitée. Nous distinguons deux types de conditions aux limites : pour l'advection et pour la diffusion.

Remarque : effet de la condition initiale

Nous étudions l'impact de l'état initial sur la qualité des résultats à l'échelle continentale du modèle POLAIR3D utilisé avec le mécanisme EMEP. Nous présentons, dans les figures 2.5 et 2.6, l'erreur relative moyenne sur la France, la Grande Bretagne et l'Allemagne, pendant 3 mois (Janvier, Février et Mars) de l'intégration, avec un pas de temps de 900 secondes. On remarque la décroissance très rapide.

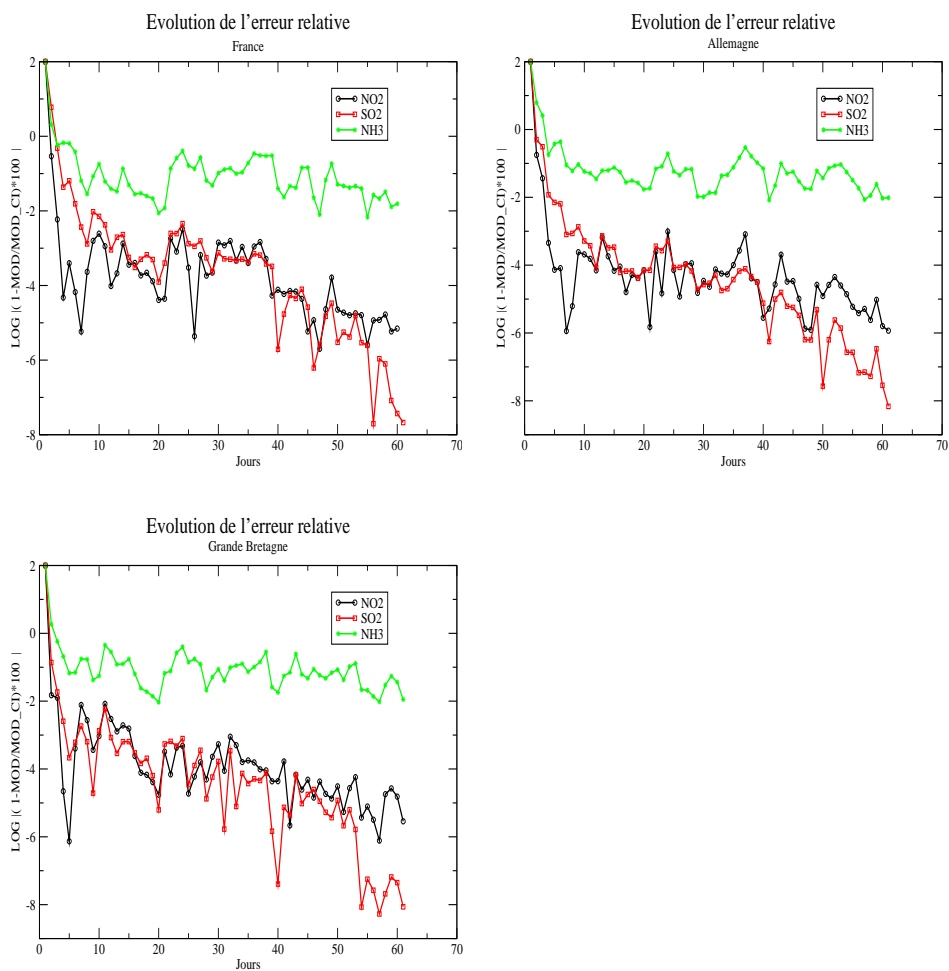


Figure 2.5: Effet de la condition initiale (France, Allemagne et Grande Bretagne) pour les espèces NO_2 , SO_2 et NH_3 .

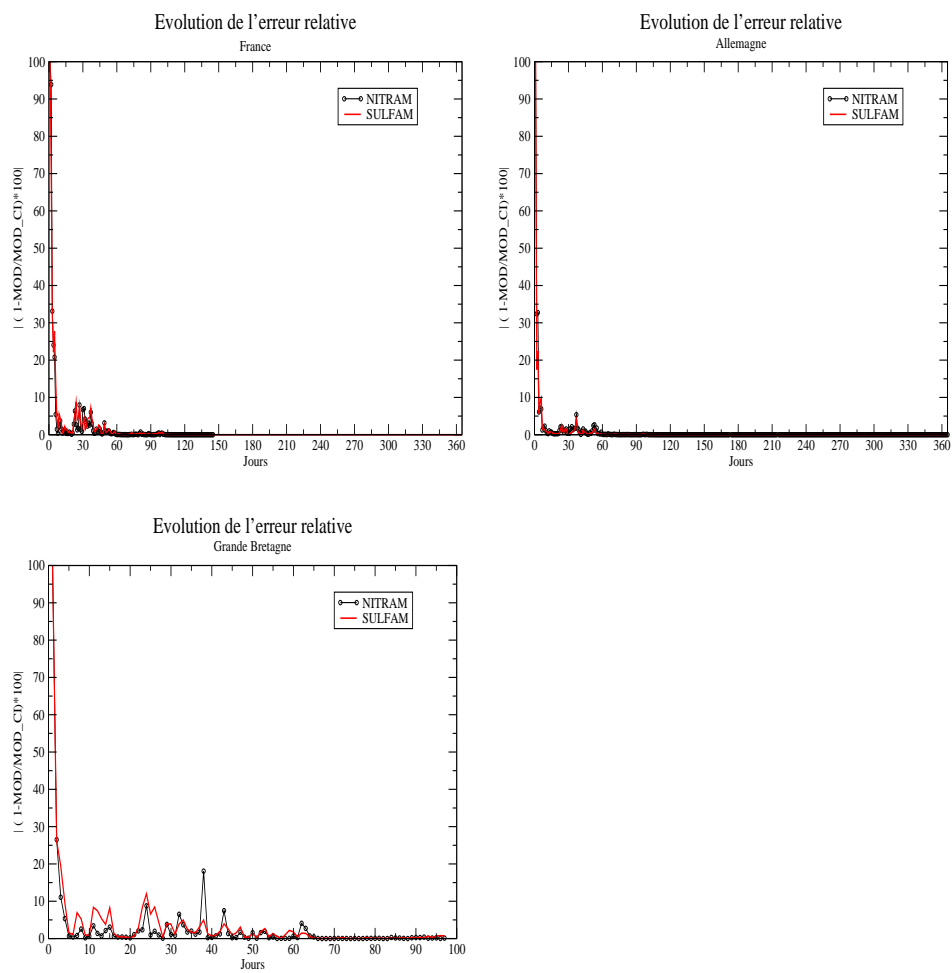


Figure 2.6: Effet de la condition initiale (France, Allemagne et Grande Bretagne) pour les espèces NITRAM et SULFAM.

2.2.7.1 Diffusion

Au sommet du domaine (limite supérieure), nous prenons un flux nul (atmosphère libre) :

$$K \frac{\partial c_i}{\partial n} = K \nabla c_i \cdot n = 0 \quad (2.29)$$

où n est le vecteur normal au domaine (orienté vers l'intérieur).

Au sol (limite inférieure du domaine) :

$$-K \frac{\partial c_i}{\partial n} = K \nabla c_i \cdot n = E_i(x, t) - v_i^{dep}(x, t) c_i \quad (2.30)$$

où v_i^{dep} est la vitesse de dépôt sec et $E_i(x, t)$ est le facteur d'émission surfacique pour l'espèce i . Dans les modèles Euleriens de qualité de l'air les processus de dépôt affectent les concentrations dans la couche la plus basse du domaine comme un flux de condition au limite.

Nous nous référons à la prochaine section pour le calcul des vitesses de dépôt secs.

Aux frontières latérales, nous avons un flux nul :

$$K \nabla c_i \cdot n = 0 \quad (2.31)$$

2.2.7.2 Advection

Pour les frontières le long desquelles $n \cdot v > 0$ (n est le vecteur normal orienté vers l'intérieur), nous utilisons la condition aux limites suivante :

$$n \cdot Vc = Flux \quad (2.32)$$

où V est le vecteur vent selon la direction x, y ou z . $Flux$ est typiquement donné par un modèle global. Pour les applications du modèle POLAIR3D à l'échelle continentale, les flux entrant sont calculés par le modèle MOZART [13]. Pour les applications à l'échelle régionale, les flux entrant sont calculés par POLAIRD3D à l'échelle continentale.

2.2.8 Processus de perte

2.2.8.1 Dépôt sec

Lors de leur transport, les espèces chimiques peuvent subir le processus de dépôt sec : il s'agit d'un transfert et d'une absorption par les surfaces naturelles (sol, végétation, mer, neige,...) d'une substance venant de l'atmosphère. Il dépend de :

- la taille des particules (si l'on considère des aérosols) ;

- la nature chimique (solide, gaz,...) et physique des espèces (solubilité, réactivité,...);
- la nature du terrain (cultivé, sauvage, eau, ...) et son état (sec, humide,...).

Le terme de dépôt intervient comme un terme de perte dans l'équation générale de Chimie-Transport. Il est défini sous forme d'un flux de perte de matière au niveau du sol, par un produit de la vitesse de dépôt à une certaine hauteur de référence Z_{ref} et la concentration de l'espèce chimique à cette hauteur, selon :

$$\frac{\partial c}{\partial t} = -\frac{\partial F_{D_{sec}}}{\partial z} ; \quad F_{D_{sec}} = \frac{\partial(V_d C)}{\partial Z_{ref}} \quad (2.33)$$

où V_d est la vitesse de dépôt sec en $m s^{-1}$. Elle est calculée de manière classique, dans le cas d'une chimie gazeuse, à partir des résistances des 3 couches représentatives des processus intervenant dans le dépôt sec :

$$v_d = \frac{1}{R_a + R_b + R_c} \quad (2.34)$$

où

- R_a ($s.m^{-1}$) est la résistance aérodynamique représentative de la couche de surface dominée par la diffusion turbulente.
- R_b ($s.m^{-1}$) est la résistance de surface ou quasi laminaire, relative à la couche laminaire dominée par la diffusion moléculaire malgré la taille de cette couche qui reste très fine, quelques millimètres.
- R_c ($s.m^{-1}$) est la résistance de transfert dans la canopée : elle caractérise le phénomène d'absorption et de solubilité de l'espèce chimique par la canopée.

Pour quantifier les valeurs de ces résistances, il existe plusieurs paramétrisations, qui diffèrent sur la façon de calculer la résistance R_c , mais en général équivalentes pour l'évaluation des résistances R_a et R_b . Dans le modèle Polair3D les deux paramétrisations "Baer et Nester" ([38]) et "Wesely" ([39]) sont disponibles. Dans la Figure 2.7 nous schématisons l'évolution de la vitesse du dépôt avec les deux paramétrisations différentes, durant l'année 2001, la fréquence temporelle est de 6h (4 valeurs par jour).

Il est à noter l'extrême sensibilité de la vitesse de dépôt à la saison (du fait de l'état des sols) et les disparités de résultat selon les paramétrisations utilisées. De manière générale, les résultats d'un modèle de Chimie-Transport sont très sensibles au calcul de la vitesse de dépôt.

2.2.8.2 Lessivage humide

Contrairement au dépôt sec, le lessivage humide correspond à une paramétrisation des échanges entre la phase gazeuse (soluble) et l'eau liquide présente dans

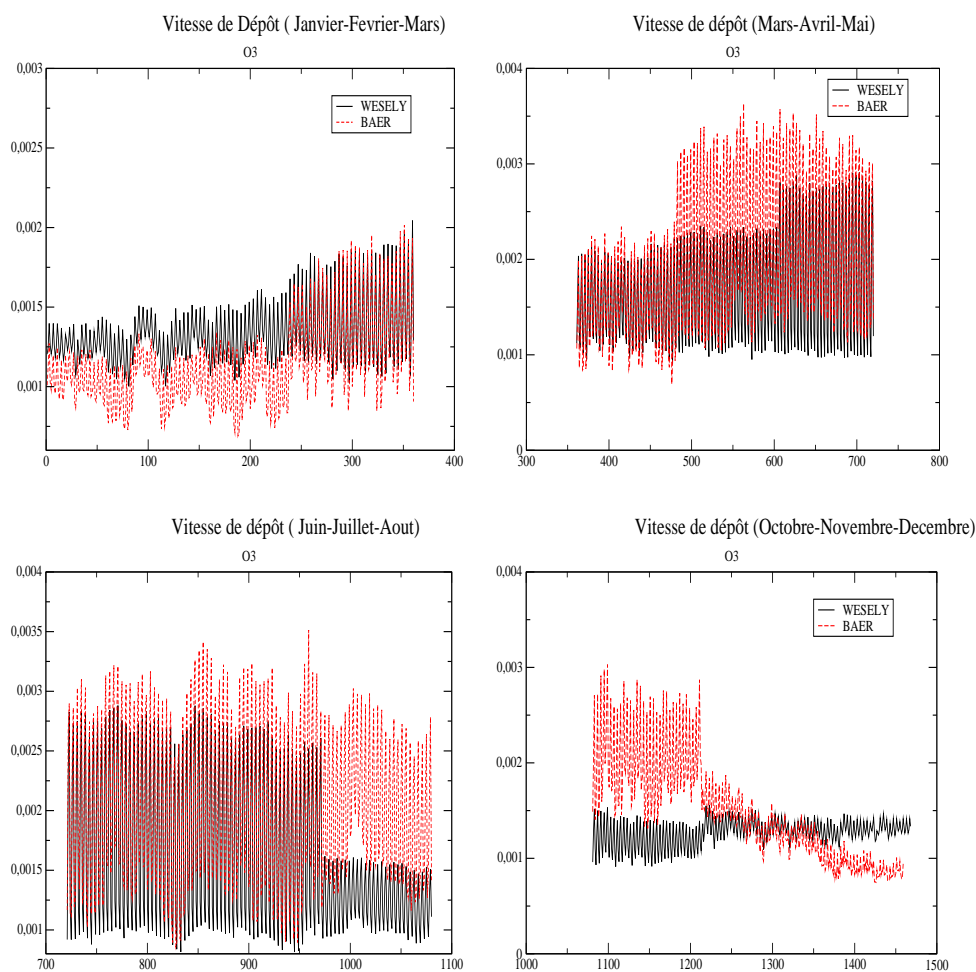


Figure 2.7: Profil de la vitesse de dépôt en ms^{-1} . Moyenne européenne selon les quatre trimestres de l'année 2001. En ligne la paramétrisation de Weesley, en pointillé celle de Baer-Nester. En abscisse est indiqué le jour, en ordonnée la vitesse.

l’atmosphère. Il caractérise l’échange de masse entre l’espèce chimique en phase gazeuse et l’eau nuageuse (“In cloud scavenging”) ou précipitante (“Below cloud scavenging”). Le terme de lessivage humide est pris comme un terme de puits dans l’équation de Chimie-Transport. Sa paramétrisation est exprimée selon que l’espèce chimique se trouve à l’intérieur du nuage ou à l’extérieur sous la forme :

$$L_{wet}(c_i) = -\Lambda_i(t)c_i \quad (2.35)$$

où Λ_i est une paramétrisation du lessivage humide calculée en fonction des données météorologiques. Dans le cas d’un lessivage par eau précipitante le terme Λ_i est exprimé selon ([40]) :

$$\Lambda_i = \frac{10^{-6} p_0}{3.6 h_{rain}} H_i R T (1 - \exp(-\frac{h_{rain} H_i R T U_{rain}}{6 D K_i})) \quad (2.36)$$

où p_0 l’intensité de pluie (en mm.hr⁻¹), h_{rain} le sommet de la colonne précipitante (en m), H_i le coefficient d’Henry pour l’espèce i (en mol.l⁻¹.atm⁻¹), R la constante des gaz parfaits (en atm.l.mol⁻¹.K⁻¹), U_{rain} la vitesse de pluie (en m.s⁻¹), D le diamètre de la goutte de pluie (en centimètre) et K_i ($s-1$) est un coefficient de transfert. Pour le lessivage par eau nuageuse nous nous référons à [40] pour une description plus détaillé de ce processus, des modèles de cinétique chimique en phase aqueuse étant habituellement pris en compte.

Dans la version du modèle POLAIR3D (utilisée pour ces travaux) , nous ne tenons en compte que le lessivage par eau précipitante.

2.3 Description de la plateforme numérique de Chimie-Transport, POLAIR3D

Les modèles récents de Chimie-Transport CTM sont basés sur la conservation de la masse et sont résolus sur une grille eulérienne.

Nous supposons à ce stade qu’un modèle a été bien choisi en prenant en compte tous les processus physique et chimiques important dans la dispersion des espèces chimiques, sur la base des paramétrisations appropriées comme explicitées dans les sections précédentes.

La simulation des équations de dispersion réactive est particulièrement difficile et nous essayerons dans cette section de mettre l’accent sur quelques points spécifiques comme :

- l’utilisation de méthodes de splitting (découplage des processus) ;
- l’utilisation des solveurs raides appropriés pour des problèmes multi-échelles comme la chimie atmosphérique ;

– la résolution à proprement parler de l'advection, de la chimie et de la diffusion. De manière plus générale, on peut se référer à [2].

2.3.1 Géométrie du modèle

2.3.1.1 Coordonnées

La complexité des modèles de chimie transport provient aussi de la variabilité spatiale et temporelle des phénomènes simulés. Nous distinguons différentes applications :

- une application *régionale* (photochimique), pour laquelle nous utiliserons des coordonnées cartésiennes (x, y, z) ;
- une application *continentale* (dispersion de polluants à l'échelle de l'Europe) ou globale (à l'échelle du globe), pour laquelle on utilise souvent les coordonnées sphériques (λ, ϕ) en horizontal et en vertical la coordonnée pression, hybride ou la hauteur z .

En coordonnée sphérique l'équation du modèle s'écrit ainsi :

$$\begin{aligned}
& \frac{\partial c_i}{\partial t} + \frac{1}{r \cos \phi} \left(\frac{\partial u c_i}{\partial \lambda} + \frac{\partial v c_i \cos \phi}{\partial \phi} \right) + \frac{1}{r^2} \frac{\partial w c r^2}{\partial z} \\
& = \frac{\partial}{\partial \lambda} \left(\frac{K_\lambda}{r^2 \cos^2 \phi} \frac{\partial c}{\partial \lambda} \right) + \frac{1}{r^2 \cos \phi} \frac{\partial}{\partial \phi} \left(K_\phi \cos \phi \frac{\partial c}{\partial \phi} \right) + \frac{\partial}{\partial z} \left(K_z \frac{\partial c}{\partial z} \right) \\
& + \chi_i(c, \lambda, \phi, z, t) \\
& + E_i(\lambda, \phi, t) + S_i(\lambda, \phi, z, t) - v_{dep}^i(\lambda, \phi, t) \cdot c_i - \Lambda_i(t) c_i + F(c_i)
\end{aligned} \tag{2.37}$$

où u, v, w sont les composants de la vitesse du vent dans le système de coordonnées (λ, ϕ, z) , λ et ϕ désignent respectivement la longitude et la latitude. z est la coordonnées verticales, $r \simeq 6378$ kilomètres est le rayon de la terre, supposé constant.

Il peut être intéressant de ne disposer que d'une seule formulation pour les deux applications. C'est le choix que nous avons fait afin de n'avoir qu'un seul et même modèle de manière aisée.

Après un changement de variable [41] :

$$\begin{cases} x = r \cdot \lambda \\ y = r \cdot \sin(\phi) \\ z = Z - Z_{sol} \end{cases} \tag{2.38}$$

L'équation en coordonnées sphériques s'écrit dans le nouveau système de coordonnées curvilignes (x,y,z) selon :

$$\boxed{\frac{\partial c_i}{\partial t} + \left(\frac{\partial U^* c_i}{\partial x} + \frac{\partial V^* c_i}{\partial y}\right) + \frac{\partial w c_i}{\partial z} = \frac{\partial}{\partial x} \left(K_x \frac{\partial c_i}{\partial x}\right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial c_i}{\partial y}\right) + \frac{\partial}{\partial z} \left(K_z \frac{\partial c_i}{\partial z}\right) + \chi_i(c, x, y, z, t) + E_i(x, y, t) + S_i(x, y, t) - v_{dep}^i(x, y, t) \cdot c_i - \Lambda_i(t) c_i + F(c_i)}$$
(2.39)

où

$$U^* = u / \cos \phi \quad V^* = v * \cos \phi \quad (2.40)$$

$$K_x = K_\phi / \cos^2 \phi \quad K_y = K_\lambda * \cos^2 \phi \quad (2.41)$$

L'équation (2.39) est l'équation générale qui gouverne le modèle POLAIR-3D, et elle est valable pour une grille dite cartésienne ou une grille sphérique (modulo le changement de variable 2.38 et les modifications des variables d'entrée 2.40 et 2.41).

Remarque : vitesse verticale

Afin d'assurer la continuité du mouvement, la vitesse verticale w est rediagnostiquée à partir du vent horizontale (U^*, V^*) sur la grille du modèle, pour respecter :

$$\text{div } U = 0 \quad (2.42)$$

2.3.1.2 Grille du modèle

Nous avons fait le choix d'une approche de type "différences finies" pour la discrétisation spatiale. L'équation d'évolution du modèle (2.39) est intégrée sur une grille de de type 'cell-center' avec N_x points selon la direction x , N_y points selon la direction y et N_z niveaux verticaux. Le calcul est fait sur une grille irrégulière composée des points de centres de mailles, schématisée sur les figures 2.8 (grille horizontale) et 2.9 (grille verticale).

- Les données relatives à l'advection sont données en cohérence avec le choix du schéma d'advection et de diffusion, sur les facettes de la maille au point $x_{i+\frac{1}{2}}$, $y_{i+\frac{1}{2}}$ et $z_{i+\frac{1}{2}}$ respectivement pour U , V et le couple (W, K_z) .
- Les autres données nécessaires aux modules de diffusion (émissions, sources,...), à la chimie (température, humidité, pression, les constantes photolytique) , du dépôt (vitesse de dépôt) et la concentration des espèces, sont données aux centres de mailles.

Remarque : Prise en compte du relief

Les données météorologiques issues des modèles numériques sont souvent calculées sur une grille sphérique en latitude, longitude et en coordonnée verticale hybride η , afin de mieux prendre en compte les caractéristiques du relief. La relation clé est pour la pression :

$$p(\lambda, \phi, \eta, t) = A(\eta) + B(\eta)p_g(\lambda, \phi, t) \quad (2.43)$$

où p et p_g sont respectivement la pression et la pression au niveau du sol. A et B sont des fonctions données.

Nous avons fait le choix de travailler sur une grille verticale en hauteur relative (en mètres) par rapport au relief ($h(x, y)$), ($Z_h(x, y)$).

En utilisant l'approximation hydrostatique, nous pouvons retrouver les hauteurs (en mètres) de chaque niveau hybride (www.ecmwf.int) et effectuer les interpolations nécessaires des données météo sur la grille verticale (en hauteur) du modèle de Chimie-Transport.

2.3.2 Méthodes de séparation d'opérateurs : Splitting

Une fois discrétisées spatialement (voir points suivants), l'équation d'évolution de la concentration c_i s'écrit :

$$\frac{dc_i}{dt} = A(c_i) + B(c_i) + D(c_i) + E(c_i) \quad (2.44)$$

La résolution couplée du système 2.44 aux dérivées partielles décrivant le modèle CTM avec une discrétisation spatiale tridimensionnelle demande la résolution d'un système non-linéaire raide d'une grande dimension. Elle nécessitera l'utilisation de solveurs implicites et l'inversion de systèmes de grande dimension, typiquement le produit du nombre de points de grille par le nombre d'espèces chimiques. Pour une chimie d'une centaine d'espèces et avec une application européenne discrétisée selon ($N_x = 47$, $N_y = 32$, $N_z = 20$), une telle résolution demanderait d'inverser un système de taille supérieur à 10^7 et cela à chaque pas de temps de simulation du modèle.

Avec l'état de l'art des moyens de calcul actuels une telle approche est non utilisable en pratique, surtout dans un contexte opérationnel.

Il existe plusieurs méthodes pour résoudre ce problème ([42]), fondées sur un découplage des processus. C'est cette méthode que nous allons détailler dans la section suivante et celle utilisée dans le modèle POLAIR3D.

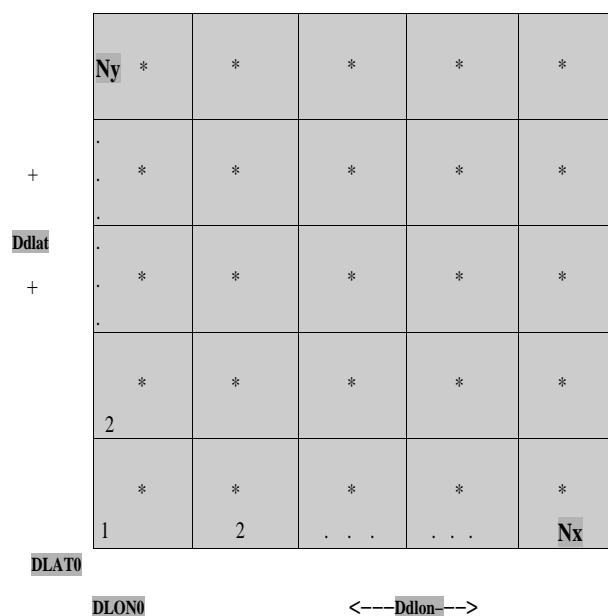


Figure 2.8: Grille horizontale où $Ddlon$, $Ddlat$ sont les résolutions spatiales en degrés selon respectivement la longitude et la latitude. $Dlon0$ et $Dlat0$ désignent la position du premier point de la discrétisation

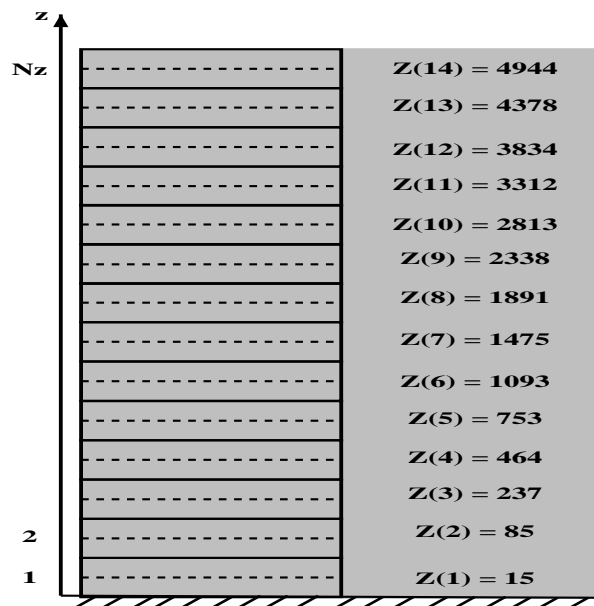


Figure 2.9: Grille verticale du modèle (Application EMEP à l'échelle Européenne) avec 14 niveaux verticaux distribués irrégulièrement. Z désigne la hauteur des centres de mailles en mètre.

La méthode de séparation des processus (dite méthodes de splitting) est habituellement utilisée ([43, 44]) afin de découpler les phénomènes réactifs locaux (un modèle chimique de boîte par cellule de grille) et le transport passif (un problème d'advection-diffusion par espèce chimique).

Un autre point important avec l'utilisation des méthodes de splitting est la possibilité d'employer des sous-modèles en tant que boîtes noires. Dans la pratique, l'advection, la diffusion et la chimie sont résolues dans un ordre séparé.

Pour illustrer les différentes méthodes de splitting nous prenons le cas linéaire simple avec deux processus (A et B). Nous considérons l'équation suivante d'évolution :

$$\frac{dc}{dt} = Ac + Bc, \quad c(0) = c_0 \quad (2.45)$$

avec A et B deux processus.

2.3.2.1 Splitting standard : 1er ordre

La méthode de splitting de premier ordre "AB" est définie par les deux étapes successives suivantes :

- première étape : intégration de l'opérateur A

$$\frac{dc^*}{dt} = Ac^* \quad \text{sur} \quad [t_n, t_n + \Delta t], \quad c^*(t_n) = c_n \quad (2.46)$$

- deuxième étape : intégration de l'opérateur B

$$\frac{dc^{**}}{dt} = Bc^{**} \quad \text{sur} \quad [t_n, t_n + \Delta t], \quad c^{**}(t_n) = c^*(t_n + \Delta t) \quad (2.47)$$

ce qui définit la valeur $c_{n+1} = c^{**}(t_n + \Delta t)$.

La solution est finalement :

$$c_{n+1} = e^{B\Delta t} e^{A\Delta t} c_n \quad (2.48)$$

à comparer à la solution exacte

$$c_{n+1} = e^{(A+B)\Delta t} c_n \quad (2.49)$$

l'erreur locale due au splitting est alors en premier ordre en Δt :

$$Err_n = \frac{1}{2} \Delta t (AB - BA) c_n + O(\Delta t^2) \quad (2.50)$$

L'erreur est donc associée au défaut de commutation $AB - BA$. Pour une extension au cas non linéaire [45, 46, 47] l'erreur locale de splitting s'écrit :

$$Err_n = \frac{1}{2}\Delta t \left(\frac{\partial A}{\partial c} B - \frac{\partial B}{\partial c} A \right) (t_n, c_n) + O(\Delta t^2) \quad (2.51)$$

2.3.2.2 Splitting de second ordre

La méthode de séparation d'opérateurs de deuxième ordre (Strang splitting, [48, 2]) est souvent utilisée pour améliorer la précision du splitting du premier ordre. Le splitting "BAB" se déroule selon les trois étapes suivantes pour remédier au défaut de "symétrie" des méthodes de second ordre :

- Première étape : Intégration de l'opérateur B

$$\frac{dc^*}{dt} = Bc^* \quad \text{sur} \quad [t_n, t_n + \frac{\Delta t}{2}], \quad c^*(t_n) = c_n \quad (2.52)$$

- Deuxième étape : intégration de l'opérateur A

$$\frac{dc^{**}}{dt} = Ac^{**} \quad \text{sur} \quad [t_n, t_n + \Delta t], \quad c^{**}(t_n) = c^*(t_n + \frac{\Delta t}{2}) \quad (2.53)$$

- Troisième étape : intégration de l'opérateur B

$$\frac{dc^{***}}{dt} = Bc^{***} \quad \text{sur} \quad [t_n + \frac{\Delta t}{2}, t_n + \Delta t], \quad c^{***}(t_n + \frac{\Delta t}{2}) = c^{**}(t_n + \Delta t) \quad (2.54)$$

et $c_{n+1} = c^{***}(\frac{\Delta t}{2})$.

L'erreur locale de splitting associée au second ordre dans le cas d'un système non linéaire s'écrit [49] :

$$Err_n = \frac{1}{24}\Delta t^2 \left[\frac{\partial}{\partial c} \left(\frac{\partial A}{\partial c} A \right) B - 2 \frac{\partial}{\partial c} \left(\frac{\partial A}{\partial c} B \right) A + \frac{\partial}{\partial c} \left(\frac{\partial B}{\partial c} A \right) A - 2 \frac{\partial}{\partial c} \left(\frac{\partial B}{\partial c} B \right) A + 4 \frac{\partial}{\partial c} \left(\frac{\partial B}{\partial c} A \right) B - 2 \frac{\partial}{\partial c} \left(\frac{\partial A}{\partial c} B \right) B \right] (t_n, c_n) + O(\Delta t^2) \quad (2.55)$$

2.3.2.3 Source splitting

La méthode dite de "source-splitting" est une autre alternative [2]. L'idée est d'éviter les étapes transitoires en rajoutant des termes sources explicites. La première étape est classiquement :

$$\frac{dc^*}{dt} = Ac^* \quad \text{sur } [t_n, t_n + \Delta t], \quad c^*(t_n) = c_n \quad (2.56)$$

mais ne conduit pas à une modification de la condition initiale. L'opérateur B est plutôt intégré avec un terme source complémentaire :

$$\frac{dc^{**}}{dt} = Bc^{**} + \frac{c^*(\Delta t) - c_n}{\Delta t} \quad \text{sur } [t_n, t_n + \Delta t], \quad c^{**}(t_n) = c_n \quad (2.57)$$

$$c_{n+1} = c^{**}(\Delta t).$$

Notons que les conditions initiales pour la deuxième étape ne sont pas modifiées, ce qui évite toute étape transitoire si B est un modèle raide.

Remarque : Splitting dans un cas raide

Dans le cas où l'un des opérateurs de l'équation (2.45) est raide (prenons par exemple l'opérateur A tel que $||\Delta t A|| \gg 1$ [49]), il est suggéré dans [50, 2] de résoudre l'équation du système en finissant par l'opérateur le plus raide, A , dans le splitting de premier ordre ("B-A") et en commençant et finissant par l'opérateur le plus raide, A , dans le cas d'un splitting de second ordre ("A-B-A").

2.3.2.4 Splitting dans le modèle de chimie-transport POLAIR3D

Dans le cas d'un modèle de Chimie-Transport, l'équation d'évolution est sous la forme fonctionnelle suivante :

$$\boxed{\begin{aligned} \frac{\partial c_i}{\partial t} &= L_{adv}(c_i) + L_{diff}(c_i) + [L_{conv}(c)]_i + [L_{chem}(c)]_i + L_{source} \\ &+ L_{Emission} + L_{depot} \end{aligned}} \quad (2.58)$$

où chaque processus est explicité dans la section (2.1.3). Nous utilisons les différentes méthodes de séparation d'opérateurs présentées dans la dernière section précédente, le terme volumique des sources étant couplé au terme de la chimie. Comparée aux autres termes de cette équation, la chimie est l'opérateur le plus raide, le splitting se fait donc en prenant en compte cette raideur.

- **Splitting de premier ordre :**

La première approche est celle dite de premier ordre, le splitting se fait dans l'ordre suivant (Δt est le pas de temps de splitting) :

- Intégration de L_{adv} ($(L_{advx}, L_{advy}, L_{advz})$) sur $[t, t + \Delta t]$,
- Intégration L_{diffx} sur $[t, t + \Delta t]$,
- Intégration L_{diffy} sur $[t, t + \Delta t]$,
- Intégration L_{diffz} sur $[t, t + \Delta t]$,
- Intégration de L_{chem} sur $[t, t + \Delta t]$,

Après chaque intégration, les conditions initiales sont changées (section 2.3.2.1).

Nous présentons dans la figure 2.10 une comparaison des simulations effectuées par le modèle POLAIR3D avec un splitting de premier ordre :

- Diffusion puis chimie (D-C),
- Chimie puis diffusion (C-D),
- Chimie et diffusion sont traités d'une manière couplée par la méthode AMF (voir section suivante). Ce calcul est pris comme un cas de référence pour évaluer les deux approches (D-C) et (C-D).

On peut observer aisément que le calcul effectué avec un splitting (D-C) est très proche de la simulation AMF, alors que la simulation faite en changeant le sens du splitting pour un ordre Chimie -Diffusion donne des résultats plus éloignés du calcul AMF. Cela confirme notre choix dans l'ordre du splitting des opérateurs.

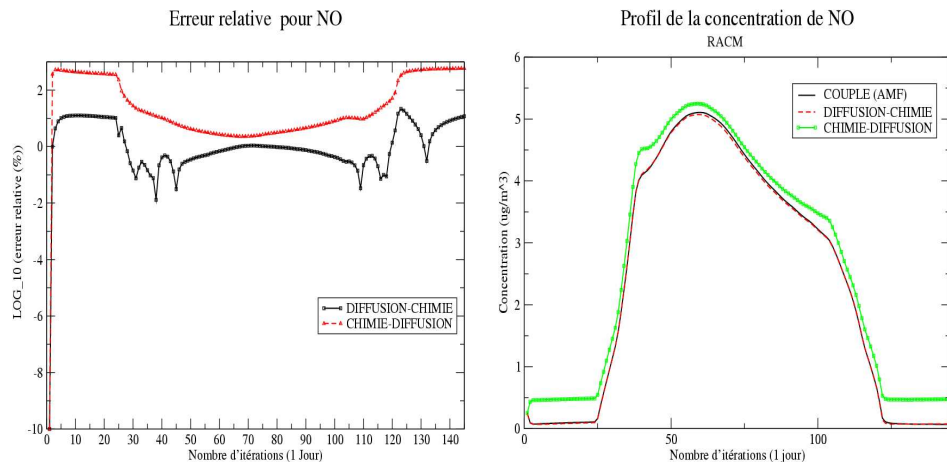


Figure 2.10: Comparaison DIFFUSION-CHIMIE vs CHIMIE-DIFFUSION : Erreur relative (en log) à gauche et le profil de NO à droite avec AMF (calcul de référence), Chimie-Diffusion et Diffusion-Chimie.

- **Strang Splitting :**

La seconde approche est celle dite de Strang, le splitting se fait dans l'ordre suivant (Δt est le pas de temps de splitting) :

- Intégrer L_{adv} les processus ($L_{advx}, L_{advy}, L_{advz}$) sur $[t, t + \frac{\Delta t}{2}]$,
 - Intégration de L_{chem} sur $[t, t + \frac{\Delta t}{2}]$,
 - Intégration L_{diffx} sur $[t, t + \frac{\Delta t}{2}]$,
 - Intégration L_{diffy} sur $[t, t + \frac{\Delta t}{2}]$,
 - Intégration L_{diffz} sur $[t, t + \Delta t]$,
 - Intégration L_{diffy} sur $[t + \frac{\Delta t}{2}, t + \Delta t]$,
 - Intégration L_{diffx} sur $[t + \frac{\Delta t}{2}, t + \Delta t]$,
 - Intégration L_{chem} sur $[t + \frac{\Delta t}{2}, t + \Delta t]$,
 - Intégration L_{adv} selon les séquences ($L_{advz}, L_{advy}, L_{advx}$) sur $[t + \frac{\Delta t}{2}, t + \Delta t]$,
- Après chaque intégration, des conditions initiales sont changées (section 2.3.2.2).

Nous prenons comme exemple le cas d'une chimie complexe non-linéaire de type RACM [51], nous présentons dans la figure 2.12 le résultat obtenu avec un splitting 'Strang' versus un splitting classique de premier ordre, avec POLAIR3D ($\Delta t = 600s$, pas de temps optimal selon la figure 2.11) sur une période d'un mois (Janvier 2001) et en moyenne sur la France. Les deux méthodes ont le même comportement pour les espèces présentées (NO , NO_2 , SO_2 et O_3), même si quelques différences sont à noter du fait des erreurs de splitting de ces deux méthodes.

- **AMF : Couplage des termes L_{chem} et L_{diffz}**

La troisième approche est la méthode de Strang avec un couplage du terme de production chimique et le terme de la diffusion verticale puisque la séparation de ces processus est la principale origine de l'erreur ([52, 50]) de splitting :

- Intégration L_{adv} sur $[t, t + \frac{\Delta t}{2}]$,
- Intégration L_{diffx} sur $[t, t + \frac{\Delta t}{2}]$,

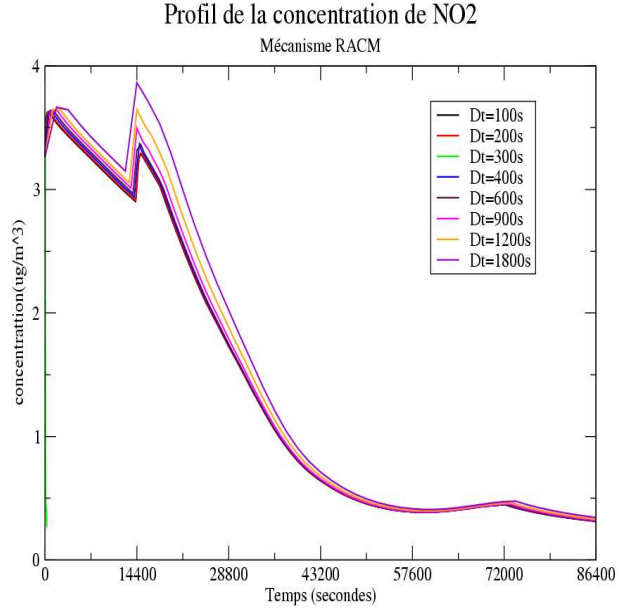


Figure 2.11: Profil de NO_2 selon plusieurs pas de temps de chimie, obtenu avec le modèle POLAIR0D

- Intégration L_{diffy} sur $[t, t + \frac{\Delta t}{2}]$,
- Intégration $L_{diffz} + L_{chem}$ sur $[t, t + \Delta t]$,
- Intégration L_{diffy} sur $[t + \frac{\Delta t}{2}, t + \Delta t]$,
- Intégration L_{diffx} sur $[t + \frac{\Delta t}{2}, t + \Delta t]$,
- Intégration L_{adv} sur $[t + \frac{\Delta t}{2}, t + \Delta t]$,

Les méthodes d'intégration des termes $L_{diffz} + L_{chem}$ sont présentées dans la section suivante.

Cette approche est basée sur un splitting au niveau algébrique du système :

$$f(c) = L_{diff}(c) + L_{chem}(c) , \quad A = A_{diff} + A_{chem} \quad (2.59)$$

avec l'approximation AMF (Approximate Matrix Factorization) :

$$(I - \gamma_i \Delta t A) \sim (I - \gamma_i \Delta t A_{chem})(I - \gamma_i \Delta t A_{diff}) + O(\Delta t^2) \quad (2.60)$$

ce qui nous permet d'éviter l'inversion du système dont la dimension est le produit du nombre de points de grille et le nombre d'espèces chimiques.

Dans la figure 2.13 nous présentons une comparaison des résultats obtenus

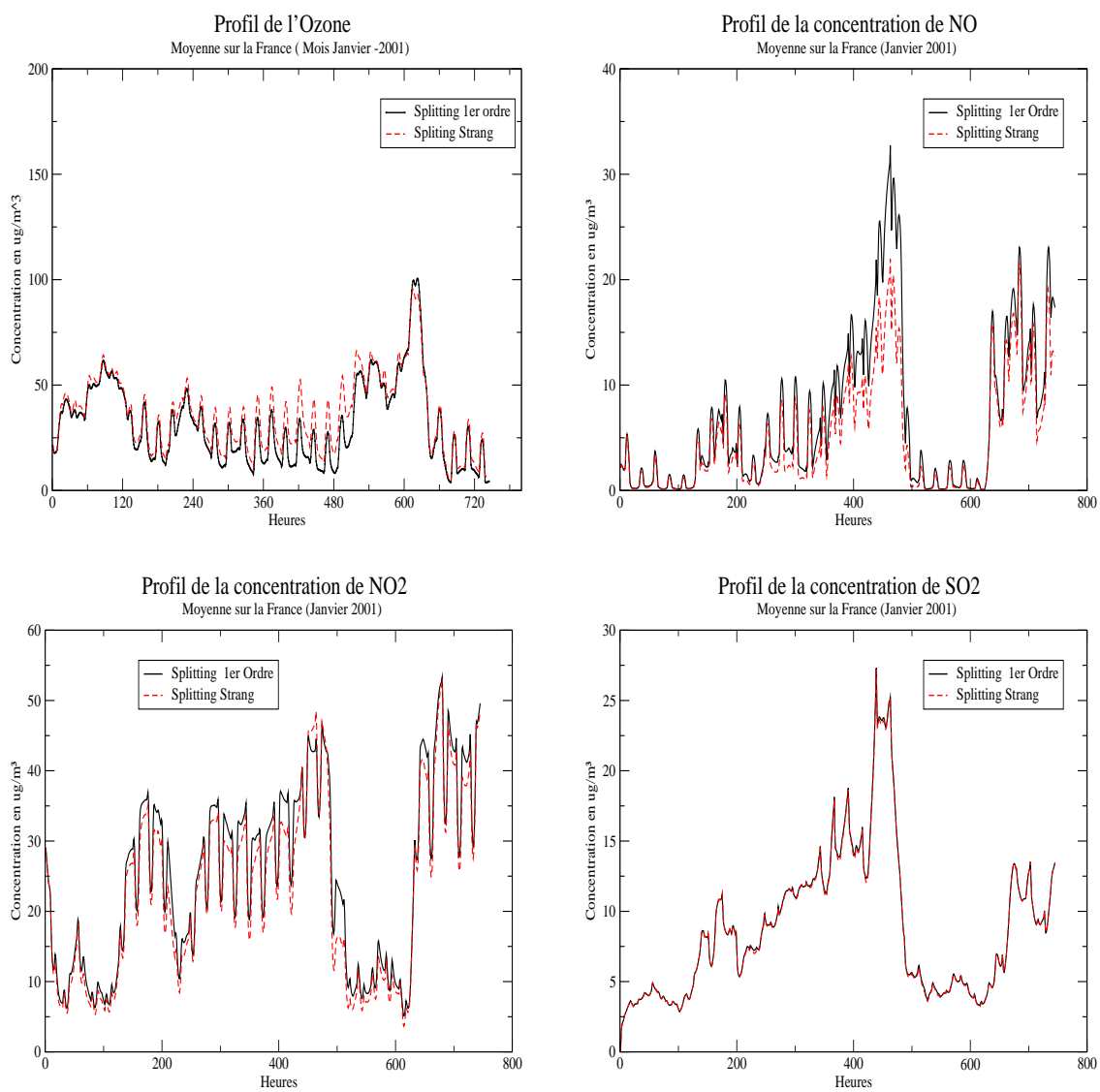


Figure 2.12: Comparaison Strang versus 1er ordre (Janvier 2001)

(en un point donnée situé sur la France) avec les deux méthodes AMF et la version où les termes de chimie et diffusion sont séparés. Le test est effectué avec le modèle POLAIR3D en utilisant le mécanisme chimique RACM et un pas de temps de 100 secondes pendant une intégration d'un jour. On peut constater que les résultats sont très semblables, l'erreur relative est inférieure à 1 %. Le tableau 2.2 regroupe les résultats obtenu avec les deux méthodes pour les espèces NO_2 , SO_2 , NO et l'ozone, les résultats sont très proches en moyenne pour toutes les espèces et l'erreur relative n'atteint pas 1%.

Nous pouvons nous référer à [44, 53], par exemple, pour plus de détails.

	AMF	Diffusion-Chimie
SO_2	max = 139.384 mean = 1.30863	139.377 1.30695
NO_2	max = 84.2543 mean = 3.81169	84.0437 3.80232
NO	max = 78.9695 mean = 0.727317	79.0501 0.725399
O_3	max = 177.199 mean = 78.9784	177.199 78.9891

Table 2.2: Comparaison AMF versus Diffusion-Chimie avec le modèle POLAIR3D. Valeur maximale et moyenne spatio-temporelles.

2.3.3 Intégration de l'advection L_{advec}

Dans un modèle CTM, le module d'advection par le vent est parmi les processus les plus importants. Il est donc crucial d'utiliser un schéma d'advection performant, qui permette notamment de minimiser la diffusion numérique.

La résolution numérique de l'équation d'advection peut se faire par plusieurs méthodes. On peut citer par exemple :

- les méthodes de "Différences finies", la méthode la plus utilisée dans les CTM ;
- les méthodes spectrales, utilisées dans les modèles météorologiques mais peu dans les modèles de chimie-transport ;
- la méthode des volumes finis ;
- les approches semi-lagrangiennes [44].

Dans notre modèle nous avons choisi d'utiliser le schéma aux différences finies 'DST' (Direct Space Time) [44, 49], un schéma explicite de premier ordre en temps et de troisième ordre en espace. Il fait intervenir les trois points de la discretisation

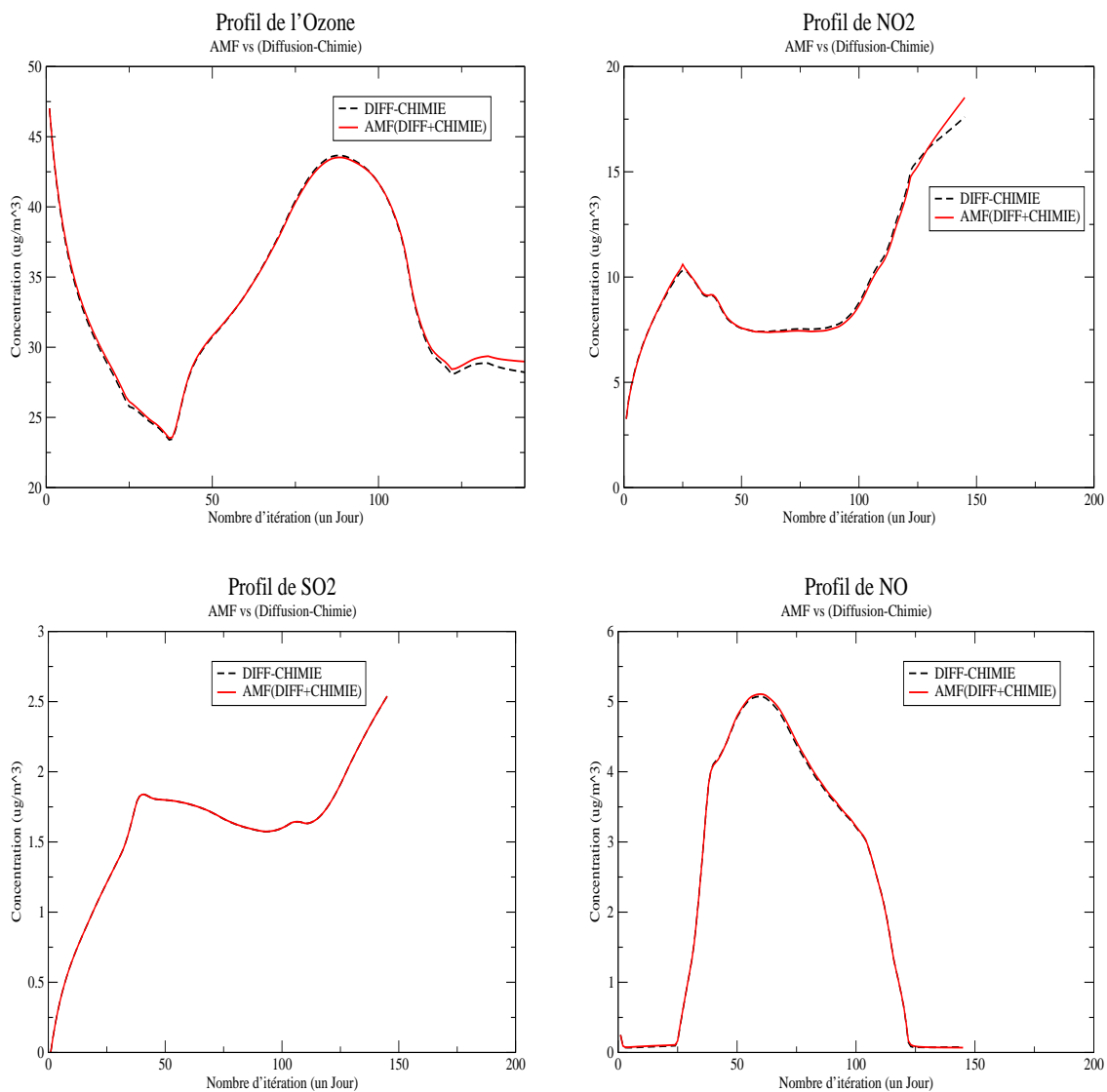


Figure 2.13: AMF versus Diffusion-Chimie (séparés) en un point situé sur la France, simulation effectuée avec POLAIR3D.

avec une fonction de limiteur de flux, ce qui permet de le positionner comme un schéma très peu diffusif.

2.3.3.1 Formulation

Pour un modèle unidimensionnel selon la direction x :

$$\frac{dc}{dt} = L_{adv_x} \quad (2.61)$$

nous utilisons le schéma de 3ème ordre Direct Space Time (DST) avec une fonction limiteur de flux de type Koren-Sweby [54] :

$$c_i^{n+1} = c_i^n + (F_{i-\frac{1}{2}} - F_{i+\frac{1}{2}}) \quad (2.62)$$

où :

$$- u_{i+\frac{1}{2}} \geq 0 :$$

$$F_{i+\frac{1}{2}} = \nu_{i+\frac{1}{2}} (c_i + \psi(\nu_{i+\frac{1}{2}}, \theta_i)(c_{i+1} - c_i)) \quad (2.63)$$

$$- \text{pour } u_{i+\frac{1}{2}} < 0 :$$

$$F_{i+\frac{1}{2}} = -\nu_{i+\frac{1}{2}} (c_{i+1} + \psi(\nu_{i+\frac{1}{2}}, \frac{1}{\theta_{i+1}})(c_i - c_{i+1})) \quad (2.64)$$

2.3.3.2 Fonction de limiteur de flux

L'équation d'advection (2.62) peut produire des valeurs de concentration d'espèces chimiques négatives. Quand l'advection est couplée avec de la chimie, ceci peut mener à une instabilité de modèle. Une manière simple de résoudre ce problème est de remettre les concentrations négatives à zéro (procédure classique de "clipping"), mais ceci peut influencer la conservation de la masse. Une autre approche plus robuste est d'introduire une fonction de limiteur de flux. Par exemple, le limiteur de Sweby [54] est donné par :

$$\psi(\nu, \theta) = \max(0, \min(1, d_0(\nu) + d_1(\nu)\theta, \mu\theta)) \quad (2.65)$$

où

$$d_0(\nu) = \frac{1}{6}(2 - \nu)(1 - \nu) , \quad d_1(\nu) = \frac{1}{6}(1 - \nu^2) , \quad (2.66)$$

$$\nu_{i+\frac{1}{2}} = \frac{\Delta t}{\Delta x} \left| u_{i+\frac{1}{2}} \right| , \quad \theta_i = \frac{c_i - c_{i-1}}{c_{i+1} - c_i} \quad (2.67)$$

et le paramètre μ dans (2.65) et pris comme dans [44] selon :

$$\mu = \frac{1 - \nu}{\nu} \quad (2.68)$$

Pour ce schéma la condition de Courant-Friedrichs-Lewy (CFL) est :

$$\nu_{i+\frac{1}{2}} \leq 1 \quad (2.69)$$

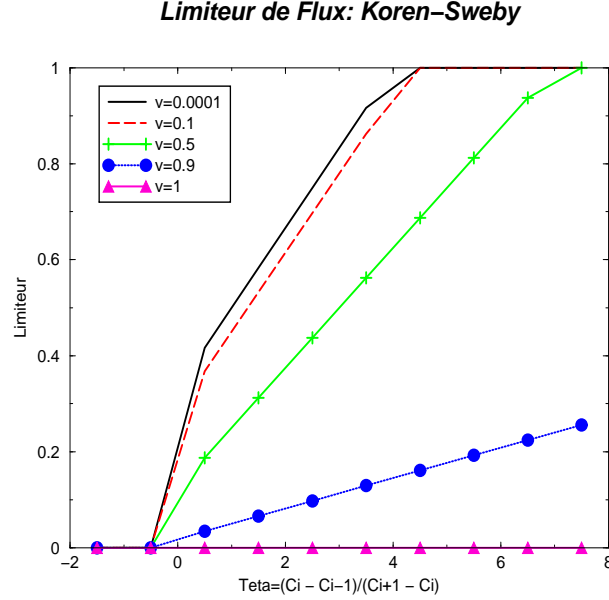


Figure 2.14: Fonction de limiteur de flux pour les différentes valeurs de ν

Dans le cas du modèle POLAIR3D, l'advection schématisée par la figure 2.15, est discrétisée selon :

$$c_{ijk}^{n+1} = c_{ijk}^n + ((F_{i-\frac{1}{2}}^{jk} - F_{i+\frac{1}{2}}^{jk}) + (G_{j-\frac{1}{2}}^{ik} - G_{j+\frac{1}{2}}^{ik}) + (H_{k-\frac{1}{2}}^{ij} - H_{k+\frac{1}{2}}^{ij})) \quad (2.70)$$

où c_{ijk}^{n+1} est la concentration à l'instant ($t_{n+1} = t_n + \Delta t$) au point de discrétisation (i,j,k), $F_{i-\frac{1}{2}}^{jk}, F_{i+\frac{1}{2}}^{jk}$ (resp. $G_{j-\frac{1}{2}}^{ik}, G_{j+\frac{1}{2}}^{ik}$ et $H_{k-\frac{1}{2}}^{ij}, H_{k+\frac{1}{2}}^{ij}$) sont les flux selon la direction x (resp. y et z) au point i (resp. j et k).

2.3.3.3 Limitation du pas de temps d'advection :

La diffusion numérique du schéma d'advection est dépendante du nombre de Courant défini par le pas de temps de l'advection (Δt_{advec}). La valeur maximale du pas de temps de l'advection est calculée pour une grille de résolution Δx et pour une vitesse du vent U^* , selon :

$$\Delta t_{max} \leq \frac{\Delta x}{\min(U_i^*)} \quad (2.71)$$

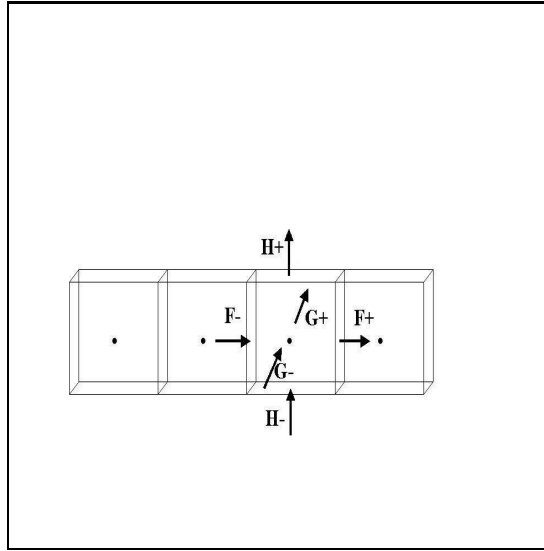


Figure 2.15: Flux d'advection

Cette expression est évaluée pour chaque point de grille i selon l'horizontale et la verticale. Dans la figure (2.16) nous présentons, à titre d'exemple, l'évolution du pas de temps maximal calculé par (2.3.3.3) pour l'advection selon x , y et z pour une application européenne avec une résolution spatiale de $\Delta x = \Delta y = 1.125^\circ$ et une grille verticale de 14 niveaux (figure 2.9).

Comme nous pouvons le constater, sous de telles conditions, le Δt maximal exigé par la CFL ne peut pas dépasser les 1000s.

2.3.3.4 Test de Molenkamp-Crowley

Pour illustrer la faible diffusion numérique associée à la méthode DST, nous testons le schéma sur un exemple classique. Nous résolvons le problème en deux dimensions de Molenkamp-Crowley ([44]) :

$$\frac{\partial c}{\partial t} + \text{div}(U.c) = 0 \quad t > 0 \quad \Omega = [0, 1]^2 \quad (2.72)$$

Nous considérons un cylindre à l'état initial de rayon 0.15 centré au point $(\frac{1}{2}, \frac{3}{4})$, de hauteur 1. Ce test est habituellement présenté comme très sélectif du fait des fortes discontinuités à simuler.

Le vent U est caractérisé par :

$$u(x, y) = 2\pi(y - \frac{1}{2}) \quad , \quad v(x, y) = -2\pi(x - \frac{1}{2}) \quad (2.73)$$

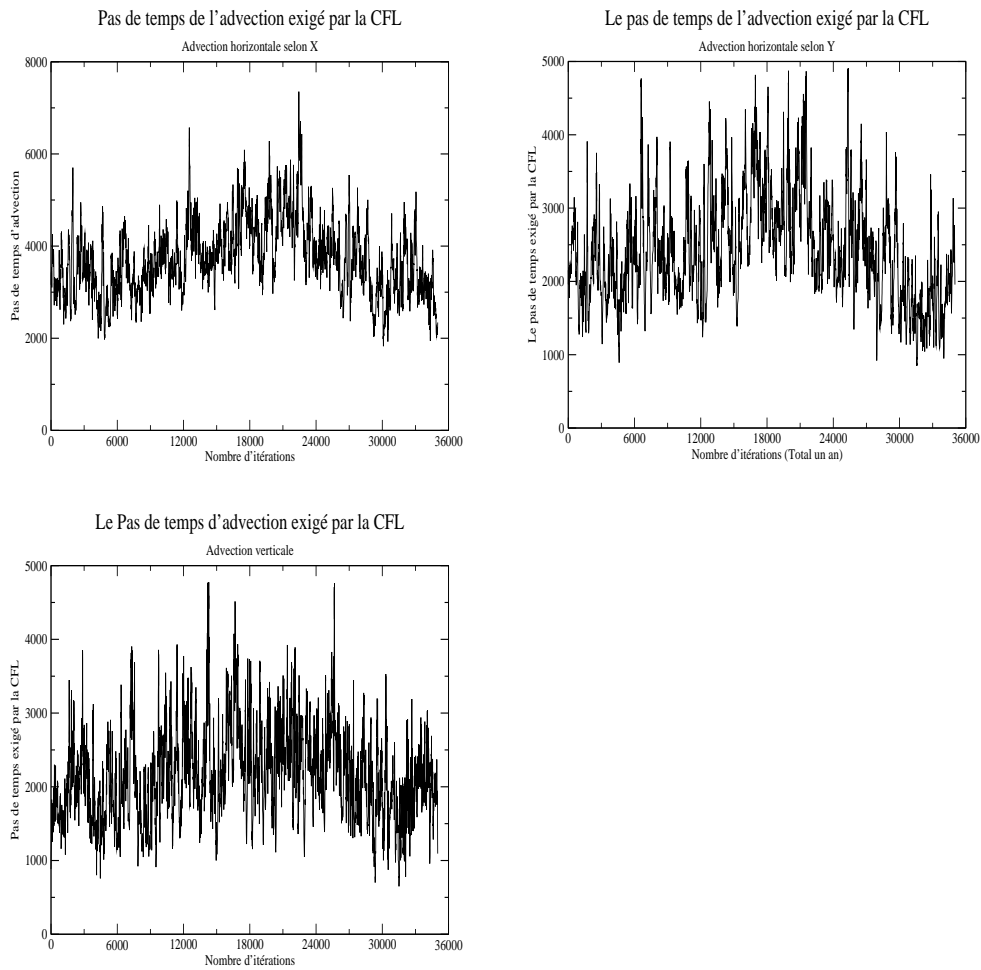


Figure 2.16: Pas de temps exigé par la CFL de l'advection horizontale et verticale

Le cylindre est défini par sa trace au sol $(x(t), y(t))$ selon :

$$x(t) = a \sin(2\pi t) + b \cos(2\pi t) + \frac{1}{2}, \quad y(t) = -b \sin(2\pi t) + a \cos(2\pi t) + \frac{1}{2} \quad (2.74)$$

avec $(a, b) = (\frac{1}{2}, \frac{1}{2})$. Nous faisons pivoter le cylindre autour du point $(\frac{1}{2}, \frac{1}{2})$, le champ de vitesse définissant la rotation avec une période $T = 1$. Comme le montre la figure 2.17, le profil du cylindre est diffusé mais les résultats sont convenables (on peut citer par exemple le benchmark réalisé dans [44] qui montre l'intérêt d'utiliser ce schéma).

2.3.4 Discrétisation du terme de la diffusion

Dans l'équation de diffusion :

$$\frac{\partial c}{\partial t} = \nabla(K \cdot \nabla c) = L_{diff} \quad (2.75)$$

le terme de la diffusion est discrétisé selon le schéma classique à trois points. Par exemple pour la direction x au point x_i :

$$div(K_x \nabla c) \sim \frac{K_x(x_{i+\frac{1}{2}}) \frac{c(x_{i+1}) - c(x_i)}{x_{i+1} - x_i} - K_x(x_{i-\frac{1}{2}}) \frac{c(x_i) - c(x_{i-1})}{x_i - x_{i-1}}}{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}} \quad (2.76)$$

La dépendance aux autres directions est omise dans l'équation ci-dessus. Les conditions aux limites sont prises en compte selon le schéma classique :

Pour $i = 1$:

$$K_x(x_{i-\frac{1}{2}}) \frac{c(x_i) - c(x_{i-1})}{x_i - x_{i-1}} = 0 \quad (2.77)$$

Pour $i = N_x$:

$$K_x(x_{i+\frac{1}{2}}) \frac{c(x_{i+1}) - c(x_i)}{x_{i+1} - x_i} = 0 \quad (2.78)$$

où N_x est le nombre de centres de maille selon la direction x . Cette discrétisation est appliquée de la même manière à la direction y .

Pour la diffusion verticale nous rappelons les conditions aux limites :

Au sol ($i = 1$) :

$$-K_z(z_{i-\frac{1}{2}}) \frac{c(z_i) - c(z_{i-1})}{z_i - z_{i-1}} = E - v^{dep} c(z_i) \quad (2.79)$$

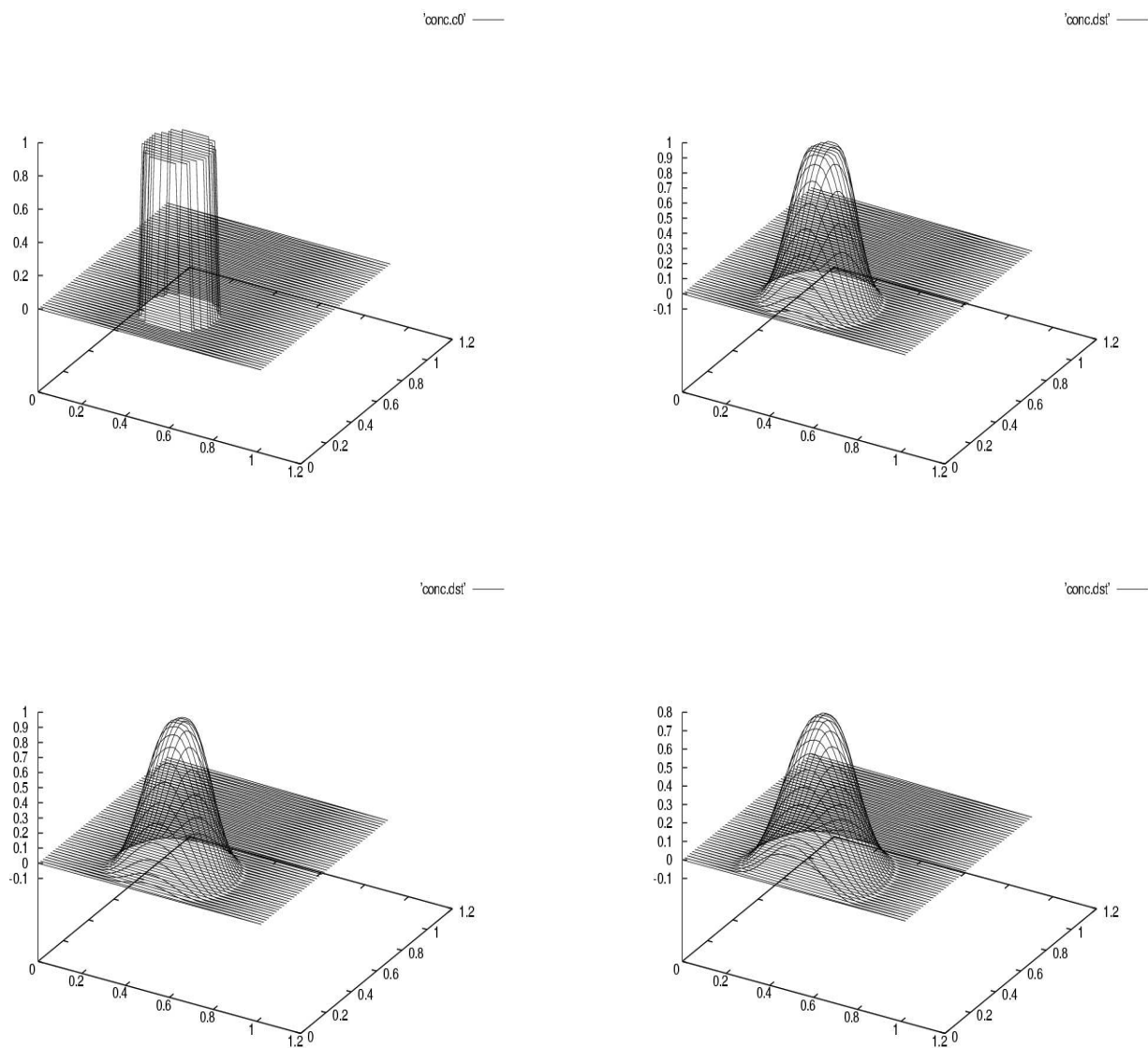


Figure 2.17: Test de Molenkamp-Crowley : en haut à gauche (instant zéro), en haut à droite (après un tour), en bas à gauche (après 2 tours) et en bas à droite (après 5 tours).

Au sommet du domaine ($i = n_z$) :

$$-K_z(z_{i+\frac{1}{2}}) \frac{c(z_{i+1}) - c(z_i)}{z_{i+1} - z_i} = 0 \quad (2.80)$$

où N_z est le nombre de centres de maille selon la direction z .

2.3.5 Intégration de la chimie (L_{chem}) et la diffusion (L_{diffz})

2.3.5.1 Raideur numérique de la chimie

Nous considérons à présent le système chimique décrit par l'équation différentielle en toute généralité :

$$\frac{dc}{dt} = f(c, t) \quad (2.81)$$

où f est la fonction L_{chem} et c le vecteur de concentrations.

La cinétique chimique fait intervenir un grand nombre de réactions et d'espèces chimiques dont les échelles de temps ont des ordres de grandeur très différents les uns des autres. C'est cette dispersion des échelles en temps qui définit la notion numérique de **raideur** du système. Dans ce cas, le pas de temps d'intégration de l'équation (2.81) est déterminé par le plus petit temps caractéristique du système (de l'ordre de la milliseconde, pour les radicaux de la chimie atmosphérique).

Le module de chimie, à travers sa raideur et sa non-linéarité, est le module le plus coûteux en temps de calcul, pouvant représenter jusqu'à 80% du coût total d'un modèle de chimie-transport (lorsqu'on ne tient pas compte des aérosols).

L'usage de grands pas de temps lors de l'intégration de l'équation (2.81) ne peut se faire qu'en utilisant des solveurs avec de bonnes propriétés de stabilité [2]. Parmi les solveurs existants nous citons des schémas "explicites" adaptés au cas de la cinétique chimique, par exemple :

- le solveur QSSA (Quasi Steady State Approximation) [55] ;
- le solveur EBI (Euler Backward Iterative) [56] ;
- le solveur TWOSTEP variante du solveur BDF (Backward Difference Formula [42, 57]).

Une autre catégorie de solveurs qui utilise des pas de temps élevés est fournie par les méthodes implicites :

- le solveur LSODE (Livermore Solver for Ordinary Differential Equations) [58] ;
- le solveur VODE (Variable coefficient ODES solver) [58] ;
- le solveur semi-implicite Rosenbrock de type Runge-Kutta implicite. C'est cette dernière méthode que nous avons choisi d'utiliser dans le modèle POLAIR3D.

2.3.5.2 Solveur numérique de second ordre Rosenbrock (ROS2)

Pour l'intégration du système :

$$\frac{dc}{dt} = f(t, c) \quad (2.82)$$

la méthode Rosenbrock autonome de second ordre est définie par les étapes suivantes ([59]) de l'instant t_n à t_{n+1} :

$$c^{n+1} = c^n + \sum_{i=1}^2 b_i k_i \quad (2.83)$$

où

$$\begin{cases} k_i = \Delta t f(t_n + \alpha_i \Delta t, c^n + \sum_{j=1}^{i-1} \alpha_{ij} k_j) + \Delta t J \sum_{j=1}^i \gamma_{ij} k_j \\ \alpha_i = \sum_{j=1}^{i-1} \alpha_{ij} \quad , \quad \gamma_i = \sum_{j=1}^i \gamma_{ij} \end{cases} \quad (2.84)$$

J dénote la matrice Jacobienne $J = \partial f / \partial c$ et les coefficients b_i, α_{ij} et γ_{ij} sont choisis afin d'obtenir l'ordre désiré de consistance et de stabilité (convergence).

La fonction de stabilité [49] de cette méthode calculée à partir de $(\frac{dc}{dt} = \lambda c)$ est :

$$R(z) = \frac{1 + (1 - 2\gamma)z + (\frac{1}{2} - 2\gamma + \gamma^2)z^2}{(1 - \gamma z)^2} \quad (2.85)$$

où $z = \lambda \Delta t$, La fonction de stabilité est A-stable ($|R(z)| < 1$) pour z assez grand) si $\gamma \geq 1/4$.

Afin d'assurer la L-stabilité ($R(z)=0$ pour z assez grand), la valeur de γ doit être prise égale à $1 + 1/\sqrt{2}$ ou $1 - 1/\sqrt{2}$.

Le schéma autonome de ROS2 est défini par les coefficients :

$$\begin{cases} \gamma = 1 + \frac{1}{\sqrt{2}} \quad , \quad \gamma_1 = \gamma_{11} = \gamma \quad , \quad \gamma_{21} = -2\gamma \quad , \quad \gamma_{22} = \gamma \\ \gamma_2 = \gamma_{21} + \gamma_{22} = -\gamma \quad , \quad \alpha_1 = 0 \quad , \quad \alpha_2 = \alpha_{21} = 1 \quad , \quad b_1 = b_2 = \frac{1}{2} \end{cases} \quad (2.86)$$

qui définissent les deux pas d'intégration :

$$\begin{cases} (1 - \gamma \Delta t A)k_1 = f(t_n, c_n) \\ (1 - \gamma \Delta t A)k_2 = f(t_{n+1}, c_n + \Delta t k_1) - 2k_1 \end{cases} \quad (2.87)$$

où $\gamma = 1 + 1/\sqrt{2}$ est priconisé (garantit la positivité pour le cas linéaire) et A désigne une approximation de la matrice Jacobienne $A \simeq \partial f / \partial c$.

Pour des raisons de non autonomie on peut rajouter, dans certains cas, aux paramètres k_1 et k_2 , le terme $\gamma_i \Delta t^2 \partial f / \partial t$ [3] :

$$\begin{cases} (1 - \gamma \Delta t A) k_1 = f(t_n, c_n) + \gamma \Delta t \frac{\delta f(t_n, c_n)}{\delta t} \\ (1 - \gamma \Delta t A) k_2 = f(t_{n+1}, c_n + \Delta t k_1) - 2k_1 - \gamma \Delta t \frac{\delta f(t_n, c_n)}{\delta t} \end{cases} \quad (2.88)$$

Ce terme est nul en l'absence de la photolyse ou de contenu en eau variable pour la chimie en phase aqueuse [3].

Pour plus de détails nous nous référons à [59, 49] pour l'utilisation des méthodes de Rosenbrock dans le contexte général des modèles de chimie-transport.

Nous comparons dans la figure 2.18 les résultats obtenus en utilisant le solveur ROS2 et le solveur classique LSODE, calcul fait avec la version 0D de POLAIR3D en utilisant la chimie CBM-IV (Annexe A), le pas de temps est de 900 secondes, la simulation est faite sur une journée. Les deux solveurs donnent des résultats très proche, l'erreur relative entre la calcul avec ROS2 et LSODE avec une valeur de $r_{tol} = 10^{-10}$ étant négligeable.

2.3.5.3 Clipping

Pour des systèmes chimiques réels, la positivité n'est pas garantie. La positivité du schéma ROS2 est forcée par clipping (les valeurs de c_{n+1} négatives sont mises à zéro). Dans un exemple de simulation continentale ($N_x=47$, $N_y=32$, $N_z=9$) avec le modèle POLAIR3D et la chimie RACM en utilisant un pas de temps de 600 secondes sur une journée de calcul, nous avons observé en moyenne spatiale et temporelle sur toutes les espèces chimiques un pourcentage de clipping ne dépassant pas les 0.3%.

Le clipping interfère en toute rigueur avec la propriété de conservation de la masse, mais, en pratique, on n'observe pas de perte notable sur la précision des résultats.

2.3.5.4 Caractère "creux" des matrices jacobienne de chimie

Avec le schéma de Rosenbrock de second ordre (ROS2), il est demandé, à chaque pas de temps, d'évaluer la matrice jacobienne $A = f'(c_n)$ et de résoudre deux systèmes linéaires. La résolution du système linéaire demande une factorisation de matrice par une décomposition LU et deux stockages en mémoire ce qui exige un

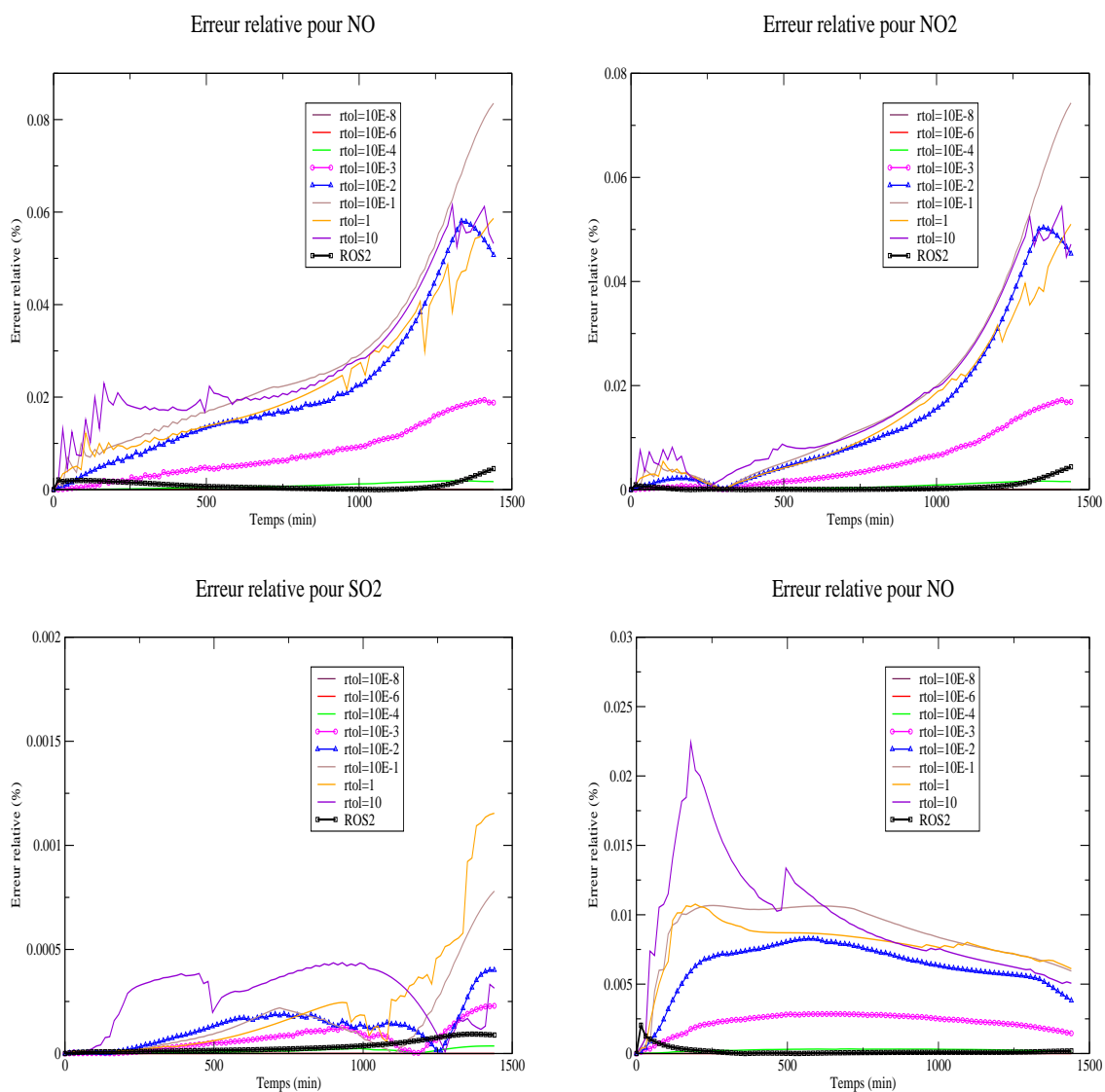


Figure 2.18: Comparaison ROS2/LSODE : l'erreur relative entre les deux méthodes en % (le calcul LSODE avec $r_{tol} = 10^{-10}$ est pris comme référence), plusieurs valeurs de r_{tol} sont comparées.

temps CPU important. Pour les modèle de chimie, le Jacobien est une matrice creuse (contenant un grand nombre de zéros). Ceci peut être exploité pour réduire le coût de ces calculs d'algèbre linéaire. La méthode consiste en un rangement adéquat des lignes et colonnes composantes de la matrice jacobienne A [60]. Pour des systèmes complexes de chimie, le gain peut être important.

Le tableau 2.3 montre le pourcentage des éléments de la matrice jacobienne non nuls, pour la chimie CBMIV (Annexe A) et RACM (Annexe C), qui font intervenir respectivement 32 espèces et 76 espèces chimiques.

Le préprocesseur SPACK fournit la routine calculant la décomposition LU pour la matrice non creuse à partir des réactions du système chimique. La comparaison du coût calcul des deux approches est faite à base d'un calcul chimie-transport 3D de POLAIR3D en utilisant la chimie RACM (Annexe C), avec un pas de temps de 100s et pendant 100 itérations. Nous obtenons un speedup de 2.9 avec l'utilisation astucieuse de la matrice creuse (tableau 2.4).

	Dimension	Nombre de non zéros	Pourcentage des zéros
CBM-IV	32^2	230	77.5%
RACM	76^2	902	84.4%

Table 2.3: Taux de remplissage de la matrice jacobienne de chimie

	FULL	SPARSE
RACM	766 s	263s

Table 2.4: Comparaison du temps CPU par les approches Full et Sparse

Chapitre 3

Evaluation du modèle Chimie-Transport POLAIR3D par comparaison aux données du réseau EMEP pour l'été 2001

Résumé

Le modèle de Chimie-Transport POLAIR3D a été développé dans l'optique de fournir une plate-forme numérique modulaire, multi-fonctions et multi-chimies pour des études de dispersion atmosphérique. La validation d'un tel modèle est bien sûr primordiale afin de déterminer sa fiabilité, ses performances et aussi ses limitations et mieux comprendre l'importance relative des processus physiques pris en compte dans le modèle. Plusieurs études ont déjà été effectuées pour évaluer le modèle dans les cas passif ([61]) et réactif lors d'une validation à l'échelle régionale ([62]).

Dans ce chapitre, nous présentons les performances du modèle POLAIR3D pour une simulation à l'échelle européenne de la période de Juillet-Août-Septembre de l'année 2001 avec la chimie RACM. La performance du modèle est évaluée par comparaison des résultats obtenus pour l'ozone aux mesures du réseau EMEP, sur le continent européen.

La qualité des résultats obtenus justifie l'utilisation de POLAIR3D pour le reste de nos travaux.

3.1 La géométrie du modèle

Le modèle POLAIR3D est utilisé pour fournir des simulations des concentrations de l'ozone pendant la période du 23 Juin au 30 septembre 2001 sur un domaine géographique qui s'étend sur toute l'Europe, de -11.8125° W au 39.9375° E ; et de 36.5625° au $71^{\circ}4375^{\circ}$ N, avec une résolution horizontale de $1,125^{\circ} \times 1,125^{\circ}$, et un maillage de 47 points selon l'axe méridional et 33 points selon l'axe zonal. La simulation est effectuée avec un maillage vertical de 9 niveaux verticaux, avec une grille irrégulière, le premier niveau du modèle est situé à 15.5 m, suivi des niveaux : 85.5, 237, 464.5, 897.5, 1437.5, 1894.5, 2338.5 et le dernier est situé à 2813m. Le domaine géographique sur lequel notre étude est effectuée, est schématisé par la figure 3.1.

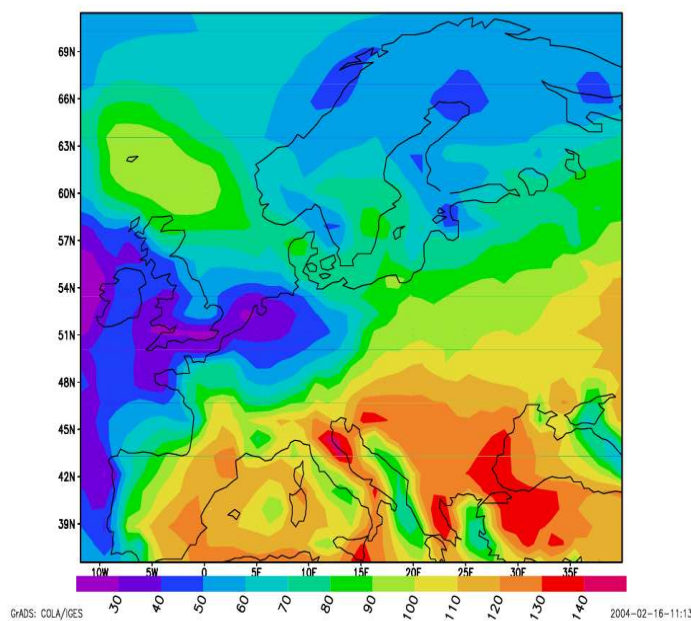


Figure 3.1: Domaine de simulation (Ozone simulé par POLAIR3D le 24 Juin 2001 à 14h GMT sur l'Europe)

3.2 Les données météorologiques

Pour les processus de transport et de chimie, le modèle POLAIR3D a besoin tout au long de l'intégration et sur tout le domaine géographique des données météorologiques. Elles sont fournies (dans notre cas) par les modèles de prévision numérique du ECMWF (European Center of Medium-Range Weather Forcecasts) sur une coordonnée verticale de type hybride de 61 niveaux, et une grille horizontale en latitude/longitude de 0.5 degrés de résolution horizontale. Elles sont disponibles avec une fréquence de 3 heures. Notons de plus que :

- La diffusion horizontale est prise constante ($50000 \text{ m}^2\text{s}^{-1}$).
- La diffusion verticale est calculée selon le schéma de Louis ([31]).
- Le vent vertical est diagnostiqué pour conserver la continuité du mouvement ($\text{div } \mathbf{U}=0$).
- Le vent horizontal est transformé selon :

$$U = u * \cos(\Phi) \quad V = v / \sin(\Phi) \quad (3.1)$$

- Les données 3D de température, d'humidité et de pression et les données 2D de précipitation sont de plus utilisées.

Un prétraitement est nécessaire afin de convertir les données en coordonnées Z (en hauteurs) et sur la grille horizontale du modèle POLAIR3D. Les données du vent et de la diffusion sont données sur les facettes du maillage contrairement aux données de température, d'humidité, de pression et de pluies qui sont données sur les centres de mailles.

Une interpolation temporelle est effectuée dans le modèle au moment de la lecture des données afin de calculer les données à chaque pas de temps de l'advection, de la diffusion et de la chimie (dans notre application le pas de temps est de 600 secondes).

3.3 La chimie

La chimie utilisée dans cette validation est le mécanisme RACM (Regional Atmospheric Chemistry Modeling) de Stockwell et al. ([51]). Le schéma met en jeu 237 réactions chimiques et utilise 72 espèces chimiques (annexe [?]).

Pour chaque pas de temps, POLAIR3D calcule la concentration de chacune des espèces chimique après les différents processus physiques de transport et de diffusion. Le pas de temps du modèle est pris à 600 secondes.

3.4 Les conditions initiales et aux limites

Afin d'initialiser les espèces chimiques du modèle nous avons besoin des conditions initiales qui peuvent être prises à partir des réseaux de mesures par le biais d'interpolation, ou des résultats du modèle ultérieur ou encore à partir d'autres modèles qui couvrent le domaine d'étude.

Les conditions aux limites (latérales, au sommet et au sol) sont nécessaires au modèle à cause du schéma numérique qui exigent des données aux points extérieur au domaine de l'étude. Elles sont disponibles soit :

- à partir du même modèle sur un domaine plus large couvrant le domaine d'étude.
- à partir d'un autre modèle de CTM 3D qui couvre notre domaine de calcul (typiquement des modèles globaux par exemple).
- ou à partir de simple fonctions qui prescrivent les concentrations des espèces en terme de latitude et en fonction du temps dans l'année ou du jour.

Dans notre étude, les conditions aux limites sont disponibles pour 25 espèces chimiques, les conditions initiales pour 26 espèces chimiques. Elles sont fournies par les simulation du modèle Global Mozart (Model for OZone And Related chemical Tracers, "<http://acd.ucar.edu/models/MOZART/m2/>").

3.5 Les émissions

3.5.1 Les émissions dans le modèle

Le modèle POLAIR3D, avec la chimie RACM, a besoin des émissions de 20 espèces chimiques. Elles sont traitées comme des conditions aux limites basses de la diffusion verticale. Ces émissions sont fournies à partir de source anthropogéniques et biogéniques classées en classes dites des 'SNAP'. Dans la base de données des émissions de MSC-W de EMEP, il existe 11 SNAP dont une pour les émissions biogénique et le reste pour les anthropogéniques. Dans le tableau 3.1 sont représentés les différents secteurs économiques responsables des émissions des différents espèces : NO_x , SO_x , CO, CH_4 et les Composés Organiques Volatiles, et leurs niveaux verticaux d'émissions (www.emep.int). Les données d'émissions proviennent de la base de données EMEP (webdab.emep.int), et sont délivrées sous formes de totaux annuels pays par pays.

Afin d'utiliser ces données brutes dans le modèle CTM POLAIR3D, une procédure de modulation temporelle est utilisée pour disposer des données mensuelles, quotidiennes et horaires. Un travail de spatialisation est nécessaire afin de projeter

N°	Secteur SNAP	Hauteur d'émission
1	Industrie de transformation d'énergie	324-781
2	Usine non industrielle	0-184
3	Industrie manufacturières	324-781
4	Procédés de production	0-100
5	Extraction, distribution de fuels fossiles et énergie géothermique	0-200
6	Utilisation des solvants	0-100
7	Transport routier	0-100
8	autre Transport	0-100
9	Traitement des déchets et Mise en décharge	184-522
10	agriculture	0-100
11	biogénique	0-100

Table 3.1: Secteurs d'émissions.

ces données sur la grille du modèle. Puis une étape d'agrégation est utilisée à partir des 32 classes qui composent les COV pour produire les espèce modèles selon la méthode de Middleton [63].

3.5.2 Inventaire d'émissions et distribution spatiale

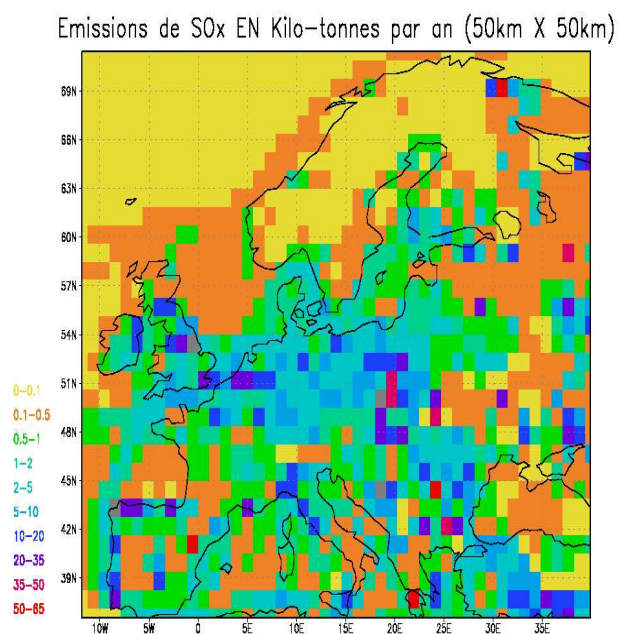
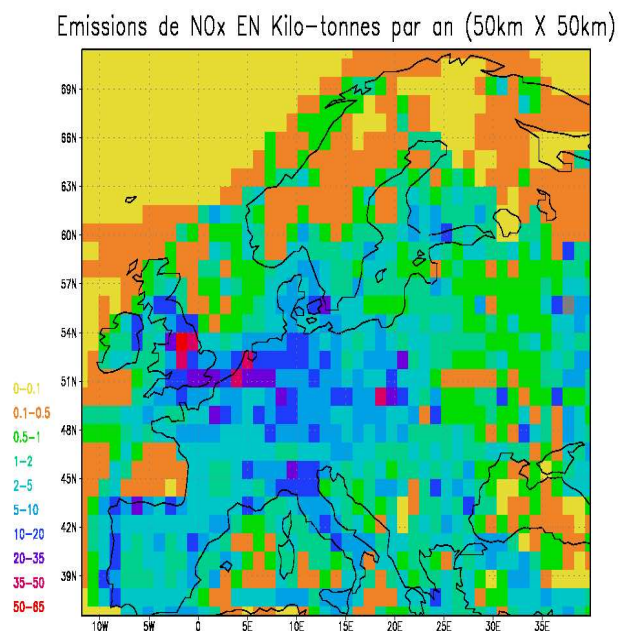
Les données d'émissions sont disponibles sous forme de totaux annuels en tonne émis par chaque pays en plusieurs stations, les données étant représentées sur une grille de 50km x 50km. Nous présentons un exemple d'inventaire de NO_x , SO_x et de CO et $NMVOC$ au titre de l'année 2001 dans le Tableau 3.2, et dans la figure 3.2 une spatialisation des émissions proposée par (www.emep.int) que nous avons projeté sur le domaine du modèle.

3.5.3 Distribution temporelle

Les émissions sont fournies comme données d'entrée au mécanisme RACM du modèle POLAIR3D pour chaque pas de temps du modèle et pour les 20 espèces (NO_2 , NO , SO_2 , CO , $HONO$, HC_3 , CH_4 ,...). Cependant les émissions dont nous disposons sont des données annuelles pour NO_x , SO_2 , $NMVOC$ et CO . Après une agrégation en espèces modèles, nous utilisons une spéciation de NO_x en NO et

Pays	SO_x	NO_x	$NM VOC$	CO	Pays	SO_x	NO_x	$NM VOC$	CO
Ireland	131.	125.	90.2	279.5	Luxembourg	3.	17.	15.	49.
Kola Karelia	296.3	55.2	33.7	199.6	Netherlands	89.	410.2	270.9	659.
France	609.8	1411.	1673.6	6364.6	Bosnia	419.	55.	42.	193.
North Sea	454.	648.	15.	59.	Black Sea	57.	86.	2.	8.
Italy	758.	1372.	1464.	5207.2	Norway	24.7	220.7	375.8	548.2
Belarus	150.7	134.8	215.4	710.7	Serbia	394.	158.	129.	553.
Armenia	4.4	13.3	28.3	104.2	Poland	1564.	805.	577.	3528.
Portugal	286.	377.	468.	1004.	Turkey	2112.	951.1	725.6	3778.2
Romania	912.	319.	638.	2325.	Kazakhstan	237.	50.2	50.	278.5
Slovenia	96.	58.	40.	68.	Estonia	91.7	37.7	33.2	177.4
Spain	1394.	1302.6	1439.8	2740.7	United Kingdom	1125.3	1680.3	1336.	3737.1
Croatia	58.	76.8	79.8	402.	Georgia	9.	30.1	19.	221.9
Sweden	56.7	248.	303.3	808.4	Kaliningrad	12.4	18.	20.3	147.5
Azerbaijan	14.7	43.2	8.5	293.1	Latvia	13.3	42.	80.8	381.5
Albania	58.	29.	34.	101.1	Cyprus	48.3	18.	14.4	84.5
Ukraine	1029.	561.	269.5	2671.8	Lithuania	48.7	55.	70.6	228.5
Austria	36.7	199.4	232.2	859.7	Czech Republic	251.	332.	220.	649.
Switzerland	21.	91.1	147.	409.4	Slovakia	128.5	105.7	89.7	286.8
Belgium	161.8	316.6	251.5	1026.8	Denmark	25.3	204.	123.9	587.3
Moldova	12.	16.9	22.14	100.2	Iceland	27.	27.7	10.	39.8
Germany	649.6	1621.6	1605.6	4796.2	Finland	85.2	221.8	157.1	605.
Bulgaria	845.9	163.6	122.	521.3	Baltic Sea	228.	352.	8.	29.
Greece	485.	331.	267.	1366.					
Hungary	400.4	184.5	165.7	591.8					

Table 3.2: Inventaire d'émission de NO_x , SO_2 et $NM VOC$ et CO en tonne par an (2001).

Figure 3.2: Distribution spatiale de NO_x et SO_x (2001)

NO_2 respectivement en 95% et 5% en équivalent masse molaire.

La distribution saisonnière des émissions de SO_x et NO_x est effectuée selon une matrice qui fournit pour chaque pays la proportion d'émissions mensuelles (tableau 3.3). Pour chaque mois correspond un coefficient de proportion à partir duquel sont interpolés les coefficients journaliers. La figure 3.3 présente la distribution, selon les mois, de SO_x pour la France, l'Allemagne, l'Espagne, le Royaume Uni et la Pologne.

La distribution horaire des émissions est ensuite effectuée en fonction de l'espèce émise et en fonction du jour dans la semaine. Les figures 3.5 et 3.4 donnent, respectivement, un aperçu sur le profil de la distribution journalière de NO_x et SO_x utilisé dans le modèle.

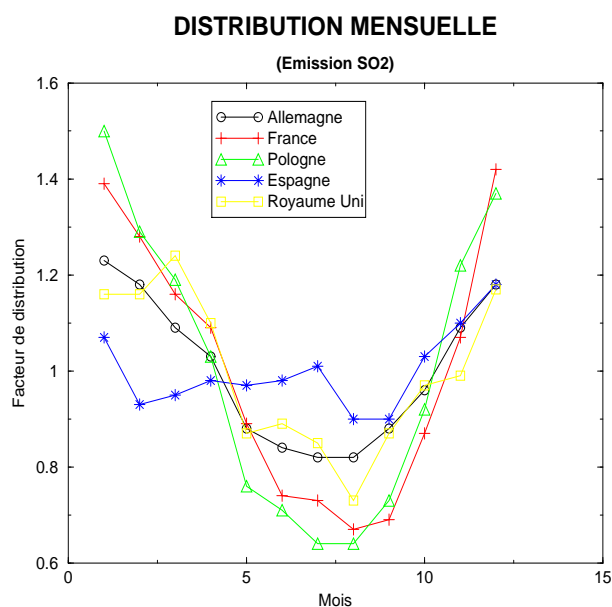


Figure 3.3: Distribution mensuelle de SO_x

3.6 Méthodologie d'analyse

3.6.1 Les mesures du réseau EMEP

Pour analyser les performances du modèle POLAIR3D, nous avons utilisé les données de mesures du réseau d'EMEP comme référence. Pour plus de détails sur

Pays	1	2	3	4	5	6	7	8	9	10	11	12
Albania	2.06	1.83	1.57	1.26	1.08	0.42	0.35	0.35	0.38	0.38	0.93	1.39
Austria	1.38	1.26	1.16	1.11	0.94	0.78	0.66	0.61	0.76	0.90	1.12	1.32
Belgium	1.17	1.14	1.08	1.04	0.93	0.91	0.81	0.74	0.88	1.01	1.11	1.18
Bulgaria	1.25	1.30	1.21	1.08	1.00	0.85	0.78	0.78	0.80	0.83	1.01	1.11
Switzerland	1.34	1.22	1.15	1.10	0.94	0.82	0.71	0.60	0.74	0.90	1.13	1.35
Rep.Tchec	1.42	1.25	1.13	1.02	0.85	0.79	0.72	0.72	0.78	0.94	1.16	1.22
Germany	1.23	1.18	1.09	1.03	0.88	0.84	0.82	0.82	0.88	0.96	1.09	1.18
Denmark	1.43	1.24	1.08	0.97	0.81	0.70	0.60	0.70	0.86	0.95	1.20	1.45
Spain	1.07	0.93	0.95	0.98	0.97	0.98	1.01	0.90	0.90	1.03	1.10	1.18
Finland	1.30	1.25	1.20	1.09	0.93	0.81	0.66	0.64	0.76	0.95	1.14	1.27
France	1.39	1.28	1.16	1.09	0.89	0.74	0.73	0.67	0.69	0.87	1.07	1.42
Greece	1.09	1.14	1.06	0.97	0.94	0.98	0.98	0.98	0.94	0.95	1.01	0.96
Hungary	1.78	1.44	1.25	0.91	0.69	0.59	0.55	0.57	0.66	0.88	1.28	1.40
Ireland	1.19	1.19	1.09	1.10	0.98	0.77	0.80	0.73	0.77	1.00	1.14	1.24
Italy	1.09	1.15	1.11	1.04	0.96	0.91	0.91	0.84	0.89	1.01	1.04	1.05
Luxembourg	0.96	0.98	0.99	1.00	0.95	0.94	0.98	0.97	0.97	0.99	1.16	1.11
Netherlands	1.11	1.05	1.03	1.04	0.97	0.95	0.91	0.89	0.92	0.97	1.04	1.12
Norway	1.09	1.09	1.12	1.05	0.91	0.89	0.84	0.78	0.93	1.04	1.11	1.15
Poland	1.50	1.29	1.19	1.03	0.76	0.71	0.64	0.64	0.73	0.92	1.22	1.37
Portugal	1.01	0.81	0.86	0.97	1.03	1.04	1.05	1.00	1.01	1.05	1.05	1.12
Roumania	1.20	1.20	1.16	1.02	0.97	0.90	0.84	0.85	0.85	0.90	1.05	1.06
Sweden	1.29	1.20	1.19	1.09	0.91	0.83	0.63	0.61	0.78	0.99	1.18	1.30
Russian Fed	1.31	1.18	1.13	1.02	0.84	0.82	0.77	0.79	0.84	0.95	1.13	1.22
Turkey	1.12	1.07	1.04	0.98	0.95	0.92	0.87	0.93	0.94	0.99	1.09	1.10
UK	1.16	1.16	1.24	1.10	0.87	0.89	0.85	0.73	0.87	0.97	0.99	1.17
Yugoslavia	1.18	1.25	1.17	1.06	0.96	0.84	0.83	0.85	0.91	0.93	1.00	1.02

Table 3.3: Distribution mensuelle de SO_x autour de la moyenne annuelle

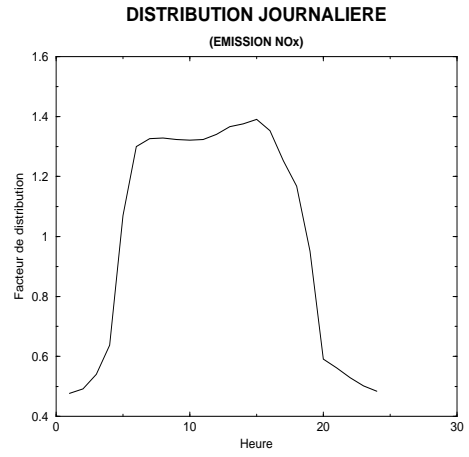
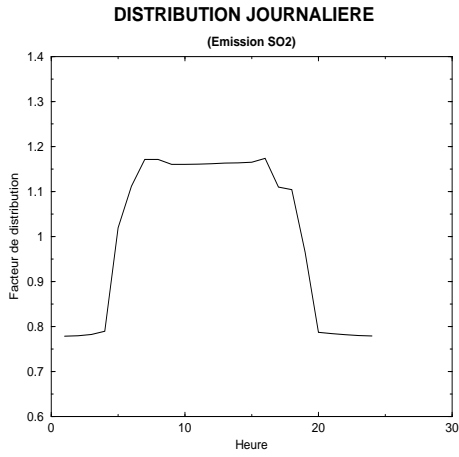


Figure 3.4: Distribution journalière : SO_x Figure 3.5: Distribution journalière : NO_x

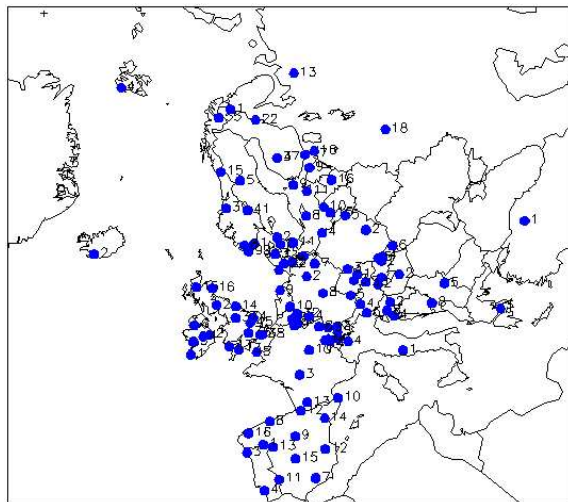


Figure 3.6: Réseau de mesure européen d'EMEP

ce réseau on peut se référer au site-web “www.emep.net”. Le réseau contient 151 stations étendues sur l'Europe (figure 3.6).

Les stations fournissent des mesures horaires O_3 , à quelques exceptions près. Ainsi, nous ne prenons en compte pour notre analyse que les stations dont la disponibilité de mesures dépasse 12 mesures par jour (50%). Le nom des stations et leurs altitudes sont présentés dans l'annexe D.

3.6.2 Mesures statistiques de l'erreur

Il existe plusieurs façons pour évaluer les performances d'un modèle par rapport à un champ de référence. Dans notre cas, le champ de référence est donné par les mesures d'observations issues du réseau EMEP.

Nous calculons les paramètres suivants pour évaluer les performances du modèle :

- **Le biais** défini par $\frac{1}{N} \sum_{i=1}^N (C_{Mod} - C_{Mes})$.

N désigne le nombre de points où la comparaison est faite, autrement dit les points où les mesures sont disponibles.

Cet indicateur nous permet de voir si le modèle sur-estime (dans le cas où le biais est positif) ou sous-estime (dans le cas où le biais est négatif) les valeurs observées. Si la valeur est nulle cela ne signifiera pas que le modèle est parfait mais juste l'absence d'une erreur répétitive entre les valeurs estimées par le modèle et les mesures.

- **L'erreur quadratique moyenne ou RMS (Root Mean Square)** de l'erreur, défini par :

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N (C_{Mes} - C_{Mod})^2} \quad (3.2)$$

Ce paramètre nous permet d'estimer l'erreur commise (en valeur absolu) par le modèle.

- **La corrélation** entre les mesures et les résultats du modèle, définie entre 0 et 1, par :

$$\begin{cases} Corr(Mes, Mod) = \frac{cov(Mes, Mod)}{\sigma_{Mes} * \sigma_{Mod}} \\ cov(Mes, Mod) = \frac{1}{N} \sum_{i=1}^N (C_{Mes} - \bar{C}_{Mes}) * (C_{Mod} - \bar{C}_{Mod}) \end{cases} \quad (3.3)$$

\bar{C}_{Mes} et \bar{C}_{Mod} sont les valeurs moyennes respectivement des données de mesure et de modèle. σ_{Mes} et σ_{Mod} sont les écarts type relatif, respectivement des mesures et des résultats issus du modèle.

- **Le nuage de points** (Scatter-plots) autour de la droite $Model - Mesure = 0$, et compris entre les deux droites $Model = Mesure * 2$ et $Model = Mesure/2$.
- **La distribution de fréquence** qui représente le pourcentage de stations de mesures dont la différence entre les valeurs mesurées et simulées par le modèle POLAIR3D est inférieure en valeur absolue à 50, 40, 30, 20, 10, 5, 2 et 1 $\mu g/m^3$.

3.6.3 Résultats de l'évaluation

Le modèle POLAIR3D est utilisé avec la chimie RACM sur l'Europe pendant la période du **24 Juin au 30 Septembre 2001**, avec un pas de temps de 600s et une sauvegarde horaire des résultats. Les concentrations d'ozone obtenues sont interpolées sur la grille des observations (figure 3.6) et nous ne nous intéressons qu'aux concentrations au premier niveau modèle à 15.5m. Pour cela nous avons calculé **la moyenne journalière** et **le maximum journalier** simulés par le modèle, qui seront comparés aux résultats des mesures.

La moyenne journalière ainsi que le maximum journalier observé de l'ozone sur toutes les stations sont assez variables d'un jour à l'autre comme peuvent le représenter les figures 3.7 et 3.8. La moyenne sur toutes les stations est de 69.43 ($\mu g/m^3$) simulée contre 67.09 ($\mu g/m^3$) observée pour la moyenne journalière et 86.70 ($\mu g/m^3$) simulée contre 92.91 ($\mu g/m^3$) observée pour la maximum journalier.

Dans les tableaux 3.4 et 3.5 sont représentées les statistiques relative à la moyenne journalière, modèle et mesure à chacune des stations de mesures. Pour le biais entre les mesures et le modèle, sa valeur est de +2.3 $\mu g/m^3$ en moyenne sur toutes les stations, ce qui signifie que le modèle a tendance à surestimer par rapport aux observations, et dans 90% des cas le biais ne dépasse pas 28 $\mu g/m^3/m^3$ (figure 3.13). La RMS moyenne sur tous les points de mesures est de 22.32($\mu g/m^3$).

Pour les résultats obtenus pour le maximum journalier, les statistiques de toutes les stations du réseau sont présentées dans les tableaux 3.6 et 3.7. Le biais moyen est de l'ordre de $-6.21(\mu\text{g}/\text{m}^3)$ ce qui veut dire que le modèle a plutôt tendance à sous-estimer ce pic journalier, et dans 90% des cas le biais ne dépasse pas $27 \mu\text{g}/\text{m}^3$ (Figure 3.13). La corrélation est sensiblement meilleure par rapport à celle obtenue pour la moyenne journalière, de l'ordre de de 0.505. La RMS moyenne sur tous les points de mesures est de $25.3\mu\text{g}/\text{m}^3$, valeur plus importante que celle associée à la moyenne journalière.

Les figures 3.9 et 3.10 présentent les séries temporelles de la moyenne journalière et du maximum journalier pour quelques stations du réseau européen EMEP. Les chutes brutales à des valeurs nulles signifient que les données de mesures ne sont pas disponibles ces jours là. En général, le modèle respecte assez bien la variabilité temporelle et spatiale des concentrations de l'ozone.

Les figures 3.11 et 3.12 schématisent respectivement pour la moyenne journalière et le maximum journalier, le nombre de stations autour de la droite $Y = X$, et $\frac{X}{2} < Y < 2X$ (où Y est le modèle et X est la mesure), on constate qu'un grand nombre de stations sont bien distribuées autour de la droite $Y = X$, ce qui montre une autre fois la qualité du modèle à représenter raisonnablement bien la distribution spatiale des concentrations de l'ozone.

86 Chapitre 3. Evaluation du modèle Chimie-Transport POLAIR3D par comparaison aux données du réseau EMEP pour l'été 2001

Stat	MOY MODELE	MOY MESURES	BIAIS	RMS	CORR	Stat°	MOY MODELE	MOY MESURES	BIAIS	RMS	CORR
1	81.10	75.73	5.37	14.54	0.70	53	72.96	71.39	1.5799	15.85	0.095
3	79.77	89.29	-9.92	19.67	0.70	59	105.44	88.85	16.594	27.36	0.0075
4	79.77	82.31	-2.54	15.65	0.70	60	81.53	56.04	25.49	31.873	0.10
5	77.05	76.32	0.72	14.43	0.75	61	92.01	98.88	-7.01	20.74	0.39
7	77.93	96.63	-18.70	26.46	0.70	62	102.49	102.95	-0.519	18.37	0.47
8	81.36	79.19	1.818	17.56	0.60	63	86.79	65.31	21.47	28.61	0.11
9	79.77	108.59	-29.58	33.23	0.66	64	111.53	93.68	17.86	23.96	0.38
10	79.03	101.70	-22.66	27.92	0.66	65	84.60	88.75	-4.197	19.30	0.30
11	79.77	110.18	-30.40	33.98	0.69	66	101.49	87.09	14.21	20.91	0.52
12	81.73	99.74	-18.01	24.44	0.72	67	95.42	105.31	-9.892	22.96	0.14
13	73.37	87.62	-14.25	22.87	0.72	68	68.65	66.13	2.235	17.06	0.28
14	75.95	70.50	5.444	15.26	0.73	70	70.93	74.39	-3.473	15.70	-0.0067
15	77.05	82.45	-5.393	17.97	0.72	71	77.00	69.52	7.776	12.80	0.53
16	81.73	101.50	-18.59	26.30	0.65	72	71.29	50.73	20.55	25.17	0.22
17	77.05	74.89	2.15	15.74	0.72	73	76.94	56.65	20.29	24.86	0.32
18	77.39	70.20	7.18	17.35	0.65	74	64.95	90.96	-26.29	37.17	0.54
19	77.39	75.18	2.20	16.25	0.66	75	51.14	68.18	-17.03	28.47	0.53
20	53.87	65.03	-11.15	23.32	0.55	76	70.71	60.93	9.42	19.83	0.61
21	44.83	64.87	-20.03	28.07	0.59	78	87.62	88.61	-1.751	18.40	0.44
22	42.61	54.15	-11.53	20.40	0.63	79	83.13	65.82	19.66	26.12	0.40
23	73.97	67.59	6.37	18.02	0.60	80	73.97	63.13	10.84	20.30	0.65
24	76.04	67.51	8.52	18.85	0.65	81	44.69	42.19	2.494	13.86	0.51
25	73.97	95.52	-21.54	31.14	0.61	82	43.53	36.63	6.81	20.35	0.19
26	76.04	93.159	-17.81	25.81	0.70	83	42.39	54.17	-11.47	19.91	0.25
28	70.86	75.26	-4.40	19.45	0.61	84	45.59	56.32	-10.46	18.60	0.61
29	67.82	66.97	0.18	14.48	0.73	85	56.61	58.85	-0.67	15.18	0.23
30	57.04	74.85	-14.90	20.90	0.37	86	38.32	59.84	-21.517	27.48	0.32
31	59.28	60.53	-1.12	16.37	0.70	87	41.96	46.93	-4.965	16.22	0.64
32	64.97	96.24	-31.27	39.85	0.56	88	52.34	48.68	3.668	16.32	0.33
33	56.43	74.62	-18.19	29.47	0.57	89	41.38	37.45	4.465	16.39	0.50
34	73.37	86.85	-13.53	22.92	0.73	91	39.42	49.26	-10.40	20.62	0.45
35	62.43	53.30	9.13	16.54	0.69	92	38.92	48.07	-13.09	19.59	0.45
36	60.24	87.41	-27.17	36.91	0.59	93	37.73	60.66	-23.03	32.47	0.22
37	63.43	61.76	1.67	11.34	0.77	94	49.18	49.45	-0.27	16.50	0.47
39	53.28	49.39	3.23	15.23	0.72	96	39.00	56.62	-16.59	22.79	0.27
42	65.81	60.59	4.47	12.85	0.74	97	38.76	56.97	-19.31	25.33	0.30
44	65.32	63.37	3.51	16.63	0.67	98	46.98	50.13	-1.90	16.75	0.59
46	59.19	54.19	6.75	14.63	0.69	102	82.75	74.27	9.01	18.99	0.52
47	59.11	67.62	-0.349	19.71	0.42	103	38.66	56.80	-18.07	26.60	-0.08
48	61.81	61.28	0.527	13.28	0.55	104	91.37	78.57	12.23	25.91	0.22
49	60.82	60.59	0.226	15.50	0.34	105	79.52	61.09	18.53	25.28	0.71
50	65.31	48.84	16.3	19.54	0.68	106	76.61	62.32	14.32	22.56	0.26
51	65.31	55.67	10.307	16.02	0.55	107	73.97	43.52	29.37	32.47	0.24
52	76.81	57.84	18.978	22.99	0.35	110	46.55	48.06	-0.60	13.74	0.54

Table 3.4: Numéro de la station(Stat°), Moyenne journalière modèle (MOY MODELE), Moyenne journalière mesurée (MOY MESURE), le biais, la RMS et le coefficient de corrélation (CORR) pour l'ozone

Stat	MOY Modele	MOY Mesures	BIAIS	RMS	CORR
111	39.77	47.01	-6.94	17.69	0.59
112	70.67	45.71	24.951	28.05	0.36
113	61.19	52.66	8.53	14.72	0.20
115	72.12	37.59	34.52	37.47	0.15
116	76.80	45.76	31.10	34.10	0.47
118	71.76	60.84	10.90	18.89	0.33
120	74.85	55.05	19.79	23.67	0.43
122	70.38	43.97	26.40	29.14	0.43
124	63.96	62.49	1.47	14.00	0.12
125	63.48	48.24	15.24	20.54	0.08
126	76.80	46.88	29.63	32.80	0.45
127	78.40	57.54	20.86	25.10	0.36
128	68.59	79.92	-11.32	25.48	0.42
129	77.19	66.46	10.72	16.82	0.60
130	80.13	56.35	23.77	27.45	0.41
131	69.67	69.80	-0.47	23.62	0.10
134	79.99	62.57	18.	24.80	0.02

Stat	MOY Modele	MOY Mesures	BIAIS	RMS	CORR
134	79.99	62.57	18.	24.80	0.02
135	78.33	38.60	41.1	43.16	0.17
136	69.15	62.46	6.52	17.40	0.34
137	70.03	63.17	6.86	14.73	0.69
138	81.11	59.32	22.15	25.98	0.47
139	66.31	52.99	13.31	19.09	0.15
140	80.64	68.36	12.23	16.34	0.68
141	73.21	46.03	27.18	31.79	0.17
142	81.24	47.27	33.96	35.87	0.61
143	84.03	64.52	19.61	26.49	0.35
144	84.03	92.74	-8.30	18.05	0.61
145	84.03	109.60	-25.60	30.29	0.65
146	84.03	89.91	-6.14	20.47	0.65
148	76.22	50.76	25.43	31.54	0.42
149	80.34	59.30	21.03	26.38	0.09
150	78.45	49.20	23.05	26.37	0.71
120 stat°	69.43	67.10	2.32	22.32	0.47

Table 3.5: Suite de Table 3.4

88 Chapitre 3. Evaluation du modèle Chimie-Transport POLAIR3D par comparaison aux données du réseau EMEP pour l'été 2001

Stat°	MAX MODELE	MAX MESURES	BIAIS	RMS	Corr	Stat	MOY MODELE	MOY MESURES	BIAIS	RMS	Corr
1	106.59	110.23	-3.64	20.48	0.72	49	70.81	81.92	-11.11	26.37	0.24
3	102.41	107.10	-5.31	18.63	0.76	50	77.30	70.5	6.55	17.42	0.48
4	102.41	108.52	-6.10	19.598	0.74	51	77.30	80.66	-2.62	18.18	0.50
5	100.59	104.75	-4.16	21.441	0.73	52	87.24	86.41	0.83	18.85	0.22
7	100.47	116.83	-16.36	26.248	0.75	53	80.60	89.07	-8.43	21.39	0.10
8	105.77	105.59	-0.26	17.777	0.73	59	125.50	113.80	11.69	27.74	0.035
9	102.41	123.58	-21.96	30.738	0.63	60	97.93	71.62	26.30	33.76	0.26
10	101.11	117.63	-16.51	26.00	0.68	61	112.13	133.73	-21.87	34.75	0.39
11	102.41	125.25	-22.83	30.66	0.65	62	121.48	126.67	-5.32	24.95	0.50
12	107.26	116.48	-9.21	19.04	0.77	63	107.96	90.24	17.72	30.02	0.046
13	95.28	112.96	-17.68	26.79	0.76	64	128.45	119.05	9.40	21.62	0.38
14	98.91	103.76	-4.85	18.64	0.80	65	105.81	114.61	-8.89	23.88	0.33
15	100.59	104.52	-3.92	18.93	0.78	66	121.16	113.09	7.93	18.36	0.60
16	107.26	122.55	-13.40	23.67	0.70	67	118.07	122.95	-4.88	23.53	0.12
17	100.59	112.60	-12.01	25.32	0.77	68	85.87	98.64	-13.08	28.47	0.25
18	101.14	108.	-6.85	23.23	0.71	70	77.39	89.83	-12.49	22.59	-0.03
19	101.14	107.09	-5.94	21.28	0.73	71	85.79	96.62	-11.60	18.26	0.35
20	73.76	93.43	-19.67	34.09	0.60	72	77.51	66.12	11.41	19.82	0.30
21	64.87	95.72	-30.85	41.77	0.63	73	85.41	75.77	9.63	18.20	0.37
22	63.40	93.59	-30.19	42.47	0.62	74	86.59	113.29	-26.89	40.89	0.63
23	95.19	106.11	-10.91	24.25	0.71	75	70.47	95.03	-24.55	38.10	0.60
24	98.61	106.53	-7.92	25.27	0.72	76	92.18	86.87	4.95	20.79	0.70
25	95.19	113.77	-18.58	30.84	0.67	78	108.57	109.01	-1.50	22.1	0.49
26	98.61	114.18	-16.23	28.79	0.72	79	103.63	89.64	16.74	25.54	0.57
28	93.47	94.08	-0.60	21.29	0.70	80	95.19	89.71	5.47	22.830	0.71
29	90.52	93.91	-4.16	20.07	0.75	81	55.68	60.44	-4.76	18.83	0.35
30	68.08	99.69	-27.931	37.20	0.17	82	56.54	57.19	-0.73	17.75	0.44
31	79.20	96.34	-16.825	28.37	0.72	83	53.95	75.11	-20.92	32.03	0.26
32	86.89	115.18	-28.29	39.37	0.66	84	62.08	76.95	-14.57	24.59	0.61
33	77.56	98.95	-21.39	33.91	0.68	85	67.37	69.52	-1.75	16.467	0.20
34	95.28	107.27	-12.20	22.86	0.79	86	51.18	77.33	-26.14	36.95	0.32
35	83.10	86.60	-3.50	19.27	0.75	87	61.48	79.07	-17.58	30.36	0.61
36	80.63	107.57	-26.94	35.89	0.72	88	63.57	65.75	-2.18	17.662	0.36
37	81.38	84.35	-2.966	15.20	0.77	89	56.22	67.69	-10.76	22.86	0.55
39	71.90	83.03	-10.88	26.13	0.74	91	57.16	78.36	-21.48	36.04	0.53
42	82.58	88.44	-6.59	19.99	0.71	92	55.98	66.59	-15.53	25.90	0.40
44	87.92	83.93	5.91	20.90	0.69	93	52.47	86.59	-34.29	48.45	0.15
46	77.11	85.41	-5.78	20.64	0.70	94	65.51	77.83	-12.32	27.06	0.52
47	81.10	109.44	-17.81	32.20	0.59	96	48.64	74.27	-24.57	34.60	0.18
48	76.20	80.73	-4.52	17.404	0.59	97	51.25	82.15	-32.07	40.68	0.32

Table 3.6: Numéro de la station, Maximum journalier simulé (MAX MODELE), Maximum journalière mesurée (MAX MESURE), le biais, la RMS et le coefficient de corrélation (CORR) pour l'ozone

Stat	MOY Modele	MOY Mesures	BIAIS	RMS	CORR	Stat ^o	MOY MODELE	MOY Mesures	BIAIS	RMS	CORR
98	65.89	85.46	-18.25	30.72	0.66	129	92.16	91.20	0.95	15.62	0.62
102	109.01	117.14	-6.86	23.48	0.61	130	97.19	84.97	12.21	22.11	0.50
103	50.25	66.89	-16.51	25.26	0.11	131	85.32	98.85	-13.89	31.92	0.08
104	107.25	152.09	-45.46	62.58	0.22	134	91.69	90.56	2.00	26.46	0.03
105	106.23	128.16	-21.78	42.71	0.70	135	101.76	84.11	20.48	27.09	0.49
106	90.70	87.03	3.73	20.41	0.33	136	78.91	86.86	-8.16	22.53	0.43
107	85.01	69.42	14.84	21.89	0.38	137	82.56	85.15	-2.58	17.02	0.67
110	63.83	73.62	-8.51	22.25	0.59	138	92.94	82.82	10.57	17.85	0.60
111	59.38	86.15	-26.65	42.13	0.63	139	71.96	65.12	6.84	15.78	0.26
112	79.35	68.64	10.70	17.75	0.38	140	93.48	85.13	8.28	15.69	0.67
113	66.01	67.65	-1.64	14.07	0.22	141	78.92	67.48	11.43	20.94	0.14
115	78.80	62.16	16.64	23.15	0.21	142	92.46	70.37	22.09	25.95	0.66
116	86.37	69.55	16.90	22.38	0.50	143	108.02	117.51	-9.50	24.13	0.60
118	81.43	83.40	-2.057	19.25	0.36	144	108.02	115.27	-6.7	20.24	0.63
120	85.10	74.47	10.63	19.76	0.41	145	108.02	128.	-20.07	28.58	0.63
122	79.00	65.01	13.96	20.46	0.49	146	108.02	113.46	-5.86	21.16	0.66
124	71.48	79.96	-8.488	19.07	0.21	148	99.23	84.15	15.06	28.72	0.47
125	68.79	59.47	9.315	17.23	0.18	149	100.91	91.182	9.73	23.83	0.17
126	86.37	68.67	17.46	24.40	0.50	150	102.65	81.82	13.33	21.76	0.76
127	98.28	92.30	5.984	23.12	0.41						
128	91.07	102.75	-11.68	28.97	0.48						
						120 stat^o	86.70	92.91	-6.2	25.3	0.505

Table 3.7: Suite de Table 3.6

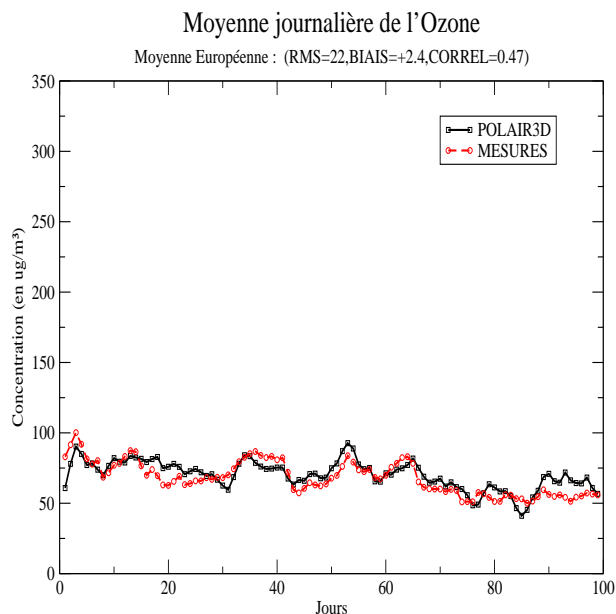


Figure 3.7: Comparaison du profil de la **moyenne journalière** pendant la période de validation du 24 Juin au 30 Septembre entre POLAIR3D et les mesures : moyenne sur toutes les stations du réseau EMEP

Conclusion

Dans le premier chapitre de cette partie nous avons donné une description du modèle Eulérien de Chimie-Transport POLAIR3, développé dans l'objectif de fournir une plateforme numérique pour les études de dispersion atmosphérique. Le modèle se caractérise par :

- **Sa modularité :**

Plusieurs mécanismes chimiques peuvent être employés sur la base d'une norme uniforme pour SPACK le préprocesseur chimique de POLAIR3D. Ils incluent, selon l'application, la chimie en phase gazeuse (EMEP, CBM4, RACM,...), la chimie multiphasique (chimie aqueuse et des aérosols), métaux lourds, mercure, radionucléides, etc...

- **Sa capacité multi-fonctions :**

POLAIR3D fournit l'évolution en temps et en espace pour un ensemble d'espèces chimiques données (**modèle direct**). L'adjoint et les versions tangent linéaires de POLAIR3D (dits **modèles dérivés**) sont disponibles automatiquement en utilisant le différentiateur automatique Odyssee [27]. Les modèles adjoint et tangent linéaire de POLAIR3D sont nécessaires pour les méthodes variationnelles dans l'assimilation de données, et pour les études d'analyse de

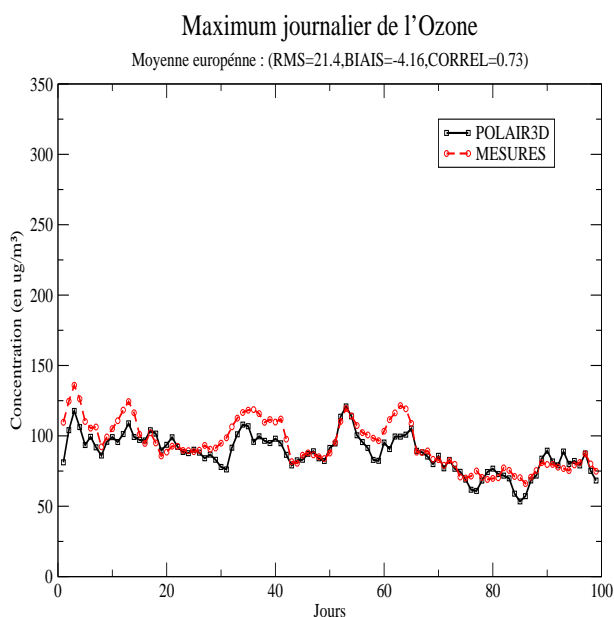


Figure 3.8: Comparaison du profil de la **maximum journalier** pendant la période de validation du 24 Juin au 30 Septembre entre POLAIR3D et les mesures : moyenne sur toutes les stations du réseau EMEP

sensibilité.

- **Sa capacité multi-échelles :**

POLAIR3D peut être employé à l'échelle régionale/urbaine (en général 100 km x 100 km) ou/et à l'échelle régionale/continentale (typiquement sur l'Europe).

L'objectif dans ce deuxième chapitre était de quantifier les performances du modèle POLAIR3D développé et décrit lors du premier chapitre. Nous avons présenté un exemple d'étude préliminaire de validation, sur une période de 3 mois et sur l'Europe en se focalisant sur la concentration de l'ozone.

Nous avons pu vérifier que le modèle obtient des résultats satisfaisants et très réalistes en reproduisant des concentrations de l'ozone comparables à celles mesurées (biais= $-6.2(\mu\text{g}/\text{m}^3)$, RMS= $25.3(\mu\text{g}/\text{m}^3)$ pour le maximum journalier).

Pour compléter ce travail de validation, il est à noter que d'autres travaux sont en cours de réalisation pour valider le modèle à d'autres échelles (régionale, urbain) en comparaison à des résultats de campagnes de mesures récentes avec une focalisation, en plus de l'ozone sur d'autres espèces chimiques telles que le NO_2 , NO , SO_2 , CO et les particules. De plus, des paramétrisations ont été amélio-

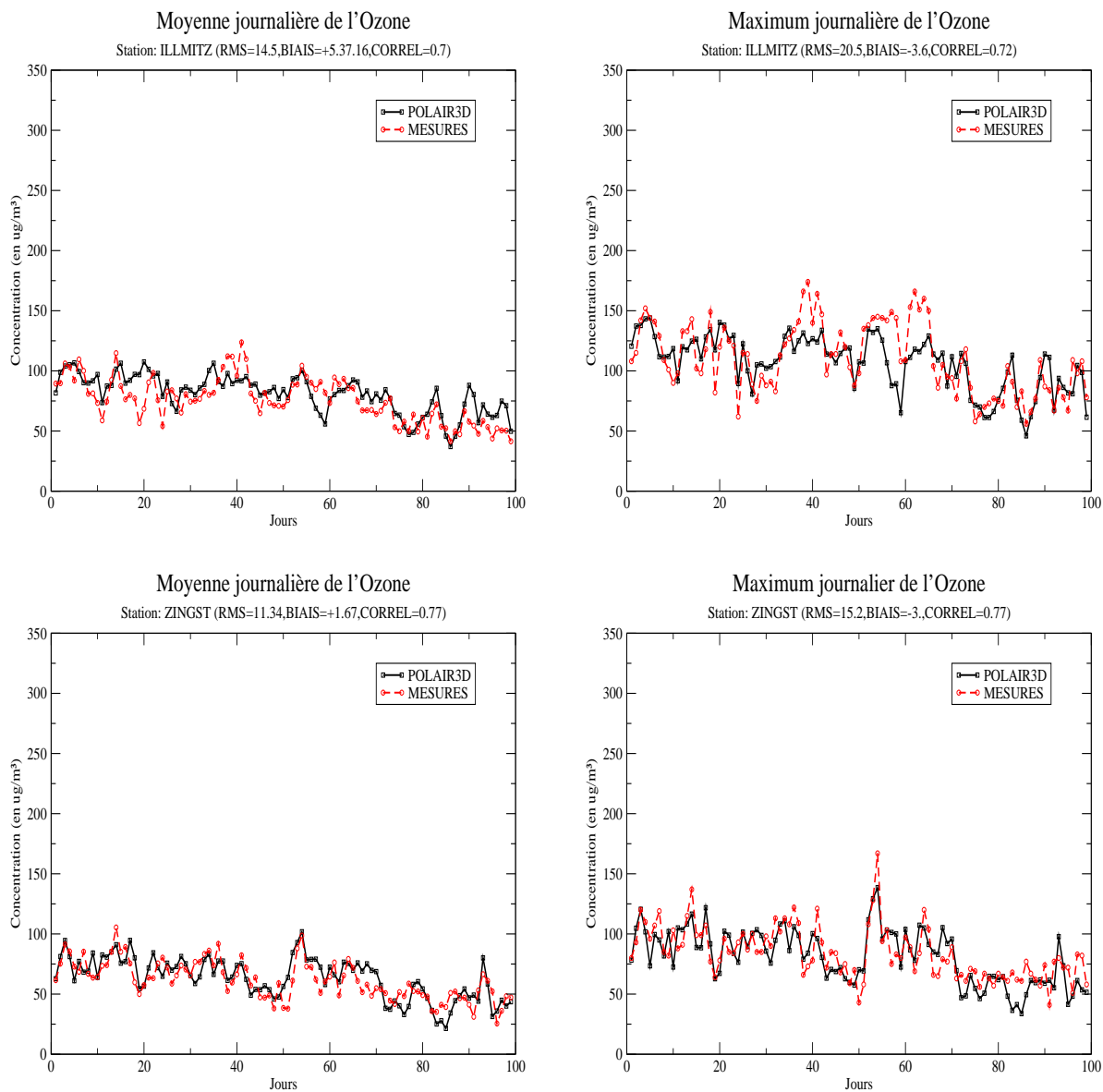


Figure 3.9: Comparaison du profil de la **moyenne journalière** et du **maximum journalier** pendant la période de validation du 24 Juin au 30 Septembre entre POLAIR3D et les mesures, sur quelques stations du réseau EMEP

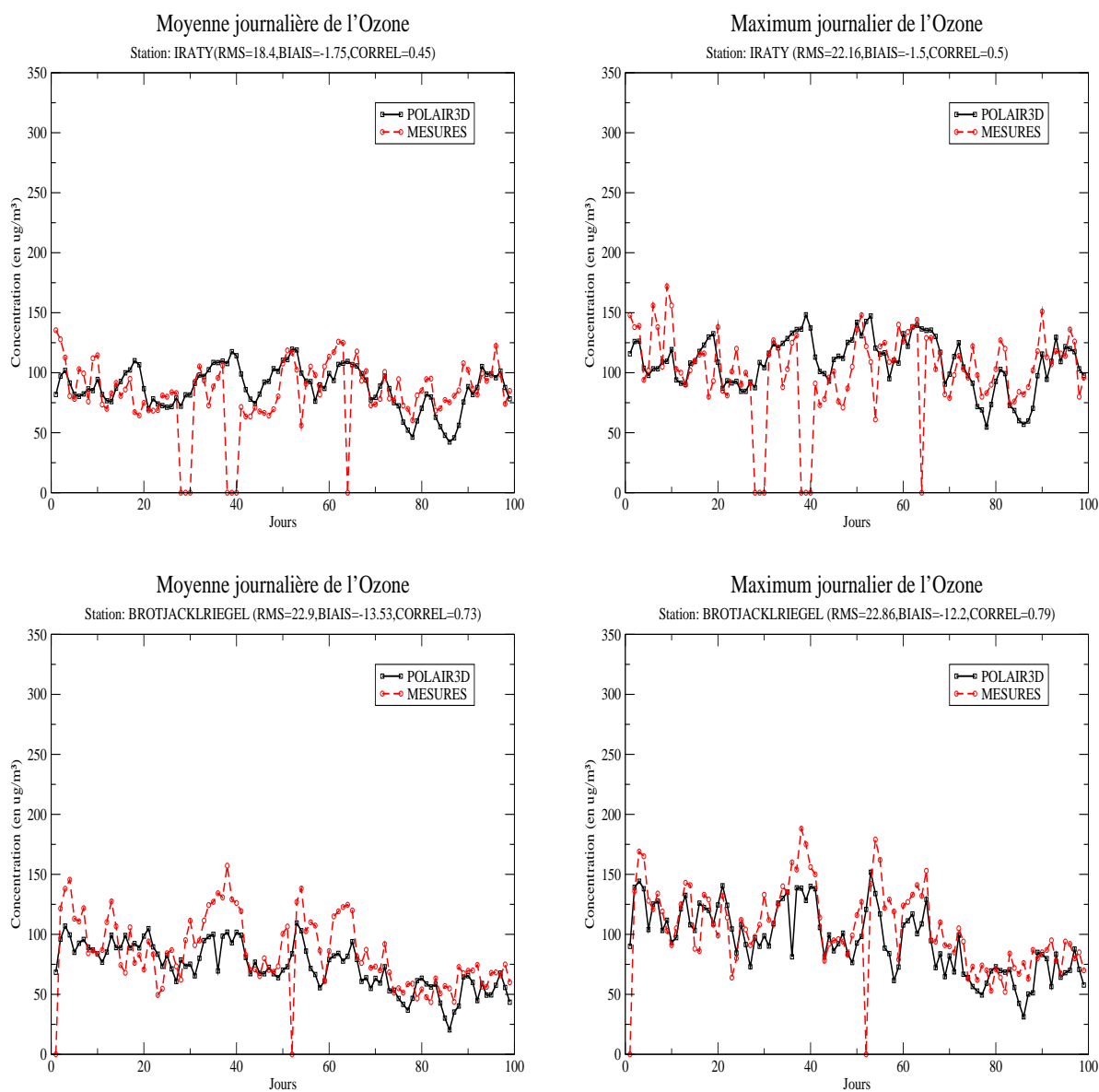


Figure 3.10: Comparaison du profil de la **moyenne journalière** et du **maximum journalier** pendant la période de validation du 24 Juin au 30 Septembre entre POLAIR3D et les mesures, sur quelques stations du réseau EMEP

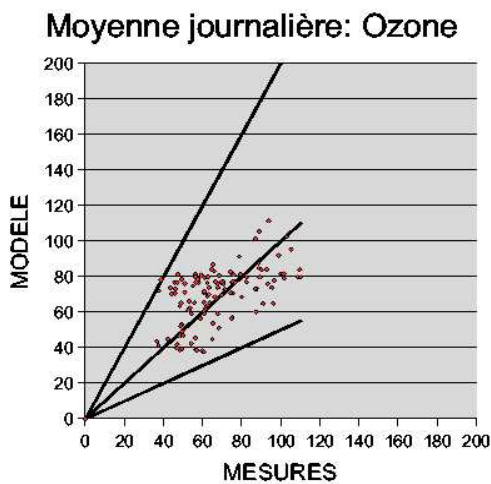


Figure 3.11: Nuage de points Mesures/Modele :moyenne journalière

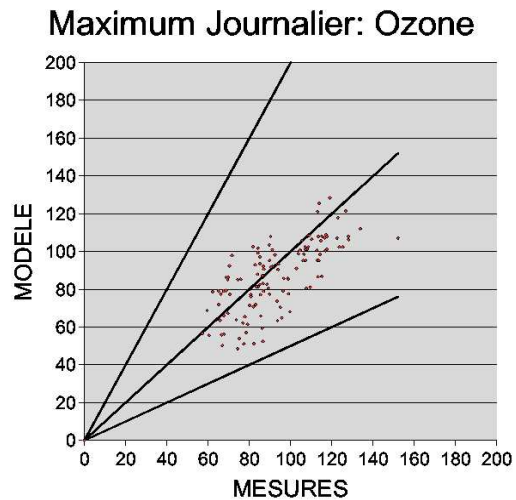


Figure 3.12: Nuage de points Mesures/Modele :maximum journalier

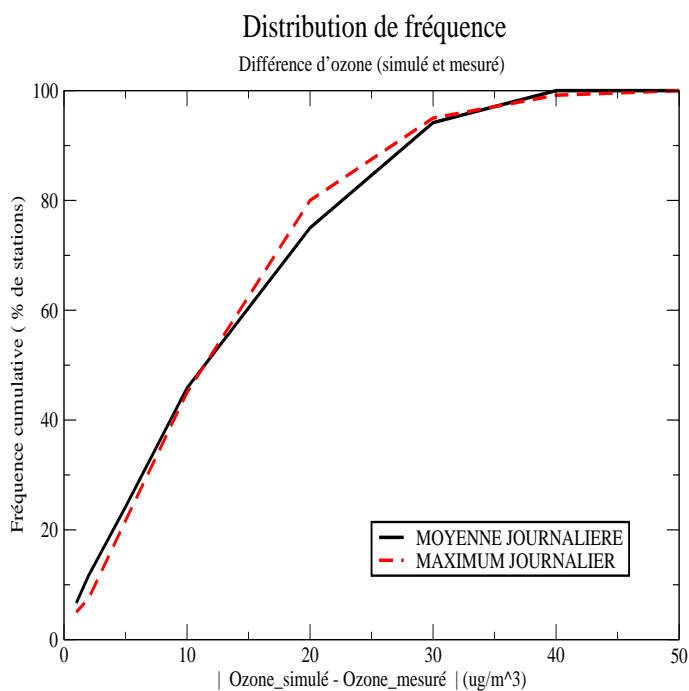


Figure 3.13: Distribution de fréquence

rées afin d'augmenter ces performances (ce qui n'était pas l'objectif de notre travail).

Pour la suite, nous estimerons que le modèle est capable de bien représenter les processus de Chimie-Transport à l'échelle continentale. Ainsi tout au long de cette thèse, tous nos travaux sont effectués en utilisant le modèle POLAIR3D.

Partie III

Réduction de modèle

Chapitre 4

Tabulation réduite avec la méthode HDMR (High Dimensional Model Representation)

Résumé

L'objectif de ce chapitre est d'utiliser les méthodes de tabulation réduite HDMR (High Dimensional Representation Model), dans le contexte de réduction de la cinétique chimique.

Nous présentons ici l'algorithme et le code qui ont été développés, indépendamment du modèle utilisé. Une première application à la cinétique a été effectuée avec les mécanismes EMEP, MOCA et CBM4. A précision conservée, des gains importants en temps calcul ont été effectués.

Une première application à un cas tridimensionnel est enfin présenté avec le modèle POLAIR3D.

Introduction

La cinétique chimique fait intervenir un grand nombre de réactions et d'espèces chimiques dont les échelles de temps ont des ordres de grandeur très différents les unes des autres. Cette dispersion des échelles en temps induit la raideur des équations différentielle qui décrivent la chimie. Ces équations sont, en plus, fortement couplées entre elles et sont non-linéaires, ce qui rend plus difficile et plus coûteux la résolution d'un modèle de cinétique chimique. Ce problème exige l'utilisation de techniques numériques appropriées, les meilleurs algorithmes numériques restant extrêmement chers pour résoudre de telles équations ([3]).

Dans les modèles 3D de chimie-Transport, la cinétique chimique peut consommer entre 50% et 90% du temps CPU total de la simulation. Cette situation devient encore plus critique quand des espèces et des réactions chimiques sont à rajouter dans les modèles pour les rendre plus complets et plus réalistes(par exemple : la prise en compte de plusieurs phases).

Une première approche pour répondre à cette difficulté est liée à la réduction dynamique de tels systèmes. En se fondant sur le caractère lent/rapide des système différentiels résultants, des modèles réduits de plus petite taille peuvent être construits en filtrant les composantes rapides [15].

Une seconde approche est liée au comportement statistique de ces systèmes et est fondée sur une décomposition en modes propres des solutions [16].

Une troisième approche, basée sur des développements en polynômes paramétrisés et sur des interpolations des tables pré-calculées (look-up table) peut être utilisée comme modèle de représentation pour remplacer les solveurs de chimie. On se réfère à [17] pour plus de détails. Pour un modèle ayant x_1, x_2, \dots, x_n variables d'entrée, pour une résolution en $1/s$ (s points d'échantillonnage par direction), la construction de la table traditionnelle d'interpolation nécessite l'exécution de s^n fois le modèle initial. En plus des difficultés liées aux interpolations à n dimensions, ces approches restent prohibitives pour des modèles chimiques qui comprennent un très grand nombre de degrés de liberté(une centaine d'espèces chimiques typiquement).

Récemment une méthode nommée "High Dimensional Model representation" (HDMR), a été présentée dans [23] pour améliorer l'efficacité de déduction de la relation existante entre les variables d'entrée et celles de sortie dans un modèle complexe de haute dimension.

L'objectif de notre étude est d'appliquer dans le contexte des modèles 3D de Chimie-Transport en tant que substitut aux solveurs de la cinétique chimique afin d'améliorer la rapidité de leur exécution et rendre ainsi possible des études d'impact à scénarii multiples.

La méthode HDMR est un développement en n séries de fonctions de variables d'entrée corrélées qui représentent efficacement les liens existants entre les entrées et

sorties du modèle. L'ordre k de l'expansion correspond à l'effet coopératif résiduel (une fois déduits ceux des ordres précédents) des k uplets de variables d'entrée.

Si aucune troncature n'est effectuée, le développement est exact. L'ansatz, sur lequel se fonde la méthode HDMR, est qu'*a priori*, il est suffisant de se limiter aux interactions entre couples voire triplets de variables pour restituer l'information contenue dans les modèles physiques. Ceci est vérifié en pratique et on se limitera ici au développement à l'ordre 2.

La méthode HDMR se caractérise par :

- des opérations basées seulement sur des manipulations algébriques très rapides et stables, ce qui rend la méthode efficace ;
- une exactitude comparable aux solveurs de la cinétique chimique, tout en atteignant une rapidité significative de temps CPU ;
- une prise en compte de toutes les variables du système de dimension élevée.

Elle peut être appliquée pour différentes applications, notamment :

- la construction d'un modèle opérationnel entièrement équivalent au modèle d'origine (il calculera directement les concentrations d'espèces de sortie à partir des concentrations initiales d'espèce d'entrée et des paramètres de forçage),
- l'identification des variables d'entrée principales du modèle,
- l'étude de l'incertitude des sorties de modèle en fonction de celles des variables d'entrée. HDMR étant beaucoup plus rapide en temps d'exécution que le modèle initial, une approche de type Monte Carlo est utilisable,
- l'identification des corrélations entre variables, ce qui permet d'avoir une meilleure compréhension du comportement qualitatif du système.

Nous présentons ici, les principales propriétés de cette approche (section 4.1), notre implémentation algorithmique (section 4.2), des exemples simples d'illustration (section 4.3), une application à des cinétiques chimiques utilisées pour l'étude de la pollution atmosphérique troposphérique (section 4.4) et des études de sensibilités de HDMR au point de référence et au nombre de points d'échantillonnage (section 4.5).

4.1 Présentation de la méthode HDMR

4.1.1 Formulation

La plus grande difficulté des tabulations traditionnelles des relations d'entrée/sortie d'un système complexe de grande dimension vient du nombre de calculs qui augmente exponentiellement avec le nombre d'entrée n et le nombre de points d'échantillonnage en s^n .

Considérons la variable d'entrée $x = (x_1, x_2, \dots, x_n) \in \Omega$ avec n le nombre de

variables d'entrée (de l'ordre de 100 ou plus dans nos applications à la cinétique chimique), et $g(x)$ la sortie du modèle.

L'influence des variables d'entrée sur les variables de sortie peut être indépendante et/ou simultanée. Le principe d'HDMR est d'isoler les effets coopératifs entre variables : HDMR exprime ainsi la sortie $g(x)$ comme un développement exact, lorsqu'il n'est pas tronqué, de fonctions de variables d'entrée selon :

$$g(x) = f_0 + \sum_{i=1}^{i=n} f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) + \sum_{1 \leq i < j < k \leq n} f_{ijk}(x_i, x_j, x_k) + \dots + f_{1,2,\dots,n}(x_1, x_2, \dots, x_n) \quad (4.1)$$

où apparaissent successivement dans le développement :

- f_0 qui représente un **effet moyen** sur $g(x)$,
- les fonctions de premier ordre, $f_i(x_i)$, qui représentent l'effet de la variable d'entrée x_i agissant indépendamment, bien que généralement le système est non-linéaire, sur la sortie $g(x)$ (figure 4.1),

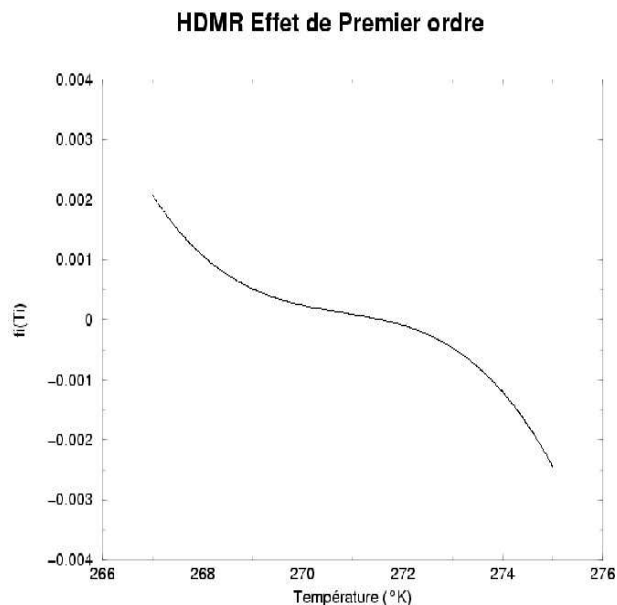


Figure 4.1: Effet de premier ordre sur la sortie NO : le profil de la fonction f_i relative à la variable d'entrée T (Température), mécanisme EMEP

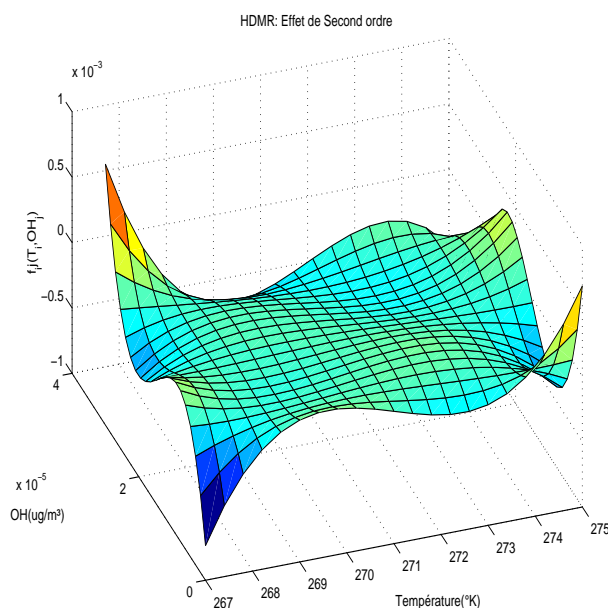


Figure 4.2: Effet de deuxième ordre sur la sortie NO : le profil de la fonction f_{ij} relative à la variable d'entrée T (Température) et la concentration OH (espèce forcée), mécanisme EMEP

- les fonctions de deuxième ordre, $f_{ij}(x_i, x_j)$, qui décrivent l'effet des variables d'entrée x_i et x_j agissant ensemble sur la sortie $g(x)$ (figure 4.2),
- ...,
- le dernier terme $f_{1,2,\dots,n}(x_1, x_2, \dots, x_n)$ représente le terme résiduel de toutes les variables d'entrée agissant ensemble sur la sortie $g(x)$.

L'intérêt de la structure HDMR est que chaque fonction du développement exprime la corrélation physique et l'effet des variables d'entrée (en agissant seule ou avec d'autres) sur la variable de sortie. En analysant les comportements des fonctions du développement f_0, f_i, f_{ij}, \dots , on peut déterminer :

- les variables qui ont un effet important sur une sortie du modèle,
- les variables d'entrée qui agissent seules et celles qui agissent de manière corrélée sur la sortie.

Plusieurs études [17, 23, 24, 64] prouvent qu'une expression de HDMR au deuxième ordre :

$$g(x) \approx f_0 + \sum_{i=1}^{i=n} f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) \quad (4.2)$$

fournit souvent une description satisfaisante de $g(x)$ pour des applications de dimension élevée liées à des modèles physiques. Ainsi, la détermination de la table d'entrée/sortie d'un modèle de grand dimension devient plus efficace et plus rapide. Notons que la prise en compte d'un faible nombre de fonctions dans le développement n'implique pas nécessairement une prise en compte d'un nombre petit de variables d'entrée significatives et conserve la nature non-linéaire de la relation entrée-sortie.

4.1.2 Coût de calcul des tables HDMR

Dans la figure 4.3 nous présentons une comparaison entre le coût de construction des tables HDMR et la méthode traditionnelle de tabulation. Pour cette dernière son coût de génération des tabulations croît d'une manière très forte en fonction du nombre d'entrées du modèle contrairement à la méthode HDMR.

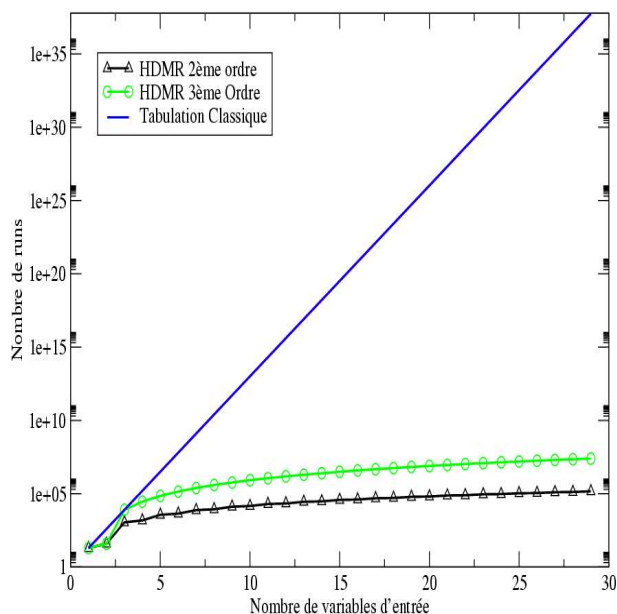


Figure 4.3: Coût de génération des tables de HDMR

4.1.3 Détermination des fonctions composantes de HDMR

Nous allons à présent préciser le mode de calcul des fonctions du développement, qui est directement relié au fondement théorique de la méthode.

4.1.3.1 Décomposition HDMR

Etant donnée une mesure $\mu(dx) = \prod_i \mu_i(dx_i)$ définie sur Ω , on note V l'espace des fonctions μ -intégrables.

Pour des raisons de simplicité de présentation, les variables d'entrée x_i sont transformées afin que $0 \leq x_i \leq 1$. La fonction de sortie $g(x)$ est alors définie dans le cube unitaire $K^n = \{(x_1, x_2, \dots, x_n) \mid 0 \leq x_i \leq 1, i = 1, 2, \dots, n\}$.

Le lemme suivant est fondamental et donne une décomposition en sous-espaces vectoriels de l'espace vectoriel V :

$$V = V_0 \oplus \sum_i V_i \oplus \sum_{i < j} V_{ij} \oplus \dots \sum_{i_1 < i_2 < \dots < i_l} V_{i_1 i_2 \dots i_l} \oplus \dots \oplus V_{12 \dots n} \quad (4.3)$$

avec :

- $V_0 = \{f \in V, f = C \in R\}$,
- $V_i = \{f \in V, f = f_i(x_i), \int f_i(x_i) d\mu_i(x_i) = 0\}$,
- $V_{ij} = \{f \in V, f = f_{ij}(x_i, x_j), \int f_{ij}(x_i, x_j) d\mu_k(x_k) = 0, k = i, j\}$,
- $V_{12 \dots n} = \{f \in V, \int f(x) d\mu_k(x_k) = 0, k = 1, 2 \dots n\}$.

La décomposition HDMR est directement reliée à cette décomposition en sous-espaces, dont la preuve constructive est la suivante :

- $f_0 = \int g(x) \mu(dx)$,
- $f_i(x_i) = \int g(x) \prod_{k \neq i} \mu_k(dx_k) - f_0$,
- $f_{ij}(x_i, x_j) = \int g(x) \prod_{k \neq i, j} \mu_k(dx_k) - f_i(x_i) - f_j(x_j) - f_0$,
- ...

On se réfère par exemple à [23, 24, 65]

4.1.3.2 Erreur de troncature

Notons par ailleurs que le développement HDMR assure que *l'erreur de troncature est minimale*. En effet, on peut montrer que le problème de minimisation suivant :

$$\begin{aligned} \min J(u) &= \int_{\Omega} (g(x) - u)^2 \mu(dx) , \\ u &\in V_0 \oplus \sum_i V_i \oplus \sum_{i < j} V_{ij} \oplus \dots \sum_{i_1 < i_2 < \dots < i_l} V_{i_1 i_2 \dots i_l} \end{aligned} \quad (4.4)$$

a pour solution les premiers ordres du développement HDMR :

$$u = f_0 + \sum_{i=1}^{i=n} f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) \dots + \sum_{i_1 < i_2 < \dots < i_l} f_{i_1 i_2 \dots i_l}(x_{i_1}, x_{i_2}, \dots, x_{i_l}) \quad (4.5)$$

La preuve se fait directement en écrivant le système d'optimisation associé par annulation de la dérivée première de u .

Ceci justifie biensûr la pertinence à utiliser des troncatures de cette distribution.

4.1.3.3 Choix de la mesure μ

Pour finir cette présentation, notons que le choix des poids μ est crucial et détermine de manière unique le développement HDMR.

Nous allons présenter deux méthodes associées à des choix différents de μ et expliciter les fonctions de développement de HDMR :

- la méthode Cut-HDMR est associée au choix d'un point de référence au sein du domaine Ω et la mesure choisie est alors :

$$\mu_{cut}(dx) = \Pi_i \delta_{\bar{x}_i}(x_i) \quad (4.6)$$

- la méthode RS-HDMR (Random Sampling) est associée à une distribution équiprobable avec :

$$\mu_{RS}(dx) = \Pi_i dx_i \quad (4.7)$$

Dans la suite de ce chapitre nous nous focalisons sur la méthode cut-HDMR, la deuxième approche de déterminations des fonctions composante de HDRM, dite RS-HDMR, fera l'objet du chapitre 7 dans le cadre d'une étude de sensibilité du modèle POLAIR3D.

4.1.4 Méthode Cut-HDMR

4.1.4.1 Cut-HDMR

La méthode Cut-HDMR est déterminée par le choix du point de référence $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ dans le domaine Ω . Dans toute la suite, on utilise s points d'échantillonnage des variables d'entrée dans chaque direction (ce nombre peut bien sûr varier dans chaque direction, mais on présente le cas constant par souci de clarté).

Les fonctions composantes de cut-HDMR $f_i(x_i)$, $f(x_i, x_j)$, ... sont déterminées numériquement par des coupes de droites, plans, volumes, etc. à travers le point de référence \bar{x} du domaine Ω , d'où le nom de Cut-HDMR. Elles sont données par [65, 66] :

$$f_0 = g(\bar{x}) \quad (4.8)$$

$$f_i(x_i) = g(x_i, \bar{x}^i) - f_0 \quad (4.9)$$

$$f_{ij}(x_i, x_j) = g(x_i, x_j, \bar{x}^{ij}) - f_i(x_i) - f_j(x_j) - f_0 \quad (4.10)$$

où :

- f_0 la fonction constante calculée à partir de $g(x)$, sortie du modèle évaluée au point de référence (toutes les variables d'entrée sont prises au point de référence \bar{x}).
- Les fonctions de premier ordre $f_i(x_i)$ sont calculées alors à partir de $g(x_i, \bar{x}^i)$ en mettant toutes les variables d'entrée au point de référence sauf la variable x_i qui est à sa valeur prise de l'intervalle choisi ($g(x_i, \bar{x}^i) = g(\bar{x}_1, \dots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \dots, \bar{x}_n)$).
Le terme f_0 est soustrait de chaque sortie du modèle utilisé avec x_i variant dans son intervalle de départ, comme montré dans l'équation (4.9). Pour s points utilisés pour chaque variable d'entrée, afin de couvrir l'espace des données d'entrée, $(s - 1)$ exécutions de modèle sont nécessaires pour calculer les fonctions de premier ordre, l'exécution du modèle au point de référence n'est pas nécessaire vu que la fonction est égale à zéro selon la loi d'orthogonalité (4.11). Si nous disposons de n variables d'entrée, nous aurons alors n fonctions de premier ordre $f_i(x_i)$, et donc nous aurons besoin d'exécuter le modèle $n(s - 1)$ fois.
- Les fonctions de second ordre $f_{ij}(x_i, x_j)$ sont calculées à partir de $g(x_i, x_j, \bar{x}^{ij})$ en mettant toutes les variables d'entrée à leur point de référence sauf les variables x_i and x_j qui sont prises sur leur intervalle respectif couvrant tout l'espace possible des données d'entrée du modèle. Comme chaque variable d'entrée est représentée par s points, $(s-1)^2$ exécutions de modèle sont nécessaires pour calculer la fonction de second-ordre $f_{ij}(x_i, x_j)$ pour chaque variable d'entrée, les fonctions aux points de référence \bar{x}_i, \bar{x}_j valent zéro et donc l'exécution du modèle n'est pas nécessaire (4.11). Les fonctions à un ordre supérieur sont calculées de la même manière, mais en pratique on se contente de l'ordre inférieur 2.
- Le dernier terme $f_{12\dots n}(x_1, x_2, \dots, x_n)$ est déterminé par la différence entre $g(x)$ et les fonctions déjà calculées avec l'équation (7.1).

Il est aisé de vérifier qu'une fonction composante de cut-HDMR est nulle lorsqu'une de ses variables prend sa valeur en une composante du point de référence \bar{x} , c'est à dire :

$$f_{i_1 i_2 \dots i_l}(x_{i_1}, x_{i_2}, \dots, x_{i_l}) \Big|_{x_s = \bar{x}_s} = 0, s \in \{i_1, i_2, \dots, i_l\} \quad (4.11)$$

ce qui définit une relation d'orthogonalité entre deux fonctions composantes de cut-HDMR selon :

$$f_{i_1 i_2 \dots i_l}(x_{i_1}, x_{i_2}, \dots, x_{i_l}) f_{j_1 j_2 \dots j_k}(x_{j_1}, x_{j_2}, \dots, x_{j_k}) \Big|_{x_s = \bar{x}_s} = 0. \quad (4.12)$$

$$s \in \{i_1, i_2, \dots, i_l\} \cup \{j_1, j_2, \dots, j_k\}$$

La génération des tables de HDMR est schématisée par la figure 4.4 et basée sur :

- le choix des variables d'entrée,

- le choix d'un point de référence dans l'espace des variables d'entrée (nos tests indiquent une faible sensibilité à ce choix),
- la définition des variables de sorties, pour lesquelles HDMR sera appliqué,
- l'exécution du modèle initial pour calculer les différentes fonctions du développement HDMR pour chaque sortie définie,
- le calcul de coefficients d'interpolations linéaires et bilinéaires des fonctions de développement HDMR dans l'espace des variables d'entrée définie,
- la sauvegarde des tables des coefficients pour chaque variable de sortie définie.

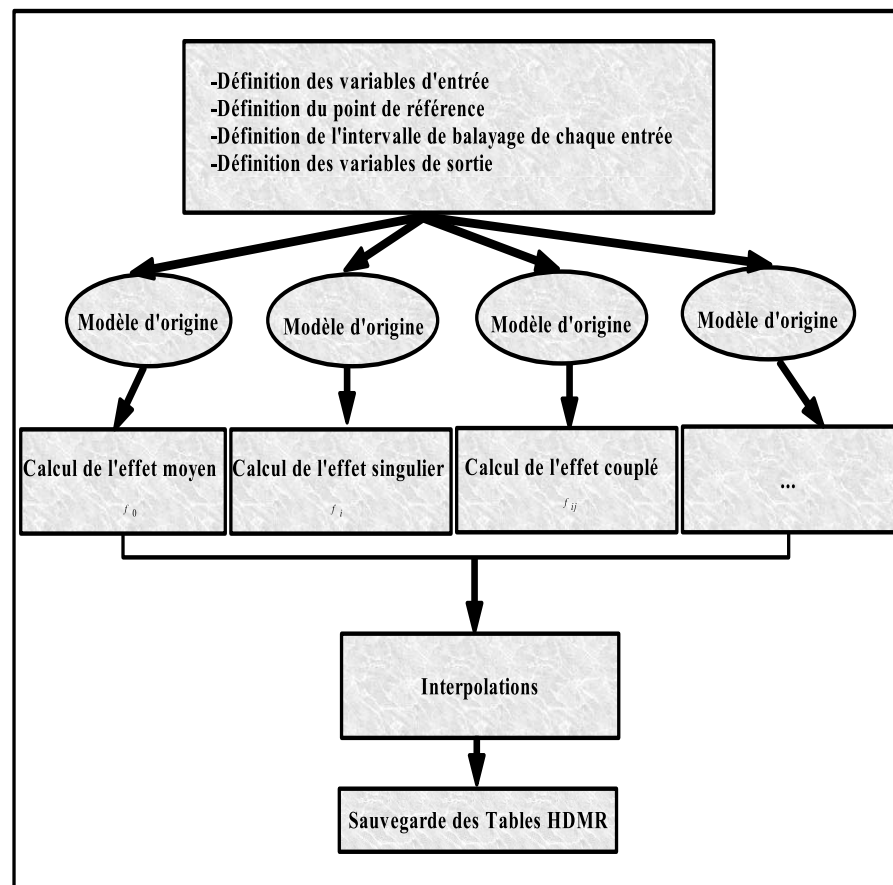


Figure 4.4: Schéma de génération des tables de HDMR

Les procédures de minimisation et le choix optimal des fonctions de développement de Cut-HDMR à un ordre petit rendent la construction des tables efficace et rapide. Ainsi la valeur du modèle $g(x)$ pour un point arbitraire x est obtenue à partir

de ces tables en effectuant seulement des interpolations de dimension inférieure sur les fonctions $f^i(x_i)$, $f(x_i, x_j)$, \dots

Précisons pour finir le coût du calcul de la table. Si chaque variable d'entrée prend s valeurs dans l'espace des données d'entrée, pour construire les tables de $f^i(x_i)$, $f(x_i, x_j)$ le modèle est utilisé :

$$\sum_{i=0}^L \frac{n!}{(n-i)!i!} (s-1)^i \quad \text{fois} \quad (4.13)$$

pour un ordre L général ($1 \leq L \leq n$). Dans le cas du second ordre, le modèle initial est utilisé :

$$1 + n(s-1) + \frac{n(n-1)(s-1)^2}{2} \quad \text{fois} \quad (4.14)$$

ce qui est à comparer au calcul des tables traditionnelles qui demande s^n exécutions de modèles.

4.1.5 Quelques propriétés des fonctions de décomposition

Le développement HDMR possède des propriétés intéressantes, notamment pour les aspects suivants :

- *Convergence* :

Comme mentionné ci-dessus, les développements en série de fonctions de HDMR convergent rapidement, et en général l'ordre 2 est suffisant pour représenter le comportement des variables d'entrée vis-à-vis des variables de sortie [65].

- *Conservation des lois Physiques* :

Si un ensemble de sorties $\{f^1(x), f^2(x), f^3(x), \dots, f^s(x)\}$ obéit à un ensemble de lois linéaire de conservation (de masse, d'énergie, de moment,...) leurs approximations par HDMR, à n'importe quel ordre, obéissent également à ces lois de conservation [65].

- *Décomposition en variance par RS-HDMR* :

En utilisant la propriété d'orthogonalité des fonctions composantes de RS-hdmr, la variance totale σ^2 de $g(x)$ induites par toutes les variables d'entrée

peut être décomposée en une série de variances des contributions des variables d'entrée [65] :

$$\sigma^2 = \sum_{i=1}^n \sigma_i^2 + \sum_{1 \leq i < j \leq n} \sigma_{ij}^2 + \dots \quad (4.15)$$

Cette propriété est utile pour l'analyse d'incertitude (la décomposition ci-dessus est valable sur tout le domaine). Selon les importances de σ_i^2 , σ_{ij}^2 , etc..., il est facile de suivre comment l'incertitude des variables de sortie est influencée par les incertitudes des variables d'entrée.

4.1.6 Méthode MPcut-HDMR

En pratique, on a vu que l'on se limite à l'effet des variables agissant seules et/ou en combinaison avec d'autres variables (troncature au 2ème ordre).

Néanmoins, dans quelques cas, la prise en compte du deuxième ordre n'est pas suffisante, et on peut devoir aller plus loin dans la troncature du développement HDMR, jusqu'à l'ordre 3, 4 ou plus.

La prise en compte d'ordres supérieurs exige l'utilisation d'un grand nombre de simulations de modèle et une grande disponibilité de mémoire pour la construction des tables HDMR. Dans le tableau 4.1 sont présentés, le nombre de simulations et la mémoire nécessaire pour la production des fonctions de développement HDMR en fonction de l'ordre retenu.

Ainsi pour un nombre de variable d'entrée $N=100$ et le nombre d'échantillonnage $N_s = 10$, la mémoire demandée s'élève à 65 GO. Avec l'état de l'art des stations de travail actuelles, cette puissance demandée reste prohibitive, encore plus quand l'ordre 2 et 3 ne sont pas suffisants pour représenter le modèle à réduire par HDMR.

Dans ce contexte et afin de réduire les ressources demandées et le nombre de simulations pour construire la table HDMR qui nécessiterait une prise en compte d'ordre supérieur à 2, une méthode dite '**MP-Cut HDMR**' a été présentée dans [67]. L'idée de base de cette méthode est d'approcher les fonctions d'ordre supérieures (f_{i_1, i_2, \dots, i_l}) par des fonctions d'ordre inférieur. Dans le cas de l'ordre 3, l'idée est d'estimer les fonctions de développement HDMR de troisième ordre f_{ijk} à partir des fonctions de second ordre, d'une manière optimale.

Ordre de troncature	1	2	3
Nombre de Runs	$1+N * N_s$	$1+N * N_s + \frac{N(N-1)(N_s-1)^2}{2}$	$1+N * N_s + \frac{N(N-1)(N_s-1)^2}{6} + \frac{N(N-1)(N-2)(N_s-1)^3}{6}$
Mémoire demandée (octet)	$(1+N * N_s)*N_{out}*4$	$(1+N * N_s + \frac{N(N-1)(N_s-1)^2}{2}) * N_{out} * 4$	$(1+N * N_s + \frac{N(N-1)(N_s-1)^2}{6} + \frac{N(N-1)(N-2)(N_s-1)^3}{6}) * N_{out} * 4$

Table 4.1: Nombre de simulations et mémoire demandée pour le calcul des fonctions de Cut-HDMR

$$\begin{aligned}
f_{ijk}(x_i, x_j, x_k) \approx & \phi_{ijk}(x_i, x_j, x_k)\bar{f}_0 + \phi_{jk}(x_j, x_k)\bar{f}_i(x_i) \\
& + \phi_{ik}(x_i, x_k)\bar{f}_j(x_j) + \phi_{ij}(x_i, x_j)\bar{f}_k(x_k) \\
& + \phi_k(x_k)\bar{f}_{ij}(x_i, x_j) + \phi_j(x_j)\bar{f}_{ik}(x_i, x_k) \\
& + \phi_i(x_i)\bar{f}_{jk}(x_j, x_k)
\end{aligned} \tag{4.16}$$

où $\phi_i(x_i)$, $\phi_j(x_j)$, $\phi_k(x_k)$, ..., $\phi_{ijk}(x_i, x_j, x_k)$ sont des polynômes spécifiques (voir plus loin) et \bar{f}_0 , $\bar{f}_i(x_i)$, ..., $\bar{f}_{jk}(x_j, x_k)$ sont des fonction cut-HDMR pour une fonction particulière $\bar{f}(x)$ qui dépend de la fonction $g(x)$ qu'on cherche à approcher par HDMR.

Supposons que $g(x)$ est approchée par Cut-HDMR avec une troncature d'ordre 2, au point de référence \bar{x} . L'erreur faite avec cette approximation est :

$$\begin{aligned}
r_2(x) &= g(x) - f_0 - \sum_{i=1}^{i=n} f_i(x_i) - \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) \\
r_2(x) &\approx \sum_{1 \leq i < j < k \leq n} f_{ijk}(x_i, x_j, x_k)
\end{aligned} \tag{4.17}$$

Les fonctions $f_{ijk}(x_i, x_j, x_k)$ de développement HDMR peuvent être prises comme des sommes des termes du développement de Taylor à un point de référence $x^0 =$

(x_i^0, x_j^0, x_k^0) qui contiennent seulement les variables (x_i, x_j, x_k) selon :

$$\begin{aligned} f_{ijk}(x_i, x_j, x_k) &= f_{ijk}(x^0) + \sum_{i=1}^3 \frac{\partial f_{ijk}(x^0)}{\partial x_i} * (x_i - x_i^0) \\ &+ \sum_{i,j=1}^3 \frac{1}{2!} \frac{\partial^2 f_{ijk}(x^0)}{\partial x_i \partial x_j} * (x_i - x_i^0)(x_j - x_j^0) + \dots \end{aligned} \quad (4.18)$$

Si nous considérons un sous-ensemble $I = (i_1, i_2, i_3)$ de $1, 2, 3, \dots, n$ et $x_I = x_{i_1}, x_{i_2}, x_{i_3}$, $\bar{x}_I = \bar{x}_{i_1}, \bar{x}_{i_2}, \bar{x}_{i_3}$ point de référence associé à x_I et $\bar{z}_I = \bar{z}_{i_1}, \bar{z}_{i_2}, \bar{z}_{i_3}$ un deuxième point de référence choisi dans $x = (x_1, x_2, \dots, x_n)$ différent de \bar{x}_I , la méthode MPcut HDMR donne l'approximation des fonctions HDMR d'ordre supérieure ([67]), pour une fonction d'ordre m ($m > 2$), $m = l + 1$, où l est l'ordre de la troncature HDMR, $f_{i_1 i_2 i_3 \dots i_m}(x_{i_1}, x_{i_2}, x_{i_3}, \dots, x_{i_m})$ selon :

$$\begin{aligned} &f_{i_1 i_2 i_3 \dots i_m}(x_{i_1}, x_{i_2}, x_{i_3}, \dots, x_{i_m}) \\ &= r_l(x_I, \bar{x}^I) \\ &\approx \prod_{s=1}^m \frac{(x_{i_s} - \bar{x}_{i_s})}{(\bar{z}_{i_s} - \bar{x}_{i_s})} r_l(\bar{z}_I, \bar{x}^I) \\ &+ \sum_{s=1}^m \left[\prod_{r=1, r \neq i_s}^m \frac{(x_{i_r} - \bar{x}_{i_r})}{(\bar{z}_{i_r} - \bar{x}_{i_r})} r_l(x_{i_s}, \bar{z}_I^{i_s}, \bar{x}^I) - \prod_{r=1}^m \frac{(x_{i_r} - \bar{x}_{i_r})}{(\bar{z}_{i_r} - \bar{x}_{i_r})} r_l(\bar{z}_I, \bar{x}^I) \right] \\ &+ \sum_{1 \leq r < s \leq m} \left[\prod_{t=1, t \neq i_r, i_s}^m \frac{(x_{i_t} - \bar{x}_{i_t})}{(\bar{z}_{i_t} - \bar{x}_{i_t})} r_l(x_{i_r}, x_{i_s}, \bar{z}_{i_r i_s}, \bar{x}^I) \right. \\ &- \prod_{t=1, t \neq i_r}^m \frac{(x_{i_t} - \bar{x}_{i_t})}{(\bar{z}_{i_t} - \bar{x}_{i_t})} r_l(x_{i_r}, \bar{z}_{i_r}, \bar{x}^I) - \prod_{t=1, t \neq i_s}^m \frac{(x_{i_t} - \bar{x}_{i_t})}{(\bar{z}_{i_t} - \bar{x}_{i_t})} r_l(x_{i_s}, \bar{z}_{i_s}, \bar{x}^I) \\ &\left. + \prod_{s=1}^m \frac{(x_{i_t} - \bar{x}_{i_t})}{(\bar{z}_{i_t} - \bar{x}_{i_t})} r_l(\bar{z}_I, \bar{x}^I) \right]. \end{aligned} \quad (4.19)$$

où

- $\bar{z}_I^{i_t}$ (resp. $\bar{x}_I^{i_t}$) est le vecteur référence \bar{z}_I (resp. \bar{x}_I) sans l'élément \bar{z}_{i_t} (resp. \bar{x}_{i_t}),
- $\bar{z}_I^{i_r i_s}$ (resp. $\bar{x}_I^{i_r i_s}$) est le vecteur référence \bar{z}_I (resp. \bar{x}_I) sans les éléments \bar{z}_{i_r} et \bar{z}_{i_s} (resp. \bar{x}_{i_r} et \bar{x}_{i_s}),
- $r_l(x_I, \bar{x}^I)$ est la valeur du résidu calculé au point de référence \bar{x}_I sauf aux éléments de I ,

Pour le cas de $m=3$, et $I = (x_{i_1}, x_{i_2}, x_{i_3})$, les résidus sont exprimés selon :

$$\begin{aligned}
r_2(\bar{z}_I, \bar{x}^I) &= g(\bar{z}_{i_1}, \bar{z}_{i_2}, \bar{z}_{i_3}, \bar{x}^{i_1 i_2 i_3}) - g(\bar{z}_{i_1}, \bar{z}_{i_2}, \bar{x}^{i_1 i_2}) \\
&\quad - g(\bar{z}_{i_1}, \bar{z}_{i_3}, \bar{x}^{i_1 i_3}) - g(\bar{z}_{i_2}, \bar{z}_{i_3}, \bar{x}^{i_2 i_3}) \\
&\quad + g(\bar{z}_{i_1}, \bar{x}^{i_1}) + g(\bar{z}_{i_2}, \bar{x}^{i_2}) + g(\bar{z}_{i_3}, \bar{x}^{i_3}) - g(\bar{x})
\end{aligned} \tag{4.20}$$

$$\begin{aligned}
r_2(x_{i_1}, \bar{z}_I^{i_1}, \bar{x}^I) &= g(x_{i_1}, \bar{z}_{i_2}, \bar{z}_{i_3}, \bar{x}^{i_1 i_2 i_3}) - g(x_{i_1}, \bar{z}_{i_2}, \bar{x}^{i_1 i_2}) \\
&\quad - g(x_{i_1}, \bar{z}_{i_3}, \bar{x}^{i_1 i_3}) - g(\bar{z}_{i_2}, \bar{z}_{i_3}, \bar{x}^{i_2 i_3}) \\
&\quad + g(x_{i_1}, \bar{x}^{i_1}) + g(\bar{z}_{i_2}, \bar{x}^{i_2}) + g(\bar{z}_{i_3}, \bar{x}^{i_3}) - g(\bar{x})
\end{aligned} \tag{4.21}$$

$$\begin{aligned}
r_2(x_{i_1}, x_{i_2}, \bar{z}_I^{i_1 i_2}, \bar{x}^I) &= g(x_{i_1}, x_{i_2}, \bar{z}_{i_3}, \bar{x}^{i_1 i_2 i_3}) - g(x_{i_1}, x_{i_2}, \bar{x}^{i_1 i_2}) \\
&\quad - g(x_{i_1}, \bar{z}_{i_3}, \bar{x}^{i_1 i_3}) - g(x_{i_2}, \bar{z}_{i_3}, \bar{x}^{i_2 i_3}) \\
&\quad + g(x_{i_1}, \bar{x}^{i_1}) + g(x_{i_2}, \bar{x}^{i_2}) + g(x_{i_3}, \bar{x}^{i_3}) - g(\bar{x})
\end{aligned} \tag{4.22}$$

Le calcul des fonctions ordre supérieur de MPcut-HDMR ne fait intervenir que des fonctions de premier ordre $r_2(x_{i_1}, \bar{z}_I^{i_1}, \bar{x}^I)$ et de deuxième ordre $r_2(x_{i_1}, x_{i_2}, \bar{z}_I^{i_1 i_2}, \bar{x}^I)$ respectivement. Cette approche réduit notablement le nombre d'échantillonnages pour le calcul des fonctions d'ordre supérieur en comparaison de la méthode Cut-HDMR, et aussi les ressources mémoire nécessaires pour le calcul et le stockage de ces fonctions d'ordre supérieur. L'apport de gain en nombre de simulations et en mémoire sont représentés dans le tableau 4.2.

HDMR	Nombre de simulations exigées pour le 3ème ordre	Mémoire
MPcut-HDMR	$3C_n^3 * N_s + 3C_n^3 * N_s^2$	$(N * (N_s - 1) + \frac{N(N-1)(N_s-1)^2}{2}) * N_{out} * 4$
Cut-HDMR (3ème Ordre)	$C_n^3 * N_s^3$	$\frac{N(N-1)(N-2)(N_s-1)^3}{6} * N_{out} * 4$
Apport	$\frac{3 + C_n^3 * N_s}{N_s^2}$	$\frac{(N-2)(N_s-1)}{3}$
$N_s = 20$	Apport=6.35	$\approx 6.67 * (N-2) = 54$ (pour $N = 10$)

Table 4.2: Nombre de runs et mémoire demandés pour le calcul des fonctions de Cut-HDMR

4.2 Implémentation

Dans cette section, nous présentons l'implémentation de la méthode HDMR dans un modèle de cinétique chimique. Nous considérons un modèle, sous sa forme discrétisée donnée par l'équation suivante :

$$C_{n+1} = F(C_n, K_n) \quad (4.23)$$

avec :

- C_n le vecteur de concentration à l'instant t_n .
- F la fonction associée à la cinétique chimique permettant de passer du temps t au temps $t + \Delta t$.

En dehors des concentrations, elle dépend de plusieurs paramètres de forçage selon le schéma chimique qu'on utilise : température, humidité, angle zénithal,... L'ensemble des paramètres forme le vecteur K_n .

Notons que K_n n'est pas la donnée des paramètres aux temps t_n . Dans le cadre d'une discrétisation en temps par un schéma numérique d'intégration d'une équation aux dérivées partielles, K_n comprend les valeurs des paramètres aux instants utilisés par l'algorithme.

L'objectif est de remplacer ce modèle par un modèle réduit efficace basé sur l'utilisation spécifique des tables calculées par la méthode HDMR. Ce modèle appelé FEOM (pour "Fully Equivalent Operational model"), est utilisé pour calculer directement les sorties du modèle à partir de ses entrées en tant qu'intégrateur en substitution du modèle initial donné par l'équation (4.32).

L'utilisation de FEOM commence par

- la définition des variables d'entrée : concentration initiale des espèces, des conditions aux limites, et toute autre variable auxiliaire (température, pression,...),
- la définition des variables de sortie du modèle (concentrations d'espèces chimiques) qu'on cherche à calculer à l'instant $t + \Delta t$,

L'évaluation se fait ensuite par une simple lecture des tables (une fois au début de l'intégration) et une sommation des différentes fonctions composantes de HDMR selon l'équation (7.1) (voir la figure 4.5).

L'utilisation de HDMR pour la réduction d'un modèle de cinétique chimique, peut se faire selon l'algorithme suivant, que l'on explicite pour cut-HDMR :

- **Construction de la table :**

- Définir des variables d'entrée et les variables de sortie du modèle qu'on cherche à l'instant $t + \Delta t$.
- Choisir un intervalle I_i qui couvre l'évolution de chaque variable d'entrée i . Si nous ne disposons pas d'informations préalables sur la variabilités des paramètres d'entrée, on pourra les déduire à partir d'un calcul préliminaire du modèle d'origine.
- Choisir un nombre s_i de point de l'intervalle I_i pour chaque variable d'entrée i .
- Définir une valeur de référence \bar{x}_i dans l'intervalle I_i pour chaque variable d'entrée i , parmi les s_i points choisis.
- Calculer les fonctions d'ordre 0 (f_0) pour chaque variable de sortie.
- Calculer la série de fonctions d'ordre 1 $\{f_i\}$, de chaque point de variable d'entrée i , pour chaque variable de sortie.
- Sélectionner les variables d'entrée qui ont un effet, en agissant seul indépendamment des autres variables d'entrée, sur la sortie de modèle, pour chaque variable de sortie.
- Calculer la série de fonctions d'ordre 2 f_{ij} , de chaque couple de variables d'entrée i et j , pour chaque variable de sortie.
- Sélectionner les couples de variables d'entrée qui ont un effet sur la sortie de modèle, pour chaque variable de sortie.
- Calculer les coefficients d'interpolations linéaires (a et b) relatifs aux fonctions de premiers ordre f_i et bilinéaires (α , β , γ et δ) relatifs aux fonctions de second ordre f_{ij} .
- Sauvegarder les coefficients d'interpolations pour toutes les variables de sortie dans un fichier binaire.
- **Calcul avec FEOM :**
 - (a) Lire les tables de coefficients d'interpolations.
 - (b) Situer les différentes composantes du vecteur de variables d'entrée à l'instant t , dans l'espace des variables, et la sélection des coefficients d'interpolations appropriés.
 - (c) Calculer la sortie de modèle à l'instant $t + \Delta t$ par simple sommation des différentes fonctions de développement HDMR selon l'équation (7.1)

La rapidité de FEOM provient de deux facteurs :

- Les interpolations nécessaires au calculs des fonctions de développement de HDMR (7.1) sont très rapides et basées uniquement sur des opérations algébriques.

- Les pas de temps du modèle Δt peuvent être pris arbitrairement grands (et indépendamment du pas de temps du modèle initial contraint par la raideur). Par exemple, pour des applications à des CTM, le pas de temps Δt est typiquement celui qui caractérise le forçage météo (900 secondes) pour indication, à comparer au pas de temps d'intégration numérique de la cinétique chimique (100 secondes par exemple avec un schéma de type Rosenbrock pour des cinétiques raides).

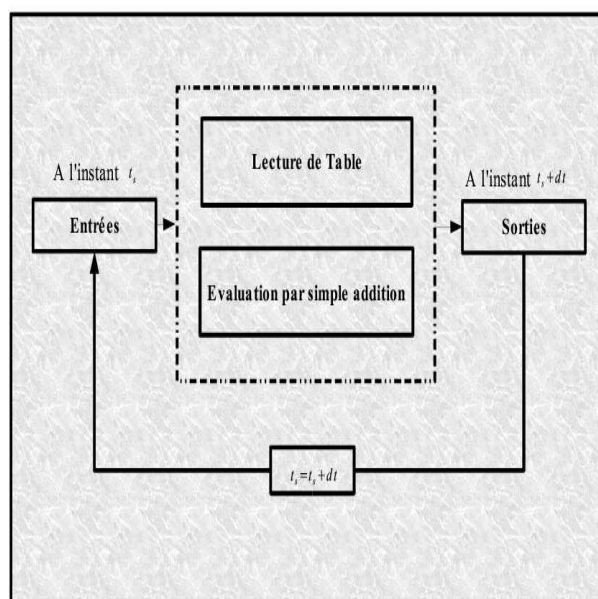


Figure 4.5: Schéma d'implémentation de FEOM dans un modèle

4.3 Modèles simples d'illustration

4.3.1 Premier exemple d'illustration

Notre première application, a été effectuée sur un modèle académique de 3 variables d'entrée x_1, x_2, x_3 , définies dans l'intervalle $[0,1]$, et une variable de sortie (non linéaire) définie selon :

$$Y = x_1 + x_2 + x_3 + x_2^2 + x_3^2 + x_2 * x_3 \quad (4.24)$$

l'objectif étant d'essayer d'approcher ce modèle par la méthode cut-HDMR présenté auparavant.

$$Y_{HDMR} = Y_0 + \sum_{i=1}^{i=n} f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) \quad (4.25)$$

où n est égal à 3 dans notre cas. Le point de référence étant pris au milieu de l'intervalle des paramètres d'entrée et le nombre de points qui composent l'échantillonnage est de $N_s = 20$.

Une fois la représentation Cut-HDMR calculée, nous avons exécuté le modèle réduit 10000 fois avec 3 variables d'entrée différentes tirées aléatoirement dans l'intervalle $[0,1]$ et nous avons calculé l'erreur relative (ER) entre le modèle réduit (Y_{HDMR}) et le modèle exact (Y).

$$ER = \left| \frac{Y_{HDMR} - Y}{Y} \right| * 100 \quad (4.26)$$

Dans le tableau 4.3 nous présentons le pourcentage de fois où $ER < 5\%$, $ER < 10\%$, $ER < 20\%$, $ER < 100\%$ et $ER > 100\%$ obtenus avec l'ordre de troncature ($r=1,2$). L'ordre 1 de HDMR fournit un résultat assez satisfaisant puisque 97% des simulations HDMR présentent une erreur relative inférieure à 20%, 79% des cas inférieurs à 5% et que 0.22% des simulations qui donnent des erreurs supérieures à 100%. Quant à l'ordre 2, il fournit dans 99% des cas des erreurs relatives au dessous de 5%.

Notons que comme on a un modèle quadratique, l'ordre 2 est suffisant bien sûr.

ER	1 ^{er} ordre	2 nd ordre
< 5%	79 %	99 %
< 10%	93 %	99 %
< 20%	97 %	100 %
< 100%	99.78 %	100 %
> 100%	0.22 %	0 %

Table 4.3: Pourcentage de réussite obtenu avec Cut-HDMR selon l'ordre de la troncature pour l'exemple 4.3.1

4.3.2 Deuxième exemple d'illustration

Supposons la réaction chimique suivante :



où X , Y , Z sont trois espèces chimiques. k est la constante de réaction qu'on peut définir par exemple sous la forme suivante :

$$k = e^{-1/T} \quad (4.28)$$

T peut être vu comme un terme de température dans le cas d'une vraie cinétique chimique. Si on s'intéresse à l'espèce Y , son équation d'évolution peut s'écrire selon :

$$\frac{dY}{dt} = -k.Y.X \quad (4.29)$$

ayant pour solution numérique par un schéma explicite :

$$Y(t + \delta t) = Y(t) - K(t).Y(t).X(t).\delta t$$

En supposant que T , X et Y varient entre 0 et 1, le pas de temps de la résolution est pris égal à $\delta t = \frac{1}{100}$. Nous avons appliqué la méthode Cut-HDMR à ce modèle qui représente d'une manière simplifiée le type d'équations différentielles qui gouvernent la cinétique chimique. Le point de référence est pris au milieu des intervalles de variation des paramètres d'entrée et le nombre de points qui composent l'échantillonnage est $N_s = 20$.

De la même manière que dans l'exemple précédent, après le calcul des fonctions représentatives de Cut-HDMR, nous avons exécuté le modèle réduit 10000 fois avec 3 variables d'entrée différentes tiré aléatoirement dans l'intervalle $[0,1]$ et nous avons calculé l'erreur relative (ER) définie selon (4.26) entre le modèle réduit et le modèle exact. Dans le tableau 4.4 nous présentons le pourcentage de fois où $ER < 5\%$, $ER < 10\%$, $ER < 20\%$, $ER < 100\%$ et $ER > 100\%$ obtenus avec l'ordre de troncature ($r=1,2,3$). Le premier ordre du développement HDMR fournit d'assez bons résultats puisque 62.6% des 10000 simulations effectuées présentent une erreur relative inférieure à 5% et 80% des simulations donnent des erreurs inférieures à 10%. Le deuxième ordre capture très bien le comportement du modèle de référence puisque le nombre de cas où l'erreur relative est inférieure à 5% est de plus de 85% des cas. Le troisième ordre présente un résultat excellent puisque les 10000 simulations présentent une erreur relative inférieure à 5%.

4.4 Application à la cinétique chimique

Dans cette section nous présentons les résultats d'application de cette méthode HDMR, en utilisant le modèle FEOM à la place du solveur Rosenbrock de la version 0D du modèle POLAIR.

ER	1 ^{er} ordre	2 nd ordre	3 ^{me} ordre
< 5%	62.68 %	85.65 %	100 %
< 10%	80.0 %	90.72 %	100 %
< 20%	88.56 %	94.40 %	100 %
< 100%	97.13%	98.69 %	100 %
> 100%	2.87 %	1.31 %	0 %

Table 4.4: Pourcentage de réussite obtenu avec Cut-HDMR selon l'ordre de la troncature pour l'exemple 4.3.2

4.4.1 Comparaison des profils de concentrations

Nous avons appliqué cette méthode à trois schémas chimiques différents :

- EMEP [7] avec 17 variables d'entrée dont dix correspondent aux concentrations d'espèces, trois aux concentrations d'espèces forcées, une à la température, une à l'humidité, une à la pluie, une à l'angle zénithal ; les 10 espèces du schéma correspondent aux sorties de modèle.
- CBM4 [68] avec 31 variables d'entrée qui correspondent aux concentrations des espèces du schéma et 31 variables de sorties Pour ce cas (et pour des raisons de disponibilité des données lors de ces tests), on a figé les conditions cinétiques.
- MOCA [69] avec 69 variables d'entrée dont 66 correspondent aux concentrations d'espèces chimique et les trois autres à la température, humidité et angle zénithal.

Les intervalles des variables d'entrée utilisés dans ces tests proviennent de simulations antérieures qui donnent une idée de la variation prévisible des valeurs d'entrée dans le modèle durant l'intégration. Le point de référence pour chaque variable d'entrée est pris au centre de l'intervalle de définition, le nombre de points qui constituent l'échantillonnage est pris à $N_s = 20$.

Les figures 4.6 et 4.7 présentent les résultats obtenus avec le modèle FEOM construit à base des tables HDMR et le modèle de référence POLAIR résolu avec le solveur numérique Rosenbrock respectivement pour le schéma EMEP et CBM-IV. Les profils de concentrations sont très semblables et l'erreur est minime.

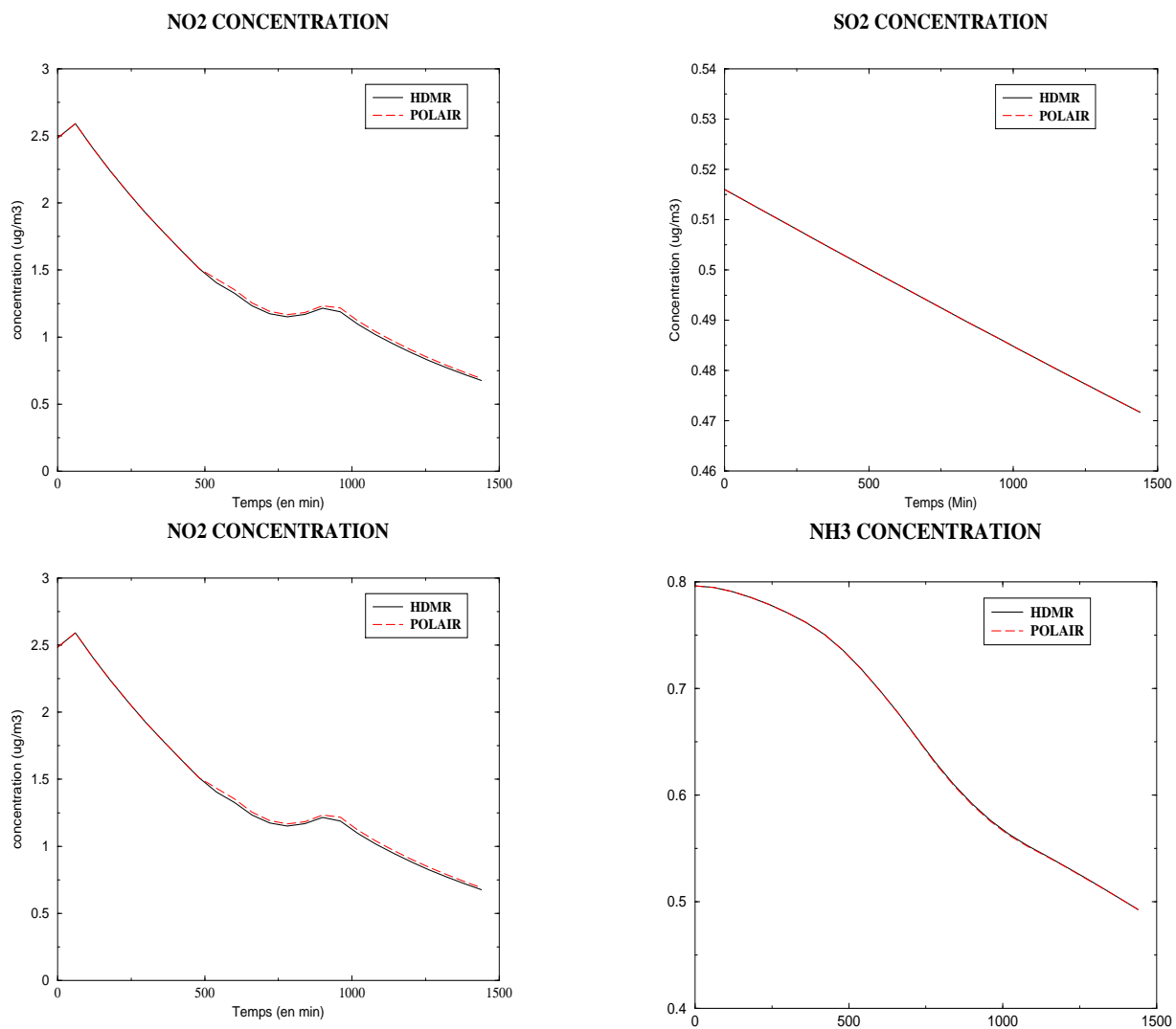


Figure 4.6: Comparaison des profils des 4 espèces du schéma EMEP. En (-) résultat obtenu avec le modèle d'origine POLAIR0D et en (ligne) résultat obtenu avec le modèle FEOM

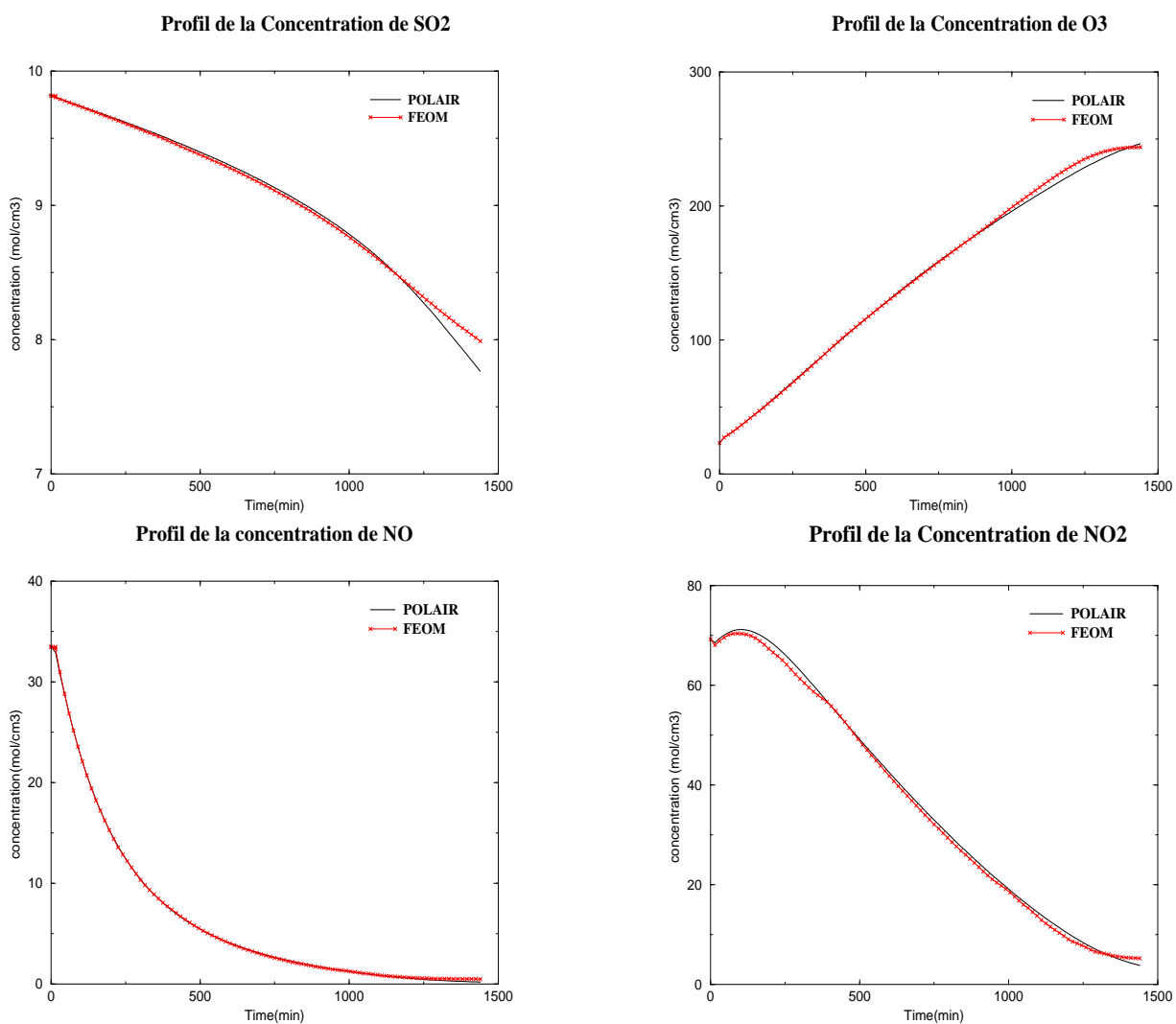


Figure 4.7: Comparaison des profils des 4 espèces du schéma CBM4. En (ligne) résultat obtenu avec le modèle d'origine POLAIR0D et en (-*-) résultat obtenu avec le modèle FEOM

	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-5}$
<hr/>		
Nbre de f_{ij}	8577	34803
<hr/>		

Table 4.5: Nombre de f_{ij} retenu en fonction du filtre ε (cas du schéma MOCA)

4.4.2 Filtrage des fonctions HDMR

Dans le cas d'un modèle de haute dimension, comme c'est le cas du mécanisme MOCA à 69 variables d'entrée, la construction et le stockage de la table nécessitent des places mémoires importantes. Pour remédier à ce problème, nous avons filtré les couples de $f_{ij}(x_i, x_j)$ qui n'ont pas d'effet sur la sortie, selon le critère suivant :

$$\text{Max}_{(x_i, x_j)} \left| \frac{f_{ij}(x_i, x_j)}{f_0 + f_i(x_i) + f_j(x_j)} \right| < \varepsilon \quad (4.30)$$

où $x_i = (x_1, x_2, \dots, x_{s_i})$ et $x_j = (x_1, x_2, \dots, x_{s_j})$, s_i et s_j désignent le nombre de points qui constituent l'échantillonnage des variables d'entrée x_i et x_j , ($s_i = s_j = 10$ à 20).

L'utilisation d'un tel filtre permet un gain en stockage de la table HDMR et en utilisation de place mémoire, sans perte de précision.

Dans notre application de HDMR au mécanisme MOCA, nous utilisons $n=69$ variables d'entrée, $N_{out} = 66$ variables de sorties et $N_s = 11$ points d'échantillonnage. Initialement le nombre total de couples $f_{ij}(x_i, x_j)$ nécessaires pour HDMR est de l'ordre de $\frac{n * (n - 1)}{2} * N_{out}$ qui vaut 154836. En utilisant le filtre ci dessus avec un ε de l'ordre de 10^{-3} ou 10^{-5} nous gardons un nombre très réduit de couples par rapport au nombre initial (tableau 4.4.2).

Le filtrage des fonctions de premier ordre $f_i(x_i)$ induit une perte dans la précision des résultats, de plus la place mémoire requise par ces fonctions n'est pas très coûteuse comparée à celle nécessaire pour le stockage des couples $f_{ij}(x_i, x_j)$. Pour ces raisons, dans le cas de MOCA, nous n'avons utilisé le filtre ci dessus que pour les composantes $f_{ij}(x_i, x_j)$ de HDMR, les résultats montrés dans la Figure 4.8 sont obtenus avec un filtre de $\varepsilon = 10^{-5}$.

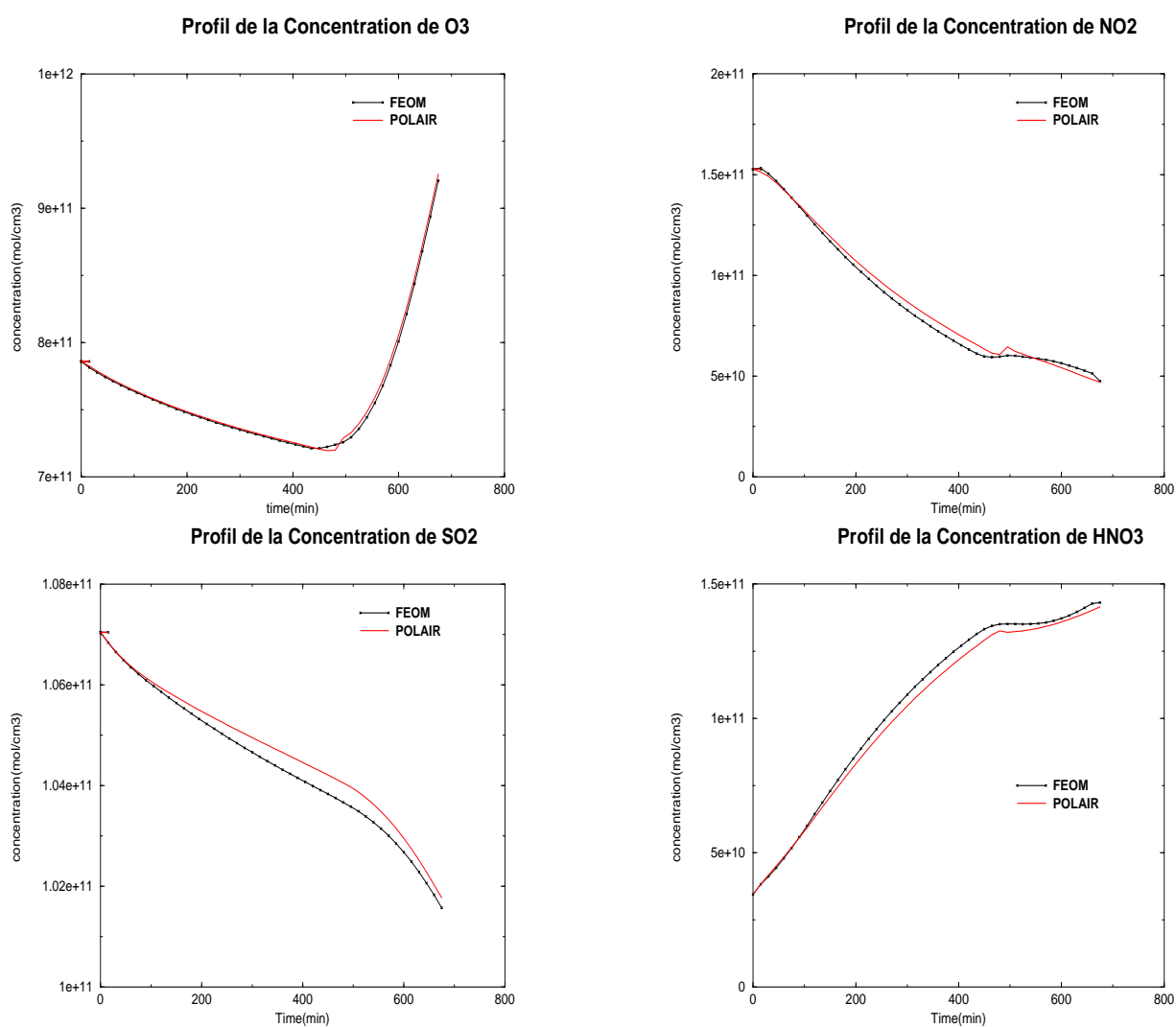


Figure 4.8: Comparaison des profils des 4 espèces du schéma MOCA, en (ligne) résultat obtenu avec le modèle d'origine POLAIR0D et en (-*-) résultat obtenu avec le modèle FEOM.

4.4.3 Test d'efficacité et coût de calcul

Un autre test d'efficacité de la méthode HDMR a été effectué, en utilisant l'approche dans le contexte d'analyse d'incertitude. Nous propageons l'incertitude supposée sur les variables d'entrée à travers le modèle réduit construit par HDMR, en développant les fonctions de probabilité (PDF) des variables de sortie du modèle.

Les variables d'entrée sont supposées varier selon une loi uniforme à l'intérieur des intervalles représentatifs de chacune des variables. Cette approche est comparée à la méthode classique de propagation d'incertitude de Monte Carlo qui utilise un grand nombre d'exécutions du modèle à partir des variables d'entrée tirées aléatoirement selon leurs distributions.

Dans notre cas les distributions sont supposées uniformes.

Le test est effectué sur le mécanisme chimique CBM IV avec 31 variables d'entrée qui sont composées des concentrations des espèces chimiques. La comparaison est faite avec un nombre d'échantillons de variables d'entrée $N=10000$. POLAIR est exécuté avec un pas de temps de 600s et sur une itération.

Nous comparons le résultat obtenu avec le modèle HDMR de premier ordre. Le nombre de points d'échantillonnage qui ont servi à construire la table HDMR est de 15 points.

Les espèces de sorties auxquelles nous nous sommes intéressés sont O_3 , NO_2 , NO , SO_2 et CO . Les résultats sont présentés dans la figure 4.9. Les deux PDF de sorties (avec le modèle réduit HDMR et le modèle POLAIR) pour chacune des espèces chimiques obtenues, sont très proches.

Les simulations sont effectuées sur un PC avec un processeur Pentium PIV de 2.8 Gega Hz et un Gega-octet de mémoire. Le temps CPU demandé par la méthode FEOM est de 0.19 secondes pour les 10000 simulations, tandis que la simulation utilisant POLAIR de référence a nécessité 1.84 s, soit un speedup de l'ordre de 10.

Remarque : coût calcul dans un cas 3D

Dans le cadre du modèle chimie-transport POLAIR3D où la chimie est la plus coûteuse en temps calcul, nous utilisons la méthode HDMR afin de réduire le temps CPU nécessité par la chimie. Ce processus est très raide et exige des pas de temps Δt_{chem} plus petit par rapport aux pas de temps des autres processus que nous notons Δt_{trans} (sous-cyclage).

Après identification des paramètres d'entrée du module chimie et leur variations,

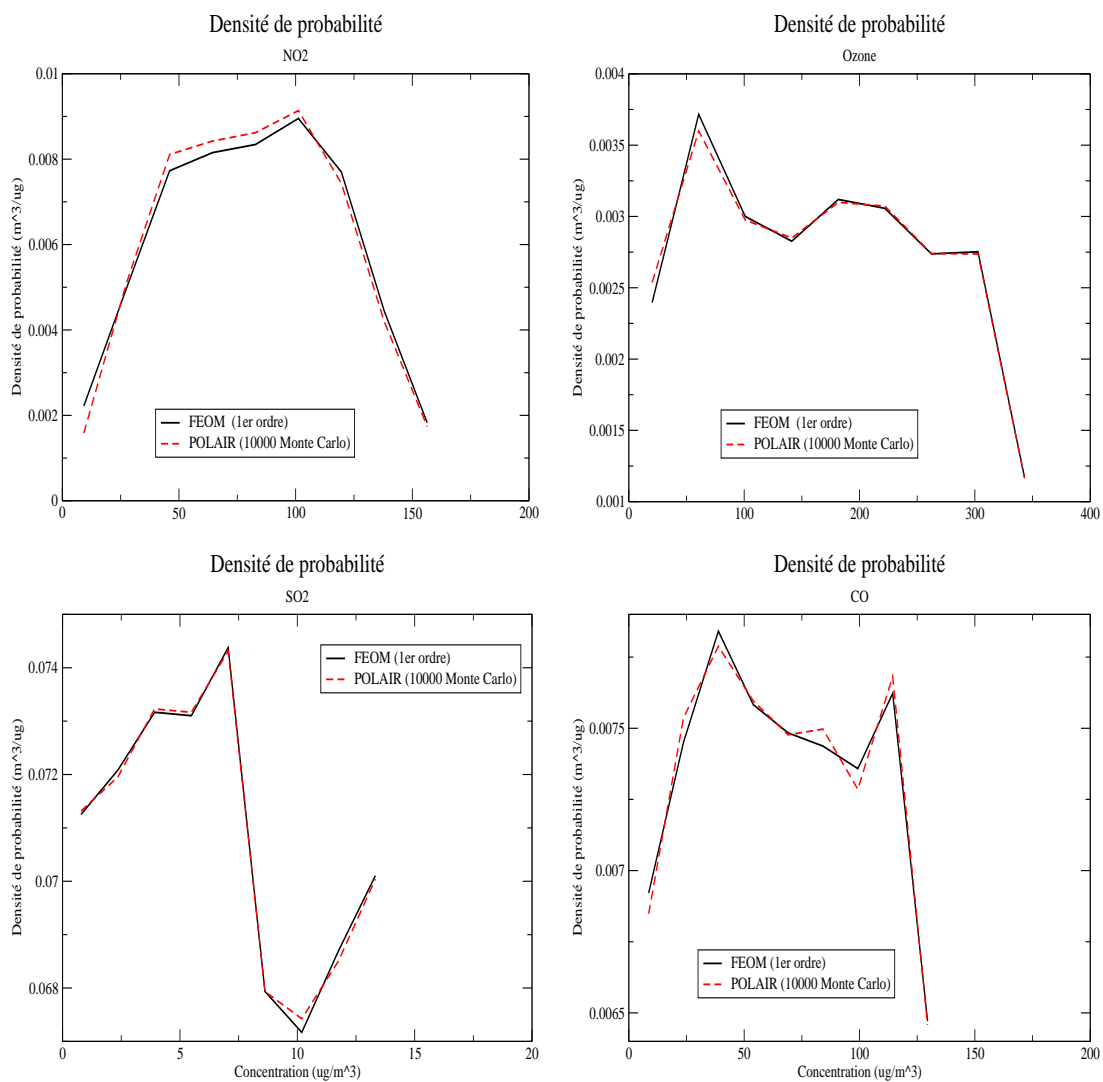


Figure 4.9: Comparaison des distributions de probabilités, en (ligne) résultat obtenu avec le modèle d'origine Monte Carlo et en (—) résultat obtenu avec le modèle FEOM/HDMR (1er ordre)

Pas de temps de la Chimie	FEOM (1er ordre)	FEOM (2ème ordre)	Pas de temps de la Chimie	FEOM (1er ordre)	FEOM (2ème ordre)
100s	180	18	100s	360	36
200s	90	9	200s	180	18
300s	60	6	300s	120	12
600s	30	3	600s	60	6
900s	20	2	900s	40	4

(a) (b)

Table 4.6: Speed-up du modèle FEOM, le pas de temps transport est supposé de 1800s : (a) cas où le modèle utilise un splitting de premier ordre, (b) cas où le modèle utilise un splitting de second ordre

les tables HDMR sont précalculées une fois pour toute, et stockées pour une utilisation ultérieure par le modèle FEOM qui remplacera le module chimie. Dans les cas où $\Delta t_{chem} \ll \Delta t_{trans}$, on peut procéder à un sous cyclage du module de chimie en prenant pour le modèle FEOM le pas de temps du transport Δt_{trans} . Le speedup S est donc défini selon :

$$\begin{aligned}
 S &= \frac{CPU_{chimie}}{CPU_{feom}} \\
 &= \frac{\Delta t_{trans}}{\Delta t_{chem}} * \frac{CPU_{chimie}}{CPU_{hdmr}}
 \end{aligned}
 \tag{4.31}$$

où CPU_{chimie} est le coût calcul (CPU) demandé par une itération chimie et CPU_{hdmr} l'équivalent demandé par une itération de calcul FEOM. Pour des raisons de sous-cyclage, $CPU_{hdmr} = \frac{\Delta t_{chem}}{\Delta t_{trans}} * CPU_{feom}$.

Nous rappelons que le modèle FEOM est basé uniquement sur des opérations arithmétiques simples. Le coût de calcul dépend du nombre de sorties, du nombre de paramètres d'entrée retenus et du coût d'extraction des fonctions $\{f_{ij}\}$ des tables HDMR.

A l'échelle continentale et pour des simulations de longue durée, le pas de temps du transport peut aller jusqu'à 3600s, or la chimie nécessite des pas de temps très faibles de 100s, 200s dans les meilleurs des cas. En utilisant le modèle FEOM avec un sous-cyclage pendant la durée du transport, un gain en temps de calcul assez important peut être attendu (tableau 4.6).

4.5 Sensibilité de Cut-HDMR au point de référence et au nombre d'échantillonnage

Lors de l'utilisation de la méthode cut-HDMR, pour calculer les fonctions de développement HDMR f_0 , $\{f_i\}_{i=1}^n$, $\{f_{ij}\}_{1 \leq i < j \leq n}, \dots$, on définit pour chaque variable d'entrée x_i un intervalle de variation $I_i = [x_i^s, x_i^f]$ discrétisé selon N_s points et un point de référence \bar{x}_i qui représente la valeur moyenne de la variable x_i .

Nous étudions ici la sensibilité du résultat de la méthode cut-HDMR :

- au choix du point de référence \bar{x}_i : nous prenons 4 valeurs qui correspondent au premier quart de l'intervalle $(3x_i^s + x_i^f)/4$, au point milieu de l'intervalle $(x_i^s + x_i^f)/2$ et au dernier quart de l'intervalle I_i $(x_i^s + 3x_i^f)/4$. Le nombre de points d'échantillonnage N_s est pris égal à 20 points.
- au nombre N_s de points qui constituent l'échantillonnage, de l'intervalle I_i . Nous avons testé les valeurs 5, 10, 15, 20, 25 et 30 points pour la variable N_s . Les points de référence sont pris aux points milieux des intervalles.

L'étude est effectuée avec le modèle POLAIR et le schéma chimique utilisé dans cette étude est CBM IV ; un pas de temps de 900 secondes, pendant un jour de simulation, est utilisé.

Dans la figure 4.10 nous présentons les différents résultats obtenus avec les différentes tables HDMR calculées avec les 3 points de références que nous venons de définir, comparé au modèle de référence POLAIR. Nous remarquons que le point de référence correspondant à $(x_i^s + 3x_i^f)/4$ fournit des fonctions de développement qui ne représentent pas assez bien le comportement du modèle POLAIR, puisque l'erreur relative est parfois supérieur à 100%, c'est le cas de l'espèce NO_2 et NO_3 , quant aux autres points de référence ils permettent d'avoir des résultats meilleurs. Cela pourrait s'expliquer par le choix de l'intervalle I qui est très grand par rapport à la variation exacte des paramètres d'entrée.

Dans la figure 4.11 sont représentés les résultats des différents modèles réduits associés aux table de HDMR calculées avec les 6 nombres d'échantillonnage qui composent les intervalles de variations des paramètres d'entrée (5, 10, 15, 20, 25, 30), pour les espèces NO_2 , O_3 , NO et NO_3 . A part le modèle correspondant à $N_s = 5$, les résultats sont très semblables, ce qui permet d'envisager des nombres N_s relativement faibles (typiquement 10) pour gagner du temps CPU dans le calcul des tables HDMR.

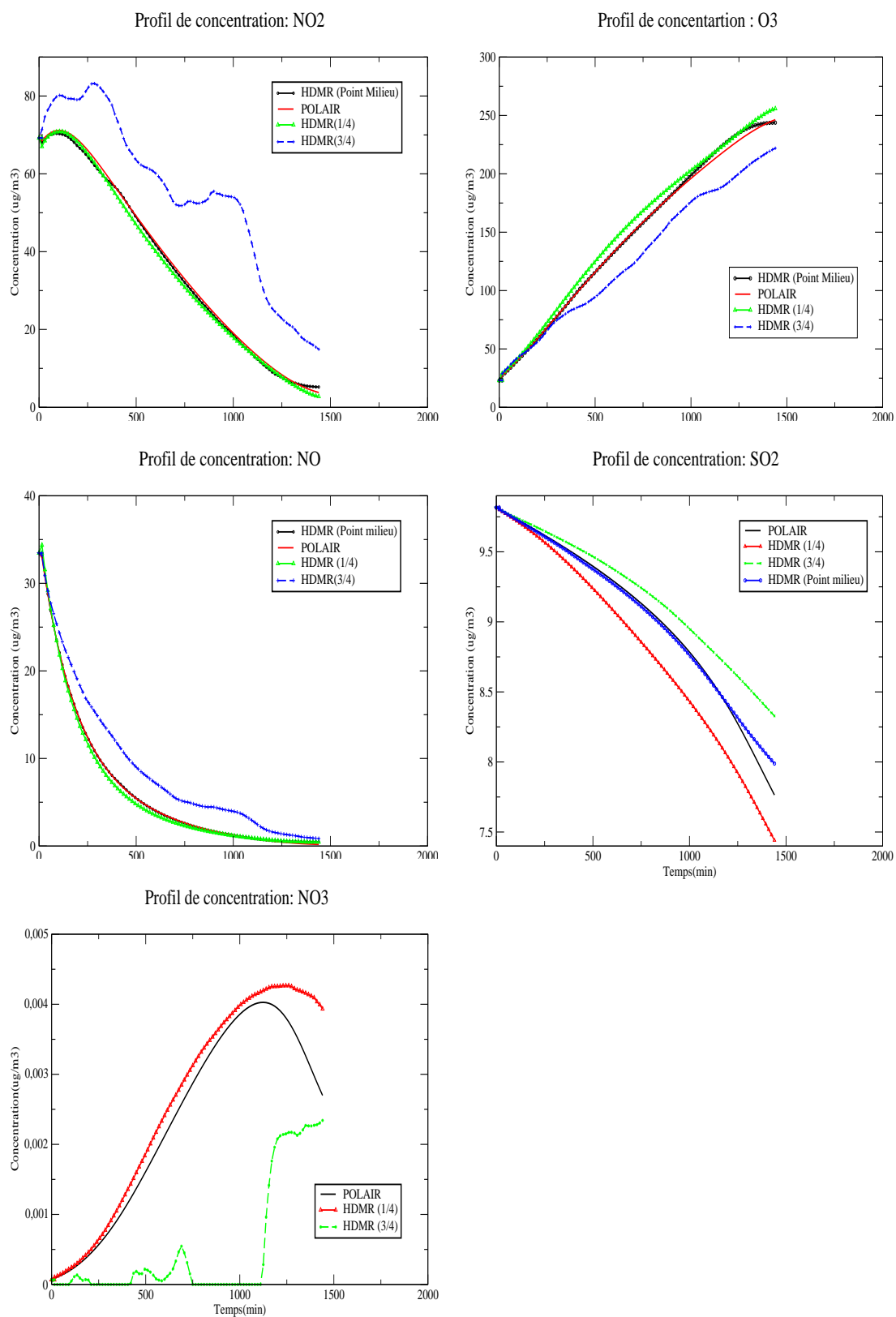


Figure 4.10: Sensibilité de Cut-HDMR au choix du point de référence ; le nombre de points d'échantillonnage est 20 points.

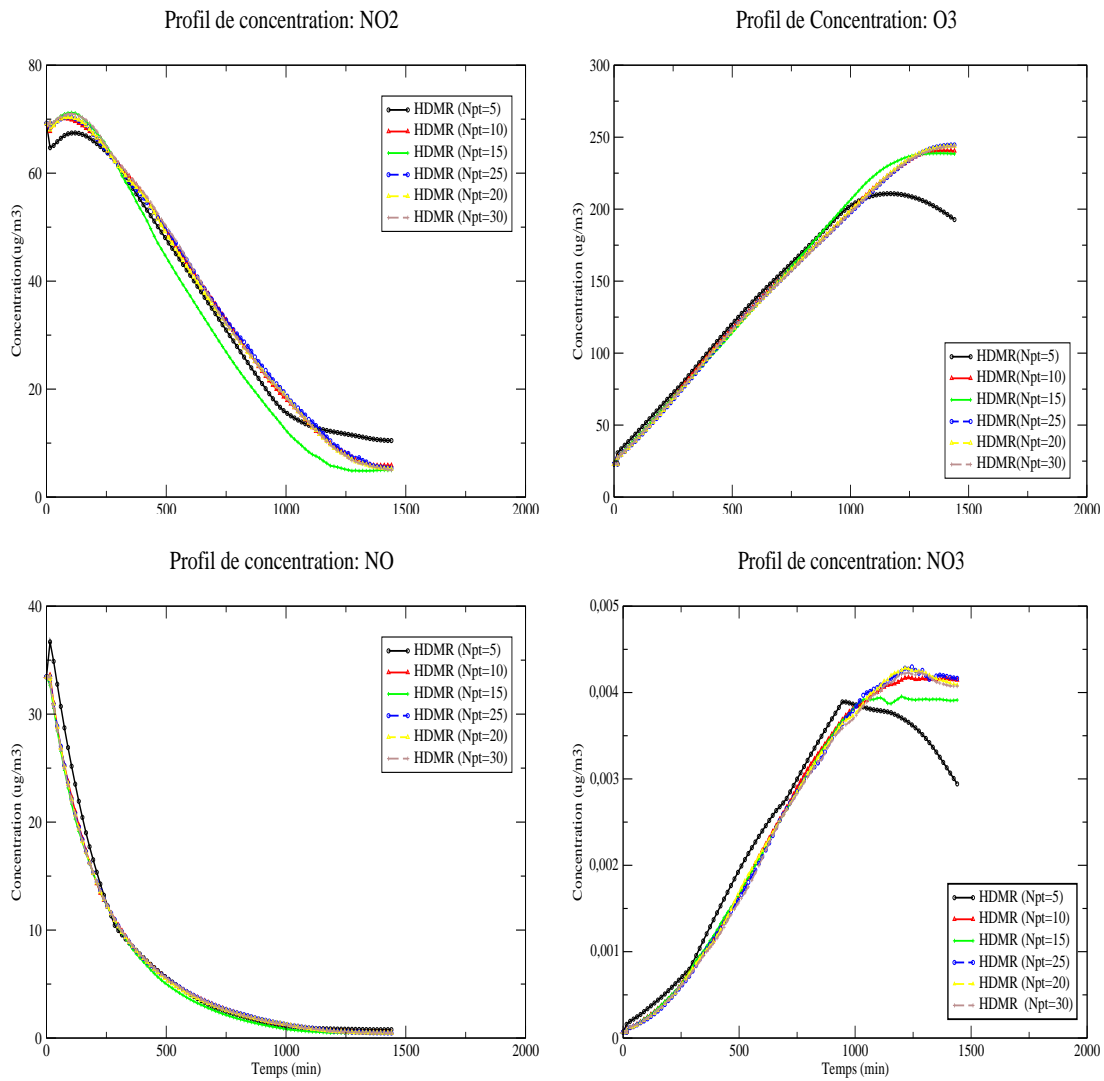


Figure 4.11: Sensibilité de Cut-HDMR au nombre de points d'échantillonnage N_s , les point de référence sont les centres d'intervalles de variations des paramètres d'entrée.

4.6 Application de HDMR au modèle POLAIR3D

Après les bons résultats obtenus en utilisation de HDMR en 0D avec différentes chimies (sections 4.3 et 4.4), nous avons étudié la faisabilité de son implémentation dans le contexte des modèles 3D de Chimie-Transport en tant que substitut au solveur de la cinétique chimique, avec une application au modèle POLAIR3D pour une chimie d'acidification EMEP sur l'Europe.

4.6.1 Le modèle FEOM (Fully Equivalent Operationel Model)

Nous considérons un modèle de cinétique chimique, sous sa forme discrétisée donnée par l'équation suivante :

$$C_{n+1} = F(C_n, K_n) \quad (4.32)$$

avec :

- C_n le vecteur de concentration à l'instant t_n .
- F la fonction associée à la cinétique chimique permettant de passer du temps t au temps $t + \Delta t$.

Après la construction des tables HDMR suivant l'algorithme explicité dans le chapitre 4, l'utilisation de FEOM se fait par une simple lecture des tables (une fois au début de l'intégration) et une sommation des différentes fonctions composantes de HDMR selon l'équation (7.1)).

L'implémentation du modèle FEOM dans un modèle de chimie transport est alors schématisée dans la figure 4.12. Nous utilisons la méthode cut-HDMR, en utilisant le modèle FEOM à la place du solveur Rosenbrock pour le module chimique du modèle POLAIR3D (figure 4.12).

Le modèle utilisé est POLAIR3D ([40], chapitres 2 et 3), modèle 3D eulérien de chimie-transport. Le modèle résout l'équation déjà décrite :

$$\frac{\partial c_i}{\partial t} = L_{adv}(c_i) + L_{diff}(c_i) + [L_{chem}(c)_i + S_i] \quad (4.33)$$

Le calcul est effectué sur le domaine européen, avec une résolution horizontale de $150\text{km} \times 150\text{km}$ (47×32 points de grilles horizontales) et 14 niveaux verticaux (le dernier niveau est à 5000 mètres de hauteur).

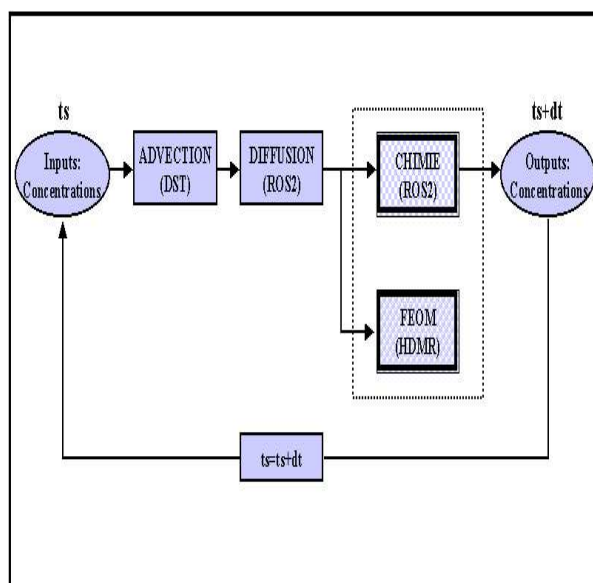


Figure 4.12: Schéma d'implémentation de FEOM dans un modèle

Le schéma chimique utilisé dans cette application 3D est celui de l'acidification d'EMEP avec 10 espèces et 3 espèces forcées.

EMEP comporte deux parties : une partie linéaire (voir mécanisme associé dans l'annexe B) et l'autre hétérogène. La deuxième partie n'est pas traitée par le solveur Ros2 et ne sera donc pas tabulée par la méthode HDMR.

Le modèle est résolu avec un pas de temps modèle Δt de 900s pour une durée de 1 jour (96 itérations). L'objectif est de remplacer la module de chimie qui coûte cher en temps de calcul, par le solveur FEOM basée sur les interpolations à partir des tables HDMR (figures 4.4 et 4.12).

4.6.2 Variables d'entrée/sortie

Afin de reproduire plus précisément le système chimique, les variables d'entrée qui décrivent les processus les plus important doivent être pris en compte. Le schéma chimique EMEP dépend essentiellement de :

- l'angle zénital solaire θ calculé selon :

$$\cos(\theta) = \sin(\phi) \sin(\delta) + \cos(\phi) \cos(\delta) \cos(t_h) \quad (4.34)$$

où : ϕ est la latitude, δ est l'angle de déclinaison solaire et t_h est l'heure solaire. La déclinaison du soleil est donnée par :

$$\begin{aligned} \delta(\text{radian}) = & 0.006918 - 0.399912 \cos(t) + 0.070257 \sin(t) - 0.006758 \cos(2t) \\ & + 0.000907 \sin(2t) - 0.002697 \cos(3t) + 0.001480 \sin(3t) \end{aligned} \quad (4.35)$$

où t est le temps dans l'année en seconde.

- la température Temp en ($^{\circ}\text{K}$),
- l'humidité Hum (%),
- le taux de pluie (en mm),
- les concentrations forcées de O_3 , OH et de ROH nécessaires pour le schéma EMEP,
- les concentrations initiales des 7 espèces chimiques : NO, NO_2 , HNO_3 , NO_3 , SO_2 , SO_4 , NH_3 ,
- le temps en jour, représenté dans la constante cinétique :
 $k_9 = 310^{-6} + 210^{-6} \sin(2\pi(\frac{\tau - 80}{365}))$, où τ est le temps en jour $\tau = \frac{t}{86400}$, t en seconde.

Les variables de sortie sont les concentrations des 7 espèces chimiques que nous faisons évoluer par le solveur ROS2 : NO, NO_2 , HNO_3 , NO_3 , SO_2 , SO_4 , NH_3 .

4.6.3 Résultats

4.6.3.1 FEOM global

Dans un premier temps, nous avons essayé de représenter la relation input/output avec le même FEOM sur tout le domaine à l'échelle européenne. Ainsi la construction des tables HDMR a été faite avec le modèle POLAIR0D, module de cinétique chimique, avec un pas de temps de 900 secondes, avec 15 variables d'entrée chacune étant définie dans un intervalle et discrétisée selon un nombre d'échantillonnage $s = 20$. Le point de référence correspond au centre de l'intervalle de chaque variable d'entrée (tableau 4.7).

Après une intégration du modèle POLAIR3D avec le module FEOM comme substitut au module de cinétique chimique, nous avons obtenu des résultats très insuffisants, comme on peut le constater sur les figures 4.13 et 4.14 qui représentent la comparaison des champs de concentrations NO_2 au premier niveau de modèle prévu par FEOM et par POLAIR3D, respectivement après 12h et 24h d'intégration. Le résultat est aussi médiocre sur les autres niveaux verticaux comme on peut le voir sur les figures 4.15 et 4.16 qui schématisent la moyenne verticale de NO_2 par

Variable d'entrée	Limite inférieure	Limite supérieure	Point référence
$\cos(\theta)$	-1	0.45	-0.275
Temp(°k)	230	292	261
Humidité (%)	1.	100	50.5
pluie(mm)	0.	1.	0.5
O_3 ($\mu g/m^3$)	60	83	71.5
OH (E-07) ($\mu g/m^3$)	2.30284826	3.12443319	2.713640725
ROH (E-05) ($\mu g/m^3$)	4.10552311	0.00040365	2.05296338
NO ($\mu g/m^3$)	0	20	10
NO2 ($\mu g/m^3$)	0	231	115.5
HNO_3 ($\mu g/m^3$)	0	64	32
NO_3 ($\mu g/m^3$)	0	83	41.5
SO_2 ($\mu g/m^3$)	0	418	209
SO_4 ($\mu g/m^3$)	0	960	480
NH_3 ($\mu g/m^3$)	0	26	13
k_9 (E-06)	1.31876319	1.32637974	1.32200582

Table 4.7: Variation des 15 variables d'entrée de HDMR.

FEOM et POLAIR3D, respectivement après 12h et 24h d'intégration.

L'ordre 2 n'est donc pas suffisant pour capturer la variation dans tout le domaine et le choix de point de référence devient très important.

NO₂ (LEVEL =1) POLAIR3D

NO₂ (LEVEL =1) HDMR/FEOM

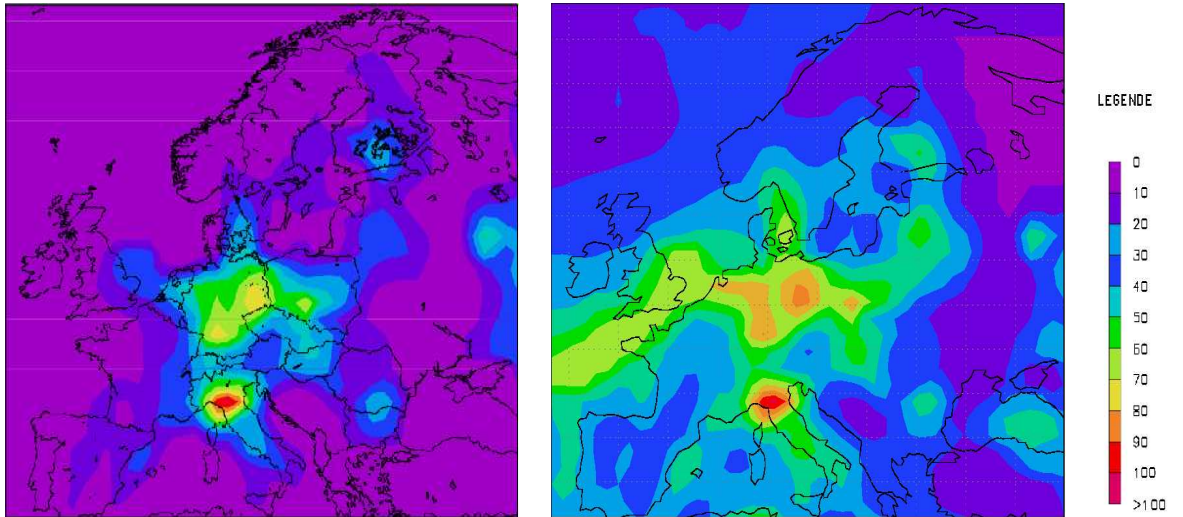


Figure 4.13: FEOM-Global : Comparaison de NO_2 (au 1er niveau), après 12h d'intégration par FEOM et POLAIR3D en $\mu g/m^3$

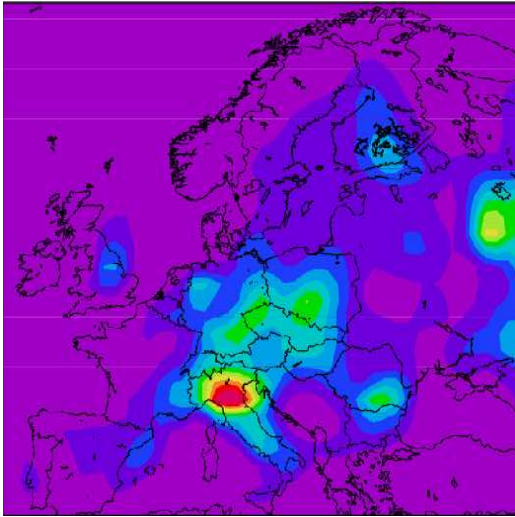
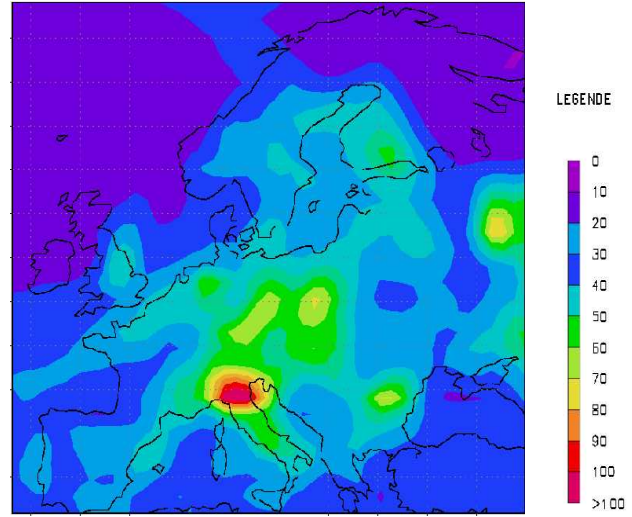
NO₂ (LEVEL =1) POLAIR3DNO₂ (LEVEL =1) HDMR/FEOM

Figure 4.14: FEOM-Global : Comparaison de NO_2 (au 1er niveau), après 24h d'intégration par FEOM et POLAIR3D en $\mu g/m^3$

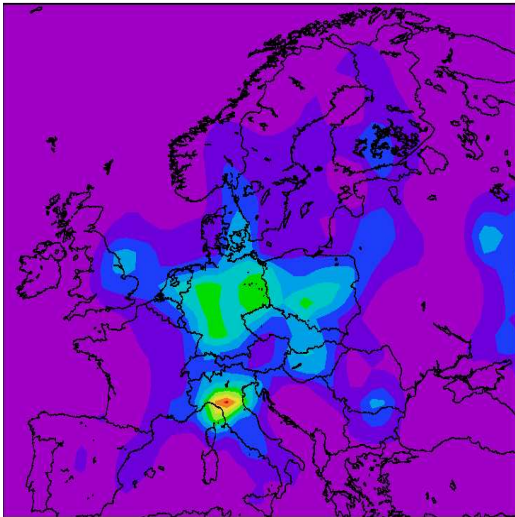
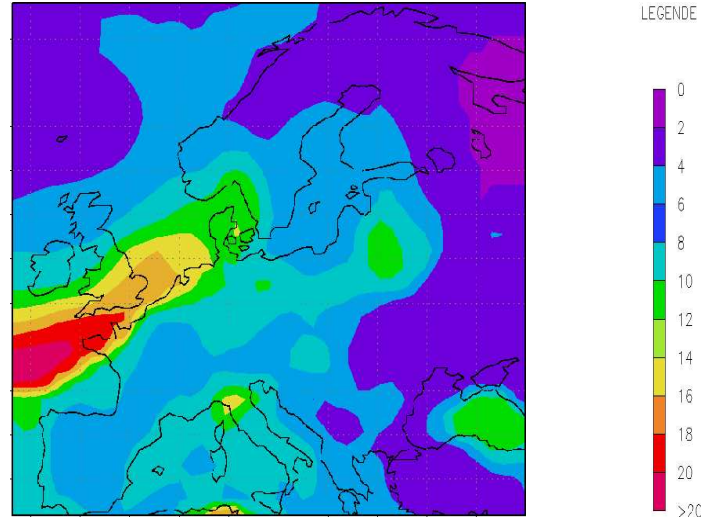
NO₂ (Vertical Average) POLAIR3DNO₂ (Vertical Average) HDMR/FEOM

Figure 4.15: FEOM-Global : Comparaison de NO_2 (en moyenne verticale), après 12h d'intégration par FEOM et POLAIR3D en $\mu g/m^3$

NO₂ (Vertical Average) POLAIR3D NO₂ (Vertical Average) HDMR/FEOM

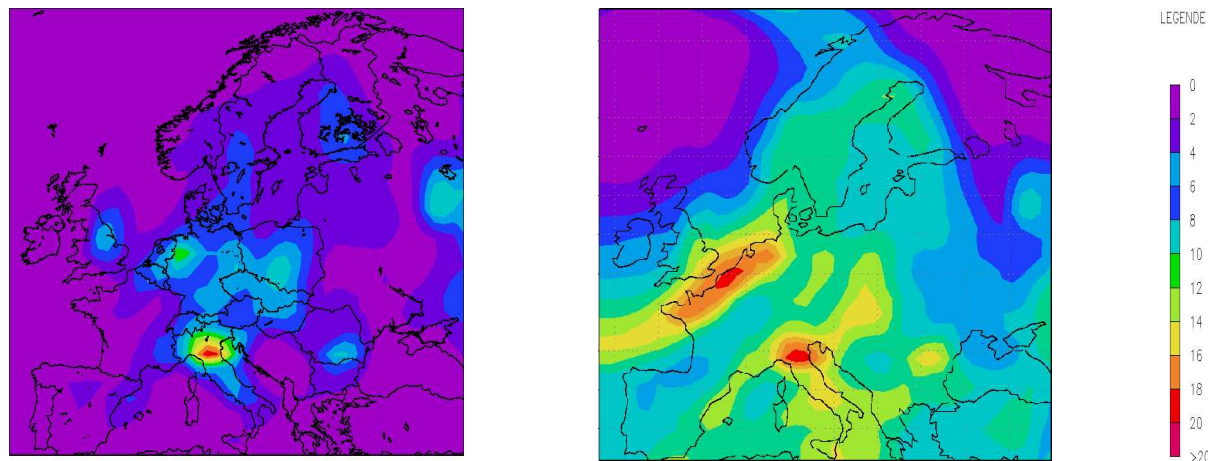


Figure 4.16: FEOM-Global : Comparaison de NO_2 (en moyenne verticale), après 24h d'intégration par FEOM et POLAIR3D en $\mu g/m^3$

4.6.3.2 Multiple-FEOM

L'utilisation d'un seul FEOM sur tout le domaine de calcul POLAIR3D avec les mêmes intervalles de variation des variables d'entrée (tableau 4.7) est très limitée comme en témoignent les résultats obtenus ci-dessus. Cela est vraisemblablement dû aux fortes variations des variables d'entrées (température, humidité, concentrations d'espèces chimiques) selon la verticale et les directions zonale et méridionale. On peut se référer à [21, 70] pour des constatations de même type pour l'application de la méthode HDMR à la simulation de l'ozone avec un modèle global de Chimie-Transport.

Pour remédier à ce problème, nous avons découpé le domaine en plusieurs sous-domaines avec la même chimie dans chacune des régions. Pour chaque sous-domaine nous avons construit un modèle FEOM spécifique. Notons que ceci est assez logique sur un plan physique puisque cela correspond à l'existence (avérée) de plusieurs régimes chimiques dans l'atmosphère.

Nous avons effectué un découpage en 12 sous-domaines : un découpage vertical en 3 régions, un découpage en 2 selon l'axe de la longitude et en 2 selon l'axe de la latitude (figure 4.17). Il est à noter que ce découpage est relativement arbitraire dans notre expérience même s'il est clair, par exemple que les régimes chimiques sont différents selon que l'on se trouve proche du sol ou en altitude.

Les paramètres d'entrée pour chacune des régions sont détaillés dans le tableau 4.8. Avec cette méthode de découpage en 12 sous domaines nous avons obtenu les résultats suivants :

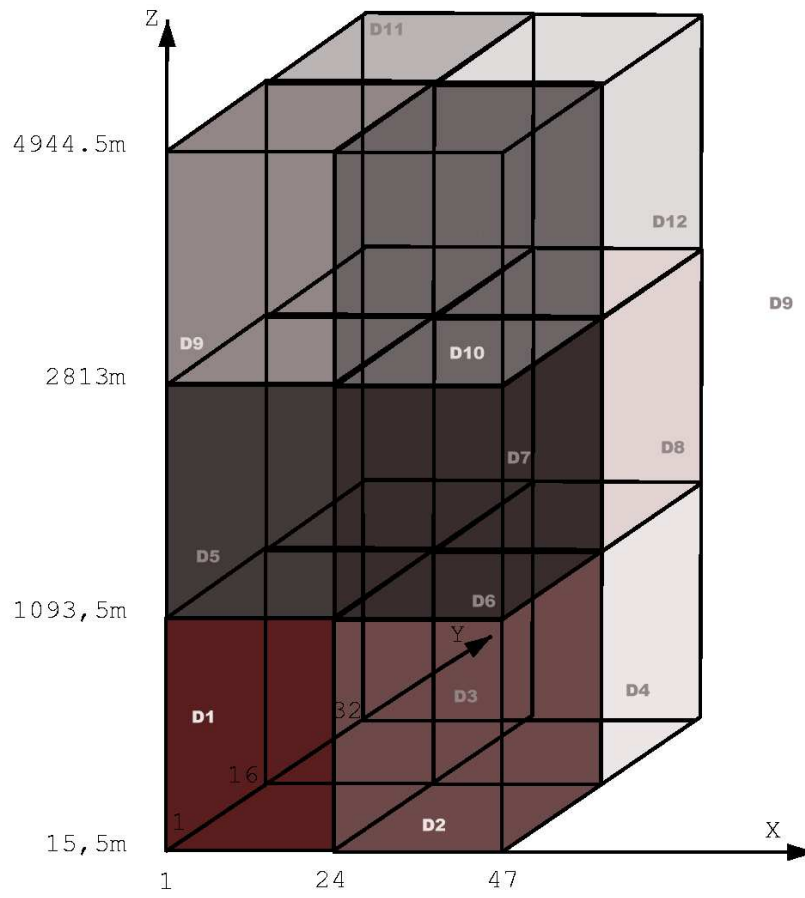


Figure 4.17: Schéma de découpage en sous-domaines

- figure 4.18 : une comparaison entre les champs de NO_2 au premier niveau du modèle, calculés par FEOM et par POLAIR3D, après 12h d'intégration.
- figure 4.19 : une comparaison entre les champs de moyenne verticale de NO_2 , calculés par FEOM et par POLAIR3D, après 12h d'intégration.
- figure 4.20 : une comparaison entre les champs de NO_2 au premier niveau du modèle, calculés par FEOM et par POLAIR3D, après 24h d'intégration.
- figure 4.21 : une comparaison entre les champs de moyenne verticale de NO_2 , calculés par FEOM et par POLAIR3D, après 24h d'intégration.

Comme on peut l'observer sur les différentes figures, les résultats obtenus avec FEOM sont très proches de ceux du modèle POLAIR3D et nettement meilleurs que ceux obtenus avec un FEOM global sur tout le domaine.

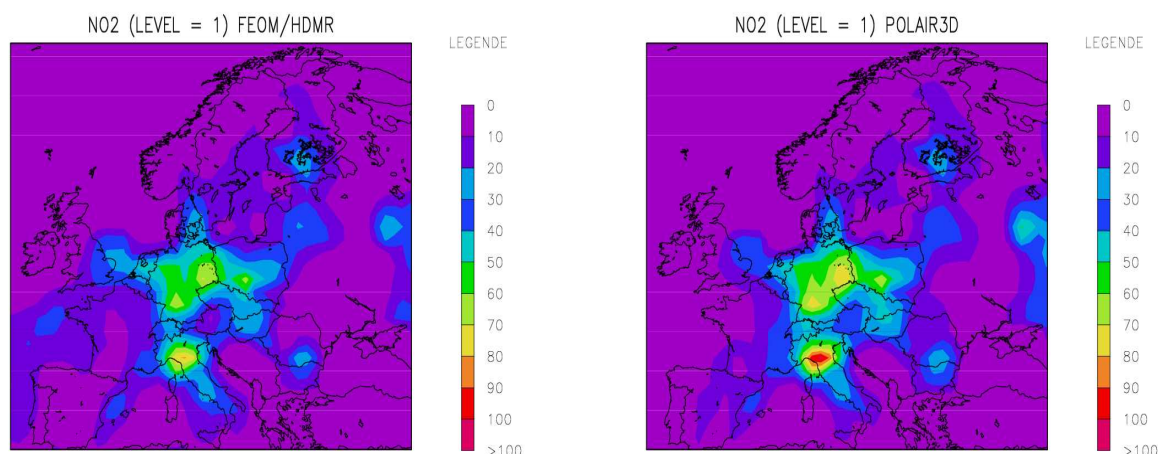


Figure 4.18: Multi-FEOM : Comparaison de NO_2 (au 1er niveau), après 12h d'intégration par FEOM et POLAIR3D en $\mu g/m^3$

Entrées	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}	D_{11}	D_{12}
$\cos(\theta)$	-1.	-0.8	-1.	-0.8	-1.	-0.8	-1.	-0.8	-1.	-0.8	-1.	-0.8
	0.	-0.1	0.	-0.1	0.	-0.1	0.	-0.1	0.	-0.1	0.	-0.1
T(°K)	262.	264.	259.	245.	256.	252.	246.	250.	243.	236.	234.	235.
	290.	284.	288.	279.	281.	274.	277.	274.	271.	263.	268.	263.
Hum(%)	4.	35.	16.	29.	3.	5.	5.	2.	1.	2.	1.	2.
	99.	99.	98.	100.	95.	92.	94.	97.	91.	89.	87.	88.
Pluie (mm)	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	0.8	0.7	0.2	0.3	0.8	0.7	0.2	0.3	0.8	0.7	0.2	0.3
O_3 (μ/m^3)	74.	62.	74.	62.	74.	62.	74.	62.	74.	62.	74.	62.
	82.	73.	82.	73.	82.	73.	82.	73.	82.	73.	82.	73.
OH (E-05)(μ/m^3)	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	1.6.	1.	3.	1.7	1.6	1.	3.	1.7	1.6	1.	3.	1.7
ROH (E-04)(μ/m^3)	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	2.1	1.8	3.6	2.7	2.2	1.8	3.6	2.7	2.2	1.8	3.6	2.7
NO (μ/m^3)	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	9.	2.5	4.5	2.	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1
NO_2 (μ/m^3)	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	230.1	52.4	150.7	69.3	13.	11.1	3.9	8.2	5.8	6.7	3.2	5.7
HNO_3 (μ/m^3)	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	59.	2.	58.	6.	36.	1.	21.5	6.	11.5	0.8	10.7	4.6
NO_3 (μ/m^3)	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	111.	13.5	77.6	52.	5.	5.4	13.	13.7	3.	4.	12.	9.7
SO_2 (μ/m^3)	7.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	416.	102.4	451.	113.	74.	24.7	54.	24.6	23.	14.5	35.	16.5
SO_4 (μ/m^3)	2.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	747.	137.	956.	159.7	43.4	17.3	68.4	26.	15.7	8.7	49.4	11.3
NH_3 (μ/m^3)	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
	1.	0.2	23.	1.	1.5	0.2	7.2	1.9	2.	0.2	6.	2.
K_9 (E-06)	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33

Table 4.8: Variation des 15 variables d'entrée de HDMR. Pour chaque variable d'entrée, la première ligne correspond à la valeur minimale et la deuxième à la valeur maximale de chaque sous-domaine.

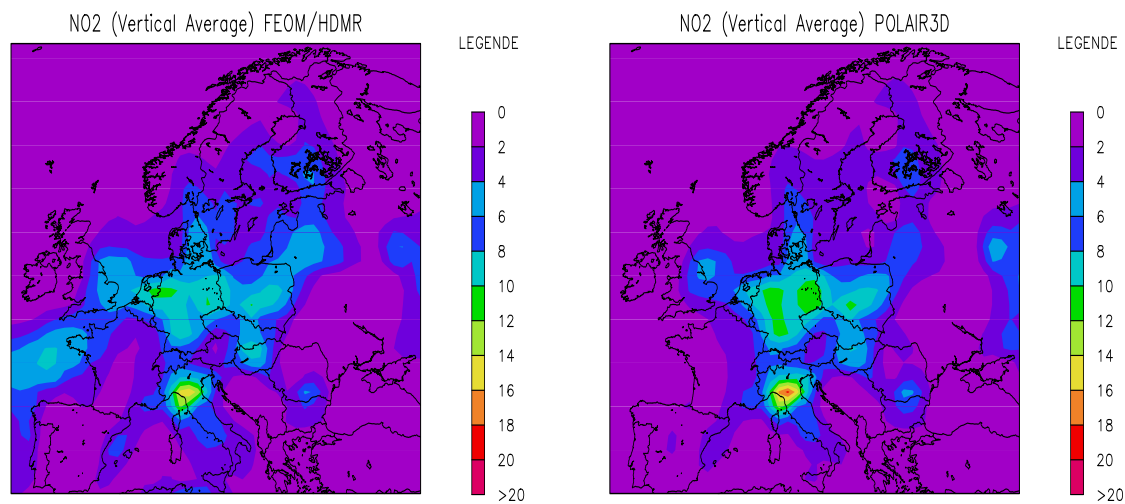


Figure 4.19: Multi-FEOM : Comparaison de NO_2 (en moyenne verticale), après 12h d'intégration par FEOM et POLAIR3D en $\mu g/m^3$

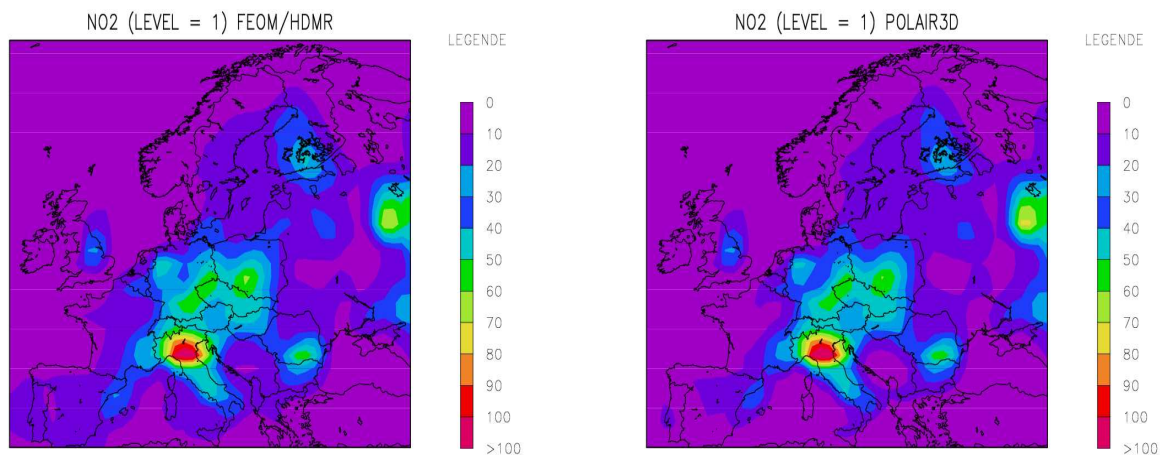


Figure 4.20: Multi-FEOM : Comparaison de NO_2 (au 1er niveau), après 24h d'intégration par FEOM et POLAIR3D en $\mu g/m^3$

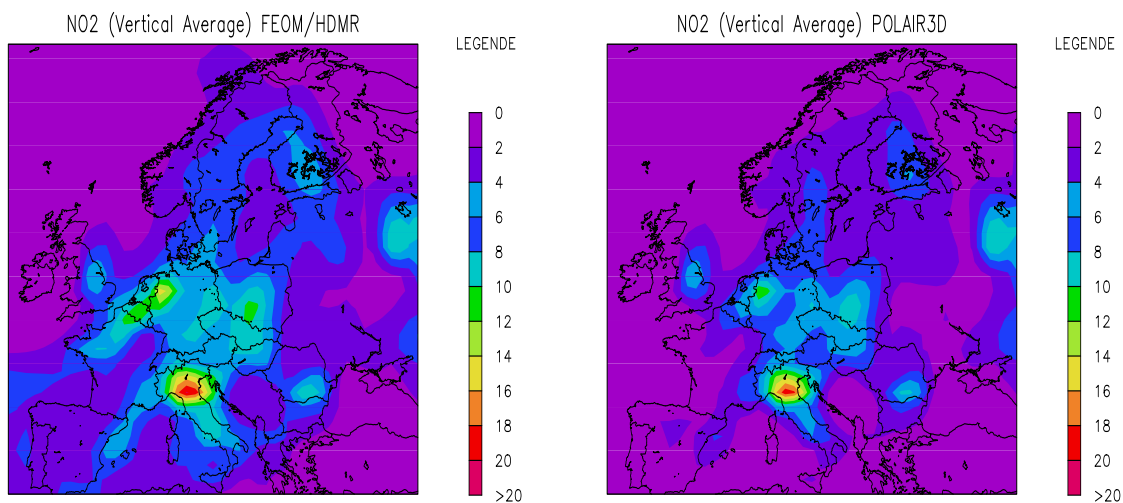


Figure 4.21: Multi-FEOM : Comparaison de NO_2 (en moyenne verticale), après 24h d'intégration par FEOM et POLAIR3D en $\mu g/m^3$

4.6.4 Gain en temps de calcul

Dans le tableau 4.9 nous présentons le speed-up obtenu avec le modèle multi-FEOM pour différents pas de temps Δt_{FEOM} . Le modèle POLAIR3D est utilisé avec un pas de temps $\Delta t_{MODEL} = 900s$.

Δt_{FEOM}	Speed Up
900s	1.33
1800s	2.66
2700s	4.
3600s	5.32

Table 4.9: Speed-up du modèle FEOM.

Conclusion

Nous avons présenté dans ce chapitre la méthode de tabulation réduite HDMR (High Dimensional Model Representation). Après quelques exemples d'illustration et une implémentation de la méthode dans le modèle POLAIR, nous avons présenté quelques applications (notamment la méthode de cut-HDMR) à des cinétiques chimiques, ce qui a montré une conservation de la précision et un fort gain en temps de calcul. Quelques études de sensibilité au point de référence et au nombre de points d'échantillonnages utilisés dans Cut-HDMR, ont révélé une forte importance du choix du point de référence et donc des intervalles définissant les variations potentielles des variables d'entrée du modèle.

Nous avons ensuite présenté une application de la méthode de réduction HDMR au modèle POLAIR3D à l'échelle européenne avec une chimie d'acidification EMEP, au long d'une journée de calcul. Le module chimique a été ainsi remplacé par une tabulation. La méthode a nécessité un découpage du domaine global en sous-domaines afin de construire plusieurs modèles FEOM et garantir une bonne précision de modèle à comparer au modèle d'origine POLAIR3D.

Pour le coût de calcul, la construction de la table HDMR prend évidemment plus de temps comparé à un FEOM global, mais ce calcul fait partie de la phase de prétraitement. Le Speed Up obtenu avec FEOM peut atteindre 5 pour la chimie EMEP qui reste une chimie très simplifiée.

Ces tests préliminaires en 3D sont à prendre comme des tests de faisabilité. Un travail plus profond doit être fait sur la façon de découper le domaine sur la base d'une climatologie détaillée des espèces chimiques dans le domaine de calcul (Europe pour des problématiques d'acidification transfrontalière). La qualité des résultats HDMR est très dépendante du choix du point de référence \bar{x} dans l'espace des variables d'entrée. Pour cela on peut envisager un calcul des tables HDMR selon les saisons.

Chapitre 5

Application de la méthode POD (Proper Orthogonal Decomposition) à la cinétique chimique

Résumé

Dans ce chapitre nous présentons une autre méthode de réduction, la méthode POD. La décomposition orthogonale propre (POD), décrite initialement par Lumley (1970) pour des applications à la turbulence, également connue sous le nom de Empirical Orthogonal Function (EOF) dans le domaine de la géophysique, est une méthode qui permet de transformer un système non linéaire d'équations aux dérivées partielles de grande dimension en un système d'équations différentielles non linéaires de dimension réduite, sur la base d'une analyse statistique de la trajectoire. Nous présentons ici l'application de cette méthode à une cinétique chimique 0D.

Introduction

La méthode POD (Proper Orthogonal Decomposition) est une technique d'analyse de données multidimensionnelle. Elle fournit une base orthonormale représentative d'un jeu de données bien précis. Cela permet de donner une approximation optimale de dimension faible aux données initiales qui contiennent un nombre de degrés de liberté très élevé.

L'approche POD est parfois combinée avec la méthode de projection Galarkin ([19]) pour générer un système réduit du modèle initial qui comporte une large dimension. Une des principales propriétés de cette méthode est son application à des systèmes non linéaires, et la génération d'un modèle réduit non linéaire après projection du système initial sur la base tronquée POD.

Cette méthode est fortement utilisée pour des applications de traitement de signal pour la compression de données, traitement d'images, dans la théorie de contrôle, dans la turbulence, dans les méthodes inverses et dans la dynamique des fluides. Pour plus de détails on peut se référer à([71, 72, 42, 16]).

Le jeu de données à partir duquel la base réduite est construite peut être soit un jeu de données expérimentales, soit un jeu de résultats de simulation numérique (ce qui sera notre cas).

Comme nous l'avons souligné auparavant, les équations différentielles qui décrivent la chimie du modèle POLAIR3D sont non-linéaires et font intervenir un grand nombre de réactions et d'espèces chimiques dont les échelles de temps sont très dispersées. Nous rappelons que le module de chimie consomme jusqu'à 90 % du temps CPU total d'un modèle de chimie-transport. Nous avons donc utilisé la méthode POD dans le contexte de réduction du système chimique de POLAIR3D indépendamment des autres modules (advection, diffusion, ...)

Nous présentons dans ce chapitre le fondement théorique et les principales propriétés de la méthode POD, notre implémentation algorithmique et des applications à des cinétiques chimiques (CBM4 et RACM) utilisées dans POLAIR3D.

5.1 Fondement théorique de la méthode POD

Le module de chimie du modèle de chimie-transport POLAIR3D peut s'écrire sous la forme fonctionnelle suivante :

$$\frac{dc}{dt} = f(c) \quad (5.1)$$

où c est le vecteur des concentrations des espèces chimiques de dimension N_{esp} ($C \in R^{N_{esp}}$), $f(c)$ est la fonction chimie. Le système à résoudre est donc de dimension N_{esp} pour un point donné de l'espace.

L'objectif de la méthode POD est de trouver une base réduite de dimension $N_r \ll N_{esp}$ qui contient toutes les informations sur le comportement de la solution exacte du modèle (5.1). Cette base réduite peut être calculée à partir d'une série de données mesurées ou de données issues d'une simulation. Cette dernière approche basée sur la trajectoire du modèle est classiquement appelée "méthode de snapshots".

L'objectif de la méthode de réduction est de construire un autre modèle non linéaire d'équation différentielle :

$$\frac{d\bar{c}}{dt} = \bar{f}(\bar{c}) \quad (5.2)$$

où $\bar{c} \in R^{N_r}$ et $N_r < N_{esp}$, tel que les comportements des deux systèmes restent similaires.

5.1.1 Méthode des Snapshots

La méthode des Snapshots consiste à calculer une trajectoire exacte du modèle (5.1) durant une période T composée de N_s pas de temps $(t_1, t_2, \dots, t_{N_s})$, que nous notons $\mathcal{C} = c(t_1), c(t_2), \dots, c(t_{N_s})$. \mathcal{C} peut être vu aussi comme un ensemble de données de mesure ou d'expérimentations.

L'objectif est d'extraire le plus d'informations possible de cette trajectoire. Bien évidemment toute cette information est contenue dans une base de dimension d , $\phi' = \{\phi'_1, \phi'_2, \dots, \phi'_d\}$. La base étant supposée orthonormale par rapport à un produit scalaire donné $\langle \cdot, \cdot \rangle$, on a :

$$\forall u \in \mathcal{C}, u = \sum_{i=1}^{i=d} \langle \Phi'_i, u \rangle \Phi'_i \quad (5.3)$$

où $\langle \Phi'_i, \Phi'_i \rangle = \|\Phi'_i\|^2 = 1$.

La méthode de POD est basée sur un choix particulier et optimal de la base orthonormale $\{\phi'\}$ pour que $c(t_j)$ puisse être approché de manière optimale par une somme tronquée. La base Φ'_i est donc calculée de la manière suivante.

Pour tout k tel que $1 \leq k \leq d$, $(\Phi'_1, \dots, \Phi'_k)$ minimise la fonction suivante :

$$J_k(\Phi'_1, \dots, \Phi'_k) = \int_0^T \|c(t_j) - \sum_{i=1}^{i=k} \langle \Phi'_i, c(t_j) \rangle \Phi'_i\|^2 dt \quad (5.4)$$

où $\|\cdot\|$ est la norme correspondante au produit scalaire $\langle \cdot \rangle$.

La solution de la minimisation de la fonction (5.4) $\{\Phi'_i\}_{i=1}^k$ est appelée la base POD d'ordre k .

5.1.1.1 Base POD

La minimisation de la fonction J_k demande la construction de la matrice de corrélation définie par :

$$K_{ij} = \frac{1}{N_s} \langle c(t_i), c(t_j) \rangle \quad (5.5)$$

Comme détaillé dans [71], les éléments de la base POD Φ' sont en réalité solution de l'équation aux valeurs propres :

$$K\Phi' = \lambda\Phi' \quad (5.6)$$

qui correspondent aux vecteurs propres de la matrice de corrélation K de la snapshot \mathcal{C} .

Remarque :

En pratique la matrice de corrélation est calculée en prenant en compte la valeur

moyenne de c , $\bar{c} = \frac{1}{N_s} \sum_{i=1}^{i=N_s} c_i$

$$K_{ij} = \frac{1}{N_s} \langle c(t_i) - \bar{c}, c(t_j) - \bar{c} \rangle \quad (5.7)$$

K est une matrice positive de rang r et ses valeurs propres peuvent être ordonnées : $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \geq 0$. On note les vecteurs propres V_1, V_2, \dots, V_r associés normalisés selon :

$$\begin{cases} \langle V_i, V_j \rangle = \frac{1}{N_s \lambda_j} & \text{si } (i = j) \\ \langle V_i, V_j \rangle = 0 & \text{si } (i \neq j) \end{cases} \quad (5.8)$$

Partant des principaux modes de POD correspondant à la plus grande valeur propre, on va se donner une valeur de coupure λ_{N_r} telle que :

$$\sum_{i=1}^{N_r} \lambda_i \approx \sum_{i=1}^r \lambda_i \quad (5.9)$$

On peut alors montrer que le résidu de l'erreur de la troncature associée à $(\Phi'_1, \dots, \Phi'_{N_r})$ est donnée par :

$$\int_0^T \|c(t_j) - \sum_{i=1}^{i=N_r} \langle \Phi'_i, c(t_j) \rangle \Phi'_i\|^2 dt = \sum_{i=N_r+1}^r \lambda_i \quad (5.10)$$

Le rapport donne le pourcentage de l'erreur quadratiques moyenne quand on élimine les N_r modes de POD dans la base réduite [72].

Le choix de N_r dimension de la base réduite de POD est géré par la définition d'un ε^{POD} tel que :

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{N_r} \geq \varepsilon^{POD} \dots \geq \lambda_r > 0 \quad (5.11)$$

La base réduite de dimension N_r étant obtenue, on a alors pour tout i , $1 \leq i \leq N_r$:

$$\Phi'_i = \frac{1}{\sqrt{\lambda_i}} \sum_{j=1}^{N_s} (V_i)_j c(t_j) \quad (5.12)$$

Les principales questions qu'on peut se poser dans la méthode POD sont les suivantes :

- Comment peut-on choisir un ensemble de Snapshot capable de fournir une approximation globale de la solution (modèle d'origine) ?
- Comment choisir les éléments de la base réduite qui peuvent capturer l'ensemble des propriétés et informations de l'ensemble d'origine ?

5.1.2 Projection du modèle d'origine sur la nouvelle base réduite

On suppose ici que la base réduite de POD a déjà été calculée à partir d'une snapshot du modèle en off-line, $\Phi' = \{\Phi'_1, \Phi'_2, \dots, \Phi'_{N_r}\}$, avec $N_r \leq N_s$. N_r est donc le nombre de degrés de liberté retenu. Nous cherchons ici à projeter le système initial sur la nouvelle base réduite :

$$\frac{dc}{dt} = f(c), \quad c(0) = c_0 \in R^n \quad (5.13)$$

$c(t)$ peut s'écrire dans le nouveau système sous la forme

$$c(t) = \sum_{i=1}^{N_r} c_i(t) \Phi'_i, \quad c_i \in R \quad (5.14)$$

et donc :

$$\frac{dc}{dt} = \sum_{i=1}^{N_r} \frac{dc_i(t)}{dt} \Phi'_i \quad (5.15)$$

ce qui nous amène à :

$$\frac{dc_i}{dt} = \langle f(c), \Phi'_i \rangle \quad (5.16)$$

le modèle réduit est donc pour tout i , $1 \leq i \leq N_r$:

$$\frac{dc_i}{dt} = \left\langle f\left(\sum_{i=1}^{N_r} c_i(t) \Phi'_i\right), \Phi'_i \right\rangle, \quad c_i(0) = \langle c(0), \Phi'_i \rangle \quad (5.17)$$

La résolution du système réduit 5.17 et l'Equation 5.14 permettront de déduire la concentration des espèces chimiques composantes du vecteur c . Notons que dans le cas général, et pour notre cas également on ne dispose pas d'un contrôle à priori de l'erreur induite par projection sur la base POD.

5.2 Implémentation algorithmique

L'implémentation de la méthode POD dans un modèle de cinétique passe par deux étapes : la première consiste à calculer une trajectoire du modèle à partir de laquelle la base POD sera produite, la seconde à projeter le modèle initial sur la base ainsi calculée. Dans la figure 5.1 est présenté un schéma d'illustration de la méthode POD. L'algorithme de l'implémentation de POD est détaillé dans ce qui suit.

Algorithme 1 : Construction de la base POD

- Calcul de la Snapshot \mathcal{C} correspondant à la trajectoire du modèle pendant t_1, t_2, \dots, t_{N_s} pas de temps.
- Calcul de la moyenne de trajectoire, pour une utilisation d'une matrice de corrélation centrée autour de la moyenne.
- Calcul de la matrice de corrélation K
- Calcul du spectre de la matrice K : valeurs propres et vecteurs propres associés.
- Tri des valeurs propres dans l'ordre décroissant.

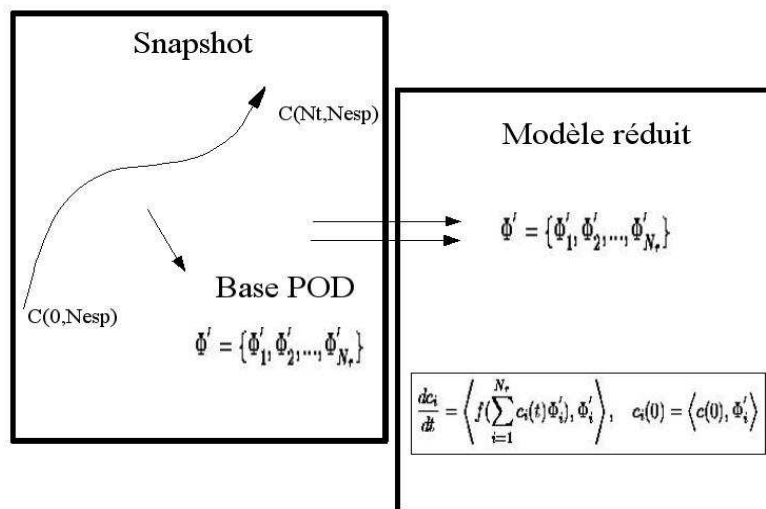


Figure 5.1: Schéma d'illustration de la méthode POD, on distingue deux parties : une pour la construction de la base l'autre pour la construction du modèle réduit

- Choix d'un ε^{POD} pour filtrer les valeurs propres inférieures à cette valeur : on ne garde que les valeurs propres importantes qui serviront à calculer la base réduite.
- Calcul des éléments de la base réduite POD conformément à l'équation (5.12).
- Sauvegarde de la base POD.

Algorithme 2 : Projection du système initial sur la base POD

- Lecture de la base POD réduite.
- Calcul de la condition initiale dans la nouvelle base réduite.
- Calcul du nouveau terme de production de l'équation de modèle réduit.
- Résolution de l'équation du modèle réduit par Rosenbrock (Solveur numérique utilisé dans le cas du modèle POLAIR).

Notons qu'un point clé est le choix du produit scalaire $\langle \cdot, \cdot \rangle$, qui détermine la base POD. Dans un premier temps, nous avons pris le produit scalaire euclidien.

5.3 Application à des cinétiques chimiques

5.3.1 Application au mécanisme CBM IV

Nous avons appliqué cette technique pour réduire le mécanisme chimique du modèle POLAIR3D, le schéma CBMIV, composé de 31 espèces et 68 réactions. Le temps de calcul de la base POD est de 4h.

Le nombre de modes retenu dans cette application est de 3 comme montré sur le spectre de la matrice de corrélation figure (5.2). Dans la figure (5.3) nous comparons les résultats obtenus avec le modèle réduit et le modèle exact (POLAIR0D).

Spectre de la matrice de corrélation

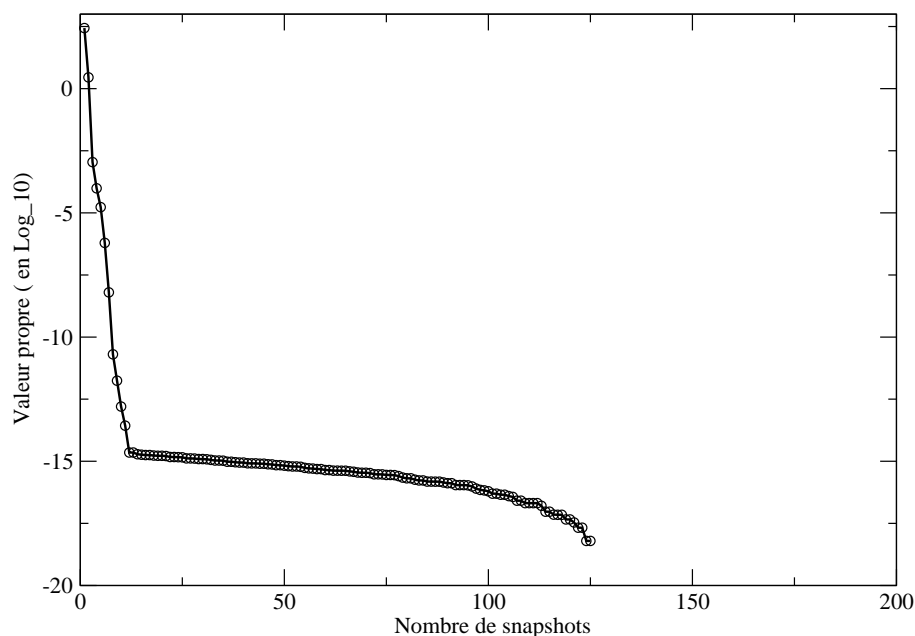


Figure 5.2: Spectre de la matrice de corrélation (CBM4), Snapshot de période 4h

Nous pouvons observer que le modèle réduit obtenu avec POD est très performant durant les 500 minutes, comparé aux 240 minutes, période de la snapshot. Le modèle réduit est donc valable localement. Nous notons que le nombre de degré de liberté retenu est $p = 3$, à comparer à la dimension initiale du système (31).

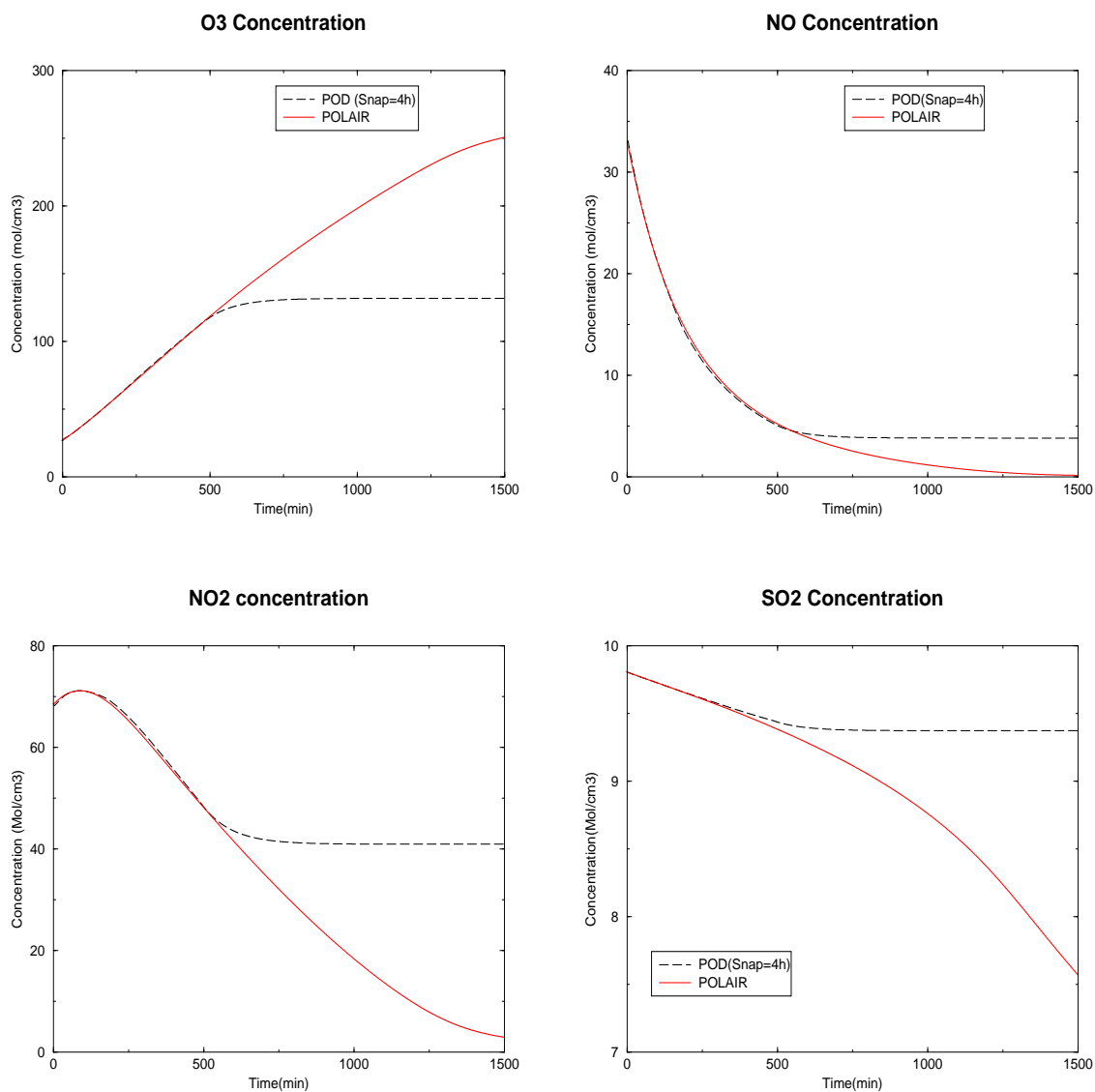


Figure 5.3: Comparaison des profils de concentrations de O_3 , NO_2 , SO_2 et CO simulées avec le modèle réduit POD et le modèle initial de référence POLAIR. Le mécanisme chimique utilisé est CBM-IV

5.3.2 Application au mécanisme RACM

Dans cette section nous présentons une autre application de la méthode POD au mécanisme RACM ([51]) qui fait intervenir 72 espèces et 273 réactions chimiques. La base POD est calculée à partir d'une trajectoire du modèle POLAIR d'une durée T avec un nombre N_T d'itérations pour un pas de temps de δt . On notera la trajectoire $\{C = (C_1, \dots, C_{Nesp})\}_{i=0}^{N_T}$.

Pour ce cas test, la base POD est calculée à partir des résultats du modèle POLAIR simulé pendant 3h et avec un pas de temps de 600 secondes. le nombre de modes retenu par la méthode POD est de 3 (pour un $\varepsilon^{POD} = 10^{-3}$), soit 3 degrés de libertés au lieu de 72 initialement (figure 5.4), que l'on note (Φ_1, Φ_2, Φ_3) .

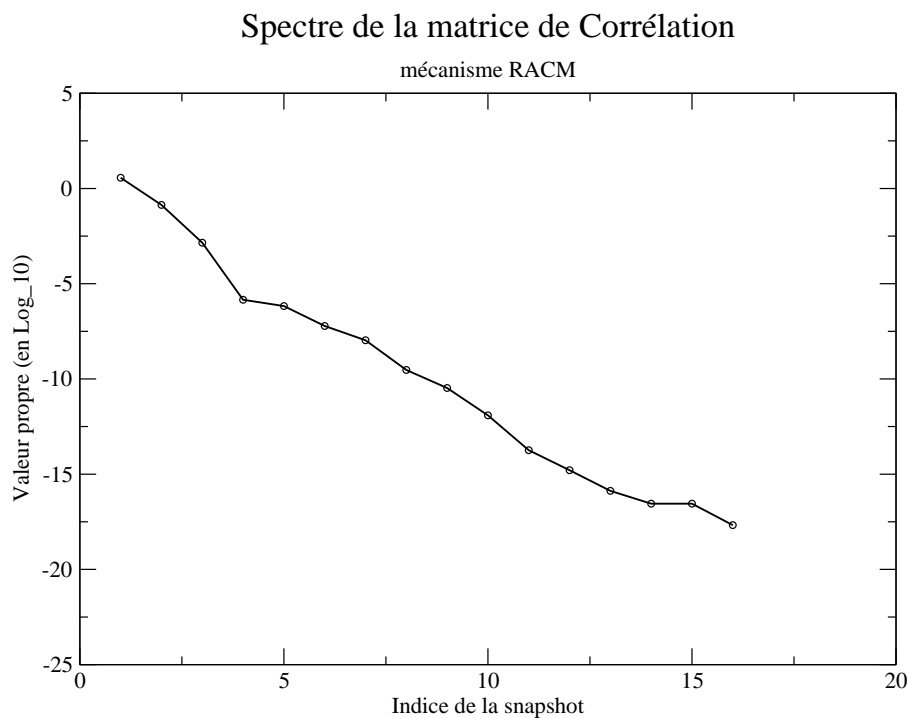


Figure 5.4: Spectre de la matrice de corrélation (RACM), Snapshot de période 3h

Un résultat préliminaire est schématisé dans la figure 5.5 qui présente une comparaison entre le modèle d'origine POLAIR et le modèle réduit construit avec la base POD (Φ_1, Φ_2, Φ_3) , pour les espèces chimiques O_3 , NO_2 , SO_2 et CO . La simulation est faite avec les mêmes conditions initiales qui ont servi à construire la trajectoire pour le calcul de la base POD.

Nous avons effectué une simulation de 7 heures, le modèle réduit POD reproduit très bien le comportement et les valeurs des concentrations simulées par le modèle POLAIR pendant et au delà des 3h, durée de la snapshot. Le modèle réduit reste valable jusqu'à 300 minutes d'intégration soit 5h de calcul et au delà le modèle POD donne un comportement assez différent de celui du modèle d'origine, comme on peut le constater sur la carte de l'ozone : POD a tendance à diminuer les valeurs de l'ozone alors que POLAIR simule l'inverse.

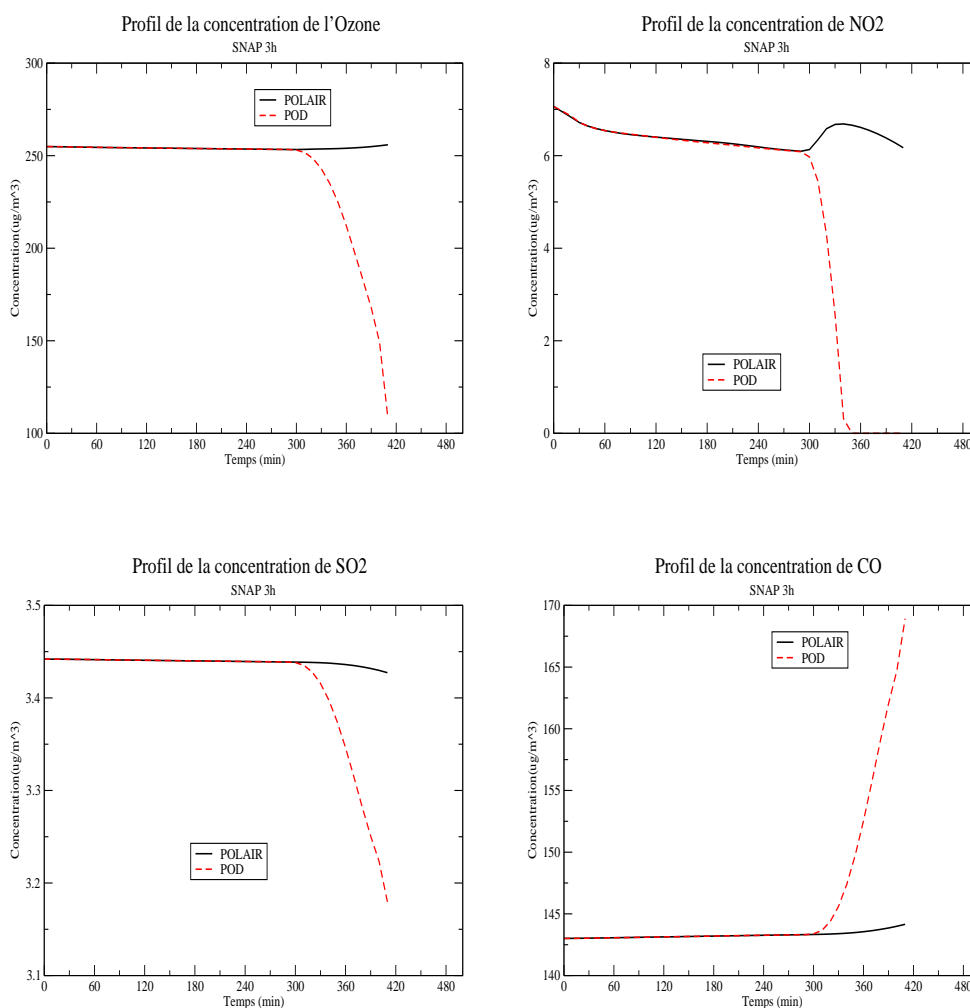


Figure 5.5: Comparaison des profils de concentrations de O_3 , NO_2 , SO_2 et CO simulées avec le modèle réduit POD et le modèle initial de référence POLAIR, le mécanisme chimique utilisé est RACM, la période de la snapshot $T=3$ heures et $Dt=600s$ ($N_T = 18$)

5.3.3 Etudes de sensibilité POD-RACM

Pour mieux comprendre les limites de la méthode POD nous approfondissons l'étude de sensibilité des paramètres qui pourraient avoir un effet sur la base POD et la qualité du résultat du modèle réduit en analysant l'influence :

- de la période de la snapshot T : nous avons pris plusieurs trajectoires de 3, 4, 6, 9 et 12 heures.
- du pas de temps de la trajectoire : δt , 100s, 300s et 600s.
- de l'état initial de la trajectoire $\{C\}_0$ sur la base POD calculée, nous avons testé plusieurs scénarios de C_0 allant de la valeur initiale à une variation de 10, 20, 30, 40, 45 ou 50%.
- La base POD est calculée pour une trajectoire donnée C , nous étudierons aussi l'effet de la condition initiale sur la qualité du modèle réduit obtenu, Dans ce cas nous avons fait varier C_0 de sa valeur initiale de 5, 10, 20%.

5.3.3.1 Sensibilité au pas de temps de la snapshot δt et à la durée de la snapshot T

Sensibilité au pas de temps de la Snapshot δt

Le nombre de modes retenus par la technique POD avec une snapshot de 3 heures, pour différents pas de temps utilisé dans le calcul de la trajectoire 100s, 300s et 600s est de 3 dans les trois cas. Nous notons $\{\Phi\}_{100s}$, $\{\Phi\}_{300s}$, et $\{\Phi\}_{600s}$ les trois différentes bases réduites.

La validation des résultats du modèle réduit POD construit avec les 3 trajectoires correspondantes aux pas de temps 100s, 300s, et 600s est faite en comparaison au modèle POLAIR3D pendant une période d'intégration de 7h. Les modèles réduit POD associés aux trois bases $\{\Phi\}_{100s}$, $\{\Phi\}_{300s}$ et $\{\Phi\}_{600s}$ donnent de bons résultats en gardant un bon comportement par rapport au modèle POLAIR3D durant une période de 5h d'intégration période qui dépasse largement la période 3h de la snapshot (figures 5.7, 5.6 et 5.5). Quelques fortes explosions sont à remarquer au delà de la période de la snapshot (après 6 heures d'intégration) pour les deux modèles réduits associés à $\{\Phi\}_{100s}$ et $\{\Phi\}_{300s}$ respectivement pour les espèces SO_2/CO et NO_2 .

Sensibilité à la durée de la Snapshot T

La table 5.1 synthétise le nombre de modes retenus par la technique POD pour différentes périodes de la snapshot 3, 4, 6 et 9 heures. Le nombre de modes

Durée de la Snapshot	Nombre de modes
3h	3
4h	3
6h	4
9h	3
12h	3

Table 5.1: Nombre de modes retenus en fonction de la durée de la snapshot T , le pas de temps de la simulation snapshot est de 600s (Mécanisme RACM).

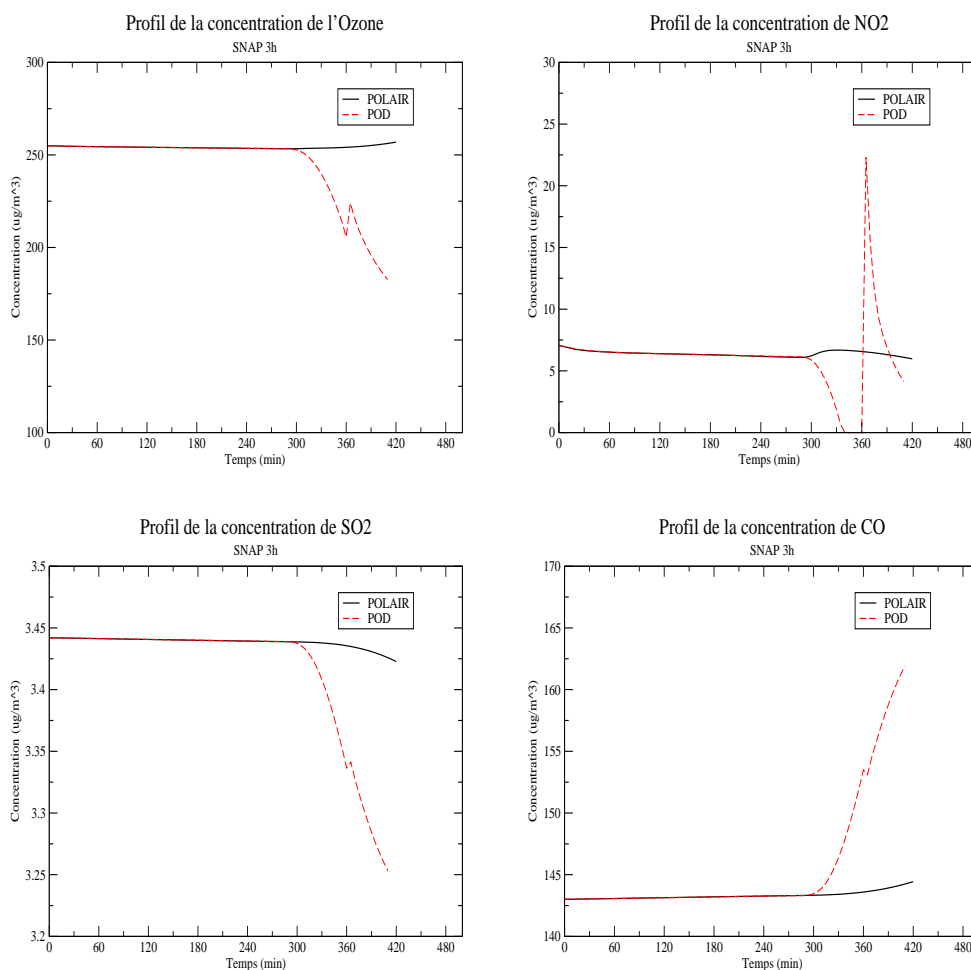


Figure 5.6: Comparaison des profils de concentrations de O_3 , NO_2 , SO_2 et CO simulées avec le modèle réduit POD et le modèle initial de référence POLAIR, le mécanisme chimique utilisé est RACM, la période de la snapshot $T=3$ heures et $Dt=300s$ ($N_T = 36$).

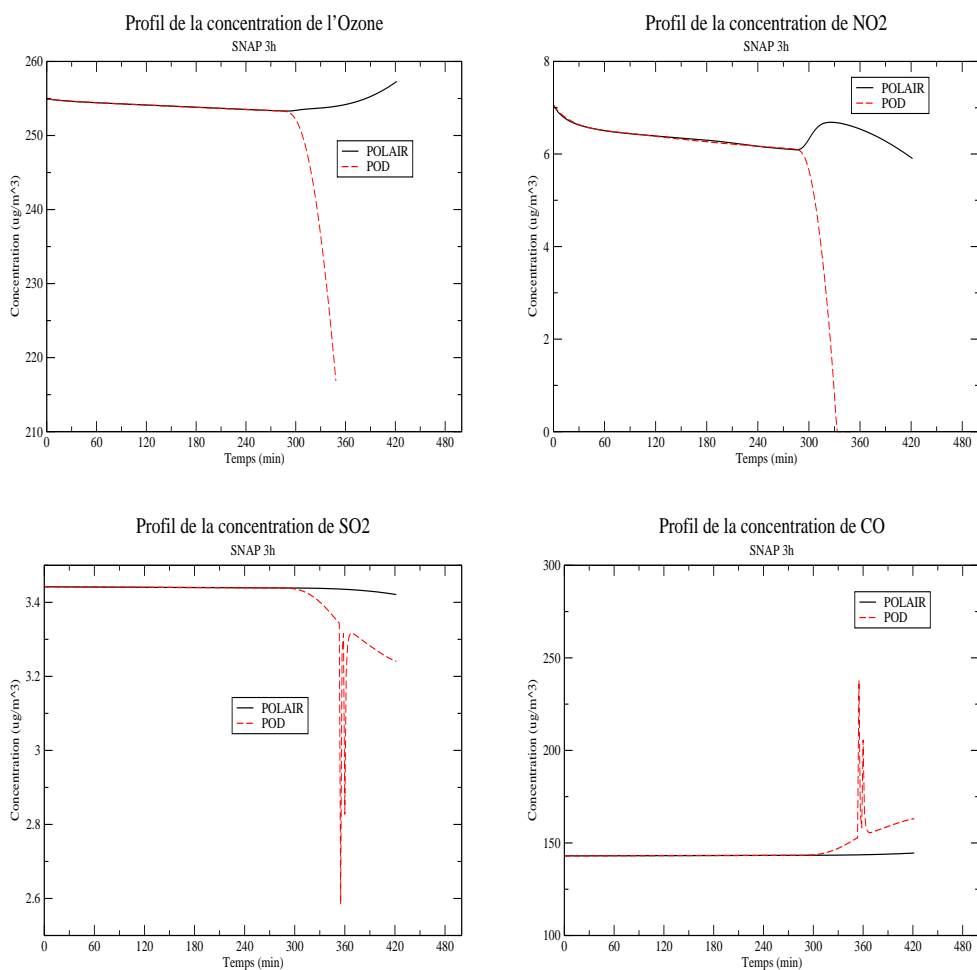


Figure 5.7: Comparaison des profils de concentrations de O_3 , NO_2 , SO_2 et CO simulées avec le modèle réduit POD et le modèle initial de référence POLAIR, le mécanisme chimique utilisé est RACM, la période de la snapshot $T=3$ heures et $Dt=100s$ ($N_T = 108$).

retenus dans les 4 cas est de 3 ou 4 modes. Les quatre différentes bases réduites correspondantes seront notées $\{\Phi\}_{3h}$, $\{\Phi\}_{4h}$, $\{\Phi\}_{6h}$ et $\{\Phi\}_{9h}$.

Pour la sensibilité à la durée de la snapshot T , 4 trajectoires sont calculées à partir du modèle POLAIR, 3, 4, 6 et 9 heures. Les bases POD correspondants contiennent 3 ou 4 modes (tableau 5.1). La figure 5.8 montre le résultat obtenu avec les différentes snapshots, les différents modèles réduits sont intégrés pendant 10 heures, le résultat exposé auparavant sur la validité locale de la base POD est confirmé ici. L'explosion du modèle est observé pour la base $\{\Phi\}_{3h}$, $\{\Phi\}_{4h}$, $\{\Phi\}_{6h}$, après les périodes des snapshots. Pour la base $\{\Phi\}_{3h}$, le modèle réduit diverge à partir de 300 min (5 heures) d'intégration (figure 5.5). Pour la base $\{\Phi\}_{4h}$ le modèle commence à diverger à partir de 5 heures d'intégrations (Figure 5.9), et le résultat pendant les périodes de la snapshot sont excellents dans les deux cas.

Pour la dernière simulation utilisant une trajectoire de 9h, la figure 5.10 présente un résultat du modèle réduit pas semblable au modèle POLAIR3D, pendant la période de la snapshot et l'erreur relative globale est de l'ordre de 20%. La divergence au delà de la période de la snapshot se confirme aussi dans ce cas test.

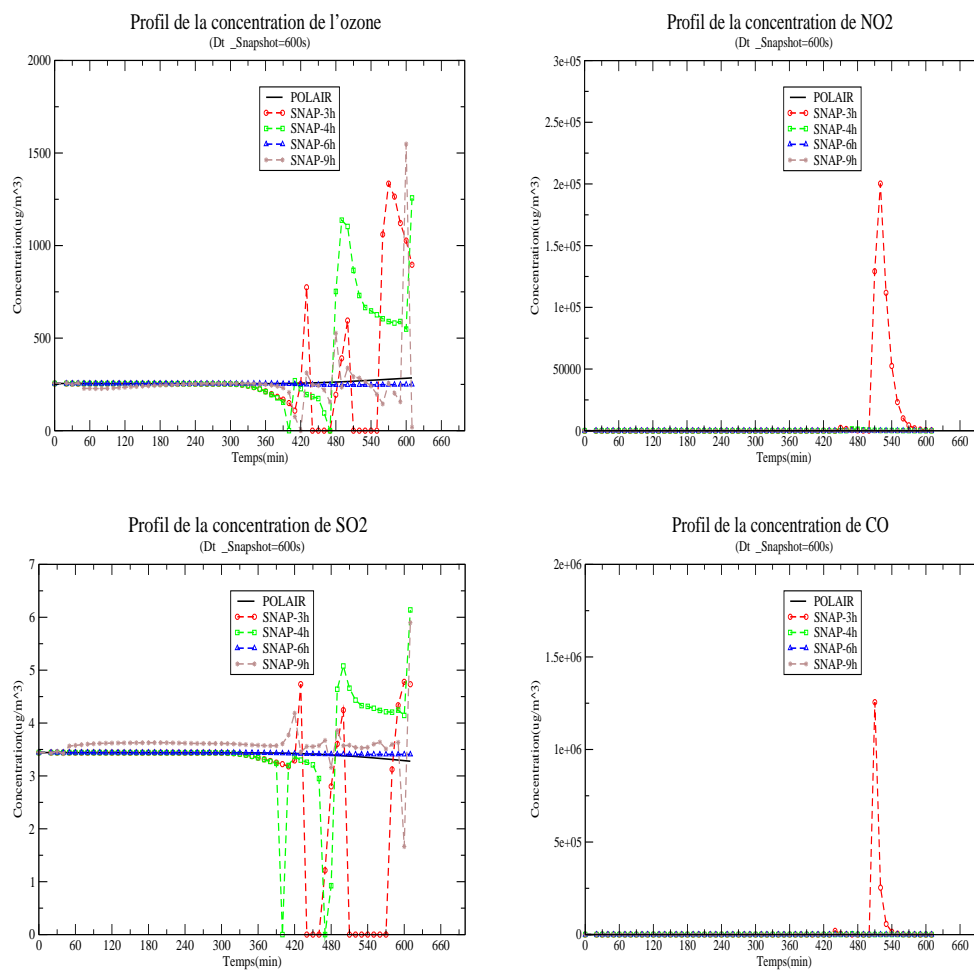


Figure 5.8: Comparaison des profils de concentrations de O_3 , NO_2 , SO_2 et CO simulées avec les différents modèles réduit POD associés aux différents types de snapshots et le modèle initial de référence POLAIR, le mécanisme chimique utilisé est RACM, la période de la snapshot $T=3/4/6/9$ heures et $Dt=600s$ ($N_T = 36$), la durée de la simulation est de 10h.

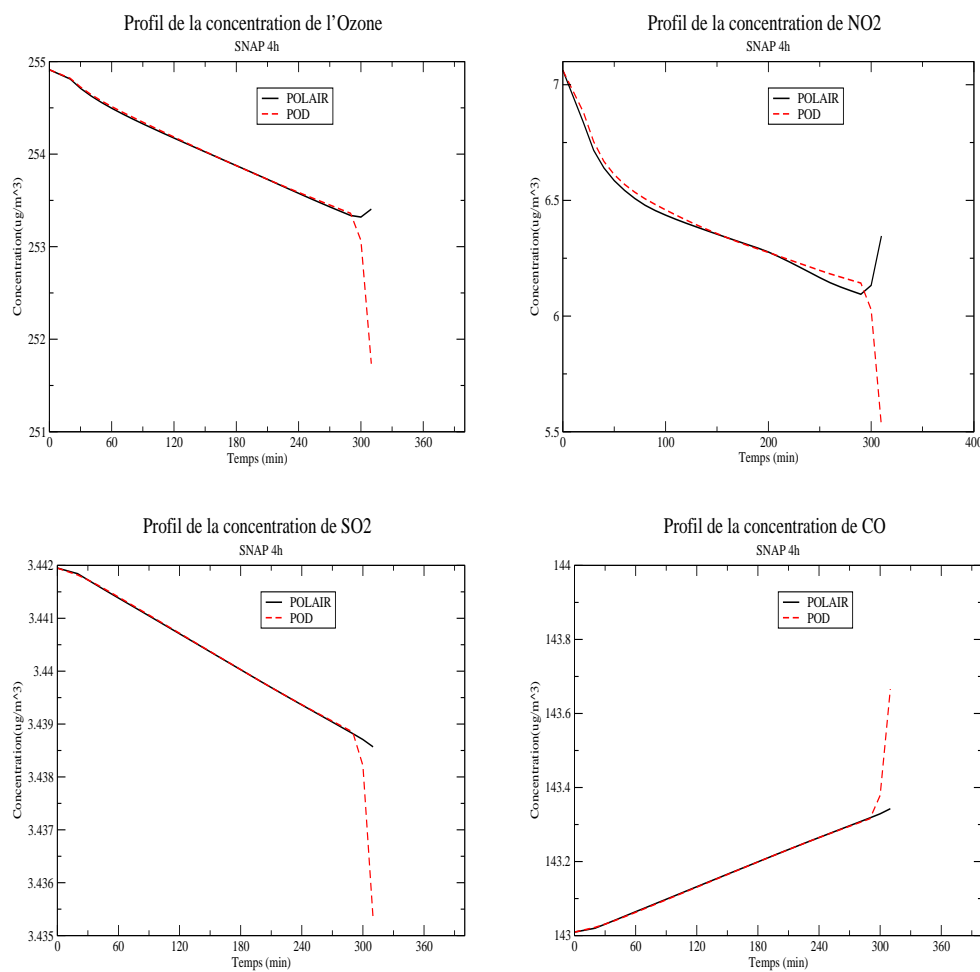


Figure 5.9: Comparaison des profils de concentrations de O_3 , NO_2 , SO_2 et CO simulées avec le modèle réduit POD et le modèle initial de référence POLAIR, le mécanisme chimique utilisé est RACM, la période de la snapshot $T=4$ heures et $Dt=600s$ ($N_T = 24$).

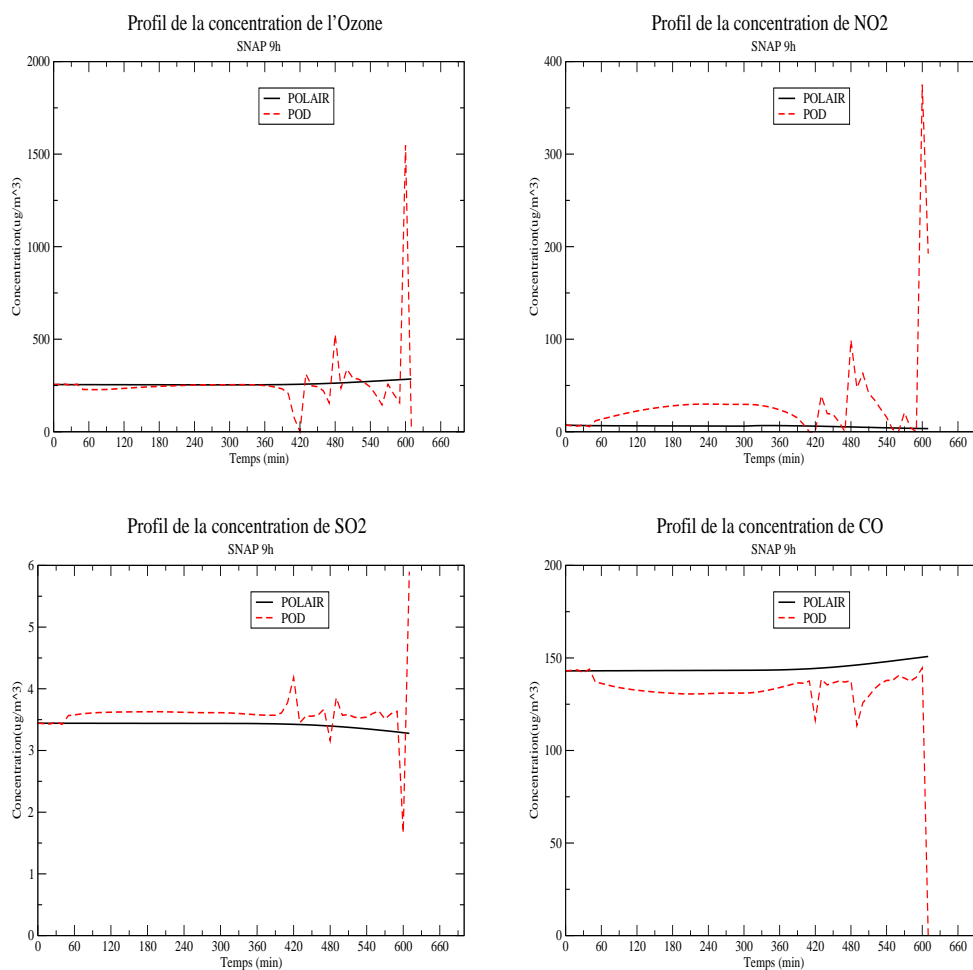


Figure 5.10: Comparaison des profils de concentrations de O_3 , NO_2 , SO_2 et CO simulées avec le modèle réduit POD et le modèle initial de référence POLAIR, le mécanisme chimique utilisé est RACM, la période de la snapshot $T=9$ heures et $Dt=600s$ ($N_T = 54$).

5.3.3.2 Sensibilité du modèle réduit POD à l'état initial du modèle

Pour un ensemble de données X calculé (ou mesuré) à partir du modèle POLAIR pendant une durée T , la méthode POD calcule une base POD réduite $\{\Phi\}_{N_r}^T$ pour approcher un autre jeu de données Y . Dans la section précédente le modèle réduit POD a été appliqué au modèle POLAIR avec la même condition initiale C_0 qui a permis le calcul de la snapshot. Afin de voir jusqu'à quel point la base POD reste valable, nous allons faire varier cette valeur C_0 selon deux scénarios présentés dans les figures 5.11 et 5.12 :

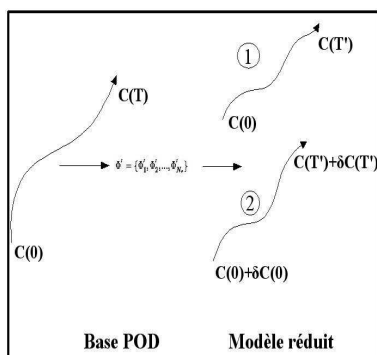


Figure 5.11: Sensibilité du modèle réduit POD à son état initial.

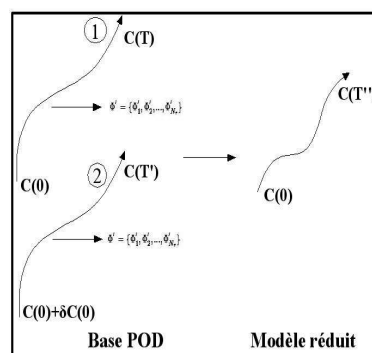


Figure 5.12: Sensibilité du modèle réduit POD à l'état initial de la snapshot.

- Variation de la condition initiale dans le modèle réduit (Figure 5.11). Le modèle construit par la méthode POD avec une snapshot de condition initiale $C(0)$, est simulé avec différentes conditions initiales $C(0)$, $C(0)+5\%C(0)$, $C(0)+10\%C(0)$ et $C(0)+20\%C(0)$, la durée de la snapshot est de 6 heures, le nombre de modes POD retenus est de 4 et la pas de temps modèle est de 600s. Les différents résultats sont schématisés sur la figure 5.13. La méthode POD a été utilisée pour des concentrations centrées autour de leurs moyennes. Les variations effectués sur la condition initiale sont autour de leur moyenne initiale, soit $(\bar{c} + \delta\bar{c}, \bar{c}, \bar{c} = c - \langle c \rangle$ et $\delta\bar{c} = 5, 10, 20\%$). Nous remarquons aisément que le modèle POD associé à une variation de la condition initiale 20% donne des erreurs importantes (espèce NO_2) et ceci pendant la période de la snapshot. Les différents modèles convergent après la période de la snapshot vers les mêmes concentrations.
- L'autre scénario consiste à faire varier la condition initiale de la snapshot selon

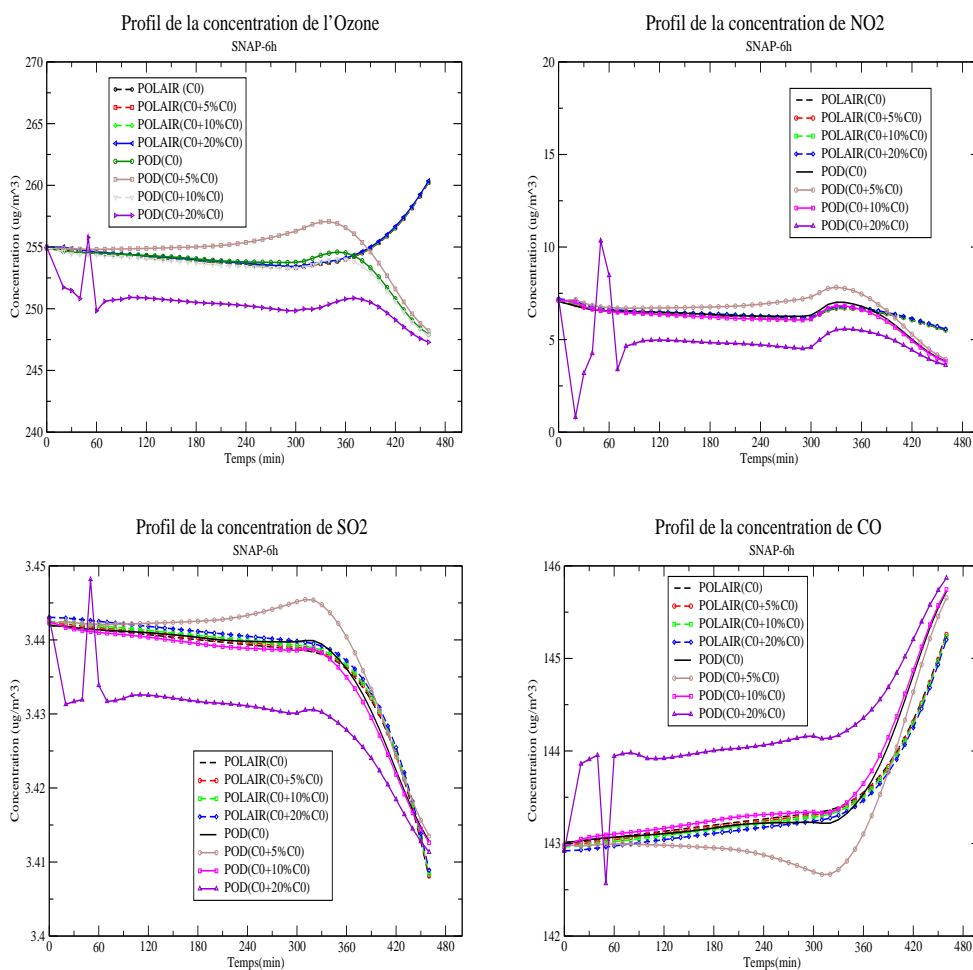


Figure 5.13: Sensibilité du modèle réduit POD à son état initial : comparaison des profils de concentrations de O₃, NO₂, SO₂ et CO simulées avec les différents modèles réduits POD associés aux **différents états initiaux (modèle)** et le modèle initial de référence POLAIR, le mécanisme chimique utilisé est RACM, la période de la snapshot **T=6 heures et Dt=600s (N_T = 36)**, la durée de la simulation est de 7 heures et demie.

la (Figure 5.12), variation effectuée selon :

$$\begin{cases} C\bar{(0)} = C\bar{(0)} + \delta C\bar{(0)} * C\bar{(0)} \\ C\bar{(0)} = C(0) - \langle C \rangle \\ \delta C\bar{(0)} = 10, 20, 30, 40, 45, 50\% \end{cases} \quad (5.18)$$

On notera les différentes snapshots $S_{10\%}$, $S_{20\%}$, $S_{30\%}$, $S_{40\%}$, $S_{45\%}$, et $S_{50\%}$. Le nombre de modes retenus avec la méthode POD dans les différents cas est de 4. Dans la Figure 5.14 sont présentés les 4 modes pour les différents cas de conditions initiales. Nous remarquons que les modes 1, 2 et 3 sont très semblables pour toutes les snapshots $S_{10\%}$, $S_{20\%}$, $S_{40\%}$ et $S_{50\%}$. Cependant le 4ème mode exprime une différence pour quelques espèces chimiques relativement à la Snapshot $S_{50\%}$, ce qui se traduit par des valeurs de concentration aberrantes avec le modèle réduit associé (Figure 5.16). Quant aux résultats obtenus avec les différents modèle réduits associés aux autres bases POD ($S_{10\%}$, $S_{20\%}$, $S_{40\%}$), ils sont très proches entre eux avec peu d'erreur (Figure 5.15). Après la durée de la snapshot il y a une convergence des différents modèles réduit POD vers les mêmes concentrations.

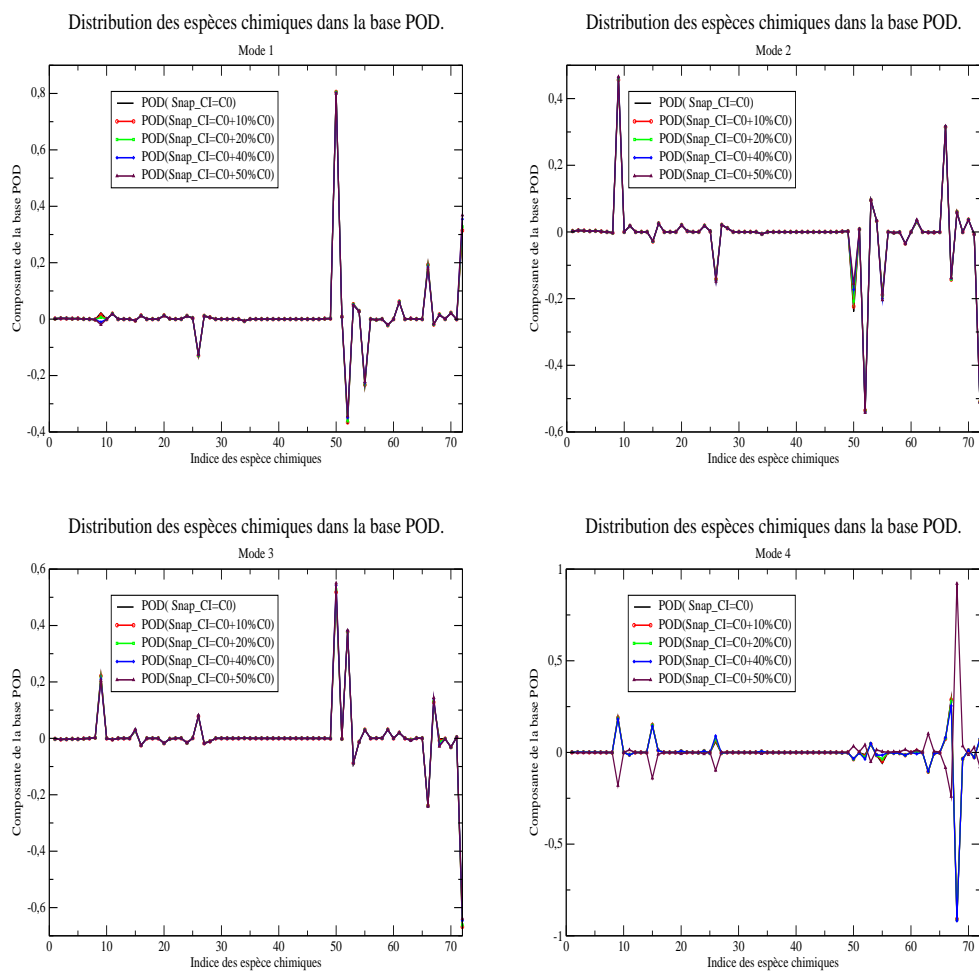


Figure 5.14: Les 4 modes retenus par la méthode POD associés aux différents états initiaux de la snapshot, le mécanisme chimique utilisé est RACM, la période de la snapshot $T=6$ heures et $Dt=600s$ ($N_T = 36$).

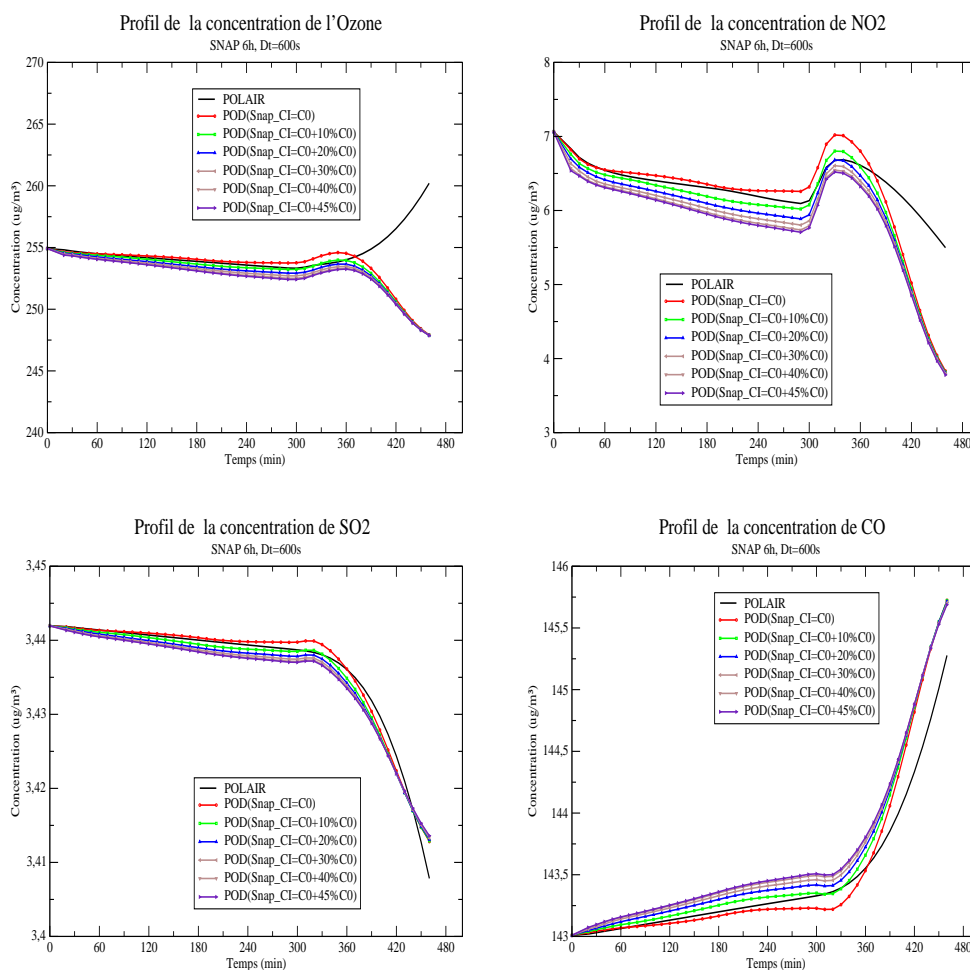


Figure 5.15: Sensibilité du modèle réduit à l'état initial de la snapshot : Comparaison des profils de concentrations de O_3 , NO_2 , SO_2 et CO simulées avec les différents modèles réduit POD associés aux **différents états initiaux de la snapshot** et le modèle initial de référence POLAIR, le mécanisme chimique utilisé est RACM, la période de la snapshot **T=6 heures et Dt=600s** ($N_T = 36$), la durée de la simulation est de 7 heures et demie.

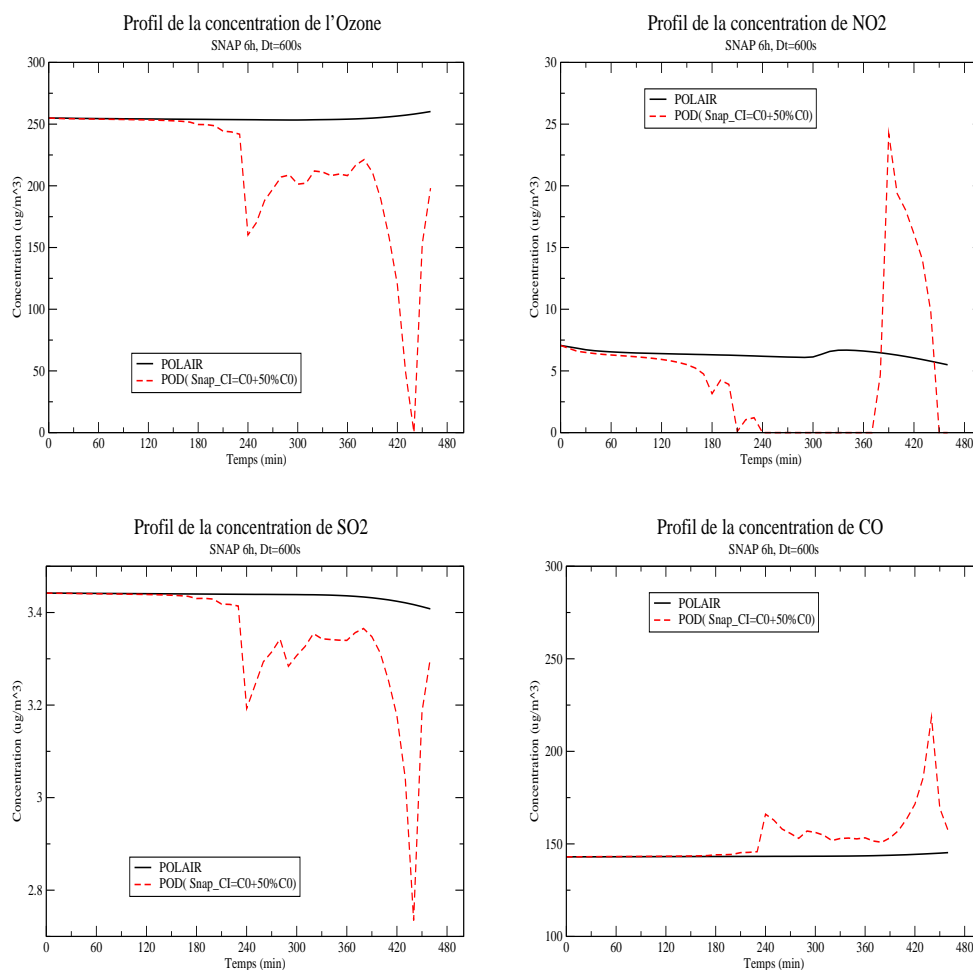


Figure 5.16: Sensibilité du modèle réduit à l'état initial de la snapshot (**suite**) : Comparaison des profils de concentrations de O_3 , NO_2 , SO_2 et CO simulées avec le modèle réduit POD associés aux **une variation de 50 % de l'état initiale de la snapshot** et le modèle initial de référence POLAIR, le mécanisme chimique utilisé est RACM, la période de la snapshot **T=6 heures et Dt=600s** ($N_T = 36$), la durée de la simulation est de **7 heures et demie**.

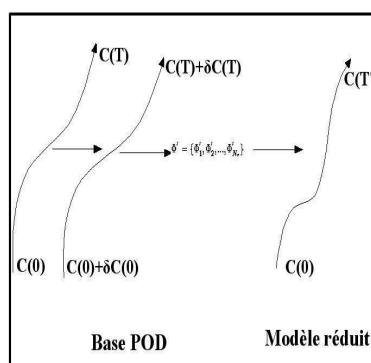


Figure 5.17: Illustration du scénario de sensibilité du modèle réduit POD à l'état initial de la snapshot

5.3.3.3 Sensibilité à la totalité de la Snapshot

Nous avons effectué un autre test de sensibilité où nous perturbons toute la trajectoire initiale qui permet de déduire la base POD (figure 5.17).

Le test consiste à calculer la base POD avec une snapshot $S + \delta S$ où S est la trajectoire initiale, contrairement au cas précédent où nous perturbons uniquement la conditions initiale de la snapshot. La période de la snapshot est de 6h avec un pas de temps de 600 secondes. La perturbation est de 0.1% et 0.3%. Comme le montre la Figure 5.18, de faibles perturbations sur la trajectoire induisent des très grandes perturbations sur la concentration simulée par le modèle réduit, et cela même pendant la période la snapshot avec des erreurs très importantes dépassant les 200%.

Remarque : Temps CPU

Le speedup ($S = \frac{CPU_{Polair}}{CPU_{POD}}$) défini comme le rapport du temps CPU du modèle POD par le temps CPU du modèle de référence, est de l'ordre de 4. Les deux modèles ont été utilisé sur une machine de type PC ayant un processeur de 1.5 GHZ et de 1GO de mémoire.

Le coût de construction de la base POD à partir d'une Snapshot par rapport au coût de calcul de la trajectoire en question (6h de POLAIR-EMEP) est 10 fois plus rapide.

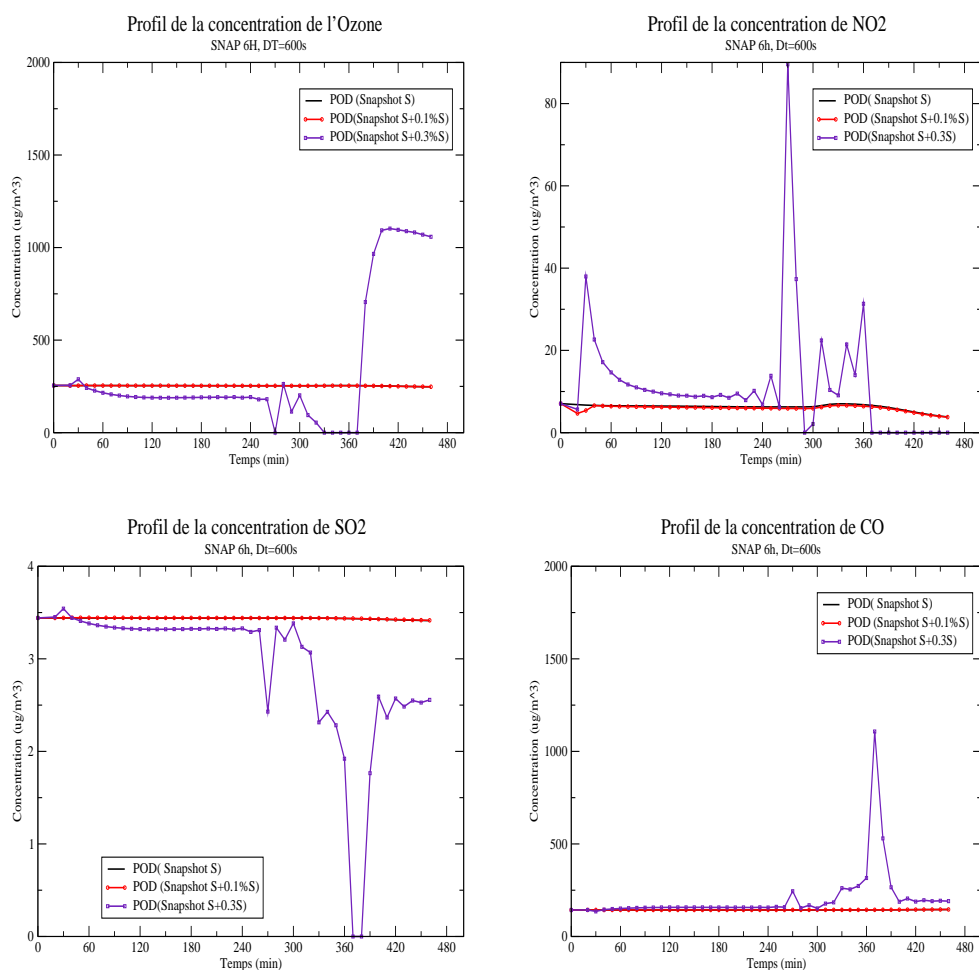


Figure 5.18: Sensibilité du modèle réduit POD à la perturbation de la totalité de la trajectoire : Comparaison des profils de concentrations de O_3 , NO_2 , SO_2 et CO simulées avec les différents modèles réduits POD associés aux **différentes perturbations de la snapshot** et le modèle initial de référence POLAIR, le mécanisme chimique utilisé est RACM, la période de la snapshot **T=6 heures** et **Dt=600s** ($N_T = 36$), la durée de la simulation est de **7 heures et demie**

Conclusion

Nous avons présenté dans ce chapitre une méthode de réduction dite POD (Proper Orthogonal Decomposition) basée sur la projection du modèle de cinétique chimique sur une base optimale réduite calculée à partir d'une trajectoire calculée au préalable avec le même modèle POLAIR. Nous avons appliqué la méthode à deux cinétiques chimiques différentes, les schémas CBM4 et RACM.

La technique capture très bien le comportement du modèle initial, mais sa validité reste locale, au delà de la période de la snapshot (quelques heures), le modèle POD a tendance à diverger de la solution exacte et peut donner dans ces cas des valeurs aberrantes. En général le nombre de modes retenus dépend des caractéristiques de la simulation : la période de la snapshot, le pas de temps de la snapshot, et des conditions initiales, mais en général le nombre de modes ne dépasse pas 5.

Quelques sensibilités ont été remarquées, notamment :

- La méthode est sensible à la période de la snapshot, pour une période de 9h , la base POD construite fournit des résultats moins précis que ceux obtenus avec des snapshot de 3h, 4h et 6h.
- La validité de la base POD reste valable quelques heures après la période de la snapshot. Le pas de temps choisi pour calculer la snapshot POD n'influence pas la qualité des résultats, puisque avec des pas de temps de 100s, 300s et 600s le modèle réduit POD donne des résultats équivalents et sa validité reste jusqu'à deux heures au delà de la période de snapshot. Quelques fortes explosions ont été remarquées avec le pas de temps de 100s.
- L'utilisation des conditions initiales dans le modèle réduit qui varient beaucoup (typiquement de 20%) par rapport à la condition initiale C_0 qui a permis le calcul de la base POD, pourrait donner des erreurs importants sur les concentrations des espèces.
- Pour les différentes bases POD calculées avec des snapshots ayant des conditions initiales différentes, le modèle réduit donne des concentrations très similaires. Cependant pour une variation de 50% et plus le modèle a tendance à donner des résultats totalement éronnés (figure 5.16).
- Le choix de la snapshot est vital puisque la base POD ne reflète que l'information qui provient de cette trajectoire, une perturbation de la trajectoire donne un modèle réduit erroné.

En conclusion, cette approche a un intérêt sur la compréhension de la complexité du système : un nombre réduit de degrés de liberté (3-4) permet d'expliquer la

dynamique chimique. Mais ces degrés de libertés sont locaux dans l'espace des phases et il est exclu d'utiliser cette méthode pour construire un modèle réduit.

Une piste possible serait un choix plus approprié/physique du produit scalaire (que nous avons pris comme le produit scalaire euclidien).

Partie IV

Modélisation de l'incertitude

Chapitre 6

Propagation d'incertitude dans un modèle de Chimie-Transport

Résumé

L'objectif de ce chapitre est d'étudier la propagation d'incertitude, dans le contexte des études d'impact en utilisant un modèle de Chimie-Transport. Nous utilisons à cette fin la méthode DEMM (Deterministic Equivalent Modeling Method).

Nous présentons ici, les fondements théoriques de la méthode, un algorithme d'implémentation et quelques applications à la cinétique chimique du mécanisme EMEP (en 0D) et au modèle POLAIR (en 3D) à l'échelle Européenne. A précision conservée, des gains importants en temps calcul sont effectués, en comparaison d'une méthode classique de type Monte Carlo.

Introduction

Les modèles de Chimie-Transport (CTM) font intervenir de nombreux processus physiques, dynamiques et chimiques. Ces processus utilisent plusieurs données et paramètres qui sont parfois connus avec une forte incertitude. On distinguera deux types d'incertitudes :

1. les incertitudes peuvent être structurelles et sont alors liées essentiellement :
 - aux paramétrisations physiques utilisées (par exemple le schéma convectif utilisé) ;
 - au mécanisme chimique choisi pour décrire les processus chimiques ;
 - aux techniques numériques employées pour la résolution des équations du modèle (algorithmes, résolution).

Ce genre d'incertitudes peut être réduit par l'utilisation de schémas numériques plus efficaces, de paramétrisations physiques améliorées, de schémas chimiques plus complets, et par l'augmentation de la résolution du modèle.

2. Le deuxième type d'incertitude est attribuée aux *valeurs* des données et des paramètres du modèle et provient essentiellement :
 - de la valeur des constantes utilisées dans le modèle ;
 - des données initiales du modèle (concentrations des espèces chimiques) ;
 - des émissions et des sources ;
 - des données de forcages (météorologiques et autres) ;
 - des conditions aux limites.

La réduction de ce type d'incertitude sur les données peut se faire par l'amélioration des systèmes de mesure et par le recours systématique à l'assimilation de données.

Nous allons nous intéresser plus spécifiquement au second type d'incertitude. L'importance des paramètres et leur effet sur les sorties du modèle peuvent être déterminés par une étude de sensibilité ou par une étude de propagation d'incertitude. Dans l'étude de sensibilité on s'intéresse au comportement du modèle par rapport à de petites variations d'un paramètre incertain en maintenant les autres paramètres fixés. L'objet de la propagation de l'incertitude est de quantifier l'incertitude du modèle en prenant en compte l'incertitude de tous les paramètres incertains par la construction d'une densité de probabilité (PDF) des sorties à partir de la PDF des entrées (figure 6.1).

Dans le contexte des études d'impact, nous nous intéresserons à la propagation d'incertitude sur les données dans un CTM.

Il existe plusieurs méthodes pour étudier la propagation d'incertitude. La méthode de Monte Carlo est l'une des méthodes la plus largement utilisée : elle est

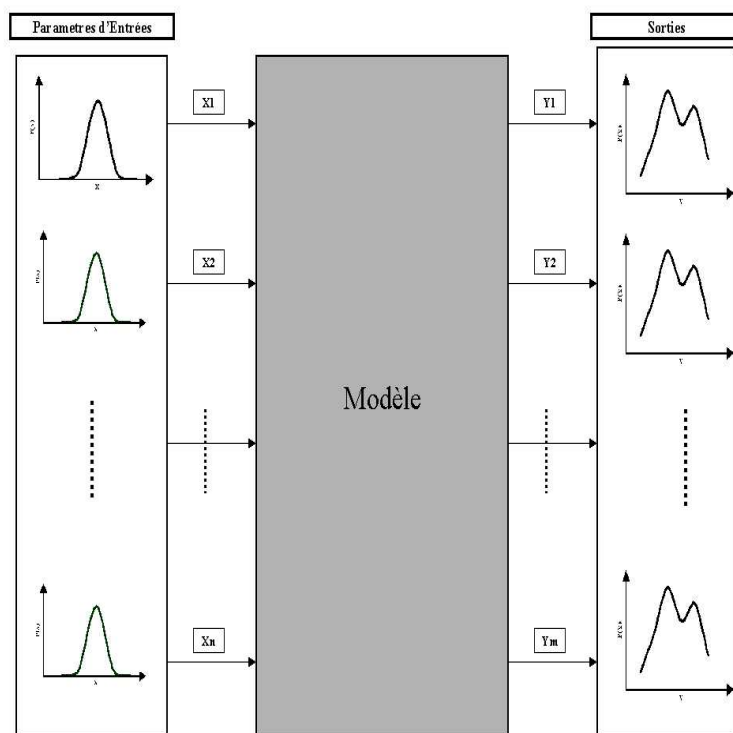


Figure 6.1: Schéma de propagation d'incertitude

basée sur un tirage aléatoire de variables d'entrée à partir de leur distribution et un grand nombre d'exécutions du modèle jusqu'à convergence de la distribution de sortie. Cette méthode demande un grand nombre de points d'entrée et un grand nombre d'exécutions de modèle. Ainsi son application reste limitée à des modèles simples. Dans le cas de modèles CTM complexes et coûteux en temps CPU, les ressources en temps de calcul nécessaires deviennent excessives, et rendent la méthode guère utilisable en l'état de l'art des moyens de calcul actuels. Par exemple un calcul continental annuel avec la chimie EMEP coûte environ 6h CPU avec POLAIR3D (sur une station bi-processeur 1GHZ et 512Mo de mémoire chacun). Pour 5000 exécutions du modèle (nécessaires typiquement pour une méthode de Monte Carlo), on arrive à près de **3 ans et demi de CPU!**

Il existe des méthodes optimisées de Monte Carlo, comme l'approche dite LHS (Latin Hypercube Sampling) qui consiste à améliorer l'échantillonnage de l'espace d'entrée en divisant l'intervalle des valeurs probables de chacune des variables d'entrée en plusieurs segments de probabilité identique et d'effectuer un échantillonnage à partir de ces segments ce qui permet un balayage complet de tous les valeurs même les peu probables. Mais ces méthodes restent encore assez coûteuses ([26]).

Dans ce chapitre, nous étudions des approches de propagation d'incertitude basées sur une utilisation efficace et réduite du nombre de simulations du modèle (contrairement à la méthode de Monte Carlo) : la méthode DEMM (Deterministic Equivalent Modeling Method) et la méthode SRSM (Stochastic Response Surface Method), dont les références sont [73, 74, 26]).

Ces méthodes ont été conçues pour permettre d'analyser la propagation d'incertitude pour les modèles complexes et coûteux avec un coût de calcul très faible. DEMM/SRSM est caractérisée par une représentation de la sortie de modèle incertaine comme un développement en polynômes orthogonaux en fonction de variables d'entrée incertaines et de coefficients inconnus qui peuvent être calculés à partir du résultat du modèle exécuté avec un nombre limité de points de variables d'entrée. Cette méthode fournit un modèle réduit basé uniquement sur un calcul polynômial rendant de fait l'utilisation de simulations Monte Carlo possibles.

Nous présentons dans ce chapitre les principales propriétés de cette approche en comparaison d'une méthode de type Monte Carlo (section 6.1 et 6.2) et notre implémentation algorithmique et un exemple simple d'illustration (section 6.3). La première application concerne un modèle 0D (version 0D de POLAIR) décrivant la cinétique chimique EMEP à 10 espèces et 9 réactions chimiques, en prenant comme variables incertaines les constantes cinétiques de la chimie. Nous analysons l'effet de cette incertitude sur l'incertitude de la sortie de modèle, les concentrations des

espèces chimiques, dans la section 6.4. Une deuxième application est relative à un calcul 3D continental (section 6.5). Les résultats confirment l'efficacité de la méthode à précision conservée.

Enfin nous présentons une approche de couplage de la méthode DEMM/SRSM à la différentiation automatique pour réduire le nombre de simulations du modèle demandé par DEMM/SRSM (section 6.6).

6.1 Contexte méthodologique

6.1.1 Méthodes classiques de propagation d'incertitudes

Les incertitudes dans un modèle sont souvent étudiées par les méthodes d'analyse de sensibilité. Leur objectif est d'estimer le taux de changement dans les sorties de modèle suite aux variations de quelques données d'entrée ou paramètres du modèle. Ces variations sont supposés être petites. Elles permettent de comprendre le comportement du système à modéliser et de déterminer les paramètres et les entrées de modèles qui nécessiteraient une plus grande précision ou une amélioration dans leur traitement. Le choix d'une méthode d'analyse de sensibilité dépend de la mesure de sensibilité utilisée, la précision désirée dans l'estimation de la mesure de sensibilité et la coût calcul demandé ([26]). Dans le tableau 6.1 sont résumées quelques mesures possibles de sensibilité pour un modèle défini par :

$$u = g(k) \tag{6.1}$$

où, k est l'ensemble des m paramètres d'entrée dont nous voulons analyser l'effet sur le modèle et u est le vecteur de n variables de sortie.

Les méthodes traditionnelles pour étudier la sensibilité et la propagation d'incertitude dans un modèle sont typiquement classées en quatre catégories [26, 8] :

1. **'Test de sensibilité' simple :**

Dans cette approche le modèle est utilisé plusieurs fois avec un ensemble de variables d'entrée ou paramètres du modèle différent. Cette méthode est facile à mettre en oeuvre mais malheureusement elle ne donne pas assez d'information sur la propagation d'incertitude.

2. **Méthodes basées sur une stratégie d'échantillonnage :**

Mesure de sensibilité	Définition
Variation arbitraire d'un paramètre	$s = u(\bar{k} + \delta k) - u(k)$
Terme normalisé	$D_i = \frac{\delta u_i}{u_i(\bar{k})}$
Terme moyen	$\overline{u_i(\bar{k})} = \frac{\int_{\bar{k}} \dots \int u_i(\bar{k}) dk}{\int_{\bar{k}} \dots \int dk}$
Espérance	$\langle u_i(k) \rangle = \int_{\bar{k}} \dots \int u_i(k) f(k) dk$
Variance	$\delta_i^2(k) = \langle u_i(k)^2 \rangle - \langle u_i(k) \rangle^2$
Minima- maxima	$\max[u_i(k)], \min[u_i(k)]$
Gradient local	$\delta u \approx [S] \delta k, S_{ij} = \frac{\delta u_i}{\delta k_j}$
Gradient normalisé	$S_{ij}^m = \frac{\bar{k}_j}{u_i(\bar{k})} \frac{\delta u_i}{\delta k_j}$

Table 6.1: Les mesures de sensibilité [26].

Cette approche ne nécessite pas une manipulation du modèle, qui peut être pris comme une boîte noire. Elle demande une exécution du modèle à plusieurs points d'échantillonnage de variables d'entrée. Une des approches les plus utilisées est la méthode FAST (Fourier Amplitude Sensitivity Test) basée sur la transformation de Fourier des paramètres incertains du modèle et la méthode de Monte Carlo. Nous donnerons plus de détails sur l'algorithme de cette dernière dans la section (6.1.3).

3. Calcul de sensibilité locale :

L'objectif est de calculer le gradient des sorties du modèle par rapport aux entrées. On dispose de plusieurs approches possibles pour calculer ce gradient :

(a) Méthode de différenciation automatique :

Ces méthodes travaillent directement sur le code source du modèle écrit généralement avec des langages performants (Fortran, C). Elles ne nécessitent pas d'information sur les équations de modèle ni sur sa structure, elles effectuent des opérations pour générer le "code dérivé" du modèle d'origine appelé 'code direct'. C'est en ce sens une approche "boîte noire". Le calcul des dérivées partielles est fait avec la précision de la machine. Il existe plusieurs logiciels de différenciation automatique : Odysée [27], ADIFOR [75], ...

(b) La différenciation analytique, avec par exemple, l'approche par développement en séries de Neumann appliquée à l'opérateur inverse du modèle et par la méthode de perturbation basée sur le développement de sa série de polynômes en fonction des petites perturbations dans les paramètres d'entrée. Ces méthodes sont utilisées pour des modèles simples et dans le cas de modèles complexe et non linéaire comme les modèles de Chimie-Transport, restent limitées.

(c) La construction des fonctions de Green pour résoudre les équations de sensibilité issues de la différenciation des équations du modèle d'origine.

(d) La méthode C/DDM (Coupled/Decoupled Direct Method) est basée sur la différenciation de l'équation du modèle, ainsi l'équation de sensibilité obtenue est résolue soit d'une manière couplée avec le modèle d'origine ou d'une manière séparée ([76]).

4. Méthode stochastique :

Elle est basée sur la représentation de processus stochastiques où les paramètres incertains sont exprimés en terme d'un développement en série de fonctions orthogonales, en fonction de variables aléatoires. Ce développement est connu sous le nom de 'développement en polynômes de Chaos'. C'est cette approche que nous allons utiliser pour analyser la propagation d'incertitude dans le modèle de Chimie-Transport POLAIR3D. Pour valider cette méthode, nous comparerons les résultats par rapport à l'autre méthode largement utilisée, la méthode de Monte Carlo.

6.1.2 Représentation de l'incertitude

Pour mener à bien une analyse de propagation d'incertitude dans un modèle, nous devons choisir une représentation efficace du paramètre incertain. Il existe plusieurs approches pour représenter l'incertitude dans un modèle. Elles diffèrent d'un domaine d'application à un autre et peuvent mener à des analyses et à des conclusions différentes. L'approche probabiliste est la méthode la plus utilisée dans le domaine des sciences physiques et de l'ingénierie pour caractériser l'incertitude, surtout quand on dispose de la distribution de probabilité des paramètres incertains. Dans cette approche, l'incertitude est associée aux données d'entrée du modèle décrites par la distribution de probabilité et l'objectif dans une étude de propagation d'incertitude est d'estimer la distribution de probabilité de la sortie du modèle. Cela nécessite donc deux étapes :

- **La détermination de la distribution des paramètres d'entrée** du modèle, et l'incorporation d'une variation aléatoire qui peut être due à une variabilité naturelle ou à des erreurs. Ceci est effectué par des estimations statistiques ou par des jugements d'expert. En cas d'estimation statistique, l'estimation de la distribution de probabilité est effectuée à partir des données de mesures disponibles, ou d'un grand nombre d'échantillonnages représentatifs. Des techniques d'estimation de la distribution de probabilité à partir des données sont disponibles dans la littérature ([25]). Par exemple dans le cas des modèles de chimie de transport, l'incertitude sur les émissions et sur les constantes cinétique de la chimie est connue sous forme d'une distribution de type log-normal [25].
- **La propagation d'incertitude à travers le modèle :**

Elle consiste à déterminer les variations détectées sur la sortie de modèle suite aux changements dans les variables d'entrée, qui seront caractérisées par une distribution de probabilité (PDF) qui dépendra de celles des entrées du modèle.

6.1.3 Méthode de Monte Carlo classique

La méthode de Monte Carlo est la plus utilisée pour l'étude de la propagation d'incertitude dans les modèles. Elle sera prise comme méthode de référence pour une comparaison avec l'approche stochastique DEMM/SRSM.

La méthode de Monte Carlo nécessite la connaissance pour chaque paramètre d'entrée incertain de sa distribution de probabilité. Elle est basée sur une génération aléatoire répétitive des variables d'entrée à partir de leur distribution, et plusieurs exécutions du modèle jusqu'à obtention d'une distribution de probabilité de sortie satisfaisante (ie convergée). Cette distribution de probabilité est calculée à partir de l'analyse statistique des différentes sorties de modèles. La méthode est schématisée de la manière suivante :

- **ETAPE 1** : Analyse du modèle et identification des paramètres incertains dans le modèle et de leur distribution.
- **ETAPE 2** : Échantillonnage de distributions des paramètres incertains et génération aléatoire de points à partir des distributions.
- **ETAPE 3** : Exécution du modèle avec le jeu de données choisi aléatoirement de la distribution à l'étape précédente.
- **ETAPE 4** : Sauvegarde du profil de la sortie.
- **ETAPE 5** : Vérification de la convergence de la distribution de sortie.
- **ETAPE 6** : Si la distribution n'est pas convergée, on recommence à partir de l'étape 2, jusqu'à convergence de la distribution.

Le vitesse de convergence de cette méthode est donnée par le contrôle de l'erreur quadratique moyenne proportionnelle à $(1/\sqrt{N})$, où N est le nombre d'exécutions de modèle. Plus précisément, pour un modèle $Y = f(X)$ où X est une variable aléatoire, l'erreur est en $(\sqrt{\text{var}(f(X))}/N)$ et le recours à des méthodes de réduction de variance ([77]) permet d'améliorer la convergence du processus.

6.2 Présentation de la méthode DEMM/SRSM

La méthode DEMM/SRSM est utilisée pour réduire le nombre de simulations effectuées par le modèle pour estimer l'incertitude d'une sortie de modèle à partir de celles des paramètres d'entrée. Elle est basée sur la représentation d'une sortie de modèle incertaine Y comme un développement en série de polynômes orthogonaux de variables aléatoires ξ sous la forme suivante :

$$Y = Y_0 + H_1(\xi) + H_2(\xi) + \dots + H_N(\xi) \quad (6.2)$$

où H_i est un polynôme orthogonal usuel : Legendre, Laguerre, Hermite... Cette formulation est dite parfois développement en "polynôme de chaos".

6.2.1 Représentation des paramètres d'entrée

La méthode DEMM/SRSM définit une représentation des paramètres d'entrée incertains en fonction d'une variable aléatoire standardisée ξ_i . On prend généralement une variable aléatoire Gaussienne de moyenne zéro et de variance égale à 1. Quand les variables incertaines d'entrée X_i sont indépendantes, elles sont directement transformées en fonction de ξ_i . Il existe plusieurs approches de transformation ; dans notre cas, nous utilisons les transformations présentées dans le tableau 6.2 ([26]).

6.2.2 Formulation

Dans le cas général, une sortie de modèle peut être influencée par plusieurs variables d'entrée. Ainsi toute représentation de l'incertitude de la variable de sortie doit prendre en compte l'incertitude de toute les variables d'entrée du modèle qui couple de manière non linéaire, dans notre cas, les variables d'entrée.

Pour un modèle déterministe disposant de n variables d'entrée aléatoires transformées en variables aléatoires standardisés $\{\xi_i\}_{i=1}^n$, sa sortie peut être représentée par le même jeu de termes $\{\xi_i\}_{i=1}^n$ [8]. Nous utilisons une représentation spécifique de la méthode DEMM/SRSM basée sur l'utilisation d'un développement en série de variables aléatoires normales en terme de Polynômes d'Hermite :

Type de distribution	Transformation
Uniforme(a,b)	$a + (b - a)\left(\frac{1}{2} + \frac{1}{2}\text{erf}(\xi/\sqrt{2})\right)$
Normal(μ,σ)	$\mu + \sigma\xi$
Lognormal(μ,σ)	$\exp(\mu + \sigma\xi)$
Gamma (a,b)	$ab\left(\xi\sqrt{\frac{1}{9a}} + 1 - \frac{1}{9a}\right)^3$
Exponentiel(λ)	$-\frac{1}{\lambda}\log\left(\frac{1}{2} + \frac{1}{2}\text{erf}(\xi/\sqrt{2})\right)$
...	...

Table 6.2: Représentation d'une variable aléatoire en fonction d'une variable aléatoire normale ξ .

$$\begin{aligned}
Y &= Y_0 + \sum_{i=1}^n Y_i^{(1)} H_1(\xi_i) + \sum_{i=1}^n Y_i^{(2)} H_2(\xi_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(1)} H_1(\xi_i) H_1(\xi_j) \\
&+ \sum_{i=1}^n Y_i^{(3)} H_3(\xi_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(2)} H_1(\xi_i) H_2(\xi_j) \\
&+ \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(3)} H_2(\xi_i) H_2(\xi_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n Y_{ijk}^{(1)} H_1(\xi_i) H_1(\xi_j) H_1(\xi_k). \\
&+ \dots
\end{aligned} \tag{6.3}$$

où :

- n est le nombre des paramètres d'entrée incertains,
- ξ_i est une variable aléatoire standard de $N(0, 1)$,
- Y est une sortie de modèle,
- $H_m(\xi_i)$ sont les séries de polynômes orthogonaux d'Hermitte de degré m données par la formule de récurrence :

$$H_m(\xi_i) = (-1)^m e^{\xi_i^2} * \frac{d^m}{d\xi_i^m} [e^{-\xi_i^2}] \tag{6.4}$$

En pratique, on rappelle l'expression des premiers polynômes :

$$H_0(\xi_i) = 1, \quad H_1(\xi_i) = \xi_i, \quad H_2(\xi_i) = \xi_i^2 - 1 \tag{6.5}$$

$$H_3(\xi_i) = \xi_i^3 - 3\xi_i, \quad H_4(\xi_i) = \xi_i^4 - 6\xi_i^2 + 3, \quad H_5(\xi_i) = \xi_i^5 - 10\xi_i^3 + 15\xi_i \tag{6.6}$$

- Dans l'équation (6.3) les coefficients $Y_0, Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_{ij}^{(1)}, Y_{ij}^{(2)}, Y_{ij}^{(3)}, Y_{ijk}^{(1)}$ etc ... sont des constantes déterministes à estimer, en exécutant le modèle un certain nombre de fois limité à quelques points d'entrée. Le choix des points d'exécution du modèle est crucial (voir plus loin). Le nombre de paramètres inconnus croît en fonction de l'ordre polynômial pris en compte, comme l'illustre le tableau 6.3. Ce nombre augmente rapidement avec le nombre de paramètres incertains à prendre en compte dans le modèle (figure 6.2) et la méthode reste efficace tant que le nombre de degrés de liberté est pas très élevé.

Remarque :

L'avantage de cette méthode est son applicabilité à un modèle quelconque.

Ordre de l'approximation	Nombre d'inconnues
1	$1+n$
2	$1+2n+\frac{n(n-1)}{2}$
3	$1+3n+\frac{3n(n-1)}{2}+\frac{n(n-1)(n-2)}{6}$

Table 6.3: Nombre de paramètres inconnus en fonction de l'ordre de l'approximation

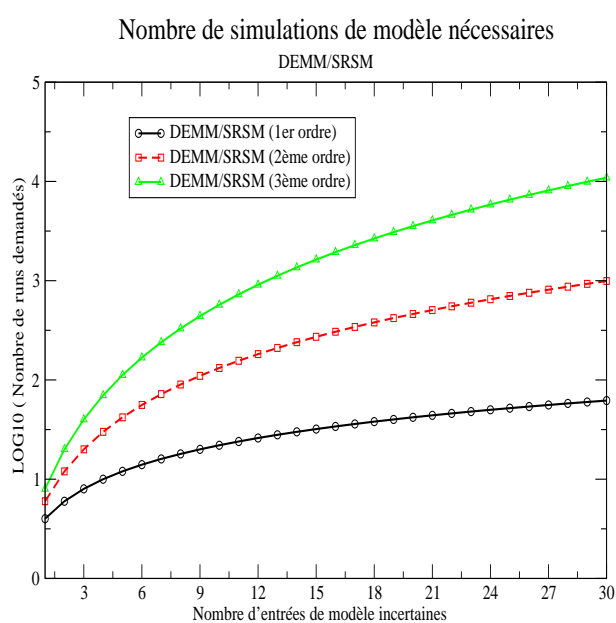


Figure 6.2: Nombre de simulations de modèles demandé par la méthode DEMM (1er, second et 3ème ordre)

Après cette présentation, nous présentons plus en détail notre implémentation algorithmique de cette approche. Le point crucial est notamment le choix des points de collocation.

Remarque : Générateur de variables aléatoires

Il existe plusieurs générateurs de variables aléatoires, différent l'un de l'autre. Dans notre cas, nous utilisons le générateur automatique du langage standard C (la fonction DRAND), qui fournit des variables aléatoires avec une distribution uniforme entre 0 et 1.

6.3 Mise en oeuvre de la méthode DEMM/SRSM

6.3.1 Algorithme

Les différentes étapes pour l'étude de la propagation d'incertitude par la méthode DEMM/SRSM sont illustrées par la figure 6.3 selon l'algorithme suivant :

- **Etape 0** : Analyse du modèle et identification des paramètres incertains dans le modèle.
- **Etape 1** : Spécification des distributions des paramètres incertains prise en compte.
- **Etape 2** : Approximation de la sortie de modèle en fonction des polynômes de chaos à l'ordre N (Equation (6.3)).
- **Etape 3** : Estimations des paramètres $Y_0, Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_{ij}^{(1)}, Y_{ij}^{(2)}, Y_{ij}^{(3)}, Y_{ijk}^{(1)}, \dots$ de l'approximation (Equation (6.3)) en exécutant le modèle d'origine en des points bien choisis (les *points de collocation*). Dans le paragraphe suivant nous présentons les deux méthodes utilisées pour la détermination de ces coefficients.
- **Etape 4** : Vérification de l'erreur de l'approximation en exécutant le modèle d'origine et le modèle réduit à différents points. Si l'erreur est inférieure à la tolérance fixée, l'approximation est supposée bonne, sinon on reprend l'étape 2 en augmentant l'ordre de l'approximation à $N+1$, jusqu'à convergence.
- **Etape 5** : Analyse statistique et calcul de la distribution de la sortie de modèle.

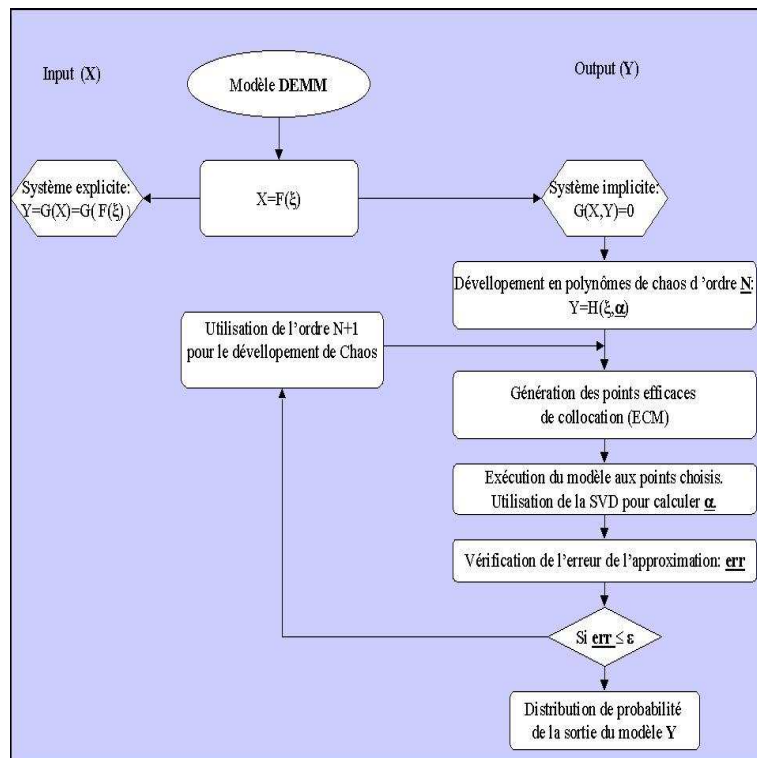


Figure 6.3: Schéma de DEMM/SRSM

6.3.2 Points de collocation

Il existe plusieurs méthodes possibles de détermination des coefficients de l'approximation (6.3) via des choix spécifiques des points de collocation. Nous les détaillons à présent.

6.3.2.1 Méthode ECM

La méthode ECM 'Efficient Collocation method' est basée sur l'utilisation des points de collocations employés dans la méthode de quadrature de Gauss pour calculer les intégrales ([78], [79]). On choisit les variables d'entrée où le modèle sera exécuté à partir des racines du polynôme d'Hermite d'ordre supérieur 'r+1' quand l'approximation est de l'ordre 'r'. Ainsi pour l'ordre 1 nous prenons les racines du polynôme d'Hermite d'ordre 2 (-1,1); pour l'ordre 2, les racines $(-\sqrt{3}, 0, \sqrt{3})$; pour l'ordre 3 nous prenons les racines du polynôme d'Hermite de l'ordre 4 à savoir $(-\sqrt{3+\sqrt{6}}, \sqrt{3+\sqrt{6}}, -\sqrt{3-\sqrt{6}}, \sqrt{3-\sqrt{6}})$, etc ...

Mais, comme on peut s'en rendre compte aisément, le nombre de possibilités de points d'entrée du modèle à partir de ses différents points de collocation est $(r+1)^n$, bien supérieur au nombre d'inconnus! En absence de critère bien défini, le choix d'un jeu de variables d'entrée pour le modèle devient délicat. Pour résoudre ce problème, dans un premier temps, nous avons utilisé la méthode ECM (Efficient Collocation Method, [26]) basée sur les critères suivants :

- On prend tout d'abord, pour le coefficients Y_0 , le point de variables d'entrée correspondant à $\xi_i = 0$ pour toutes les variables d'entrée i de 1 à n , qui correspond au point le plus probable pour une distribution gaussienne de moyenne 0 et de variance 1 (après transformation).
- Pour les coefficients $Y_i^{(1)}$ et $Y_i^{(2)}$ toutes les variables d'entrée sont prises à partir des racines du polynôme d'ordre supérieur (r+1), où r est l'ordre de l'approximation, à l'exception de la variable i où nous prenons la valeur 0. Et on procède de la même façon pour les coefficients $Y_{ij}^{(1)}$, sauf que toutes les variables sont prises égales aux racines à l'exception des variables i et j qui seront prises égales à zéro.
- Pour les termes Y_{ijk}^1 , toutes les variables d'entrée sont calculées à partir des racines du polynôme d'ordre supérieur pour chaque variable i , j et k (si nous allons jusqu'à l'ordre 3 dans l'approximation) et 0 pour les autres entrées. Dans le cas où l'on dispose encore de plus de racines que de points nécessaires pour le calcul des coefficients, nous prenons les racines les plus proches de zéro (le point le plus probable au sens d'une distribution gaussienne).

- Une fois épuisé l'ensemble de ces points, si il y a besoin, nous prenons les points de collocation de telle manière que la distribution globale des points de collocation soit la plus symétrique possible par rapport au point origine. Si après ce stade le nombre de points choisis reste encore inférieur aux coefficients inconnus, nous choisissons les variables d'entrée d'une manière aléatoire à partir des combinaisons possibles restantes.

Nous avons appliqué cette méthode dans un cas académique à une cinétique chimique (section 6.4) et le résultat est très instable.

6.3.2.2 Méthode SRSM

Cette méthode (Stochastic Response Surface Method) est basée sur la même formulation que DEMM sauf qu'elle apporte une solution à la limitation de la méthode DEMM pour le choix des points de collocation. SRSM utilise une méthode de régression en résolvant un système linéaire à travers la méthode SVD (Singular value decomposition). Elle fournit une façon plus robuste pour calculer les coefficients du développement polynomial.

Cette approche est une extension de la méthode efficace de collocation présentée avant. Elle est basée sur un nombre de choix N de jeux de variables d'entrée, choisis de la même façon que dans la méthode ECM, où N est bien supérieur au nombre de coefficients inconnus dans le développement (6.3). D'une certaine manière, on peut s'attendre à que cela "stabilise" notre calcul des coefficients inconnus de (6.3). Il est recommandé que N soit largement supérieur au nombre de coefficients inconnus (typiquement le double, [26]). Le modèle est ensuite évalué en ces différents points choisis et ainsi on obtient un système linéaire à résoudre. Ce système d'équations est résolu par une méthode de type SVD (Singular Value Decomposition).

6.3.3 Un modèle simple

Pour illustrer cette méthode nous prenons le modèle académique suivant :

$$Y = X_1^2 + X_2^3 \quad (6.7)$$

avec deux variables d'entrée incertaines X_1 supposée distribuée uniformément entre 1.0 et 10.0 et X_2 ayant une distribution Gaussienne de moyenne 2.0 et variance 1.0.

Après la transformation respective de X_1 , X_2 en ξ_1 et ξ_2 variables aléatoires normales ($N(0,1)$), la formulation en polynôme de chaos du modèle Y à l'ordre 3 peut s'écrire sous la forme :

$$\begin{aligned} \tilde{Y} = & Y_0 + Y_1\xi_1 + Y_2\xi_2 + Y_3(\xi_1^2 - 1) + Y_4(\xi_2^2 - 1) + Y_5(\xi_1^3 - 3\xi_1) \\ & + Y_6(\xi_2^3 - 3\xi_2) + Y_7\xi_1\xi_2 + Y_8\xi_1(\xi_2^2 - 1) + Y_9\xi_2(\xi_1^2 - 1) \end{aligned} \quad (6.8)$$

Pour estimer les 10 coefficients Y_0, Y_1, \dots, Y_9 ,

- 1. nous choisissons selon le méthode SRSM , $N=2^*10$ points $:(\xi_1^{(1)}, \xi_2^{(1)}), (\xi_1^{(2)}, \xi_2^{(2)}), \dots, (\xi_1^{(N)}, \xi_2^{(N)})$,
- 2. nous exécutons le modèle aux points correspondants : $(X_1^{(1)}, X_2^{(1)}), (X_1^{(2)}, X_2^{(2)}), \dots, (X_1^{(N)}, X_2^{(N)})$, ce qui nous donne un jeu de sorties : $Y^{(1)}, Y^{(2)}, \dots, Y^{(N)}$.
- 3. les coefficients Y_0, Y_1, \dots, Y_9 sont calculés en résolvant le système linéaire :

$$M^T * \vec{Y} = \vec{B} \quad (6.9)$$

avec :

$$M = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ \xi_1^{(1)} & \xi_1^{(2)} & \xi_1^{(3)} & \dots & \xi_1^{(N)} \\ \xi_2^{(1)} & \xi_2^{(2)} & \xi_2^{(3)} & \dots & \xi_2^{(N)} \\ (\xi_1^2)^{(1)} & (\xi_1^2)^{(2)} & (\xi_1^2)^{(3)} & \dots & (\xi_1^2)^{(N)} \\ (\xi_2^2)^{(1)} & (\xi_2^2)^{(2)} & (\xi_2^2)^{(3)} & \dots & (\xi_2^2)^{(N)} \\ (\xi_1^3)^{(1)} & (\xi_1^3)^{(2)} & (\xi_1^3)^{(3)} & \dots & (\xi_1^3)^{(N)} \\ (\xi_2^3)^{(1)} & (\xi_2^3)^{(2)} & (\xi_2^3)^{(3)} & \dots & (\xi_2^3)^{(N)} \\ (\xi_1\xi_2)^{(1)} & (\xi_1\xi_2)^{(2)} & (\xi_1\xi_2)^{(3)} & \dots & (\xi_1\xi_2)^{(N)} \\ (\xi_1(\xi_2^2 - 1))^{(1)} & (\xi_1(\xi_2^2 - 1))^{(2)} & (\xi_1(\xi_2^2 - 1))^{(3)} & \dots & (\xi_1(\xi_2^2 - 1))^{(N)} \\ (\xi_2(\xi_1^2 - 1))^{(1)} & (\xi_2(\xi_1^2 - 1))^{(2)} & (\xi_2(\xi_1^2 - 1))^{(3)} & \dots & (\xi_2(\xi_1^2 - 1))^{(N)} \end{pmatrix}$$

et M^T , est la matrice transposée de M . De plus :

$$\vec{B} = \begin{pmatrix} \tilde{Y}^{(1)} \\ \tilde{Y}^{(2)} \\ \tilde{Y}^{(3)} \\ \vdots \\ \tilde{Y}^{(N)} \end{pmatrix}, \quad \vec{Y} = \begin{pmatrix} Y_0 \\ Y_1 \\ Y_2 \\ \vdots \\ Y_9 \end{pmatrix} \quad (6.10)$$

Cette méthode nous a permis d'obtenir les valeurs suivantes :

$$Y_0 = 96.0, \quad Y_1 = 18.0, \quad Y_2 = 15.0, \quad Y_3 = 81.0, \quad Y_4 = 6.0$$

$$Y_5 = 0.0, \quad Y_6 = 1.0, \quad Y_7 = 0.0, \quad Y_8 = 0.0, \quad Y_9 = 0.0$$

Commentons rapidement ces résultats. Les valeurs nulles de Y_5 , Y_7 , Y_8 et Y_9 , montrent bien :

- l'effet nul de la première variable d'entrée X_1 à l'ordre 3 sur la sortie.
- l'indépendance entre la variable X_1 et X_2 au premier ordre et au deuxième ordre dans notre modèle ($Y = X_1^2 + X_2^3$).

Ces valeurs de coefficients donnent bien, après une simple vérification, la formulation initiale de notre modèle.

6.4 Application à un modèle 0D

Dans cette section nous présentons les résultats d'application de cette méthode DEMM/SRSM au modèle POLAIR0D .

6.4.1 Modèle

Le modèle pris pour procéder à la validation la méthode DEMM/SRSM est la version 0D du modèle POLAIR avec une chimie EMEP à 10 espèces et 9 réactions (Tableau 6.4).

	Réaction	
$\text{NO} + \text{O}_3$	$\xrightarrow{k_1}$	$\text{NO}_2 + \text{O}_2$
$\text{NO}_2 + \text{h}\nu$	$\xrightarrow{k_2}$	$\text{NO} + \text{O}$
$\text{NO}_2 + \text{OH}$	$\xrightarrow{k_3}$	HNO_3
$2\text{NO}_2 + \text{O}_3 + \text{H}_2\text{O}$	$\xrightarrow{k_4}$	$2\text{NO}_3^- + 2\text{H}^+$
$\text{NO}_2 + \text{CH}_3\text{COO}_2(\text{ROH})$	$\xrightarrow{k_5}$	$\text{CH}_3\text{COO}_2\text{NO}_2$ (PAN)
PAN	$\xrightarrow{k_6}$	$\text{NO}_2 + \text{ROH}$
HNO_3	$\xrightarrow{k_7}$	NO_3^-
NO_3^-	$\xrightarrow{k_8}$	HNO_3
SO_2	$\xrightarrow{k_9}$	SO_4^{2-}

Table 6.4: Schéma chimique EMEP

Pour notre test académique, nous avons supposé une incertitude sur les constantes cinétiques k_3 , k_5 , k_7 et k_8 avec une distribution gaussienne d'écart type de 20% selon la figure 6.4. Les sorties de modèle supposées incertaines sont les concentrations d'espèces chimiques du schéma. Le modèle POLAIR0D est intégré pendant

2h30min avec un pas de temps de 900s.

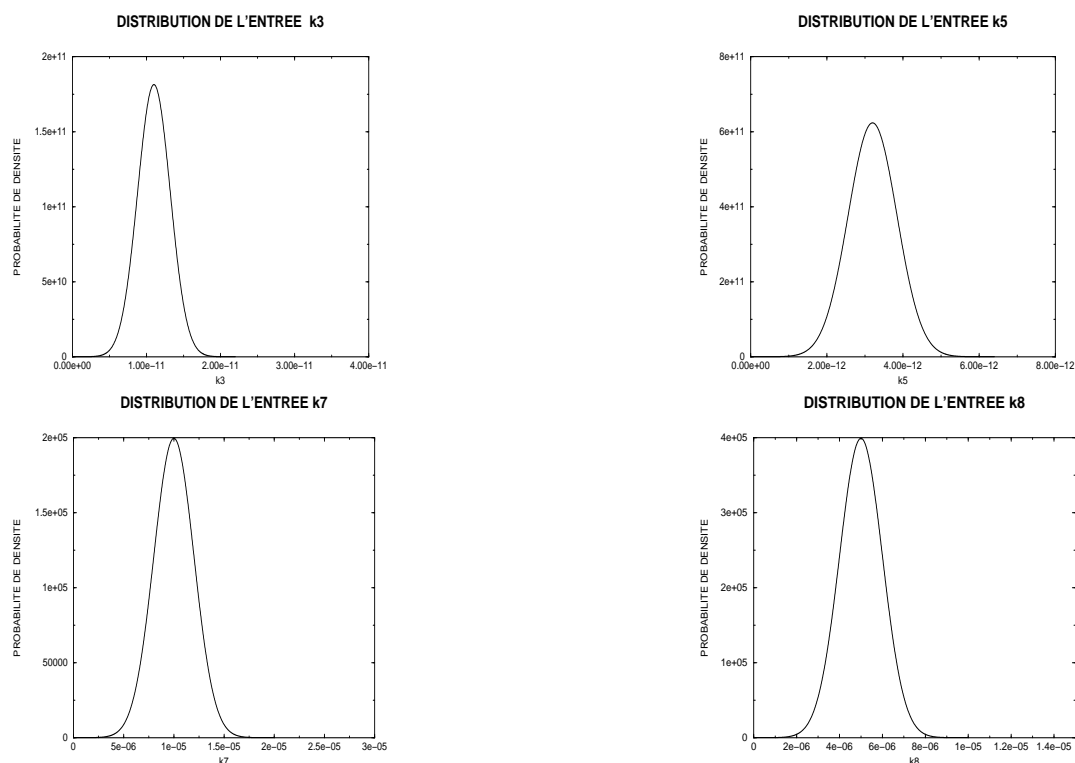


Figure 6.4: Distribution des paramètres incertains k_3, k_5, k_7 et k_8

6.4.2 Résultats

Nous présentons dans la figure 6.5 la comparaison des résultats obtenus d'une part avec DEMM et la méthode de collocation SRSM avec ceux obtenus d'autre part par simulation de type Monte Carlo. Le nombre de runs nécessaire à la convergence est $N=10000$.

La méthode DEMM/SRSM apparaît comme suffisamment précise dès le second ordre.

Notons de plus que l'utilisation de la méthode DEMM/SRSM permet d'utiliser un nombre limité de simulations de modèle initial. Le tableau 6.5 montre le gain obtenu.

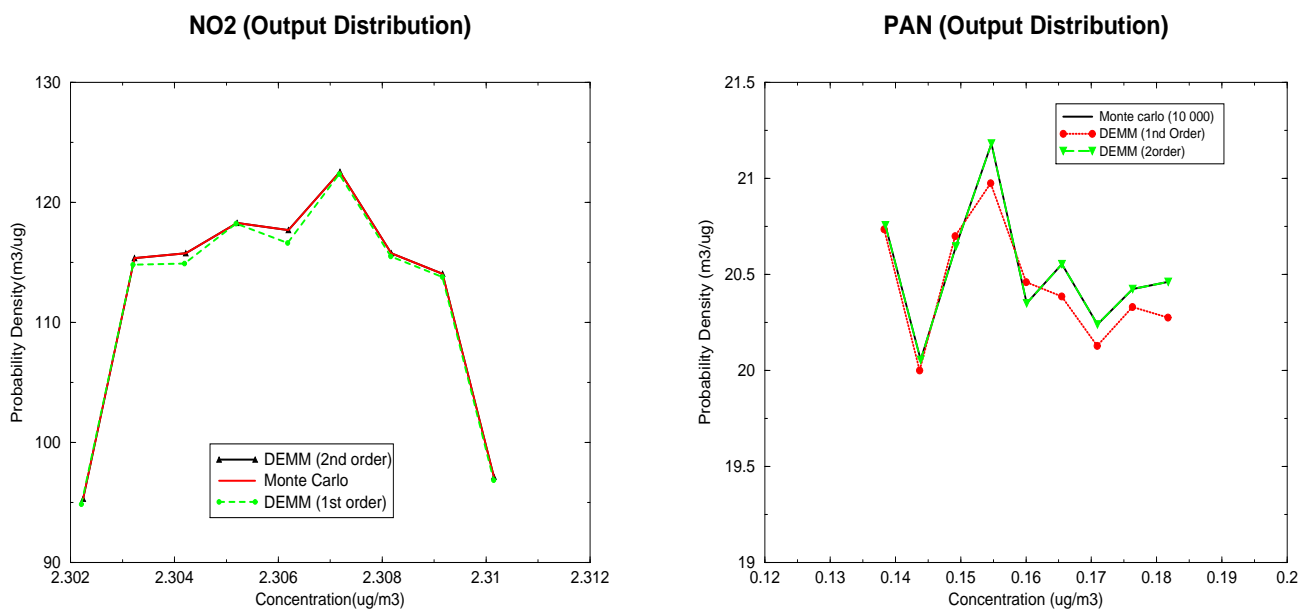


Figure 6.5: Comparaison des distributions de probabilité obtenues avec DEMM/SRSM et Monte Carlo

Ordre de l'approximation	DEMM/SRSM	Monte Carlo
1	$2(1+n)=10$	N=10000
2	$2 + 3n + n^2 = 30$	N=10000
3	$2+3n+3n^2+\frac{n(n-1)(n-2)}{3} =70$	N=10000

Table 6.5: Coût calcul (nombre d'évaluations du modèle initial)

6.5 Application à un modèle 3D sur l'Europe dans le cadre d'une étude d'impact

Dans cette partie nous appliquons la méthode DEMM-SRSM au modèle POLAIR3D décrit dans la première partie, avec une chimie EMEP. Le but est de voir l'effet des incertitudes des émissions sur la qualité des résultats, en d'autre terme d'étudier comment une incertitude sur les émissions se propage jusqu'aux concentrations des espèces chimiques à travers le modèle. Les paramètres de sortie auxquels nous nous sommes intéressés sont les concentrations au sol des espèces NO, NO_2 , SO_2 , Nitrate d'ammonium et Sulfate d'ammonium moyennées temporellement pendant la durée de l'intégration (1 jour) et spatialement sur toute l'Europe.

Les paramètres incertains d'entrée sont les émissions de NO_x , SO_2 , et NH_3 , émissions du réseau EMEP sur l'Europe. Ces émissions sont issues des bilans annuels de chaque pays européen. Leur utilisation dans le modèle est faite après un travail d'interpolation spatiale, temporelle et une agrégation en espèce modèle (Chapitre 3). L'inventaire des émissions utilisé dans cette application est résumé dans le Tableau 6.7.

L'incertitude est prise égale pour tous les pays qui figurent dans la base EMEP, mais différente pour chacune des espèces émises, selon :

$$E_i(w) = \alpha(w)e_i(w) \quad (6.11)$$

où w désigne l'espèce chimique émise, $e_i(w)$ la valeur d'émission de l'espèce w en un point $i = (x, y)$ du domaine et $\alpha(w)$ est le paramètre supposé incertain suivant une distribution normale de moyenne 1 et d'écart type de 20%.

Nous avons supposé une incertitude constante sur tous les pays européens émetteurs.

Espèce émise incertaine	Type de distribution	écart type
NO_x	Normale	20%
SO_2	Normale	20%
NH_3	Normale	20%

Table 6.6: Incertitude sur les variables d'entrée.

La sortie du module DEMM/SRSM est un ensemble de N (N typiquement de l'ordre de 10000) valeurs pour chacune de variables de sortie en chaque point de maillage. A partir de ces valeurs et pour chacune des sorties on peut calculer les probabilités de densités (PDF) et les mesures statistiques usuelles. Pour une sortie

Indice	Pays	NO_x	SO_2	NH_3
1	Albania	24000.02	71999.94	31999.99
2	Remaining Asian areas	211999.96	869000.05	303000.17
3	Austria	201800.03	90740.0	79860.03
7	Belarus	284999.99	636999.98	142000.05
5	Belgium	320580.01	357000.01	107300.
4	Bosnia and Herzegovina	80000.03	480000.	31000.01
6	Bulgaria	361000.02	2008000.02	143999.91
18	Croatia	87599.99	179999.98	37100.01
9	Czech Republic	742000.	1875999.99	156000.02
11	Denmark	276930.02	180617.02	132189.04
12	Estonia	67700.	252099.99	24250.01
14	Finland	300000.01	259999.94	37999.98
15	France	1898999.96	1340999.98	763000.05
10	Germany	2705999.95	5321000.06	764999.99
17	Greece	310999.99	479000.02	79000.
19	Hungary	238000.02	1009999.96	124000.03
20	Ireland	118100.01	185699.98	112400.
21	Italy	1937999.98	1651000.02	465999.99
23	Latvia	92280.	119244.01	43850.
22	Lithuania	158000.05	221999.98	84000.04
25	Netherlands	573799.99	202349.99	226749.98
28	North Africa	95999.84	413000.61	234999.61
27	Norway	226449.64	52544.85	22727.18
29	Poland	1279999.99	3210000.01	511999.99
30	Portugal	317000.01	359400.	104600.01
24	Republic of Moldova	99999.98	265000.	49000.01
31	Romania	545999.99	1311000.07	300000.04
32	Russian Federation	3599999.35	4671000.89	1190999.64
35	Slovakia	215000.	542000.01	63000.02
34	Slovenia	63000.	196000.01	24000.02
13	Spain	1279000.05	2167000.03	471999.91
33	Sweden	348850.	111129.97	50999.89
8	Switzerland	153690.	41959.96	71500.
25	The former Yugoslav Republic	30400.02	104999.99	16999.98
36	Turkey	643660.02	1590019.99	321000.05
37	Ukraine	1097000.01	3782000.	729000.04
16	United Kingdom	2763000.01	3720999.93	341000.24

Table 6.7: Inventaire d'émission de NO_x , SO_2 et NH_3 en tonne par an (1990).

Y au point (x,y,z), nous disposons de N valeurs issues de DEMM/SRSM (ou Monte Carlo) et on peut calculer les mesures suivantes :

$$\left\{ \begin{array}{l} \bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i \\ \min_y = \min\{Y_i\}_{i=1,N} \\ \max_y = \max\{Y_i\}_{i=1,N} \\ \sigma_y = \sqrt{VAR(Y)} \quad , \quad VAR(Y) = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2 \end{array} \right. \quad (6.12)$$

Nous présentons ici deux résultats :

- le premier (**Simulation 1**) est relatif à une simulation DEMM/SRSM où nous nous sommes intéressés à la moyenne spatiale (Europe) et temporelle (pendant 6 heures d'intégration) pour chacune des concentrations des espèces NO, NO₂, SO₂, Nitrate d'ammonium et Sulfate d'ammonium.
- le deuxième résultat est associé à une simulation DEMM/SRSM (**Simulation 2**) ayant pour variables de sortie les concentrations moyennes journalières des 37 pays européens (6.7). Dans les deux cas la comparaison des résultats est effectuée par rapport à l'approche classique de propagation d'incertitude Monte Carlo avec N=10000 simulations.

- **Simulation 1** La pdf correspondante à la variable moyenne spatio-temporelle pour chacune des concentrations (NO, NO₂, SO₂, Nitrate et Sulfate) est présentée sur la Figure 6.6, en comparaison avec les résultats obtenus avec la méthode Monte Carlo. Les PDFs sont très semblables pour les deux méthodes. La méthode DEMM reproduit très bien les résultats obtenus avec la méthode de Monte Carlo qui est très coûteuse.

- **Simulation 2** Nous présentons ici les résultats statistiques obtenus avec les deux approches DEMM/SRSM (1er et 2ème ordre) et la méthode Monte Carlo. Pour chaque pays européen et pour chacune des mesures statistiques (*mes*) définies par L'équation (6.12), nous avons calculer l'erreur relative de ces mesures ($\frac{|mes_{DEMM} - mes_{Mcarlo}|}{mes_{Mcarlo}} * 100$) obtenues avec DEMM/SRSM (1er et 2ème ordre) par rapport à celles obtenues avec la méthode de Monte Carlo. Dans les figures 6.7, 6.8, 6.9 et 6.10 nous présentons l'erreur relative (en Logarithme décimal) respectivement pour le minimum, le maximum, moyenne et l'écart type obtenus pour les quatre espèces NO₂, SO₂, NITRAM

et SULFAM. Nous constatons que la méthode DEMM/SRSM (1^{er} ordre) fournit de très bons résultats puisque les erreurs relatives en général sont très faibles et atteignent rarement les 10% d'erreur, et ceci pour toutes les mesures calculées. Quant au deuxième ordre de DEMM/SRSM, l'erreur relative des différentes mesures est très faible et n'atteint pas les 1% d'erreur. Ce résultat est confirmé aussi par le tableau 6.9 qui compare la moyenne sur les 37 pays européens des différentes mesures statistiques, obtenue avec DEMM/SRSM (1^{er} et 2^{ème} ordre) et Monte Carlo.

Nous donnons un autre exemple de sortie de la méthode DEMM/SRSM (second ordre) dans les Tableaux 6.10 et 6.11.

L'utilisation de la méthode DEMM/SRSM a donné de très bon résultats par rapport à la méthode classique de Monte Carlo, avec un gain impressionnant en temps de calcul. Les temps CPU demandés par DEMM/SRSM (1^{er} et 2^{ème} ordre) et la méthode Monte Carlo sont présentés dans le tableau 6.8. Le speedup obtenu avec le 2^{ème} ordre est de l'ordre de 262, et de plus de 10^3 pour le premier ordre de DEMM/SRSM.

DEMM/SRSM 1 ^{er} ordre	DEMM/SRSM 2 ^{ème} ordre	Monte Carlo (10000 Runs)
343 s	1355 s	355242

Table 6.8: Comparaison du temps CPU pour la méthode DEMM/SRSM (1^{er} ordre et 2^{ème} ordre) versus méthode classique de Monte Carlo (10000 simulations).

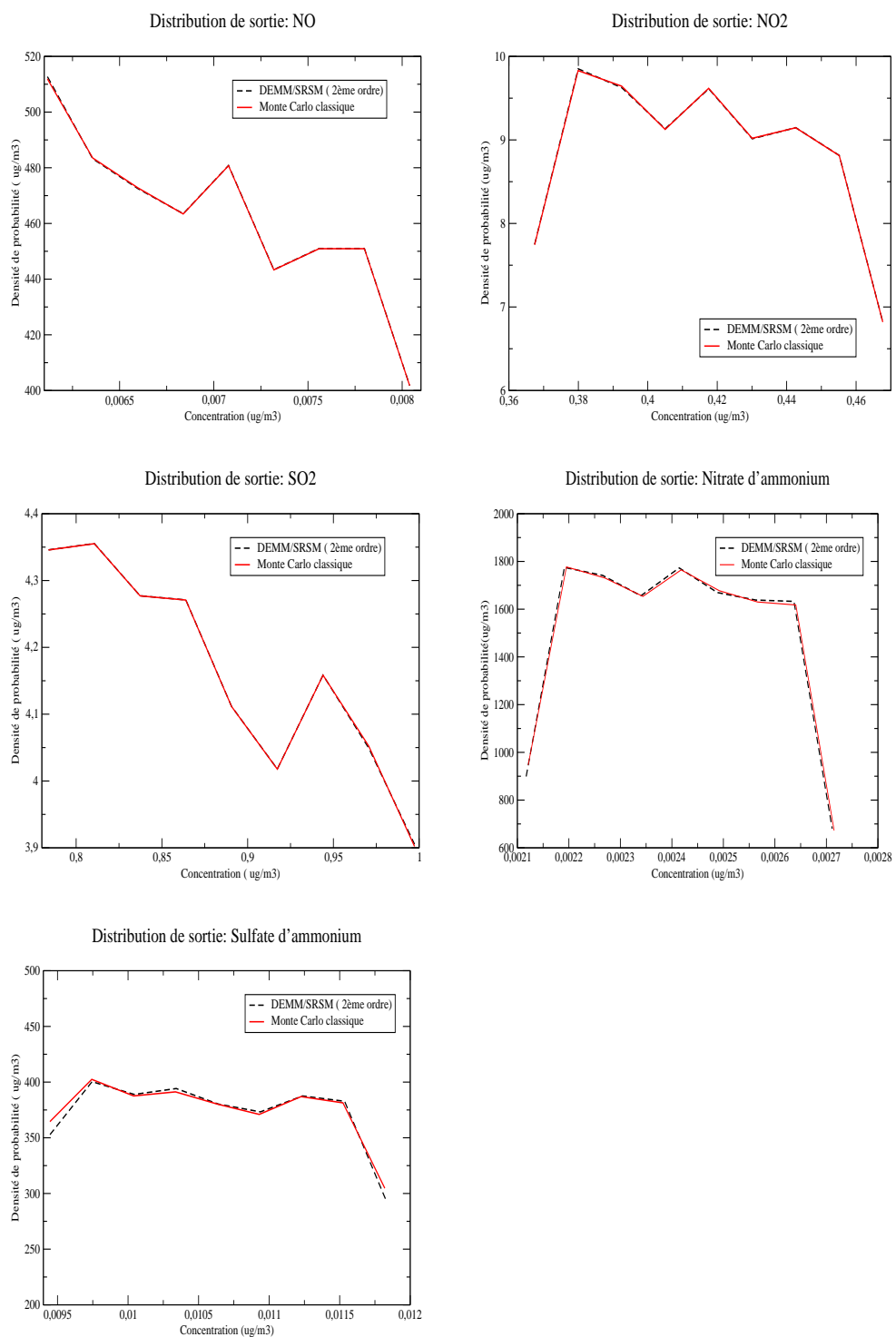


Figure 6.6: Comparaison des distributions de sortie obtenues avec DEMM/SRSM et Monte Carlo classique.

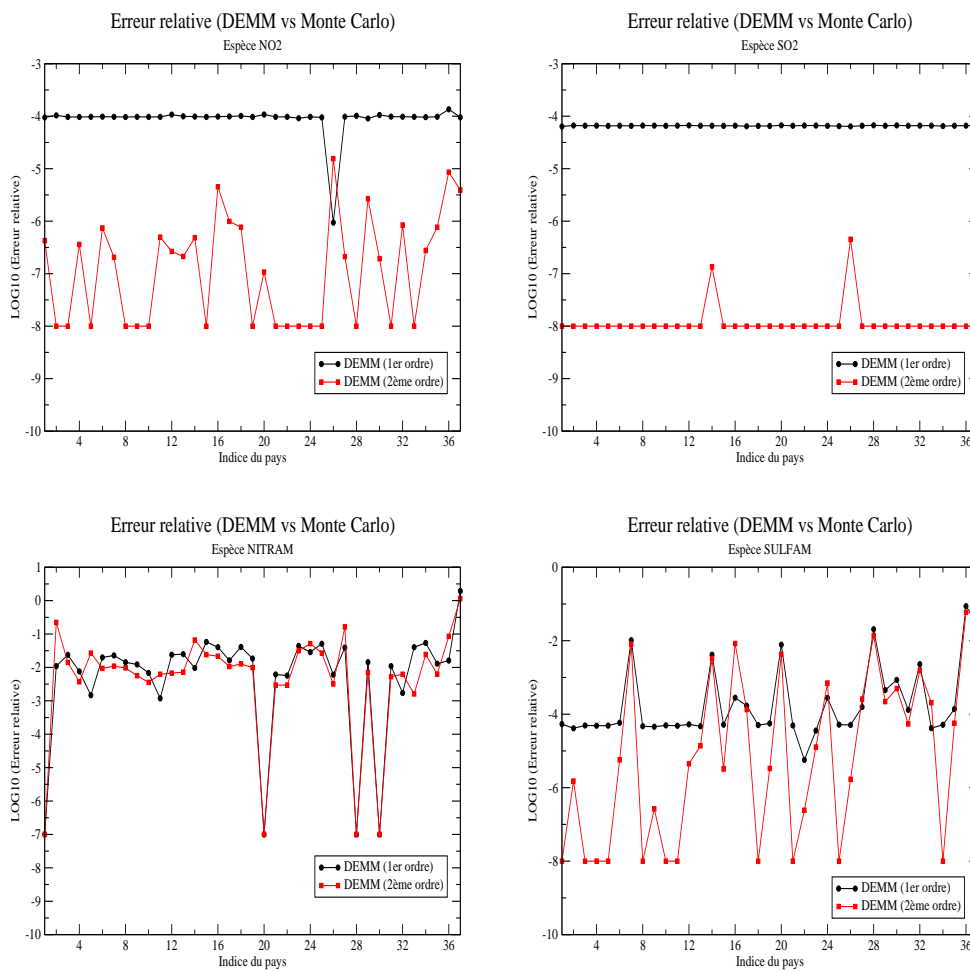


Figure 6.7: **Minimum.** Erreur relative associée à la méthode DEMM/SRSM (1^{er} ordre, 2^{ème} ordre) versus Monte Carlo classique (10000 runs).

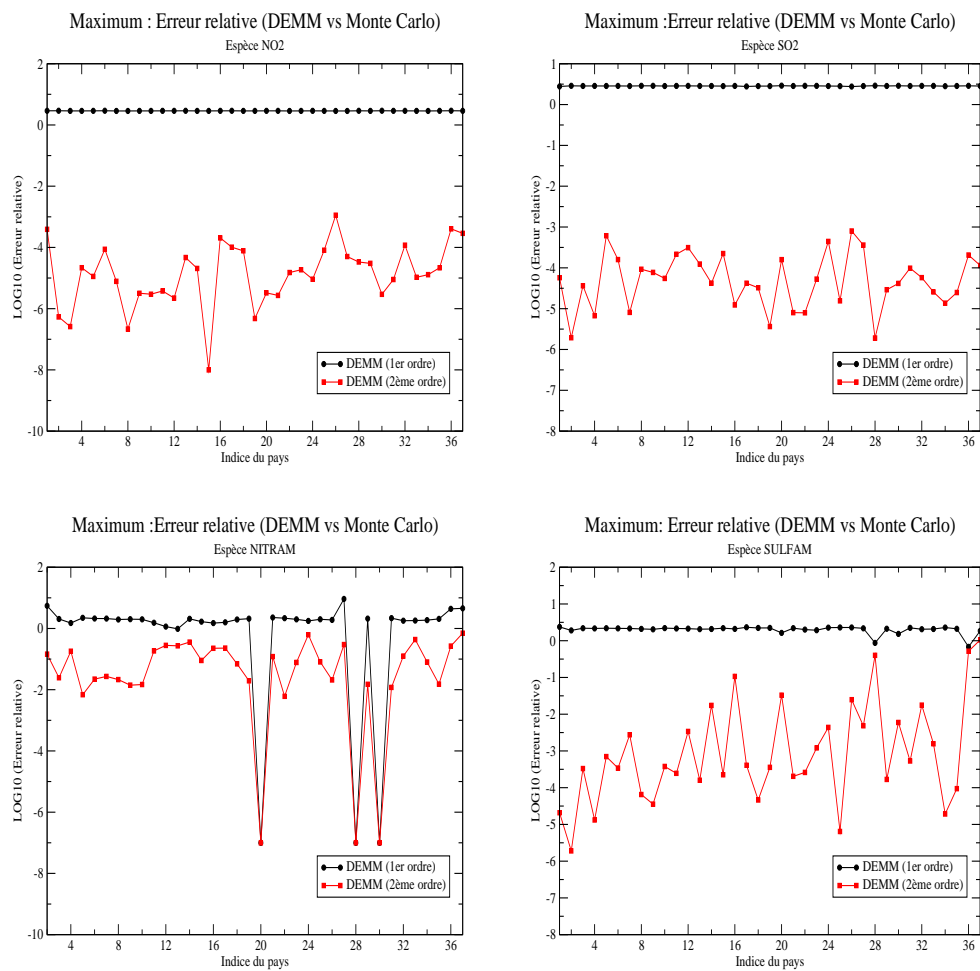


Figure 6.8: **Maximum.** Erreur relative associée à la méthode DEMM/SRSM (1^{er} ordre, 2^{ème} ordre) versus Monte Carlo classique (10000 runs).

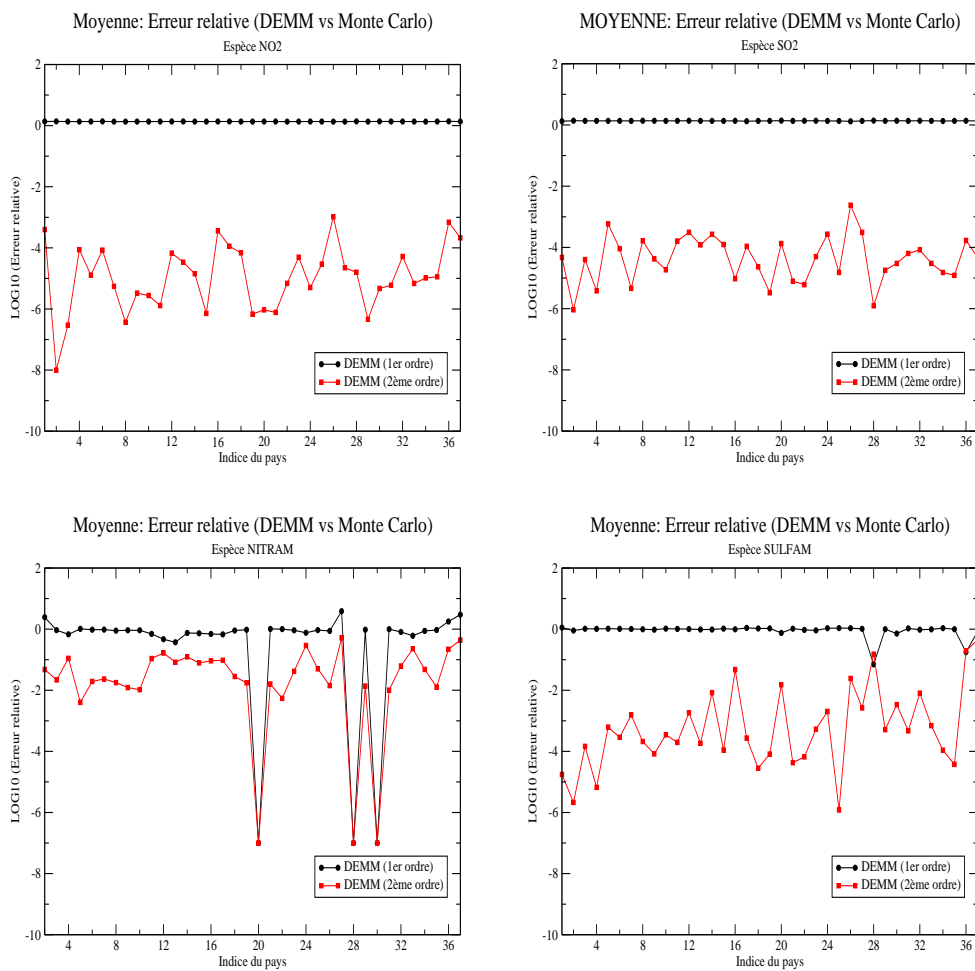


Figure 6.9: **Moyenne.** Erreur relative associée à la méthode DEMM/SRSM (1^{er} ordre, 2^{ème} ordre) versus Monte Carlo classique (10000 runs).

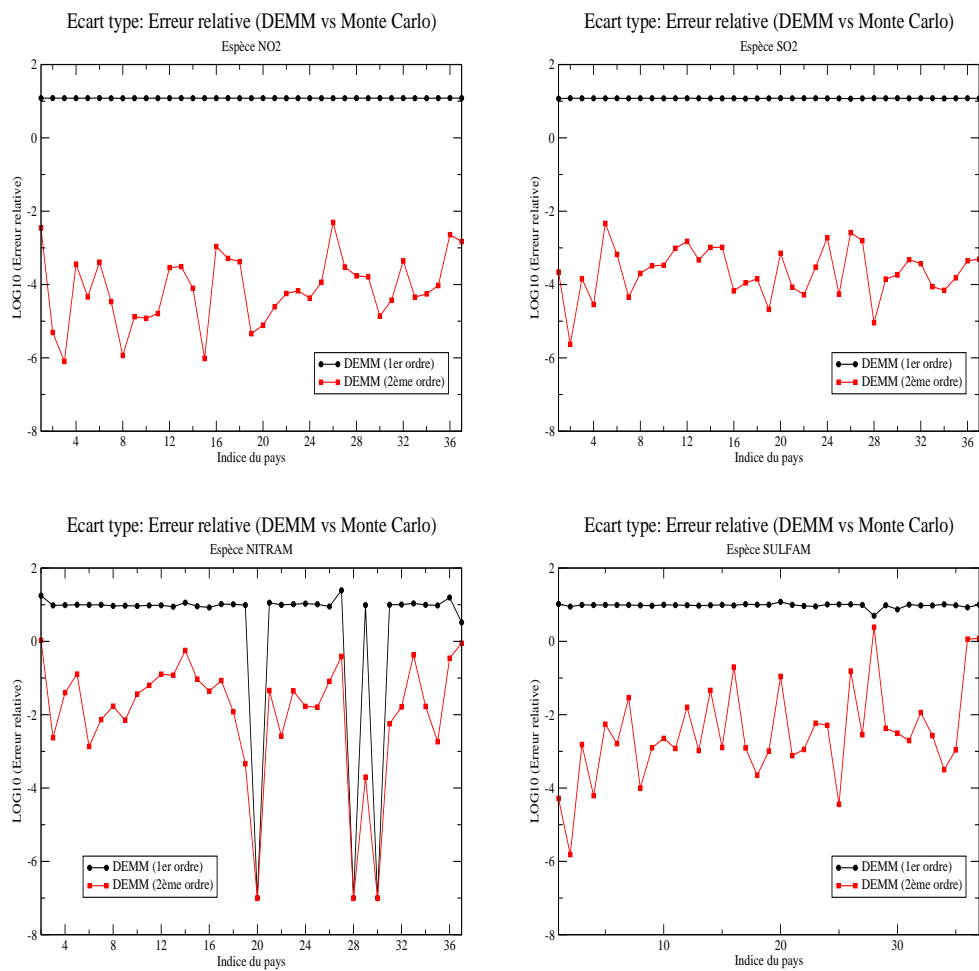


Figure 6.10: **Ecart type.** Erreur relative associée à la méthode DEMM/SRSM (1^{er} ordre, 2^{ème} ordre) versus Monte Carlo classique (10000 runs).

Mesure statistique	DEMM (1er ordre)	DEMM (2ème ordre)	Monte Carlo (10000 runs)
Min			
NO2	7.86714956	7.86715653	7.86715665
SO2	12.3424447	12.3424528	12.3424528
NITRAM	0.0470433891	0.0469806707	0.0469987725
SULFAM	0.13994733	0.139950552	0.139943812
Max			
NO2	10.0278916	10.3254829	10.3254722
SO2	15.7340748	16.195702	16.195698
NITRAM	0.059641833	0.0607311295	0.0606694665
SULFAM	0.175752402	0.179725325	0.179583968
Moyenne			
NO2	8.94276381	9.06619905	9.06618886
SO2	14.0471433	14.2407213	14.2407207
NITRAM	0.0533173907	0.0537177881	0.0537086085
SULFAM	0.15794331	0.159597478	0.159543724
Ecart type			
NO2	0.241696319	0.245032407	0.245032131
SO2	0.379652522	0.384884359	0.384884344
NITRAM	0.00144101056	0.00145183211	0.00145158401
SULFAM	0.0042687381	0.00431344535	0.00431199253

Table 6.9: Comparaison de la moyenne des différentes mesures statistiques (le minimum, le maximum, la moyenne et l'écart type) sur les 37 pays européen pris en compte, entre DEMM/SRSM (1er ordre , 2ème ordre) et Monte Carlo à 10000 simulations (**Moyenne sur les 37 pays européens**).

	NO_2				SO_2				SULFATE			
	min_y	max_y	\bar{Y}	σ_y	min_y	max_y	\bar{Y}	σ_y	min_y	max_y	\bar{Y}	σ_y
1	0.0937382	0.1230781	0.1080477	0.0080575	0.6984450	0.9141987	0.8047747	0.0623057	0.0000000	0.0000063	0.0000019	0.0000015
2	1.4158997	1.8599621	1.6324639	0.1214882	4.7097422	6.1859462	5.4368764	0.4263008	0.0067360	0.0097398	0.0081627	0.0006896
3	5.8655998	7.6958274	6.7583331	0.5019745	6.4073256	8.4084377	7.3931395	0.5778848	0.0347479	0.0445506	0.0395586	0.0026792
4	2.7741929	3.6408428	3.1969021	0.2377400	29.5170589	38.7386304	34.0598647	2.6630221	0.0208702	0.0272959	0.0240199	0.0016936
5	10.7843535	14.1550620	12.4283886	0.9245673	11.3192618	14.8510171	13.0591787	1.0199067	0.0467211	0.0600628	0.0532309	0.0035692
6	13.4909125	17.7197317	15.5532664	1.1600708	41.3836839	54.3141002	47.7535402	3.7340689	0.1365993	0.1756117	0.1557410	0.0106600
7	4.8818003	6.4042468	5.6244221	0.4174537	13.1508161	17.2493820	15.1700624	1.1835913	0.0371455	0.0478179	0.0423831	0.0029129
8	7.0384268	9.2330596	8.1089340	0.6019151	1.0546939	1.3849233	1.2173607	0.0953643	0.0312762	0.0399673	0.0355446	0.0023746
9	26.3311770	34.5546220	30.3422127	2.2553175	33.4069343	43.8668943	38.5593948	3.0206460	0.1471316	0.1881798	0.1672741	0.0112266
10	28.2795942	37.1048551	32.5842838	2.4205876	44.2336175	58.0321787	51.0315407	3.9847735	0.1422350	0.1815787	0.1615317	0.0107744
11	8.0688860	10.5937690	9.3003247	0.6926554	0.5318144	0.6979863	0.6136752	0.0479874	0.0453348	0.0590287	0.0520423	0.0036056
12	3.7719714	4.9511468	4.3471008	0.3234183	1.1969070	1.5714408	1.3814016	0.1081585	0.0225908	0.0303769	0.0264061	0.0019537
13	4.6953585	6.1657498	5.4124817	0.4031666	7.6920918	10.0945301	8.8756141	0.6937805	0.0167410	0.0226899	0.0196437	0.0014425
14	2.0734283	2.7208700	2.3892243	0.1775044	0.7451161	0.9776429	0.8596704	0.0671495	0.0074067	0.0104538	0.0088732	0.0007148
15	12.1141706	15.8984461	13.9599541	1.0382255	6.2509762	8.1989947	7.2107130	0.5625523	0.0544479	0.0698100	0.0619828	0.0041288
16	31.0304464	40.7373458	35.7647714	2.6614171	51.9033274	68.1293190	59.8965218	4.6857710	0.1038270	0.1325418	0.1178742	0.0076269
17	1.0111889	1.3280676	1.1657301	0.0868575	1.5257042	1.9969882	1.7579668	0.1360983	0.0075484	0.0100883	0.0087869	0.0006453

Table 6.10: Sortie DEMM/SRSM (2ème ordre) pour les 37 pays Européens

	NO_2				SO_2				SULFATE			
	min_y	max_y	\bar{Y}	σ_y	min_y	max_y	\bar{Y}	σ_y	min_y	max_y	\bar{Y}	σ_y
18	3.9075195	5.1265870	4.5021486	0.3343838	10.0960398	13.2408027	11.6454051	0.9081504	0.0273418	0.0355424	0.0313478	0.0021902
19	6.3818854	8.3756075	7.3543237	0.5469043	10.4549804	13.7125119	12.0598875	0.9407159	0.0492933	0.0633054	0.0561695	0.0038288
20	0.9297226	1.2204201	1.0715094	0.0793882	0.7039850	0.9254080	0.8130381	0.0639429	0.0000000	0.0000000	0.0000000	0.0000000
21	11.2414852	14.7584425	12.9567933	0.9647010	2.9524289	3.8741129	3.4064887	0.2661656	0.0653477	0.0869569	0.0757880	0.0054839
22	10.0803151	13.2244556	11.6139603	0.8622005	6.5417558	8.5900349	7.5507152	0.5915057	0.0659114	0.0845966	0.0750648	0.0051183
23	6.0010819	7.8736250	6.9144551	0.5135447	2.3181839	3.0435163	2.6754832	0.2094628	0.0373338	0.0485648	0.0428218	0.0029981
24	4.6619000	6.1193371	5.3727471	0.3997652	13.1419809	17.2381231	15.1600310	1.1828914	0.0405092	0.0537777	0.0469670	0.0033801
25	1.7414634	2.2852420	2.0066951	0.1492591	14.2301420	18.6566827	16.4111209	1.2783045	0.0235456	0.0305469	0.0269799	0.0018936
26	21.2865053	27.9317879	24.5279327	1.8230090	4.4363300	5.8057173	5.1112647	0.3954543	0.1060206	0.1349957	0.1202383	0.0079369
27	0.4735901	0.6212449	0.5456141	0.0404771	0.1124577	0.1475376	0.1297401	0.0101305	0.0005416	0.0008817	0.0006987	0.0000728
28	0.0539968	0.0709209	0.0622507	0.0046333	0.2783024	0.3657893	0.3213913	0.0252646	0.0000000	0.0000000	0.0000000	0.0000000
29	18.7228887	24.5716870	21.5756463	1.6043607	32.7076641	42.9207424	37.7390102	2.9493512	0.1367841	0.1755166	0.1557837	0.0105915
30	0.5175669	0.6798579	0.5967153	0.0443698	0.2108375	0.2769833	0.2434179	0.0191017	0.0000000	0.0000000	0.0000000	0.0000000
31	10.1912168	13.3802042	11.7465537	0.8748193	29.0490835	38.1138036	33.5148072	2.6177263	0.0990426	0.1272842	0.1128840	0.0077303
32	3.5497387	4.6601376	4.0913137	0.3043169	4.5632709	5.9910806	5.2666104	0.4123254	0.0237492	0.0310258	0.0273022	0.0019130
33	2.1580684	2.8323279	2.4869378	0.1849053	0.4783787	0.6279897	0.5520785	0.0432049	0.0064744	0.0093103	0.0078444	0.0006529
34	3.6084731	4.7333710	4.1571840	0.3086339	2.8517872	3.7382700	3.2885728	0.2560001	0.0277350	0.0358953	0.0317393	0.0022032
35	13.0325836	17.1078371	15.0202237	1.1177114	34.6337482	45.4504043	39.9623985	3.1236535	0.0906669	0.1161447	0.1031639	0.0069704
36	1.1647981	1.5304077	1.3430958	0.1001067	2.4192456	3.1769228	2.7924668	0.2188034	0.0027508	0.0041688	0.0034260	0.0003138
37	7.6588470	10.0526847	8.8264229	0.6566041	28.7626352	37.7319287	33.1814637	2.5901689	0.0738774	0.0987379	0.0862813	0.0059556

Table 6.11: Sortie DEMM/SRSM (2ème ordre) pour les 37 pays Européens (suite)

6.6 Couplage de DEMM/SRSM et de la différentiation automatique

La méthode DEMM/SRSM nécessite un nombre important de choix possibles des points qui vont intervenir dans la résolution du système linéaire (6.9), et en absence de critères objectifs et efficaces, cette tâche pourrait s'avérer difficile. Afin de réduire le nombre de sélection à effectuer pour trouver les différents coefficients de l'approximation DEMM/SRSM, nous pouvons utiliser les méthodes de gradients pour évaluer les termes dérivés $\frac{\partial \tilde{Y}}{\partial \xi_l}$ pour chacune des possibilités.

Un différentiateur automatique fournit la dérivée d'un modèle automatiquement en fournissant un modèle retour capable d'évaluer les dérivées. Dans le cas d'une application au modèle POLAIR3D nous utilisons le différentiateur automatique Odysée ([27]).

Le couplage de la méthode DEMM/SRSM avec une méthode de gradient passe par les mêmes étapes qu'une méthode DEMM/SRSM : Si nous reprenons l'approximation (6.13)

$$\begin{aligned}
Y &= Y_0 + \sum_{i=1}^n Y_i^{(1)} H_1(\xi_i) + \sum_{i=1}^n Y_i^{(2)} H_2(\xi_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(1)} H_1(\xi_i) H_1(\xi_j) \\
&+ \sum_{i=1}^n Y_i^{(3)} H_3(\xi_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(2)} H_1(\xi_i) H_2(\xi_j) \\
&+ \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(3)} H_2(\xi_i) H_2(\xi_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n Y_{ijk}^{(1)} H_1(\xi_i) H_1(\xi_j) H_1(\xi_k). \\
&+ \dots
\end{aligned} \tag{6.13}$$

Nous pouvons exprimer les termes dérivées selon :

$$\begin{aligned}
\frac{\partial Y}{\partial \xi_l} &= \sum_{i=1}^n Y_i^{(1)} \frac{\partial H_1(\xi_i)}{\partial \xi_l} + \sum_{i=1}^n Y_i^{(2)} \frac{\partial H_2(\xi_i)}{\partial \xi_l} \\
&+ \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(1)} (H_1(\xi_i) \frac{\partial H_1(\xi_j)}{\partial \xi_l} + \frac{\partial H_1(\xi_i)}{\partial \xi_l} H_1(\xi_j)) \\
&+ \sum_{i=1}^n Y_i^{(3)} \frac{\partial H_3(\xi_i)}{\partial \xi_l} + \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(2)} (\frac{\partial H_1(\xi_i)}{\partial \xi_l} H_2(\xi_j) + H_1(\xi_i) \frac{\partial H_2(\xi_j)}{\partial \xi_l}) \\
&+ \sum_{i=1}^n Y_i^{(3)} \frac{\partial H_3(\xi_i)}{\partial \xi_l} + \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(2)} (\frac{\partial H_1(\xi_i)}{\partial \xi_l} H_2(\xi_j) + H_1(\xi_i) \frac{\partial H_2(\xi_j)}{\partial \xi_l}) \quad (6.14) \\
&+ \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_{ij}^{(3)} (\frac{\partial H_2(\xi_i)}{\partial \xi_l} H_2(\xi_j) + H_2(\xi_i) \frac{\partial H_2(\xi_j)}{\partial \xi_l}) \\
&+ \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n Y_{ijk}^{(1)} (\frac{\partial H_1(\xi_i)}{\partial \xi_l} H_1(\xi_j) H_1(\xi_k) \\
&+ H_1(\xi_i) \frac{\partial H_1(\xi_j)}{\partial \xi_l} H_1(\xi_k) + H_1(\xi_i) H_1(\xi_j) \frac{\partial H_1(\xi_k)}{\partial \xi_l}) \\
&+ \dots
\end{aligned}$$

Ces dérivées ($\frac{\partial Y}{\partial \xi_l}$) de premier ordre sont calculées par le modèle adjoint généré par Odyssee.

Sélection des points

Le sélection des points à utiliser pour pouvoir déterminer les coefficients de l'approximation sont choisis de la même manière que celle présentée dans la section (6.3), ce qui diffère étant juste le nombre de sélections à faire.

Pour illustrer cette méthode nous travaillons sur l'exemple d'illustration simple présente dans la section (6.3.3) avec 2 paramètres d'entrée incertains et une variable de sortie, exprimée par DEMM/SRSM de 3ème ordre selon :

$$\begin{aligned}
\tilde{Y} &= Y_0 + Y_1 \xi_1 + Y_2 \xi_2 + Y_3 (\xi_1^2 - 1) + Y_4 (\xi_2^2 - 1) + Y_5 (\xi_1^3 - 3\xi_1) \\
&+ Y_6 (\xi_2^3 - 3\xi_2) + Y_7 \xi_1 \xi_2 + Y_8 \xi_1 (\xi_2^2 - 1) + Y_9 \xi_2 (\xi_1^2 - 1) \quad (6.15)
\end{aligned}$$

$$\begin{aligned}\frac{\partial \tilde{Y}}{\partial \xi_1} &= Y_1 + 2Y_3\xi_1 + 3Y_5(\xi_1^2 - 1) + Y_7\xi_2 + Y_8(\xi_2^2 - 1) + 2Y_9\xi_2\xi_1 \\ \frac{\partial \tilde{Y}}{\partial \xi_2} &= Y_2 + 2Y_4\xi_2 + 3Y_6(\xi_2^2 - 1) + Y_7\xi_1 + 2Y_8\xi_1\xi_2 + Y_9(\xi_1^2 - 1)\end{aligned}\quad (6.16)$$

Pour estimer les 10 coefficients Y_0, Y_1, \dots, Y_9 , Contrairement à la méthode DEMM/SRSM initiale, nous avons besoin de $N = 7$ points (au lieu de 20) où le modèle sera exécuté avec les points correspondants $(X_1^{(1)}, X_2^{(1)})$, $(X_1^{(2)}, X_2^{(2)})$, ..., $(X_1^{(N)}, X_2^{(N)})$, ce qui nous donne un jeu de sorties : $Y^{(1)}, Y^{(2)}, \dots, Y^{(N)}$. Les coefficients Y_0, Y_1, \dots, Y_9 seront calculés de la même manière que dans la section (6.3.3), en résolvant le système linéaire :

$$M^T * \vec{Y} = \vec{B} \quad (6.17)$$

avec : $M = (M_1 \ M_2)$

$$M_1 = \begin{pmatrix} 1 & 0 & 0 & \dots \\ \xi_1^{(1)} & 1 & 0 & \dots \\ \xi_2^{(1)} & 0 & 1 & \dots \\ (\xi_1^2)^{(1)} & 2\xi_1^{(1)} & 0 & \dots \\ (\xi_2^2)^{(1)} & 0 & 2\xi_2^{(1)} & \dots \\ (\xi_1^3)^{(1)} & 3(\xi_1^2 - 1)^{(1)} & 0 & \dots \\ (\xi_2^3)^{(1)} & 0 & 3(\xi_2^2 - 1)^{(1)} & \dots \\ (\xi_1\xi_2)^{(1)} & \xi_2^{(1)} & \xi_1^{(1)} & \dots \\ (\xi_1(\xi_2^2 - 1))^{(1)} & (\xi_2^2 - 1)^{(1)} & 2(\xi_1\xi_2)^{(1)} & \dots \\ (\xi_2(\xi_1^2 - 1))^{(1)} & 2(\xi_1\xi_2)^{(1)} & (\xi_1^2 - 1)^{(1)} & \dots \end{pmatrix}$$

$$M_2 = \begin{pmatrix} \dots & 1 & 0 & 0 \\ \dots & \xi_1^{(N)} & 1 & 0 \\ \dots & \xi_2^{(N)} & 0 & 1 \\ \dots & (\xi_1^2)^{(N)} & 2\xi_1^{(N)} & 0 \\ \dots & (\xi_2^2)^{(N)} & 0 & 2\xi_2^{(N)} \\ \dots & (\xi_1^3)^{(N)} & 3(\xi_1^2 - 1)^{(N)} & 0 \\ \dots & (\xi_2^3)^{(N)} & 0 & 3(\xi_2^2 - 1)^{(N)} \\ \dots & (\xi_1\xi_2)^{(N)} & \xi_2^{(N)} & \xi_1^{(N)} \\ \dots & (\xi_1(\xi_2^2 - 1))^{(N)} & (\xi_2^2 - 1)^{(N)} & 2(\xi_1\xi_2)^{(N)} \\ \dots & (\xi_2(\xi_1^2 - 1))^{(N)} & 2(\xi_1\xi_2)^{(N)} & (\xi_1^2 - 1)^{(N)} \end{pmatrix}$$

et M^T , est la matrice transposée de M . De plus :

$$\vec{B} = \left(\tilde{Y}^{(1)}, \frac{\partial \tilde{Y}^{(1)}}{\partial \xi_1^{(1)}}, \frac{\partial \tilde{Y}^{(1)}}{\partial \xi_2^{(1)}}, \tilde{Y}^{(2)}, \frac{\partial \tilde{Y}^{(2)}}{\partial \xi_1^{(1)}}, \frac{\partial \tilde{Y}^{(2)}}{\partial \xi_2^{(1)}}, \tilde{Y}^{(3)}, \dots, \tilde{Y}^{(N)}, \frac{\partial \tilde{Y}^{(N)}}{\partial \xi_1^{(1)}}, \frac{\partial \tilde{Y}^{(N)}}{\partial \xi_2^{(1)}} \right)^T$$

$$\vec{Y} = (Y_0, Y_1, Y_2, \dots, Y_9)^T$$

D'une manière générale pour un système (6.15) ayant K coefficients à déterminer pour P variables d'entrée incertaines, le code adjoint fournira le calcul de P dérivées de première ordre et au total pour chaque point d'échantillonnage nous disposons de $P + 1$ équations linéaires. Le nombre requis d'échantillonnage pour déterminer les coefficients de l'approximation est donc de l'ordre de $\frac{2K}{P+1}$.

L'application de cette approche a ce modèle simple d'illustration a donné les mêmes résultats que ceux obtenus dans la section 6.15.

Conclusion

Nous avons présenté dans ce chapitre une méthode de propagation d'incertitude (DEMM/SRSM) basée sur une utilisation réduite de la méthode de Monte Carlo . Une première application a été effectuée dans le cadre 0D pour la cinétique chimique (EMEP) avec des incertitudes sur les constantes cinétiques. Nos tests témoignent de la conservation de précision avec un gain important en temps de calcul.

Dans l'application au CTM POLAIR3D, nous avons supposé une incertitude sur les émissions et plus précisément sur les bilans annuels de chaque pays figurant dans le réseau EMEP. Nous avons pris une incertitude différente pour chacune des espèces émises mais partout égale pour chaque pays émetteur. La méthode a montré un excellent résultat et cela dès le premier ordre de DEMM/SRSM, avec un speedup de 10^3 par rapport à la méthode classique de Monte Carlo.

En pratique cette hypothèse d'incertitude identique pour tous les pays, n'est pas vérifiée et nous l'avons utilisée à cause des contraintes imposées par la méthode DEMM/SRSM, car même si la méthode montre des gains très importants, en terme de représentation de l'incertitude, le nombre d'entrée supposées incertaines est nécessairement restreint (figure 6.2).

Si on suppose une incertitude différente pour chacune des espèces émises (4 espèces) et pour chaque pays européen (25 pays), ceci se traduit par un nombre de

paramètres incertains de l'ordre de 100, ce qui demanderait, comme le montre la figure 6.2, un nombre de simulations de modèle très élevé.

En supposant une incertitude homogène pour toutes les espèces mais différente en fonction du pays émetteur, cela baisserait le nombre des paramètres incertains à 25, ce qui demanderait une cinquantaine de simulation pour la méthode DEMM/SRSM de premier ordre, quelques 700 simulations pour le second ordre et 6500 simulations pour le troisième ordre (figure 6.2).

Chapitre 7

Analyse de sensibilité par la méthode RS-HDMR (Random Sampling - High Dimensional Model Representation)

Résumé

Ce chapitre présente une méthode de tabulation réduite, la méthode RS-HDMR (version Random Sampling de HDMR), fondée sur des développements fonctionnels différents de ceux présentés au Chapitre 4. Cette méthode a l'avantage de mieux "couvrir" l'espace des entrées du modèle à représenter et se prête bien à des calculs de sensibilité.

Nous présentons ici l'algorithme et le code qui a été développé, indépendamment du modèle utilisé. Une première application à la cinétique chimique a été effectuée avec le mécanisme EMEP ainsi qu'une analyse de sensibilité des sorties de ce mécanisme aux paramètres d'entrée du modèle.

Introduction

La méthode “High Dimensional Model representation” (HDMR), présentée dans [80, 81], est une méthode de tabulation réduite. Elle représente d’une manière efficace la relation qui existe entre les variables de sortie et les variables d’entrée d’un modèle, par une décomposition en n séries de fonctions de variables d’entrée.

Il existe deux méthodes de développement HDMR [80] :

- la méthode Cut-HDMR, associée au choix particulier d’un point de référence au sein du domaine Ω des variables d’entrée [80] : cette méthode a été présentée et utilisée au Chapitre 4.
- la méthode RS-HDMR (Random Sampling) qui dépend de la moyenne d’une sortie de modèle sur tout le domaine Ω .

C’est cette dernière approche que nous détaillerons dans ce chapitre. L’objectif est d’étudier la faisabilité d’une utilisation de l’approche RS-HDMR dans le contexte 3D de Chimie-Transport en tant que substitut aux solveurs de la cinétique chimique et/ou en tant que préprocess pour analyser la sensibilité du modèle à ces paramètres d’entrée.

Ce chapitre est organisé de la manière suivante. Nous présentons la méthode en section 1 et l’appliquons à un modèle académique simple. Cette application montre la nécessité d’utiliser des méthodes de réduction de variance afin de diminuer le nombre de simulations Monte Carlo nécessaire. Nous appliquons ensuite la méthode avec réduction de variance à la chimie EMEP 0D en section 2 et à une étude de sensibilité dans la section 3.

7.1 Méthode RS-HDMR

7.1.1 Formulation

On part de la représentation HDMR de la fonction $y = g(x)$, présentée au chapitre 4 :

$$\begin{aligned} g(x) = & f_0 + \sum_{i=1}^{i=n} f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) \\ & + \sum_{1 \leq i < j < k \leq n} f_{ijk}(x_i, x_j, x_k) + \dots + f_{1,2,\dots,n}(x_1, x_2, \dots, x_n) \end{aligned} \quad (7.1)$$

Dans la méthode dite RS-HDMR, les fonctions qui composent le développement HDMR sont déterminées à partir d’un ensemble de points aléatoirement prélevés

dans le domaine Ω .

Pour RS-HDMR, il est nécessaire de faire une transformation des variables d'entrée x_i de Ω sur le domaine unitaire ($K^n = [0, 1]^n$), telle que : $0 \leq x_i \leq 1$ pour tout i . Les fonctions composantes de RS-HDMR possèdent alors les formes suivantes [82] :

$$f_0 = \int_{K^n} \prod_{k=1}^{k=n} w_k(x_k) g(x) dx \quad (7.2)$$

$$f_i(x_i) = \int_{K^{n-1}} \prod_{k=1, k \neq i}^{k=n} w_k(x_k) g(x) dx^i - f_0 \quad (7.3)$$

$$f_{ij}(x_i, x_j) = \int_{K^{n-2}} \prod_{k=1, k \neq i, j}^{k=n} w_k(x_k) g(x) dx^{ij} - f_i(x_i) - f_j(x_j) - f_0, \quad (7.4)$$

...

où :

– $w_i(x_i)$ est la densité de probabilité de la variable x_i définie en respectant :

$$\begin{cases} w_i(x_i) \geq 0 & (0 \leq x_i \leq 1) \\ \int_0^1 w_i(x_i) dx_i = 1 & (i = 1, 2, \dots, n) \end{cases} \quad (7.5)$$

– $dx^i = \prod_{k \neq i} dx_k$,

– $dx^{ij} = \prod_{k \neq i, j} dx_k, \dots$

– f_0 est réellement la valeur moyenne de $g(x)$ sur le domaine entier K^n cube unitaire, contrairement à f_0 de Cut-HDMR qui est la valeur de $g(x)$ au seul point de référence choisi \bar{x} .

Remarque :

Pour des raisons de clarté et de simplicité, nous supposons que les variables d'entrée sont distribuées d'une manière uniforme sur leur intervalle respectif ($w_i(x_i) = 1$, pour toute variable d'entrée $\{x_i\}_{i=1, n}$). Cette approche est valable aussi pour des variables distribuées non uniformément ([83]).

On vérifie conformément à la définition du développement HDMR que :

$$\int_0^1 f_{i_1 i_2 \dots i_l}(x_{i_1}, x_{i_2}, \dots, x_{i_l}) dx_s = 0, \quad s \in \{i_1, i_2, \dots, i_l\} \quad (7.6)$$

ce qui définit la relation d'orthogonalité entre deux différentes fonctions composantes de Rs-hdmr :

$$\int_{K^n} f_{i_1 i_2 \dots i_l}(x_{i_1}, x_{i_2}, \dots, x_{i_l}) f_{j_1 j_2 \dots j_l}(x_{j_1}, x_{j_2}, \dots, x_{j_l}) dx = 0, \quad (7.7)$$

$$\{i_1, i_2, \dots, i_l\} \neq \{j_1, j_2, \dots, j_l\}$$

Le calcul des intégrales intervenant dans les fonctions de RS-HDMR peut être effectué par la méthode de Monte Carlo. En pratique, pour le calcul de f_0 par exemple :

$$f_0 \approx \frac{1}{N} \sum_{s=1}^N g(X^s) \quad (7.8)$$

où N est le nombre de points choisis aléatoirement, $x^s = (x_1^s, x_2^s, \dots, x_n^s)$ est tiré aléatoirement et $g(x^s)$ est la sortie de modèle correspondante.

La détermination directe de toutes les composantes de RS-HDMR à différentes valeurs x_i, x_j, \dots par un tirage Monte Carlo demande un grand nombre d'échantillons aléatoires. Par exemple pour déterminer $f_i(x_i)$, différents jeux d'échantillonnages aléatoires de $f(x_i, x^i)$ à $(x_i, x^i)^{(s)} = (x_1^{(s)}, x_2^{(s)}, \dots, x_{i-1}^{(s)}, x_i, x_{i+1}^{(s)}, \dots, x_n^{(s)})$ avec différentes valeurs de x_i sont indispensables :

$$f_i(x_i) = \int_{K^{n-1}} g(x) dx^i - f_0$$

$$\approx \frac{1}{N} \sum_{s=1}^N g((x_i, x^i)^{(s)}) - \frac{1}{N} \sum_{s=1}^N g(X^s) \quad (7.9)$$

Si la discrétisation selon l'axe de la variable x_i prend m valeurs, alors $m * N$ échantillonnages aléatoires sont nécessaires pour construire la table numérique $f_i(x_i)$ relativement à la seule variable d'entrée x_i .

De la même façon, pour calculer les $f_{ij}(x_i, x_j)$, différents jeux d'échantillonnages aléatoires de $f(x_i, x_j, x^{ij})$ à $(x_i, x_j, x^{ij})^{(s)} = (x_1^{(s)}, x_2^{(s)}, \dots, x_{i-1}^{(s)}, x_i, x_{i+1}^{(s)}, \dots, x_{j-1}^{(s)}, x_j, x_{j+1}^{(s)}, \dots, x_n^{(s)})$ avec différentes valeur de (x_i, x_j) sont nécessaires :

$$f_{ij}(x_i, x_j) = \int_{K^{n-2}} g(x) dx^{ij} - f_i(x_i) - f_j(x_j) - f_0$$

$$\approx \frac{1}{N} \sum_{s=1}^N g((x_i, x_j, x^{ij})^{(s)}) - \frac{1}{N} \sum_{s=1}^N g((x_i, x^i)^{(s)})$$

$$- \frac{1}{N} \sum_{s=1}^N g((x_j, x^j)^{(s)}) + \frac{1}{N} \sum_{s=1}^N g(X^s) \quad (7.10)$$

Si la discrétisation selon les axes des variables x_i et x_j prennent m valeurs dans chacune des directions, alors $m^2 * N$ échantillonnages aléatoires sont nécessaires pour la construction de la table $f_{ij}(x_i, x_j)$ relativement au seul couple de variables d'entrée (x_i, x_j) . Le nombre d'échantillonnages aléatoires nécessaires pour le calcul des fonctions composantes de RS-HDMR, croît rapidement avec l'ordre de troncature du développement RS-HDMR. Par ailleurs, les intégrales convergent pour un nombre de simulations Monte Carlo très élevé ($N \gg 1$). Le coût de construction de la table est donc particulièrement élevé.

7.1.2 Projection sur des bases de polynômes orthogonaux

Une approche alternative fondée sur l'utilisation d'un développement en bases de polynômes orthogonaux est proposée dans [84]-[85] et permet de diminuer le nombre de calculs. Elle exprime les fonctions $f_i(x_i), f_{ij}(x_i, x_j)$ de l'équation (7.1) de HDMR comme un développement de fonctions de type polynômes orthogonaux, fonctions splines ou tout autre type de fonctions polynômiales $\phi_k(x_i)$ selon :

$$f_i(x_i) \approx \sum_{k=1}^r c_k^i \phi_k(x_i) \quad (7.11)$$

$$f_{ij}(x_i, x_j) \approx \sum_{k=1}^r \sum_{l=1}^r c_{kl}^{ij} \phi_k(x_i) \phi_l(x_j) \quad (7.12)$$

.....

Dans la plupart des cas [84], la prise en compte de l'ordre 3 (ϕ_1, ϕ_2, ϕ_3) fournit une description satisfaisante de $g(x)$. Les coefficients c_k^i et c_{kl}^{ij} sont des constantes à déterminer. En utilisant ces formules, l'équation (7.1) peut s'écrire selon :

$$g(x) \approx f_0 + \sum_{i=1}^n \sum_{k=1}^r c_k^i \phi_k(x_i) + \sum_{1 \leq i < j \leq n} \sum_{k=1}^r \sum_{l=1}^r c_{kl}^{ij} \phi_k(x_i) \phi_l(x_j) + \dots \quad (7.13)$$

Chacun des coefficients c_k^i et c_{kl}^{ij} peut être déterminé par minimisation de la fonction J :

$$J(c_k^i, c_{kl}^{ij}, \dots) = \int_{K^n} [g(x) - f_0 - \sum_{i=1}^n \sum_{k=1}^r c_k^i \phi_k(x_i) - \sum_{1 \leq i < j \leq n} \sum_{k=1}^r \sum_{l=1}^r c_{kl}^{ij} \phi_k(x_i) \phi_l(x_j) - \dots]^2 dx \quad (7.14)$$

Les coefficients c_k^i et c_{kl}^{ij} sont donc obtenus en résolvant le système algébrique suivant :

$$Ay = b \quad (7.15)$$

d'inconnue $y = (c_1^i, c_2^i, \dots, c_r^i)^T$ et où A est une matrice constante carré d'ordre r, dont les éléments sont :

$$A_{k,l} = \int_0^1 \phi_k(x_i) \phi_l(x_i) dx_i, \quad (k, l = 1, 2, \dots, r) \quad (7.16)$$

$$b = \begin{pmatrix} \int_{K^n} g(x) \phi_1(x) dx \\ \int_{K^n} g(x) \phi_2(x) dx \\ \vdots \\ \int_{K^n} g(x) \phi_r(x) dx \end{pmatrix} \approx \frac{1}{N} \sum_{s=1}^N \begin{pmatrix} g(x^{(s)}) \phi_1(x_i^{(s)}) dx \\ g(x^{(s)}) \phi_2(x_i^{(s)}) dx \\ \vdots \\ g(x^{(s)}) \phi_r(x_i^{(s)}) dx \end{pmatrix} \quad (7.17)$$

La détermination de b nécessite le même jeu aléatoire de variables de sortie $g(x^s)$ utilisé initialement dans le calcul de f_0 . Tous les coefficients c_k^i et c_{kl}^{ij} sont donné par $A^{-1}b$. Cela rend la méthode efficace et rapide comparée au coût initial de calcul des $f_0, f_i, f_{ij} \dots$ par Monte Carlo.

Dans notre cas, nous utilisons les polynômes orthonormaux comme bases de fonction dans l'approximation de RS-HDMR. Les polynômes $\phi_k(x)$ définis sur un intervalle $[a, b]$ sont dits orthonormaux, quand ils satisfont :

$$\int_a^b \phi_k(x) dx = 0 \quad (7.18)$$

$$\int_a^b \phi_k^2(x) dx = 1 \quad (7.19)$$

$$\int_a^b \phi_k(x) \phi_l(x) dx = 0 \quad (k \neq l) \quad (7.20)$$

Sur l'intervalle $[0, 1]$ les polynômes orthonormaux peuvent être obtenus facilement à partir des formules (7.18), (7.19) et (7.20) :

$$\phi_1(x) = \sqrt{3}(2x - 1) \quad (7.21)$$

$$\phi_2(x) = 6\sqrt{5}(x^2 - x + \frac{1}{6}) \quad (7.22)$$

$$\phi_3(x) = 20\sqrt{7}(x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20}) \quad (7.23)$$

.....

En utilisant cette base de fonctions, les composantes $f_0, f_i, f_{ij} \dots$ de RS-HDMR, sont représentées dans les équations (7.11), (7.12) avec :

$$\phi_{kl}(x_i, x_j) = \phi_k(x_i)\phi_l(x_j) \quad (7.24)$$

En utilisant la propriété d'orthonormalité de ces polynômes, la matrice A est une identité et les coefficients c_k^i, c_{kl}^{ij} peuvent s'exprimer comme :

$$c_k^i \approx \frac{1}{N} \sum_{s=1}^N g(x^s) \phi_k(x_i^s) \quad (7.25)$$

$$c_{kl}^{ij} \approx \frac{1}{N} \sum_{s=1}^N g(x^s) \phi_k(x_i^s) \phi_l(x_j^s) \quad (7.26)$$

7.1.3 Application : Modèle Simple

Afin d'illustrer cette approche, notre première application a été effectuée sur un modèle académique de n variables d'entrée x_1, x_2, \dots, x_n , définies dans l'intervalle $[0,1]$:

$$Y = \sum_{i=1}^n x_i \quad (7.27)$$

l'objectif étant d'essayer d'approcher ce modèle par la méthode RS-HDMR selon :

$$Y_{RS\text{HDMR}} \approx f_0 + \sum_{i=1}^n \sum_{k=1}^r c_k^i \phi_k(x_i) + \sum_{1 \leq i < j \leq n} \sum_{k=1}^r \sum_{l=1}^r c_{kl}^{ij} \phi_{kl}(x_i, x_j) + \dots \quad (7.28)$$

En jouant sur le nombre d'entrée n , nous allons essayer de montrer l'influence de la précision de l'approximation par tirage Monte Carlo.

Les polynômes orthonormaux ϕ_k, ϕ_{kl} ($r = 1, 2, 3, \dots$) sont définis selon (7.21), (7.22), (7.23) et (7.24).

Une fois la représentation RSHDMR calculée, nous avons exécuté le modèle réduit 100 fois pour n variables d'entrée différentes et nous avons calculé l'erreur relative (ER) entre le modèle réduit et le modèle exact. Dans le tableau 7.1 nous présentons les résultats obtenus avec ($r=3$) et $N=10000$ simulations Monte Carlo utilisées pour le calcul des coefficients c_k^i et c_{kl}^{ij} .

Nombre de Variable d'entrée n	ER <5%	ER <10%	ER <20%	ER >100%
2	77	96	100	0
3	63	86	100	0
4	42	71	96	0
10	17	39	68	0
20	12	26	43	4

Table 7.1: Effet du nombre des variables d'entrée sur RS-HDMR (modèle académique)

Comme on peut le remarquer, la qualité des résultats décroît avec l'augmentation du nombre de variables d'entrée pris en compte!

La précision des coefficients c_k^i et c_{kl}^{ij} des fonctions f_i et f_{ij} et de f_0 composantes de l'approximation de RSHDMR dépend donc de la qualité de la simulation Monte Carlo.

Comme le nombre de tirages Monte Carlo est nécessairement contraint pour des raisons de coût calcul, des approches permettant de réduire ce nombre, tout en conservant une bonne précision, doivent être utilisées. Ceci fait l'objet de la sous-section suivante avec l'utilisation de méthodes de réduction de variance.

7.1.4 Réduction de variance

7.1.4.1 Fondement de la méthode Monte Carlo

En général, le calcul des intégrales multidimensionnelles ne peut être fait analytiquement, et il doit être déterminé numériquement. Des méthodes classiques déterministes (rectangles, trapèze,...) sont utilisées, mais la méthode de Monte Carlo est particulièrement efficace et la plus utilisée pour le calcul de ces intégrales.

Pour estimer l'intégrale :

$$I = \int_0^1 f(x) dx \quad (7.29)$$

La méthode de Monte-Carlo consiste à écrire cette intégrale sous forme d'une espérance $I = E(f(\xi))$ où ξ est une variable aléatoire qui suit une loi uniforme sur l'intervalle $[0,1]$. En utilisant *la loi des grands nombres* :

si $(\xi_i)_{i \geq 1}$ est une suite de variables aléatoires indépendantes suivant toutes la même loi uniforme sur $[0,1]$ que ξ , alors :

$$E(f(\xi)) = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{i=1}^N f(\xi_i) \quad (7.30)$$

ce qui nous mène à

$$I = \frac{1}{N} \sum_{i=1}^N f(\xi_i) \quad (7.31)$$

En pratique la taille de l'échantillon simulé est finie et la vitesse de convergence de la méthode de Monte-Carlo est donnée par *le théorème de la limite centrale*[86] :

Soit $(\xi_i)_{i \geq 1}$ une suite de variables aléatoires indépendantes suivant toutes la même loi uniforme sur $[0,1]$ que ξ . Si σ^2 est la variance de ξ , alors $\frac{\sqrt{n}}{\sigma} \left(\frac{1}{n} \sum_{i=1}^n f(\xi_i) - E(f(\xi)) \right)$ converge en loi vers une variable aléatoire gaussienne centrée réduite G :

$$P\left(a \leq \frac{\sqrt{N}}{\sigma} \left(\frac{1}{N} \sum_{i=1}^N f(\xi_i) - E(f(\xi)) \right) \leq b \right) \xrightarrow{N \rightarrow +\infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{\xi^2}{2}} d\xi \quad (7.32)$$

La convergence de la méthode de Monte-Carlo se fait à la vitesse de $\frac{\sigma}{\sqrt{N}}$, N étant le nombre d'échantillonnage et σ l'écart type de f . Cette méthode est donc très coûteuse en temps de calcul.

Pour améliorer la précision de calcul de cette méthode on utilise, en général, des méthodes de réduction de variance, qui consiste à diminuer la valeur de la variance de la fonction f .

7.1.4.2 Réduction de variance par la méthode de corrélation

Considérons un des coefficients c_k^i de l'équation (7.25) et (7.26) :

$$c_k^i = \int_{K^n} g(x) \phi_k(x_i) dx. \quad (7.33)$$

Pour réduire la variance de $g(x)\phi_k(x_i)$ dans K^n , nous utilisons la méthode dite de corrélation [77] qui consiste à introduire une fonction $h(x)$ et à décomposer :

$$c_k^i = \int_{K^n} [g(x) - h(x)] \phi_k(x_i) dx + \int_{K^n} h(x) \phi_k(x_i) dx. \quad (7.34)$$

Le choix de la fonction h est déterminé par les deux critères suivants :

- le terme $g(x)-h(x)$ doit rester petit ou avec des faibles variations sur le domaine K^n afin de garantir une faible variance de $(g - h)$
- l'intégrale $\int_{K^n} h(x) \phi_r(x_i) dx = \bar{c}_k^i$ doit être connue analytiquement.

Nous utilisons l'approximation RS-HDMR (7.13) comme fonction $h(x)$:

$$h(x) = f_0 + \sum_{i=1}^n \sum_{k=1}^r \bar{c}_k^i \phi_k(x_i) + \sum_{1 \leq i < j \leq n} \sum_{k=1}^r \sum_{l=1}^r \bar{c}_{kl}^{ij} \phi_k(x_i) \phi_l(x_j) + \dots \quad (7.35)$$

où les coefficients \bar{c}_k^i et \bar{c}_{kl}^{ij} sont déterminés par la méthode de Monte Carlo directe donnée par les équations (7.25) et (7.26).

Algorithme de RS-HDMR avec réduction de variance

L'algorithme de cette méthode se présente selon :

- **Etape 1** : Run par Monte Carlo direct, il fournit la valeur de f_0 et les différents coefficients $\bar{c}_k^i, \bar{c}_{kl}^{ij}, \dots$
- **Etape 2** : la valeur de g approchée par RS-HDMR (7.13) est prise comme fonction $h(x)$ pour la réduction de variance.
- **Etape 3** : les nouveaux coefficients c_k^i et c_{kl}^{ij} peuvent être alors approchés par un tirage Monte Carlo (le même tirage utilisé dans **Etape 1**) selon :

$$c_k^i \approx \frac{1}{N} \sum_{s=1}^N [g(x^{(s)}) - h(x^{(s)})] \phi_k(x_i^{(s)}) + \bar{c}_k^i \quad (7.36)$$

$$c_{kl}^{ij} \approx \frac{1}{N} \sum_{s=1}^N [g(x^{(s)}) - h(x^{(s)})] \phi_k(x_i^{(s)}) \phi_l(x_j^{(s)}) + \bar{c}_{kl}^{ij} \quad (7.37)$$

à précision équivalente aux approximations fournies par (7.25) et (7.26) pour un nombre N moins élevé.

- **Etape 4** : Les coefficients c_k^i et c_{kl}^{ij} peuvent être réutilisés pour le calcul d'une nouvelle fonction $h(x)$ pour réduire l'écart $g(x) - h(x)$.

Ainsi les étapes 2,3 et 4 deviennent une procédure itérative avec une convergence une fois que l'écart de $h(x)$ à $g(x)$ est très petit et le nombre d'échantillonnage N assez grand.

Nous avons appliqué cette méthode de réduction de variance au modèle simple présenté dans le paragraphe précédent, avec le test où le nombre de variables d'entrée $n=20$, ce qui conduisait des résultats peu satisfaisants (voir tableau 7.1). Les résultats obtenus avec cette méthode de réduction de variance sont très bons comme on peut l'observer sur le tableau 7.2, qui représente le nombre de fois où l'erreur relative entre le modèle exact et le modèle RS-HDMR est inférieure à 5% en utilisant la méthode de réduction de variance, comparé aux résultats précédents.

Nombre de Simulation Monte Carlo	Sans Réduction de Variance	Avec Réduction de Variance
10000	12%	21%
50000	16%	74%
100000	36%	97%
500000	47%	100%

Table 7.2: Effet de réduction de variance sur RSHDMR (% de runs avec une erreur relative $\leq 5\%$)

7.2 Application : cinétique chimique de POLAIR3D

Dans cette section nous présentons les résultats d'application de cette méthode (RS-HDMR avec réduction de variance) au modèle POLAIR0D . Le modèle pris pour procéder à la validation de la méthode RS-HDMR avec réduction de variance est la version 0D du modèle POLAIR3D [40] avec une chimie EMEP à 10 espèces et 9 réactions (Annexe B), la durée de l'intégration est 12h avec un pas de temps de 900 secondes.

Les résultats sont représentés dans la figure 7.1 qui montre une comparaison des profils de concentrations calculés avec POLAIR et ceux calculés par RS-HDMR, pour les espèces chimiques : NO_3 , NO_2 , SO_2 et NH_3 . Les résultats sont très semblables et montrent la validité de la méthode.

Ce résultat est obtenu avec une méthode de réduction de variance, exécutée avec un nombre d'itérations égal à 4.

7.3 Etude de sensibilité par RS-HDMR

L'analyse de sensibilité d'un modèle complexe consiste à identifier l'effet de la variation des paramètres d'entrée du modèle sur les sorties. La méthode traditionnelle utilisée pour ce genre d'étude est la méthode de Monte Carlo, et demande un grand nombre d'exécutions de modèle. Dans le cas des modèles de chimie transport cette simulation peut être difficile à réaliser.

Il existe deux types d'approches de propagation d'incertitude efficace :

- L'approche DEMM/SRSM qui consiste à réduire et optimiser le nombre d'exécution de modèle. Cette approche a été traitée dans le Chapitre 6.
- L'approche FEOM/HDMR en remplaçant le modèle d'origine par un modèle réduit, efficace et plus rapide (chapitres 4 et 5).

Dans notre cas, nous utiliserons la combinaison de ces deux approches. Dans la méthode RSHDMR le modèle est approché selon :

$$Y_{RS-HDMR} \approx f_0 + \sum_{i=1}^n \sum_{k=1}^r c_k^i \phi_k(x_i) + \sum_{1 \leq i < j \leq n} \sum_{k=1}^r \sum_{l=1}^r c_{kl}^{ij} \phi_{kl}(x_i, x_j) + \dots \quad (7.38)$$

où les coefficients c_k^i , c_{kl}^{ij} sont calculés selon la méthode de réduction de variance présentée auparavant.

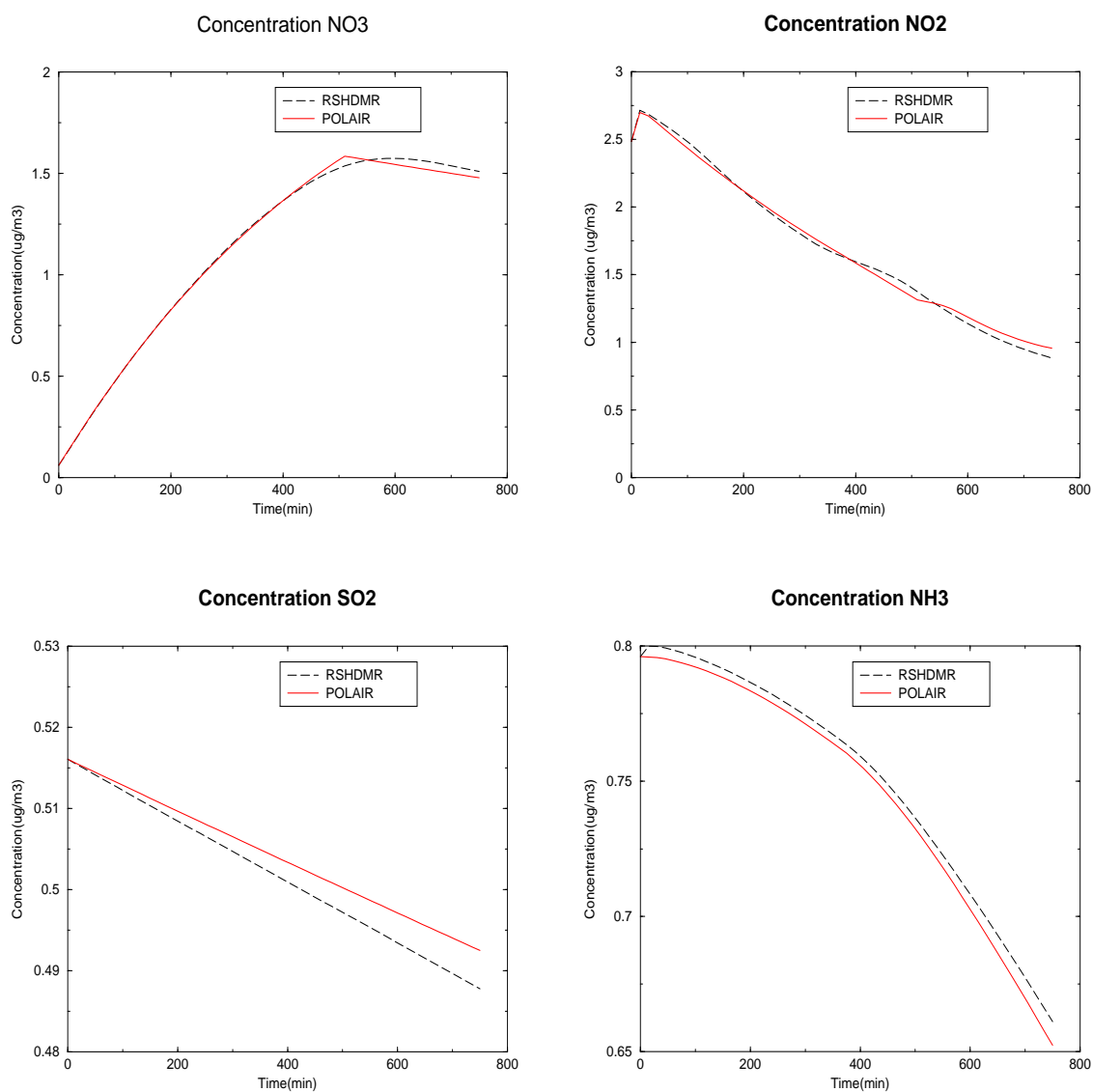


Figure 7.1: Comparaison des profils des 4 espèces du schéma EMEP, en (ligne) résultat obtenu avec le modèle d'origine POLAIR0D et en (-*-) résultat obtenu avec RSHDMR

En pratique, l'effet de la variation des paramètres d'entrée (x_i) sur les sorties peut être quantifié par un calcul de variance. En utilisant l'approximation HDMR de second ordre :

$$g(x) \approx f_0 + \sum_{i=1}^{i=n} f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) \quad (7.39)$$

la variance de la sortie de modèle σ_g^2 peut être décomposée selon :

$$\begin{aligned} \sigma_g^2 &= \int_{C^n} [g(x) - \bar{g}]^2 dx = \int_{C^n} [g(x) - f_0]^2 dx \\ &= \sum_{i=1}^n \int_0^1 f_i(x_i)^2 dx_i + \sum_{1 \leq i < j \leq n} \int_0^1 \int_0^1 f_{ij}(x_i, x_j) dx_i dx_j \\ &= \sum_{i=1}^n \sigma_i^2 + \sum_{1 \leq i < j \leq n} \sigma_{ij}^2 \end{aligned} \quad (7.40)$$

\bar{g} désigne la valeur moyenne de $g(x)$ sur le domaine Ω , les σ_i^2 et σ_{ij}^2 sont simplement :

$$\sigma_i^2 = \sum_{k=1}^r (c_k^i)^2, \quad (7.41)$$

$$f_0 = \int_{C^n} g(x) dx \approx \frac{1}{N} \sum_{i=1}^N g(x^s) \quad (7.42)$$

$$\sigma_g^2 = \int_{C^n} [g(x) - f_0]^2 dx \approx 1/N \sum_{i=1}^N g^2(x^s) - f_0^2 \quad (7.43)$$

$$\sigma_{ij}^2 = \sum_{k=1}^r \sum_{l=1}^r (c_{kl}^{ij})^2, \quad (7.44)$$

Nous notons que le calcul de ces variances ne nécessite qu'un jeu de variables aléatoires x ($g(x^s)$) pour le calcul de f_0 , $f_i(x_i)$, $f_{ij}(x_i, x_k)$, et par conséquent le calcul des variances σ_g^2 , σ_i^2 , σ_{ij} ,...

Dans ce cas la mesure de sensibilité est le facteur $S_{i_1 i_2 \dots i_l} = \frac{\sigma_{i_1 i_2 \dots i_l}^2}{\sigma_g^2}$ désignant la part de contribution de l'ensemble des variables d'entrée $\{i_1, i_2, \dots, i_l\}$ dans la variance de sortie σ_g^2 .

7.3.1 Test : modèle simple

Notre premier cas de test est un modèle académique simple faisant intervenir 4 variables d'entrée (x_1, x_2, x_3, x_4) et ayant une seule variable de sortie définie par :

$$Y = x_1 + x_2 + x_3 + x_2^2 + x_3^2 + x_2 * x_3 \quad (7.45)$$

Nous calculons la variance produite par les 4 variables d'entrée selon l'équation (7.43) puis la somme des différentes variances produites par chacune des variables d'entrée ($\sum_{i=1}^n \sigma_i^2$) obtenues à partir du développement RS-HDMR de premier ordre. Si la différence entre ces deux valeurs est petite, cela signifie que l'influence couplée des variables d'entrée sur la sortie de modèle est inexistante ou très négligeable.

Dans le tableau 7.3 est représentée la variance globale σ_g^2 et les résultats obtenus par RS-HDMR premier ordre et deuxième ordre. La variance produite par les variables d'entrée agissant seules (RS-HDMR premier ordre) est très proche de la valeur globale (une erreur relative de l'ordre de 0.3%) ce qui implique que l'effet des variables d'entrée agissant en couple est très faible (comme on peut le constater dans l'équation 7.45, l'effet couplé n'intervient qu'une fois par l'intermédiaire de $x_2 * x_3$).

Dans le tableau 7.4 nous synthétisons les variances produites par chacune des quatre variables d'entrée (x_1, x_2, x_3, x_4). La variance associée à la 4 ème variable d'entrée a une valeur nulle, ce qui est justifié par l'équation 7.45 qui ne fait apparaître aucun terme relatif à la variable x_4 . Les variables x_2 et x_3 sont symétriques dans l'équation 7.45 et ont un effet important sur la sortie du modèle comme on peut le constater avec les valeurs de variances relatives à x_2 et x_3 (tableau 7.4).

Dans le tableau 7.5 qui contient les variances σ_{ij}^2 associées aux variables d'entrée x_i et x_j sur la sortie du modèle, on peut constater que seules les variables x_2 et x_3 agissent ensemble sur la sortie du modèle (la seule variance importante est σ_{23}^2 , les autres sont nulles), comme on peut l'observer dans l'équation analytique du modèle initial (7.45).

7.3.2 Cinétique chimique EMEP

Dans cette section nous analysons la sensibilité du modèle de cinétique chimique EMEP, présenté dans l'annexe B . Le modèle fait intervenir 18 variables d'entrée dont 10 correspondant aux concentrations des espèces chimiques et qui représentent

Sortie de modèle	RSHDMR (1er Ordre)	RSHDMR (2ème ordre)	σ_g^2
1	1.06588518	1.0691377	1.06958092

Table 7.3: Apport du second ordre : impact de l'effet couplé

Entrée de modèle	σ_i^2
1	0.0833332566
2	0.526388579
3	0.526389383
4	$8.2196E^{-13}$

Table 7.4: Effet de chacune des variables d'entrée agissant seule

	1	2	3	4
1	0.	2.81621045E-12	4.0345556E-13	2.1591361E-12
2	0.	0.	0.00694435973	8.57555753E-13
3	0.	0.	0.	9.53847473E-13
4	0.	0.	0.	0.

Table 7.5: Effet de chacune des variables d'entrée agissant en couple

aussi les variables de sortie du modèle (tableau 7.6).

Angle zénithal	Température	Humidité	pluie	O_3 (Forcé)	ROH (Forcé)	OH (Forcé)	NO	NO ₂
HNO ₃	PAN	NO ₃	SO ₂	SO ₄	NH ₃	NITRAM	SULFAM	K_9

Table 7.6: Variables d'entrée et de sortie de RS-HDMR (**les variables de sortie sont les espèces chimiques de EMEP (en gras)**). K_9 est la constante cinétique de la 9ème réaction du schéma EMEP (Tableau 6.4)

Dans le tableau 7.7 sont représentées les variances globales de chaque concentration de sortie et les valeurs obtenues avec la méthode RS-HDMR de premier et second ordre. Nous pouvons aisément constater que les sorties 1 (NO), 2 (NO_2), 4 (PAN), 5 (NO_3), 6 (SO_2), 9 (NITRAM ¹) et 10 (SULFAM ²) ne sont affectées que par les variables d'entrée agissant seules. Cependant les variables de sorties 3 (HNO_3), 7 (SO_4) et 8 (NH_3) sont dépendantes, en plus, de l'effet couplé de quelques variables d'entrée.

Le tableau 7.8 montre l'effet des variances de premier ordre des variables d'entrée sur les sorties du modèle. La matrice présentée est composée des variables de sortie qui ne sont affectées que par les variables d'entrée agissant seules. Par exemple pour la sortie 1 qui correspond à la concentration de NO, cette variable n'est affectée directement que par les entrées 1, 2, 5, 8 et 9 qui correspondent à l'angle zenithal, la température, la concentration forcée de l'ozone, celles de NO et de NO_2 ; les autres valeurs de variances restent faibles par rapport à la variance globale de l'ordre de 1.2. La sortie 5 (NO_3) n'est affectée que par la variable d'entrée 12 correspondant à la concentration de (NO_3) et légèrement par la variable d'entrée 10 (HNO_3).

Dans les tableaux 7.9, 7.10, 7.11 nous montrons l'effet couplé des variables d'entrée sur les sorties 3 (HNO_3), 7 (SO_4) et 8 (NH_3) respectivement. Pour la sortie 3 correspondant à la concentration de l'espèce HNO_3 , seuls les couples de variables d'entrée (Température, humidité), (humidité, NITRAM), (HNO_3 , SO_4), (HNO_3 , NH_3) et (SO_4 , NH_3) agissent fortement d'une manière dépendante sur l'espèce HNO_3 . Cela est confirmé par les réactions et les constantes chimiques qui interviennent pour la production et la transformation de HNO_3 (Annexe B). De la même manière, la sortie 7 (SO_4) n'est affectée, en dehors de l'action singulière des variables d'entrée, que par l'effet couplé de (SO_4 , NH_3). Quant à la sortie 8 (NH_3), elle est sensible à l'action couplée des couples (température, humidité), (humidité,

¹Nitrate d'ammonium

²Sulfate d'ammonium

Chapitre 7. Analyse de sensibilité par la méthode RS-HDMR (Random Sampling - 7.3. Etude de sensibilité par RS-HDMR High Dimensional Model Representation)

NITRAM), (HNO_3 , SO_4), (HNO_3 , NH_3) et (SO_4 , NH_3). Cela est confirmé par la chimie de la partie non-linéaire de EMEP (B).

Sortie de modèle	RSHDMR(1er Ordre) $\sqrt{\sum_{i=1}^n \sigma_i^2}$	σ_g	RSHMDR(2ème ordre) $\sqrt{\sum_{i=1}^n \sigma_i^2 + \sum_{1 \leq i < j \leq n} \sigma_{ij}^2}$
1	0.956623274	1.12046086	1.09618684
2	13.4931462	13.6130956	13.6010458
3	29.7101369	39.9201073	33.9921463
4	74.3173305	74.4446513	74.4332415
5	117.336635	117.412959	117.33696
6	239.297392	239.320704	239.297401
7	19.4774511	23.3960017	23.3076971
8	10.9218003	13.3914728	12.1600265
9	107.257358	112.42354	109.288789
10	67.6476789	69.5239973	69.6077554

Table 7.7: Apport du second ordre : impact de l'effet couplé

Conclusion

Nous avons présenté dans ce chapitre la méthode de tabulation réduite RS-HDMR avec une utilisation de la méthode de réduction de variance afin de diminuer le nombre de simulations Monte Carlo nécessaire. L'approche a été appliquée à un exemple simple pour l'illustration de la méthode et à la cinétique chimique EMEP dans deux contextes différents :

- le premier afin d'étudier l'efficacité de la méthode à reproduire les résultats du modèle d'origine POLAIR ;
- le second dans le contexte d'analyse de sensibilité. La méthode a montré une bonne conservation de la précision comparée à celle du modèle 0D de POLAIR3D. L'étude de sensibilité a permis :
 - d'identifier les paramètres d'entrée importants dans la cinétique chimique pour chacune des espèces qui évolue ;
 - d'identifier l'action des paramètres qui agissent seuls ou en liaison avec d'autres variables.

variable	1	2	4	5	6	9	10
1	0.83228	1.47018	0.00072	0.06777	3.023E-09	0.062537	0.0008515
2	0.022859	6.91767	51.8567	0.10253	9.217E-09	3.7404	0.0001558
3	1.315E-06	4.916E-06	1.9629E-05	7.9760E-06	9.8645E-09	116.617	0.000457
4	3.1995E-06	1.8821E-05	8.8278E-05	1.2626E-05	1.4548E-08	0.0170	0.00031
5	0.00157	0.00078	8.891E-06	0.00174	8.3E-09	0.0219	0.00034
6	1.0201E-06	0.0073	2.1483E-05	1.090E-06	4.637E-09	0.0329	0.0001
7	2.740E-06	0.00374	0.02673	1.777E-06	1.41E-09	0.01668	0.00065
8	0.00873	0.5689	0.00025	0.00022	9.128E-09	0.013	0.00116
9	0.04955	172.457	0.03189	0.05322	1.40E-08	0.00947	0.000139
10	1.724E-06	6.846E-06	4.822E-05	0.0825	1.5467E-09	380.143	2.30E-05
11	9.376E-05	0.6389	5471.14	0.000319	4.724E-09	0.01985	0.000453
12	2.26E-06	1.36751E-05	2.797E-05	13767.577	8.5638E-10	0.1029	0.00063
13	2.163E-06	1.453E-06	3.277E-05	1.0865E-06	57263.22	0.02262	0.02527
14	2.786E-06	2.7E-06	1.777E-05	2.548E-06	1.172E-09	123.748	1093.398
15	1.6299E-06	4.1739E-06	8.14E-05	2.853E-06	5.14E-09	693.9	464.70
16	7.60E-06	1.25E-05	0.0001565	2.54E-06	1.068E-09	10185.52	0.
17	2.493E-06	2.975E-05	7.268E-05	8.402E-06	1.0875E-08	0.0277	3018.04
18	4.577E-07	2.355E-06	3.058E-05	2.442E-06	0.013	0.10985	0.03082
σ_g^2	1.25543254	185.316371	5542.00611	13785.8031	57274.3995	12639.0523	4833.58621

Table 7.8: Effet de chacune des variables d'entrée

var	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0.	0.02743	0.31895	0.03984	0.03245	0.03334	0.03033	0.03294	0.03817	0.05190	0.04688	0.04558	0.03552	0.04048	0.04771	0.10615	0.01402	0.05623
2	0.	0.	23.79863	0.03348	0.01773	0.02146	0.02556	0.03180	0.02742	0.12218	0.03477	0.04094	0.02339	0.04605	0.03869	0.45688	0.03992	0.04647
3	0.	0.	0.	0.15268	0.19804	0.26317	0.09932	0.06501	0.19891	0.80348	0.20151	0.04200	0.27027	0.57589	1.17656	18.18191	0.13205	0.74307
4	0.	0.	0.	0.	0.09193	0.03153	0.02945	0.04103	0.04520	0.01370	0.05796	0.11501	0.01176	0.04389	0.02082	0.06350	0.04155	0.02973
5	0.	0.	0.	0.	0.	0.08330	0.03590	0.03565	0.04283	0.05747	0.03471	0.01990	0.04892	0.02298	0.03540	0.05112	0.03019	0.03702
6	0.	0.	0.	0.	0.	0.	0.03997	0.06695	0.06328	0.05767	0.01417	0.03049	0.03180	0.02973	0.02109	0.07792	0.01416	0.04319
7	0.	0.	0.	0.	0.	0.	0.	0.03252	0.03850	0.02644	0.03582	0.07395	0.03859	0.02745	0.02590	0.09583	0.04932	0.03235
8	0.	0.	0.	0.	0.	0.	0.	0.	0.07828	0.03223	0.03314	0.07197	0.03437	0.02834	0.00689	0.10878	0.04914	0.02635
9	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.02756	0.03966	0.03043	0.05220	0.08144	0.01543	0.02729	0.05771	0.04112
10	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.02190	0.03104	0.01582	23.86894	142.46345	0.04285	0.06117	0.05358
11	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.05265	0.02550	0.03932	0.01386	0.07182	0.01995	0.04083
12	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.02277	0.04676	0.06592	0.03395	0.03246	0.05002
13	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.01454	0.10608	0.02054	0.02469	0.06707
14	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	53.12644	0.06908	0.03246	0.03719
15	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.04628	0.06453	0.05087
16	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.02472	0.10421
17	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.02146
18	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.

Table 7.9: σ_{ij}^2 Effet de chacune des variables d'entrée (agissant en couple) sur la sortie 3 (HNO_3)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0.	0.0004	0.0004	0.0015	0.0007	0.001	0.0005	0.0009	0.0004	0.00158	0.0003	0.00058	0.0006	0.00150	0.00032	0.0007	0.00069	0.00093
2	0.	0.	0.001	0.0005	0.0004	0.0007	0.0011	0.0006	0.0005	0.00128	0.0005	0.00025	0.0003	0.00116	0.00054	0.0007	0.00039	0.00069
3	0.	0.	0.	0.0005	0.0002	0.0008	0.0004	0.0010	0.0014	0.00231	0.0011	0.00038	0.0010	0.00057	0.00038	0.0007	0.00073	0.00076
4	0.	0.	0.	0.	0.0015	0.0004	0.0008	0.0007	0.0007	0.00078	0.0004	0.00033	0.0009	0.00126	0.00109	0.001	0.00072	0.00098
5	0.	0.	0.	0.	0.	0.0006	0.0004	0.001	0.0008	0.00055	0.0003	0.00157	0.0010	0.00065	0.00092	0.0005	0.00038	0.00143
6	0.	0.	0.	0.	0.	0.	0.0007	0.0006	0.0008	0.00058	0.0008	0.00053	0.0009	0.00063	0.00135	0.0010	0.00034	0.00057
7	0.	0.	0.	0.	0.	0.	0.	0.0003	0.0004	0.00090	0.0009	0.00075	0.0003	0.00042	0.00065	0.0007	0.00039	0.00077
8	0.	0.	0.	0.	0.	0.	0.	0.	0.0005	-1.00095	0.0006	0.00067	0.0008	0.00160	0.00077	0.0006	0.00063	0.00056
9	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00106	0.0004	0.00055	0.0004	0.00057	0.00049	0.0003	0.00036	0.00033
10	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.0007	0.00078	0.0005	0.00082	0.00131	0.0005	0.00039	0.00041
11	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00076	0.0004	0.00112	0.00094	0.0001	0.00058	0.00092
12	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.0005	0.00053	0.00062	0.0005	0.00146	0.00105
13	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00077	0.00513	0.0010	0.00023	0.00149
14	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	163.76	0.0002	0.00094	0.00120
15	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.001	0.00109	0.00270
16	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00056	0.00056
17	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00019
18	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.

Table 7.10: σ_{ij}^2 : Effet de chacune des variables d'entrée (agissant en couple) sur la sortie 7 (SO_4)

var	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0.	0.00205	0.02197	0.00306	0.00230	0.00272	0.00203	0.00234	0.00250	0.00323	0.00317	0.00282	0.00239	0.00266	0.00329	0.00804	0.00093	0.00454
2	0.	0.	1.72856	0.00248	0.00130	0.00143	0.00231	0.00215	0.00179	0.00756	0.00235	0.00302	0.00195	0.00265	0.00281	0.03442	0.00310	0.00299
3	0.	0.	0.	0.01061	0.01391	0.01759	0.00765	0.00409	0.01556	0.05739	0.01337	0.00274	0.02064	0.04063	0.08571	1.31430	0.00993	0.05367
4	0.	0.	0.	0.	0.00607	0.00223	0.00185	0.00306	0.00299	0.00103	0.00402	0.00795	0.00117	0.00307	0.00155	0.00478	0.00257	0.00204
5	0.	0.	0.	0.	0.	0.00575	0.00239	0.00256	0.00279	0.00417	0.00232	0.00157	0.00383	0.00183	0.00248	0.00343	0.00219	0.00309
6	0.	0.	0.	0.	0.	0.	0.00289	0.00472	0.00478	0.00395	0.00125	0.00200	0.00216	0.00187	0.00137	0.00569	0.00100	0.00313
7	0.	0.	0.	0.	0.	0.	0.	0.00216	0.00287	0.00176	0.00285	0.00517	0.00269	0.00186	0.00183	0.00748	0.00334	0.00215
8	0.	0.	0.	0.	0.	0.	0.	0.	0.00533	0.00257	0.00268	0.00512	0.00268	0.00228	0.00031	0.00790	0.00412	0.00207
9	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00219	0.00272	0.00234	0.00369	0.00606	0.00119	0.00198	0.00414	0.00273
10	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00149	0.00212	0.00129	1.74691	10.40647	0.00307	0.00438	0.00392
11	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00349	0.00196	0.00296	0.00113	0.00516	0.00157	0.00272
12	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00159	0.00320	0.00445	0.00233	0.00196	0.00334
13	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00104	0.00823	0.00132	0.00171	0.00383
14	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	12.57457	0.00501	0.00210	0.00254
15	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00380	0.00417	0.00393
16	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00163	0.00717
17	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.00146
18	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.

Table 7.11: σ_{ij}^2 . Effet de chacune des variables d'entrée (agissant en couple) sur la sortie 8 (NH_3)

L'intérêt de ce types d'études est d'utiliser RS-HDMR en préalable à l'utilisation de la méthode cut-HDMR pour construire le modèle FEOM qui remplacera le module de la cinétique chimique dans le modèle POLAIR3D. Elle permet d'identifier les variables d'entrée à prendre en compte pour construire les tables HDMR et d'identifier les espèces agissant en couples sur les sorties de modèle, afin de ne garder que les couples agissant d'une manière sensible. Le problème exposé dans le chapitre 4 relatif à la mémoire et l'espace de sauvegarde demandés par la construction des tables HDMR, dans le cas d'une utilisation pour une cinétique complexe faisant intervenir un grand nombre d'espèces et de paramètres chimiques, peut ainsi se résoudre sans passer par un filtre numérique comme utilisé dans le Chapitre 4.

Partie V

Matrices de Transfert et Modèles moyens

Chapitre 8

Calcul des matrices de transfert par modèle adjoint

Résumé

Nous présentons dans ce chapitre une méthode pour le calcul des matrices de transfert, basée sur le modèle adjoint.

Dans un premier temps nous étudions la non-linéarité du modèle POLAIR3D avec une chimie EMEP. Nous présentons ensuite notre approche de calcul de matrices transfert par le modèle adjoint en utilisant le Différentiateur Automatique Odysée. Nous finissons par une application de cette méthode à un cas d'étude d'impact en utilisant le modèle de Chimie-Transport POLAIR3D à l'échelle Européenne.

Introduction

Dans un système intégré d'étude d'impact (8.1) le but est de développer des stratégies de contrôle des émissions afin de réduire leur impact sur la pollution de l'air. Ce système, comme schématisé dans la figure 8.1, comporte :

- un modèle d'émission,
- un système de simulation de Chimie-Transport,
- des indicateurs d'impact,
- un modèle de type dose/effet,
- et des modules macro-économiques d'analyse du coût engendré.

Le système utilise et combine les différentes informations issues de ces modules et effectue un traitement d'optimisation afin de trouver la solution optimale pour un problème environnemental donné avec un minimum de coût et/ou un équilibre entre le coût de réduction des émissions et le bénéfice acquis avec la réduction d'impact. Le choix d'une approche d'optimisation reste bien sûr fondamentalement basé sur des choix politiques et économiques.

Parmi les informations essentielles dans ce système intégré, la relation qui lie la réduction des émissions à la réduction sur les espèces chimiques déposées au sol ou dans l'air est crucial. Ces informations sont fournies généralement sous forme de matrices, dites **Matrices de Transfert**.

Celles-ci jouent un rôle très important dans la modélisation des impacts, en fondant des stratégies de réductions des émissions, par exemple sur l'Europe, Elles fournissent les variations en concentration et en dépôt sur une zone de réception donnée, induites par les changements dans les émissions d'une zone émettrice. Leur construction se fait sur la base d'un calcul sur plusieurs années (environ 10 ans) en utilisant un modèle de chimie transport avec le même scénario d'émissions des principales espèces émises. Ainsi, à partir de ces simulations et de leur moyenne, on calcule par simple division (dépôt/emission) les différents éléments de la matrice ([87]).

L'utilisation des matrices de transfert dans les système intégrés d'études d'impact est donc basée sur l'hypothèse de linéarité qui suppose une relation linéaire entre la réduction des émissions et leur impact. Cependant cette hypothèse n'est pas forcément vérifiée dans les modèles de chimie transport, les mécanismes de chimie utilisés étant de plus en plus complexes et non linéaires.

Sans l'utilisation de ces matrices de transfert (ie sans l'hypothèse de linéarité sous-jacente), l'étude des différents scénarios d'impact, deviendrait très coûteuse

voire impraticable. Si nous envisageons d'analyser l'effet de réduction de 20%, 40%, 69%, 80% et 100% des émissions actuelles de 2004 (émissions de référence) d'une espèce chimique (NO_x , SO_2 , NH_3 ,...) d'un pays européen donné sur la concentration dans l'air ou le dépôt d'une espèce chimique (NO_2 , SO_2 , Sulfate d'ammonium, nitrate d'ammonium,...) sur l'Europe en 2014 (10 ans plus tard), nous supposons que la concentration des espèces transportées pour des émissions données resteront peu éloignées de la concentration moyenne sur les 10 années dont nous disposons. De plus, les champs météo dont nous disposons correspondent à la période 1994-2004, les champs météorologiques pour l'année 2014 n'étant évidemment pas disponibles. Ce qui signifie qu'il est indispensable d'effectuer les 10 années de calculs avec le même jeu d'émissions pour chaque scénario de réduction, pour chaque espèce émise (auxquelles on s'intéresse) et pour chaque pays émetteur. Cela se traduit, de point de vue ressource de calcul nécessaire, par un coût calcul de l'ordre :

$$\text{Coût} = N_{scen} * N_{pays} * N_{emis} * N_{an} * T_{POLAIR} \quad (8.1)$$

où N_{scen} est le nombre de scénario de réduction (ici 5), N_{pays} le nombre de pays concernés par la réduction des émissions (de l'ordre de 40 pays Européens), N_{emis} le nombre d'espèces chimiques émises à réduire (NO_x , SO_x , NH_3 , CO, VOC), N_{an} est le nombre d'années nécessaires pour effectuer le calcul et T_{POLAIR} est le temps CPU du modèle de chimie transport. Dans le cas de POLAIR3D et avec une chimie simplifiée de EMEP, T_{POLAIR} est de 6 heures pour un an de calcul. Pour effectuer cette étude de de scénario de réduction des émissions, il faudra environ **6 ans de calcul** (on ne tient pas compte ici de l'apport possible du parallélisme)...

Pour surmonter cette difficulté, nous proposons dans ce chapitre, une autre approche de calcul de matrices de transfert basée sur les modèles adjoints. Nous présentons également une application dans le cadre d'une étude d'impact utilisant le modèle de Chimie-Transport (POLAIR3D) à l'échelle Européenne.

8.1 Non linéarité du modèle

En général, le calcul des matrices de transfert se heurte à des difficultés liées à la non-linéarité du modèle (que ce soit pour le schéma chimique ou les méthodes numériques utilisées, par exemple pour l'advection).

Pour illustrer cette non-linéarité nous nous sommes intéressés à la chimie d'acidification d'EMEP sur l'Europe avec les émissions annuelles de 1994. Nous étudions l'effet de la réduction des émissions françaises mensuelles (Janvier) de NO_x , SO_2 et NH_3 sur le reste de l'Europe, en étudiant les concentrations de sortie de NO_2 , SO_2 et NH_3 . Plusieurs scénarios sont étudiés et sont résumés dans le Tableau 8.1.

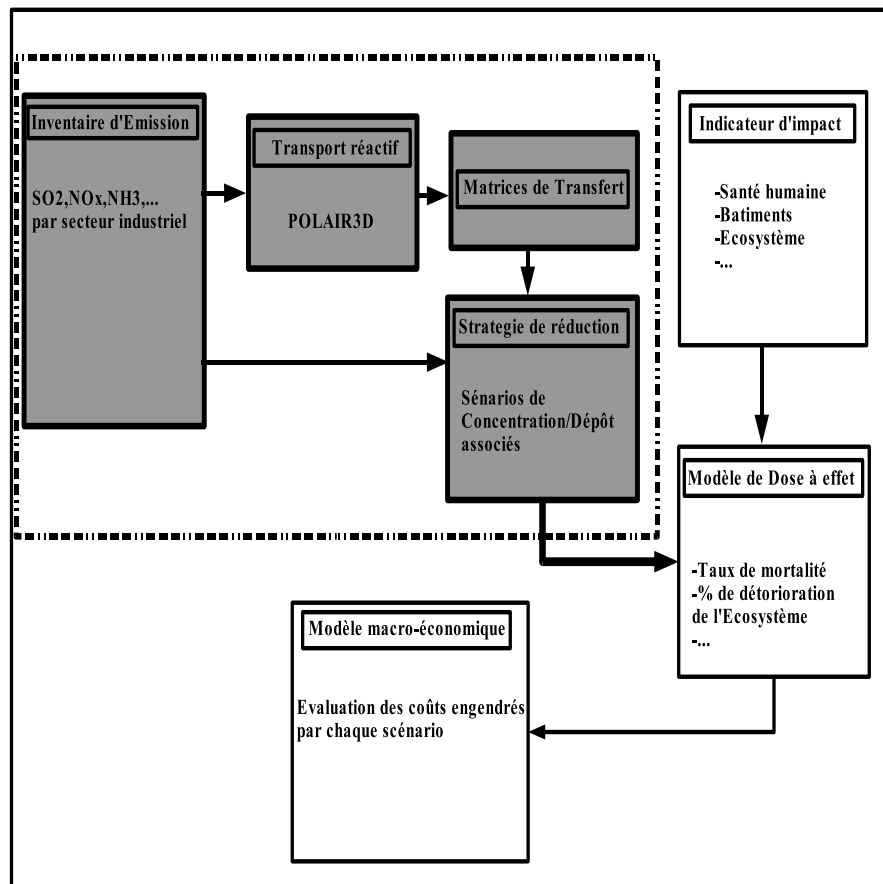


Figure 8.1: Structure d'un système intégré d'aide à la décision pour une étude d'impact

Dans un premier temps nous avons étudié la non linéarité qui existe entre les mêmes espèces émises et transportées, puis ensuite entre deux espèces différentes.

Scénario A : Réduction des émissions mensuelles de la France en NO_x	Scénario B : Réduction des émissions totales mensuelles de la France en SO_x	Scénario C : Réduction des émissions totales mensuelles de la France en NH_3
A.1 :20% de réduction de NO_x	B.1 :20% de réduction de SO_x	C.1 :20% de réduction de NH_3
A.2 :40% de réduction de NO_x	B.2 :40% de réduction de SO_x	C.2 :40% de réduction de NH_3
A.3 :60% de réduction de NO_x	B.3 :60% de réduction de SO_x	C.3 :60% de réduction de NH_3
A.4 :80% de réduction de NO_x	B.4 :80% de réduction de SO_x	C.4 :80% de réduction de NH_3
A.5 :100% de réduction de NO_x	B.5 :100% de réduction de SO_x	C.5 :100% de réduction de NH_3

Table 8.1: Scénarios de réductions

8.1.1 Non linéarité des effets inter-pays, mono-espèce

Nous analysons ici l'effet de réduction des émissions de NO_x , SO_2 et NH_3 sur les concentrations en l'air respectivement des espèces NO_2 , SO_2 et de NH_3 . Comme exemple nous prenons les scénarios de réduction présentés dans le Tableau 8.1 avec comme pays émetteur la France, et comme pays récepteur l'Allemagne et la Pologne. Nous rappelons que, dans cette expérience, les émissions des autres pays ne sont pas modifiées afin de voir l'effet marginal de la réduction des émissions en France.

Dans la figure 8.2 est présenté l'effet de réduction des émissions de NO_x (Scénario-A) de la France sur la réduction de la concentration de NO_2 dans l'air en France et en Allemagne. L'effet sur la France peut être considéré comme linéaire même si une réduction de 100% des émissions de NO_2 se traduirait par une diminution de 92% de NO_2 dans l'air. L'effet sur l'Allemagne est faible puisqu'une réduction de 100% des émissions de la France en NO_x ne donnerait qu'une diminution avoisinant les 2.5%.

Pour le Scénario-B, une réduction de SO_x (figure 8.2) sur la France, représente une non-linéarité plus importante puisqu'une réduction, par exemple, de 60% des émissions sur la France engendrerait une réduction d'environ 40% dans la concentration de l'air sur la France en SO_2 , et moins de 2 % sur l'Allemagne. Une réduction de 100% des émissions (SO_x) de la France, donnerait une réduction de 70% sur la concentration en SO_2 de la France et environ de 3.5% sur l'Allemagne. La figure 8.4 (Scénario-C) montre l'effet de réduction des émissions de NH_3 sur la France et l'Allemagne : pour la France on peut observer une linéarité, sauf pour une

réduction de 100% que donnerait à peu près 95% de réduction de la concentration de NH_3 sur la France. Cependant pour l'Allemagne, la réduction est d'environ 5% (resp. 7.5%) pour une réduction de 40% (resp. 100%) des émissions de France en NH_3 .

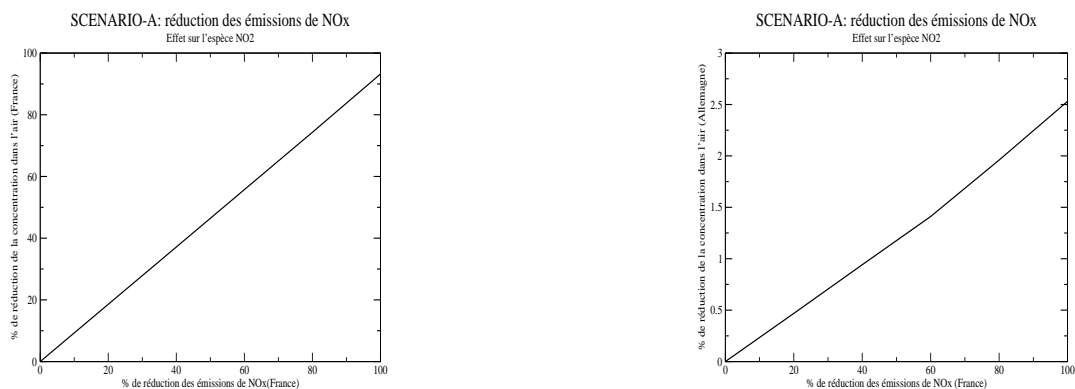


Figure 8.2: Non-linéarité : Emissions de NO_x (France), Réception NO_2 France et Allemagne

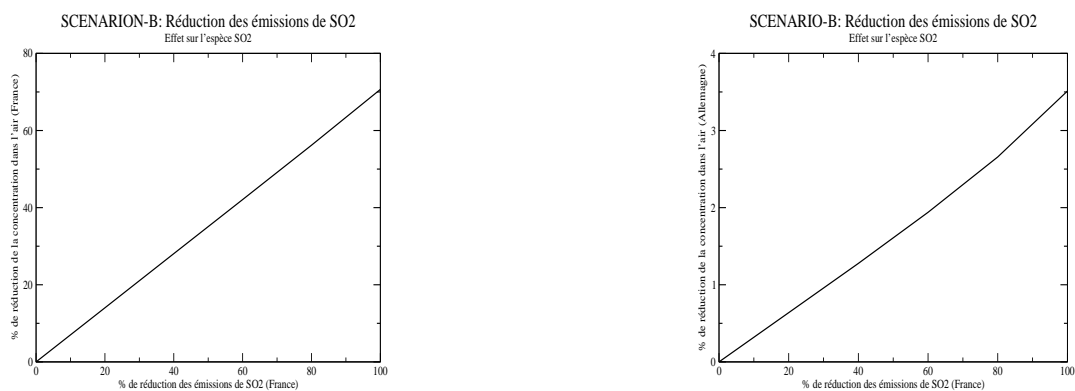


Figure 8.3: Non linéarité entre mêmes espèces : Emissions de SO_2 (France), Réception SO_2 France et Allemagne

8.1.2 Non linéarité des effets inter-espèces

Dans le calcul et dans les applications de la matrice de transfert ([88, 87]), il est supposé qu'il n'existe aucune dépendance entre les espèces émises et les espèces déposées (ou dans l'air). Dans le cas d'EMEP ([88]) par exemple, il est considéré que :

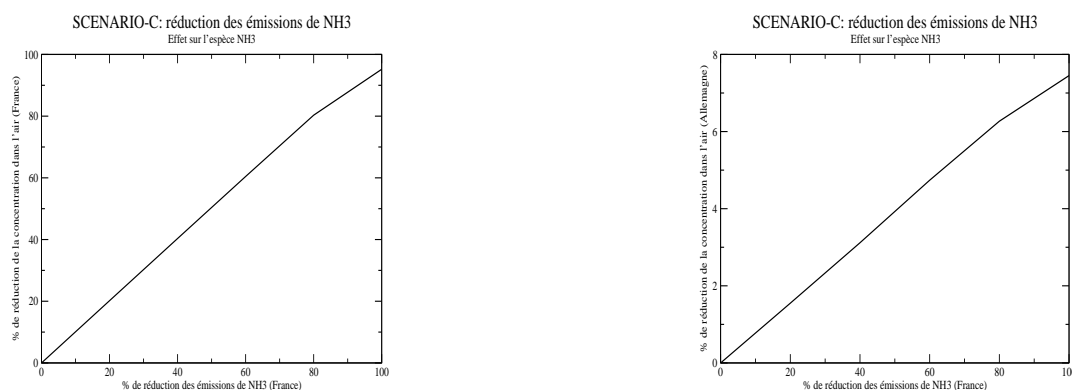


Figure 8.4: Non linéarité entre mêmes espèces : Emissions de NH_3 (France), Réception NH_3 France et Allemagne

1. Le dépôt de SO_2 est indépendant des émissions de NO_X et NH_3 ;
2. Le dépôt de NO_2 est indépendant des émissions de SO_2 et NH_3 ;
3. Le dépôt de NH_3 est indépendant des émissions de NO_X et SO_2 ;

Nous étudions ici l'effet croisé de la réduction des émissions de la France en NO_X , SO_2 et de NH_3 sur les concentrations dans l'air de NO_X , SO_2 et de NH_3 sur la France et l'Allemagne selon les effets suivant :

- Effet de réduction des émissions de la France en SO_2 sur la concentration dans l'air de NO_2 et NH_3 en France et en Allemagne.
- Effet de réduction des émissions de la France en NO_x sur la concentration dans l'air de SO_2 et NH_3 en France et en Allemagne.
- Effet de réduction des émissions de la France en NH_3 sur la concentration dans l'air de SO_2 et NO_2 en France et en Allemagne.

Dans la figure 8.5 est présentée la relation entre la réduction des émissions de la France en NO_X (Scénario A) et la concentration de NH_3 dans l'air. Nous pouvons constater qu'une diminution des émissions de NO_X induit une augmentation de la concentration de NH_3 sur la France et sur l'Allemagne, des augmentations inférieures à 10% pour la France (environ 5.5% pour une réduction de 100% des émissions de NO_x) et une augmentation inférieure à 2% sur l'Allemagne. L'impact de cette réduction sur la concentration dans l'air de SO_2 est nulle (aucun changement par rapport à l'expérience de référence effectuée sans scénario de réduction).

Dans le Scénario B (Figure 8.6), une réduction des émissions de SO_2 se traduit par une légère augmentation dans la concentration de l'air en NH_3 sur la

France (maximum de 1.7% pour une élimination totale des émissions de SO_2) et sur l'Allemagne (un maximum de 0.75% d'augmentation de NH_3 en cas d'une réduction de 100% des émissions de SO_2). L'effet de réduction des émissions de SO_2 sur la concentration de l'air de NO_2 est nul.

Le Scénario C, correspondant à une stratégie de réduction des émissions de NH_3 , n'affecte pas la concentration de SO_2 ni celle de NO_2 dans l'air.

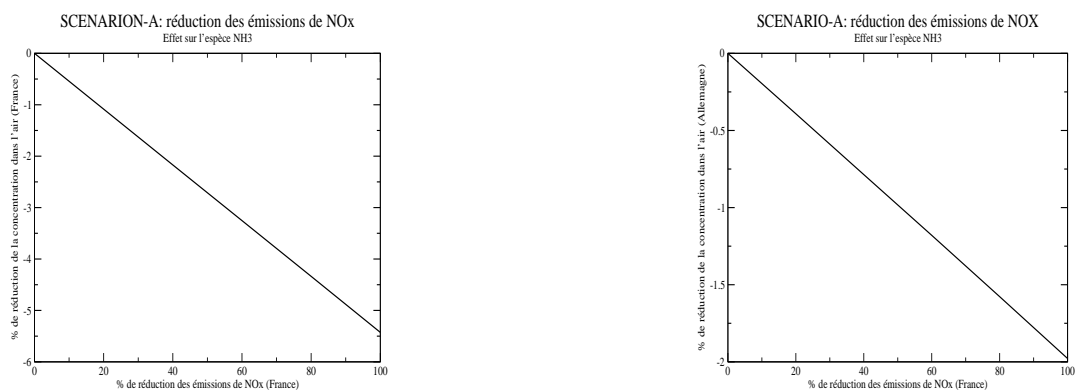


Figure 8.5: Non linéarité entre espèces différentes : Emissions de NO_x (France), Réception NH_3 France et Allemagne

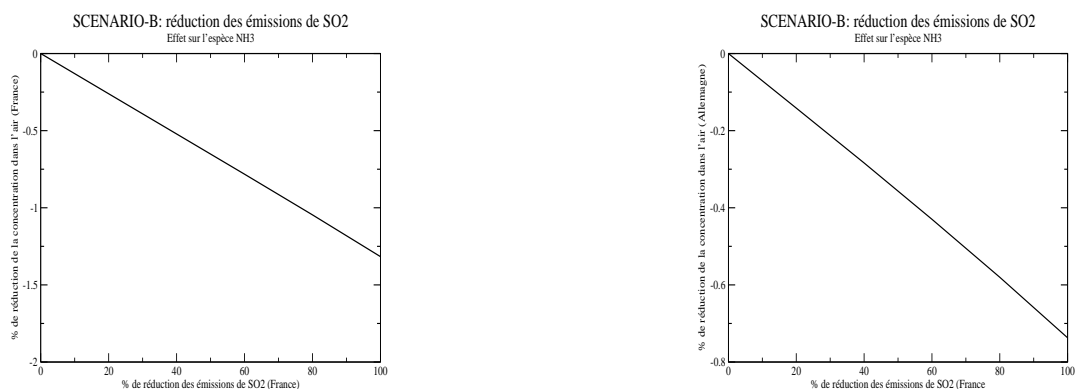


Figure 8.6: Non linéarité entre espèces différentes : Emissions de SO_2 (France), Réception NH_3 France et Allemagne

Cette étude, à caractère purement illustratif, permet de montrer néanmoins les limites de l'hypothèse linéaire.

8.2 Définition de la matrice de transfert

Les matrices de transfert "émission/récepteur" matrices décrivant la relation entre les émissions et le dépôt des espèces chimiques (ou les concentrations dans l'air) sur différents échelles de temps et d'espace.

Différents types de matrice de transfert émission/récepteur sont utiles dans la modélisation d'impact, à savoir ([88, 87]) :

- la matrice "Pays/point de grille" : l'objectif est d'évaluer l'impact de la réduction des émissions d'un pays sur le dépôt sur un point de grille bien spécifique,
- la matrice "Pays/Pays", qui représente l'effet de réduction d'émission d'un pays sur le dépôt dans un autre pays,
- la matrice "point de grille/pays" : le but ici est d'évaluer l'impact de réduction des émissions d'un point de grille (typiquement une source d'émissions) sur le dépôt/concentration d'un pays,
- la matrice ponctuelle liée à une activité industrielle "point de grille/point de grille", où l'on étudie par exemple l'impact des émissions d'une source sur un autre point (une ville avoisinante par exemple).

En général, on calcule des matrice journalières, mensuelles ou annuelles, selon le besoin de l'étude. Néanmoins, les matrices annuelles sont les plus utilisées dans le cadre des études d'impact. Elles peuvent être calculées pour les concentrations en l'air ou pour les concentrations déposées (sèches et humide) au sol. Une matrice annuelle (ou mensuelle) Pays/Pays peut ainsi être définie selon :

$$(D_X)_i = \sum_{j=1}^{j=N_{srce}} M_{ij} * (E_Y)_j \quad (8.2)$$

Où

- X représente les espèces chimiques du modèle avec $(D_X)_i$ le terme de concentration déposée (ou dans l'air) de l'espèce X sur un point (Pays) récepteur i .
- Y symbolise l'espèce chimique émise avec $(E_Y)_j$ terme des émissions de l'espèce émise Y du point (pays) j . N_{srce} dans ce cas est le nombre de points (Pays) émetteurs.
- M_{ij} est l'élément correspondant de la matrice de transfert.

Dans notre application nous avons illustré l'utilisation des matrices de transfert dans le cas d'études d'impact des émissions Françaises sur le reste des pays européens.

8.2.1 Modèle RAINS

RAINS (Regional Air pollution INformation and Simulation) est un outil développé par IIASA comme outil intégré d'aide à la décision pour l'optimisation des stratégies de réduction des émissions afin de maîtriser le dépôt acide sur l'Europe et l'Asie [5]. Les résultats de ce modèle sont utilisés pour le développement de nouveaux protocoles, comme celui de Goeteborg en 1999 sur la pollution transfrontalière et pour des directives européennes sur les émissions nationales (Directive 2001). On peut par exemple se référer au site web www.iiasa.ac.at/rains.

Dans le modèle RAINS, les éléments M_{ij} de la matrice de transfert mettent en oeuvre :

- le calcul du dépôt annuel sur le domaine d'étude (Europe dans notre cas) après plusieurs années (5-10 ans) de calcul du modèle de Chimie-Transport (EMEP) ;
- le calcul du dépôt annuel pour chaque pays de l'Europe (D_i) ;
- le calcul des éléments de la matrice M_{ij} en divisant le dépôt annuel d'un pays D_i par les émissions annuelles de chaque pays E_j .

Dans le modèle RAINS, la matrice est calculée pour représenter la relation émission/réception en perturbant les émissions annuelles d'un pays ou un groupe de sources dans l'exécution du modèle de chimie transport (EMEP).

Pour la calcul de la matrice de transfert, le modèle RAINS, basé sur des études préliminaires sur la linéarité [88] qui pourrait exister entre les émissions d'une espèce et le dépôt pour chaque ou une combinaison d'espèce, procède pour chaque pays de la manière suivante :

- Run1 : perturbation des émissions de SO_2 et NO_x .
- Run2 : perturbation des émissions de NH_3 .

Avec deux runs du modèle CTM par pays, la relation émission/réception pour le dépôt de SO_2 et NO_2 peut être produite.

Notons qu'une telle méthode ne peut être utilisée pour étudier l'impact des sources ponctuelles de chaque pays sur le reste des pays européens (matrice de type Point de grille/ Pays) car cela nécessiterait un coût calcul de l'ordre de :

$$Cout_{RAINS} = 2 * N_{an} * N_{pays} * N_{srce} * N_{emis} * T_{POLAIR} \quad (8.3)$$

où N_{an} est le nombre d'années retenu pour construire la matrice de transfert (de l'ordre de 5 à 10 années), N_{pays} le nombre de pays, N_{srce} le nombre de sources émettrice dans chaque pays, N_{emis} le nombre d'espèce émises retenues et supposées ayant un effet environnemental et T_{POLAIR} est le temps CPU effectué par le modèle POLAIR3D pour une simulation d'une année ou d'un mois selon le type de matrice de transfert à construire. Évidemment, le temps demandé pour calculer ces matrices de transfert dans de telles conditions s'avère élevé et quasiment impraticable dans un contexte opérationnel.

8.3 Calcul de la matrice de transfert par modèle adjoint

8.3.1 Modèles dérivés

Dans cette section nous présentons une autre approche de calcul des éléments de la matrice de transfert en se basant sur le calcul des dérivées partielles. On suppose que la concentration dans l'air ou déposée (D_i) sur un pays i est liée aux émissions (E) de tous les points sources d'un pays ou un ensemble de pays (N_{srce}) par la relation fonctionnelle :

$$D_i = f_i(E) \quad (8.4)$$

Le terme D_i peut alors être exprimé au premier ordre d'un développement de Taylor selon :

$$D_i = f(E) = f_i(E^0) + \sum_{j=1}^{j=N_{srce}} \frac{\partial f_i}{\partial E_j}(E^0) * (E_j - E_j^0) \quad (8.5)$$

Pour un jeu de perturbations (ΔE_j), on alors :

$$\Delta D_i = \sum_{j=1}^{j=N_{srce}} M_{ij} * \Delta E_j \quad (8.6)$$

avec

$$M_{ij} = \frac{\partial f_i}{\partial E_j} \quad (8.7)$$

Notons que ce calcul linéarisé est valide pour un modèle de Chimie-Transport quelconque (ie sans hypothèse de linéarité). Dans notre cas, f_i est la sortie du modèle de chimie transport sur le pays i (POLAIR3D).

8.3.2 Calcul des dérivées partielles

Le calcul des termes de la matrice de transfert, est donc basé sur le calcul des dérivées de la fonction f_i par rapport aux émissions.

Pour une fonction quelconque $x \mapsto f(x)$, le calcul de $\frac{\partial f}{\partial x}$ peut se faire par plusieurs méthodes [89, 90] :

1. Différence finie :

$$\left\{ \begin{array}{l} \frac{\partial f}{\partial x}(x_i) \approx \frac{f(x_i + \Delta x) - f(x_i - \Delta x)}{2\Delta x} \quad \text{centrale} \\ \frac{\partial f}{\partial x}(x_i) \approx \frac{f(x_i + \Delta x) - f(x_i)}{\Delta x} \quad \text{droite} \\ \frac{\partial f}{\partial x}(x_i) \approx \frac{f(x_i) - f(x_i - \Delta x)}{\Delta x} \quad \text{gauche} \end{array} \right. \quad (8.8)$$

cette méthode reste très coûteuse en temps CPU car elle nécessite un très grand nombre d'exécutions du modèle f (le nombre de variables d'entrée, ie la dimension de c) qui peut être difficile à effectuer avec un choix de Δx qui doit être très petit ($\Delta x \rightarrow 0$).

2. Différentiation analytique :

Elle est basée sur le calcul du gradient d'une manière analytique en dérivant l'équation mathématique du modèle initial, ainsi l'équation obtenue du gradient est discrétisée et résolue avec un autre schéma numérique. Cette méthode reste très difficile à mettre en oeuvre surtout pour des modèles complexes tels que les modèles de chimie transport [89, 90].

3. Différentiation automatique :

Dans cette approche, le modèle calculant les gradients est fourni automatiquement par un outil de différentiation du code associé au modèle d'origine (f). Elle est basée sur le fait que chaque fonction f représentant un modèle est résolue numériquement comme un ensemble de m séquences d'opérations

élémentaires (en pratique les instructions successives du code) $\{f_k\}_{k=1}^{k=m}$ (d'addition, de soustraction, de multiplication et d'autres fonctions élémentaires : sinus, cosinus, valeur absolue, racine carré,...) :

$$f = f_m \circ f_{m-1} \circ \cdots \circ f_1 \quad (8.9)$$

Le points clé est l'utilisation de la règle de dérivation de la composition de deux fonctions :

$$\frac{\partial f(g(x))}{\partial x} \Big|_{x=x_0} = \left(\frac{\partial f(s)}{\partial s} \Big|_{s=g(x_0)} \right) \left(\frac{\partial g(x)}{\partial x} \Big|_{x=x_0} \right) \quad (8.10)$$

En utilisant la relation (8.10) à chaque utilisation des opérations élémentaires, le différentiateur automatique peut calculer les dérivées de f , d'une manière correcte selon la précision de la machine.

La Différentiation Automatique comporte deux modes, le mode linéaire tangent et le mode adjoint.

- Dans le **mode linéaire tangent**, la dérivée de f par rapport à x s'exprime selon :

$$\begin{aligned} f'(x) = & f'_m \circ f_{m-1} \circ \cdots \circ f_1(x) \\ & * f'_{m-1} \circ f_{m-2} \cdots \circ f_1(x) \\ & * \cdots \\ & * f'_1(x) \end{aligned} \quad (8.11)$$

Notons que pour calculer le gradient d'une composante de f par rapport à x , ceci nécessite n évaluations de ce modèle ($x \in \mathbb{R}^n$).

- Dans le **mode adjoint**, on remarque que pour une fonction scalaire J de $f(x)$, $\nabla_x J$ est défini pour toute perturbation δx par la relation au premier ordre :

$$\delta J \approx (\nabla_x J, \delta x) = (\nabla_f J, \delta f) = (\nabla_f J, \nabla_x f \cdot \delta x) = ((\nabla_x f)^T \nabla_f J, \delta x) \quad (8.12)$$

soit

$$\nabla_x J = (\nabla_x f)^T \nabla_f J \quad (8.13)$$

Ce qui est donc intéressant à calculer est le transposé de la dérivée de la

fonction f par rapport à x selon :

$$\begin{aligned}
 f'^T(x) = & f'_1{}^T(x) \\
 & * f'_2{}^T \circ f_1(x) \\
 & * \dots \\
 & * f'_{m-1}{}^T \circ f_{m-2} \circ \dots \circ f_1(x) \\
 & * f'_m{}^T \circ f_{m-1} \circ \dots \circ f_1(x)
 \end{aligned}
 \tag{8.14}$$

Notons que ceci permet d'obtenir le gradient d'une sortie scalaire par une seule application du modèle adjoint.

A présent, si nous avons à calculer le gradient de p sorties par rapport à n entrées, le coût calcul est de p modèles adjoints ou n modèles linéaires tangents. Les deux coûts "unitaires" étant équivalents, le choix entre les deux modèles dépendra donc de n/p .

La différentiation automatique est basée sur l'utilisation d'un outil de calcul capable d'interpréter le code direct et de produire un code linéaire tangent et/ou adjoint. Le principal avantage de cette méthode est qu'elle peut être utilisée pour des modèles numériques complexes qui peuvent être pris comme "boîtes noires".

Les méthodes de différentiation automatique sont largement utilisées pour l'assimilation de données et l'analyse de sensibilité des modèles de chimie transport. Il existe plusieurs outils de différentiation automatique valable pour des modèles écrits en Fortran(77 et/ou 90), en C/C++. On peut citer dans le cas d'un modèle en Fortran :

- Odyssee outil développé par l'INRIA, utilisée pour des codes en Fortran 77 ([27]).
- Tapede développé par l'INRIA également, valable pour des modèles en Fortran 77 et 90 (www-sop.inria.fr/tropics/tapede.html).
- ADIFOR développé au USA par 'Mathematics and Computer Science Division of Argonne National Laboratory' en commun avec 'the Center for Research on Parallel Computation at Rice University', utilisé pour des modèles écrits en Fortran 77 ([91])

Dans nos applications nous avons utilisé le différentiateur automatique Odyssee, comme déjà indiqué au chapitre 6.

8.3.3 Adjointisation du modèle POLAIR

Le calcul de la matrice de transfert a été effectuée à l'aide du modèle de chimie transport POLAIR3D, en calculant la relation qui lie les émissions du modèle et les concentrations dans l'air ou déposées des espèces chimiques.

Si on représente le modèle POLAIR3D de la façon fonctionnelle suivante :

$$\begin{cases} C^n = f(S^n(E_0), S^{n-1}(E_0), C^{n-1}, \alpha^n) \\ C_0 \longrightarrow C_1 \longrightarrow C_2 \cdots C_{n-1} \longrightarrow C_n \end{cases} \quad (8.15)$$

où

- C^n est la concentration à l'instant t_n en un point (i, j, k) du domaine, sortie du modèle POLAIR3D, C^{n-1} est la concentration à l'instant t_{n-1} .
- S^n et S^{n-1} sont les termes sources (volumique ou surfacique) respectivement à l'instant t_n et t_{n-1} , termes déduit des émissions annuelles E_0 , après une distribution spatiale et temporelle (présenté dans le chapitre (3) sur la validation du modèle POLAIR3D). Dans l'application qui va suivre, nous utiliserons les termes de sources volumiques.
- α^n est l'ensemble des autres paramètres qui interviennent dans le modèle pour le calcul de C^n et qui sont indépendants des émissions E_0 .

L'objectif est de calculer le gradient de f ou d'une fonction coût J ($J(C^n)$) par rapport à E_0 .

Nous choisissons comme fonction coût la moyenne temporelle et spatiale sur un pays (p) de la concentration dans l'air (au niveau du sol) d'une espèce chimique C :

$$\begin{cases} J = \frac{1}{N} \sum_{n=1}^N J^n, & J^n = \frac{1}{N_{x_p} * N_{y_p}} \sum_{j=1}^{N_{y_p}} \sum_{i=1}^{N_{x_p}} C^n(i, j) \\ \frac{\partial J}{\partial E_0} = \frac{1}{N} \sum_{n=1}^{N_{stop}} \frac{\partial J^n}{\partial E_0} \end{cases} \quad (8.16)$$

La fonction coût J étant décomposée en sous fonctions $J^n = F(C_n)$, son gradient s'écrit sous forme :

$$\frac{\partial J^n}{\partial E_0} = \frac{\partial C^n}{\partial E_0} \cdot \frac{\partial J^n}{\partial C^n} \quad (8.17)$$

On peut vérifier aisément que $\frac{\partial J^{nT}}{\partial C^n} = \frac{1}{N_{x_p} * N_{y_p}} * I$ (I matrice identité). N_{x_p}, N_{y_p} désignent le nombre de points de grille qui composent le pays 'p', N est le nombre d'itérations effectuées durant l'intégration du modèle.

Nous utiliserons le mode adjoint pour calculer $\frac{dC^n}{dE_0}$ à chaque instant t_n la sortie de modèle POLAIRD3D qui est liée directement à la concentration à l'instant t_{n-1} ainsi qu'à la source à l'instant t_n et t_{n-1} ($C^n = F(C^{n-1}, S^{n-1}, S^n)$), si on pose pour tout n , $\lambda^n = (\frac{dS^n}{dE_0})^T$. Les termes $(\frac{dS^n}{dE_0})^T$ et $(\frac{dS^{n-1}}{dE_0})^T$ sont calculés d'une manière analytique à partir de la procédure Emissions/sources, on les notera respectivement λ^n et λ^{n-1} .

Nous obtenons :

$$\frac{dC^{nT}}{dE_0} = \lambda^n \cdot \frac{\partial C^{nT}}{\partial S^n} + \lambda^{n-1} \cdot \frac{\partial C^{nT}}{\partial S^{n-1}} + \frac{dC^{n-1T}}{dE_0} \cdot \frac{\partial C^{nT}}{\partial C^{n-1}} \quad (8.18)$$

on posera ($\alpha^n = \lambda^n \cdot \frac{\partial C^{nT}}{\partial S^n} + \lambda^{n-1} \cdot \frac{\partial C^{nT}}{\partial S^{n-1}}$) ce qui donnera par une procédure itérative :

$$\frac{dC^{nT}}{dE_0} = \sum_{j=0}^{n-1} \alpha^j \cdot \left[\frac{\partial C^{j+1T}}{\partial C^j} \cdot \frac{\partial C^{j+2T}}{\partial C^{j+1}} \cdots \frac{\partial C^{nT}}{\partial C^{n-1}} \right] + \alpha^n \quad (8.19)$$

$$\nabla_{E_0} J = \frac{1}{N} \sum_{n=1}^N \left[\sum_{j=0}^{n-1} \alpha^j \cdot \left[\frac{\partial C^{j+1T}}{\partial C^j} \cdot \frac{\partial C^{j+2T}}{\partial C^{j+1}} \cdots \frac{\partial C^{nT}}{\partial C^{n-1}} \right] + \alpha^n \right] \quad (8.20)$$

Ce qui peut s'écrire selon :

$$\begin{aligned} \nabla_{E_0} J = & \frac{1}{N} \alpha^N + \alpha^{N-1} \cdot \left[1 + \frac{\partial C^{NT}}{\partial C^{N-1}} \right] + \alpha^{N-2} \cdot \left[1 + \frac{\partial C^{N-1T}}{\partial C^{N-2}} \cdot \left[1 + \frac{\partial C^{NT}}{\partial C^{N-1}} \right] \right] + \\ & \alpha^{N-2} \cdot \left[1 + \frac{\partial C^{N-2T}}{\partial C^{N-3}} \cdot \left[1 + \frac{\partial C^{N-1T}}{\partial C^{N-2}} \cdot \left[1 + \frac{\partial C^{NT}}{\partial C^{N-1}} \right] \right] \right] + \\ & \cdots + \\ & \alpha^0 \cdot \left[1 + \frac{\partial C^{1T}}{\partial C^0} \cdot \left[1 + \frac{\partial C^{2T}}{\partial C^1} \cdot \left[1 + \frac{\partial C^{3T}}{\partial C^2} \cdot [\cdots] \cdot \left[1 + \frac{\partial C^{NT}}{\partial C^{N-1}} \right] \right] \right] \right] \end{aligned} \quad (8.21)$$

Nous utiliserons le modèle adjoint de POLAIR pour calculer pour tout vecteur

Y :

$$\left\{ \begin{array}{l} \alpha^n \cdot Y \\ \hat{a}^n = \left[\frac{\partial C^n}{\partial S^n} \right]^T, \quad \hat{b}^n = \left[\frac{\partial C^n}{\partial S^{n-1}} \right]^T \\ \alpha^n = \lambda^n * \hat{a}^n + \lambda^{n-1} * \hat{b}^n. \\ Y = \left[1 + \frac{\partial C^N}{\partial C^{N-1}} \right]^T, \left[1 + \frac{\partial C^{N-1}}{\partial C^{N-2}} \right]^T \cdot \left[1 + \frac{\partial C^N}{\partial C^{N-1}} \right]^T, \dots \end{array} \right. \quad (8.22)$$

L'utilisation de l'adjoint exige pour un modèle non linéaire la connaissance de la trajectoire du modèle C^0, C^1, \dots, C^N , ce qui oblige la sauvegarde de cette trajectoire une fois au début de la procédure, au lieu d'un calcul au fur et à mesure de l'algorithme. Ceci nécessite des ressources en espace disque importante.

Le code POLAIR3D est adjointisé par le différentiateur ODYSSEE ([27]) de l'INRIA. Nous représentons dans la figure 8.7 un schéma de structure d'adjointisation.

Algorithme d'adjointisation

Ce que nous venons de décrire se traduit algorithmiquement par :

- Initialisation : $\hat{C}_N = 0, \nabla_{E_0} J = 0, \hat{a}^N = 0$ et \hat{b}^N
- Boucle rétrograde sur $n = N - 1$ à 0
 - lire dans le fichier la valeur de C_{n-1}
 - $\hat{C}_{n-1} = \hat{C}_n + I$
 - $\alpha^{n-1} : \hat{b}^{n-1} = \hat{a}^n ; \hat{a}^{n-1} = 0$
 - Calcul par adjoint ($\alpha^n : \hat{a}^{n-1} * \hat{C}_{n-1}, \hat{b}^{n-1} * \hat{C}_{n-1}$) et \hat{C}_{n-1}
 - $\nabla_{E_0} J = \nabla_{E_0} J + \lambda^{n-1} \hat{b}^{n-1}$
- Fin de la boucle
- $\nabla_{E_0} J = \nabla_{E_0} J + \lambda^0 \cdot \hat{a}^0$
- Sortie : $\nabla_{E_0} J = \nabla_{E_0} J * \frac{1}{N * N_{x_p} * N_{y_p}}$

8.3.3.1 Test de Taylor

La validation numérique du modèle adjoint se fait par test de Taylor [92]. Considérons une fonction coût J de paramètre E_0 et $\nabla_{E_0} J$ le gradient calculé avec ODYSSEE; nous comparons ce gradient à celui calculé d'une manière approchée par différence finie ∇J_{ref} .

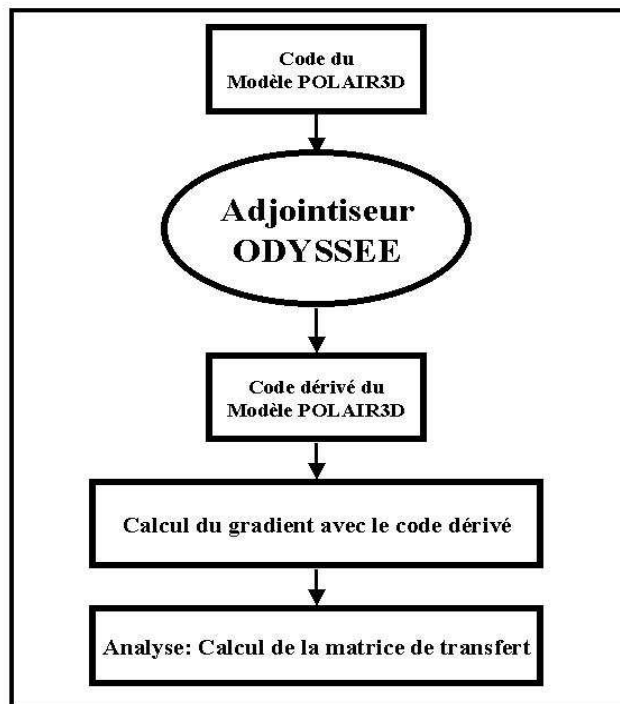


Figure 8.7: Schéma d'adjointisation du modèle POLAIR3D par Odyssee (Application : calcul de matrices de transfert).

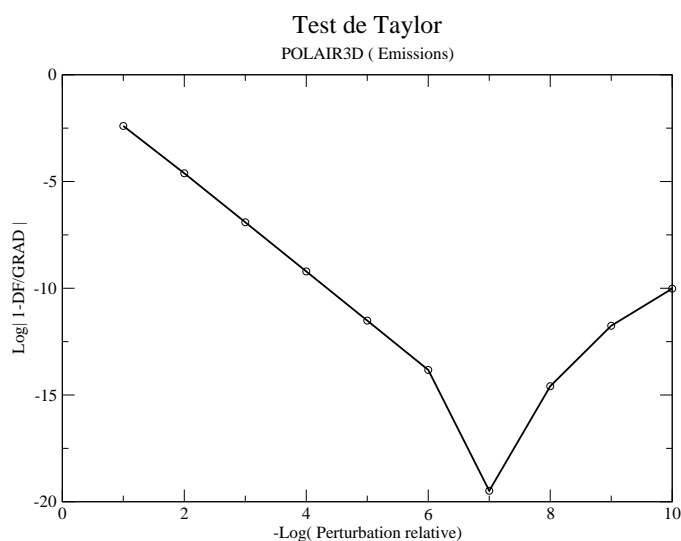


Figure 8.8: Test de Taylor pour POLAIR3D

Si pour toute petite perturbation δE_0 de E_0 on a :

$$\lim_{\alpha \rightarrow 0} \frac{J(E_0 + \alpha \delta E_0) - J(E_0)}{\alpha \cdot (\nabla J_{ref}(E_0) \cdot \delta E_0)} = 1 \quad (8.23)$$

alors le test de Taylor est considéré vérifié et la différentiation du code direct est correcte. La figure 8.8 présente le résultat de ce test pour notre cas d'étude avec le modèle POLAIR3D. La figure montre la convergence du modèle adjoint vers le résultat par différence finie avec une précision de l'ordre de 10^{-19} pour une valeur de perturbation de l'ordre de 10^{-7} , au delà de cette valeur de perturbation, l'importance des erreurs d'arrondies dans le modèle apparaît ce qui explique la légère augmentation de l'écart entre les deux approches (deuxième partie du graphe).

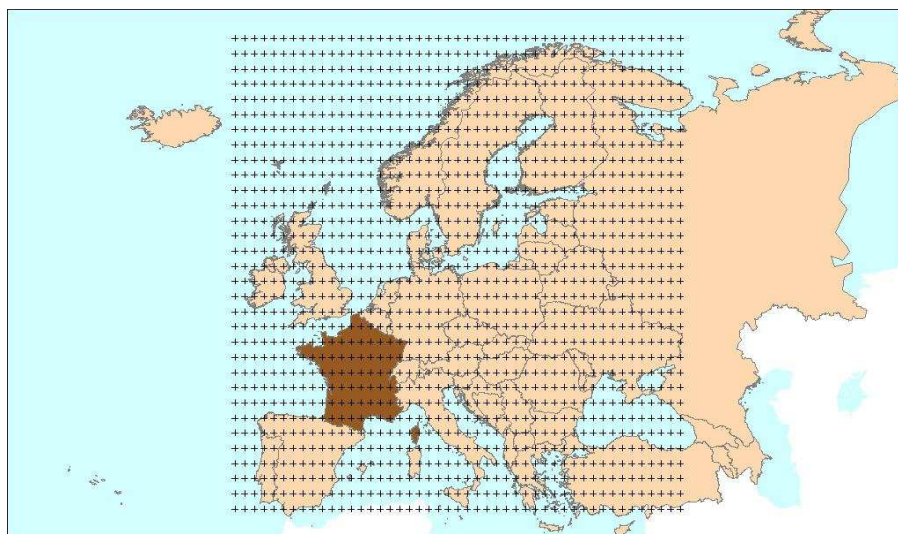


Figure 8.9: Domaine Européen maillé (47*32 points de grilles) : Impact des émissions des points maillant la France sur le reste de l'Europe

8.3.4 Application et résultats

Nous avons calculé par adjointisation pour chaque espèce émise X des matrices de transfert (de type Point de grille- Pays) associées à chaque espèce déposée Y , M^{XY} . Chacune de ces matrices est de dimension, $N_{pays} * (N_x * N_y)$, où N_{pays} est le nombre de pays récepteurs et $(N_x * N_y)$ est le nombre de points émetteurs.

Nous appliquons la méthode détaillée dans la section précédente sur le calcul de matrice de transfert par adjointisation à l'étude de l'impact de réduction des émissions de NO_x , SO_2 et de NH_3 mensuelle (Janvier 1994) de tous les points sources de la France (leur nombre est de 49, comme on peut le vérifier sur la Figure 8.9) sur la concentration dans l'air (ou déposée) mensuelle prévue sur les autres pays européens.

Nous avons calculé la matrice de transfert qui représente l'effet de chacun des points de grille (47*32) du domaine sur chacun des pays de l'europe, par l'adjoint du modèle POLAIR3D. L'utilité de cette matrice est de pouvoir identifier l'effet de la réduction des émissions d'un point du domaine (représentant par exemple une centrale thermique) sur le bilan annuel (mensuel) des concentrations dans l'air et du dépôt annuel (mensuel) sur un pays bien spécifique. Nous présentons quelques matrices (avec des valeurs arrondies pour des raisons de lisibilité) relatives à la France comme pays émetteur, pour illustration dans les tableaux suivants :

- les tableaux 8.3 et 8.4 représentent la matrice de transfert associée à l'émission

de l'espèce SO_2 sur la France et son impact sur la concentration dans l'air de SO_2 sur chaque pays européen (les codes ISO correspondant à chaque pays sont synthétisé dans le tableau 8.2).

AL	AT	BA	BE	BG	BY	CH	CZ	DE
Albania	Austria	Bosnia	Belgium	Bulgaria	Belarus	Switzerland	Republi Czech	Germany
DK	EE	ES	FI	FR	GB	GR	HR	HU
Denmark	Estonia	Spain	Finland	France	United Kingdom	Greece	Croatia	Hungary
IE	IT	LT	LV	MD	MK	NL	NO	PL
Ireland	Italy	Lithuania	Latvia	Moldova	Macedonia	Netherlands	Norway	Poland
PT	RO	RU	SE	SI	SK	TR	UA	YU
Portugal	Romania	Russia	Sweden	Slovenia	Slovakia	Turkey	Ukraine	Serbia and Montenegro

Table 8.2: Codes ISO correspondant à chaque pays Européen récepteur dans les matrices de transfert

- les tableaux 8.5 et 8.6 représentent la matrice de transfert associée à l'émission de l'espèce SO_2 sur la France et son impact sur la concentration dans l'air de NH_3 sur chaque pays européen. Nous constatons que tous les éléments de la matrice sont négatif, ce qui signifie qu'une diminution dans les émissions françaises de SO_2 engendrerait une augmentation dans les concentrations dans l'air de NH_3 sur plusieurs pays européen. cela s'explique par les deux réactions du schéma EMEP : $SO_2 \rightarrow SO_4^{2-}$ et $3NH_3 + 2SO_4^{2-} \rightarrow 2[(NH_4)_{1.5}SO_4]$ ($(NH_4)_{1.5}SO_4$: **sulfate d'ammonium**).
- les tableaux 8.7 et 8.8 représentent la matrice de transfert associée à l'émission de l'espèce NO_2 sur la France et son impact sur la concentration dans l'air de NO_2 sur chaque pays européen.
- les tableaux 8.9 et 8.10 représentent la matrice de transfert associée à l'émission de l'espèce NO_2 sur la France et son impact sur la concentration dans l'air de NH_3 sur chaque pays européen.
- les tableaux 8.11 et 8.12 représentent la matrice de transfert associée à l'émission de l'espèce NH_3 sur la France et son impact sur la concentration dans l'air de NH_3 sur chaque pays européen. On peut remarquer également que tous les éléments de la matrice sont négatif ce qui voudrait dire qu'une diminution des émissions de NO_2 se traduirait pas une augmentation dans la concentration dans l'air des pays correspondant (Réactions : $NO_2 + OH \rightarrow HNO_3$ et $NH_3 + HNO_3 \rightleftharpoons NH_4NO_3$ (NH_4NO_3 : Nitrate d'ammonium)).

Une première interprétation des ces matrices de transfert est de pouvoir détecter facilement les pays qui ne sont pas affectés par les stratégies de réduction des émissions d'un pays (dans notre cas le pays émetteur est la France). Par exemple pour les émissions de NO_x , leur réduction n'affecte pas les concentration dans l'air ou déposées en NO_2 des pays AL, BG, DK, EE, GR, HR, PL (Tableaux 8.7 et 8.8). Cependant la concentration en NH_3 de ces pays sera affectée par la réduction des émissions de NO_x (tableaux 8.9 et 8.10).

Afin d'identifier la précision et la validité de ces matrices de transfert nous avons repris les tests effectués auparavant dans la section 8.1 pour comparer les différents scénarios A, B et C en utilisant les matrices de transfert par rapport à la méthode traditionnelle où le modèle direct est simulé pour chaque scénario. L'utilisation de ces matrices se fait pour un jeu d'émission de référence $E_{ref}(N_{pays})$ de chaque pays européen selon :

$$\delta D_i(Y) = \sum_{j=1}^{N_{pays}} M_{ij}^{XY} \cdot \delta E_j(X) \quad (8.24)$$

Nous validons ces matrices en comparant aux résultats du test présenté dans la section 8.1 qui consiste à réduire les émissions de NO_x , SO_x et NH_3 de la France de 20%, 40%, 60%, 80% et 100% et à étudier l'impact sur les autres pays européen.

- La figure 8.10 représente l'impact de réduction des émissions françaises en NO_x , SO_2 et NH_3 sur la concentration dans l'air en France respectivement de NO_2 , SO_2 et NH_3 , et cela avec deux méthodes différentes : la méthode directe et par différentiation automatique. Les deux méthodes sont assez proches.
- De la même manière nous présentons dans la figure 8.11 l'impact de réduction des émissions françaises en NO_x , SO_2 sur la concentration de NH_3 dans l'air en France. Une stratégie de réduction des émissions NO_x , SO_2 a un impact très faible sur la concentration moyenne en France de NH_3 . La réduction de la concentration dans l'air est inférieure à 6% pour une réduction d'émissions françaises de NO_x et à 2% dans le cas du scénario de réduction des émissions françaises de SO_2 .
- La figure 8.12 montre l'impact de réduction des émissions françaises en NO_x , SO_2 et NH_3 sur la concentration dans l'air en Allemagne respectivement de NO_2 , SO_2 et NH_3 . Plus la réduction dans les émissions est importante plus l'écart entre les deux approches est relativement important, mais ca reste raisonnable. Les deux approches affichent des réductions dans la concentration en Allemagne inférieure à 4% pour le SO_2 et NO_2 et à 10 % pour le NH_3 .
- Dans la figure 8.13 est présenté l'impact de réduction des émissions françaises en NO_x , SO_2 sur la concentration de NH_3 dans l'air en Allemagne. Les deux méthodes montrent le même comportement sur la concentration de NH_3 ; une

diminution des émissions de NO_x et SO_2 Françaises engendre une augmentation dans la concentration de NH_3 sur l'Allemagne, mais les valeurs restent faibles inférieures à 2%.

8.3.5 Coût calcul de la matrice de transfert

Selon qu'on utilise le gradient ou le tangent linéaire, le coût calcul de l'adjointisation est différent :

– **Premier cas :**

si on s'intéresse au calcul de la dérivée de f_i sortie du modèle f au point i par rapport à la variable X en tout point du domaine ($\nabla_X f_i$) le temps CPU du modèle adjoint est de l'ordre de 5 fois le temps CPU du modèle direct.

– **Deuxième cas :**

Si on s'intéresse au calcul de la dérivée du modèle f en tout point du domaine par rapport à la variable X au point i ($\frac{\partial f}{\partial X_i}$), le coût d'adjointisation est de 2 fois le temps CPU du modèle direct f .

Dans notre cas nous nous intéressons au calcul de la dérivée d'une concentration au point i par rapport aux émissions en tout point du domaine (**premier cas**).

Si on cherche à étudier l'effet de réduction des émissions mensuelles ou annuelles d'un pays sur la concentration d'un nombre N_{esp} d'espèces dans l'air (ou déposée) à travers l'Europe, le temps de calcul demandé est alors de l'ordre :

$$T_{matrices} = N_{esp} * N_{pays} * 5 * T_{POLAIR3D} \quad (8.25)$$

où, $T_{POLAIR3D}$ est le temps CPU du code direct, N_{esp} est le nombre d'espèces de sorties, N_{pays} le nombre de pays Européens récepteurs. Cette évaluation comporte le calcul de tous les éléments de toutes les matrices de transferts M^{XY} , X espèce émise et Y espèce de sortie.

Cependant sans utilisation de cette technique de calcul de matrice de transfert par différentiation automatique, le coût de construction de telles matrices serait de :

$$T_{matrices} = N_{emis} * N_{scen} * \left(\sum_{i=1}^{N_{pays}} N_{pts}(i) \right) * T_{POLAIR3D} \quad (8.26)$$

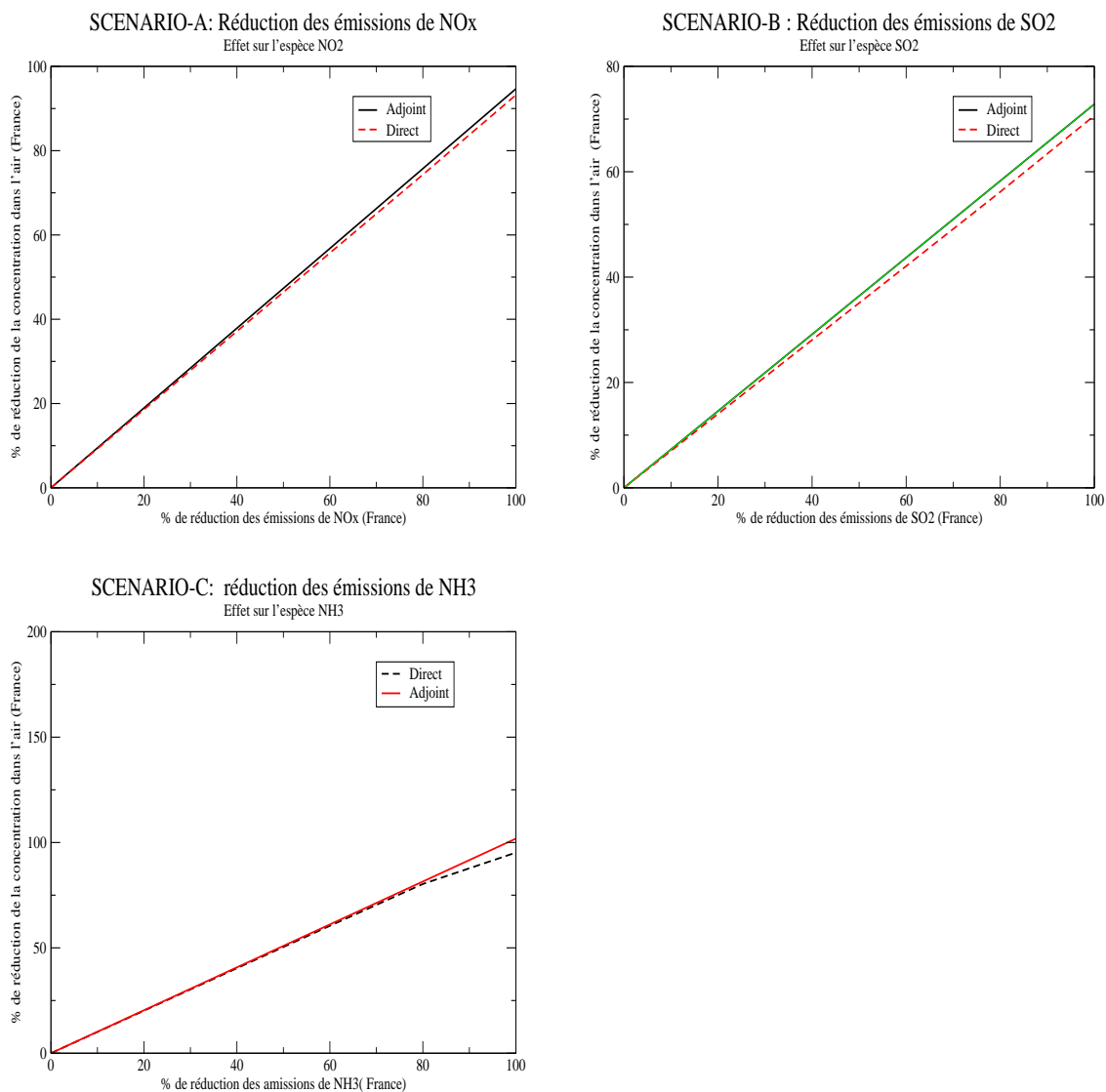


Figure 8.10: Impact de réduction des émissions Française de NO_x , SO_2 et NH_3 sur la concentration dans l'air des espèces réduites sur **la France**

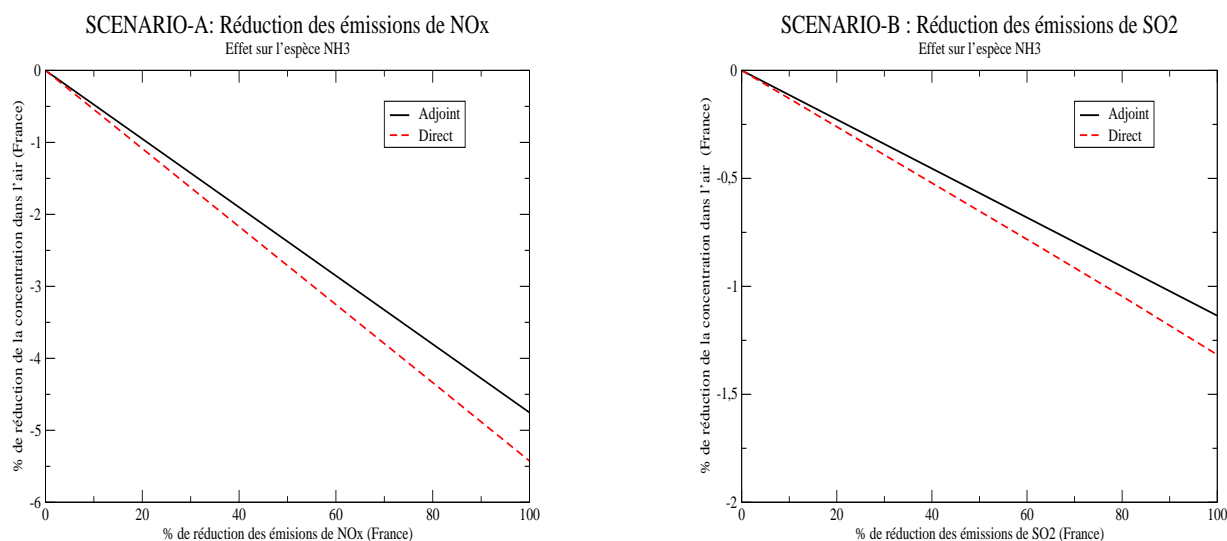


Figure 8.11: Impact de réduction des émissions Française de NOx , SO_2 sur la concentration dans l'air de NH_3 sur la France

où N_{emis} est le nombre d'espèce émises, $N_{pts}(i)$ est le nombre de points composants la pays i , N_{scen} est le nombre de scénario possibles. Nous rappelons que via l'utilisation de l'adjoint, les matrices de transfert sont calculées par rapport à tous les points (supposé sources) du domaine ($N_x * N_y$). Pour être plus objective, la comparaison du coût calcul par rapport à l'approche classique doit se faire à base de $T_{matrices} = N_{emis} * N_{scen} * (N_x * N_y) * T_{POLAIR}$, ce qui rend l'approche utilisant le calcul du gradient encore plus intéressante.

Conclusion

Nous avons présenté dans ce chapitre une méthode pour effectuer des calculs de matrices de transfert en utilisant l'adjoint du modèle avec une application au modèle Chimie-Transport POLAIR3D dans le cadre d'une étude d'impact à l'échelle européenne. Son utilisation pour estimer le taux de réduction des concentrations dans l'air sur l'Europe suite à un scénario de réduction des émissions françaises a donné de bons résultats par rapport à la version directe. Cette méthode demande des ressources de calcul importantes.

Sachant qu'une parallélisation est possible, en fonction des pays et des espèces à étudier, il est donc évident qu'une approche indispensable est de paralléliser ce calcul sur plusieurs machines pour réduire le temps d'évaluation de ces matrices à :

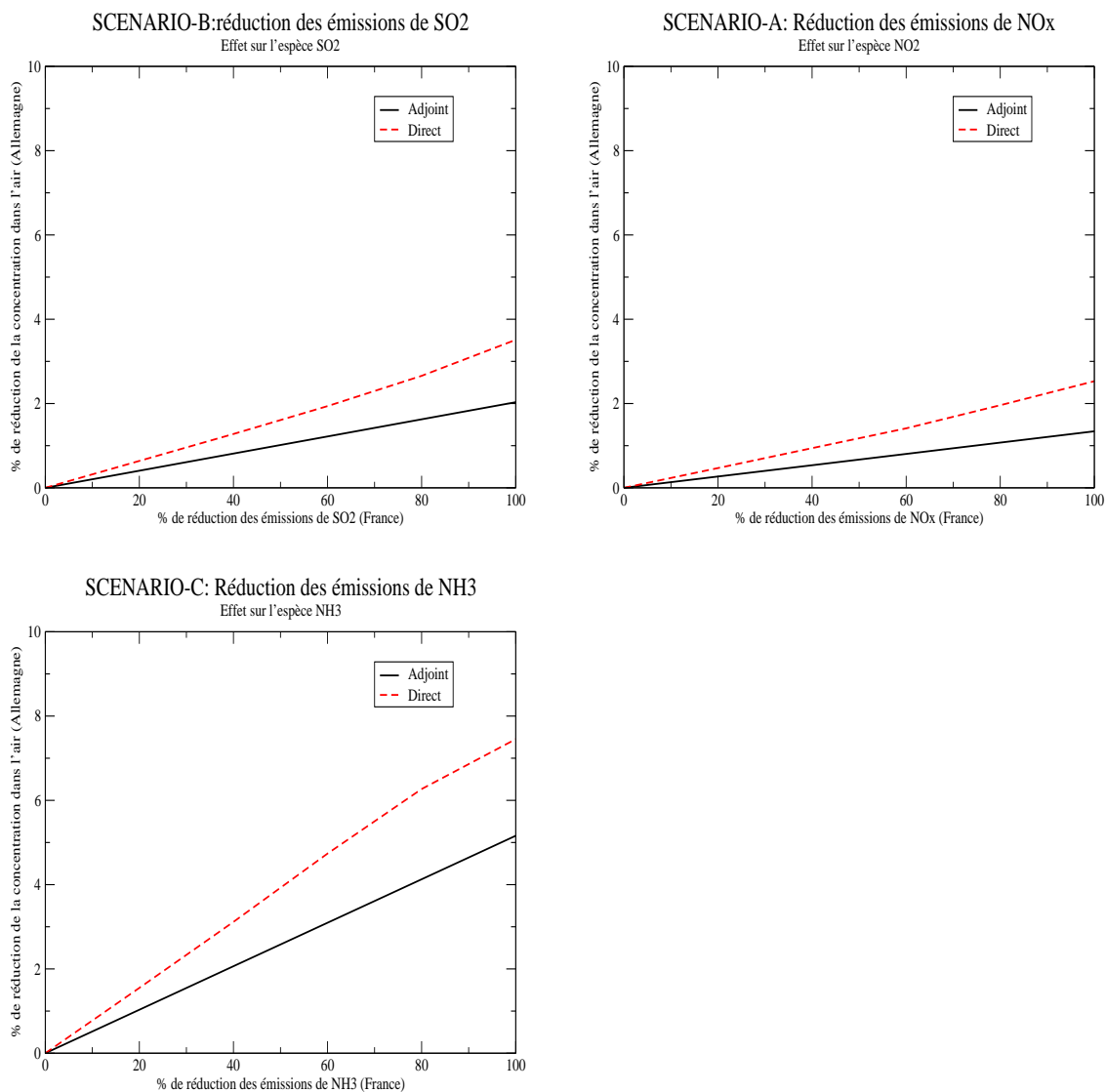


Figure 8.12: Impact de réduction des émissions Française de NO_x , SO_2 et NH_3 sur la concentration dans l'air des espèces réduites sur **l'Allemagne**

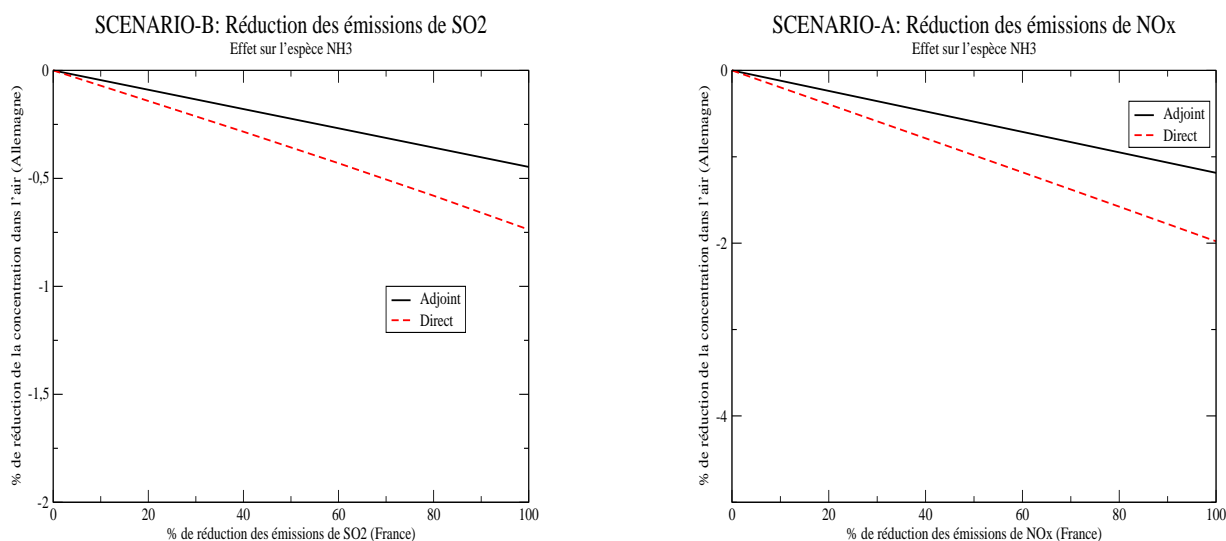


Figure 8.13: Impact de réduction des émissions Française de NO_x , SO_2 sur la concentration dans l'air de NH_3 sur l'Allemagne

$$T_{matrices} = N_{esp} * N_{pays} * 5 * \frac{T_{POLAIR3D}}{Nbre_{CPU}} \quad (8.27)$$

où $Nbre_{CPU}$ est le nombre de processeurs disponibles, nous reviendrons sur cette parallélisation dans le chapitre (9).

Dans un contexte opérationnel de modélisation intégrée d'impact, les matrices de transfert doivent être calculée sur une longue période en temps (5 à 10 ans) afin de mieux prendre en compte les variations possibles des champs de concentration d'une année à une autre. Cela rendrait très coûteux le temps de calcul POLAIR3D ($T_{POLAIR3D}$). Dans l'objectif de pouvoir réduire le temps $T_{POLAIR3D}$, nous allons nous intéresser dans le chapitre suivant à étudier la variabilité interannuelle des concentrations et la possibilité de représenter chaque mois de l'année par un jour type capable de capturer l'effet moyen du mois.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
AL	2	1	1	1	1	0	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	2	1	1
AT	3	3	3	4	6	13	3	2	3	4	6	8	15	3	2	2	3	3	4	8	2	2	2	2
BA	2	2	2	2	2	1	2	1	1	1	1	1	-1	2	2	1	1	1	1	2	2	2	2	1
BE	20	22	25	30	31	5	35	38	37	43	39	33	10	56	70	64	71	72	63	24	50	89	155	134
BG	1	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	1	1	1	1
BY	1	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1	1	1	1	1
CH	1	1	1	1	2	2	1	1	1	1	2	2	2	1	1	1	1	1	1	2	1	2	2	1
CZ	3	3	3	3	5	11	3	2	3	4	5	8	13	4	3	3	3	4	5	9	3	4	5	4
DE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ES	12	9	8	12	11	7	7	7	6	7	7	9	9	5	5	5	5	5	5	5	6	6	6	4
FI	1	1	1	1	1	1	1	1	1	1	2	2	2	2	1	1	1	2	2	2	2	2	2	2
FR	1747	2866	3087	2655	2199	1874	2144	655	713	531	189	294	819	-28	80	212	217	128	147	214	30	63	63	51
GB	18	19	17	9	6	1	30	25	19	10	5	2	1	52	33	23	19	10	5	3	87	85	65	32
GR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IT	3	2	3	7	9	13	1	2	2	5	7	11	16	2	1	1	2	4	7	10	3	2	1	1
LT	2	2	2	2	2	3	2	2	2	2	2	3	4	2	1	1	2	2	2	3	2	2	2	1
LV	1	1	1	1	2	2	2	2	1	2	2	2	3	2	2	1	2	2	2	2	2	3	3	2
MD	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1
MK	1	1	1	1	1	0	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1
NL	8	8	9	10	11	3	13	13	12	12	13	11	5	21	22	18	20	22	19	10	21	32	47	35
NO	5	4	4	4	5	3	6	6	5	5	5	5	5	8	8	7	6	5	5	5	10	11	13	8
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PT	13	6	4	3	2	1	6	4	2	2	2	2	2	5	3	2	1	1	1	1	6	6	4	2
RO	1	1	1	1	1	1	1	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1
RU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
SE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8.3: Matrice de Transfert : SO2-SO2 (Pays Emetteur : France)

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
AL	1	1	2	2	1	2	2	2	2	2	2	2	3	3	2	2	2	2	2	2	3	3	2	3	3
AT	3	4	5	10	2	2	1	2	2	2	3	4	6	14	1	1	2	3	3	3	5	9	2	2	2
BA	1	2	2	3	1	2	2	2	2	2	2	3	4	5	2	2	2	3	2	3	3	4	3	3	3
BE	128	177	171	92	33	50	102	209	327	349	342	419	230	94	84	243	358	639	602	1054	307	111	437	1047	1027
BG	1	1	1	1	1	1	1	2	2	1	1	2	2	3	1	2	2	2	2	2	2	2	2	2	3
BY	1	1	1	2	1	1	2	2	2	1	1	2	2	3	2	2	2	2	2	2	2	3	2	3	3
CH	1	2	2	3	1	1	2	2	2	2	2	2	3	4	1	2	2	3	2	3	3	5	2	2	3
CZ	4	6	8	14	2	3	3	3	4	4	6	8	13	30	2	3	4	5	5	6	9	27	3	3	3
DE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ES	3	3	3	3	4	4	4	4	3	2	2	2	1	1	3	2	2	2	1	1	0	0	1	1	1
FI	1	2	2	2	2	2	2	2	3	2	2	2	3	3	2	2	2	3	2	2	2	3	2	3	3
FR	64	48	58	72	7	7	9	19	25	15	14	9	9	7	6	9	11	11	5	3	2	1	5	5	2
GB	23	16	9	6	92	120	131	78	51	25	17	13	8	4	188	104	61	35	16	9	6	2	56	23	5
GR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IT	1	2	3	5	2	2	2	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1
LT	2	2	3	3	2	2	2	2	3	2	2	3	3	5	2	2	3	3	2	3	3	5	2	3	3
LV	1	2	2	3	2	2	3	3	3	2	2	2	3	4	3	3	3	3	2	2	3	3	3	4	4
MD	1	1	1	2	1	1	1	1	1	1	1	2	2	3	1	1	1	2	2	2	2	3	1	2	2
MK	1	1	2	2	1	1	2	2	2	2	2	2	3	3	2	2	2	2	2	2	2	3	2	3	3
NL	33	42	40	25	16	23	37	65	86	76	74	74	77	28	35	73	98	145	119	126	105	34	104	182	296
NO	8	8	7	7	10	11	12	13	14	9	8	8	9	8	13	14	16	13	9	8	9	8	16	17	15
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PT	1	1	1	1	4	4	3	2	1	0	0	0	0	0	2	1	1	0	0	0	0	0	0	0	0
RO	1	2	2	2	1	2	2	2	2	2	2	3	4	5	2	2	3	3	3	3	3	5	3	4	4
RU	0	0	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
SE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

8.3. Calcul de la matrice de transfert par modèle adjoint

Table 8.4: Matrice de Transfert : $SO_2 - SO_2$ (Pays Emetteur : France) (Suite)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
AL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AT	-1	-1	-1	-1	-1	-3	-1	-1	-1	-1	-2	-2	-4	-1	-1	-1	-1	-1	-2	-3	-1	-1	-1	-1
BA	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	-1	-1	-1	0
BE	-3	-3	-4	-4	-4	-1	-4	-5	-5	-6	-5	-4	-2	-5	-6	-7	-8	-8	-8	-5	-4	-7	-10	-10
BG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CH	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	-1	-1	-1	-1	-1	-1	-1	-1
CZ	-1	-1	-1	-1	-1	-2	-1	-1	-1	-1	-1	-2	-3	-2	-1	-1	-1	-1	-2	-2	-1	-1	-2	-1
DE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ES	-2	-1	-1	-2	-2	-1	-1	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
FI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FR	-26	-43	-51	-32	-19	-13	-21	-21	-25	-15	-8	-11	-17	-5	-7	-11	-11	-8	-9	-11	-3	-4	-6	-5
GB	-2	-2	-2	-1	-1	0	-3	-3	-2	-1	-1	0	0	-4	-3	-3	-2	-2	-1	-1	-4	-5	-4	-3
GR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IT	-1	-1	-1	-1	-2	-2	0	0	-1	-1	-2	-2	-3	-1	0	0	-1	-1	-1	-2	-1	-1	0	0
LT	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	-1	-1	0	-1	-1	-1	-1	-1	-1	-1	0
LV	-1	0	0	0	-1	-1	0	0	0	0	-1	-1	-1	-1	0	0	0	0	-1	-1	-1	-1	-1	0
MD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	-1	-1	0
MK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NL	-1	-1	-1	-2	-2	0	-2	-2	-2	-2	-2	-2	-1	-2	-2	-2	-3	-3	-3	-2	-2	-3	-4	-4
NO	-1	-1	-1	-1	-1	0	-1	-1	-1	-1	-1	-1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PT	-3	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	0	-1	-2	-2	-1	0
RO	-1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	-1	-1	-1	-1
RU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8.5: Matrice de Transfert : $SO_2 - NH_3$ (Pays Emetteur : France)

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
AL	0	0	0	-1	0	0	0	0	0	0	0	-1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0
AT	-1	-1	-2	-3	-1	-1	-1	-1	-1	-1	-1	-1	-2	-3	0	0	-1	-1	-1	-1	-1	-2	0	-1	0
BA	-1	-1	-1	-1	0	0	0	0	0	0	-1	-1	-1	-2	0	0	0	-1	-1	-1	-1	-1	0	-1	-1
BE	-12	-13	-13	-12	-3	-4	-6	-10	-15	-15	-18	-20	-17	-12	-4	-9	-13	-22	-20	-25	-16	-12	-8	-14	-15
BG	0	0	0	0	0	0	0	0	0	0	0	0	-1	-1	0	0	0	0	0	0	0	-1	0	0	0
BY	0	0	0	-1	0	0	0	0	0	0	-1	-1	-1	-1	0	0	0	0	0	0	0	-1	0	0	0
CH	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	-1	-1	-1	-1	-1	-1	-1	-2	-1	-1	-1
CZ	-2	-2	-3	-4	-1	-1	-1	-1	-1	-1	-2	-2	-3	-6	-1	-1	-1	-1	-1	-1	-2	-5	-1	-1	-1
DE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ES	-1	-1	-1	0	-1	-1	-1	-1	-1	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
FI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FR	-5	-5	-5	-6	-1	-1	-1	-2	-3	-2	-2	-1	-1	-1	-1	-1	-1	-1	-1	0	0	0	0	0	0
GB	-3	-2	-2	-1	-3	-4	-4	-3	-3	-2	-2	-2	-1	-1	-4	-3	-2	-2	-2	-2	-1	-1	-1	-1	0
GR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IT	0	-1	-1	-1	-1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LT	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
LV	0	0	-1	-1	-1	-1	-1	-1	-1	0	0	-1	-1	-1	-1	-1	-1	-1	0	0	-1	-1	0	-1	-1
MD	0	-1	-1	-1	0	0	-1	-1	-1	-1	-1	-1	-1	-1	0	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
MK	0	0	0	-1	0	0	0	0	0	0	0	-1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0
NL	-4	-5	-5	-4	-1	-2	-3	-4	-6	-5	-6	-7	-6	-4	-2	-4	-6	-9	-7	-7	-6	-4	-4	-7	-8
NO	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PT	0	0	0	0	-1	-1	-1	-1	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
RO	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	-1	-1	-1	-1	-1	-1	-1	-2	-1	-1	-1
RU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8.6: Matrice de Transfert : $SO_2 - NH_3$ (Pays Emetteur : France) (Suite)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
AL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AT	0	0	0	0	0	1	0	0	0	0	0	1	2	0	0	0	0	0	1	1	0	0	0	0
BA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BE	2	2	3	3	3	1	5	5	6	6	9	10	3	11	16	17	22	20	23	11	14	25	39	36
BG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CZ	0	0	0	0	0	1	0	0	0	0	1	1	2	0	1	0	1	1	1	2	0	0	1	1
DE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ES	1	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	0	0	0	0	1	1	1	1
FI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FR	2216	3434	3494	3665	3367	2971	3110	571	350	350	94	176	490	125	20	37	60	40	55	94	7	17	15	14
GB	3	4	3	2	1	0	8	6	5	2	1	1	0	16	13	11	6	4	2	1	44	37	30	15
GR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
LV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NL	1	1	1	1	1	0	2	2	2	2	3	3	1	3	4	4	6	6	6	4	6	9	11	8
NO	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	1
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PT	2	1	1	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0
RO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8.7: Matrice de Transfert : $NO_x - NO_2$ (Pays Emetteur : France)

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
AL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AT	0	0	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	1	2	0	0	0
BA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BE	44	25	81	46	11	20	41	91	129	137	208	305	120	52	51	151	211	350	382	1360	228	57	408	1171	1076
BG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CH	0	0	0	1	0	0	0	0	1	0	1	1	1	2	0	0	0	1	1	1	1	2	0	1	1
CZ	1	1	2	3	0	0	0	1	1	1	1	2	4	7	0	1	1	1	1	2	3	6	1	1	1
DE	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
DK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ES	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	1	1	1	0	1	1
FR	15	10	12	17	1	2	1	4	5	3	1	0	0	0	1	2	2	2	1	1	0	0	1	1	1
GB	10	8	5	4	64	81	99	58	36	18	12	9	6	3	180	113	57	32	15	7	5	2	73	28	11
GR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LT	0	0	1	1	0	0	0	0	1	0	1	1	1	2	0	0	1	1	1	1	1	1	2	1	1
LV	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1
MD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NL	8	8	12	11	5	8	15	28	35	32	35	24	28	16	17	44	60	85	71	37	51	19	81	155	231
NO	1	1	2	2	2	2	2	3	3	2	2	2	2	2	3	4	4	4	3	3	3	3	5	6	6
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8.8: Matrice de Transfert : $NO_x - NO_2$ (Pays Emetteur : France) (suite)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
AL	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	-1	0	0	0	0	0	-1	-1	-1	-1
AT	-1	-1	-2	-2	-3	-6	-1	-1	-1	-2	-3	-4	-7	-1	-1	-1	-1	-2	-2	-4	-1	-1	-1	-1
BA	0	0	0	0	0	0	-1	0	0	0	0	0	-1	-1	-1	-1	0	-1	-1	-1	-1	-1	-1	-1
BE	-4	-4	-6	-8	-7	-1	-5	-6	-8	-10	-9	-6	-3	-7	-9	-10	-15	-15	-13	-9	-6	-10	-16	-13
BG	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	-1	0	0	0	0	0	-1	-1	-1	-1
BY	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	-1	0	0	0	0	0	-1	-1	-1	-1
CH	-1	-1	-1	-1	-1	-2	-1	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
CZ	-1	-1	-1	-2	-2	-4	-2	-1	-1	-2	-3	-3	-5	-2	-2	-2	-2	-2	-3	-4	-2	-2	-3	-2
DE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ES	-2	-2	-2	-2	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	-1
FI	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	-1	-1	-1	0
FR	-42	-69	-78	-63	-37	-24	-32	-29	-30	-23	-14	-17	-23	-10	-11	-13	-15	-11	-12	-15	-4	-6	-9	-7
GB	-3	-4	-4	-3	-1	0	-5	-5	-4	-2	-1	-1	0	-6	-5	-5	-4	-3	-2	-1	-6	-7	-8	-5
GR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IT	0	0	0	-1	-1	-1	0	0	0	-1	-1	-1	-2	0	0	0	0	0	-1	-1	0	0	0	0
LT	-1	-1	-1	-1	-1	-2	-1	-1	-1	-1	-2	-2	-2	-1	-1	-1	-1	-1	-1	-2	-1	-1	-1	-1
LV	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
MD	-1	-1	-1	-1	-1	-1	-1	0	0	-1	-1	-1	-1	-1	-1	-1	0	-1	-1	-1	-1	-1	-1	-1
MK	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	-1	0	0	0	0	0	-1	-1	-1	-1
NL	-2	-2	-2	-3	-3	-1	-2	-2	-3	-4	-3	-2	-1	-3	-3	-3	-5	-6	-5	-3	-3	-4	-6	-5
NO	-1	-1	-1	-1	-1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	-1
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PT	-4	-2	-1	-1	-1	-1	-2	-1	-1	-1	-1	-1	-2	-1	-1	-1	-1	-1	-1	-1	-2	-2	-2	-1
RO	-1	0	0	0	-1	0	-1	0	0	0	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
RU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8.9: Matrice de Transfert : $NO_x - NH_3$ (Pays Emetteur : France)

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
AL	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	-1	-1	-1	-1	-1	-1	-1	-2	-1	-1	-1
AT	-2	-2	-3	-4	-1	-1	-1	-1	-1	-1	-1	-2	-2	-3	0	-1	-1	-1	-1	-1	-2	-3	-1	-1	-1
BA	-1	-1	-2	-2	-1	-1	-1	-1	-1	-1	-1	-1	-2	-2	0	-1	-1	-1	-1	-1	-2	-2	-1	-1	-1
BE	-18	-20	-23	-21	-4	-7	-10	-17	-25	-23	-28	-33	-28	-22	-8	-16	-21	-33	-28	-42	-25	-18	-14	-22	-25
BG	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-1	-1	-2	-2	-1	-1	-2
BY	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-1	-1	-2	-2	-1	-1	-2
CH	-1	-1	-1	-2	-1	-1	-1	-2	-2	-1	-2	-2	-3	-3	-1	-1	-2	-2	-2	-2	-3	-4	-1	-2	-2
CZ	-3	-3	-4	-5	-1	-1	-1	-2	-2	-2	-3	-3	-5	-7	-1	-1	-2	-2	-2	-3	-4	-6	-1	-2	-2
DE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ES	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
FI	0	0	0	0	0	-1	-1	-1	-1	0	0	0	0	-1	-1	-1	-1	-1	0	0	0	-1	-1	-1	-1
FR	-7	-6	-6	-6	-2	-2	-2	-3	-4	-2	-2	-1	-1	0	-1	-1	-1	-1	-1	0	0	0	-1	-1	0
GB	-5	-5	-4	-3	-5	-6	-8	-8	-7	-5	-4	-5	-3	-2	-8	-6	-5	-4	-3	-3	-2	-1	-2	-2	-1
GR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IT	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LT	-1	-1	-1	-2	-1	-1	-2	-2	-2	-1	-1	-1	-2	-3	-1	-2	-2	-2	-2	-2	-2	-2	-3	-2	-2
LV	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2
MD	-1	-1	-1	-1	-1	-1	-1	-1	-2	-1	-2	-2	-2	-3	-1	-1	-1	-2	-2	-2	-2	-2	-3	-1	-1
MK	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-1	-1	-2	-2	-1	-1	-2
NL	-6	-7	-8	-7	-2	-4	-5	-7	-10	-9	-10	-11	-10	-8	-4	-8	-10	-14	-11	-11	-10	-7	-8	-12	-15
NO	-1	-1	-1	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-1	-1	-2	-2	-2	-1	-1	-1	-1	-2	-2	-2
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PT	-1	0	0	0	-1	-1	-1	-1	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
RO	-1	-1	-2	-2	-1	-1	-1	-2	-2	-2	-2	-2	-3	-4	-1	-2	-2	-2	-2	-3	-3	-4	-2	-2	-3
RU	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	-1	-1	0	0	-1	-1	0	-1	-1
SE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8.10: Matrice de Transfert : $NO_x - NH_3$ (Pays Emetteur : France) (Suite)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
AL	0	1	1	1	1	0	2	1	1	0	0	0	0	3	2	1	1	1	1	1	2	3	3	3
AT	0	6	7	9	11	26	6	5	5	8	13	24	40	8	7	6	6	8	13	25	6	7	9	8
BA	0	2	2	2	2	1	3	2	1	1	1	3	4	3	2	2	2	3	3	3	4	5	4	
BE	0	35	50	65	53	7	37	52	74	90	85	49	9	88	106	134	172	186	171	90	72	133	222	227
BG	0	1	1	1	1	0	1	1	0	0	0	0	0	2	2	1	1	1	1	1	2	2	3	2
BY	0	1	1	1	1	0	1	1	0	0	0	0	0	2	2	1	1	1	1	1	2	3	3	2
CH	0	4	4	4	5	5	4	3	4	4	5	7	8	5	4	3	3	4	6	7	5	7	8	5
CZ	0	7	7	8	9	18	10	8	7	8	12	20	29	14	12	10	10	13	18	28	10	13	18	14
DE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ES	0	9	10	9	9	5	8	9	8	7	7	7	7	10	8	7	5	5	4	4	9	10	10	5
FI	0	1	1	1	1	2	2	1	1	1	2	2	2	2	2	1	1	1	2	2	2	3	3	2
FR	0	7073	7667	7265	5841	4290	5904	1604	1012	930	322	496	1306	181	226	267	326	194	216	279	48	73	126	106
GB	0	35	35	26	12	1	41	47	37	23	9	2	0	72	61	48	37	26	15	8	92	94	98	59
GR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IT	0	2	3	6	8	10	2	2	2	4	6	10	12	3	2	2	2	3	5	7	6	4	3	2
LT	0	4	4	5	6	6	4	4	4	5	7	9	11	6	5	4	4	5	7	9	6	7	8	5
LV	0	2	2	3	3	3	3	3	3	3	4	5	6	4	3	3	3	3	4	5	5	5	5	3
MD	0	2	2	2	2	1	2	2	2	2	2	2	2	3	3	2	2	2	3	2	3	4	5	4
MK	0	1	1	1	1	0	1	1	0	0	0	0	0	3	2	1	1	1	1	1	2	3	3	3
NL	0	13	17	22	19	3	13	18	22	27	26	14	5	31	29	34	47	54	51	31	29	46	62	59
NO	0	6	6	6	6	3	6	7	7	7	6	5	4	9	8	8	8	8	7	7	10	11	13	10
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PT	0	13	9	6	5	2	12	9	6	4	4	3	3	14	8	4	3	3	3	2	13	13	10	4
RO	0	2	2	2	2	1	2	2	1	1	1	1	0	4	3	2	2	2	2	2	4	5	5	4
RU	0	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	2	2	1	1
SE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8.11: Matrice de Transfert : $NH_3 - NH_3$ (Pays Emetteur : France)

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
AL	0	1	1	1	1	0	2	1	1	0	0	0	0	3	2	1	1	1	1	1	2	3	3	3	
AT	0	6	7	9	11	26	6	5	5	8	13	24	40	8	7	6	6	8	13	25	6	7	9	8	
BA	0	2	2	2	2	1	3	2	1	1	1	1	3	4	3	2	2	2	3	3	3	4	5	4	
BE	0	35	50	65	53	7	37	52	74	90	85	49	9	88	106	134	172	186	171	90	72	133	222	227	
BG	0	1	1	1	1	0	1	1	0	0	0	0	0	2	2	1	1	1	1	1	2	2	3	2	
BY	0	1	1	1	1	0	1	1	0	0	0	0	0	2	2	1	1	1	1	1	2	3	3	2	
CH	0	4	4	4	5	5	4	3	4	4	5	7	8	5	4	3	3	4	6	7	5	7	8	5	
CZ	0	7	7	8	9	18	10	8	7	8	12	20	29	14	12	10	10	13	18	28	10	13	18	14	
DE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
DK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
EE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
ES	0	9	10	9	9	5	8	9	8	7	7	7	7	10	8	7	5	5	4	4	9	10	10	5	
FI	0	1	1	1	1	2	2	1	1	1	2	2	2	2	2	1	1	1	2	2	2	3	3	2	
FR	0	7073	7667	7265	5841	4290	5904	1604	1012	930	322	496	1306	181	226	267	326	194	216	279	48	73	126	106	
GB	0	35	35	26	12	1	41	47	37	23	9	2	0	72	61	48	37	26	15	8	92	94	98	59	
GR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
HR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
HU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
IE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
IT	0	2	3	6	8	10	2	2	2	4	6	10	12	3	2	2	2	3	5	7	6	4	3	2	
LT	0	4	4	5	6	6	4	4	4	5	7	9	11	6	5	4	4	5	7	9	6	7	8	5	
LV	0	2	2	3	3	3	3	3	3	3	4	5	6	4	3	3	3	3	4	5	5	5	5	3	
MD	0	2	2	2	2	1	2	2	2	2	2	2	2	3	3	2	2	2	3	2	3	4	5	4	
MK	0	1	1	1	1	0	1	1	0	0	0	0	0	3	2	1	1	1	1	1	2	3	3	3	
NL	0	13	17	22	19	3	13	18	22	27	26	14	5	31	29	34	47	54	51	31	29	46	62	59	
NO	0	6	6	6	6	3	6	7	7	7	6	5	4	9	8	8	8	8	7	7	10	11	13	10	
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
PT	0	13	9	6	5	2	12	9	6	4	4	3	3	14	8	4	3	3	3	2	13	13	10	4	
RO	0	2	2	2	2	1	2	2	1	1	1	1	0	4	3	2	2	2	2	2	4	5	5	4	
RU	0	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	2	2	1	1	
SE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
SI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
SK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
TR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
UA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
YU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

8.3. Calcul de la matrice de transfert par modèle adjoint

273

Table 8.12: Matrice de Transfert : $NH_3 - NH_3$ (Pays Emetteur : France) (suite)

Chapitre 9

Modèles statistiques moyens : Journées et années représentatives

Résumé

L'objectif de ce chapitre est l'étude de la variabilité interannuelle et intermensuelle de la qualité de l'air à l'échelle européenne, dans l'optique de construire éventuellement des jours et des années types représentatifs.

Introduction

Le calcul des matrices de transfert se fait sur la base de simulation 3D du modèle de chimie transport utilisé sur de longues périodes allant de 5 à 10 ans ([88, 87]). La construction de ces matrices demanderait des ressources de calcul très importantes. Ceci nous amène à bien analyser les simulations préalables à la construction de ces matrices afin de pouvoir minimiser le nombre de calculs à effectuer. Dans cette optique, l'objectif de ce chapitre est :

- d'étudier la variabilité des concentrations dans l'air et des concentrations déposées calculées par le modèle POLAIR3D d'une année à une autre en gardant le cadastre des émissions constant. Si la variabilité d'une année à une autre des concentrations n'est pas très importante, cela nous permettrait de travailler sur uniquement une année, la plus proche de la moyenne.
- d'extraire pour chaque mois de l'année le jour "type" le plus représentatif du comportement moyen du mois de la concentration dans l'air/déposée calculée par POLAIR3D. Avec cette méthode, une simulation d'une année avec le modèle POLAIR3D, s'effectuera uniquement sur quelques jours (les jours représentatifs de chaque mois).

Pour effectuer cette étude, nous disposons de 9 années de données météorologiques issues de ECMWF de 1993 à 2001 sur le domaine européen. Nous utilisons le modèle POLAIR3D avec une chimie EMEP pour les neuf années. Nous nous intéressons aux concentrations dans l'air et déposées de chacune des espèces qui rentrent en jeu dans le mécanisme chimique.

Les émissions sont issues de la base de données d'EMEP, prétraitées par la suite pour une utilisation dans le modèle POLAIR3D (chapitre 3). Les espèces émises et utilisées dans le modèle sont NO_x , SO_x et NH_3 . Pour l'étude de variabilité interannuelle, nous avons supposé les émissions constantes et pris l'année 94 comme année de référence du point de vue des émissions sur l'Europe.

Dans un premier temps, nous avons étudié la faisabilité de représenter chaque mois par un jour type capable de capturer le comportement chimique moyen du mois ; dans un second temps, l'objectif était d'analyser la variabilité des concentrations des espèces chimiques d'une année à une autre et de pouvoir identifier pour chaque espèce l'année la plus proche de la moyenne des résultats obtenus avec les 9 années de calcul.

9.1 Les simulations annuelles de POLAIR3D

Les neuf années de calcul de 1993 à 2001 sont effectuées avec les conditions initiales et aux limites décrites dans le chapitre 3, avec le même jeu d'émissions de 1994, le pas de temps utilisé est de 900 secondes. Une moyenne des 9 années de calcul est également calculée.

Dans les figures 9.1, 9.2 et 9.3 nous présentons les profils journaliers des espèces NO_2 , SO_2 , NH_3 et sulfate d'ammonium sur respectivement, la France, la Grande Bretagne et une moyenne spatiale sur toute l'Europe (domaine de la simulation). Nous pouvons constater la forte variabilité d'une année à une autre, les variations dépassant souvent le double de la concentration moyenne et cela pour toutes les espèces présentées.

Au niveau annuel, la moyenne des concentrations dans l'air sur la France, la Grande Bretagne et une moyenne sur toute l'Europe sont exposées respectivement dans les figures 9.4, 9.5 et 9.6 pour les espèces NO_2 , SO_2 , Nitrate d'ammonium et sulfate d'ammonium. Comme on peut le constater, les variations d'une année à une autre du bilan annuel des concentrations, sont faible pour les espèces NO_2 et SO_2 , sur l'Europe elles sont inférieurs à 3.5% pour le NO_2 et 16% pour le SO_2 par rapport à la moyenne des 9 années. Par contre, ces variations sont plus importantes pour le nitrate et sulfate d'ammonium, sur l'Europe elles peuvent atteindre 70% pour le nitrate d'ammonium et 33% pour le sulfate d'ammonium, par rapport à la moyenne des 9 années.

Nous nous sommes intéressés aussi à la moyenne mensuelle des concentrations dans l'air sur tout le domaine européen. Dans la figure 9.7 est présentée la variation de la concentration mensuelle de NO_2 et SO_2 et cela pendant les 9 années de 1993 à 2001. Globalement on retrouve les mêmes profils et des variations qui peuvent atteindre d'une année à une autre les 30% pour le SO_2 et moins de 20% pour le NO_2 , il est à remarquer que les fortes variabilités mensuelles se produisent pendant la période d'automne et d'hiver.

9.2 Le choix d'une année type

Dans cette partie nous étudions la possibilité de déterminer à partir des 9 années de simulations l'année qui se rapproche le plus possible de la moyenne. Ainsi pour chaque champ de moyenne annuelle d'une espèce X, nous calculons la distance qui la sépare par rapport au centre de gravité (moyenne des 9 années de calcul) défini

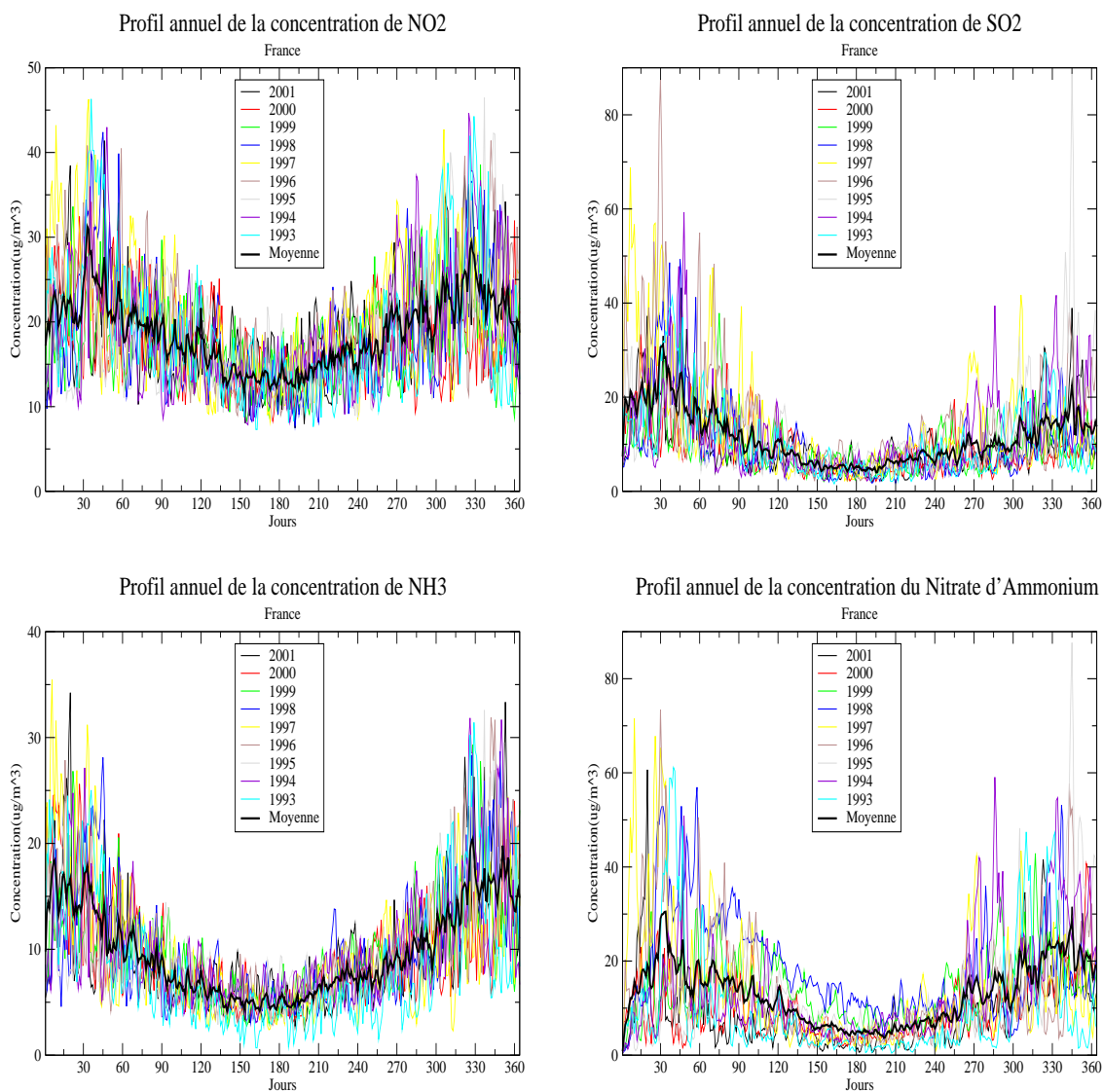


Figure 9.1: Profil de la concentration dans l'air annuelle de NO_2 , SO_2 , NH_3 et Sulfate d'ammonium sur la **France**

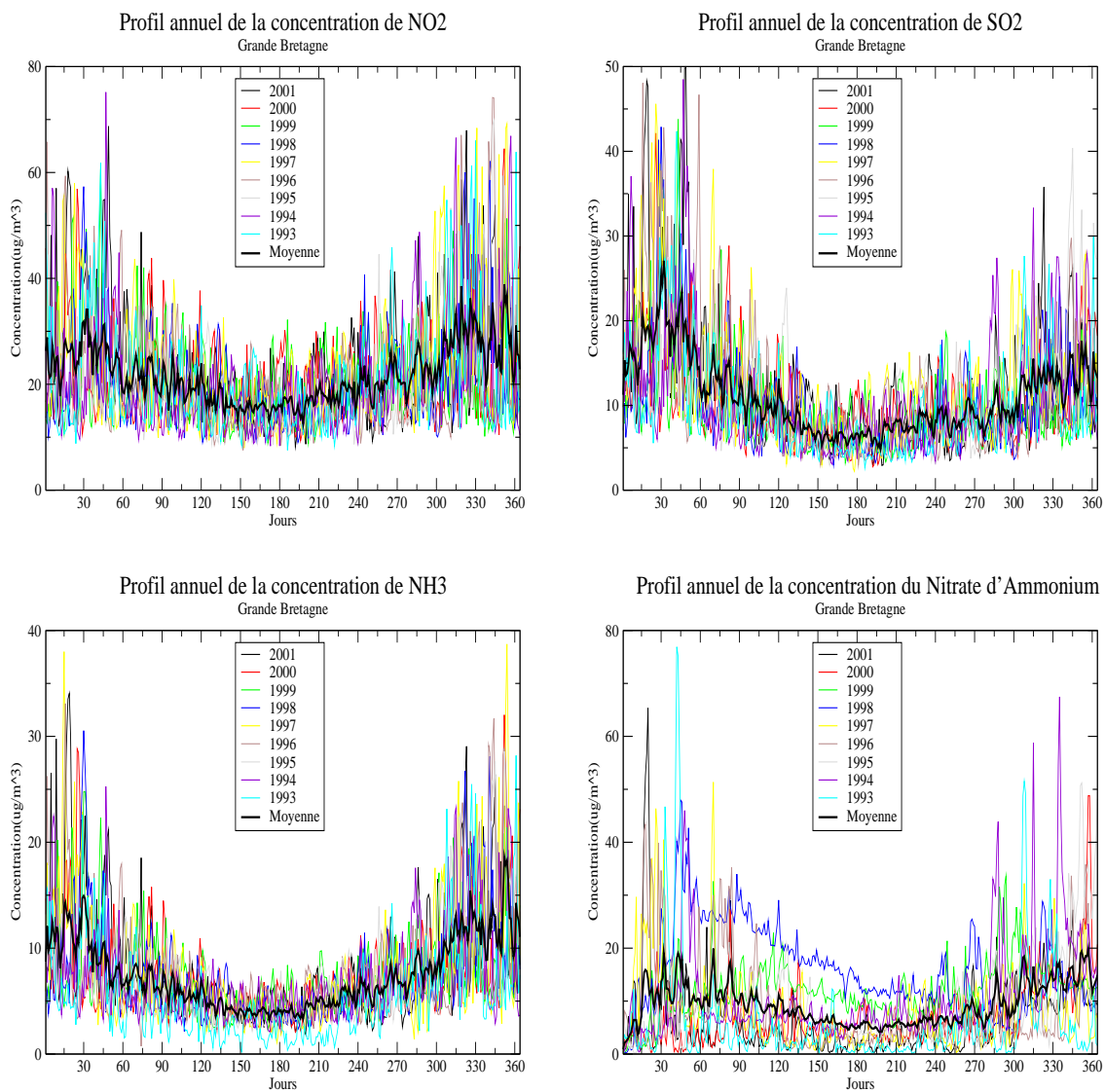


Figure 9.2: Profil de la concentration dans l'air annuelle de NO_2 , SO_2 , NH_3 et Sulfate d'ammonium sur la **Grande Bretagne**

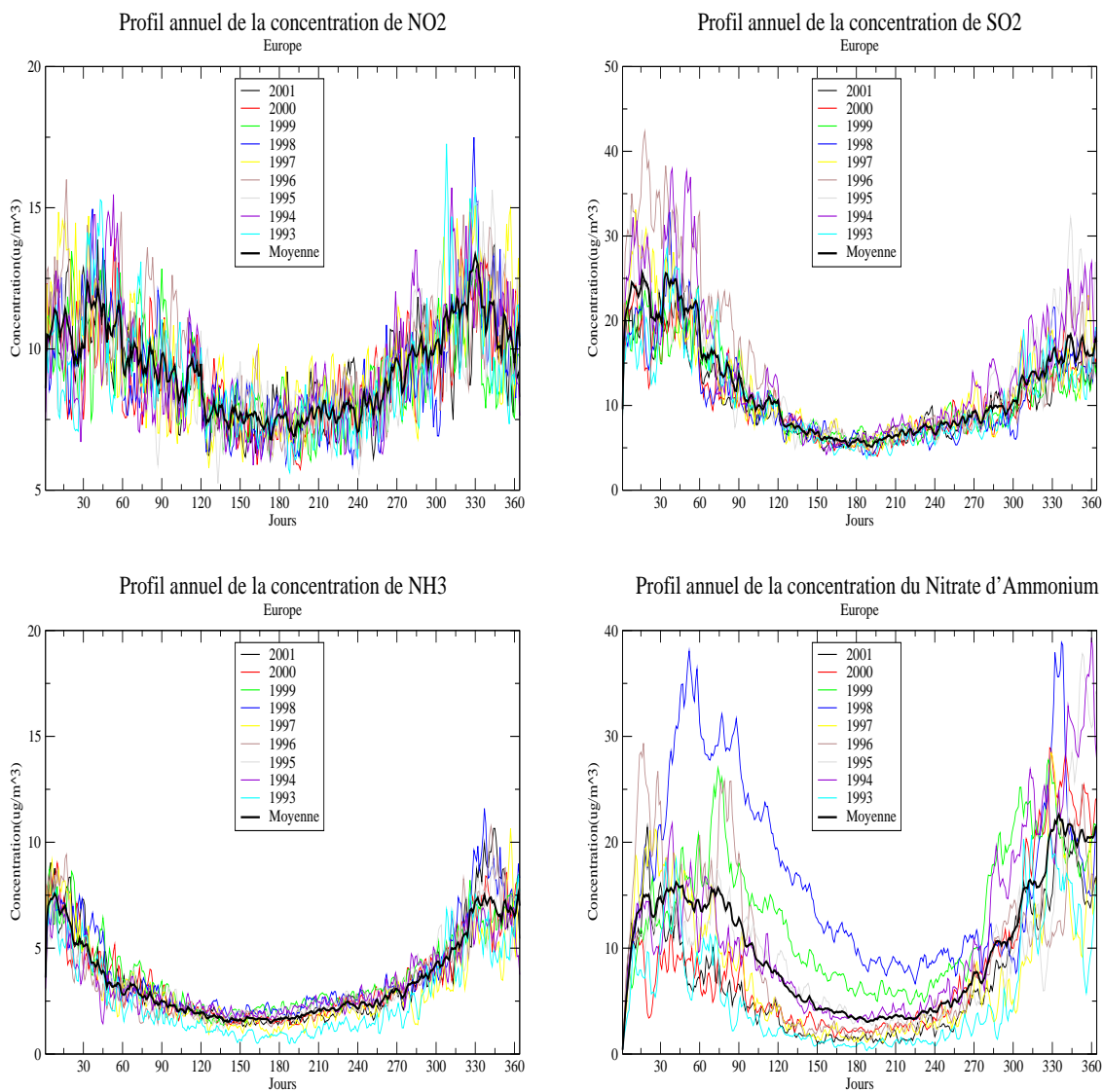


Figure 9.3: Profil de la concentration dans l'air annuelle de NO_2 , SO_2 , NH_3 et Sulfate d'ammonium sur la **Toute l'Europe**

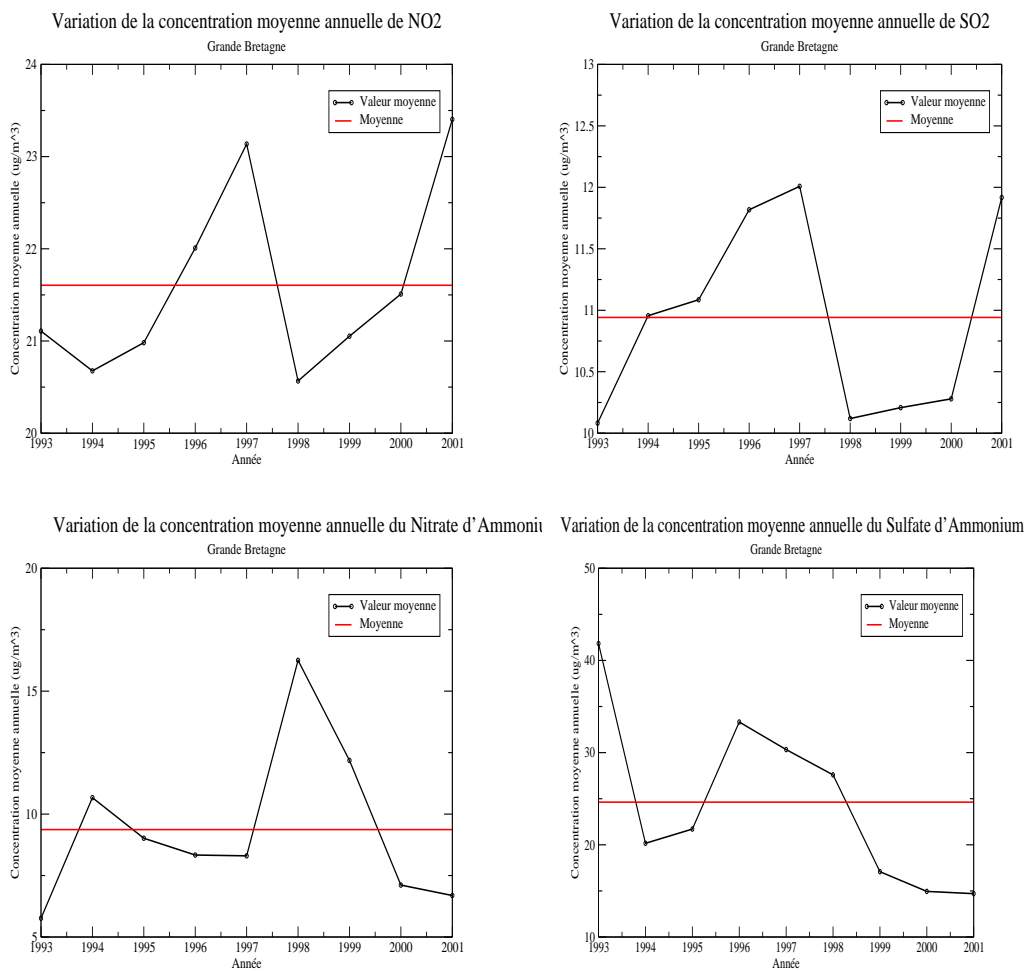


Figure 9.4: Variation de la concentration moyenne annuelle de NO_2 , SO_2 , Nitrate d'ammonium et Sulfate d'ammonium sur la **Grande Bretagne**

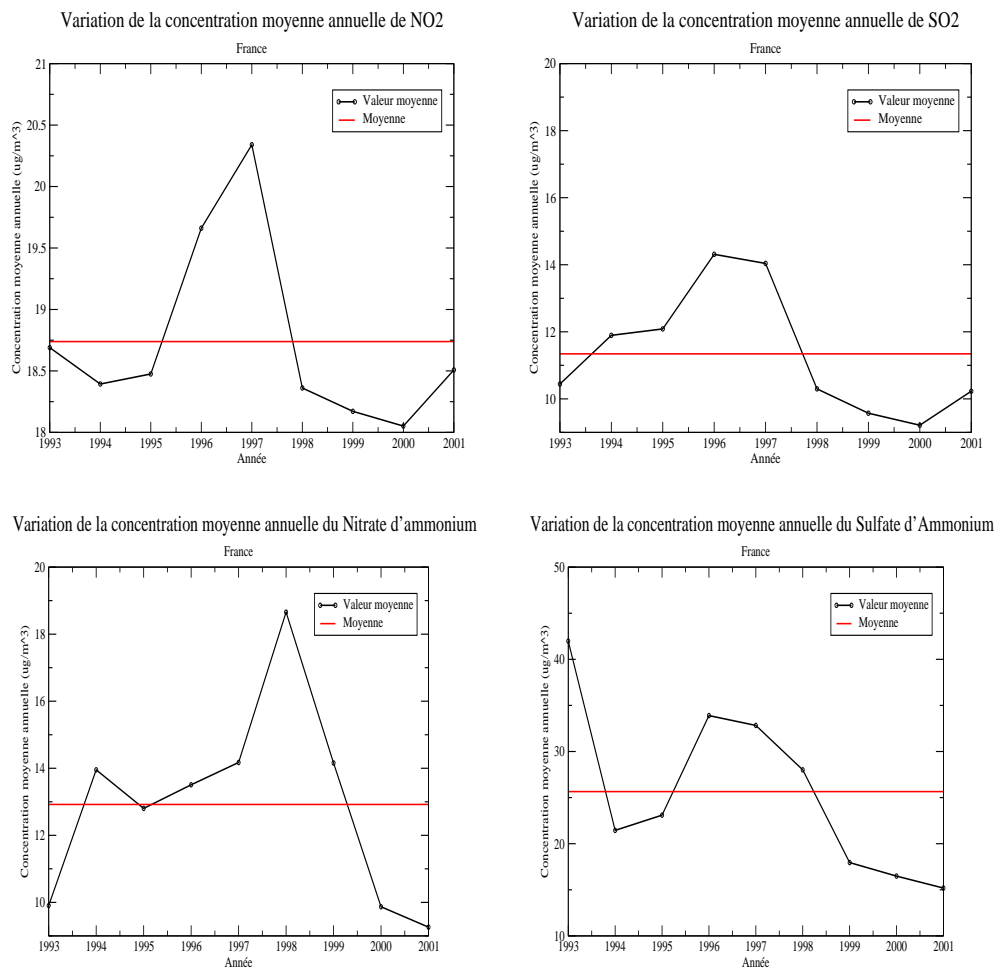


Figure 9.5: Variation de la concentration moyenne annuelle de NO_2 , SO_2 , Nitrate d'ammonium et Sulfate d'ammonium sur la **France**

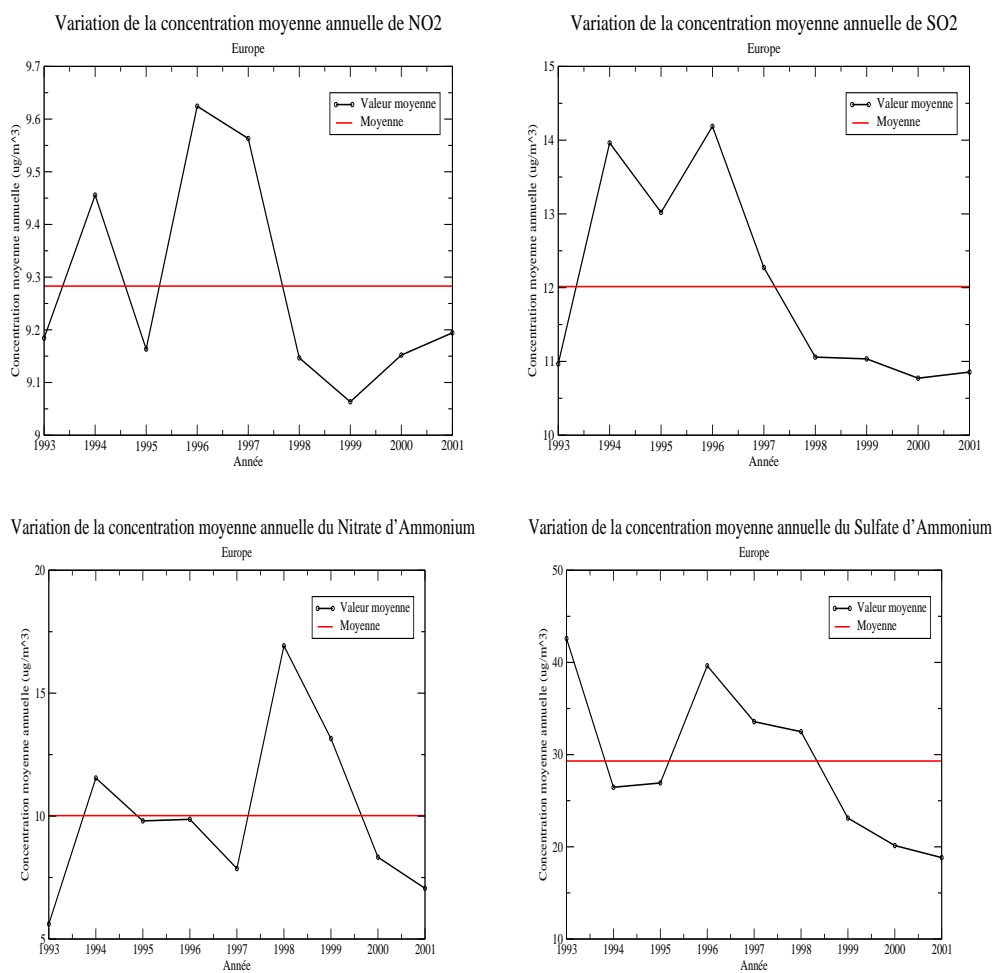


Figure 9.6: Variation de la concentration moyenne annuelle de NO_2 , SO_2 , Nitrate d'ammonium et Sulfate d'ammonium sur la l'Europe

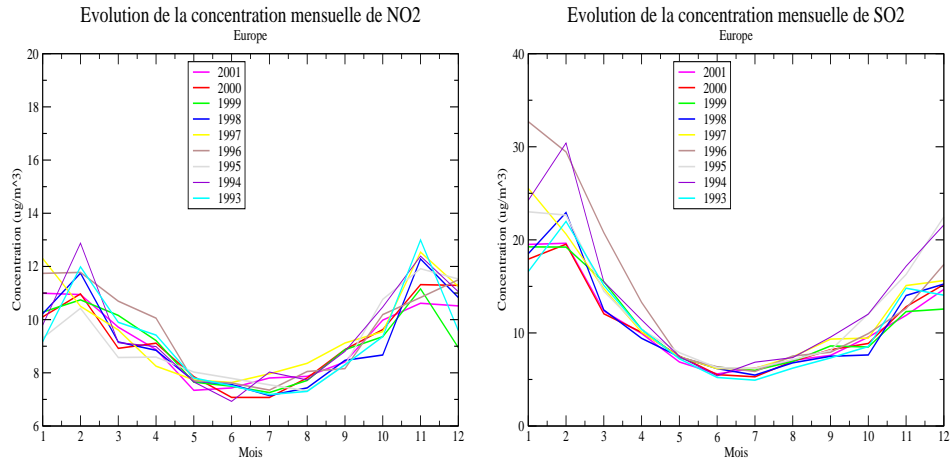


Figure 9.7: Variation de la concentration mensuelle de NO_2 , SO_2 sur l'Europe de 1993-2001

par :

$$CG_X = \frac{1}{N} \sum_{K=1}^{K=N} X(K) \quad (9.1)$$

Il existe plusieurs types de distances pour mesurer l'éloignement d'un champs par rapport à une valeur moyenne. Dans notre étude nous utilisons la distance euclidienne définie pour $X(N_x, N_y)$ (concentration journalière d'une espèce chimique), par :

$$d(X, CG_X) = \sqrt{\left[\sum_{j=1}^{N_y} \sum_{i=1}^{N_x} |X(i, j) - CG_X(i, j)|^2 \right]} \quad (9.2)$$

Pour chaque espèce chimique transportée par le modèle POLAIR3D nous avons procédé au calcul de la distance de sa concentration annuelle (sur l'Europe) des 9 années utilisées (1993-2001) par rapport à la moyenne des simulations. Dans la figure 9.8 est schématisée l'évolution de la distance de la concentration annuelle européenne par rapport à la moyenne des 9 années de simulation, pour NO , NO_2 , SO_2 et NH_3 . Dans le tableau 9.1, nous présentons pour chaque espèce l'année la plus proche du centre de gravité pour la distance choisie. Nous remarquons que l'année 1995 est souvent la plus proche de la moyenne des simulations, pour les 9 espèces du schéma EMEP. Pour l'espèce SO_4^{2-} , l'année 1996 est l'année la plus proche de la moyenne.

Ce résultat nous permet de dire que le choix d'une année type parmi N_{an} nombre de simulations à émissions annuelle constante, s'avère très difficile, puisque la varia-

bilité d'une année à une autre de la concentration d'une espèce est très importante. En présence d'une mesure de distance (type distance euclidienne) la détermination d'une année la plus proche de la moyenne des N_{an} années de simulations, dépend de l'espèce chimique à laquelle on s'intéresse, ce qui nous pousse à définir des espèces agrégées ou des indicateurs spécifiques qui seront pris directement comme sortie du modèle POLAIR3D.

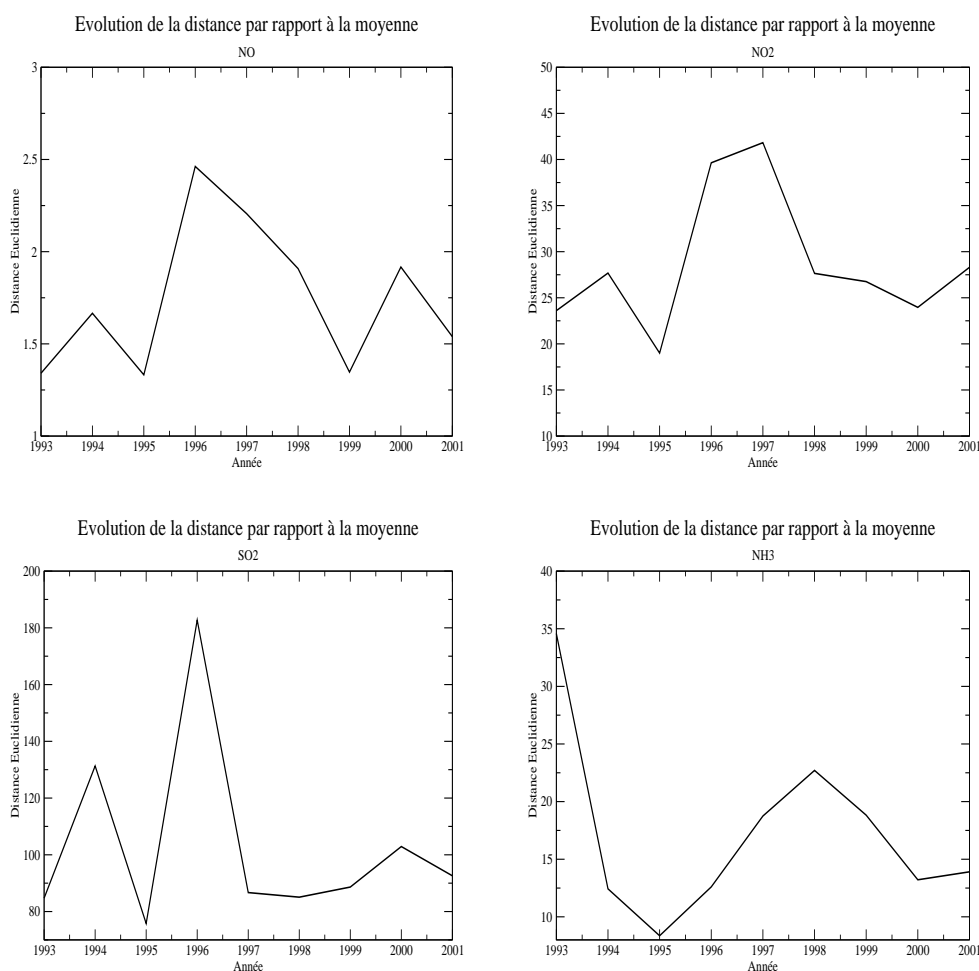


Figure 9.8: Evolution de la distance (euclidienne) de la concentration moyenne annuelle sur l'Europe par rapport à la moyenne de 1993-2001

Espèce	Année	Distance
<i>NO</i>	1995	1.33212543
<i>NO₂</i>	1995	19.0010376
<i>HNO₃</i>	1995	6.13531876
<i>PAN</i>	1995	31.7140484
<i>NO₃</i>	1995	28.1273613
<i>SO₂</i>	1995	75.7711105
<i>SO₄</i>	1996	17.2893772
<i>NH₃</i>	1995	8.36053276
<i>NITRAM</i>	1995	22.508213
<i>SULFAM</i>	1995	95.4573898

Table 9.1: l'année la plus proche de la moyenne des 9 années de simulations (1993-2001) pour chacune des espèces transportées du schéma EMEP. NITRAM =Nitrate d'ammonium et SULFAM= Sulfate d'ammonium.

9.3 Choix d'un jour "type" par mois

Dans cette partie nous étudions la possibilité de déterminer à partir des 9 années de simulations, pour chaque mois le jour le plus représentatif.

Il existe plusieurs méthode de classification statistique dans la littérature, on peut citer la méthode de Ward [93, 94], la méthode du "plus proche voisin" [?] ou la méthode classique d'Analyse en Composante Principale (ACP). Un aperçu plus détaillé sur les différentes techniques statistiques utilisées pour la classification automatique des champs est disponible dans la thèse de C. Philippe [95].

Dans notre cas, nous avons utilisé une méthode proche de la méthode dite du "plus proche voisin". Pour chaque champ de concentration moyenne journalière des espèces transportées d'une espèce X, nous calculons la distance (définie par l'équation 9.2) qui la sépare par rapport au centre de gravité mensuel (moyenne mensuelle des 9 années de calcul) défini pour chaque mois par :

$$CG_X = \frac{1}{31} \frac{1}{N} \sum_{i=1}^{i=N} \sum_{j=1}^{j=31} X(i, j) \quad (9.3)$$

Notons que, a contrario de l'approche choisie par [95], nous ne déterminons pas un jour type sur des critères météorologiques (les entrées) mais sur des critères chimiques (les sorties). Cette approche nous semble plus appropriée du fait du caractère complexe de l'influence complexe des champs météo sur la chimie. En reprenant les notations déjà utilisées au chapitre 4, on note $C = F(K)$ la relation entre les champs de forçage K (météo) et les concentrations C.

L'objectif est de trouver \bar{K} tel que :

$$J_1(\bar{K}) = \min_K J_1(K) = \|F(K) - \int F(K')P(K')dK'\|^2 \quad (9.4)$$

avec $P(K)$ la densité de probabilité supposée connue de K. C'est exactement ce que réalise notre algorithme en se restreignant à chercher \bar{K} parmi les jours "réels". Une approche de classification sur les champs météo revient à choisir \tilde{K} tel que :

$$J_2(\tilde{K}) = \min_K J_2(K) = \|K - \int P(K')dK'\|^2 \quad (9.5)$$

avec un choix approprié de la norme afin de pondérer les variables météo. Il n'y a aucune relation évidente entre \bar{K} et \tilde{K} .

Avant de commencer l'analyse proprement dit, nous commençons par une présentation de l'évolution journalière de la concentration dans l'air sur la France et une moyenne sur toute l'Europe pour le NO_2 et SO_2 . Les différents mois choisis Janvier-Avril-Juillet-Octobre correspondent aux quatre saisons Hiver-Printemps-Ete-Automne. Les résultats sont présentés dans les figures 9.9, 9.10 respectivement de NO_2 et SO_2 pour la France, les mêmes résultats sur toute l'europe (une moyenne spatiale) sont dans les figures 9.11, 9.12 pour NO_2 et SO_2 . La variabilité d'une année à une autre est aisément constatée, le mois de Juillet est le mois où la variabilité est moins importante et cela pour les deux espèces étudiées sur la France et sur toute l'Europe.

Pour calculer le jour type, le plus représentatif d'un mois donné et pour une espèce chimique donnée, il faut :

- calculer la moyenne mensuelle pour chaque espèce selon l'équation 9.3,
- pour chaque jour du mois M , calculer la distance de ce jour (en terme de concentration) qui le sépare au centre de gravité du mois selon l'équation 9.2,
- trouver parmi les 30 (ou 31) jours du mois, le jour J_{type}^M où la distance calculée est minimale. Ce jour J_{type}^M sera supposé le jour qui représente le mieux la variation de la concentration de l'espèce étudiée au cours du mois M . Il est à noter que ce jour Type sera dépendant de l'espèce chimique (tableaux 9.2 et 9.3) et du domaine géographique d'intérêt. Si nous nous intéressons à la France (tableaux 9.2, 9.3) , à la Grande Bretagne (tableaux 9.4, 9.5) ou à l'Europe (tableaux 9.6, 9.7), le jour type ne sera pas forcément identique. Donc il est important de souligner que **le choix d'un jour type dépend de l'espèce chimique étudiée et de la région de calcul**. Pour surmonter cet obstacle, il serait judicieux de définir des indicateurs agrégés de plusieurs espèces chimiques qui peuvent servir comme sorties finales du modèle POLAIR3D.

Une fois calculé le jour type pour chacune des espèces, nous avons évalué la qualité de notre choix en comparant les résultats des moyennes mensuelles obtenus avec une simulation du modèle tout au long de l'année 2001 et les résultats obtenus avec les jours types des différentes espèces (NO_2 , SO_2 , Nitrate d'Ammonium et Sulfate d'Ammonium) . Cette comparaison est faite sur tout le domaine européen, les résultats sont présentés dans la figure 9.13 ; nous constatons que les deux méthodes suivent le même profil, l'écart entre les deux méthodes peut atteindre 20% pour le NO_2 , 12% pour le SO_2 , 50% pour le Nitrate d'Ammonium et 30% pour le Sulfate d'Ammonium.

La même comparaison est faite sur l'Europe mais en utilisant le jour type

calculé à base de la France (le calcul des distance est fait sur le domaine spatial de la France). Cela permettrait d'identifier les capacités du jour type calculé pour une espèce donnée sur le France à représenter les concentrations sur toute l'Europe. Le résultat est présenté dans la figure 9.14 ; généralement le jour type respecte assez bien l'évolution mensuelle du modèle et l'écart est comparable à celui obtenu avec la comparaison précédente.

Remarque

Lors d'une étude d'impact à plusieurs scénarios et en supposant que le choix du jour "type" par mois a été fait avec succès et qu'il soit le même (cas idéal) pour toutes les espèces, le problème de la détermination des conditions initiales pour chacune des 12 simulations reste à résoudre. Une solution raisonnable semble être l'utilisation de données climatologiques moyennes.

Mois	2001	2000	1999	1998	1997	1996	1995	1994	1993
1	28	21	17	24	17	14	13	10	28
2	18	24	2	16	28	1	2	25	26
3	27	11	26	22	14	14	9	7	7
4	13	2	28	14	19	19	25	27	9
5	8	17	25	20	23	17	21	29	26
6	26	20	26	26	8	29	6	25	16
7	13	27	8	23	6	16	12	15	23
8	11	23	11	29	4	6	10	20	1
9	20	23	6	7	11	18	13	20	25
10	27	8	20	18	28	13	21	17	26
11	2	15	8	30	15	15	27	6	12
12	23	30	29	6	19	18	30	18	3

Table 9.2: Jour le plus représentatif de chaque mois pour les différentes années (1993-2001) pour NO_2 sur la France

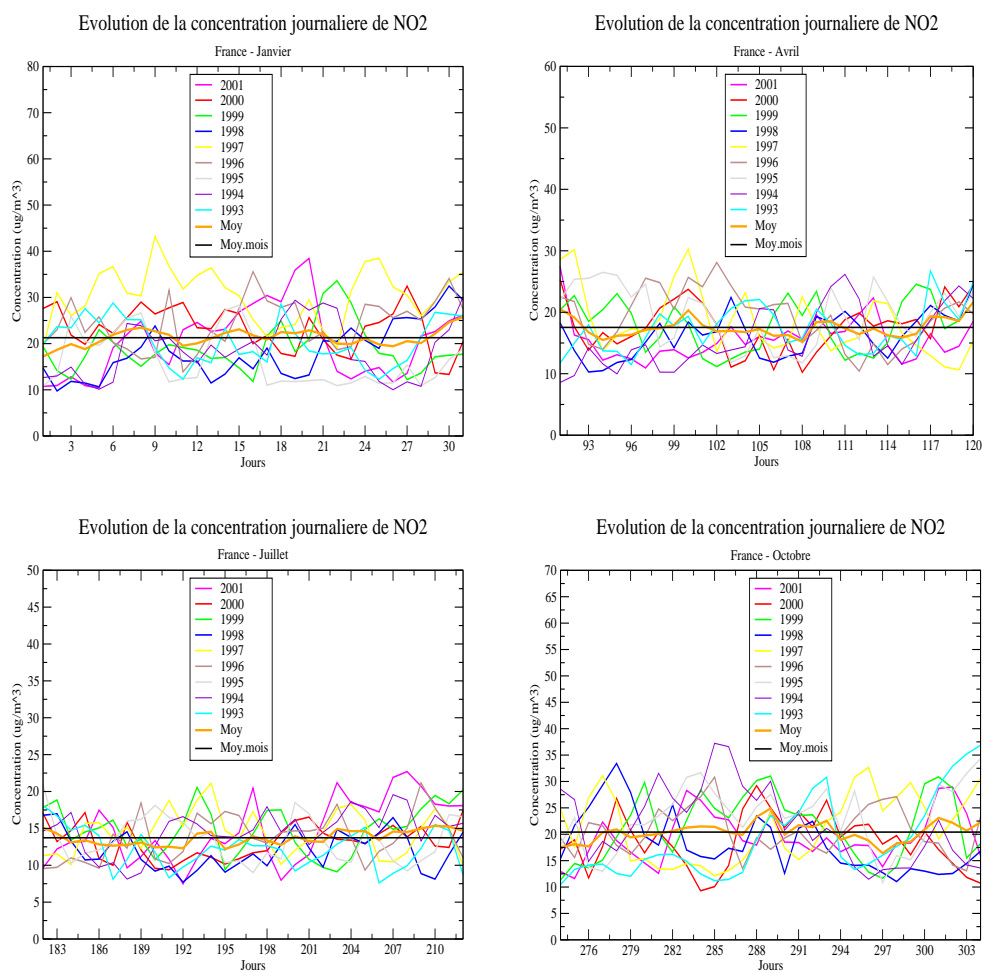


Figure 9.9: Evolution de la concentration journalière de NO_2 sur la France au cours du mois de Janvier, Avril, Juillet et Octobre de 1993-2001.

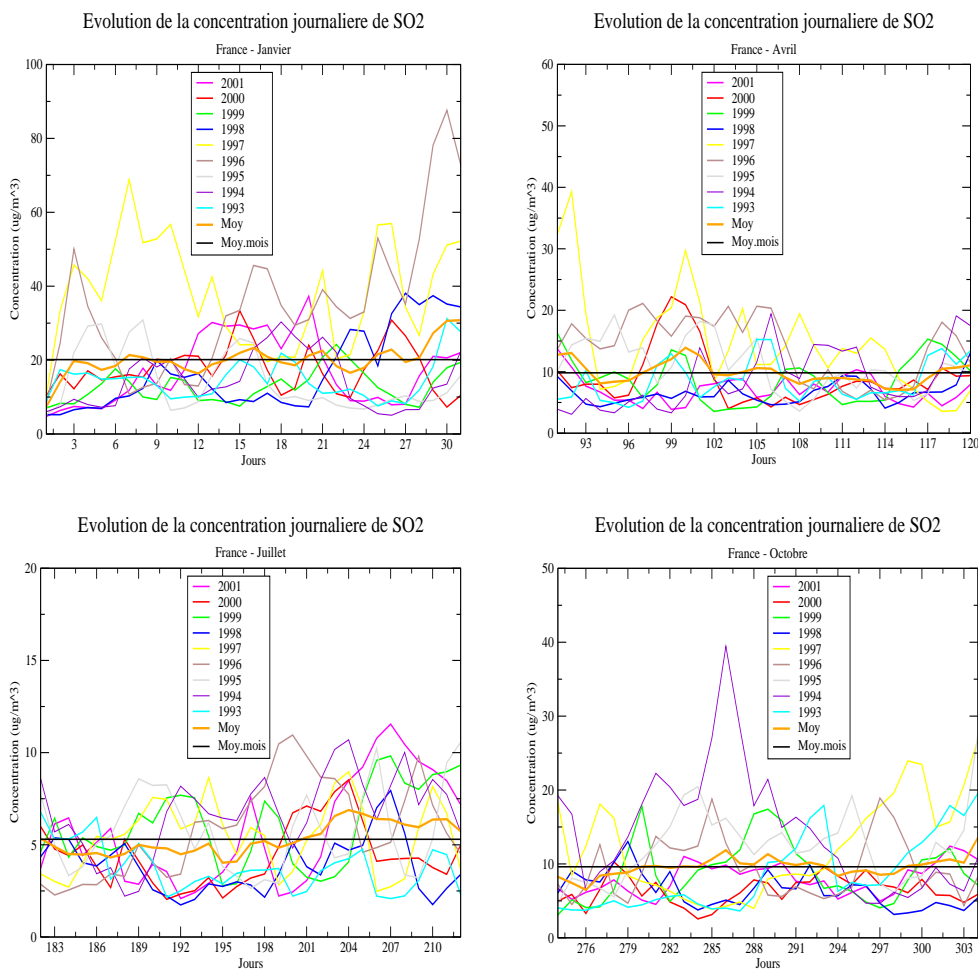


Figure 9.10: Evolution de la concentration journaliere de SO_2 sur la France au cours du mois de Janvier, Avril, Juillet et Octobre de 1993-2001.

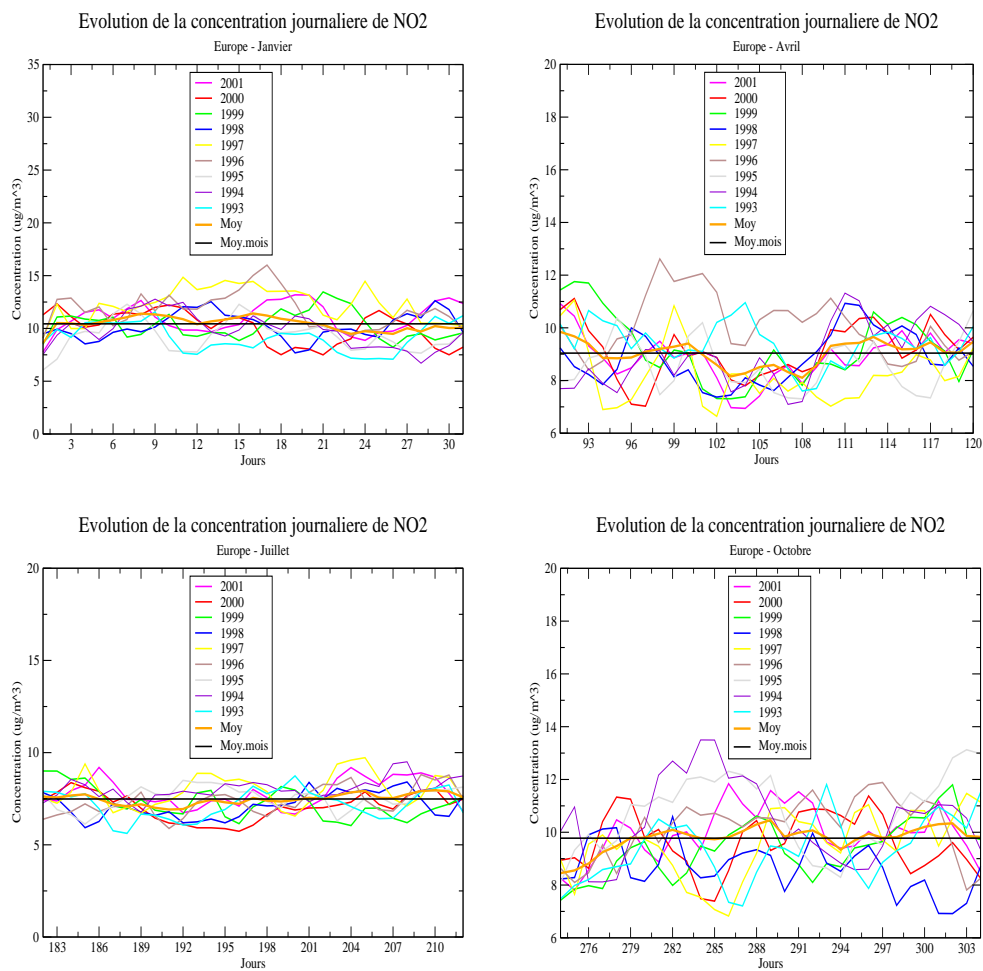


Figure 9.11: Evolution de la concentration journalière de NO_2 sur toute l'Europe au cours du mois de Janvier, Avril, Juillet et Octobre de 1993-2001.

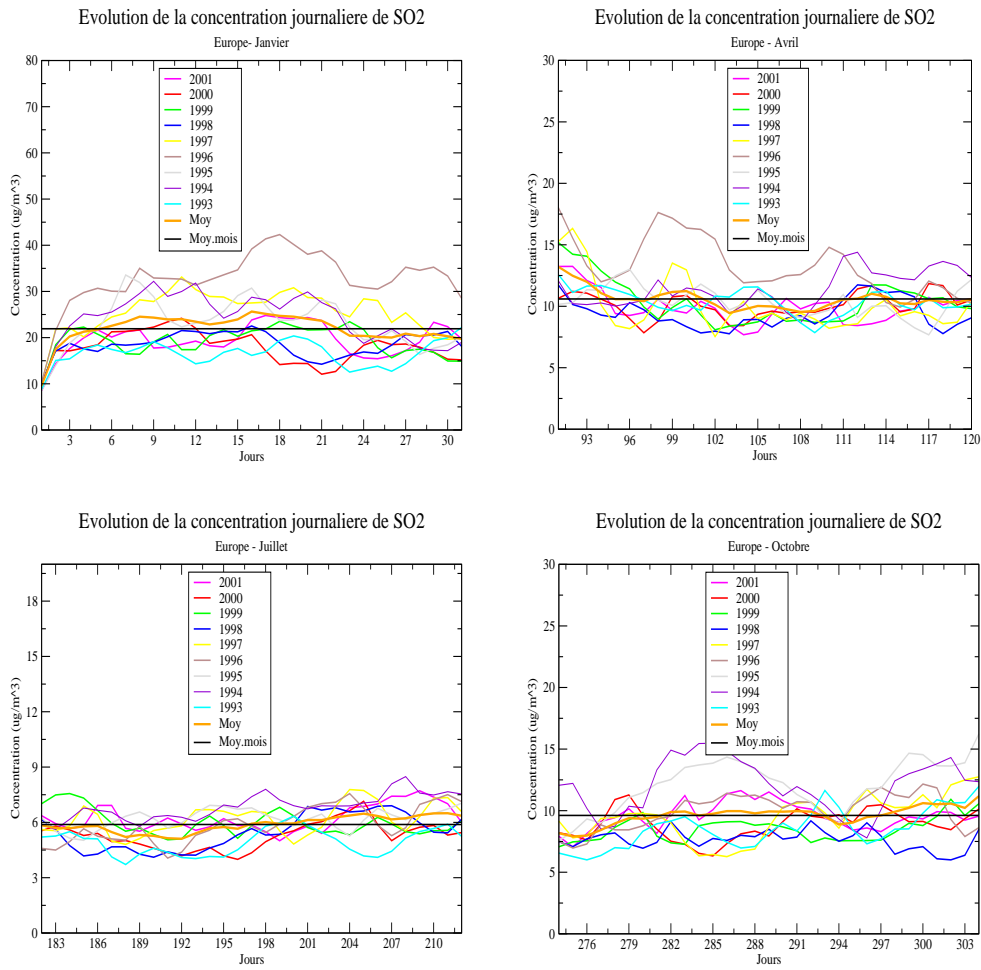


Figure 9.12: Evolution de la concentration journalière de SO_2 sur toute l'Europe au cours du mois de Janvier, Avril, Juillet et Octobre de 1993-2001.

Mois	2001	2000	1999	1998	1997	1996	1995	1994	1993
1	30	9	23	9	22	10	6	10	15
2	17	4	15	11	16	9	4	11	17
3	3	29	19	27	14	24	10	5	11
4	23	1	5	21	20	20	30	17	10
5	29	17	18	30	13	2	15	7	3
6	19	6	26	24	6	29	14	11	20
7	5	7	4	3	6	26	29	4	3
8	17	11	25	20	25	27	6	14	16
9	2	27	7	18	7	13	24	25	17
10	16	20	19	19	20	15	6	28	25
11	13	16	9	7	25	23	22	6	7
12	7	26	22	16	15	14	25	13	5

Table 9.3: Jour le plus représentatif de chaque mois pour les différentes années (1993-2001) pour SO_2 sur la France

Mois	2001	2000	1999	1998	1997	1996	1995	1994	1993
1	13	3	9	13	17	5	30	11	7
2	26	21	14	19	7	14	25	28	18
3	11	6	20	28	6	25	8	26	28
4	2	25	9	16	21	14	12	19	3
5	17	3	30	11	25	27	30	22	26
6	26	15	26	1	12	25	22	2	12
7	31	13	19	3	2	12	19	1	2
8	3	14	24	19	30	6	18	25	30
9	28	27	19	12	26	12	15	18	20
10	29	12	24	20	8	10	27	5	6
11	1	7	7	6	28	23	2	9	2
12	29	22	14	15	22	28	22	20	1

Table 9.4: Jour le plus représentatif de chaque mois pour les différentes années (1993-2001) pour NO_2 sur la Grande Bretagne

Mois	2001	2000	1999	1998	1997	1996	1995	1994	1993
1	6	18	24	12	18	27	27	10	7
2	21	14	8	11	14	22	7	8	18
3	19	26	20	17	5	14	30	26	4
4	26	19	26	23	25	27	2	30	12
5	7	11	14	10	22	29	16	12	1
6	4	5	19	18	13	29	21	26	25
7	7	21	4	16	28	14	5	27	22
8	4	19	18	2	30	12	23	2	15
9	23	18	22	3	20	23	22	12	6
10	19	13	26	3	28	24	15	5	8
11	24	21	7	27	28	19	29	3	10
12	14	16	27	23	1	30	23	4	20

Table 9.5: Jour le plus représentatif de chaque mois pour les différentes années (1993-2001) pour SO_2 sur la Grande Bretagne

Mois	2001	2000	1999	1998	1997	1996	1995	1994	1993
1	15	27	9	17	30	26	14	17	30
2	1	11	18	22	10	6	11	27	28
3	3	27	1	22	3	31	5	7	16
4	10	11	10	19	8	30	4	11	11
5	2	13	7	26	25	8	21	12	9
6	23	29	26	19	19	1	29	29	3
7	21	31	17	28	26	31	24	5	15
8	9	25	17	15	15	3	15	6	7
9	20	9	5	4	11	24	7	6	2
10	9	7	31	3	7	8	23	6	7
11	20	24	15	10	3	25	13	26	21
12	15	8	21	24	25	26	21	13	3

Table 9.6: Jour le plus représentatif de chaque mois pour les différentes années (1993-2001) pour NO_2 sur toute l'Europe

Mois	2001	2000	1999	1998	1997	1996	1995	1994	1993
1	5	12	17	11	2	2	23	26	31
2	17	23	15	3	2	17	18	28	25
3	12	14	3	30	26	30	17	30	11
4	30	24	27	21	24	26	8	7	27
5	10	12	9	3	4	19	11	17	23
6	14	20	4	11	29	16	29	27	5
7	11	25	12	19	11	18	19	1	30
8	10	11	8	10	27	29	4	10	20
9	27	30	19	2	11	2	10	12	28
10	31	7	30	19	8	20	20	21	19
11	30	13	27	30	15	16	8	1	21
12	27	1	28	17	8	23	25	30	28

Table 9.7: Jour le plus représentatif de chaque mois pour les différentes années (1993-2001) pour SO_2 sur toute l'Europe

9.4 Stratégie proposée pour le calcul des matrices de transfert

Avec le résultat obtenu dans les sections précédentes sur la variabilité d'une année à une autre de la concentration dans l'air (à émission annuelle constante) et sur la possibilité de représenter un calcul mensuel de POLAIR3D par un calcul sur quelques jours "types", nous proposons la stratégie suivante pour le calcul des matrices de transfert avec différentiation automatique.

- Identifier les espèces émises responsables et qui pourraient avoir un impact négatif (NO_x , SO_2 , NH_3 , CO, COV,...), et choisir une année de référence pour les émissions.
- Choisir les espèces de sortie à étudier, dans le cas de la problématique d'acidification transfrontalière à l'échelle européenne, on s'intéresse aux espèces : NO_2 , SO_2 , Nitrate d'ammonium et Sulfate d'ammonium.
- Préciser les pays récepteurs et les pays émetteurs objet de l'étude d'impact.
- Effectuer N_{an} (environ 10) années de simulation avec le modèle POLAIR, à émission annuelle constante.
- Pour chaque espèce de sortie, identifier l'année la plus proche de la moyenne des N_{an} années de simulations.

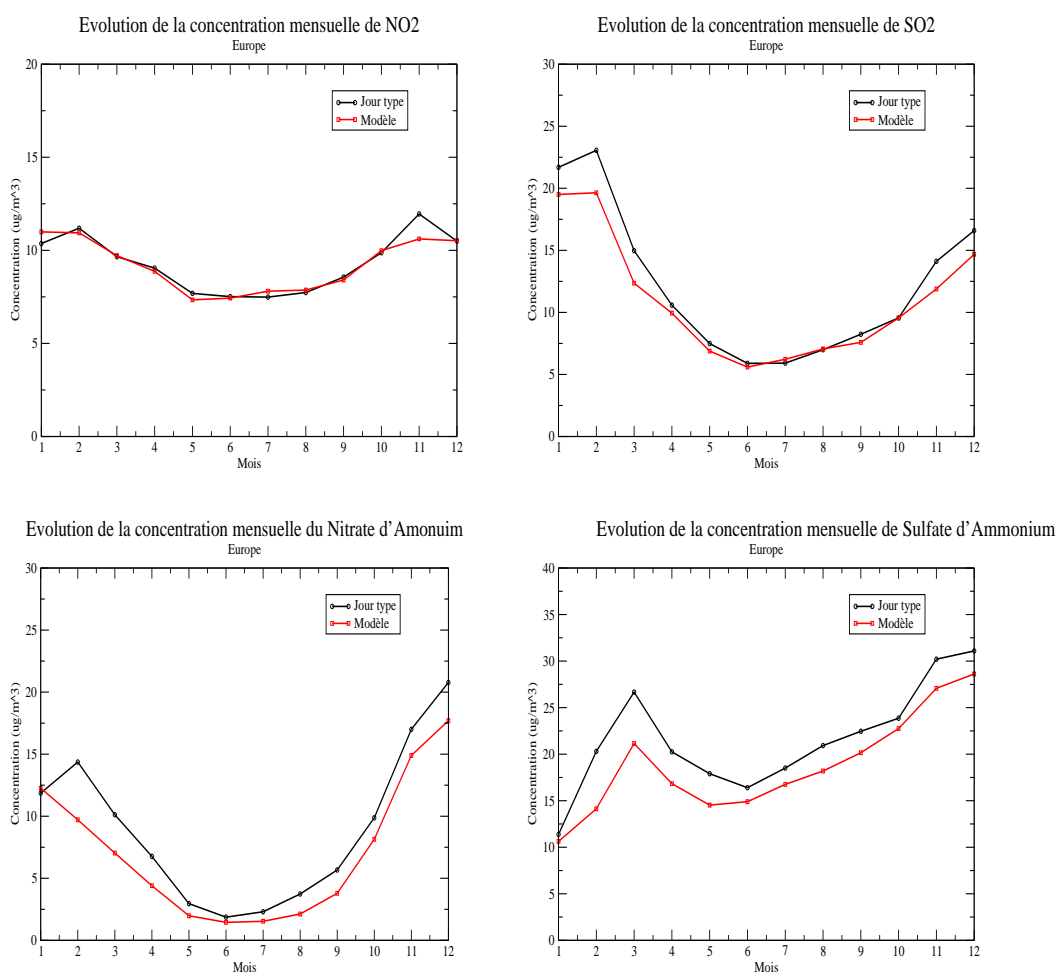


Figure 9.13: Comparaison du résultat "Jour-Type" (Europe) et "Modèle" de l'évolution de la concentration mensuelle de NO_2 , SO_2 , Nitrate d'Ammonium et du Sulfate d'Ammonium sur l'Europe (année 2001). Pour chacune des espèces est pris son jour type calculé

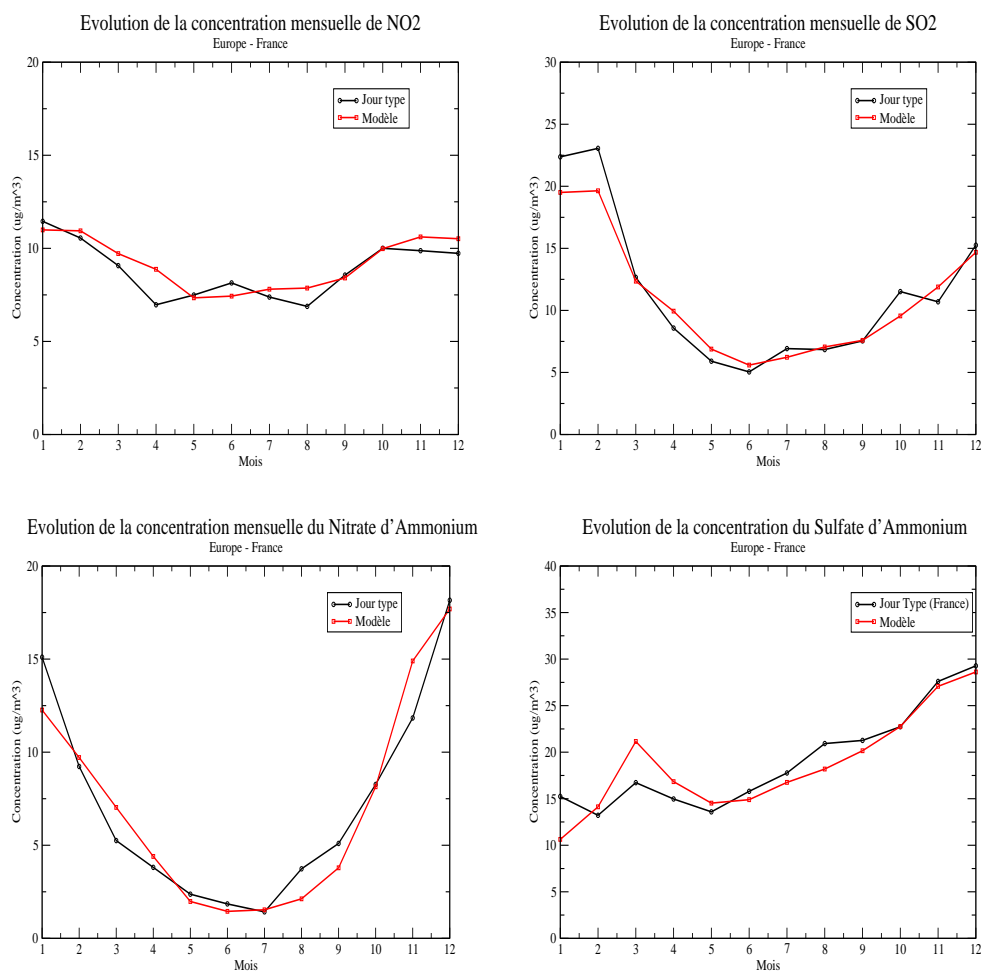


Figure 9.14: Comparaison du résultat "Jour-Type" (France) et "Modèle" de l'évolution de la concentration mensuelle de NO_2 , SO_2 , Nitrate d'Ammonium et du Sulfate d'Ammonium sur l'Europe (année 2001). Pour chacune des espèces est pris son jour type calculé

- Pour chaque espèce de sortie, calculer les matrices de transferts (selon la méthode de différentiation automatique) avec l'année correspondante (la plus proche de la moyenne). Cela peut se faire d'une manière parallèle, une parallélisation par pays récepteur et/ou par espèce de sortie.

Si nous disposons de N_{CPU} processeurs, de N_{pays} pays récepteurs et N_{out} espèces chimique de sortie ($N_{out} \leq N_{esp}$ nombre total des espèces chimique du mécanisme utilisé), le coût calcul de la totalité des matrices de transfert est de $\frac{N_{pays} * N_{out}}{N_{CPU}} * 5 * T_{POLAIR}$. Dans le tableau 9.8 est présenté le coût CPU total pour $N_{out} = 5$ et $N_{pays} = 40$. T_{POLAIR} désigne le temps CPU demandé par une simulation du modèle POLAIR3D.

10 CPU	20 CPU	30 CPU	40 CPU
$100 * T_{POLAIR}$	$50 * T_{POLAIR}$	$33 * T_{POLAIR}$	$25 * T_{POLAIR}$

Table 9.8: Coût calcul des matrices de transfert selon le nombre de processeurs ($N_{out} = 5$ et $N_{pays} = 40$)

Une autre alternative pour diminuer le nombre de simulations à effectuer est :

1. d'étudier l'impact de réduction des émissions des points composant un pays sur une région groupant plusieurs pays ou un ensemble de pays, afin de diminuer la valeur de N_{pays} .
2. si on peut diminuer le nombre des espèces de sortie, en s'intéressant pas exemple à des indicateurs d'impact capable d'agréger un ensemble de sortie brutes du modèle POLAIR3D.

Cela peut baisser énormément le nombre de simulations que nécessite le calcul des matrices de transfert, comme on peut l'observer sur la figure 9.15 (N_{pays} désigne ici, le nombre de régions au lieu du nombre de pays et N_{out} est le nombre d'indicateur d'impact agrégés à la place du nombre d'espèces sorties du modèle POLAIR3D.)

Remarque

Si nous utilisons la méthode qui consiste à représenter une année de simulation par 12 jours chacun représentatif d'un mois de l'année, le temps CPU demandé par une simulation POLAIR3D pour une période d'un année T_{POLAIR} est pratiquement divisé par 30 (365/12).

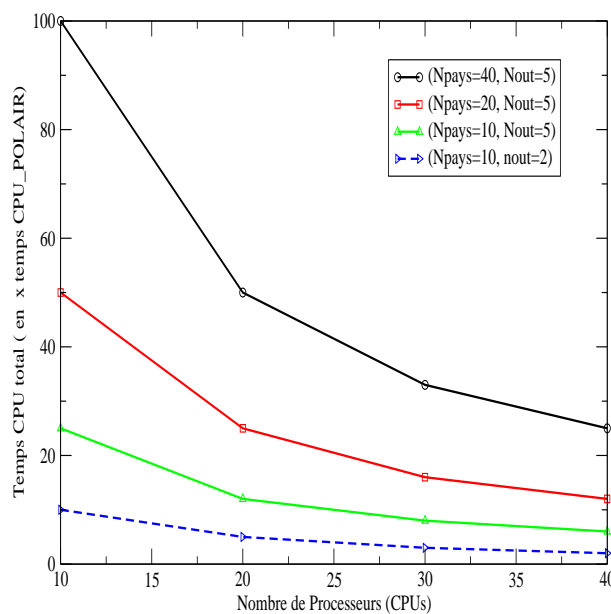


Figure 9.15: Coût CPU total pour le calcul des matrices de transfert avec différenciation automatique en réduisant le nombre de pays et le nombre d'espèces de sortie (le coût est exprimé en nombre de fois du temps CPU du modèle direct POLAIR3D)

Conclusion

Nous avons étudié la variabilité interannuelle et intermensuelle des concentrations chimiques sur l'Europe. Les approches par journée type représentative, choisie sur la base de critères chimiques, permettent d'obtenir de bons résultats. Ceci permet par ailleurs d'envisager une réduction forte du calcul des matrices de transfert.

Partie VI

Conclusions et perspectives générales

Chapitre 10

Conclusions générales et perspectives

Lors de cette thèse, nous nous sommes intéressés, dans un contexte global de modélisation intégrée des impacts, à étudier d'une part les méthodes de réduction du coût calcul de la cinétique chimique et, d'autre part, les méthodes performantes de propagation d'incertitude à travers un modèle de Chimie-Transport. Pour des besoins applicatifs, nous avons étudié la faisabilité de calcul des matrices de transfert d'une manière à la fois efficace et rapide.

10.1 Modèle de Chimie-Transport (POLAIR3D)

Dans un premier temps, nous nous sommes attachés à comprendre les modèles de Chimie-Transport et au développement du modèle POLAIR3D. Ce modèle est modulaire capable d'utiliser différents mécanismes, comme la chimie gazeuse (EMEP, CBM-IV, RADM, RACM, MOCA,...), les métaux lourds, le mercure et les radionucléides. La version aérosol du modèle POLAIR3D est en phase de finalisation.

Pour des études de sensibilité et d'assimilation de données, le modèle dispose de son code dérivé (adjoint et tangent linéaire) via le différentiateur automatique Odyssee. Le modèle POLAIR3D peut être utilisé pour différentes échelles spatiales régionale/urbaine et régionale/continentale.

Les premières études de validation ont donné des résultats assez satisfaisants. Une validation sur l'échelle européenne avec les données du réseau EMEP a donné des RMS de l'ordre de $25\mu\text{g}/\text{m}^3$ et un biais de l'ordre de $-6.2\mu\text{g}/\text{m}^3$ pour l'ozone. D'autres travaux de validation sur différentes campagnes de mesures (par exemple

la compagnie Escompte) sont en cours de réalisation. Il est question aussi que le modèle POLAIR3D intègre la plate-forme "Prév'air" "<http://prevair.ineris.fr>", dans un objectif de prévision opérationnelle.

10.2 Réduction de modèle

L'autre thème abordé lors de ce travail, est la réduction du modèle, plus particulièrement la cinétique chimique, module le plus consommateur en temps CPU dans un modèle de Chimie-Transport. L'objectif était de remplacer le solveur de chimie par une méthode de réduction plus rapide à précision conservée. Nous avons utilisé les méthodes suivantes :

10.2.1 Proper Orthogonal Decomposition (POD)

La méthode Proper Orthogonal Decomposition (POD) basée sur la réduction des degrés de liberté du système chimique, en le projetant sur une base réduite calculée d'une manière optimale à partir des simulations du modèle (trajectoire/snapshot). Dans un contexte 0D, cette méthode a donné des résultats très proches du modèle d'origine, mais sa validité reste très locale (typiquement pendant ou légèrement après la période de la snapshot qui a permis l'obtention de la base optimale POD). Une forte sensibilité de la base POD aux conditions initiales de la trajectoire, a été constatée.

En l'état actuel, l'intérêt de cette méthode ne réside que dans la compréhension qualitative de la complexité du système.

10.2.2 High Dimensional Model Representation (HDMR)

La méthode High Dimensional Model Representation (HDMR) décrit la sortie de modèle comme un développement en série de fonctions de variables d'entrée, ces fonctions étant déterminées par deux approches : CUT-HDMR basée sur le choix d'un point de référence dans l'intervalle des variables d'entrée et l'approche RS-HDMR (Random Sampling HDMR) basée sur un calcul à partir des points tirés aléatoirement de l'intervalle des entrées de modèle.

La méthode HDMR est composée de deux étapes, la première est consacrée à la prétabulation optimale des simulations POLAIR3D, et la seconde à construire le solveur FEOM (Fully Equivalent Operational Model) qui remplace

le solveur que nous souhaitons réduire (dans notre cas le solveur ROS2 de la chimie).

L'utilisation de la méthode CUT-HDMR appliquée aux différentes cinétiques chimiques potentiellement utilisées dans le modèle POLAIR (EMEP, CBM IV, MOCA et RACM), a donné d'excellents résultats en 0D, l'ordre 2 du développement HDMR étant suffisant pour représenter le modèle de chimie. Le gain en temps CPU est important, il peut atteindre un "speed up" de 18 pour une utilisation de POLAIR3D avec une chimie raide de pas de temps de 100s et 1800s pour le module de transport, dans le cas d'une utilisation du splitting de premier ordre. Dans une simulation utilisant le splitting du second ordre, le "speed up" est de l'ordre de 36 pour les mêmes pas de temps. Dans un test effectué dans un contexte d'analyse d'incertitude, l'utilisation de méthodes de Monte Carlo en utilisant le modèle réduit FEOM d'un côté et le modèle POLAIR de l'autre, a donné des PDFs (Fonctions de densité de probabilité) très similaires et un "speed up" de 10.

Une grande précaution doit être prise, au moment du calcul de la table HDMR, dans le choix des intervalles des variables d'entrée. Une utilisation d'intervalles qui ne représentent pas assez bien les variations simulées par le modèle d'origine, donnerait des divergences très importantes dans les résultats du modèle FEOM par rapport au modèle d'origine POLAIR0D.

L'utilisation de HDMR dans le cas tridimensionnel a été effectuée sur l'Europe en utilisant la chimie EMEP avec le modèle POLAIR3D. Le calcul "FEOM-global" effectué avec des tables HDMR sur la base des mêmes intervalles de variations des variables d'entrée sur tout le domaine, a donné des résultats très divergents du modèle POLAIR3D. Cela est dû essentiellement à la diversité et la variabilité spatiale (horizontale et verticale) des concentrations chimiques à l'échelle continentale sur l'Europe, ce qui engendre un choix du point de référence pour HDMR peu représentatif surtout dans les régions de faibles variation des régimes chimiques. Le point de référence qui est censé représenter l'effet moyen de toutes les variables d'entrée sur une sortie de modèle.

Pour surmonter cette difficulté, nous avons effectué un découpage du domaine géographique en plusieurs sous domaines tridimensionnels (12 sous-domaines) et utilisé de multiples FEOM, ce qui a donné des résultats nettement meilleurs que ceux obtenus avec un FEOM global. Un travail plus profond doit être fait sur la manière de découper le domaine de calcul, en se basant par exemple sur des simulations de modèle sur plusieurs années.

10.3 Propagation d'incertitude

Le troisième thème abordé dans cette thèse est la propagation d'incertitude, dans un contexte de réduction du nombre de simulations demandé par la méthode classique de Monte Carlo. Deux approches ont été utilisées (RS-HDMR) et (DEMM/SRSM).

10.3.1 Random Sampling HDMR (RSHDMR)

La méthode RS-HDMR est une autre approche de calcul des fonctions de développement de HDMR, basée sur un échantillonnage aléatoire sur tout le domaine des variables d'entrée. Elle peut être utilisée comme solveur mais aussi comme une technique d'étude de sensibilité, en décomposant les variances des sorties du modèle en fonction des contributions des variables d'entrée. Les fonctions RS-HDMR se basent sur une intégration par la méthode de Monte Carlo, et une réduction de variance par la méthode de corrélation.

L'utilisation de cette méthode, dans le cadre d'étude de sensibilité, nous permet de dégager les paramètres pertinents qui rentrent en jeu dans un modèle de Chimie-Transport, pour chacune des espèces chimiques. Son application à POLAIR3D avec la cinétique chimique EMEP, nous a permis de dégager les paramètres et les espèces chimiques qui ont un grand effet sur chacune des espèces chimiques sorties de modèle. Cette application a montré les grandes capacités de la méthode et sa validité pour des études de sensibilité.

10.3.2 Deterministic Equivalent Modeling Method/ Stochastic Response Surface Method (DEMM/SRSM)

La méthode DEMM/SRSM est une technique récente de propagation d'incertitude dans un modèle dont l'objectif classique est de pouvoir quantifier d'une manière efficace les Fonctions de Densité de Probabilité (PDFs) des sorties du modèle en fonction des PDFs des variables d'entrée. Cette approche est basée sur un développement en polynômes de variables d'entrée supposées incertaines et prises aléatoirement à partir de leur distribution de probabilité. Le développement DEMM/SRSM est composé d'un certain nombre d'inconnues qui sont calculées par des exécutions du modèle d'origine (POLAIR3D) à des points choisis astucieusement. La méthode utilisée pour le choix de ces points est la méthode ECM (Efficient Collocation Method) couplée avec la technique SRSM (Stochastic Response Surface Method).

L'utilisation de cette méthode au modèle POLAIR3D a été effectuée dans deux applications ; une application 0D a mis l'accent sur les incertitudes liées aux constantes cinétiques du mécanisme chimique EMEP ; dans la deuxième effectuée sur le modèle tridimensionnel sur l'Europe, nous avons supposé une incertitude sur les émissions (NO_x , SO_2 , NH_3) des pays européens suivant une loi normale. Comparées à la méthode classique de propagation d'incertitude Monte Carlo, ces deux applications ont donné des résultats excellents avec des gains en temps CPU très importants (un "speed up" de l'ordre de 300). Cette dernière application nous a permis d'identifier l'incertitude des concentrations des espèces de sortie du mécanisme d'acidification (NO_x , SO_2 , nitrate et sulfate d'ammonium) sur une application 3D sur l'Europe en supposant des incertitudes sur les émissions des pays européens.

Néanmoins une difficulté reste présente dans cette méthode, avec le choix des points de collocation comme entrées pour le modèle POLAIR3D afin de calculer les variables inconnues composantes du développement DEMM/SRSM.

10.4 Matrices de transfert

Parmi les composantes essentielles d'un système intégré de modélisation d'impact, les matrices de transfert donnent la relation qui pourrait lier une réduction des émissions à la réduction des espèces chimiques déposées au sol ou mesurées dans l'air sur différentes échelles de temps et d'espace. On peut calculer plusieurs types de matrices selon leurs applications, les plus utilisées étant les matrices "Pays/Point de grille", "Pays/Pays", "Point de grille/Point de grille" et "Point de grille/Pays". Pour leur calcul, nous avons proposé une méthode basée sur le modèle adjoint du modèle POLAIR3D, obtenu par le différentiateur automatique Odyssee.

Pour la validation de cette technique, nous l'avons appliquée au calcul des matrices de transfert par adjointisation pour l'étude d'impact pour la réduction des émissions de NO_x , SO_2 et NH_3 du mois de Janvier 1994 de l'ensemble des émissions de la France sur la concentration dans l'air (ou déposé) mensuelle prévue sur les autres pays européens. Cette simulation a été faite dans le cadre d'une réduction des émissions françaises selon 5 scénarios (20%, 40%, 60%, 80% et 100% de réduction) avec une comparaison à l'approche directe. Dans un contexte opérationnel cette dernière (approche directe) est infaisable avec l'état de l'art des ressources de calcul. Les écarts entre les deux méthodes existent bien évidemment à cause de la linéarisation utilisée dans le modèle dérivé pour le calcul des matrices de transfert, mais cette approche donne des résultats assez satisfaisants. De point de vue temps de calcul, avec une parallélisation selon les pays émetteurs, le coût de calcul d'une matrice de transfert peut encore être réduit.

10.5 Modèle statistique moyen : journée et année représentatives

Généralement le calcul d'impact se fait sur des longues périodes temporelles qui peuvent aller de 5 à 10 années, ce qui le rend impraticable avec des modèles de chimie transport complexe. Nous avons donc étudié la possibilité, à partir d'une base de données composée de plusieurs années de simulation à émissions constantes, d'identifier une année type capable de représenter l'évolution moyenne de plusieurs années de simulation d'un champ de concentrations. Dans la même logique, nous avons également étudié la possibilité et la faisabilité d'identifier un jour type capable de représenter l'évolution mensuelle d'un champs de concentrations chimiques.

Dans notre application, la base de donnée est alimentée par 9 années de calcul POLAIRD avec une chimie d'acidification EMEP sur l'Europe de 1993 à 2001. Pour le bilan annuel des concentrations, de faibles variations ont été observées pour NO_2 et SO_2 , de fortes variations pour le nitrate et sulfate d'ammonium. L'année 1995 s'avère l'année la plus proche de la moyenne des 9 années pour presque toutes les espèces du mécanisme chimique hormis l'espèce SO_4^{2-} qui se distingue pour avoir l'année 1996 comme année la plus proche de la moyenne. L'année type dépend donc de l'espèce chimique étudiée. Pour le bilan mensuel des concentrations dans l'air, sur les 9 années de calcul, les profils de variations mensuelles sont semblables et présentent des variations interannuelles inférieures à 20% pour le NO_2 et 30% pour le SO_2 .

La méthode du choix d'un jour type par mois, à partir des 9 années de calcul POLAIR3D, a montré une variation en fonction de l'espèce. La validation de la méthode a été effectuée en comparant les résultats de l'évolution mensuelle durant l'année 2001 moyenne sur l'Europe obtenus avec le modèle POLAIR3D en simulation complète et la méthode basée sur les résultats des 12 jours types représentatifs des 12 mois de l'année 2001. La comparaison a donné des résultats assez satisfaisants.

10.6 Perspectives et futurs travaux

Suite à ce travail, plusieurs points restent ouverts. Nous continuerons nos recherches dans les directions suivantes.

10.6.1 Développement de la version parallèle du modèle de Chimie Transport POLAIR3D

Le modèle actuel de chimie transport POLAIR3D, ne dispose pas de version parallèle. Cela peut être très utile pour une réduction du temps CPU demandé par une intégration séquentielle du modèle. L'utilisation de la parallélisation des modèles de chimie-transport a donné des résultats assez prometteurs [96, 97]. Plusieurs questions se poseront sur le type de machine à utiliser pour une telle parallélisation : machine à mémoire distribuée ou partagée, machine de type super ordinateur ou un ensemble d'ordinateurs connecté sous un réseau sous forme de clusters. Une autre interrogation se posera sur la façon de paralléliser le code du modèle POLAIR3D car il existe plusieurs approches. La technique de décomposition de domaine est souvent la plus utilisée pour la parallélisation d'un modèle de Chimie-Transport [97, 98, 99, 100]. Dans cette approche, chaque processeur s'occupe de la résolution du modèle sur un sous-domaine du découpage. Cette parallélisation nécessite des communications importantes entre les processeurs pour la partie transport à cause des schémas numériques utilisés dans le modèle. L'interrogation se posera, donc, sur la façon optimale de décomposer le domaine, plusieurs approches de découpage étant possibles [97].

A cela se rajoutent d'autres interrogations relatives aux modules du modèle CTM, qui doivent être parallélisés : doit-on paralléliser uniquement la chimie ? ou les deux modules de chimie et transport ?

Une des préoccupations permanentes dans un travail de parallélisation est d'optimiser au maximum le temps de communication entre les processeurs et le temps d'attente d'un processeur (load balancing) qui limitent généralement les capacités des techniques de parallélisation [100].

10.6.2 Réduction de modèle

- **Cut-HDMR** : Comme suite à notre travail sur l'utilisation de HDMR en 3D dans le contexte d'une modélisation intégrée, nous étudierons plus profondément le découpage sur l'Europe afin de garantir les bons intervalles de variations des paramètres d'entrée de HDMR.

Dans la nouvelle version du modèle POLAIR3D, une chimie complexe d'aérosols comportant un module de thermodynamique (ISOROPIA) très coûteux en temps CPU est intégrée. Nous comptons utiliser le code POLAIR-FEOM basé sur les développements HDMR, pour réduire le temps CPU du module d'aérosols.

- **Multiple Cut-HDMR** : La méthode de cut-HDMR est basée sur l'utilisation d'un point de référence. Dans nos travaux nous avons utilisé un point de référence $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ par intervalle de variation pour chacune des n variables d'entrée de modèle. Une autre idée consiste à utiliser pour chaque variable d'entrée plusieurs points de référence $\bar{x}^{(i)} = (\bar{x}_1^{(i)}, \bar{x}_2^{(i)}, \dots, \bar{x}_n^{(i)})$, i allant de 1 à m nombre de points de référence à utiliser. Dans ces cas le développement cut-HDMR [65] s'écrit selon :

$$f(x) \approx \sum_{k=1}^m \omega_k(x) [f_0^{(k)} + \sum_{i=1}^n f_i^{(k)}(x_i) + \dots + \sum_{i_1 i_2 \dots i_l} f_{i_1 i_2 \dots i_l}^{(k)}(x_{i_1}, x_{i_2}, \dots, x_{i_l})]. \quad (10.1)$$

où $\omega_k(x)$ sont des fonctions poids bien spécifique. Cette méthode peut s'avérer performante dans le cas d'une utilisation de HDMR à l'échelle tridimensionnelle où le choix du point de référence joue un rôle important.

- **POD** : L'utilisation de la méthode de réduction statistique POD, a été effectuée dans l'espace des phases sur des trajectoires de quelque heures, les bases POD calculées restent valables localement après la période de la snapshots. Nous envisageons d'appliquer cette technique aux données de simulation de POLAIR3D, pour chacune des espèces, sur une période de 900s (équivalent d'un pas de temps d'intégration) et sur tout le domaine de calcul, afin de déduire les différents modes capable de représenter l'espace physique de dimension $(N_x * N_y)$ [101].

Une nouvelle approche sera étudiée, en introduisant la méthode de **"balanced truncation"** et un produit scalaire défini par le grammien [102], afin de lever les incertitudes sur le choix du produit scalaire.

- **ISAT** : la méthode de réduction ISAT (In Situ Adaptive tabulation) est largement utilisée dans le domaine de la combustion [103], son utilisation ayant donné des gains en temps de calcul très importants. La méthode est basée sur une tabulation in situ pour la résolution de l'équation (représentative de la cinétique chimique) :

$$\frac{dC}{dt} = F(C(t)) \quad (10.2)$$

Pour un pas de temps Δt fixé, ISAT tabule la solution de l'équation (10.2) $R(C)$ après une intégration de Δt depuis la condition initiale C . Au point initial C^m , ISAT tabule $R^m = R(C^m)$ et la gradient $A^m = A(C^m)$ où :

$$A_{ij} = \frac{\partial R_i(C)}{\partial C_j} \quad (10.3)$$

L'approximation linéaire de $R(C)$ pour le point C proche de C^m , pour une erreur ne dépassant pas l'erreur de tolérance demandée ε_{tol} , se fait selon l'équation :

$$R(C) \approx R^m + A^m(C - C^m) \quad (10.4)$$

Si l'approximation linéaire (10.4) est insuffisante, ISAT calcule une nouvelle trajectoire pour la valeur C pendant une intégration de Δt . Une description détaillée de la méthode est dans [103].

10.6.3 Propagation d'incertitude

- **RS-HDMR** : Les premières applications ont été effectuées sur la chimie EMEP en supposant que les variables d'entrée du modèle varient uniformément. Une des applications futures est d'étendre cette application à des distributions non uniformes des variables d'entrée [83] et à une chimie plus complexe RACM, pour dégager les paramètres et les espèces chimiques pertinents pour les principales espèces du schéma.
- **DEMM/SRSM** : cette méthode, parmi ses avantages, est capable d'être utilisée pour des modèles en "boîte noire". L'un des points faibles de la méthode réside dans la difficulté à choisir les points de collocation pour trouver les paramètres inconnus qui composent le développement DEMM/SRSM. Une des alternatives est de coupler cette méthode à un **modèle dérivé** par différentiation automatique afin de réduire le nombre de points de collocation à choisir [26].

L'application de la méthode DEMM/SRSM au modèle POLAIR3D déjà réduit par la technique HDMR pourrait diminuer le temps de calcul demandé par DEMM/SRSM.

L'utilisation de cette méthode peut être étendue à d'autres applications comme par exemple l'étude d'impact à l'échelle européenne du mercure et des métaux lourds.

10.6.4 Matrices de transfert

Le calcul des matrices de transfert est très demandeur en ressource de calcul. Comme nous avons pu le constater sa parallélisation est automatique et elle n'a pas besoin de modifications dans le coeur des codes direct et dérivé. Cependant une définition plus justifiée d'indicateurs plus agrégés par espèces chimiques ou par région pourrait être efficace dans la réduction du temps CPU demandé par le calcul

de telles matrices.

10.6.5 Modèle statistique moyen

La méthode du calcul du jour "type" par mois et de l'année "type" parmi plusieurs années de simulations de POLAIR3D sur la base d'un jeu d'émissions constant, nécessite plus d'approfondissement pour mieux comprendre sa faisabilité dans un contexte opérationnel de modélisation d'impact.

10.6.6 Modélisation intégrée

Un de nos objectifs futurs est de se doter d'un outil performant d'étude d'impact, basé sur des fondements scientifiques, outil capable de fournir (en sortie) des scénarios d'impact en fonction des émissions des espèces chimiques (en entrée), à différentes échelles spatio-temporelles, de l'échelle régionale à l'échelle continentale pour une année à une dizaine d'années. Cet outil nécessite des moyens de calcul importants. Un axe futur est probablement de regrouper les différents maillons développé dans le cadre de cette thèse (POLAIR3D, Cut-HDMR, RS-HDMR, DEMM/SRSM, calcul des matrices de transfert, calcul d'un modèle moyen) afin de mettre en place une chaîne opérationnelle d'impact. Cela passera par un couplage avec d'autres modèles, tel que les modèles d'exposition de dose à effet, des modèles d'évaluation de dommages et des modèles macro-économique pour l'évaluation des coûts engendrés par une stratégie de réduction d'émission donnée.

Bibliography

- [1] R.L.G.M. Eggels. *Modeling of combustion processes and NO formation with reduced reaction mechanisms*. PhD thesis, Un. Eindhoven, NL, 1996.
- [2] B. Sportisse. *Contribution à la modélisation des écoulements réactifs : réduction des modèles de cinétique chimique et simulation de la pollution atmosphérique*. PhD thesis, Ecole Polytechnique, April 1999.
- [3] R. Djouad, B. Sportisse, and N. Audiffren. Numerical simulation of aqueous-phase atmospheric models : use of a non-autonomous rosenbrock method. *Atmos.Environ.*, 36 :873–879, 2002.
- [4] J. M. Bouchereau and B. Oudart. Lutte contre la pollution atmosphérique transfrontalière en europe : les nouveaux objectifs de la france. *Les données de l'environnement (air)*, Institut francais de l'ENvironnement, 53, 2000.
- [5] Website for RAINS Model. <http://www.iiasa.ac.at/rains/>.
- [6] Website for the International Institute for Applied Systems Analysis (IIASA). <http://www.iiasa.ac.at>.
- [7] K. Barrett and E. Berge. Transboundary air pollution in europe. msc-w status report 1/96. *EMEP Status Report*, 1996.
- [8] M.A. Tatang. *PhD thesis*. PhD thesis, MIT, 1995.
- [9] J. Boutahar, S. Lacour, V. Mallet, D. Quélo, Y. Roustan, and B. Sportisse. Development and validation of a fully modular platform for the numerical modeling of air pollution : Polair. *International Journal of Environment and Pollution*, 22(1-2), 2004.
- [10] Website for EURAD Model. <http://www.eurad.uni-koeln.de>.
- [11] Website for EMEP Model. <http://www.emep.int>.
- [12] Website for CMAQ Model. <http://www.epa.gov/asmdnerl/models3/>.
- [13] Website for MOZART Model. <http://acd.ucar.edu/models/mozart/>.
- [14] Website for CHIMERE Model. <http://euler.lmd.polytechnique.fr/chimere/>.
- [15] B. Sportisse and R. Djouad. Reduction of chemical kinetics in air pollution modelling. *J.Comp.Phys.*, 164 :354–376, 2000.

- [16] R. Sportisse, B. Djouad. Use of proper orthogonal decompositions for the reduction of atmospheric chemistry. *In revision for Linear Algebra and its Applications. Special issue "Model reduction"*, 2004.
- [17] J.A. Shorter and Precila C.Ip. An efficient chemical kinetics solver using high dimensional model representation. *Journal of Physical Chemistry (1999)*, 103 :7192–7198, 1999.
- [18] K. Kunisch and S. Volkwein. Control of burgers' equation by a reduced order approach using proper orthogonal decomposition. *Journal of Optimization Theory and Applications*, 102 :345 :371, 1999.
- [19] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *To appear in Numerische Mathematik*, 2001.
- [20] G. Berkooz, P. Holmes, and J.L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annu. Rev. Fluid Mech.*, pages 539–575, 1993.
- [21] S.W. Wang, H. Levy II, G. Li, and H. Rabitz. Fully equivalent operational models for atmospheric chemical kinetics within global chemistry-transport models. *J. Geophys. Res.*, 104(D23) :30417–30426, December 1999.
- [22] S.W. Wang, H. Levy, G. Li, and H. Rabitz. Fully equivalent operational models for atmospheric chemical kinetics within global chemistry-transport models. *Journal of Geophysical Research*, 104 :30 :417–30 :426, 1999.
- [23] H. Rabitz, O.F. Alis, J.A. Shorter, and K. Shim. Efficient input-output model representations. *Journal of Computer Physics Communication*, 117 :11–20, 1999.
- [24] H. Rabitz and O.F. Alis. General foundations of high-dimensional model representations. *Journal of Mathematical Chemistry*, 25 :197–223, 1999.
- [25] S. R. Hanna, Chang. J. C., and Fernau. M. E. Monte carlo estimations of uncertainties in prediction by a photochemical grid model (uam-iv) due to uncertainties in input variables. *J. Atmospheric Environment*, 21 :3619–3628, 1998.
- [26] S.S. Isukapalli. *PhD thesis*. PhD thesis, MIT, 1999.
- [27] C. Faure and Y. Papegay. Odyssee user's guide. version 1.7. Technical Report RT-0224, INRIA, 1998.
- [28] J.H. Seinfeld and S.N. Pandis. *Atmospheric chemistry and Physics*. Wiley-interscience, 1998.
- [29] P. Zannetti. *Air Pollution Modeling : theories, computational methods and available software*. Computational Mechanics Publications, 1990.
- [30] B. Sportisse. Box models versus eulerian models in air pollution modeling. *Atm. Environ.*, 35 :173–178, 2001.

- [31] J.F. Louis. A parametric model of vertical eddy fluxes in the atmosphere. *Boundary Layer Met.*, 17 :187 :202, 1979.
- [32] M. Tiedke. A comprehensive mass flux scheme for cumulus parameterization in large scale models. *Monthly Weather Review*, 117 :1779–1800, 1989.
- [33] B. Aumont. *Modélisation de la chimie de la base troposphère continentale : Développement et tests d'un modèle chimique condensé*. PhD thesis, Université Paris VII, 1994.
- [34] O. Wild, X. Zhu, and M.J. Prather. Fast-j : accurate simulation of in and below cloud photolysis in tropospheric chemical models. *J.Atm.Chem.*, 37 :245 :282, 2000.
- [35] B. Sportisse. Spack user's guide : a simplified preprocessor for atmospheric chemical kinetics. Technical report, ENPC-CEREA, 2003.
- [36] R. Yamartino, J. Scire, G.R. Carmichael, and Y.S. Chang. The calgrid mesoscale photochemical grid model. *Atm. Env.*, 26A :1493–1512, 1992.
- [37] M. Heimann. The global atmospheric tracer model tm2. Technical report, Max Planck Inst. fur Meteorologie, 1996.
- [38] M. Baer and K. Nester. Parametrization of trace gas dry deposition velocities for a regional mesoscale diffusion model. *Annales Geophysicae*, 10 :912–923, 1992.
- [39] M.L. Wesely. Parameterization of surface resistance to gaseous-dry deposition in regional scale numerical models. *Atmos. Environ.*, 23 :1293 :1304, 1989.
- [40] B. Sportisse, J. Boutahar, E. Debry, D. Quélo, and K. Sartelet. Some tracks in air pollution modeling. *RACSAM*, 96(3) :507–528, novembre 2002. In revision for RACSAM (Journal of the Spanish Science Academy, Real Academia de Ciencias de España).
- [41] D. Wendum. Three long range transport models compared to the etex experiments : a performance study. *Atmos. Environ.*, 32 :4297 :4305, 1998.
- [42] R. Djouad. *Contribution à la modélisation et à la simulation de la chimie atmosphérique multiphasique*. PhD thesis, Université de Rouen, 2001.
- [43] G.J. McRae, W.R. Goodin, and J.H. Seinfeld. Numerical solution of the atmospheric diffusion equation for chemically reacting flows. *J.Comp.Phys.*, 45 :1–42, 1982.
- [44] J.G. Verwer, W.H. Hundsdorfer, and J.G. Blom. Numerical time integration for air pollution models. In *Proceedings of the Conference APMS'98*. ENPC-INRIA, October 26-29 1998.
- [45] W.H. Hundsdorfer. Numerical solution of advection-diffusion-reaction equations. Technical Report NM-N9603, CWI, 1996.

- [46] L. Lanser and J.G. Verwer. Analysis of operator splitting for advection-diffusion-reaction problems from air pollution modelling. In *Proceedings 2nd Meeting on Numerical methods for differential equations*. Coimbra, Portugal, February 1998.
- [47] J.M. Sanz-Serna. *The State of the art in numerical analysis*, chapter Geometric integration, pages 121–143. Clarendon Press, Oxford, 1997.
- [48] G. Strang. On the construction and comparison of difference schemes. *SIAM J.Numer.Anal.*, 5 :506–517, 1968.
- [49] W. Hundsdorfer and J.G. Verwer. *Numerical solution of Time-Dependent Advection-Diffusion-Reaction Equations*. Springer Verlag, 2003.
- [50] B. Sportisse. An analysis of operator splitting techniques in the stiff case. *J. Comp. Phys.*, 161 :140–168, 2000.
- [51] W. R. Stockwell, F. Kirchner, M. Kuhn, and S. Seefeld. A new mechanism for regional atmospheric chemistry modeling. *Journal of Geophysical Research*, 102 :25,847–25,879, 1997.
- [52] J. Graf and N. Moussiopoulos. Intercomparison of two models for the dispersion of chemically reacting pollutants. *Beitr.Phys.Atmosph.*, 64(1) :13–25, Febr. 1991.
- [53] M.A. Botchev. Improving approximate matrix factorizations for implicit time integration in air pollution modelling. *J. Comp. Appl. Math*, 157 :309–327, 2003.
- [54] E.J. Spee. *Numerical methods in global transport-chemistry models*. PhD thesis, Univ. Amsterdam, 1998.
- [55] E. Hesstvedt, O. Hov, and I.S.A. Isaksen. Qssas in air pollution modelling : comparison of two numerical schemes for oxydant prediction. *Int.J.Chem.Kinet.*, 10 :971–994, 1978.
- [56] O. Hertel, R. Berkowicz, and J. Christensen. Test of two numerical schemes for use in atmospheric transport-chemistry models. *Atmos.Environ.*, 27A(16) :2591–2611, 1993.
- [57] J.G. Verwer. Gauss-seidel iteration for stiff odes from chemical kinetics. *SIAM J.Sci.Comput.*, 15 :1243–1250, 1994.
- [58] A.C. Hindmarsh. *Scientific computing*, chapter ODEPACK : a systematized collection of ODE solvers, pages 55–74. North Holland, 1983.
- [59] J.H. Verwer, E.J. Spee, J.G. Blom, and W.H. Hundsdorfer. A second order rosenbrock method applied to photochemical dispersion problem. *SIAM J. SCI. COMPUT.*, 20(4) :1456–1480, 1999.

- [60] A. Sandu, F.A. Potra, G.R. Carmichael, and V. Damian. Efficient implementation of fully implicit methods for atmospheric chemical kinetics. *J. Comp. Phys.*, 129 :101–110, 1996.
- [61] J.-P. Issartel and J. Baverel. Inverse transport for the verification of the comprehensive test ban treaty. *Atmos. Chem. Phys*, 3 :475–486, 2003.
- [62] K.N. Sartelet, J. Boutahar, D. Quélo, I. Coll, P. Plion, and B. Sportisse. Development and validation of a 3d chemistry-transport model, polair3d, by comparison with data from esquif campaign. In *Proceedings of the 6th Gloream workshop : Global and regional atmospheric modelling*. Aveiro, Portugal. P140–146., 2002.
- [63] P. Middleton and W. R. Stockwell. A new mechanism for regional atmospheric chemistry modeling. *Journal of Atmospheric Environment*, 24A :1107–1133, 1990.
- [64] J. Shorter and P. Ip. Radiation transport simulation by means of a fully equivalent operational model. *J. Geophys. Res.*, 27(21) :3485–3488, November 2000.
- [65] G. Li, C. Rosenthal, and H. Rabitz. High-dimensional model representations. *Journal of Physical Chemistry*, 105-33 :7765–7777, 2001.
- [66] S.-W. Wang, P. G. Georgopoulos, G. Li, and J. H. Rabitz. Computationally efficient atmospheric chemical kinetic modeling by means of high dimensional model representation (hdmr). *Springer Verlag, Lecture Notes in Computer Science*, pages 326–333, 2001.
- [67] G Li, S.-W. Wang, C. Rosenthal, and J. H. Rabitz. High dimensional model representations generated from low dimensional data samples. i. mp-cut-hdmr. *J. Math. Chem*, 30 :1–30, 2001.
- [68] M.W. Gery, G.Z. Whitten, J.K. Killus, and M.C. Dodge. A photochemical kinetics mechanism for urban and regional scale computer modeling. *J. Geophys. Research*, 94(D10) :12925–12956, 1989.
- [69] B. Aumont. Module de chimie moca : notice descriptive. Technical report, LISA, 1998.
- [70] C. Taylor. *The Development of a Fast Chemical Scheme for General Circulation Models*. PhD thesis, Un. Reading, 2003.
- [71] M. Rathinam and R. Petzold. A new look at proper orthogonal decomposition. *SIAM J. Math. Anal.*, 41(5) :1893–1925, 2003.
- [72] W. C. Rowley, T. Colonius, and M. R. Murray. Model reduction for compressible flows using pod and galerkin projection. *Elsevier Science*, To appear 2003.
- [73] M.A. Tatang, W. Pan, R.G. Prinn, and McRae G.J. An efficient implementation for parametric uncertainty analysis of numerical geophysical models. *Journal of geophysical research*, 102 :21.925–21.932, 1997.

- [74] Phenix B.D., Dinario J.L., M.A. Tatang, tester J.W., Howard J.B., and McRae G.J. Incorporation of parametric uncertainty into complex kinetic mechanisms : Application to hydrogen oxidation in supercritical water. *Combustion and flame*, 112 :132,146, 1998.
- [75] Ch. Bischof, A. Carle, G. Corliss, A. Griewank, and P. Hovland. Adifor generating derivative codes from fortran programs. Technical report, Mathematics and Computer Science Division, Argonne National Laboratory, 1992.
- [76] Y. J. Yang, J. G. Wilkinson, and A. G. Russel. Fast direct sensitivity analysis of multidimensional photochemical models. *Environ. Sci. Technol.*, 31 :2859–2868, 1997.
- [77] G. Li, H. Rabitz, S.W. Wang, and P.G Georgopoulos. Correlation method for variance reduction of monte carlo integration in rs-hdmr. Technical report, Princeton University and EOHS Institute, 2002.
- [78] W.H. Press, Flannery B.P., S.A. Teukolsky, and Vetterling W.T. *Numerical Recipes (Fortran Version)*. Cambridge University Press.
- [79] M. Webster, M.A. Tatang, and G.J. McRae. Application of the probabilistic collocation method for uncertainty analysis of a simple ocean model. Technical report, Cambridge, MA, MIT Joint Program, 1996.
- [80] J. Boutahar and B. Sportisse. Rapport de contrat edf 2. méthodes de tabulation pour la réduction de modèles atmosphériques dans le contexte des études d'impacts. application à la cinétique chimique. Technical report, ENPC-CEREVE, 2002.
- [81] J. Boutahar and B. Sportisse. Reduction methods for atmospheric chemistry. In *Proceedings of the 6th Gloream workshop : Global and regional atmospheric modelling*. Aveiro, Portugal. P133-139, 2002.
- [82] O.F. Alis and H. Rabitz. Efficient implementation of high dimensional model representations. *Journal of Mathematical Chemistry*, 29-2 :127–142, 2001.
- [83] S.-W. Wang, P. G. Georgopoulos, and J. H. Rabitz. Random sampling-high dimensional model representations (rs-hdmr) with nonuniformly distributed variables : Application to integrated multimedia/multipathway exposure and dose model for trichloroethylene. *J. Phys. Chem*, 107 :4707–4716, 2003.
- [84] G. Li, H. Rabitz, S. Wang, P. Jaffe, S.W. Wang, and G. Georgopoulos. Efficient sensitivity/uncertainty analysis using the hdmr method :application to the princeton groundwater model. Technical report, CERM/CRESP-CEEA Technical Report CCL-TR-01.01, 2002.
- [85] S.W. Wang G.Li and H. Rabitz. Practical approaches to construct rs-hdmr component functions. *Journal of Physical Chemistry*, 106 :8721–8733, 2002.
- [86] J. Neveu. Introduction aux probabilités. Cours Ecole Polytechnique, 1994.

- [87] J. Bartnicki. Nonlinear effects in the source-receptor matrices computed with emep eulerian acid deposition model. *EMEP Status Report*, (ISSN 0332-9879), 2000.
- [88] L. Tarrason, J. E. Jonson, and Fagerli H. Transboundary acidification, eutrophication and ground level ozone in europe. part iii (source-receptor relationships). *EMEP Status Report*, (ISSN 0806-4520), 2003.
- [89] A. Sandu, G.R. Carmichael, and Potra. Sensitivity analysis for atmospheric chemistry models via automatic differentiation. *J. Atmospheric Environment*, 31-3 :475–489, 1997.
- [90] P. Eberhard and C. Bischof. Automatic differentiation of numerical integration algorithms. *Mathematic of computation*, 68 :717–731, 1999.
- [91] C. Bischof, A. Carle, G. Corliss, A. Griewank, and P. Hovlana. Adifor - generating derivative codes from fortran programs,. *Scientific Programming*, (1) :9–29, 1992.
- [92] B. Sportisse and D. Quélo. Data assimilation and inverse modeling of atmospheric chemistry. *Proc. of Indian National Science Academy. Part A Physical Sciences*, 69, November 2003.
- [93] J.R. Brook, P.J. Samson, and S. Sillman. Aggregation selected three day periods estimated annual seasonal wet deposition totals for sulfule nitrate acidity part i-ii. *Journal of Applied meteorology*, 34(2) :297–325,326–339, 1995.
- [94] M.E. Fernau and P.J. Samson. use cluster analysis define periods similar meteorology precipitation chemistry eastern north america part i-ii. *Journal of Applied meteorology*, 29(8) :735–750,751–761, 1990.
- [95] C. Philippe. *Analyse de la pollution atmosphérique aux échelles locale et régionale. Modélisation spatiale et temporelle à l'aide d'une méthode de scénarios épisodiques*. PhD thesis, Institut National des sciences appliquées de Rouen, 2004.
- [96] Ostromsky. T. and Zlatev. Z. Flexible two-level parallel implementation of a large air pollution model. *Springer Verlag, lecture notes in Computer Science*, pages 545–554, 2003.
- [97] P. Miehe, A. Sandu, and G. R. Carmichael. A communication library for the parallelization of air quality models on structured grids. *Journal of Atmospheric Environment*, 36 :3917–3930, 2002.
- [98] P. C.S. Lee, R. A. Zaveri, R. C. Easter, and L. K. Peters. Technical note on the parallelization of a global climate-chemistry modeling system. *Atm. Env.*, 33 :675–681, 1999.
- [99] M. Martin, O. Oberson, B. Chopard, F. Mueller, and A. Clappier. Atmospheric pollution transport : the parallelization of a transport and chemistry code. *Atm. Env.*, 33 :1853–1860, 1999.

-
- [100] H. Elbern. Parallelization and load balancing of a comprehensive atmospheric chemistry transport model. *Journal of Atmospheric Environment*, 31(21) :3561–3574, 1997.
- [101] M. F. Arlene and J. D. Jacob. Application of empirical orthogonal functions to evaluate ozone simulations with regional and global models. *Journal of geophysical research*, 108 :10.1–10.10, 2003.
- [102] S. Lall, J. E. Marsden, and S. Glavaski. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *Journal of Robust and Nonlinear control*, 12 :519–535, 2002.
- [103] S.B. Pope. Computationally efficient implementation of combustion chemistry using in situ adaptive tabulation. *Combustion Theory Modelling*, 1 :41–63, 1997.
- [104] W.R. Stockwell, P. Middleton, J. Chang, and X. Tang. The second regional acid deposition model chemical mechanism for regional air quality modeling. *J. of Geophysical Research*, 95(D10) :16343–16367, September 1990.

Partie VII

Annexes

Appendix A

Mécanisme CBM-IV

Le mécanisme chimique CBM-V est composé de 31 espèces et 68 réactions chimiques ; dans les tableaux suivants sont présentées les différentes réactions chimiques représentées par ce schéma [56].

Table A.1: Mécanisme CBM-IV

	réaction	constante cinétique (*)
1	$NO_2 \rightarrow NO + O^3P$	0.58(-2)
2	$NO + O_3 \rightarrow NO_2$	0.44(-03)
3	$HO_2 + NO \rightarrow NO_2 + OH$	20.5(-2)
4	$HCHO \rightarrow 2. HO_2 + CO$	0.1434(-04)
5	$HCHO \rightarrow CO$	0.1367(-04)
6	$HCHO + OH \rightarrow HO_2 + CO$	0.25(0)
7	$ALD \rightarrow XO_2 + 2. HO_2 + CO + HCHO$	0.2167(-05)
8	$ALD + OH \rightarrow C_2O_3$	0.4(0)
9	$C_2O_3 + NO \rightarrow NO_2 + XO_2 + HCHO + HO_2$	2.75(-1)
10	$C_2O_3 + NO_2 \rightarrow PAN$	0.15(0)
11	$PAN \rightarrow C_2O_3 + NO_2$	3.67(-4)
12	$XO_2 + NO \rightarrow NO_2$	0.2(0)
13	$NO_2 + OH \rightarrow HNO_3$	27.17(-2)
14	$O^3P \rightarrow O_3$	0.8(5)
15	$O_3 \rightarrow .205 OH + .897 O^3P$	58.33(-7)
16	$O_3 \rightarrow O^3P$	24.49(-5)
17	$SO_2 + OH \rightarrow H_2SO_4 + HO_2$	20.67(-3)
18	$NO_3 \rightarrow .15 NO + .85 NO_2 + .85 O^3P$	17.84(-2)
19	$NO_2 + O_3 \rightarrow NO_3$	0.79(-6)
20	$NO_3 + NO_2 \rightarrow NO + NO_2$.98(-5)
21	$NO_3 + NO_2 \rightarrow N_2O_5$	29.67(-3)
22	$N_2O_5 \rightarrow NO_3 + NO_2$.52(-1)
23	$N_2O_5 + 2. HNO_3 \rightarrow HNO_3$	4.75(-4)
24	$NO + NO \rightarrow 2. NO_2$	25.33(-10)
25	$OH + CO \rightarrow HO_2$.6667(-2)
26	$O^3P + NO_2 \rightarrow NO$	0.23(0)
27	$O^3P + NO_2 \rightarrow NO_3$	38.67(-3)
28	$O^3P + NO \rightarrow NO_2$.52(-1)
29	$O_3 + OH \rightarrow HO_2$	16.67(-5)
30	$O_3 + HO_2 \rightarrow OH$	0.5(-4)

(*) $A(B)$ read $A \cdot 10^B$. Units are in *ppb*.

Table A.1 (suite)

N_r	réaction	constante cinétique
31	$NO_3 + NO \rightarrow 2.NO_2$	46.83(-2)
32	$NO + NO_2 \rightarrow 2.HNO_2$	0.4(-11)
33	$HNO_2 + HNO_2 \rightarrow NO + NO_2$	0.2(-9)
34	$HNO_2 \rightarrow NO + OH$	10.54(-4)
35	$NO + OH \rightarrow HNO_2$	16.28(-2)
36	$OH + HNO_2 \rightarrow NO_2$	16.28(-2)
37	$HO_2 + HO_2 \rightarrow H_2O_2$	12.36(-2)
38	$H_2O_2 \rightarrow 2.OH$	41.41(-7)
39	$H_2O_2 + OH \rightarrow HO_2$	50.15(-3)
40	$HCHO + O^3P \rightarrow OH + HO_2 + CO$	3.95(-3)
41	$HCHO + NO_3 \rightarrow HO_2 + CO + HNO_3$	1.55(-5)
42	$ALD + O^3P \rightarrow C_2O_3 + OH$	3.95(-3)
43	$ALD + NO_3 \rightarrow C_2O_3 + HNO_3$	6.17(-5)
44	$ALD + HO_2 \rightarrow XO_2O + 2.HCHO$	8.33(-5)
45	$C_2O_3 + C_2O_3 \rightarrow 2.XO_2 + 2.HCHO + 2.HO_2$	61.67(-3)
46	$C_2O_3 + HO_2 \rightarrow$	0.16(0)
47	$MGLY \rightarrow C_2O_3 + HO_2 + CO$	11.66(-5)
48	$MGLY + OH \rightarrow XO_2 + C_2O_3$	43.33(-2)
49	$OH \rightarrow XO_2 + HCHO + HO_2$	0.35(0)
50	$PAR + OH \rightarrow 1.49 XO_2 + .067 XO_2N + .93 HO_2$ $+ .45 ALD + .75 RXPAR$	19.17(-3)
51	$O^3P + OLE \rightarrow .95 ALD + .35 HO_2 + .2 XO_2 + .15 CO$ $+ .05 HCHO + .05 C_2O_3 + .35 RXPAR$	98.67(-3)
52	$OH + OLE \rightarrow HCHO + ALD + XO_2 + HO_2$.7(0)
53	$O_3 + OLE \rightarrow .5 ALD + .66 HCHO + .212 CO$ $+ .28 HO_2 + .144 XO_2 + .08 OH + RXPAR$.3(-6)
54	$NO_3 + OLE \rightarrow .91 XO_2 + .91 HO_2 + .09 XO_2N + ONIT$.19(-3)
55	$O^3P + ETH \rightarrow HCHO + XO_2 + CO + 2.HO_2$.18(-1)
56	$OH + ETH \rightarrow XO_2 + 2.HCHO + HO_2$.2(0)
57	$O_3 + ETH \rightarrow HCHO + .37 CO + .13 HO_2$.45(-7)
58	$TOL + OH \rightarrow HO_2 + .64 XO_2 + 1.13 HCHO$ $+ .56 MGLY + .36 PHEN + .36 PAR + 1.13 CO$	16.25(-2)
59	$PHEN + NO_3 \rightarrow PHO + HNO_3$.23(0)
60	$PHO + NO_2 \rightarrow ONIT$.33(0)

Table A.1 (suite)

N_r	réaction	constante cinétique
61	$XYL + OH \rightarrow HO_2 + .72 XO_2 + .67 CO + 1.33 MGLY + .28 PHEN + .67 HCHO + .56 PAR$.6(0)
62	$RXPAR + PAR \rightarrow$	2.(0)
63	$XO_2N + NO \rightarrow ONIT$	16.67(-3)
64	$XO_2 + HO_2 \rightarrow$	83.33(-3)
65	$XO_2 + C_2O_3 \rightarrow XO_2 + HO_2 + HCHO$.4(-1)
66	$HO_2 + NO_2 \rightarrow HNO_4$	39.26(-3)
67	$HNO_4 \rightarrow HO_2 + NO_2$	85.36(-2)
68	$HNO_4 + OH \rightarrow NO_2$	1233.64(-4)

Appendix B

La chimie d'acidification : EMEP

La chimie d'acidification "EMEP" utilisée lors de cette thèse est tirée du rapport [7]. Elle met en jeu les espèces chimiques suivantes (figure B.1) : NO , NO_2 , HNO_3 , $PAN = CH_3COO_2NO_2$, NO_3 , SO_2 , SO_4^{2-} , NH_3 , Sulfate d'ammonium ($(NH_4)_{1.5}SO_4$) et Nitrate d'ammonium (NH_4NO_3). Pour chaque pas de temps du modèle, POLAIR3D calcule la concentration de chacune de ces espèces après les différents processus physiques de transport et de diffusion. Outre ces espèces chimiques, d'autres espèces sont prises en compte comme l'ozone O_3 et les radicaux : hydroxyle (OH) et peroxyacétyle (CH_3COO_2). Ces espèces sont prescrites au modèle comme des conditions au limites.

Le schéma comporte deux parties, une réservée aux transformations linéaires et l'autre aux réactions thermodynamiques non linéaires.

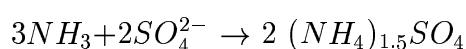
Dans la première section (tableau B.1) nous présentons les différentes réactions de la partie linéaire du schéma EMEP et dans la deuxième section une présentation succincte de la partie thermodynamique. Pour plus de détails on peut se référer à [7].

B.1 Réactions de EMEP ; Partie linéaire

Les réactions chimiques qui composent le schéma EMEP sont présentées dans le tableau B.1.

B.2 Réaction de EMEP : Partie non linéaire

– Sulfate d'Ammonium ($(NH_4)_{1.5}SO_4$)



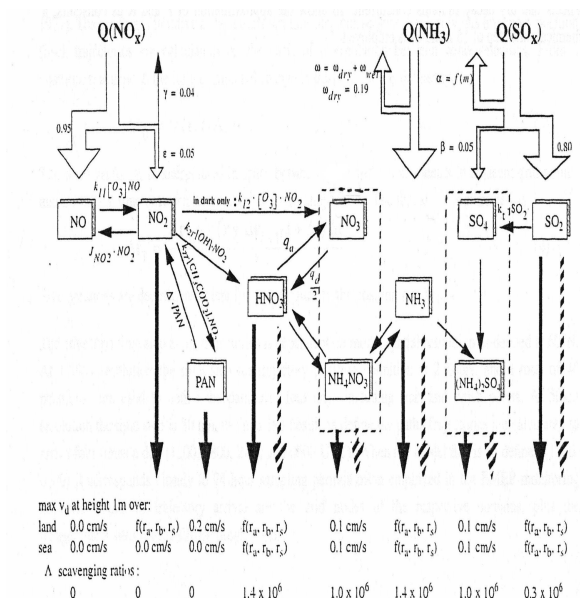


Figure B.1: Chimie EMEP [7]

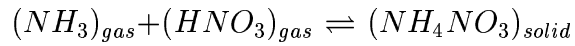
Réaction	
$NO + O_3$	$\xrightarrow{k_1} NO_2 + O_2$
$NO_2 + hv$	$\xrightarrow{k_2} NO + O$
$NO_2 + OH$	$\xrightarrow{k_3} HNO_3$
$2NO_2 + O_3 + H_2O$	$\xrightarrow{k_4} 2NO_3^- + 2H^+$
$NO_2 + CH_3COO_2 (ROH)$	$\xrightarrow{k_5} CH_3COO_2NO_2 (PAN)$
PAN	$\xrightarrow{k_6} NO_2 + ROH$
HNO_3	$\xrightarrow{k_7} NO_3^-$
NO_3^-	$\xrightarrow{k_8} HNO_3$
SO_2	$\xrightarrow{k_9} SO_4^{2-}$

Table B.1: Schéma chimique EMEP : partie linéaire. Les constantes cinétiques sont détaillées dans [7]

Si $3(SO_4^{2-}) - 2(NH_3) > 0$, NH_3 disparaît complètement.

Si $3(SO_4^{2-}) - 2(NH_3) < 0$, SO_4^{2-} disparaît complètement.

– **Nitrate d'Ammonium** (NH_4NO_3)



avec une constante d'équilibre thermodynamique de la réaction $K_{eq} = NH_3 * HNO_3$.

Appendix C

Mécanisme chimique RACM

Le mécanisme RACM (Regional Atmospheric Chemistry Mechanism) proposé dans [51], est une version récente du mécanisme RADM2 [104]. Le schéma RACM fait intervenir 237 réactions et 77 espèces chimiques : 17 espèces stables inorganiques, 4 espèces inorganiques intermédiaire, 32 espèces organiques stables et 24 espèces organiques intermédiaires. Nous présentons dans cette annexe les différentes espèces et réactions chimiques qui interviennent dans ce mécanisme, pour plus de détails on peut se référer à [51].

Numéro de réaction	Réaction	Fréquence de photolyse S^{-1}
R1	$NO_2 \rightarrow O_3P + NO$	0.760031E-02
R2	$O_3 \rightarrow O_1D + O_2$	0.172587E-04
R3	$O_3 \rightarrow O_3P + O_2$	0.423344E-03
R4	$HONO \rightarrow HO + NO$	0.165516E-02
R5	$HNO_3 \rightarrow HO + NO_2$	0.470402E-06
R6	$HNO_4 \rightarrow 0.65HO_2 + 0.65NO_2 + 0.35HO + 0.35NO_3$	0.320567E-05
R7	$NO_3 \rightarrow NO + O_2$	0.235172E-01
R8	$NO_3 \rightarrow NO_2 + O_3P$	0.187348
R9	$H_2O_2 \rightarrow HO + HO$	0.603528E-05
R10	$HCHO \rightarrow H_2 + CO$	0.354845E-04
R11	$HCHO \rightarrow HO_2 + HO_2 + CO$	0.224649E-04
R12	$ALD \rightarrow MO_2 + HO_2 + CO$	0.376612E-05
R13	$OP_1 \rightarrow HCHO + HO_2 + HO$	0.414587E-05
R14	$OP_2 \rightarrow ALD + HO_2 + HO$	0.414587E-05
R15	$PAA \rightarrow MO_2 + HO$	0.159652E-05
R16	$KET \rightarrow ACO_3 + ETHP$	0.685160E-06
R17	$GLY \rightarrow 0.13HCHO + 1.87CO$	0.582325E-04
R18	$GLY \rightarrow 0.45HCHO + 1.55CO + 0.80HO_2$	0.203173E-04
R19	$MGLY \rightarrow ACO_3 + HO_2 + CO$	0.951654E-04
R20	$DCB \rightarrow TCO_3 + HO_2$	0.430840E-03
R21	$ONIT \rightarrow 0.20ALD + 0.80KET + HO_2 + NO_2$	0.220592E-05
R22	$MACR \rightarrow HCHO + ACO_3 + CO + HO_2$	0.135225E-05
R23	$HKET \rightarrow ACO_3 + HCHO + HO_2$	0.685160E-06

Table C.1: Réactions photolytiques

Numéro de réaction	Réaction	A, C $cm^3 s^{-1}$
R24	$O3P + O2 \rightarrow O3O3P \rightarrow O3$	Table C.9
R25	$O3P + O3 \rightarrow 2.OO2$	8.00E-12 2060.0
R26	$O1D + N2 \rightarrow O3P + N2$	1.80E-11 -110.0
R27	$O1D + O2 \rightarrow O3P + O2$	3.20E-11 -70.0
R28	$O1D + H2O \rightarrow 2.HO$	2.20E - 100.0
R29	$O3 + HO \rightarrow HO2 + O2$	1.60E-12 940.0
R30	$O3 + HO2 \rightarrow HO + 2.OO2$	1.10E-14 500.0
R31	$HO + HO2 \rightarrow H2O + O2$	4.80E-11 -250.0
R32	$H2O2 + HO \rightarrow HO2 + H2O$	2.90E-12 160.0
R33	$HO2 + HO2 \rightarrow H2O2 + O2$	Table C.9
R34	$HO2 + HO2 \rightarrow H2O2 + O2 + H2O$	Table C.9
R35	$O3P + NO \rightarrow NO2$	Table C
R36	$O3P + NO2 \rightarrow NO + O2$	6.50E-12 -120.0
R37	$O3P + NO2 \rightarrow NO3$	Table C
R38	$HO + NO \rightarrow HONO$	Table C
R39	$HO + NO2 \rightarrow HNO3$	Table C
R40	$HO + NO3 \rightarrow NO2 + HO2$	2.20E-11
R41	$HO2 + NO \rightarrow NO2 + HO$	3.70E-12 -250.0
R42	$HO2 + NO2 \rightarrow HNO4$	Table C
R43	$HNO4 \rightarrow HO2 + NO2$	Table C
R44	$HO2 + NO3 \rightarrow 0.3HNO3 + 0.7NO2 + 0.7HO$	3.50E-12
R45	$HO + HONO \rightarrow NO2$	1.80E-11 390.0
R46	$HO + HNO3 \rightarrow NO3$	Table C.9
R47	$HO + HNO4 \rightarrow NO2$	1.30E-12 -380.0
R48	$O3 + NO \rightarrow NO2 + O2$	2.00E-12 1400.0
R49	$O3 + NO2 \rightarrow NO3 + O2$	1.20E-13 2450.0
R50	$NO + NO \rightarrow NO2 + NO2$	3.30E-39 -530.0
R51	$NO3 + NO \rightarrow NO2 + NO2$	1.50E-11 -170.0
R52	$NO3 + NO2 \rightarrow NO + NO2 + O2$	4.50E-14 1260.0
R53	$NO3 + NO2 \rightarrow N2O5$	Table C
R54	$N2O5 \rightarrow NO2 + NO3$	Table C
R55	$NO3 + NO3 \rightarrow NO2 + NO2 + O2$	8.50E-13 2450.0
R56	$HO \rightarrow HO2 + H2O$	5.50E-12 2000.0
R57	$HO + SO2 \rightarrow SULF + HO2$	Table C
R58	$CO + HO \rightarrow HO2 + CO2$	Table C.9
R59	$ISO + O3P \rightarrow 0.86OLT + 0.05HCHO + 0.02HO + 0.01CO + 0.13DCB + 0.28HO2 + 0.15XO2$	6.00E-11
R60	$MACR + O3P \rightarrow ALD$	1.59E-11 -13.0
R61	$CH4 + HO \rightarrow MO2 + H2O$	table 2c
R62	$ETH + HO \rightarrow ETHP$	Table C.8
R63	$HC3 + HO \rightarrow .583HC3P + .381HO2 + .335ALD + .036ORA1 + .036CO + .036GLY + .036HO + .010HCHO$	5.26E-12 260.0
R64	$HC5 + HO \rightarrow .750HC5P + .250KET + .250HO2 + H2O$	8.02E-12 155.0
R65	$HC8 + HO \rightarrow .951HC8P + .025ALD + .024HKET + .049HO2 + H2O$	1.64E-11 125.0
R66	$ETE + HO \rightarrow ETEP$	1.96E-12 -438.0
R67	$OLT + HO \rightarrow OLTP$	5.72E-12 -500.0
R68	$OLI + HO \rightarrow OLIP$	1.33E-11 -500.0

Table C.2: Réactions gazeuses

R71	$API + HO \rightarrow APIP$	1.21E-11 -444.0
R72	$LIM + HO \rightarrow LIMP$	1.71E-10
R73	$TOL + HO \rightarrow 0.90ADDT + 0.10XO2 + 0.10HO2$	1.81E-12 -355.0
R74	$XYL + HO \rightarrow 0.90ADDX + 0.10XO2 + 0.10HO2$	7.30E-12 -355.0
R75	$CSL + HO \rightarrow 0.85ADDC + 0.10PHO + 0.05HO2 + 0.05XO2$	6.00E-11 0.0
R76	$HCHO + HO \rightarrow HO2 + CO$	1.00E-11
R77	$ALD + HO \rightarrow ACO3$	5.55E-12 -331.0
R78	$KET + HO \rightarrow KETP$	Table C.8
R79	$HKET + HO \rightarrow HO2 + MGLY$	3.00E-12
R80	$GLY + HO \rightarrow HO2 + 2.0CO$	1.14E-11
R81	$MGLY + HO \rightarrow ACO3 + CO$	1.72E-11
R82	$MACR + HO \rightarrow 0.51TCO3 + 0.41HKET + 0.08MGLY +$ $0.41 CO + 0.08 HCHO + 0.49 HO2 + 0.49 XO2$	1.86E-11 -175.0
R83	$HO + DCB \rightarrow 0.50TCO3 + 0.50HO2 + 0.50XO2 + 0.35UDD + 0.15GLY + 0.15MGLY$	2.80E-11 -175.
R84	$HO + UDD \rightarrow 0.88ALD + 0.12KET + HO2$	2.70E-10
R85	$OP1 + HO \rightarrow .65MO2 + .35HCHO + 0.35HO$	2.93E-12 -190.0
R86	$OP2 + HO \rightarrow 0.44HC3P + 0.08ALD + 0.49HO + 0.07XO2 + 0.41KET$	3.40E-12 -190.0
R87	$PAA + HO \rightarrow 0.65ACO3 + 0.35HO2 + 0.35HCHO + 0.35XO2$	2.93E-12 -190.0
R88	$PAN + HO \rightarrow HCHO + NO3 + XO2$	4.00E-14
R89	$TPAN + HO \rightarrow 0.60HKET + 0.60NO3 + 0.40PAN + 0.40HCHO + 0.40HO2 + XO2$	3.25E-13 -500.0
R90	$ONIT + HO \rightarrow HC3P + NO2$	5.31E-12 260.0
R91	$HCHO + NO3 \rightarrow HO2 + HNO3 + CO$	3.40E-13 1900.0
R92	$ALD + NO3 \rightarrow ACO3 + HNO3$	1.40E-12 1900.0
R93	$GLY + NO3 \rightarrow HNO3 + HO2 + 2.0CO$	2.90E-12 1900.0
R94	$MGLY + NO3 \rightarrow HNO3 + ACO3 + CO$	1.40E-12 1900.0
R95	$MACR + NO3 \rightarrow 0.20TCO3 + 0.20HNO3 + 0.80OLNN + 0.80CO$	8.27E-15 150.0
R96	$DCB + NO3 \rightarrow 0.50TCO3 + 0.50HO2 + 0.50XO2 +$ $0.25 GLY + 0.25 ALD + 0.50 NO2 + 0.03 KET + 0.25 MGLY + 0.50 HNO3$	2.87E-13 1000.0
R97	$CSL + NO3 \rightarrow HNO3 + PHO$	2.20E-11
R98	$ETE + NO3 \rightarrow 0.80OLNN + 0.20OLND$	Table C.8
R99	$OLT + NO3 \rightarrow 0.43OLNN + 0.57OLND$	1.79E-13 450.0
R100	$OLI + NO3 \rightarrow 0.11OLNN + 0.89OLND$	8.64E-13 -450.
R101	$DIEN + NO3 \rightarrow 0.90OLNN + 0.10OLND + 0.90MACR$	1.00E-13
R102	$ISO + NO3 \rightarrow 0.90OLNN + 0.10OLND + 0.90MACR$	4.00E-12 446.
R103	$API + NO3 \rightarrow 0.10OLNN + 0.90OLND$	1.19E-12 -490.0
R104	$LIM + NO3 \rightarrow 0.13OLNN + 0.87OLND$	1.22E-11
R105	$TPAN + NO3 \rightarrow 0.60ONIT + 0.60NO3 + 0.40PAN + 0.40HCHO + 0.40NO2 + XO2$	2.20E-14 500.0
R106	$ETE + O3 \rightarrow HCHO + 0.43CO + 0.37ORA1 + 0.26HO2 + 0.13H2 + 0.12HO$	9.14E-15 2580.0
R 107	$OLT + O3 \rightarrow 0.64HCHO + 0.44ALD + 0.37CO +$ $0.14 ORA1 + 0.10 ORA2 + 0.25 HO2 + 0.40 HO + 0.03 KET + 0.03 KETP +$ $0.06 CH4 + 0.05 H2 + 0.03 ETH + .006 H2O2 + 0.19 MO2 + 0.10 ETHP$	4.33E-15 1800.0
R108	$OLI + O3 \rightarrow 0.02HCHO + 0.99ALD + 0.16KET +$ $0.30 CO + .011 H2O2 + 0.14 ORA2 + 0.07 CH4 + 0.22 HO2 + 0.63 HO +$ $0.23 MO2 + 0.12 KETP + 0.06 ETH + 0.18 ETHP$	4.40E-15 845.0
R109	$DIEN + O3 \rightarrow 0.90HCHO + 0.39MACR + 0.36CO +$ $0.15 ORA1 + 0.09 O3P + 0.30 HO2 + 0.35 OLT + 0.28 HO + 0.05 H2 +$ $0.15 ACO3 + 0.03 MO2 + 0.02 KETP + 0.13 XO2 + .001 H2O2$	1.34E-14 2283.0

Table C.3: Réactions gazeuses (suite)

R110	$ISO + O_3 \rightarrow 0.90HCHO + 0.39MACR + 0.36CO +$ $0.15 ORA1 + 0.09 O_3P + 0.30 HO_2 + 0.35 OLT + 0.28 HO + 0.05 H_2 +$ $0.15 ACO_3 + 0.03 MO_2 + 0.02 KETP + 0.13 XO_2 + .001 H_2O_2$	7.86E-15 1913.0
R111	$API + O_3 \rightarrow 0.65ALD + 0.53KET + 0.14CO +$ $0.20 ETHP + 0.42 KETP + 0.85 HO + 0.10 HO_2 + 0.02 H_2O_2$	1.01E-15 732.0
R 112	$LIM + O_3 \rightarrow 0.04HCHO + 0.46OLT + 0.14CO +$ $0.16 ETHP + 0.42 KETP + 0.85 HO + 0.10 HO_2 + 0.02 H_2O_2 + 0.79 MACR +$ $0.01 ORA1 + 0.07 ORA2$	2.00E-16
R113	$MACR + O_3 \rightarrow 0.40HCHO + 0.60MGLY + 0.13ORA2 +$ $0.54 CO + 0.08 H_2 + 0.22 ORA1 + 0.29 HO_2 + 0.07 HO + 0.13 OP_2 +$ $0.13 ACO_3$	1.36E-15 2112.0
R 114	$DCB + O_3 \rightarrow 0.21HO + 0.29HO_2 + 0.66CO +$ $0.50 GLY + 0.62 MGLY + 0.28 ACO_3 + 0.16 ALD + 0.11 PAA + 0.11 ORA1 +$ $0.21 ORA2$	2.00E-18
R115	$TPAN + O_3 \rightarrow 0.70HCHO + 0.30PAN + 0.70NO_2 +$ $0.13 CO + 0.04 H_2 + 0.11 ORA1 + 0.08 HO_2 + .036 HO + 0.70 ACO_3$	2.46E-15 1700.0
R116	$PHO + NO_2 \rightarrow 0.10CSL + ONIT$	2.00E-11
R 117	$PHO + HO_2 \rightarrow CSL$	1.00E-11
R118	$ADDT + NO_2 \rightarrow CSL + HONO$	3.60E-11
R 119	$ADDT \rightarrow 0.98TOLP + 0.02CSL + 0.02HO_2$	1.66E-17 -1044.0
R120	$ADDT + O_3 \rightarrow CSL + HO$	5.00E-11
R121	$ADDX + NO_2 \rightarrow CSL + HONO$	3.60E-11
R122	$ADDX \rightarrow 0.98XYLP + 0.02CSL + 0.02HO_2$	1.66E-17 -1044.0
R123	$ADDX + O_3 \rightarrow CSL + HO$	1.00E-11
R124	$ADDC + NO_2 \rightarrow CSL + HONO$	3.60E-11
R125	$ADDC \rightarrow 0.98CSLP + 0.02CSL + 0.02HO_2$	1.66E-17 -1044.0
R126	$ADDC + O_3 \rightarrow CSL + HO$	5.00E-11
R127	$ACO_3 + NO_2 \rightarrow PAN$	Table C
R128	$PAN \rightarrow ACO_3 + NO_2$	Table C
R129	$TCO_3 + NO_2 \rightarrow TPAN$	Table C
R130	$TPAN \rightarrow TCO_3 + NO_2$	Table C
R131	$MO_2 + NO \rightarrow HCHO + HO_2 + NO_2$	4.20E-12 -180.0
R132	$ETHP + NO \rightarrow ALD + HO_2 + NO_2$	8.70E-12
R133	$HC3P + NO \rightarrow .047HCHO + .233ALD + .623KET +$ $.063 GLY + .742 HO_2 + .150 MO_2 +$ $.048 ETHP + .048 XO_2 + .059 ONIT + .941 NO_2$	4.00E-12
R134	$HC5P + NO \rightarrow .021HCHO + .211ALD + .722KET +$ $.599 HO_2 + .031 MO_2 + .245 ETHP +$ $.334 XO_2 + .124 ONIT + .876 NO_2$	4.00E-12
R135	$HC8P + NO \rightarrow .150ALD + .642KET + .133ETHP +$ $.261 ONIT + .739 NO_2 + .606 HO_2 + .416 XO_2$	4.00E-12
R136	$ETEP + NO \rightarrow 1.6HCHO + HO_2 + NO_2 + 0.2ALD$	9.00E-12
R137	$OLTTP + NO \rightarrow 0.94ALD + HCHO + HO_2 + NO_2 + 0.06KET$	4.00E-12
R138	$OLIP + NO \rightarrow HO_2 + 1.71ALD + 0.29KET + NO_2$	4.00E-12
R139	$ISOP + NO \rightarrow .446MACR + .354OLT + .847HO_2 + .847NO_2 + .153ONIT + .606HCHO$	4.00E-12
R140	$APIP + NO \rightarrow 0.80HO_2 + 0.80ALD + 0.80NO_2 + 0.80KET + 0.20ONIT$	4.00E-12
R141	$LIMP + NO \rightarrow 0.65HO_2 + 0.40MACR + 0.25OLI + 0.25HCHO + 0.65NO_2 + 0.35ONIT$	4.00E-12

Table C.4: Réactions gaseuses (suite)

R142	$TOLP + NO \rightarrow 0.95NO_2 + 0.95HO_2 + 1.20GLY + 0.50DCB + 0.05ONIT + 0.65MGLY$	4.00E-12
R143	$XYLP + NO \rightarrow 0.95NO_2 + 0.95HO_2 + 0.35GLY + 0.95DCB + 0.05ONIT + 0.60MGLY$	4.00E-12
R144	$CSLP + NO \rightarrow GLY + MGLY + NO_2 + HO_2$	4.00E-12
R145	$ACO_3 + NO \rightarrow MO_2 + NO_2$	2.00E-11
R146	$TCO_3 + NO \rightarrow NO_2 + ACO_3 + HCHO$	2.00E-11
R147	$KETP + NO \rightarrow 0.54MGLY + 0.46ALD + 0.23ACO_3 + 0.77HO_2 + 0.16XO_2 + NO_2$	4.00E-12
R148	$OLNN + NO \rightarrow ONIT + NO_2 + HO_2$	4.00E-12
R149	$OLND + NO \rightarrow .287HCHO + 1.24ALD + .464KET + 2.00NO_2$	4.00E-12
R150	$MO_2 + HO_2 \rightarrow OP1$	3.80E-13 -800.0
R151	$ETHP + HO_2 \rightarrow OP2$	7.50E-13 -700.0
R152	$HC3P + HO_2 \rightarrow OP2$	1.66E-13 -1300.0
R153	$HC5P + HO_2 \rightarrow OP2$	1.66E-13 -1300.0
R154	$HC8P + HO_2 \rightarrow OP2$	1.66E-13 -1300.0
R155	$ETEP + HO_2 \rightarrow OP2$	1.90E-13 -1300.0
R156	$OLTP + HO_2 \rightarrow OP2$	1.66E-13 -1300.0
R157	$OLIP + HO_2 \rightarrow OP2$	1.66E-13 -1300.0
R158	$ISOP + HO_2 \rightarrow OP2$	1.28E-13 -1300.0
R159	$APIP + HO_2 \rightarrow OP2$	1.50E-11
R160	$LIMP + HO_2 \rightarrow OP2$	1.50E-11
R161	$TOLP + HO_2 \rightarrow OP2$	3.75E-13 -980.0
R162	$XYLP + HO_2 \rightarrow OP2$	3.75E-13 -980.0
R163	$CSLP + HO_2 \rightarrow OP2$	3.75E-13 -980.0
R164	$ACO_3 + HO_2 \rightarrow PAA$	1.15E-12 -550.0
R165	$ACO_3 + HO_2 \rightarrow ORA_2 + O_3$	3.86E-16 -2640.0
R166	$TCO_3 + HO_2 \rightarrow OP2$	1.15E-12 -550.0
R167	$TCO_3 + HO_2 \rightarrow ORA_2 + O_3$	3.86E-16 -2640.0
R168	$KETP + HO_2 \rightarrow OP2$	1.15E-13 -1300.0
R169	$OLNN + HO_2 \rightarrow ONIT$	1.66E-13 -1300.0
R170	$OLND + HO_2 \rightarrow ONIT$	1.66E-13 -1300.0
R171	$MO_2 + MO_2 \rightarrow 1.33HCHO + 0.66HO_2$	9.10E-14 -416.0
R172	$ETHP + MO_2 \rightarrow 0.75HCHO + HO_2 + 0.75ALD$	1.18E-13 -158.0
R173	$HC3P + MO_2 \rightarrow .810HCHO + .992HO_2 + .580ALD + .018KET + .007MO_2 + .005MGLY + .085XO_2 + .119GLY$	9.46E-14 -431.0
R174	$HC5P + MO_2 \rightarrow .829HCHO + .946HO_2 + .523ALD + .240KET + .014ETHP + .245XO_2 + .049MO_2$	1.00E-13 -467.0
R175	$HC8P + MO_2 \rightarrow .753HCHO + .993HO_2 + .411ALD + .419KET + .322XO_2 + .013ETHP$	4.34E-14 -633.
R176	$ETEP + MO_2 \rightarrow 1.55HCHO + HO_2 + 0.35ALD$	1.71E-13 -708.0
R177	$OLTP + MO_2 \rightarrow 1.25HCHO + HO_2 + .669ALD + .081KET$	1.46E-13 -708.0
R178	$OLIP + MO_2 \rightarrow .755HCHO + HO_2 + .932ALD + .313KET$	9.18E-14 -708.0
R179	$ISOP + MO_2 \rightarrow .550MACR + .370OLT + HO_2 + .080OLI + 1.09HCHO$	1.36E-13 -708.0
R180	$APIP + MO_2 \rightarrow 2.00HO_2 + ALD + HCHO + KET$	3.56E-14 -708.0
R181	$LIMP + MO_2 \rightarrow 2.00HO_2 + 0.60MACR + 0.40OLI + 1.40HCHO$	3.56E-14 -708.0
R182	$TOLP + MO_2 \rightarrow HCHO + HO_2 + 0.65GLY + DCB + 0.35MGLY$	3.56E-14 -708.0
R183	$XYLP + MO_2 \rightarrow HCHO + HO_2 + 0.37GLY + DCB + 0.63MGLY$	3.56E-14 -708.0
R184	$CSLP + MO_2 \rightarrow GLY + MGLY + 2.00HO_2 + HCHO$	3.56E-14 -708.0
R185	$ACO_3 + MO_2 \rightarrow HCHO + HO_2 + MO_2$	3.21E-11 440.0

Table C.5: Réactions gazeuses (suite)

R186	$ACO_3 + MO_2 \rightarrow HCHO + ORA_2$	2.68E-16 -2510.0
R187	$TCO_3 + MO_2 \rightarrow ACO_3 + 2.00HCHO + HO_2$	3.21E-11 440.0
R188	$TCO_3 + MO_2 \rightarrow HCHO + ORA_2$	2.68E-16 -2510.0
R189	$KETP + MO_2 \rightarrow 0.40MGLY + 0.30ALD + 0.30HKET +$ $0.88HO_2 + 0.08XO_2 + 0.12ACO_3 + 0.75HCHO$	6.91E-13 -508.0
R190	$OLNN + MO_2 \rightarrow 0.75HCHO + HO_2 + ONIT$	1.60E-13 -708.0
R191	$OLND + MO_2 \rightarrow .960HCHO + .500HO_2 + .640ALD +$.149KET + .500NO ₂ + .500ONIT	9.68E-14 -708.0
R192	$ETHP + ACO_3 \rightarrow ALD + .500HO_2 + 0.5MO_2 + 0.5ORA_2$	1.03E-12 -211.0
R193	$HC3P + ACO_3 \rightarrow .724ALD + .127KET + .488HO_2 +$.508MO ₂ + .499ORA ₂ + .091HCHO + .006ETHP + .071XO ₂ + .004MGLY + .100GLY	6.90E-13 -460.0
R194	$HC5P + ACO_3 \rightarrow .677ALD + .330KET + .438HO_2 +$.554MO ₂ + .495ORA ₂ + .076HCHO + .018ETHP + .237XO ₂	5.59E-13 -522.0
R195	$HC8P + ACO_3 \rightarrow .497ALD + .581KET + .489HO_2 +$.507MO ₂ + .495ORA ₂ + .015ETHP + .318XO ₂	2.47E-13 -683.
R196	$ETEP + ACO_3 \rightarrow 0.8HCHO + 0.6ALD + 0.5HO_2 + 0.5MO_2 + 0.5ORA_2$	9.48E-13 -765.0
R197	$OLTP + ACO_3 \rightarrow .859ALD + .501HCHO + .501HO_2 +$.501MO ₂ + .499ORA ₂ + .141KET	8.11E-13 -765.0
R198	$OLIP + ACO_3 \rightarrow .941ALD + .510HO_2 + .510MO_2 +$.569KET + .490ORA ₂	5.09E-13 -765.0
R199	$ISOP + ACO_3 \rightarrow .771MACR + .229OLT + .506HO_2 +$.494ORA ₂ + .340HCHO + .506MO ₂	7.60E-13 -765.0
R200	$APIP + ACO_3 \rightarrow HO_2 + ALD + MO_2 + KET$	7.40E-13 -765.0
R201	$LIMP + ACO_3 \rightarrow HO_2 + 0.60MACR + 0.40OLI + 0.40HCHO + MO_2$	7.40E-13 -765.0
R202	$TOLP + ACO_3 \rightarrow MO_2 + HO_2 + DCB + 0.65GLY + 0.35MGLY$	7.40E-13 -765.0
R203	$XYLP + ACO_3 \rightarrow MO_2 + HO_2 + DCB + 0.37GLY + 0.63MGLY$	7.40E-13 -765.0
R204	$CSLP + ACO_3 \rightarrow GLY + MGLY + HO_2 + MO_2$	7.40E-13 -765.0
R205	$ACO_3 + ACO_3 \rightarrow 2.0MO_2$	2.80E-12 -530.0
R206	$TCO_3 + ACO_3 \rightarrow MO_2 + ACO_3 + HCHO$	2.80E-12 -530.0
R207	$KETP + ACO_3 \rightarrow 0.54MGLY + 0.35ALD + 0.11KET +$ 0.12ACO ₃ + 0.38HO ₂ + 0.08XO ₂ + 0.50MO ₂ + 0.50ORA ₂	7.51E-13 -565.0
R208	$OLNN + ACO_3 \rightarrow ONIT + 0.50MO_2 + 0.50ORA_2 + 0.50HO_2$	8.85E-13 -765.0
R209	$OLND + ACO_3 \rightarrow .207HCHO + .650ALD + .167KET +$.516MO ₂ + .484ORA ₂ + .484ONIT + .516NO ₂	5.37E-13 -765.0
R210	$OLNN + OLNN \rightarrow 2.00ONIT + HO_2$	7.00E-14 -1000.0
R211	$OLNN + OLND \rightarrow .202HCHO + .640ALD + .149KET +$.500HO ₂ + .500NO ₂ + 1.50ONIT	4.25E-14 -1000.0
R212	$OLND + OLND \rightarrow .504HCHO + 1.21ALD + .285KET + ONIT + NO_2$	2.96E-14 -1000.0
R213	$MO_2 + NO_3 \rightarrow HCHO + HO_2 + NO_2$	1.20E-12
R214	$ETHP + NO_3 \rightarrow ALD + HO_2 + NO_2$	1.20E-12
R215	$HC3P + NO_3 \rightarrow .048HCHO + .243ALD + .670KET +$.063GLY + .792HO ₂ + .155MO ₂ + .053ETHP + .051XO ₂ + NO ₂	1.20E-12
R216	$HC5P + NO_3 \rightarrow .021HCHO + .239ALD + .828KET +$.699HO ₂ + .040MO ₂ + .262ETHP + .391XO ₂ + NO ₂	1.20E-12
R217	$HC8P + NO_3 \rightarrow .187ALD + .880KET + .155ETHP +$ NO ₂ + .845HO ₂ + .587XO ₂	1.20E-12
R218	$ETEP + NO_3 \rightarrow 1.6HCHO + HO_2 + NO_2 + 0.2ALD$	1.20E-12
R219	$OLTP + NO_3 \rightarrow 0.94ALD + HCHO + HO_2 + NO_2 + 0.06KET$	1.20E-12

Table C.6: Réactions gaseuses (suite)

R220	$OLIP + NO_3 \rightarrow HO_2 + 1.71ALD + 0.29KET + NO_2$	1.20E-12
R221	$ISOP + NO_3 \rightarrow .600MACR + .400OLT + HO_2 + NO_2 + .686HCHO$	1.20E-12
R222	$APIP + NO_3 \rightarrow HO_2 + ALD + NO_2 + KET$	1.20E-12
R223	$LIMP + NO_3 \rightarrow HO_2 + 0.60MACR + 0.40OLI + 0.40HCHO + NO_2$	1.20E-12
R224	$TOLP + NO_3 \rightarrow NO_2 + HO_2 + 0.50DCB + 1.30GLY + 0.70MGLY$	1.20E-12
R225	$XYLP + NO_3 \rightarrow NO_2 + HO_2 + DCB + 0.74GLY + 1.26MGLY$	1.20E-12
R226	$CSLP + NO_3 \rightarrow GLY + MGLY + NO_2 + HO_2$	1.20E-12
R227	$ACO_3 + NO_3 \rightarrow MO_2 + NO_2$	4.00E-12
R228	$TCO_3 + NO_3 \rightarrow NO_2 + ACO_3 + HCHO$	4.00E-12
R229	$KETP + NO_3 \rightarrow 0.54MGLY + 0.46ALD + 0.23ACO_3 + 0.77HO_2 + 0.16XO_2 + NO_2$	1.20E-12
R230	$OLNN + NO_3 \rightarrow ONIT + NO_2 + HO_2$	1.20E-12
R231	$OLND + NO_3 \rightarrow .280HCHO + 1.24ALD + .469KET + 2.00NO_2$	1.20E-12
R232	$XO_2 + HO_2 \rightarrow OP_2$	1.66E-13 -1300.0
R233	$XO_2 + MO_2 \rightarrow HCHO + HO_2$	5.99E-15 -1510.0
R234	$XO_2 + ACO_3 \rightarrow MO_2$	3.40E-14 -1560.0
R235	$XO_2 + XO_2 \rightarrow$	7.13E-17 -2950.0
R236	$XO_2 + NO \rightarrow NO_2$	4.00E-12
R237	$XO_2 + NO_3 \rightarrow NO_2$	1.20E-12

Table C.7: Réactions gaseuses (suite)

Numéro de réaction	C $K^{-2}cm^3s^{-1}$	D K
R61	7.44E-18	1361
R62	1.51E-17	492.
R78	5.68E-18	-92.0
R98	4.88E-18	2282.0

Table C.8: Constante de réaction sous forme $k = T^2 C \exp(-D/T)$

Numéro de réaction	K cm^3s^{-1}
R24	$[M] * 6.00E-34 * (T/300)^{-2.3}$
R33	$2.3E-13 * \exp(600./T) + 1.7E-33 * [M] * \exp(1000./T)$
R34	$3.2E-34 * \exp(2800./T) + 2.38E-54 * [M] * \exp(3200./T)$
R46	$k_0 = 7.2E-15 * \exp(785/TEMP)$ $k_2 = 4.1E-16 * \exp(1440/TEMP)$ $k_3 = 1.9E-33 * \exp(725/TEMP) * [M]$ $K = k_0 + k_3 / (1 + k_3/k_2)$
R58	$1.5E-13 * (1. + 2.439E-20 * [M])$

Table C.9: Réactions chimiques avec des expressions spécifiques de constantes cinétiques. [M] désigne la concentration de l'air en molécules cm^{-3} , T désigne la température

Numéro de Réaction	k_0^{300} $cm^6 s^{-1}$	n	k_∞^{300} $cm^3 s^{-1}$	m
R35	9.00E-32	1.5	3.00E-11	0.0
R37	9.00E-32	2.0	2.20E-11	0.0
R38	7.00E-31	2.6	1.50E-11	0.5
R39	2.60E-30	3.2	2.40E-11	1.3
R42	1.80E-31	3.2	4.70E-12	1.4
R53	2.20E-30	3.9	1.50E-12	0.7
R57	3.00E-31	3.3	1.50E-12	0.0
R127	9.70E-29	5.6	9.30E-12	1.5
R129	9.70E-29	5.6	9.30E-12	1.5

Table C.10: Réactions de Troe.

Numéro de Réaction	k_0^{300} $cm^6 s^{-1}$	n	k_∞^{300} $cm^3 s^{-1}$	m	A	B
R43	1.80E-31	3.2	4.70E-12	1.4	4.76E+26	10900.
R54	2.20E-30	3.9	1.50E-12	0.7	3.70E+26	11000.0
R128	9.70E-29	5.6	9.30E-12	1.5	1.16E+28	13954.
R130	9.70E-29	5.6	9.30E-12	1.5	1.16E+28	13954.

Table C.11: Réactions d'équilibre Troe

Appendix D

Stations de Mesures du réseau Européen EMEP

Numéro	Code	Non de Station	Lat	lon	Alt (m)
1	AT 2R	Illmitz	47 46 0 N	16 46 0 E	117
2	AT 3R	Achenkirch	47 33 0 N	11 43 0 E	960
3	AT 4R	St. Koloman	47 39 0 N	13 12 0 E	851
4	AT 5R	Vorhegg	46 40 40 N	12 58 20 E	1020
5	AT30R	Pillersdorf bei Retz	48 43 16 N	15 56 32 E	315
6	AT31R	St. Leonhard	47 59 5 N	14 51 10 E	790
7	AT32R	Sulzberg	47 31 45 N	9 55 36 E	1020
8	AT33R	Stolzalpe bei Murau	47 7 45 N	14 12 14 E	1302
9	AT34G	Sonnblick	47 3 16 N	12 57 30 E	3106
10	AT37R	Zillertaler Alpen	47 8 13 N	11 52 12 E	1970
11	AT38R	Gerlitzten	46 41 37 N	13 54 54 E	1895
12	AT40R	Masenberg	47 20 53 N	15 52 56 E	1170
13	AT41R	Haunsberg	47 58 23 N	13 0 58 E	730
14	AT42R	Heidenreichstein	48 52 43 N	15 2 48 E	570
15	AT43R	Forsthof	48 6 22 N	15 55 10 E	581
16	AT44R	Graz Platte	47 6 47 N	15 28 14 E	651
17	AT45R	Dunkelsteinerwald	48 22 16 N	15 32 48 E	320
18	AT46R	Ganserndorf	48 20 5 N	16 43 50 E	161
19	AT47R	Stixneusiedl	48 3 3 N	16 40 36 E	240
20	BE 1R	Offagne	49 52 40 N	5 12 13 E	430
21	BE32R	Eupen	50 37 46 N	6 0 10 E	295
22	BE35R	Vezein	50 30 12 N	4 59 22 E	160
23	CH 2R	Payerne	46 49 0 N	6 57 0 E	510

Table D.1: Stations de mesures

24	CH 3R	Tanikon	47 29 0 N	8 54 0 E	540
25	CH 4R	Chaumont	47 3 4 N	6 58 50 E	1130
26	CH 5R	Rigi	47 4 10 N	8 27 56 E	1030
27	CH31R	Sion	46 13 0 N	7 20 0 E	480
28	CZ 1R	Svratouch	49 44 0 N	16 2 0 E	737
29	CZ 3R	Kosetice	49 35 0 N	15 5 0 E	534
30	DE 1R	Westerland	54 55 32 N	8 18 35 E	12
31	DE 2R	Langenbrugge	52 48 8 N	10 45 34 E	74
32	DE 3R	Schauinsland	47 54 53 N	7 54 31 E	1205
33	DE 4R	Deuselbach	49 45 53 N	7 3 7 E	480
34	DE 5R	Brotjacklriegel	48 49 10 N	13 13 9 E	1016
35	DE 7R	Neuglobsow	53 10 0 N	13 2 0 E	62
36	DE 8R	Schmucke	50 39 0 N	10 46 0 E	937
37	DE 9R	Zingst	54 26 0 N	12 44 0 E	1
38	DE11R	Hohenwestedt	54 6 0 N	9 40 0 E	75
39	DE12R	Bassum	52 51 0 N	8 42 0 E	52
40	DE14R	Meinerzhagen	51 7 0 N	7 38 0 E	510
41	DE17R	Ansbach	49 15 0 N	10 35 0 E	481
42	DE26R	Ueckerm?nde	53 45 0 N	14 4 0 E	1
43	DE31R	Wiesenburg	52 7 0 N	12 28 0 E	107
44	DE35R	Luckendorf	50 50 0 N	14 46 0 E	490
45	DE38R	Murnauer Moos	47 39 05 N	11 12 12 E	622
46	DE39R	Aukrug	54 4 29 N	9 47 34 E	15
47	DE42R	Ohringen	49 14 36 N	9 26 50 E	283
48	DK 5R	Keldsnor	54 44 0 N	10 44 0 E	10
49	DK31R	Ulborg	56 17 0 N	8 26 0 E	10
50	DK32R	Frederiksborg	55 58 0 N	12 20 0 E	10
51	DK41R	Lille Valby	55 41 13 N	12 7 34 E	10
52	EE 9R	Lahemaa	59 30 0 N	25 54 0 E	32
53	EE11R	Vilsandy	58 23 0 N	21 49 0 E	6
54	ES 1R	San Pablo de los Montes	39 32 52 N	4 20 55 W	917
55	ES 2R	Cartuja	37 12 0 N	3 36 0 W	720
56	ES 3R	Roquetas	40 49 14 N	0 29 29 E	44
57	ES 4R	Logro?o	42 27 28 N	2 30 11 W	445
58	ES 5R	Noya	42 43 41 N	8 55 25 W	683
59	ES 7R	Viznar	37 14 0 N	3 32 0 W	1265
60	ES 8R	Niembro	43 26 32 N	4 51 1 W	134
61	ES 9R	Campisabalos	41 16 52 N	3 8 34 W	1360
62	ES10R	Cabo de Creus	42 19 10 N	3 19 1 E	23
63	ES11R	Barcarrola	38 28 33 N	6 55 22 W	393
64	ES12R	Zarra	39 5 10 N	1 6 7 W	885
65	ES13R	Penausende	41 17 0 N	5 52 0 W	985
66	ES14R	Els Torms	41 24 0 N	0 43 0 E	470
67	ES15R	Risco Llamo	39 31 0 N	4 21 0 W	1241
68	ES16R	O Savinao	43 13 52 N	7 41 59 W	506
69	FI 4R	Ahtari	62 32 0 N	24 13 18 E	162
70	FI 9R	Uto	59 46 45 N	21 22 38 E	7

Table D.2: Stations de mesures (Suite 1)

Numéro	Code	Non de Station	Lat	lon	Alt (m)
71	FI17R	Violahti II	60 31 36 N	27 41 10 E	4
72	FI22R	Oulanka	66 19 13 N	29 24 6 E	310
73	FI37R	Ahtari II	62 35 0 N	24 11 0 E	180
74	FR 8R	Donon	48 30 0 N	7 8 0 E	775
75	FR 9R	Revin	49 54 0 N	4 38 0 E	390
76	FR10R	Morvan	47 16 0 N	4 5 0 E	620
77	FR11R	Bonnevaux	46 49 0 N	6 11 0 E	836
78	FR12R	Iraty	43 2 0 N	1 5 0 W	13
79	FR13R	Peyrusse Vieille	43 22 29 N	0 6 16 E	236
80	FR14R	Montandon	47 11 0 N	6 30 0 E	746
81	GB 2R	Eskdalemuir	55 18 47 N	3 12 15 W	243
82	GB 6R	Lough Navar	54 26 35 N	7 52 12 W	126
83	GB13R	Yarner Wood	50 35 47 N	3 42 47 W	119
84	GB14R	High Muffles	54 20 4 N	0 48 27 W	267
85	GB15R	Strath Vaich Dam	57 44 4 N	4 46 28 W	270
86	GB31R	Aston Hill	52 30 14 N	3 1 59 W	370
87	GB32R	Bottesford	52 55 46 N	0 48 55 W	32
88	GB33R	Bush	55 51 31 N	3 12 18 W	180
89	GB34R	Glazebury	53 27 31 N	2 27 59 W	21
90	GB35R	Great Dun Fell	54 41 2 N	2 26 6 W	847
91	GB36R	Harwell	51 34 23 N	1 19 0 W	137
92	GB37R	Ladybower Res.	53 23 56 N	1 45 12 W	420
93	GB38R	Lullington Heath	50 47 34 N	0 10 46 E	120
94	GB39R	Sibton	52 17 38 N	1 27 47 E	46
95	GB41R	Wharleycroft	54 36 9 N	2 28 6 W	206
96	GB43R	Narberth	51 14 0 N	4 42 0 W	160
97	GB44R	Somerton	51 13 52 N	3 2 53 W	55
98	GB45R	Wicken Fen	52 17 54 N	0 17 34 W	5
99	GR 1R	Aliartos	38 22 0 N	23 5 0 E	110
100	GR 2R	Finokalia	35 30 0 N	26 10 0 E	0
101	GR 3R	Livadi	40 32 0 N	23 15 0 E	850
102	HU 2R	K-pusztá	46 58 0 N	19 35 0 E	125
103	IE31R	Mace Head	53 10 0 N	9 30 0 W	15
104	IT 1R	Montelibretti	42 6 0 N	12 38 0 E	48
105	IT 4R	Ispra	45 48 0 N	8 38 0 E	209
106	LT15R	Preila	55 21 0 N	21 4 0 E	5
107	LV10R	Rucava	56 13 0 N	21 13 0 E	5
108	MT 1R	Giordan lighthouse	36 6 0 N	14 12 0 E	160
109	NL 2R	Witteveen	52 48 49 N	6 40 11 E	17
110	NL 9R	Kollumerwaard	53 20 2 N	6 16 38 E	1
111	NL10R	Vredepeel	51 32 28 N	5 51 13 E	28
112	NO 1R	Birkenes	58 23 0 N	8 15 0 E	190
113	NO15R	Tustervatn	65 50 0 N	13 55 0 E	439
114	NO30R	Jergul	69 27 0 N	24 36 0 E	255
115	NO39R	Karvatn	62 47 0 N	8 53 0 E	210
116	NO41R	Osen	61 15 0 N	11 47 0 E	440
117	NO42G	Spitsbergen, Zeppelinfjell	78 54 0 N	11 53 0 E	474
118	NO43R	Prestebakke	59 0 0 N	11 32 0 E	160
119	NO44R	Nordmoen	60 16 0 N	11 6 0 E	2

Table D.3: Stations de mesures (Suite 2)

Numéro	Code	Non de Station	Lat	lon	Alt (m)
120	NO45R	Jeloya	59 26 0 N	10 36 0 E	5
121	NO47R	Svanvik	69 27 0 N	30 2 0 E	30
122	NO48R	Voss	60 36 0 N	6 32 0 E	5
123	NO49R	Valle	59 3 0 N	7 34 0 E	250
124	NO52R	Sandve	59 12 0 N	5 12 0 E	15
125	NO55R	Karasjok	69 28 0 N	25 13 0 E	333
126	NO56R	Hurdal	60 22 0 N	11 4 0 E	3
127	PL 2R	Jarczew	51 49 0 N	21 59 0 E	180
128	PL 3R	Sniezka	50 44 0 N	15 44 0 E	1603
129	PL 4R	Leba	54 45 0 N	17 32 0 E	2
130	PL 5R	Diabla Gora	54 9 0 N	22 4 0 E	157
131	PT 4R	Monte Velho	38 5 0 N	8 48 0 W	43
132	RU 1R	Janiskoski	68 56 0 N	28 51 0 E	118
133	RU13R	Pinega	64 42 0 N	43 24 0 E	28
134	RU16R	Shepeljovo	59 58 0 N	29 7 0 E	4
135	RU18R	Danki	54 54 0 N	37 48 0 E	150
136	SE 2R	Rorvik	57 25 0 N	11 56 0 E	10
137	SE11R	Vavihill	56 1 0 N	13 9 0 E	175
138	SE12R	Aspvreten	58 48 0 N	17 23 0 E	20
139	SE13R	Esrange	67 53 0 N	21 4 0 E	475
140	SE32R	Norra-Kvill	57 49 0 N	15 34 0 E	261
141	SE35R	Vindeln	64 15 0 N	19 46 0 E	225
142	SE39R	Grimso	59 45 0 N	15 35 0 E	120
143	SI 8R	Iskrba	45 34 0 N	14 52 0 E	520
144	SI31R	Zarodnje	46 25 43 N	15 0 12 E	770
145	SI32R	Krvavec	46 17 58 N	14 32 19 E	1740
146	SI33R	Kovk	46 7 43 N	15 6 50 E	6
147	SK 2R	Chopok	48 56 0 N	19 35 0 E	28
148	SK 4R	Stara Lesna	49 9 0 N	20 17 0 E	808
149	SK 6R	Starina	49 3 0 N	22 16 0 E	345
150	SK 7R	Topolniky	47 57 36 N	17 51 38 E	113
151	TR 1R	Cubuk II	40 30 0 N	33 0 0 E	1169

Table D.4: Stations de mesures (Suite 3)