



**HAL**  
open science

# Assimilation variationnelle de données de télédétection dans des modèles de fonctionnement des couverts végétaux et du paysage agricole

Emmanuel Kwashi Kpemie

► **To cite this version:**

Emmanuel Kwashi Kpemie. Assimilation variationnelle de données de télédétection dans des modèles de fonctionnement des couverts végétaux et du paysage agricole. Sciences agricoles. Université d'Avignon, 2009. Français. NNT : 2009AVIG0629 . tel-00555416

**HAL Id: tel-00555416**

**<https://theses.hal.science/tel-00555416>**

Submitted on 13 Jan 2011

**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.



ACADEMIE D'AIX-MARSEILLE  
UNIVERSITE D'AVIGNON ET DES PAYS DE VAUCLUSE

## THESE DE DOCTORAT

Présentée par : **Emmanuel Kwashi KPEMLIE**

A l'Université d'Avignon et des Pays de Vaucluse  
Ecole Doctorale SIBAGHE  
Spécialité Sciences Agronomiques

# Assimilation variationnelle de données de télédétection dans des modèles de fonctionnement des couverts végétaux et du paysage agricole

Date de soutenance : 18 Décembre 2009

M. Olivier BANTON	Professeur à l'Université d'Avignon	Président du jury
M. Eric BLAYO	Professeur à l'Université de Grenoble	Rapporteur
Mme. Catherine OTTLE	Directrice de recherche CNRS Gif sur-Yvette	Rapporteur
M. Laurent PREVOT	Chargé de recherche INRA Montpellier	Examineur
M. Albert OLIOSO	Chargé de recherche INRA Avignon	Directeur de thèse



## *Remerciement*

Pour remercier tous ceux qui m'ont soutenus et encouragés de près et de loin dans ce travail, Albert OLIOSO, Samuel BUIS, Dominique COURAULT, Tereza KPEMLIE, tous les collègues et autres ; je vous adresse ces citations qui m'ont été contées par les sages africains. J'espère qu'elles vous inspireront et vous guideront afin de faire un bon usage de ce rapport de thèse. »

« *Les morts ne sont pas morts, seul l'oubli des vivants qui fait mourir les morts* »

« *L'erreur n'annule pas la valeur de l'effort accompli* »

« *La réflexion veillera sur toi, L'intelligence te gardera* »

« *Va vers la fourmi, paresseux; Considère ses voies, et deviens sage* »

« *Au bout de la patience, il y a le ciel* »

« *La langue qui fourche fait plus de mal que le pied qui trébuche* »

« *Le mensonge donne des fleurs mais pas de fruits* »



## Tables des matières

INTRODUCTION GENERALE.....	- 9 -
PREMIERE PARTIE : .....	- 15 -
ETAT DE L'ART ET PROPOSITION.....	- 15 -
Introduction .....	- 17 -
1. Différentes approches d'estimation de l'évapotranspiration.....	- 19 -
1.1. Approches semi empiriques directs.....	- 19 -
1.2. Méthodes résiduelles .....	- 21 -
1.2.1. SEBAL .....	- 21 -
1.2.2. SEBI, S-SEBI, SEBS .....	- 24 -
1.3. Approches déterministes .....	- 28 -
2. Assimilation de données.....	- 33 -
3. Apport de notre méthode dans l'estimation de l'évapotranspiration .....	- 37 -
3.1.1. Approche patchée « tiled approach » .....	- 38 -
3.1.2. Assimilation variationnelle de données.....	- 40 -
Conclusion.....	- 43 -
DEUXIEME PARTIE : .....	- 45 -
METHODES, MATERIELS ET OUTILS .....	- 45 -
1. Introduction .....	- 47 -
2. Description du modèle PBLs .....	- 48 -
2.1. Le modèle de couche limite .....	- 49 -
2.2. Le modèle de surface.....	- 52 -
2.2.1. Le rayonnement net $R_n$ .....	- 53 -
2.2.2. Le flux conductif dans le sol $G$ .....	- 54 -
2.2.3. Les flux convectifs de chaleur sensible $H$ et de chaleur latente $LE$ (et les résistances associées) .....	- 55 -
2.3. Utilisation du modèle PBLs .....	- 61 -
3. Assimilation variationnelle de données : théorie et application au modèle PBLs dans notre cas d'étude .....	- 64 -
3.1. Théorie .....	- 64 -
3.1.1. La fonction coût et son gradient.....	- 65 -
3.1.2. A propos de l'opérateur adjoint.....	- 67 -
3.1.3. Préconditionnement de la fonction coût.....	- 68 -
3.1.4. Algorithme de minimisation.....	- 69 -
3.2. Application au modèle PBLs .....	- 69 -
3.2.1. Cas d'un seul pixel .....	- 70 -
3.2.2. Cas à $n$ pixels.....	- 73 -
4. Données utilisées.....	- 76 -
4.1. Données Alpilles ReSeDA .....	- 76 -
4.1.1. Généralités sur le programme Alpilles ReSeDA.....	- 76 -
4.1.2. Description (succincte) des données Alpilles.....	- 77 -
4.1.3. Données utilisées pour alimenter PBLs .....	- 80 -
4.2. Les autres jeux de données.....	- 84 -
4.2.1. Description (succincte) des données du site 'Flux et Télédétection' d'Avignon (Site GRUE).....	- 85 -
4.2.2. Description (succincte) des données simulées par SiSPAT .....	- 87 -
4.2.3. Description (succincte) des données simulées par TEC.....	- 88 -
5. Mise en œuvre de la méthode d'assimilation sur les données Alpilles.....	- 90 -

5.1.	Introduction .....	- 90 -
5.2.	Expérience synthétique .....	- 91 -
5.2.1.	Généralités.....	- 91 -
5.2.2.	Mise en pratique .....	- 92 -
5.3.	Mise en œuvre sur les données réelles .....	- 97 -
5.4.	Définition des informations a priori .....	- 98 -
5.4.1.	Cas des données réelles .....	- 98 -
5.4.2.	Cas des expériences jumelles .....	- 101 -
5.5.	Evaluations des résultats .....	- 102 -
6.	Outils TAPENADE et PALM .....	- 105 -
6.1.	Le logiciel de différenciation automatique TAPENADE .....	- 106 -
6.1.1.	Principe de fonctionnement.....	- 106 -
6.1.2.	Principe d'utilisation .....	- 108 -
6.1.3.	Validation du code de l'adjoint .....	- 109 -
6.1.4.	Stratégie et outils de débogage avec TAPENADE .....	- 110 -
6.1.5.	Performances d'un code adjoint.....	- 112 -
6.2.	PALM : un coupleur de code de calculs pour l'assimilation de données ...	- 113 -
6.2.1.	Unité .....	- 114 -
6.2.2.	Branche.....	- 116 -
6.2.3.	Communication .....	- 116 -
6.2.4.	Compilation et exécution d'une application PALM .....	- 117 -
TROISIEME PARTIE : .....		- 120 -
APPLICATIONS, ANALYSES ET MODIFICATIONS DU MODELES PBLs.....		- 120 -
1.	Analyse de sensibilité du modèle PBLs .....	- 122 -
1.1.	Effet du 'tiled approach' et de la prise en compte des rétroactions sur surface hétérogène – atmosphère à l'échelle du paysage agricole.....	- 123 -
1.1.1.	Illustration sur un paysage simplifié .....	- 123 -
1.1.2.	Illustration sur un vrai paysage (données Alpilles).....	- 128 -
1.2.	Sensibilité des variables d'intérêt du modèle PBLs aux erreurs sur ses variables d'entrée.....	- 132 -
2.	Modifications du modèle PBLs .....	- 136 -
2.1.	Etude du rapport $G/R_n$ .....	- 137 -
2.1.1.	Positionnement du problème et introduction de l'analyse .....	- 137 -
2.1.2.	Cas d'un couvert végétal peu développé ou développé .....	- 140 -
2.1.2.1.	Apprentissage sur données SiSPAT.....	- 142 -
2.1.2.2.	Validation sur données Alpilles et Site 'Flux et Télédétection' .....	- 144 -
2.1.3.	Cas du sol nu .....	- 147 -
2.1.3.1.	Apprentissage sur les données TEC .....	- 151 -
2.1.3.2.	Validation du terme $A$ sur les données Grue .....	- 154 -
2.1.4.	Conclusion sur le rapport $G/R_n$ .....	- 157 -
2.2.	Etude du rapport $kB^{-1}$ .....	- 159 -
2.2.1.	Positionnement du problème et introduction de l'analyse .....	- 159 -
2.2.2.	Description des modèles étudiés .....	- 162 -
2.2.2.1.	Modèle de Hasager.....	- 162 -
2.2.2.2.	Modèle de Blümel .....	- 163 -
2.2.2.3.	Modèle de Massman.....	- 167 -
2.2.2.4.	Illustration du fonctionnement des modèles .....	- 169 -
2.2.3.	Détermination de $kB^{-1}$ sur les données expérimentales .....	- 170 -
2.2.3.1.	Méthode de calcul du paramètre $kB^{-1}$ .....	- 170 -
2.2.3.2.	Analyse de la sensibilité du calcul de $kB^{-1}$ aux mesures .....	- 173 -

2.2.3.3.	Evolution des $kB^{-1}$ calculés sur les données su site ‘Flux et Télédétection’ d’Avignon	- 175 -
2.2.4.	Application des modèles de $kB^{-1}$ sur les données expérimentales	- 178 -
2.2.4.1.	Evaluation des modèles sur les données ‘Flux et Télédétection’ d’Avignon	- 178 -
2.2.4.2.	Analyse de l’impact de la détermination de la rugosité aérodynamique $z0_m$ sur les modèles de $kB^{-1}$ et les calculs de flux $H$	- 182 -
2.2.4.3.	Evaluation des modèles sur les données Alpilles	- 184 -
2.2.5.	Conclusion sur le paramètre $kB^{-1}$	- 186 -
3.	Applications : étude comparative des versions 1 & 2 du modèle PBLs par rapport aux données Alpilles	- 187 -
QUATRIEME PARTIE :		- 192 -
IMPLEMENTATION ET VERIFICATION DU SYSTEME D’ASSIMILATION DE DONNEES		- 192 -
1.	Introduction	- 194 -
2.	Génération, validation et performances de l’adjoint	- 195 -
2.1.	Génération de l’adjoint du modèle PBLs	- 195 -
2.2.	Validation de l’adjoint du modèle PBLs	- 198 -
2.3.	Performances et optimisation de l’adjoint	- 199 -
3.	Assemblage et vérification du système d’assimilation	- 201 -
3.1.	Assemblage du système d’assimilation avec le logiciel PALM	- 201 -
3.1.1.	Unités	- 201 -
3.1.2.	Construction schéma de couplage	- 203 -
3.2.	Vérification	- 206 -
3.2.1.	Estimation des états vrais des paramètres $z0_i$ et $f2_i$	- 206 -
3.2.2.	Indépendances des résultats par rapport aux valeurs initiales	- 208 -
3.2.3.	Poids des matrices de covariances d’erreur de l’ébauche et d’observation dans l’estimation	- 210 -
3.2.4.	Convergence de la fonction coût	- 213 -
4.	Conclusion	- 215 -
CINQUIEME PARTIE :		- 218 -
PERFORMANCES DE L’ALGORITHME D’ASSIMILATION SUR LES DONNEES SYNTHETIQUES ET SUR LES DONNEES REELLES DE L’EXPERIENCE ALPILLES – ReSeDA		- 218 -
Introduction		- 220 -
1.	Performances sur les données synthétiques	- 222 -
1.1.	Résultats de l’algorithme d’assimilation avec la version 1 de PBLs	- 222 -
1.1.1.	Cas sans erreur modèle	- 222 -
1.1.2.	Cas avec erreur modèle	- 227 -
1.1.3.	Comparaison méthode résiduelle – assimilation	- 236 -
1.2.	Résultats de l’algorithme d’assimilation avec la version 2 de PBLs	- 239 -
1.2.1.	Cas sans erreur modèle	- 240 -
1.2.2.	Cas avec erreur modèle	- 244 -
1.2.3.	Comparaison méthode résiduelle – assimilation	- 248 -
1.3.	Conclusion sur les données synthétiques	- 249 -
2.	Performances sur les données réelles d’Alpilles – ReSeDA	- 253 -
2.1.	Introduction	- 253 -
2.2.	Analyse de l’assimilation par rapport aux observations	- 254 -
2.3.	Validation sur les parcelles de mesure	- 259 -
2.4.	Impact de l’information a priori de l’humidité du sol	- 265 -

2.5. Récapitulatif des résultats autour du midi solaire et cartographies de l'évapotranspiration.....	- 269 -
2.6. Conclusion sur les données réelles.....	- 275 -
Conclusion.....	- 277 -
CONCLUSION GENERALE .....	- 279 -
BIBLIOGRAPHIE .....	- 285 -
ANNEXES .....	- 312 -

## ***INTRODUCTION GENERALE***

Dans le domaine agricole, il est de plus en plus utile de connaître avec précision l'état des cultures pour mieux gérer les différentes interventions telles que les irrigations ou les traitements phytosanitaires, afin d'obtenir le meilleur rendement possible tout en respectant l'environnement. La connaissance du climat par exemple à l'échelle de la parcelle agricole est une donnée importante pour comprendre le développement de la culture. Le suivi des conditions hydriques des cultures passe par l'estimation des flux de surface (telle que l'évapotranspiration ou le flux de chaleur latente qui représente la consommation réelle en eau de la culture) et des conditions atmosphériques, les deux étant reliés par un principe de rétroaction. Les flux de surface agissent sur l'atmosphère, et notamment sur la température et l'humidité de l'air, et en retour l'atmosphère modifie les flux de surface.

Il existe dans la littérature de nombreux modèles permettant d'estimer les flux de surface allant d'une approche résiduelle du bilan d'énergie (Wiegand & Bartholic, 1970; Brown & Rosenberg, 1973) à des modèles décrivant l'ensemble des processus impliqués appelés généralement modèles TSVA (modèles de Transfert Sol-Végétation-Atmosphère, ou SVAT models en anglais pour Soil Vegetation Atmosphere Transfer models ; voir e.g. Deardoff, 1978; Dickinson et al., 1986; Sellers et al., 1996). Par ailleurs, les modèles de fonctionnement des couverts végétaux simulent les processus physiologiques à la base du développement et de la croissance des couverts (e.g. Brisson et al., 2008; Sinclair & Seligman, 1996). Ils peuvent être utilisés pour le diagnostic de l'état des cultures et la prévision de leur production (comme par exemple le modèle STICS développé à l'INRA par Brisson et al., 2008). Ils peuvent également être utilisés pour décrire le fonctionnement des différentes entités qui composent la majeure partie d'un paysage agricole : cultures, bois, jachères,... Ils permettent ainsi de mettre en place une première approche pour l'analyse du fonctionnement d'un paysage, par exemple en terme de consommation en eau des couverts végétaux, d'étude d'impact du climat sur les végétaux, ou encore de comprendre les processus d'assimilation de carbone ou l'extraction de l'eau par les plantes (e.g. Norman, 1979; Jones & Zur, 1984; Oliosio et al., 1996a; Tuzet et al., 2003).

A l'échelle du paysage, le renseignement des modèles de fonctionnement passe forcément par l'utilisation de la télédétection et de procédures d'assimilation de ces données dans les

modèles (Cayrol et al., 2000 ; Olioso et al., 2005 ; Gowda et al., 2008). La télédétection permet en effet d'accéder à des variables qui sont étroitement associées à l'évaporation telle que par exemple la température de surface avec l'utilisation des bandes spectrales dans l'infrarouge thermique. Par exemple encore, les indices foliaires simulés par un modèle peuvent être confrontés à leurs estimations à partir des mesures de télédétection ; il est alors possible de recalibrer le modèle en modifiant certains de ses paramètres qui étaient mal connus, en particulier en raison de leur variabilité spatiale à l'échelle du paysage. Une autre approche consiste, non plus à re-estimer des paramètres, mais à corriger l'évolution des variables dynamiques du modèle (indice foliaire, humidité du sol). L'un des principaux objectifs de certaines récentes recherches est de tester la faisabilité d'évaluer le bilan d'énergie et le bilan hydrique à l'échelle régionale avec des modèles utilisant des informations de télédétection (e.g. Su et al., 2008; Chehbouni et al., 2008; Courault et al., 2008; Kustas et al., 2003).

Jusqu'à présent, les travaux d'assimilation se sont essentiellement attachés à analyser l'apport des données acquises dans les domaines solaire et micro-onde, qui permettent d'accéder aux informations sur la structure de la végétation (Prévoit et al., 2003). Par contre, bien que l'intérêt des mesures de télédétection dans l'infrarouge thermique (température de surface) pour le suivi du fonctionnement hydrique des couverts, en particulier en conditions non optimales, soit démontré depuis longtemps (Jackson et al., 1977 ; Soer, 1980; Seguin et al., 1982 ; Nieuwenhuis et al., 1985 ; Vidal et al., 1987), leur assimilation dans les modèles de fonctionnement n'a été que très peu étudiée. Ceci est dû à deux difficultés principales :

- d'une part, il est difficile de relier les mesures thermiques aux variables simulées par les modèles de fonctionnement, en particulier en raison de l'inadéquation entre le pas de temps des modèles (la journée) et l'acquisition instantanée des mesures de température de surface ;
- d'autre part, la disponibilité et la précision des mesures de télédétection dans l'infrarouge thermique sont souvent insuffisantes dû en particulier à l'offre en capteurs opérationnels, à l'opacité des nuages aux rayonnements infrarouges, à la présence de l'atmosphère entre la surface observée et le capteur (surtout dans le cas d'un capteur embarqué sur un satellite) et aux caractéristiques propres de la surface (émissivité en particulier), qui modifient notablement l'information disponible et nécessitent généralement la mise en

place de méthodes de correction des données (Deschamps & Phulpin, 1980 ; Prata et al., 1995 ; Kerr et al., 1992).

Récemment, de nouveaux types de modèles de fonctionnement ont été développés, soit à l'INRA, soit dans d'autres organismes, comme Météo-France (ISBA-Ags, Calvet et al., 1998 ; couplage STICS-TSVA, Olioso et al., 2005). Ces modèles combinent des fonctionnalités classiques des modèles de fonctionnement, comme la simulation de la croissance des couverts, avec les fonctionnalités des modèles TSVA qui permettent en particulier de simuler la température de surface tout au long de la journée, et ainsi facilitent l'interfaçage avec les mesures de télédétection.

Par ailleurs, les travaux récents sur la cartographie de l'évapotranspiration à partir de la température de surface montrent qu'il est possible de limiter l'effet de la précision insuffisante des mesures en exploitant les informations contenues dans la variabilité spatiale des images de télédétection (Su, 2002 ; Jacob et al., 2002a ; Olioso et al., 2002a ; Moran et al., 1989). Il reste cependant nécessaire d'analyser cette variabilité spatiale et de pouvoir l'interpréter en terme de variabilité à l'échelle des couverts végétaux et des propriétés du sol (comme en terme d'humidité du sol, de propriétés hydrodynamiques des sols, de rugosité aérodynamique, d'indice foliaire...). Cela peut nécessiter l'utilisation combinée de l'information géographique, des données de télédétection et de l'assimilation de données (Olioso et al., 2006). Un point à ne pas sous-estimer également porte sur l'accès aux données météorologiques nécessaires à l'estimation des flux de surface. Afin de pallier aux difficultés liées à l'obtention des informations météorologiques aux échelles spatio-temporelles requises, certains modèles intègrent une description de la couche limite planétaire (PBL) ce qui leur permet de simuler l'évolution de variables comme la température de l'air et l'humidité de l'air (Brunet et al., 1991 ; Jacobs et al., 1997; Margulis & Entekhabi, 2003).

Certaines caractéristiques de surface telles que l'indice foliaire (*LAI* : *Leaf Area Index*) et l'albédo ( $\alpha$ ) peuvent être estimées par le biais des images de télédétection avec une précision de 0.05 pour *LAI* et 0.02 pour l'albédo (e.g. Asrar et al., 1984; Baret & Guyot, 1991; Lucht et al., 2000; Jacob et al., 2002b; Weiss et al., 2002; Bsaibes et al., 2009 ). D'autres caractéristiques importantes telles que l'humidité du sol dans la zone racinaire et la rugosité aérodynamique sont par contre difficiles à estimer directement et il est nécessaire de mettre en place des procédures permettant malgré tout d'accéder à leurs variabilités spatio-temporelles.

Les méthodes d'assimilation de données ont été proposées pour retrouver les paramètres du modèle ou les valeurs initiales des variables du modèle en utilisant les techniques de calibration, basées soit sur des algorithmes itératifs (Verhoef & Bach, 2003; Judge et al., 2008) ou des méthodes stochastiques (Demarty et al., 2004; Zhoua et al., 2006 ; Intsiful & Kunstmann, 2008).

Nous proposons donc dans ce travail d'analyser la possibilité d'utiliser les données de télédétection pour estimer l'évapotranspiration et les interactions surface – couche limite à l'échelle de quelques dizaines de kilomètres carrés en utilisant des outils mathématiques nouveaux dans ce domaine pour l'application des procédures d'assimilation :

- Nous avons utilisé un modèle simplifié de couplage surface – couche limite (Brunet et al., 1991) qui a été adapté pour prendre en compte la variabilité spatiale de la surface par l'approche dite 'patchée' (Avissar & Pielke, 1989; Blyth, 1995). Cette approche tient compte des différents flux verticaux provenant de plusieurs zones agricoles et nécessite d'introduire dans le modèle les diverses proportions et les caractéristiques des principaux couverts végétaux qui composent le paysage.
- Nous avons utilisé la méthode d'assimilation variationnelle de données, concept utilisé en météorologie entre autres pour les prévisions (Bouttier & Courtier, 1999). Elle consiste à minimiser une fonction mesurant l'écart entre l'état du système simulé par le modèle et les observations collectées (ici la température de surface). Un algorithme est utilisé pour estimer les variables non accessibles directement du modèle (ici rugosité aérodynamique et l'humidité normalisée du sol) en ajustant ces variables afin d'obtenir une bonne cohérence entre les sorties du modèle (températures de surface simulées) et les observations (températures de surface mesurées). Le grand nombre d'informations à retrouver (rugosité et humidité pour chacun des patchs considérés) rend complexe et coûteux le problème de minimisation. Il est donc avantageux d'utiliser la méthode adjointe habituellement appliquée pour les problèmes de grande taille. Par conséquent, nous avons développé l'adjoint du modèle surface - couche limite permettant ainsi de réduire le temps de calcul du processus de minimisation.
- Nous avons mené dans un premier temps une étude de cette possibilité d'estimer l'évapotranspiration par la méthode d'assimilation variationnelle de données sur des

données expérimentales. Les données expérimentales sont donc construites pour représenter l'environnement des Alpilles situé au sud-est de la France (Baret, 2002). L'analyse sur ces données nous a donc permis d'évaluer l'apport de l'assimilation ainsi que les impacts d'erreurs liées à la température de surface, et sur les paramètres d'albédo, de la vitesse du vent et de l'indice foliaire (qui sont des entrées de la méthode). Nous avons également pu tester l'apport d'information sur les statistiques d'erreurs de mesures dans la procédure d'assimilation et confronter les résultats de la méthode à des approches plus classiques. Ce travail fait dans un second temps l'objet de recherche, qui se veut donc résolument tournée vers une application opérationnelle.

Dans une première partie, nous donnerons une revue bibliographique sur les modèles d'estimation de l'évapotranspiration, qui nous permettra de mieux situer le modèle sur lequel nous avons choisi de travailler.

La seconde partie sera consacrée à la description du modèle, de la théorie d'assimilation variationnelle de données et de son application à notre modèle. Par la suite, nous décrirons les différents jeux de données que nous avons utilisés au cours de ce travail et présenterons brièvement des outils nécessaires pour la mise en œuvre de la méthode d'assimilation.

Des modifications ont été apportées au modèle utilisé, en particulier l'approche adoptée pour représenter l'hétérogénéité de la surface (approche patchée). Ces modifications seront présentées dans la troisième partie. Nous illustrerons aussi la sensibilité du modèle à certaines variables d'entrée importantes. Enfin, nous présenterons des études de modifications sur certaines équations du modèle pour améliorer ces sorties, notamment liées au flux de conduction dans le sol et à la rugosité thermique.

Dans la quatrième partie, nous exposerons les étapes de l'implémentation de la méthode dans notre modèle et les vérifications effectuées sur la mise en œuvre du système d'assimilation à partir de données synthétiques.

Nous évaluerons et comparerons dans une cinquième partie, les performances du système d'assimilation lorsque les conditions de mise en œuvre se dégradent. Une attention particulière sera portée sur la propagation des erreurs des variables d'entrée sur l'estimation de

l'évapotranspiration ainsi que les conséquences que peut avoir une mauvaise caractérisation des paramètres du modèle. Nous confronterons ensuite le système à une méthode classique d'estimation de l'évapotranspiration sur des données expérimentales. Nous testerons en dernier lieu la performance du système d'assimilation sur des cas réels issus de l'expérimentation Alpilles-ReSeDA. Ainsi, nous serons amenés à cartographier les flux de surface (particulièrement l'évapotranspiration), le microclimat, l'humidité normalisée du sol, et la rugosité aérodynamique à l'échelle régionale.

***PREMIERE PARTIE :***

***ETAT DE L'ART ET PROPOSITION***



## ***Introduction***

Actuellement, le terme « évapotranspiration » occupe une place particulièrement importante dans diverses publications scientifiques. Le suivi des transferts d'énergie et de masse au niveau des sols et des couverts végétaux est déterminant pour la gestion des ressources en eau et des ressources végétales. Il est également nécessaire à la bonne compréhension des systèmes hydrologiques et climatiques, ainsi qu'au suivi et à la prévision de leurs évolutions. La modélisation du fonctionnement du couvert est un moyen privilégié pour réaliser cette analyse puisqu'elle permet de décrire les interactions entre les caractéristiques des couverts (aussi bien physiques que physiologiques) et celles de l'environnement (atmosphère et sol). Des modèles ont ainsi été développés pour étudier l'évapotranspiration en utilisant différentes approches (Jackson et al., 1977 ; Anderson et al., 1997; Bastiaanssen et al., 1998a; Noilhan & Planton, 1989 ; Allen et al., 1998). Certains de ces modèles font appel aux mesures de télédétection en relation avec le fonctionnement du couvert pour leur application à différentes échelles spatiales (ex : Soer, 1980; Carlson & Lynn, 1991 ; Jackson et al., 1977 ; Anderson et al., 1997; Bastiaanssen et al., 1998a). La télédétection est en effet un outil privilégié pour réaliser le suivi de l'évapotranspiration car elle fournit des informations liées aux transferts d'énergie et de masse, et en particulier aux processus d'évaporation des sols et de transpiration par les végétaux (Vidal, 1990 ; Olioso, 1994). Les dispositifs d'acquisition de données par télédétection, tant satellitaires qu'aéroportés, offrent la possibilité de collecter des mesures radiométriques sur des zones d'extension régionale, avec une résolution spatiale suffisante pour l'identification des caractéristiques d'une parcelle agricole. La télédétection peut donc constituer un outil de suivi performant et intéressant à plus d'un titre lorsque la finalité est d'estimer avec précision l'évapotranspiration (Bastiaanssen & Bos, 1999). C'est particulièrement le cas lorsque l'on cherche à l'estimer à des échelles spatiales larges, les méthodes de mesures in situ, comme en particulier les méthodes du rapport de Bowen et des corrélations turbulentes, étant délicates à mettre en place et ne permettant pas d'explorer de grandes surfaces. Il convient toutefois de noter que la faible disponibilité actuelle des images infrarouge thermique avec une résolution spatiale adaptée à l'échelle de la parcelle (type ASTER ou LANDSAT qui ne passent que tous les 16 jours) pose problème pour un suivi temporel de l'évapotranspiration qui doit être fin. Un suivi continu de l'évapotranspiration est en effet nécessaire dans des problématiques hydrologiques ou agronomiques. Ainsi, certains auteurs (Coudert et al., 2006 ; Olioso et al., 2005 ; Moulin et al., 2002 ; Reichle & Entekhabi,

2001) ont proposé des techniques de calibrage ou de forçage des variables d'entrée de modèles de fonctionnement ou de modèles TSVA dites « d'assimilation de données » afin d'améliorer l'estimation et le suivi temporel de l'évapotranspiration.

Dans ce chapitre, sans être exhaustif nous listerons certaines méthodes d'estimation de l'évapotranspiration rencontrées dans la littérature, ainsi que les techniques mathématiques qui peuvent être utilisées pour améliorer la précision de cette variable. Pour conclure, nous introduirons la nouveauté apportée par la méthode développée dans ce travail de thèse.

## 1. Différentes approches d'estimation de l'évapotranspiration

Les modèles permettant d'accéder à l'évapotranspiration par télédétection se basent généralement sur l'estimation du flux de chaleur latente ( $LE$ ) par l'intermédiaire de la résolution de l'équation du bilan d'énergie (I.1). Cette équation permet d'exprimer la conservation de l'énergie au niveau de la surface terrestre en reliant l'énergie reçue (rayonnements incidents) aux différents flux de dissipation de cette énergie vers l'environnement, en particulier par transferts turbulents (chaleur sensible  $H$  et chaleur latente  $LE$ , ce dernier correspondant à un transfert de vapeur d'eau) et transferts par conduction dans le sol ( $G$ ) en supposant que les termes d'advection et de stockage de la chaleur dans la couche de végétation sont négligeables (Brutsaert, 1982) :

$$LE = R_n - G - H \quad (I.1)$$

où  $R_n$  est le rayonnement net (qui représente le bilan radiatif de la surface).

### 1.1. Approches semi empiriques directs

Les mesures de télédétection peuvent être utilisées directement dans des modèles semi empiriques qui calculent les flux à l'instant de la mesure. C'est probablement la forme de modèle la plus étudiée au cours des années 1980 et 1990 pour estimer l'évapotranspiration à partir de la température de surface (Jackson et al., 1977 ; Itier & Riou, 1982 ; Seguin & Itier, 1983 ; Vidal & Perrier, 1989 ; Lagouarde, 1991). En particulier, la relation simplifiée (Equation I.2 ci-dessous) a permis de cartographier l'évapotranspiration sur de vastes régions (Lagouarde & Brunet, 1994 ; Courault et al., 1994) avec plus ou moins de succès.

La méthode suppose qu'il est possible de relier directement l'évapotranspiration journalière ( $ET_d$ ) à la différence instantanée de  $(T_s - T_a)_i$  mesurée autour de midi (Récan, 1982 ; Carlson & Buffon, 1989),  $T_s$  et  $T_a$  étant les températures de surface et de l'air (dans certains cas  $T_a$  correspondait en fait à la température maximale journalière de l'air, mesure plus facilement accessible dans les réseaux météo que la température instantanée). L'évapotranspiration journalière s'écrit alors (Seguin & Itier, 1982 ; voir aussi Li et al., 2009 ; Courault et al., 2005):

$$ET_d = R_n + A - B (T_s - T_a)_i. \quad (I.2)$$

$A$  et  $B$  sont deux constantes empiriques déterminées en fonction de la situation locale de la zone. Une des hypothèses de base de la méthode considère que le rapport  $H/R_n$  est constant au cours de la journée (Itier & Riou, 1982 ; Chanzy, 1991). Le traitement statistique de cette équation et son échelle d'application journalière permettent de s'affranchir de la détermination du flux conductif dans le sol (i.e  $G_d = 0$ ) et de prendre implicitement en compte le rôle moyen du vent et des variations des variables météorologiques au cours de la journée. En plus de la température de surface, l'application de cette relation ne nécessite que deux variables : le maximum de la température de l'air (ou une température instantanée aux alentours de ce maximum ou au moment de la mesure de la température de surface) et le rayonnement net journalier. Seguin et al. (1982) et Steinmetz et al. (1989) ont montré que la précision pourrait atteindre 10-15% à l'échelle locale, mais également que les coefficients  $A$  et  $B$  variaient selon les expériences. Ils montrent qu'il est nécessaire de dissocier les cas où l'air est stable, de ceux où l'air est instable. Des études ont ainsi introduit différentes paramétrisations de ces coefficients comme fonction du vent, de la rugosité et des critères de stabilité atmosphérique (Riou et al., 1988 ; Vidal & Perrier, 1989 ; Lagouarde & McAneney, 1992).

Cependant, cette approche est limitée par deux contraintes importantes :

- elle ne peut être appliquée que dans le cas des journées claires bien ensoleillées : son utilisation pour le suivi du bilan hydrique suppose la mise au point de méthodes d'interpolation temporelle utilisant des modèles agrométéorologiques plus complets (Vidal et al., 1987 ; Courault et al., 1994) ;
- la mise au point de cette méthodologie nécessite des mesures au sol relativement contraignantes, d'autant que la dispersion des données mesurées est importante (Seguin & Itier, 1983)

## 1.2. Méthodes résiduelles

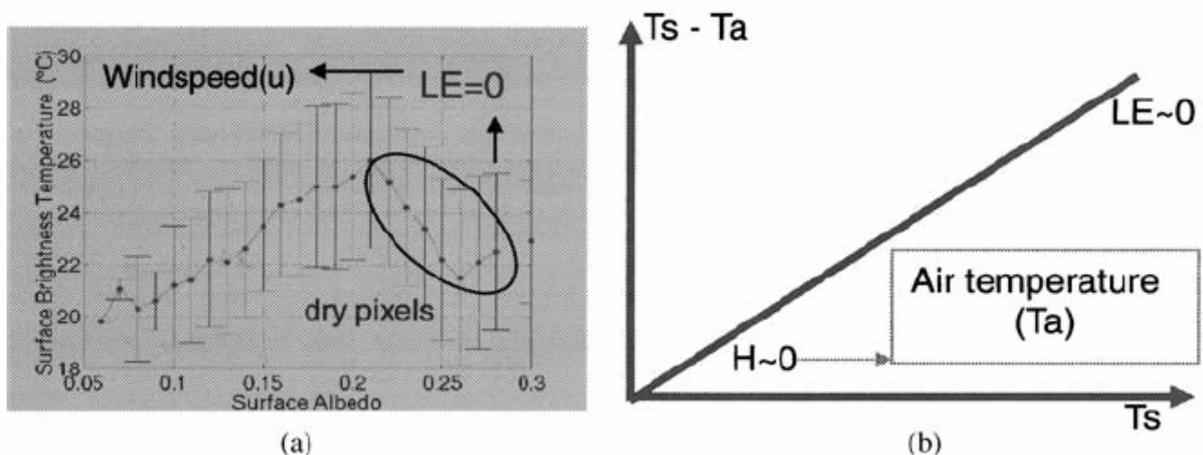
Les approches que nous qualifions d'approches résiduelles utilisent la combinaison de certaines relations empiriques et de modules physiques. La plupart des modèles opérationnels (tels que ALEXI, SEBAL, S-SEBI, METRIC) utilisent les données de télédétection pour estimer les paramètres d'entrée du modèle et ainsi l'évapotranspiration (Anderson et al., 1997; Bastiaanssen et al., 1998a; Menenti & Choudhury, 1993; Allen et al., 2007). Dans les méthodes résiduelles, la combinaison des équations de calcul des flux  $R_n$ ,  $H$ , et  $G$  à celle du bilan d'énergie permet de déterminer l'évapotranspiration, comme résultante du bilan d'énergie (Eq. I.1), à partir de la température de surface ( $T_s$ ), de variables météorologiques (rayonnement solaire incident  $R_g$ , rayonnement atmosphérique incident  $R_a$ , température de l'air  $T_a$ , et vitesse horizontale du vent  $u_a$ ) et de caractéristiques de la surface (émissivité  $\varepsilon$ , albédo  $\alpha$ , rugosités aérodynamique et thermique  $z_0$ ,  $z_{0h}$ , rapport  $G/R_n$ ). Les équations permettant de calculer les différents flux seront détaillées plus loin (description du modèle de surface dans la partie 2).

Par rapport à l'utilisation de la relation simplifiée présentée au paragraphe précédent, le développement des approches résiduelles au cours des 10 dernières années a permis des avancées significatives dans la possibilité de cartographier l'évapotranspiration. Ces avancées ont été essentiellement liées à la prise en compte dans les calculs de flux de méthodes permettant de borner les estimations de flux de chaleur sensible dans des limites physiques cohérentes, soit en les situant entre des valeurs maximales et minimales en se basant sur la variabilité spatiale au sein des images (ex. SEBAL pour lequel on définit des conditions limites humides et sèches), soit en analysant les évolutions temporelles de la température en début de matinée (ALEXI). Différentes approches sont détaillées dans la suite.

### 1.2.1. SEBAL

SEBAL est une approche intermédiaire qui se base sur l'approche résiduelle et qui utilise à la fois des relations empiriques et des paramétrisations physiques (Bastiaanssen et al., 1998a, 1998b; Jacob et al., 2002a). Ce modèle a été développé pour calculer la répartition de l'énergie à l'échelle régionale avec un minimum de données acquises au sol. Les variables atmosphériques habituellement nécessaires au calcul de l'évapotranspiration par la méthode résiduelle (température de l'air  $T_a$ , vitesse du vent  $u_a$ ) sont obtenues à partir des données de

télétection visible - proche infrarouge – infrarouge thermique et de simplifications des équations du bilan d'énergie en conditions limites sèches et humides. L'analyse de la variabilité spatiale induite par les contrastes hydrologiques et énergétiques sur les images permet la détermination des surfaces sèches et humides sur la zone étudiée (cf. Fig. I.1). Le modèle nécessite en entrée les rayonnements incidents (qui peuvent cependant être directement calculés dans certaines configurations de données), les cartes de température de surface  $T_s$ , d'indice de végétation  $NDVI$  (Normalized Difference Vegetation Index) et d'albédo. Des relations semi empiriques permettent d'estimer l'émissivité, la rugosité et le rapport  $G/Rn$  à partir du  $NDVI$  et de l'albédo. L'ensemble de ces informations est utilisé pour calculer les flux  $R_n$ ,  $G$  et  $H$  et le flux de chaleur latente  $LE$  résultante du bilan d'énergie (Eq. I.1).



**Fig. I.1 :** (a) Relation utilisée dans le modèle SEBAL entre l'albédo et la température de brillance (mesure infra rouge thermique) sur le site ALPILLES en 1997, qui permet de calculer la vitesse du vent (Jacob, 1999). (b) Relation spatiale utilisée dans le modèle SEBAL entre  $T_s$  et  $T_a$  pour estimer la température de l'air (quand  $(T_s - T_a) \sim 0$ ,  $H \sim 0$ ,  $T_a$  peut être estimée à partir des images TIR de  $T_s$ ) : issue de Courault et al., 2005

L'un des principes du modèle SEBAL, lors de l'évaluation pixel par pixel des flux de chaleur sensible et latente est d'établir des relations linéaires entre la température de surface ( $T_s$ ) et la différence entre les températures de surface et de l'air ( $dT = T_s - T_a$ ). Les coefficients des expressions linéaires sont déterminés à partir des points extrêmes secs et humides. Ainsi, la différence  $dT$  peut être estimée par une simple relation linéaire de la température de surface :

$$dT = a + bT_s \quad (I.3)$$

où  $a$  et  $b$  sont des coefficients empiriques issues des deux points d'ancrage (points secs et humides).

Sur les pixels secs, le flux de chaleur latente est supposé être égal à zéro et la différence entre les températures de surface et de l'air de ces pixels est obtenue par inversion de l'expression du flux de chaleur sensible :

$$dT_{dry} = \frac{H_{dry} \cdot r_a}{\rho c p} \quad (I.4)$$

avec  $H_{dry}$  est égal à  $R_n - G$ .

Sur les pixels humides, la valeur  $R_n - G$  est affectée au flux de chaleur latente, ce qui signifie que le flux de chaleur sensible dans cette condition est égal à zéro. De plus, la différence entre les températures de surface et de l'air sur ces pixels est aussi supposée égale à zéro ( $dT_{wet} = 0$ ). Après le calcul des différences entre les températures de surface et de l'air sur les pixels secs et humides, les coefficients  $a$  et  $b$  de l'équation (I.3) peuvent donc être déterminés.

Enfin, le flux de chaleur sensible  $H$  peut être obtenu sur chaque pixel de façon itérative suivant son expression décrite dans la deuxième partie du document. Les flux  $R_n$ ,  $G$  et  $H$  étant connus, par conséquent l'évapotranspiration peut être déterminé suivant l'équation (I.1).

Le modèle SEBAL offre une grande facilité d'utilisation et apporte des réponses concrètes aux problèmes posés par la cartographie de l'évapotranspiration. En particulier la cohérence des schémas utilisés permet de diminuer fortement les risques d'erreurs d'estimation de l'évapotranspiration (en particulier en passant par l'estimation directe de la différence  $dT$  qui évite de cumuler des erreurs d'estimation des températures de l'air et de surface). Le modèle a ainsi pu être utilisé dans différentes applications pour estimer l'évapotranspiration journalière, mensuelle et saisonnière (Droogers & Bastiaanssen, 2002 ; Bastiaanssen et al., 2005 ; Timmermans et al., 2007; Courault et al., 2009a). Différentes études de validation (cf. entre autres Bastiaanssen et al., 2005 ; Opoku-Duah et al., 2008) et en particulier celle réalisée par Jacob et al. (2002a) ont permis d'analyser et de vérifier la qualité du modèle à la fois sur les variables intermédiaires et sur les flux énergétiques de surface. L'écart absolu sur l'estimation

de l'évapotranspiration journalière est de l'ordre de 10 à 15% (Bastiaanssen et al., 2009 ; Thoreson et al., 2009). Les études récentes de Teixeira et al. (2009a ; 2009b) ont également permis de réexaminer le rôle des différentes entrées du modèle SEBAL sur l'évaluation de l'évapotranspiration et de la productivité de l'eau. Certaines études (comme Opoku-Duah et al., 2008 ; Courault et al., 2009a) ont montré le rôle de certaines hypothèses d'utilisation de SEBAL, et en particulier de celles portant sur le choix des pixels secs et humides, sur la qualité des résultats obtenus. De fait, l'hypothèse formulée dans le modèle de la présence d'un contraste hydrologique (c'est-à-dire pixels secs et humides) sur la zone étudiée apparaît déterminante. Par ailleurs, Norman et al. (2006) ont montré que l'hypothèse de linéarité entre la température de surface et le gradient de températures de surface et de l'air utilisé dans la définition du flux de chaleur sensible n'est pas valide en toute généralité et en particulier dans le cas de paysages fortement hétérogènes. Pour réduire l'impact des limites du modèle SEBAL pour cartographier l'évapotranspiration à l'échelle régionale et sur des surfaces complexes, Allen et al. (2005a ; 2005b ; 2007) ont développé le modèle METRIC qui associe certains des fondements de SEBAL à des améliorations portant sur l'intégration de certaines mesures in situ (vitesse du vent, température du point de rosée). Ces auteurs, ainsi que Gowda et al. (2008), ont pu évaluer les bonnes performances de ce nouveau modèle.

### 1.2.2. SEBI, S-SEBI, SEBS

En se basant également sur le contraste entre les zones sèches et humides, Menenti et Choudhury (1993) ont proposé une méthode pour déterminer une fraction évaporative ( $\Lambda$ ) qui exprime le rapport entre le flux de chaleur latente et l'énergie disponible exprimée par le rayonnement net. L'idée générale consiste à postuler que, pour une valeur de rayonnement net donnée et dans un espace où les paramètres atmosphériques peuvent être considérés comme constants, la température de surface est directement liée au rapport entre évaporation réelle et évaporation potentielle. On exprime donc un terme que l'on appellera évaporation relative  $\Lambda_r$  tel que (Van den Hurk, 2001) :

$$\Lambda_r = \frac{LE}{LE_p} = 1 - \frac{(T_s - T)/r_a - (T_{s,\min} - T)/r_{a,\min}}{(T_{s,\max} - T)/r_{a,\max} - (T_{s,\min} - T)} = 1 - SEBI \quad (I.5)$$

Avec  $T_{s,min}$  (respectivement  $T_{s,max}$ ) les températures de surface minimales (respectivement maximales);  $r_{a,min}$  (respectivement  $r_{a,max}$ ) les résistances aérodynamiques minimales (respectivement maximales);  $T$  la température de l'air au niveau de la couche limite atmosphérique;  $LE_p$  l'évapotranspiration potentiel. L'approche de paramétrage de SEBI a été proposée pour tenir compte de la variabilité spatiale de l'évapotranspiration réelle en raison des variations de l'albédo et de la rugosité aérodynamique (Menenti & Choudhury, 1993). Ce paramétrage est essentiellement une modification de *CWSI* (Crop Water Stress Index) proposé par Idso et al. (1981) et Jackson et al. (1987 ; 1988). Ainsi, avec  $LE_{wet}$  le flux de chaleur latente dans le cas limite des zones humides, la fraction évaporative est donnée par :

$$\Lambda = \frac{\Lambda_r . LE_{wet}}{R_n - G} \quad (I.6)$$

Le concept a été repris par Su, (2002 ; 2005) et Su et al. (2003) au sein d'un modèle plus complexe nommé **SEBS** qui permet de déterminer cette fraction évaporative ( $\Lambda$ ) en définissant:

- un ensemble d'outils pour les calculs des paramètres physiques de surface
- le calcul de la rugosité pour le transfert de chaleur ( $z0_h$ )
- le calcul de la fraction évaporative en calculant le bilan d'énergie sur un certain nombre de cas limites.

Comme pour SEBAL, dans le cas limite des zones sèches, le flux de chaleur latente  $LE$  est supposé être égal à zéro en raison de la faible humidité du sol. Dans le cas limite des zones humides, un concept différent de celui de SEBAL est développé : le flux de chaleur latente n'est limité que par l'énergie disponible et il atteint son taux potentiel  $LE_{wet}$  qui peut être calculé par une combinaison d'équations semblables à celle de Penman-Monteith (Monteith, 1965). Les flux de chaleur sensible dans les cas limites des zones sèches et humides peuvent alors s'exprimer par :

$$H_{dry} = R_n - G \quad (I.7)$$

$$H_{wet} = \frac{R_n - G - \frac{\rho c_p VDP}{\gamma r_a}}{1 + \frac{\Delta}{\gamma}} \quad (I.8)$$

Avec  $r_a$  la résistance aérodynamique dépendant de la longueur de Monin-Obukhov (Monin & Obukhov, 1954), qui à son tour, est fonction de la vitesse de frottement et du flux de chaleur sensible,  $\rho$  la masse spécifique de l'air,  $c_p$  la chaleur spécifique,  $\gamma$  la constante psychrométrique,  $\Delta$  la pente de la vapeur saturante, VDP déficit de pression de vapeur.

La fraction évaporative est donc définie dans ce modèle par l'équation (I.6) et l'expression de la fraction évaporative relative est la suivante :

$$\Lambda_r = 1 - \frac{H - H_{wet}}{H_{dry} - H_{wet}} \quad (I.9)$$

Le modèle SEBS a été validé par Jia et al. (2003) vis-à-vis de mesures du flux de chaleur sensible par scintillométrie sur trois sites situés en Espagne. Wood et al. (2003) l'ont appliqué au sein du programme ARM (Atmospheric Radiation Measurement) aux Etats-Unis et ont obtenu des estimations encourageantes des flux par rapport aux données d'un réseau de mesure du rapport de Bowen. Les résultats de Su et al. (2005) ont montré que la précision de l'évapotranspiration estimée par le modèle SEBS pouvait atteindre 10 à 15% sur une gamme de variation de la fraction évaporative de 0.5 à 0.9. Par ailleurs, Su et al. (2003) et Su (2002) ont montré que SEBS pouvait être utilisé pour estimer les évapotranspirations quotidiennes, mensuelles et annuelles ainsi qu'à des échelles locale et régionale sous tout type de régime de stabilité atmosphérique. Toutefois, le modèle requiert un grand nombre de paramètres et fait appel à une solution relativement complexe des flux de chaleur turbulents, ce qui peut être source de difficultés d'utilisation dans certaines conditions (en particulier lorsque les données ne sont pas disponibles avec une qualité suffisante).

Une méthode simplifiée de SEBI (**S-SEBI**) fut développée plus tard pour estimer les flux de surface de manière beaucoup plus directe à partir de données de télédétection (Roerink et al., 2000). Cette méthode exploite directement les images de température de surface pour déterminer la fraction évaporative sans avoir recours à des données météorologiques

supplémentaires autres que les rayonnements incidents (elle ne nécessite pas non plus la connaissance des rugosités de surface). La fraction évaporative est en effet simplement exprimée pour un pixel donné par l'écart normalisé de la température de surface aux températures des zones sèches (ici  $T_{s,max}$ ) et humides (ici  $T_{s,min}$ ) ayant un albédo similaire :

$$\Lambda = \frac{T_{s,max} - T_s}{T_{s,max} - T_{s,min}} \quad (I.10)$$

Malgré sa simplicité, ce modèle a en général fourni des résultats de bonnes qualités avec des précisions de l'ordre du mm par jour pour l'évapotranspiration journalière en tout point comparables aux performances de modèles plus complexes (Roerink et al., 2000 ; Sobrino et al., 2005; Sobrino et al., 2007 ; Gomez et al., 2005; Fan et al., 2007). Cela peut apparaître surprenant car les hypothèses du modèle sont extrêmement simplificatrices. Il y a cependant tout lieu de penser que les performances se dégraderaient si la méthode était appliquée à des paysages complexes en particulier avec de fortes variations de rugosité.

### 1.3. Approches déterministes

L'estimation de l'évapotranspiration peut se baser sur des modèles plus détaillés, dans lesquels l'ensemble des mécanismes de transfert au niveau du sol et/ou des couverts végétaux (transferts radiatifs, transferts turbulents, transferts hydriques) sont paramétrés indépendamment des mesures de télédétection. Ces modèles sont souvent qualifiés de modèles de Transfert Sol-Végétation-Atmosphère (modèles TSVA ; *SVAT models* en anglais). Leur résolution temporelle est inférieure à l'heure, en accord avec la dynamique des processus atmosphériques et de surface (Deardorff, 1978; Dickinson et al., 1986; Sellers et al., 1996). Cette résolution temporelle fine revêt également de l'importance lorsque les modèles sont utilisés en combinaison avec les données de télédétection, car celles-ci sont acquises instantanément. Nous faisons référence ici à des modèles aussi divers que MAGRET (Lagouarde, 1991 ; Courault et al., 1996), MESO-NH développé par le laboratoire d'aérodynamique et le CNRM de Toulouse (<http://www.aero.obs-mip.fr/mesonh/index2.html>; Courault et al., 2003), ISBA (Noilhan & Planton, 1989 ; Noilhan & Mahfouf, 1996), PBL (Brunet et al., 1991), SiSPAT (Braud et al., 1995 ; Demarty et al., 2002 ; Ji et al., 2009).

Dans tous les cas, les modèles TSVA donnent accès à un fonctionnement détaillé des couverts végétaux et des sols, et pas seulement à un nombre limité de variables finales telles que l'évapotranspiration, comme c'est le cas dans les approches semi empiriques et résiduelles présentées plus haut. En particulier, ils simulent des variables intermédiaires en lien avec les processus physiologiques ou les processus hydrologiques (Braud et al., 1995; Tuzet et al., 2003). Ils peuvent être utilisés pour simuler l'humidité des sols, et ainsi servir d'interface avec des modèles hydrologiques ou météorologiques. Ils ne dépendent pas des données de télédétection pour leur mise en œuvre. Cependant, ils peuvent être employés en combinaison avec ces mesures en utilisant des méthodes d'assimilation permettant ainsi leur amélioration (Olios et al., 1999a ; 1999b ; 2006 ; Calvet et al., 2001; Demarty et al., 2004).

Dans les modèles TSVA, le flux de chaleur latente n'est pas calculé comme résidu de l'équation du bilan d'énergie (I.1), mais paramétré à partir de la pression de vapeur de l'air et d'un coefficient d'échange de la vapeur d'eau (cf. partie 2). Beaucoup de modèles TSVA intègrent une description détaillée des transferts hydriques et de chaleur dans le sol, ce qui nécessite de connaître toute une série de paramètres physiques, comme les courbes de rétention, conductivités hydriques et thermiques (ex. SiSPAT, Braud et al., 1995), ou comme la capacité au champ et le point de flétrissement (ex. ISBA, Noilhan et Planton, 1989).

La description du bilan d'énergie de la surface peut faire appel à des équations plus ou moins complexes, en particulier séparant ou pas le fonctionnement de la végétation de celui de la surface du sol. On parle alors de système 'bi-couche' en opposition à un système mono-couche ne se basant que sur un seul bilan d'énergie (cf. Fig. I.2) (Norman et al., 1995 ; Anderson et al., 1997 ; Sanchez et al., 2008 ; Timmermans et al., 2007 ; Friedl, 2002). Elle permet une meilleure description du fonctionnement physiologique des couverts végétaux, mais aussi, et peut être surtout, une bien meilleure description du lien entre les transferts turbulents et la température de surface telle qu'elle peut être mesurée par télédétection. Il faut noter ici que ces descriptions bi-couches sont aussi bien utilisées au sein de modèles TSVA que pour le calcul du flux de chaleur sensible dans les approches résiduelles (par exemple Timmermans et al. 2007 ; French et al., 2005; Anderson et al., 1997). Toutefois, les hypothèses et les éléments nécessaires à la résolution des bilans d'énergie sol et végétation dans les modèles bi-couche sont complexes. Si les modèles bi-couches permettent de prendre en compte l'ensemble des interactions entre les différentes surfaces, il est nécessaire en revanche de disposer de données permettant d'approcher la complexité du couvert, comme par exemple des données radiométriques IRT multi-angulaires, permettant par exemple de séparer la réponse du sol de la réponse du couvert (Jia, 2004).

Même si en théorie cette approche apparaît plus réaliste, de nombreux auteurs ont montré qu'ils ne sont pas forcément supérieurs à une simple approche mono-couche. En effet, correctement calibrée une approche mono-couche donne souvent des résultats plus satisfaisants pour la description du bilan d'énergie qu'une approche bi-couche avec un mauvais paramétrage (Li et al., 2005; French et al., 2005 ; Kustas et al., 2007; Timmermans et al., 2007). Le point déterminant de la calibration des approches monocouches, en particulier lorsqu'elles sont utilisées en combinaison avec des mesures de télédétection dans l'infrarouge thermique repose sur la paramétrisation de la rugosité thermique (ou du rapport rugosité thermique – rugosité mécanique). Ce point sera particulièrement abordé dans le chapitre 3 de ce mémoire.

Il existe donc une grande diversité de modèles TSVA qui se caractérisent par des descriptions plus ou moins détaillées des différents compartiments représentés : sol, végétation, atmosphère. Pour illustrer cette diversité, une courte description de trois modèles utilisés dans la communauté française est donnée dans le paragraphe suivant. Il s'agit des modèles ISBA, MAGRET et SiSPAT. Un quatrième modèle, PBL (Brunet et al., 1991) qui inclut le couplage à la couche limite atmosphérique et qui va servir de support au travail présenté dans ce mémoire sera décrit en détail dans le chapitre II.

Le modèle **ISBA** simule l'évapotranspiration et la réserve hydrique du sol à partir du forçage climatique et de la prescription de l'évolution de l'indice foliaire. Il peut être utilisé pour le suivi de la consommation hydrique des couverts végétaux (Oliosio et al., 2005). A l'origine, ISBA (Noilhan & Planton, 1989) fut développé pour servir de paramétrisation de surface dans des modèles atmosphériques en permettant la simulation des échanges d'énergie, de vapeur d'eau, de quantité de mouvement et de rayonnement entre les terres émergées (en particulier couvertes de végétation) et l'atmosphère. Il se caractérise par :

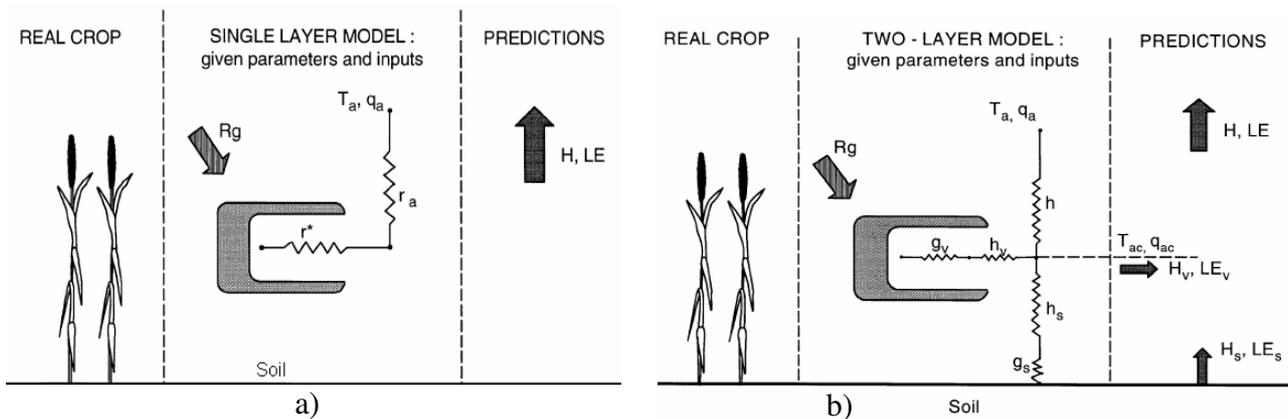
- *i*) un traitement simple des transferts dans le sol constitué de 2 ou 3 couches avec des termes d'échange entre elles,
- *ii*) un traitement mono-couche du bilan d'énergie, mais cependant une séparation de la transpiration des plantes et de l'évaporation du sol, et
- *iii*) une description simple de la conductance stomatique en fonction de différents facteurs climatiques (rayonnement, température, humidité) et de l'humidité du sol dans la couche racinaire.

L'utilisation de **MAGRET** est détaillée dans Courault et al. (1996). Il est dans bien des cas semblable au modèle ISBA, notamment par le forçage météorologique et le forçage de la végétation. Les différences concernent la manière dont l'évapotranspiration et l'humidité du sol sont calculées (ainsi que les transferts de chaleur dans le sol). A l'inverse d'ISBA, le sol est représenté par un seul grand réservoir qui ne distingue pas de couche de surface. L'évaporation et la transpiration ne sont pas séparées. L'évapotranspiration totale du couvert est obtenue en utilisant une résistance de surface regroupant différents facteurs liés à la structure de la végétation, au fonctionnement stomatique, en particulier lié au climat et à l'état hydrique du sol (la résistance stomatique est calculée suivant une méthode proche de celle utilisée dans ISBA).

Dans le modèle **SiSPAT** (Braud et al., 1995), les transferts hydriques et thermiques dans le sol et dans les plantes sont décrits avec beaucoup plus de détails que dans ISBA ou MAGRET. L'hétérogénéité verticale de la structure du sol et de la texture peut être prise en compte, ainsi que la distribution racinaire qui doit être prescrite. Les couplages des transferts de chaleur et d'humidité dans les sols partiellement saturés sont décrits en utilisant l'approche de Philip et De Vries (1957), telle que modifiée par Milly (1982). Les variables 'pronostiques' du sol sont les profils verticaux de température et de potentiel hydrique du sol. Cette approche exige des informations complexes sur les caractéristiques du sol telles que les courbes de rétention et de conductivité hydraulique en fonction de l'humidité du sol (alors que dans ISBA et MAGRET

ces informations étaient résumées dans des coefficients de transferts et surtout une caractérisation de la réserve utile). L'effet de la végétation au-dessus du sol est fondée sur la résolution de deux bilans d'énergie (bi-couche, cf. Fig. I.2b), un pour la surface du sol et un autre pour la couche de végétation. Des calculs simples de transferts radiatifs sont faits à l'intérieur du couvert afin de partitionner l'énergie entre la surface du sol et la couche de végétation (utilisant le coefficient d'atténuation). Les transferts d'eau dans les plantes sont décrits explicitement au moyen d'une série de résistance au travers des plantes.

Les trois modèles TSVA ci-dessus ont fait l'objet d'une étude comparative sur l'expérimentation Alpilles-ReSeDA d'Octobre 1996 à Novembre 1997 près d'Avignon (Oliosio et al., 2002a ; 2002b). Un des résultats principaux de cette étude montre que la complexité des modèles n'est pas forcément gage de la qualité des résultats, un facteur déterminant de cette qualité étant la qualité de l'information disponible pour alimenter le modèle. Les modèles les plus complexes nécessitent en effet la prescription d'un grand nombre de paramètres en entrée qui ne sont pas toujours facile à connaître. D'autre part, dans l'exemple traité, les équations utilisées par les modèles les plus complexes apparaissent moins robustes et plus sensibles aux imprécisions sur les paramètres en entrée.



**Fig. I.2: Schéma de la structure mono-couche (a) et bi-couche (b) des modèles TSVA basés sur un réseau de résistances pour représenter les échanges entre la végétation et l'atmosphère (issu de Oliosio et al., 1999a). ( $R_g$  : rayonnement global,  $H, LE$  : flux de chaleur sensible et latente (indice  $s$  pour sol et indice  $v$  pour végétation),  $T_a, q_a$  : température de l'air et d'humidité,  $r^*$  : résistance de surface,  $r_a$  : résistance aérodynamique,  $g_s$  résistance du sol et  $g_v$  celle de la végétation).**

Un des principaux problèmes rencontrés dans l'utilisation des modèles TSVA pour l'estimation de l'évapotranspiration est lié à la connaissance de l'ensemble des données nécessaires en entrée du modèle. Ceci est d'autant plus problématique lorsque le modèle est appliqué sur des zones ayant une certaine étendue spatiale pour lesquelles il importe d'avoir accès à la variabilité spatiale des entrées du modèle. La combinaison à des données de télédétection, en particulier par des méthodes d'assimilation peut être un moyen pour aborder ce problème.

En effet, les paramètres principaux extraits à partir des mesures de télédétection sont le taux de couverture (*veg*), l'indice foliaire *LAI*, l'albédo, l'émissivité (la plupart d'entre eux sont estimées en utilisant l'information dans le domaine solaire). Par ailleurs, la rugosité et d'autres paramètres liés à la résistance stomatique (l'humidité du sol) sont encore difficiles à estimer. Ils sont souvent estimés à la fois à partir de la connaissance du type de couvert et du stade phénologique. Certes, quelques progrès ont été faits en déterminant la rugosité à partir des données LIDAR (Menenti & Ritchie, 1994), ou par des rapports basés sur l'indice de végétation (Oliosio et al., 2002b) mais ils ne demeurent applicables que dans des conditions particulières. Les procédures d'assimilation peuvent être proposées pour pallier à ce déficit d'information sur certains paramètres. En effet, elles peuvent permettre de ré-estimer des paramètres en cherchant à améliorer la simulation des variables dépendant de ces paramètres en les rapprochant de mesures.

## 2. Assimilation de données

L'assimilation de données est un ensemble de techniques permettant de combiner de façon optimale (en un sens à définir) les propriétés théoriques qui régissent l'évolution d'un système physique donné (réunies sous la forme d'un modèle) et un ensemble d'observations de l'état de ce système afin d'estimer cet état dans l'espace et dans le temps de façon aussi précise que possible (Talagrand & Courtier, 1987 ; Bouttier & Courtier, 1999). Cet état résultant de l'application d'un algorithme d'assimilation de données est appelé état analysé ou analyse.

On utilise classiquement ces techniques pour deux types d'applications : le filtrage et le lissage. Le filtrage permet d'obtenir une estimation optimale de l'état d'un système à partir d'observations provenant du passé relativement à l'analyse. C'est une approche tout à fait appropriée pour les systèmes d'assimilation de données en temps réel. Ainsi, c'est le cas des prévisions météorologiques : après une prévision sur une fenêtre temporelle donnée (typiquement d'une durée de quelques heures), les observations collectées pendant cette période sont combinées avec cette prévision pour calculer un état initial le meilleur possible pour la prévision sur la fenêtre temporelle suivante. Le lissage a lui pour objectif de calculer aussi précisément que possible l'état d'un système sur la totalité d'une fenêtre temporelle donnée : il s'agit d'adapter la solution du modèle de façon globale à toutes les observations collectées sur cette fenêtre temporelle.

Un système d'assimilation de données est généralement composé de trois éléments: un ensemble d'observations, un modèle dynamique et une technique d'assimilation de données (Robinson & Lermusiaux., 2000). Si plusieurs familles de méthodes existent, elles sont souvent classées en deux catégories : les méthodes variationnelles, issues de la théorie du contrôle optimal, et les méthodes stochastiques, basées sur la théorie de l'estimation statistique.

Les méthodes variationnelles (3D-Var, 4D-Var, ...) consistent à minimiser une fonction mesurant l'écart entre l'état du système simulé par le modèle et les observations collectées. Un algorithme de minimisation est utilisé pour ajuster les valeurs d'un vecteur de contrôle, qui peut être composé de paramètres, conditions initiales ou conditions aux limites du modèle, afin de corriger l'état simulé du système en fonction des observations disponibles. Ce type de méthode est utilisé de façon opérationnelle dans la plupart des services météorologiques car particulièrement adapté aux modèles coûteux en temps de calcul et dont la taille du vecteur de

contrôle est grande. Nous décrirons plus précisément ce type de méthode dans la partie 2 de ce mémoire.

Les méthodes stochastiques les plus utilisées sont issues du filtre de Kalman (Kalman, 1960). Le principe de ces méthodes est de corriger la trajectoire des variables d'état du modèle aux temps pour lesquels des observations sont disponibles. Ces méthodes séquentielles sont composées d'une étape de prévision par intégration du modèle à partir de l'état analysé à l'itération précédente du filtre et jusqu'à un temps  $t$  pour lequel un jeu d'observations est disponible, puis d'une étape de correction de l'état simulé à partir de ces observations. Un des intérêts de ce type de méthode par rapport aux méthodes variationnelles est l'estimation directe de la matrice de covariance d'erreur de l'état analysé qui est une mesure de la précision de l'analyse. Toutefois, le filtre de Kalman classique est uniquement optimal pour des modèles linéaires et n'est applicable que pour des modèles dont le coût en temps de calcul n'est pas trop élevé et pour des problèmes de taille relativement réduite. Diverses variantes de ce filtre ont été élaborées pour dépasser ces verrous : on citera par exemple le filtre de Kalman étendu qui propose d'utiliser une approximation linéaire du modèle, le filtre de Kalman d'ensemble (Evensen, 2003) qui permet d'utiliser directement un modèle non linéaire, ou les filtres de Kalman de rang réduit (Pham et al., 1998; Pham, 2001) dont l'objectif est de réduire la taille du problème à résoudre. Dans la famille des méthodes de filtrage on peut également citer les filtres particuliers (Del Moral et al., 1995 ; Doucet et al., 2000) qui sont de plus en plus utilisés et dont la souplesse permet l'utilisation sur des cas où les erreurs sont supposées non gaussiennes. On trouve des applications de ces méthodes dans de nombreux domaines parmi lesquels la chimie atmosphérique, l'océanographie, l'hydrologie, ou plus récemment l'agronomie (Carmichael et al., 2008; Evensen, 2003; McLaughlin, 1995 ; Naud et al., 2007).

L'assimilation de données a d'abord été utilisée en météorologie pour la prévision opérationnelle notamment et ce depuis les années 60. Ce domaine reste l'un des plus actifs concernant la recherche et l'application des techniques d'assimilation de données. Toutefois, elles sont également largement utilisées dans d'autres domaines tels que l'océanographie ou l'hydrologie pour la prévision des courants marins ou des crues par exemple. L'utilisation des procédures d'assimilation pour l'estimation de l'évapotranspiration est plus récente. Les mesures, souvent issues de télédétection, ont longtemps été directement introduites (généralement après interpolation) en entrée des modèles (Bruckler & Witonno, 1989, Olioso et al., 1999a ; 2005), on parle alors de forçage. L'utilisation des méthodes d'assimilation, qui présente un certain nombre d'avantages par rapport au forçage, consiste généralement en

l'ajustement de certains paramètres ou le recalage de certaines variables des modèles (en particulier liés à l'état hydrique du sol ou de la végétation) (Ottlé & Vidal-Madjar, 1994 ; Taconet et al., 1995 ; Olioso et al. 1996b ; François et al., 2001 ; Wigneron et al., 2002; Inoue & Olioso, 2006 ; Muñoz-Sabater et al., 2007). Les applications d'assimilation de données pour estimer l'évapotranspiration s'appuient généralement sur un modèle numérique qui nécessite un bon nombre de forçages atmosphériques, et est aussi relativement exigeant en calcul par rapport aux modèles classiques simples (McLaughlin et al., 2006 ; Kumar et al., 2008). Le choix de la technique d'assimilation de données est alors essentiel pour parvenir à un équilibre entre faire le meilleur usage de toutes les informations disponibles (optimalité) et l'efficacité, la souplesse et la robustesse de la procédure. Toutefois, des compromis doivent être faits pour s'adapter à des objectifs précis, parce que ces critères d'évaluation sont souvent en conflit (Margulis et al., 2002).

Plusieurs études ont tenté d'utiliser les techniques d'assimilation de données associées à un modèle numérique afin d'estimer les flux turbulents de chaleur de surface à l'échelle régionale (Boni et al., 2001; Caparrini et al., 2004; Crow & Kustas, 2005 ; Margulis et al., 2005 ; Pipunic et al., 2008). Boni et al. (2001) ont développé un système d'assimilation de données variationnel pour estimer le flux de chaleur latente et l'évaporation contrôlée de surface utilisant des équations dynamiques pour la température de surface comme contrainte. Dans ce système d'assimilation, des températures de surface observées par satellite sont assimilées dans un modèle de type résiduel permettant de rechercher la fraction évaporative. Caparrini et al. (2003) ont proposé un schéma d'assimilation de données des mesures de la température de surface à une résolution de 1.1 Km, utilisant des séquences de plusieurs satellites de télédétection et des données des stations micrométéorologiques de surface pour estimer les composants du bilan énergétique de surface dans un bassin, avec différentes conditions de surface. Ils ont utilisés la même méthode que Boni et al. (2001). Margulis et Entekhabi (2003) ont examiné l'assimilation de la température de surface radiométrique dans un modèle couplé surface - couche limite atmosphérique utilisant à la fois des données synthétiques et une application unidimensionnelle à l'aide de données de terrain mesurées par un radiomètre. Ils ont utilisé un modèle dynamique local faisant intervenir différentes températures de surface (températures du couvert, du sol et en profondeur du sol). Leurs résultats ont montré que l'assimilation de la température de surface radiométrique a eu un impact positif sur les prévisions du modèle (voir aussi Margulis et al., 2005).

L'approche de l'utilisation d'un modèle TSVA couplé à une procédure d'assimilation de données pour cartographier les flux d'énergie de surface présente des avantages par rapport aux

méthodes semi empiriques (cf. section 1.2.1) ou à l'utilisation de modèles TSVA sans assimilation de données:

- Le système d'assimilation de données permet d'utiliser conjointement au modèle et de façon optimale, des informations réparties spatialement obtenues de nombreuses sources de données avec différentes résolutions, couvertures, et incertitudes (Margulis et al., 2002) ;
- La procédure d'assimilation, en particulier lorsqu'elle vise à estimer des paramètres d'entrées du modèle, permet de corriger non seulement le flux de chaleur latente mais aussi les différentes variables intermédiaires du modèle telles que celles liées aux flux turbulents de chaleur.

Un des inconvénients de l'utilisation des techniques d'assimilation des données pour estimer l'évapotranspiration à l'échelle régionale avec un modèle numérique, c'est qu'elles sont beaucoup plus exigeantes en temps de calcul que les modèles classiques d'estimation de l'évapotranspiration. Par ailleurs, la qualité des résultats d'un système d'assimilation de données, et donc son intérêt, dépend de plusieurs facteurs dont les principaux sont la qualité du modèle, la qualité des observations et le réalisme des hypothèses effectuées dans le système d'assimilation.

### **3. Apport de notre méthode dans l'estimation de l'évapotranspiration**

La connaissance de l'évapotranspiration et des variables caractéristiques associées au microclimat et à l'état hydrique du sol est une étape importante dans la mise en place de diverses connaissances à l'échelle d'un paysage agricole, en particulier pour des études agronomiques portant sur la consommation hydrique des plantes, leur développement et leur production, ainsi que le développement de diverses maladies pouvant affecter les couverts (Courault et al., 2004; Courault et al., 2009a). Elle est également importante pour diverses problématiques météorologiques et hydrologiques.

Il est cependant difficile d'accéder à ces informations aux échelles spatiales pertinentes en raison de la grande variabilité spatiale des quantités recherchées qui empêche leur mesure directe. Les méthodes plus ou moins classiques d'estimation qui combinent données de télédétection et modèles d'estimation sont également confrontées à des problèmes semblables. Comme on l'a vu précédemment si certaines données nécessaires au fonctionnement des modèles d'estimation sont assez faciles d'accès à ces échelles à partir des données de télédétection (température de surface, indice foliaire, fraction de végétation, albédo, émissivité), beaucoup d'autres informations comme les variables météorologiques, la rugosité de surface et l'état du sol ne sont pas directement accessibles. Dans ce travail nous allons chercher à résoudre ce problème en proposant d'une part l'utilisation d'un modèle de couche limite atmosphérique couplé à un modèle de surface adapté à la représentation de la variabilité spatiale ('approche patchée'), et d'autre part la mise en place de procédure d'assimilation permettant de retrouver des variables non directement accessibles.

Le travail auquel nous nous sommes attelés dans cette thèse vise à apporter des réponses au problème de la cartographie de l'évapotranspiration et des variables associées comme le microclimat et l'état hydrique du sol à l'échelle de paysage agricole. Pour cela l'accès à toutes ces variables nécessite un modèle de fonctionnement de type TSVA (processus physiques détaillés). Il faut également résoudre entre autres le problème de la variabilité spatiale du microclimat et des caractéristiques du sol et de la rugosité. En effet, la difficulté rencontrée dans les modèles TSVA est l'initialisation des paramètres d'entrée afin d'estimer des variables précises en sortie des modèles.

La méthode utilisée consiste à estimer les variables non directement accessibles du modèle (rugosité et état hydrique du sol) à partir de diverses sources d'informations (observations de variables de sorties du modèle et information a priori sur les paramètres). Nous utiliserons pour

cela une méthode variationnelle. Nous avons vu que ce concept d'assimilation de données a été introduit dans plusieurs domaines mais l'utilisation de ce type de méthode dans les modèles TSVA n'en n'est qu'à son début (Bouyssel et al., 1999; Lauvernet et al., 2002; Reichle, 2002; Margulis & Entekhabi, 2003 ; Judge et al., 2008). Seules peu d'études ont portées sur l'application de ce type de méthode à des modèles TSVA à grande échelle et beaucoup concernent des applications locales.

Par ailleurs, la mise en place du modèle PBLs et de la procédure d'assimilation de données infrarouges thermiques nous a montré que les connaissances imparfaites sur certains autres paramètres pouvaient avoir un impact important sur l'estimation des flux de surface. Il s'agit du rapport entre le flux conductif dans le sol et le rayonnement net ( $c_0 = G/R_n$ ), ainsi que la paramétrisation de la rugosité thermique  $z_{0h}$  (ou du rapport rugosité thermique – rugosité mécanique  $z_{0h} / z_{0m}$ , ou encore du paramètre aérodynamique  $kB^{-1}$ ) qui sont utilisés dans PBLs et dans bien d'autres modèles de flux de surface. Bien que de nombreuses études aient déjà porté sur ces paramètres, il n'existe pas à l'heure actuelle de consensus sur la manière de les représenter dans les modèles de surface. Dans ce travail, pour améliorer la représentation des flux de surface nous nous sommes donc attachés à préciser le déterminisme de ces paramètres pour en proposer de nouvelles représentations intégrables dans PBLs. Ainsi les procédures d'assimilation de données développées seront appliquées à la version « patchée » de base de PBLs dans laquelle les paramètres  $c_0$  et  $z_{0h} / z_{0m}$  sont considérés constants sur le domaine étudié, puis à une nouvelle version dans laquelle ces paramètres varient en fonction d'autres caractéristiques du couvert comme l'indice foliaire, l'albédo, ou la température de surface, vitesse du vent...).

### **3.1.1. Approche patchée « tiled approach »**

L'évapotranspiration et le microclimat dépendent des interactions qui s'établissent entre la surface terrestre et l'atmosphère. Les modèles couplant le fonctionnement de la surface à celui de la couche limite atmosphérique (CLA) peuvent permettre d'étudier ces interactions. De plus, ces interactions dépendent de la variabilité spatiale des caractéristiques et du fonctionnement de la surface. Ces interactions sont particulièrement sensibles lorsque l'on s'intéresse à l'échelle régionale composée d'une mosaïque de surfaces différentes. Des modèles de transferts sol – plante – atmosphère (TSVA) en 3D permettent de prendre en compte cette

variabilité de surface, comme par exemple certains modèles climatiques à mésoéchelle fine : RAMS (Regional Atmospheric Modeling System, Pielke et al., 1992 ; Mukabana & Pielke, 1996), MESO-NH (Mallet & Payart, 2008 ; Le Moigne, 2008). Une des principales contraintes d'utilisation est le coût élevé du temps de simulation. En effet, plus la maille choisie pour décrire le paysage est fine, plus le pas de temps doit être court pour simuler les phénomènes de turbulence, et seules de petites périodes peuvent être simulées.

Une alternative consiste à utiliser des modèles plus simples monodimensionnels (1D) en version patchée, dans laquelle les différents constituants du paysage sont exprimés suivant leur pourcentage de surface occupée mais dont la distribution spatiale n'est pas prise en compte, et qui calcule l'évolution de la couche limite atmosphérique (CLA) en tenant compte des différentes composantes de la surface (Blyth, 1995). Cette approche patchée consiste donc à introduire dans le modèle les diverses proportions et les caractéristiques des principaux couverts végétaux qui composent le paysage. Par exemple, une zone agricole peut être composée de 3 couverts (3 patchs) : 25% blé, 50% prairie, 25% tournesol avec pour chacune de ces classes ses propres caractéristiques de surface (hauteur du couvert, profondeur racinaire, indice foliaire...). Le modèle calcule alors un bilan d'énergie pour chaque type de végétation et moyenne ensuite les flux transmis à l'atmosphère. Ainsi, l'évolution de la couche limite dépend alors de la proportion et des caractéristiques de ces 3 patchs, mais le modèle ne permet pas de décrire les transferts latéraux entre surfaces. L'utilisation éventuelle de l'approche patchée se justifie cependant par plusieurs études récentes sur les interactions surface - couche limite au moyen de modèles utilisant le concept de Large Eddy Simulations qui montrent que les transferts latéraux à proximité de la surface du sol ont un impact limité sur la variabilité spatiale du microclimat (Albertson et al., 2001; Kustas & Albertson, 2003; Courault et al., 2007; Timmermans et al., 2008).

Dans notre travail, nous avons utilisé le modèle PBL (Planetary Boundary Layer) décrit par Brunet et al. (1991) qui utilise le couplage entre la modélisation des flux de surface suivant une formulation 'Big Leaf' (mono-couche) et une représentation très simplifiée de la couche limite atmosphérique en condition instable. Ce modèle simple permet de simuler les interactions surface - couche limite atmosphérique à l'échelle de la journée (modèle temporel), mais n'avait pas été utilisé jusqu'à présent en tenant compte de la variabilité des caractéristiques de surface pour simuler les flux et les températures. Les travaux que nous avons réalisés montrent que l'introduction des patchs dans le modèle 'PBLs' permet une amélioration des variables

recherchées, en particulier en terme de variabilité spatiale du microclimat (Courault et al., 2008 ; 2009b). L'intérêt de ce modèle est son accessibilité. Simple, il ne demande que peu de données d'entrée et de temps de calcul. Une application menée sur le site de CRAU – CAMARGUE a montré que le modèle était réaliste (Courault et al., 2008). Il permet de tester la mise en place de procédures d'assimilation de données de télédétection (telle que la température de surface) pour mieux simuler les flux de surfaces. Cependant il part d'hypothèses fortes, en particulier sur la forme des profils atmosphériques valables pour des conditions instables, le paramétrage de la rugosité thermique ( $z0_h$ ) ou le rapport  $kB^{-1}$  intervenant dans le calcul du flux sensible  $H$  (Brustsaert, 1982 ; Hasager et al., 2002a ; Massman, 1999a), ainsi que le calcul du flux conductif dans le sol  $G$  estimé à partir du rayonnement net  $R_n$  (Choudhury et al., 1987 ; Norman et al., 1995).

Par conséquent, nous avons analysé plus en détail ces derniers points ( $z0_h$  et  $G/R_n$ ) qui sont assez mal connus de façon théorique, et pour lesquels il n'existe pas de consensus malgré le nombre important d'études qui leur ont été consacrées. Nous avons évalué différents paramétrages de la rugosité thermique et du calcul du flux conductif dans le sol afin d'améliorer ces parties dans PBLs. Ceci nous a conduit à implémenter ensuite des nouvelles paramétrisations dans PBLs adaptées aux données dont nous disposons. Ces différentes modifications ont été intégrées dans une seconde version de PBLs afin d'analyser et d'évaluer l'impact de ces variables dans le processus d'estimation des flux de surface et du microclimat (cf. **chapitre III**).

### **3.1.2. Assimilation variationnelle de données**

L'utilisation du modèle PBLs sur un paysage composé de divers types de surfaces nécessite d'avoir accès à la variabilité spatiale des caractéristiques de ces surfaces. Certaines d'entre elles, telles que l'indice foliaire ( $LAI$ ) et l'albédo ( $\alpha$ ) sont facilement obtenues par le biais de la télédétection (Jacob et al., 2002b; Weiss et al., 2002; Bsaibes et al., 2009). D'autres caractéristiques importantes telles que l'humidité du sol dans la zone racinaire ( $f_2$ ) et la rugosité aérodynamique ( $z0$ ) sont par contre difficiles à estimer directement. Par ailleurs, une étude de sensibilité menée au préalable (Hoedts, 2005 ; Kpemie, 2006 ; voir aussi analyse de sensibilité du chapitre III) a montré que la température de surface ainsi que les flux de chaleurs sont sensibles à ces paramètres. Il est ainsi nécessaire de mettre en place des procédures permettant

malgré tout d'accéder à leurs variabilités spatio-temporelles, et pour résoudre ce problème, nous avons mis en place une méthode d'assimilation basée sur l'introduction de données thermiques (températures de surface) dans le modèle PBLs.

La méthode que nous avons utilisée est une méthode variationnelle (Talangrand & Courtier, 1987 ; Bouttier & Courtier, 1999). Son principe est basé sur la minimisation d'une fonction coût qui mesure l'écart entre les variables simulées par le modèle et les observations disponibles. L'algorithme de minimisation ajuste un vecteur de contrôle composé de variables d'entrées du modèle afin d'obtenir une bonne cohérence entre les variables estimées par le modèle et les observations. L'intérêt d'une telle technique est de pouvoir estimer aussi bien des conditions initiales que des conditions aux limites ou des paramètres du modèle, ou même une combinaison des trois, et d'être adaptée à des problèmes de grande taille et à des modèles coûteux en temps de calcul. Dans notre cas, une application à l'échelle du paysage peut en effet contenir plusieurs dizaines à centaines de milliers de pixels et donc de paramètres à estimer.

Ce grand nombre d'informations à retrouver (rugosité et humidité pour chacun des patches considérés) rend complexe et coûteux le problème de minimisation de la fonction coût et il est donc nécessaire d'utiliser un algorithme de minimisation efficace. On utilise classiquement dans ce cas une méthode de descente qui nécessite le calcul du gradient de cette fonction coût. Dans ce contexte, il est avantageux d'utiliser la méthode adjointe (Courtier & Talagrand, 1990) pour le calcul de ce gradient. L'implémentation de l'adjoint d'un modèle étant une tâche longue et fastidieuse, nous avons opté pour l'utilisation du logiciel de différenciation automatique TAPENADE (Hascoët & Pascual, 2004) pour générer l'adjoint de PBLs. D'autre part, le système d'assimilation a été implémenté avec le logiciel de couplage PALM qui a été développé pour l'assemblage de systèmes d'assimilation de données (Buis et al., 2006).

Le problème posé ici est donc d'estimer les deux paramètres  $f_2$  et  $z_0$  sur chaque pixel, et ce à partir d'une observation de température de surface par pixel. Il s'agit par définition d'un problème sous-déterminé. Plusieurs techniques existent pour régulariser un problème mal posé (Fletcher, 2000; Sun, 1994; Tikhonov, 1963). L'ajout d'informations a priori sur les paramètres à estimer est une solution classiquement utilisée. Les valeurs utilisées comme information a priori sont typiquement celles auxquelles auraient été fixées les paramètres s'ils n'avaient pas été estimés, et peuvent provenir de diverses sources d'information. Un autre point important dans un schéma d'assimilation de données concerne les hypothèses qui sont effectuées concernant les erreurs, et notamment celles des observations, du modèle et de l'information a priori. Le degré de réalisme de ces hypothèses peut fortement impacter la qualité des résultats

obtenus. Nous avons donc choisi d'évaluer trois stratégies dans l'implémentation du schéma d'assimilation :

- estimer qu'un seul paramètre : la rugosité aérodynamique  $z_0$  (nommé algorithme A) ;
- estimer qu'un seul paramètre : l'humidité normalisée du sol  $f_2$  (nommé algorithme B) et ;
- estimer les deux paramètres à la fois (nommé algorithme C).

Nous avons essayé dans cette étude d'illustrer ces différents aspects pour notre problématique dans le cadre d'applications synthétiques et réelles.

## ***Conclusion***

Nous pouvons noter que de nombreuses approches ou différents modèles ont été développés pour étudier le fonctionnement des couverts à l'échelle régionale et estimer l'évapotranspiration. Ils sont presque tous reliés aux données de télédétection afin d'améliorer les estimations. Les approches vont d'une simple relation empirique à des méthodes résiduelles, voire à des approches plus complexes qui décrivent les échanges entre le sol, les plantes et l'atmosphère en fonction de la physique et des processus physiologiques qui se produisent dans chaque compartiment avec en général un pas de temps fin. Des études récentes et d'autres approches de modélisation des échanges de flux turbulents de chaleur au niveau du sol ont fourni de bonnes connaissances sur l'estimation de l'évapotranspiration. Une étude détaillée de ces différentes approches est présentée par Li et al. (2009) montrant leur degré d'application, les avantages et inconvénients.

Certes, quelques progrès ont été faits dans l'acquisition des données pour alimenter les modèles à l'échelle régionale mais d'autres restent encore difficiles à obtenir. A cet égard, des procédures d'assimilation de données sont nécessaires pour déterminer ces variables à partir des observations. Cette méthode est introduite dans le modèle de surface - couche limite atmosphérique développé (le modèle PBLs) dont la description se fera dans les chapitres suivants.

Le modèle développé est décrit afin d'apporter une réponse sur les interactions qui peuvent exister entre les surfaces hétérogènes et l'atmosphère. L'approche patchée qui sera illustrée dans la partie III montrera l'effet induit sur l'évolution de la couche limite atmosphérique. La recherche des paramètres mal connus du point de vue théorique ( $G/R_n$  et  $kB^{-1}$ ) révèle un tout autre fonctionnement du modèle, notamment sur une surface végétale développée ou peu développée. La mise en oeuvre de la méthode d'assimilation dans le modèle développé est une méthodologie souple pour rechercher des valeurs optimales des variables non directement accessibles et d'améliorer la précision sur l'estimation des flux de surface (particulièrement l'évapotranspiration) et le microclimat. Le modèle PBLs implémenté avec la méthode d'assimilation variationnelle est ensuite appliqué sur les expérimentations de CRAU – CAMARGUE et d'ALPILLES. Ce sont tous ces points que nous allons détailler dans les chapitres suivants.



***DEUXIEME PARTIE :***

***METHODES, MATERIELS ET OUTILS***



## 1. Introduction

Dans cette partie, vont être présentés les différents outils, méthodes et données qui vont être utilisés au cours de ce travail. Il s'agit de présenter :

1. le modèle de transfert Sol-Végétation-Atmosphère (PBLs), dans sa nouvelle version, qui réalise le couplage entre le fonctionnement de surface d'une zone hétérogène composée de différents types d'occupation du sol et le fonctionnement de la couche limite atmosphérique
2. la théorie de la méthode d'assimilation de données qui va être utilisée pour appliquer le modèle à une petite zone agricole à partir d'images de télédétection
3. les données qui vont servir de base pour tester la méthode d'une part, et d'autre part développer de nouvelles paramétrisations au sein du modèle PBLs (concernant la rugosité thermique et le rapport  $G/R_n$ )
4. la mise en œuvre de la méthodologie sur des données simulées (expérience jumelle) et sur des données réelles
5. les outils TAPENADE, permettant le développement de la modélisation adjointe nécessaire, et PALM, permettant la mise en œuvre de la méthode d'assimilation.

## 2. Description du modèle PBLs

Le modèle d'interaction surface – couche limite PBLs (Planetary Boundary Layer – surface) décrit par Brunet et al. (1991) utilise le couplage entre la modélisation des flux de surface suivant le modèle 'Big Leaf' et une représentation très simplifiée de la couche limite atmosphérique (*CLA*) en condition instable. Le modèle de surface est basé sur des équations d'approche mono-couche (comme les équations de Penman-Monteith proposé par Monteith, 1965) et le modèle de la couche limite planétaire sur la formulation de Tennekes et Driedonks (1981). Le modèle a été initialement appliqué sur des surfaces homogènes sur une période diurne (Brunet et al., 1991; Lagouarde & Brunet, 1989; 1994). Dans notre étude, il a été adapté pour tenir compte de la variabilité spatiale des surfaces à l'échelle de la région.

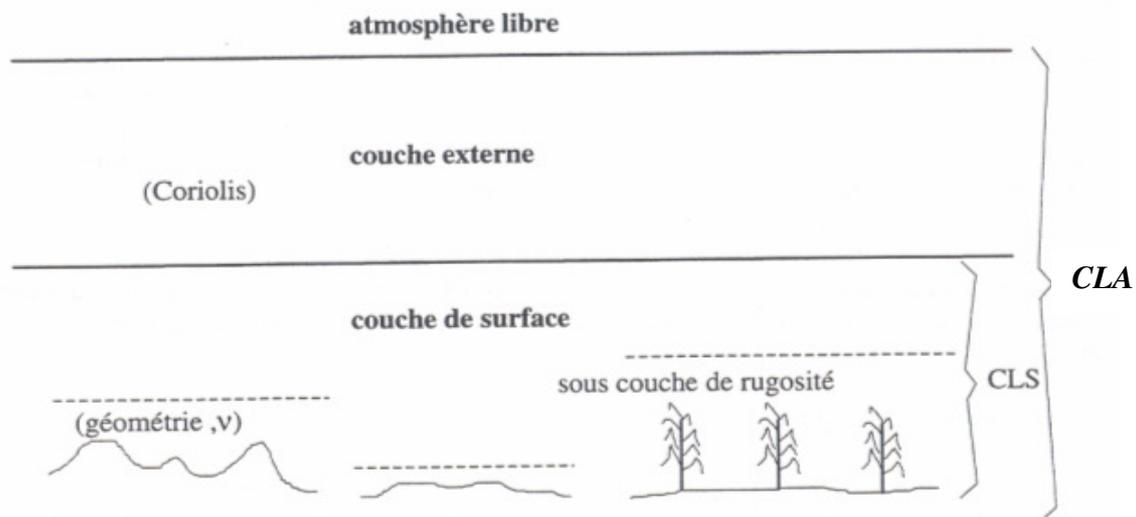
Avant d'entamer la description du modèle, il est important de définir le système auquel nous nous intéressons. Il s'agit de la surface représentée par le couvert végétal (surface du sol et végétation) ainsi que d'une couche d'air particulière située au-dessus et qu'on appelle couche limite atmosphérique (*CLA*). Celle-ci se définit comme étant une partie de l'atmosphère soumise aux influences de la surface et dont l'épaisseur peut, selon l'intensité de ces influences, varier entre 100 et 3000 m. Bien entendu, le système en question est décrit par des variables d'état que sont les températures et l'humidité pour la surface. Ces variables correspondent au vent, la pression, la température et à la masse au sens large qui peut être représentée par différents constituants de l'atmosphère (la vapeur d'eau, le CO<sub>2</sub>, etc.).

La surface exerce divers forçages qui sont fonction de ses propres caractéristiques et qui agissent sur la structure verticale de la *CLA*. Si l'on observe cette structure, il est possible d'affiner sa description en la découpant conceptuellement en deux principales couches (Raupach et al., 1980 ; voir aussi Samaali, 2002). Du bas vers le haut, on peut distinguer (cf. Fig. II.1) :

- la couche limite de surface (*CLS*) : elle correspond à la couche d'air qui est en contact direct avec la surface et qui entretient avec elle des échanges énergétiques et massiques. Son épaisseur peut varier de quelques mètres à quelques dizaines de mètres au-dessus de la surface (Brunet et al., 1995). Ces échanges se produisant donc entre le sol, la végétation et cette couche d'air sont habituellement régis par une grandeur appelée « flux » qui

représente la quantité de la variable considérée (température, par exemple) par unité de temps et par unité de surface selon une direction donnée.

- la couche de mélange (ou couche externe) : la structure verticale de cette couche varie au cours de la journée en fonction du régime convectif. Ainsi, elle est convective en conditions diurnes où les gradients thermiques verticaux sont négatifs. Par ailleurs, en conditions nocturnes, elle peut être subdivisée en deux sous-couches : une couche dite stable (en contact de la couche de surface) du fait des gradients thermiques positifs et une autre dite résiduelle qui n'est autre que le résidu de la couche convective diurne (cf. Fig. II.1).



**Fig. II.1 : Représentation schématique des différentes parties de la CLA (d'après Brutsaert, 1982)**

### 2.1. Le modèle de couche limite

La couche limite de surface (CLS) est la partie la plus simple à traiter de l'atmosphère. Très près de la surface, dans la sous-couche de rugosité, les écoulements sont dépendants de la surface, de la viscosité donc de variables très locales. Nettement plus au-dessus dans l'atmosphère, dans la couche externe (ou de mélange), les écoulements sont influencés par la rotation de la terre et font intervenir la force de Coriolis. La couche de surface se situe entre ces

deux limites. On suppose que cette zone est suffisamment éloignée de la surface pour que les mouvements soient indépendants des caractéristiques locales, mais aussi suffisamment proches pour que l'on puisse négliger les effets de rotation de la terre. Dans cette couche, on considère que tous les flux sont sensiblement constants. Les profils des différentes variables telles que la vitesse verticale du vent ou les concentrations sont de type logarithmique. La structure de cette couche dépend de l'écoulement atmosphérique et des interactions avec la surface.

Le modèle de couche limite atmosphérique utilisé ici est un modèle simplifié initialement décrit dans Tennekes et Driedonks, (1981).

On considère que (Lettau, 1951 ; Kimura & Shimizu, 1994 ; Van De Wiel et al., 2002a, b) :

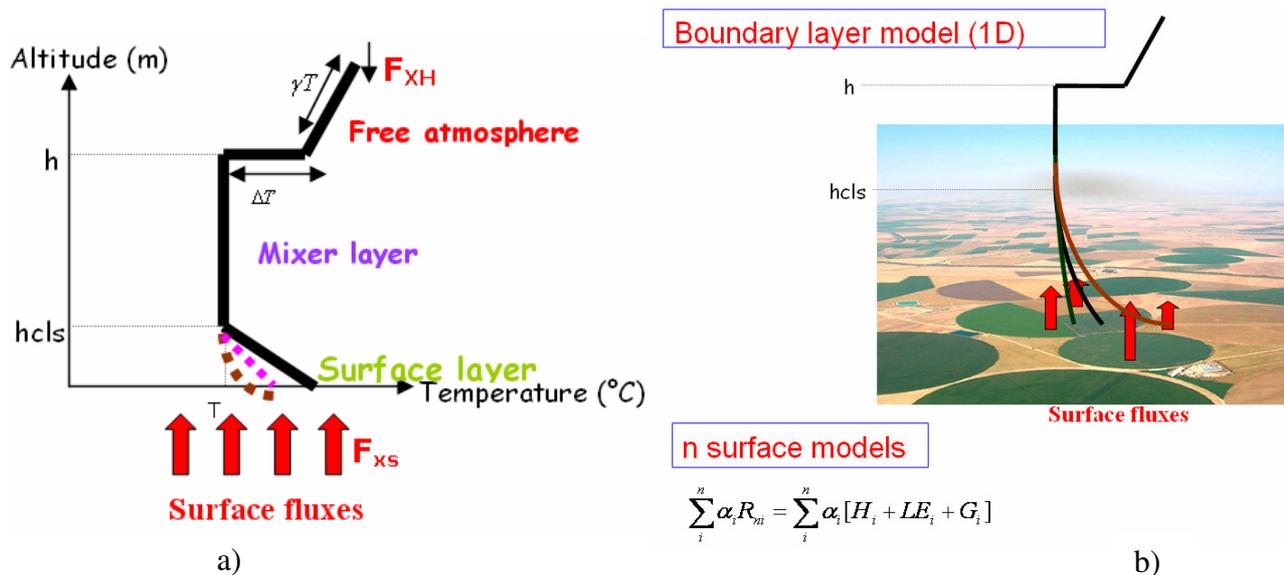
- les turbulences au sein de l'atmosphère sont assez intenses pour que les distributions verticales de vitesses et direction du vent ainsi que les températures potentielles soient uniformes ; cette hypothèse limite donc l'emploi de ce modèle à la journée ; en effet, la nuit, les turbulences sont rarement assez fortes pour maintenir une bonne uniformisation de la couche atmosphérique ;
- toutes les variables atmosphériques sont prises homogènes horizontalement ;
- la couche limite de surface étant supposée fine devant les autres, la capacité calorifique de la surface peut être négligée.

Le modèle de couche limite atmosphérique est basé sur une simple équation d'évolution (II.1) de la couche de mélange (couche centrale dans laquelle les profils de température de l'air ( $T$ ) et d'humidité de l'air ( $q$ ) sont considérés comme constants en fonction de l'altitude). L'épaisseur et l'altitude de cette couche varient en fonction des flux de surface arrivant à la base de cette couche et qui sont fournis par le schéma de surface ( $F_{XS}$ ), et d'une pression venant du dessus de la couche limite planétaire (ici schématisée suivant la pente des profils de  $T$  et  $q$ . cf. Fig. II.2a).

$$\frac{\partial X}{\partial t} = \frac{1}{h} (F_{XS} - F_{XH}) + CF_X \quad (\text{II.1})$$

Avec  $t$  le pas de temps de calcul,  $X$  la variable à évaluer (la température de l'air  $T$  ou l'humidité de l'air  $q$ ),  $h$  la hauteur de la couche de mélange à l'instant considéré.  $CF_X$  est un terme dû à la force de Coriolis qui est supposé négligeable ici. La vitesse du vent est supposée constante au

cours de la journée prise au sommet de la couche limite de surface (CLS).  $F_{XS}$  et  $F_{XH}$  sont les flux verticaux à la base et au sommet de la couche de mélange correspondant aux flux convectifs (flux  $H$  et  $LE$ ) fournis par le modèle de surface suivant la formulation 'Big Leaf' (particulièrement le calcul de  $F_{XS}$ . Le terme  $F_{XH}$  étant calculé tel que  $F_{XH} = 0.2F_{XS}$  (Brunet et al., 1991))



**Fig. II.2.** (a) Description simplifiée du profil atmosphérique adoptée dans le modèle PBLs (d'après Tennekes et Driedonks, 1981); (b) Schéma d'hypothèse de l'approche patchée «tiled approach» permettant de calculer le terme  $F_{XS}$  : ( $h_{CLS}$ ) la hauteur de la couche de surface; ( $h$ ) la hauteur de la couche de mélange; ( $\gamma$ ) la pente dans la couche atmosphérique; ( $T$ ) la température de début de la couche atmosphérique ; ( $\Delta T$ ) l'écart de température au sommet de la couche de mélange

Le modèle de couche limite atmosphérique nécessite en entrée les valeurs initiales des variables atmosphériques ( $T$  et  $q$ ), l'évolution quotidienne du rayonnement global, et la vitesse du vent. Cette dernière est entrée ici comme une constante représentant la vitesse moyenne journalière du vent. Les profils de température et d'humidité de l'air sont initialisés en utilisant des mesures issues de radiosondages (André, 1980) pour déterminer sept points caractéristiques définissant ces profils : la température de l'air  $T$ , (l'humidité spécifique  $q$ ) au début de l'évolution de la couche atmosphérique ( $h_{CLS}$ ), la pente de la température de l'air  $\gamma T$ ,

(respectivement de l'humidité  $\gamma q$ ), et l'écart de la température de l'air  $\Delta T$ , (respectivement de l'humidité  $\Delta q$ ) au sommet de la couche de mélange (cf. Fig. II.2a).

L'évolution de ces variables est discrétisée au sein du programme de calcul suivant les équations classiques différentielles au second ordre non linéaires exprimant la conservation de l'énergie, dites « équations de conservation ou de transport » (De Moor, 1978 ; Brutsaert, 1982 ; Stull, 1988 ; Jacquemin & Noilhan, 1990 ; Perego, 1998 ; Kim & Patel, 2000).

## 2.2. Le modèle de surface

Le modèle de surface est fondé sur la résolution du bilan d'énergie (I.1) pour chaque type de surface (i) composant la zone agricole modélisée (dans la version originale de Brunet et al., (1991), la surface est considérée comme homogène et un seul bilan d'énergie est résolu). Le couplage au modèle de couche limite est ici réalisé par l'introduction d'une approche patchée 'tiled approach' (Ament & Simmer, 2006; Blyth et al., 1993 ; Blyth, 1995) permettant de calculer les termes  $F_{XS}$  pour la température et l'humidité de l'air en tenant compte simultanément des différentes surfaces de la zone modélisée. Ainsi l'évolution de la couche limite est due à l'action de l'ensemble du domaine modélisé (cf. Fig. II.2a et b). La méthode consiste à moyenniser les flux de surface calculés sur chaque type de surface (Raupach, 1991 ; 1995), pondérés par leur pourcentage d'occupation de la zone agricole ( $\varphi_i$ ) et l'équation du bilan d'énergie global de la surface s'écrit alors:

$$\sum_{i=1}^n \varphi_i R_{ni} = \sum_{i=1}^n \varphi_i (H_i + LE_i + G_i) \quad (\text{II.2})$$

Avec  $n$  le nombre de surfaces composant la zone agricole. Ces surfaces peuvent représenter d'une manière simple les différents types d'occupation du sol présents (ex : blé, forêts, sols nus) ou, dans le cas de l'utilisation d'information de télédétection à leur pleine résolution, l'ensemble des pixels des images. Par convention, chaque terme de flux est positif lorsqu'il est orienté vers la surface du sol. Les termes  $F_{XH}$  et  $F_{XS}$  utilisés comme entrée du modèle de couche limite correspondent à la somme pondérée des  $n$  flux  $H_i$  et des  $n$  flux  $LE_i$ . En retour, les conditions météorologiques utilisées pour le fonctionnement du modèle de surface (le

rayonnement incident, la température de l'air, l'humidité spécifique de l'air, la vitesse moyenne du vent) sont considérées spatialement homogènes à la hauteur de la couche limite de surface ( $h_{CLS}$ ). Entre cette hauteur et la surface, des profils différents sont développés pour chacun des patchs composant la surface.

Les termes de l'équation (II.2) sont exprimés en  $W/m^2$ . Avec

- le rayonnement net ( $R_n$ ) qui regroupe toutes les composantes radiatives ;
- le flux de chaleur dans le sol ( $G$ ) qui rend compte des transferts d'énergie entre le sol et la surface (conduction thermique, flux de chaleur latente dans le sol) ;
- le flux convectif de chaleur sensible ( $H$ ) qui rend compte des échanges de chaleur entre la surface du sol et l'atmosphère lorsqu'elles sont à des températures différentes ;
- le flux convectif de chaleur latente ( $LE$ ) qui rend compte des échanges de vapeur d'eau entre la surface du sol et l'atmosphère lorsqu'elles sont à des concentrations de vapeur d'eau différentes.

La convention de signe est que les flux de chaleur sont positifs lorsqu'ils représentent des pertes par rapport à la surface du sol. La convention inverse est appliquée aux flux radiatifs.

### 2.2.1. Le rayonnement net $R_n$

Deux lois permettent de comprendre les échanges radiatifs à la surface du sol (la loi de la réflexion de la lumière et la loi de Stephan-Boltzmann). Dans le domaine des courtes longueurs d'onde, la surface reçoit un rayonnement d'origine solaire : le rayonnement global incident ( $R_g$ ), dont une partie est réfléchi par la surface correspondant à l'albédo ( $\alpha$ ). Le bilan radiatif dans les courtes longueurs d'onde est donc  $R_g (1 - \alpha)$ .

Dans le domaine des grandes longueurs d'onde, le sol émet un rayonnement en fonction de sa température de surface  $L^\uparrow = \epsilon \sigma T_s^4$  ( $\sigma$  est la constante de Stephan-Boltzmann et vaut  $5.67 \cdot 10^{-8} Wm^{-2}K^{-4}$ ) et reçoit un rayonnement d'origine atmosphérique  $R_a$  (résultant de la contribution de l'ensemble des particules air, molécules d'eau,  $CO_2$ ..., présentes dans l'atmosphère). Une partie de ce rayonnement est réfléchi au niveau de la surface avec un coefficient  $(1 - \epsilon)$ , avec  $\epsilon$  correspondant à l'émissivité (selon la loi de Kirchhoff). Le bilan radiatif dans les grandes longueurs d'onde est donc  $\epsilon(R_a - \sigma T_s^4)$ .

$R_n$  est la résultante de toutes ces composantes (Choisnel, 1977). Pour une surface (i) donnée, on peut donc exprimer le rayonnement net sous la forme :

$$R_{ni} = (1 - \alpha_i)R_g + \varepsilon_i R_a - \varepsilon_i \sigma T_{si}^4 \quad (\text{II.3})$$

### 2.2.2. Le flux conductif dans le sol $G$

Le flux conductif correspond au transfert de chaleur dans le sol par conduction (transferts décrits par la loi de Fourier). Il est positif pendant la journée, et correspond à un réchauffement du sol sous l'effet de l'apport énergétique du rayonnement solaire. Au contraire, la nuit il est orienté de la profondeur vers la surface, donc négatif, et correspond à une restitution de la chaleur accumulée dans le sol pendant le jour.

Le flux de conduction  $G$  dans le sol à une profondeur  $z$  est proportionnel au gradient thermique à cette profondeur :

$$G = -\lambda \frac{\partial T}{\partial z} \quad (\text{II.4})$$

avec  $\lambda$  la conductivité thermique du sol ( $\text{Wm}^{-1} \text{K}^{-1}$ ).

Cependant, des expériences ont montré qu'il est aussi possible d'estimer  $G$  directement à partir du rayonnement net  $R_n$  (Choudhury et al., 1987 ; Norman et al., 1995). C'est cette formulation qui a été choisie dans PBLs. Ainsi pour une surface (i) donnée, on a :

$$G_i = c_{0i} R_{ni} \quad (\text{II.5})$$

$c_{0i}$  étant une constante qui peut varier suivant le type de surface étudié. Ce coefficient varie de 10 à 40% suivant les surfaces (couvert dense ou sol nu), par exemple pour une surface agricole de blé,  $c_0$  est fixé à 15% (Courault et al., 1996).

Différentes études ont exprimé les variations de  $c_{0i}$  en fonction du développement du couvert (Kustas et al., 1993 ; Su, 2002), mais aussi en fonction de l'heure et des propriétés de surface

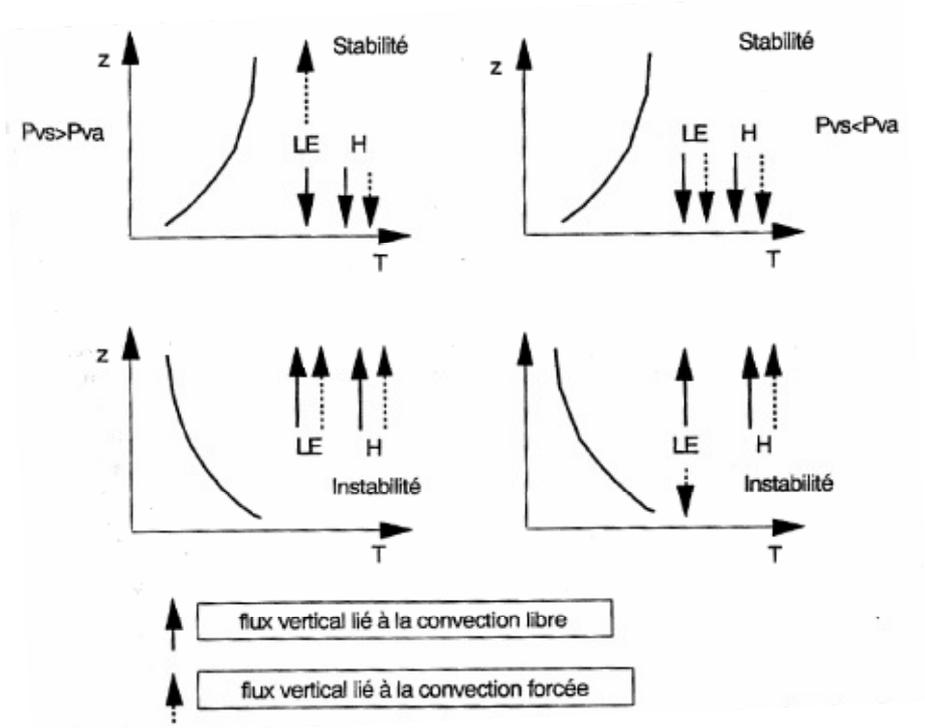
du sol (Cellier et al., 1996; Santanello & Friedl, 2003; Muray & Verhoef, 2007). Dans le cadre de notre travail nous avons analysé les variations de cette relation et décrit des approches nouvelles afin de mieux prendre en compte l'estimation du flux  $G$  dans le sol. Ces résultats sont présentés dans le **chapitre III** et intégrés dans une nouvelle version du modèle PBLs utilisée en **chapitre V**.

### **2.2.3. Les flux convectifs de chaleur sensible $H$ et de chaleur latente $LE$ (et les résistances associées)**

Les échanges de chaleur dans l'air trouvent leur origine dans la nature turbulente de l'écoulement à la surface terrestre. Les mouvements de l'atmosphère qui en résultent induisent des transports de quantités diverses: vapeur d'eau, chaleur sensible, gaz. Plusieurs régimes de convection coexistent naturellement: la convection forcée correspond aux mouvements turbulents imposés par les seuls effets dynamiques (vent), tandis que la convection libre correspond aux mouvements verticaux de l'air générés par les différences de densité liées aux différences de température. Ces déplacements de masses d'air se font de manière plus ou moins régulière et ordonnée, on parle de mouvements turbulents.

Nous nous intéressons ici au cas général de la convection mixte combinant les deux processus permettant de distinguer deux cas qui dépendent du gradient de température entre le sol et la surface : la stabilité et l'instabilité thermique (Fig. II.3).

- La stabilité est obtenue lorsque la température du sol est inférieure à celle de l'air. Dans ce cas, la convection libre s'oppose aux flux ascendants. Ceci se traduit par une stabilisation des masses d'air. Cette situation est obtenue principalement la nuit.
- L'instabilité est obtenue lorsque la température du sol est supérieure à celle de l'air. Dans ce cas, la convection libre accentue les flux ascendants. Cette situation correspond la plupart du temps à des conditions diurnes.



**Fig. II.3. Sens des flux verticaux de chaleur sensible et latente dans les cas de stabilité ou d'instabilité thermique** (issu de Chanzy, 1991).  $P_{Va}$  et  $P_{Vs}$  les pressions partielles de vapeur d'eau de l'air et à la surface du sol (Pa).

Dans la modélisation des échanges convectifs, il est donc nécessaire de prendre en compte à la fois en compte les deux types de convection, libre et forcée. La convection forcée fait intervenir la vitesse du vent et la rugosité de surface qui agit sur les turbulences. La convection libre dépend quant à elle de la stratification thermique de l'air. La modélisation de  $H$  et  $LE$  passe donc par une caractérisation des convections libre et forcée.

Ainsi le flux de chaleur sensible  $H$  se met sous la forme 
$$H = - \rho c_p K_H \frac{\partial T}{\partial z} \tag{II.6}$$

où  $K_H$  est le coefficient de transfert turbulent pour la chaleur sensible,  $T$  la température de l'air,  $\rho$  la masse spécifique de l'air, et  $c_p$  la chaleur spécifique de l'air.

De son coté, le flux de chaleur latente s'écrit 
$$LE = - \rho L K_E \frac{\partial q}{\partial z} \tag{II.7}$$

où  $K_E$  est le coefficient de transfert turbulent de masse,  $q$  désigne l'humidité spécifique de l'air,  $L$  la chaleur latente de vaporisation de l'eau, fonction de la température (en  $\text{J Kg}^{-1}$ ). On fait généralement l'hypothèse  $K_H = K_E$ .

Si on exprime les échanges convectifs entre un niveau de référence  $z$  ( $h_{CLS}$ ) dans la couche limite et la surface pour une surface (i) donnée, on a pour le flux de chaleur sensible l'expression suivante:

$$H_i = \rho c_p \left( \frac{T_{si} - T}{r_{ai}} \right) \quad (\text{II.8})$$

avec  $r_{ai}$  la résistance aérodynamique qui est calculée pour chaque surface en fonction de la vitesse moyenne du vent  $u_a$  au cours de la journée, de la stabilité atmosphérique et des rugosités aérodynamique et thermique. Le calcul des résistances ( $r_{ai}$ ) est basé sur la théorie de Monin-Obukhov au dessus du couvert (Monin & Obukhov, 1954) permettant de corriger les effets de la stabilité ou de l'instabilité de l'air dans la couche.

$$r_a = \frac{1}{k^2 u_a} \left( \ln \left( \frac{h_{CLS} - d}{z0} \right) - \Psi_m \left( \frac{h_{CLS} - d}{L_{MO}} \right) \right) \left( \ln \left( \frac{h_{CLS} - d}{z0_h} \right) - \Psi_h \left( \frac{h_{CLS} - d}{L_{MO}} \right) \right) \quad (\text{II.9})$$

avec  $k$  la constante de Von Karman fixée à 0.4,  $d$  la hauteur de déplacement,  $\Psi_m$  et  $\Psi_h$  les fonctions de correction de stabilité mécanique et thermique pour le transfert de quantité de mouvement et de chaleur.

Les valeurs caractéristiques aérodynamiques de la surface dépendent de la hauteur et aussi de la structure de la végétation. Pour des couverts continus de hauteur  $h_c$  limitée, on peut approximer la rugosité mécanique telle que  $z0 = 0.13h_c$  et  $d = 0.66h_c$  (Brutsaert, 1982 ; Tanner & Pelton, 1960 ; Zhao et al., 2002).

On choisit aussi classiquement le rapport de la rugosité thermique ( $z0_h$ ) et de la rugosité mécanique ( $z0$  ou  $z0_m$ ) comme étant  $z0_h = 0.1 z0$  (Brutsaert, 1982 ; Hasager et al., 2002b).

Ces dernières variables sont caractérisées à l'aide du paramètre  $kB^{-1}$  comme étant  $\ln(z0/z0_h)$  (Brutsaert, 1982; Troufleau et al., 1997 ; Massmann, 1999a), qui varie sensiblement selon les caractéristiques observées des surfaces: fine ou moyenne (herbe, soja , blé) ou des couverts

bien développés avec de grandes valeurs de hauteur de la végétation (vergers, forêts). Ce paramètre a fait l'objet d'une étude spécifique au cours de notre travail qui sera présentée dans le **chapitre III**. Les résultats ont été intégrés dans une nouvelle version du modèle PBLs utilisée au **chapitre V**.

$L_{MO}$  la longueur Monin-Obukhov permet de caractériser la stabilité et qui permet d'introduire la variable adimensionnelle ( $h_{CLS}/L_{MO}$ ), et s'écrit (Monin & Obukhov, 1954):

$$L_{MO} = - \frac{\rho c_p T u_*^3}{kgH} \quad (\text{II.10})$$

avec  $g$  l'accélération de la pesanteur ( $9.81 \text{ m.s}^{-2}$ ),  $H$  le flux sensible donnée par l'équation (II.8),  $k$  la constante de Von Karman (0.4), et  $u^*$  la vitesse de frottement définie par :

$$u^* = \frac{ku_a}{\ln\left(\frac{h_{CLS}-d}{z_0}\right) - \Psi_m\left(\frac{h_{CLS}-d}{L_{MO}}\right)} \quad (\text{II.11})$$

L'intégration des profils de vitesse conduit à deux fonctions de similitude  $\Psi_m$  et  $\Psi_h$  explicitées par Paulson (1970), paramétrées par le régime des échanges convectifs :

- **En régime instable** : c'est-à-dire  $H < 0$  ;  $T < T_s$  ;  $h_{CLS}/L_{MO} < 0$  ; avec  $X = (1 - 16 \frac{h_{CLS}}{L_{MO}})^{\frac{1}{4}}$

$$\Psi_m = \ln\left(\left(\frac{1+X^2}{2}\right)\left(\frac{1+X}{2}\right)^2\right) - 2 \arctan(X) + \frac{\pi}{2} \quad (\text{II.12})$$

$$\Psi_h = 2 \ln\left(\frac{1+X^2}{2}\right) \quad (\text{II.13})$$

- **En régime neutre à stable** :  $H$  légèrement positif ou nul ;  $T$  légèrement supérieur à  $T_s$  ;  $0 < h_{CLS}/L_{MO} < 1$  (Webb, 1970; Businger et al., 1971)

$$\Psi_m = \Psi_h = -5 \frac{h_{CLS}}{L_{MO}} \quad (\text{II.14})$$

- **En régime très stable** : c'est-à-dire  $H > 0$  ;  $T < T_s$  ;  $h_{CLS}/L_{MO} > 1$

$$\Psi_m = \Psi_h = -5 \left( 1 + \ln \left( \frac{h_{CLS}}{L_{MO}} \right) \right) \quad (\text{II.15})$$

Le flux de chaleur latente pour une surface (i) donnée est de la forme :

$$LE_i = \rho c_p L \frac{q_s(T_{si}) - q_a}{r_{si} + r_{ai}} \quad (\text{II.16})$$

avec  $L$  la chaleur latente de vaporisation,  $q_a$  l'humidité spécifique de l'air,  $q_s(T_{si})$  l'humidité spécifique saturante à la température de surface  $T_{si}$ , et  $r_{si}$  la résistance stomatique de la surface au transfert d'eau.

La résistance de surface ( $r_{si}$ ) est calculée pour chaque surface avec la paramétrisation utilisée dans le modèle ISBA décrit par Jacquemin et Noilhan (1990) et qui dépend à la fois des facteurs atmosphériques (rayonnement solaire incident, température de l'air et l'humidité), de la végétation ( $LAI$ ) et de l'eau disponible dans le sol (Jacquemin et Noilhan, 1990; Courault et al., 1996 ; 2008) :

$$r_{si} = \frac{r_{s \min i}}{f_{1i} f_{2i} f_{3i} f_{4i} LAI_i} \quad (\text{II.17})$$

Où  $r_{s \min i}$  désigne la valeur minimum de la résistance en l'absence de stress hydrique.

- $LAI_i$  : indice foliaire lié à la structure de la végétation
- $f_{1i}$  : la dépendance de la résistance au rayonnement solaire
- $f_{2i}$  : la dépendance à l'état hydrique de la surface
- $f_{3i}$  : la dépendance au déficit de saturation
- $f_{4i}$  : la dépendance à la température de l'air

Ainsi,  $r_{si}$  dépend de l'inverse de l'indice foliaire  $LAI$  et de conductances partielles notées  $f_{1i}$ ,  $f_{2i}$ ,  $f_{3i}$  et  $f_4$  dont les valeurs varient de 0 à 1.

Le facteur  $f_{1i}$  représente l'influence du rayonnement photosynthétique actif, supposé correspondre à 55% du rayonnement solaire, sur la résistance stomatique des plantes. Il est calculé selon Dickinson (1984) par :

$$f_{1i} = \frac{(r_{s \min i} / r_{s \max i}) + f_i}{1 + f_i} \quad (\text{II.18})$$

avec  $r_{s \max i}$  la valeur maximum de la résistance de surface ( $5000 \text{ sm}^{-1}$ ), et  $f_i$  un terme adimensionnel correspondant au rayonnement actif entrant dans le couvert, normalisé par une valeur seuil  $R_{gl}$  égale à  $30 \text{ W/m}^2$  pour les arbres et  $100 \text{ W/m}^2$  pour les cultures:

$$f_i = 0.55 \frac{2R_g}{R_{gl} LAI_i} \quad (\text{II.19})$$

En ce qui concerne le facteur du déficit de pression de vapeur  $f_{3i}$ , certains chercheurs caractérisent son apport dans le calcul de la résistance stomatique comme linéaire (Federer, 1979; Choudhury, 1983), tandis que d'autres le considèrent comme une fonction hyperbolique (Jarvis, 1976; Avissar et al., 1985). C'est la seconde approche qui est adoptée dans le modèle PBLs.

$$f_{3i} = 1 - \text{alfa} (q_s(T_{si}) - q_a) \quad (\text{II.20})$$

avec une valeur minimale  $f_{3 \min i} = 0.25$  et alfa une constante qui vaut 0.06.

La relation de l'influence de la température  $f_4$  est donnée par Dickinson (1984) avec une température optimale  $T_0$  fixée à  $298 \text{ K}$  ( $25^\circ \text{ C}$ ). Ainsi,  $f_4$  décroît quand la température de l'air s'éloigne de  $T_0$ .

$$f_4 = 1 - 1.6 \cdot 10^{-3} (T_0 - T)^2 \quad (\text{II.21})$$

Le terme  $f_{2i}$  est lié directement au contenu en eau du sol. Il varie de 0 quand le sol est sec (en dessous du point de flétrissement,  $f_{2i} = 0$ ) à 1 pour les sols humides (au dessus de la capacité au champ,  $f_{2i} = 1$ ). Nous nous intéresserons en particulier à ce facteur qui représente l'humidité du sol dans la zone racinaire dans le modèle PBLs. Noilhan et Planton (1989) ont défini ce facteur  $f_{2i}$  par l'expression suivante :

$$f_{2i} = \begin{cases} 1 & w_{Gi} > w_{fci} \\ \frac{w_{Gi} - w_{wilti}}{w_{fci} - w_{wilti}} & w_{wilti} \leq w_{Gi} \leq w_{fci} \\ 0 & w_{Gi} < w_{wilti} \end{cases} \quad (\text{II.22})$$

avec  $w_{Gi}$  l'humidité du sol dans la zone racinaire,  $w_{fci}$  la capacité au champ,  $w_{wilti}$  le point de flétrissement. Dans la suite de ce travail, le terme  $f_2$  sera aussi nommé 'humidité normalisée'. L'humidité du sol normalisée ( $f_{2i}$ ) est utilisée directement en entrée du modèle PBLs (et non  $w_{Gi}$ ). La construction des cartes de  $f_{2i}$  nécessite la connaissance de l'humidité du sol dans la zone racinaire ( $w_{Gi}$ ), du point de flétrissement, la capacité au champ et la profondeur racinaire (cf. chapitre II.4.1). D'autres formulations autres que l'équation (II.22) sont proposées pour déterminer l'humidité normalisée (cf. chapitre II.5.4 et Courault et al., 2008).

### 2.3. Utilisation du modèle PBLs

Le modèle de surface est piloté par les équations caractéristiques (II.3, II.5 ; II.8 ; II.16). Le couplage avec la CLA se fait à la hauteur de la couche limite de surface ( $h_{CLS}$ ) sous l'effet des termes  $F_{XS}$  et  $F_{XH}$  fournis par le modèle de surface. Dans la couche limite de surface, des profils ( $T$ ,  $q$ ) différents sont développés pour chacun des patches composant la surface. Le microclimat ( $T_a$  à 2 m) est déterminé à partir d'inversion des équations (II.8) et (II.11) en intégrant une hauteur de référence à 2 m dans les expressions. Quant à la température de surface, elle est déterminée par un encadrement successif en lien avec la résolution du bilan d'énergie (II.2).

Les paramètres de surface nécessaires pour l'entrée du modèle PBLs (voir Tab. II.1 ci-dessous) peuvent être obtenus à partir des classes de surface et de type de sol présents sur la

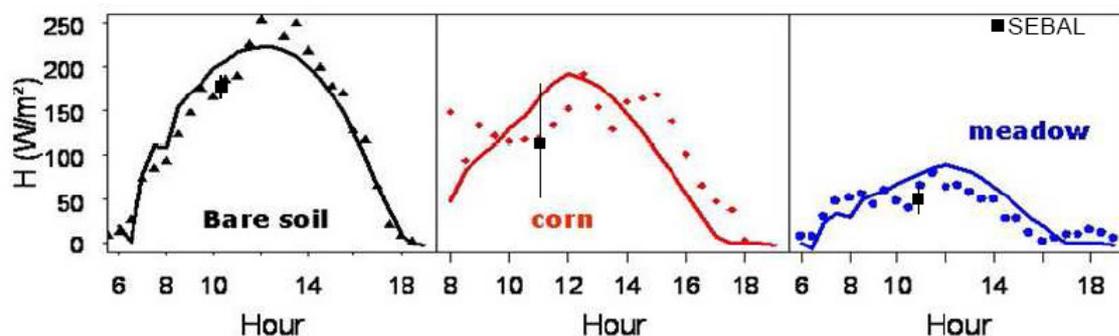
zone simulée. Lorsque les données de télédétection sont disponibles, il est possible d'utiliser le modèle en appliquant l'équation (II.2) du modèle de surface à l'ensemble de tous les pixels de l'image. Certains paramètres peuvent être directement issus de la télédétection comme l'albédo, l'émissivité et l'indice foliaire (*LAI*). Dans ce travail, nous supposons que les autres paramètres qui sont difficiles à obtenir directement à partir d'observations comme le facteur de l'humidité du sol normalisée  $f_2$  et la rugosité aérodynamique  $z_0$ , et qui ont des impacts importants sur la température de surface et les flux de surface doivent être calibrés à partir de données de télédétection de l'infrarouge thermique sur chaque pixel, en utilisant une approche d'assimilation variationnelle.

**Tab. II.1: Les entrées et sorties du modèle PBLs**

ENTREES	SORTIES
<p><b>Modèle de couche limite</b></p> <p>- <i>Propriétés atmosphériques initiales:</i>  <math>h_{CLS}</math>: hauteur de CLS: environ 250 m à 6h  <math>T</math>: température de l'air  <math>q</math>: humidité de l'air  <math>t</math>: pas de temps (20 mn par exemple)  <math>\Delta T, \Delta q, \gamma T, \gamma q</math>: pente et déplacement. de la <math>T</math> dans la CLA</p>	<p>(profil temporel)</p> <p><math>h, T, q</math> à chaque pas de temps sur la journée (entre 5h – 20h par exemple)</p>
<p><b>Modèle de surface</b></p> <p><math>z_0</math>: rugosité aérodynamique  <math>f_2</math>: humidité normalisée  <math>LAI</math>: Indice foliaire  <math>\alpha_i</math>: albédo  <math>\epsilon_i</math>: émissivité fonction de <math>LAI</math> (Francois et al., 1997)</p> <p>- <i>Paramètres de résistance de surface <math>r_s</math>:</i>  <math>R_{smin}</math>, <math>R_{smax}</math>, <math>\alpha</math>, <math>rgl</math> (Jacquemin &amp; Noihlan, 1990)  <math>c_{0j}</math>: ratio <math>G/R_n</math> fixé ici à 0.15</p>	<p>flux de surface à chaque pas de temps</p> <p>Et par pixel (de 5h à 20h par exemple)</p> <p><math>R_n</math>: rayonnement net (<math>Wm^{-2}</math>)  <math>H_i</math>: flux de chaleur sensible (<math>Wm^{-2}</math>)  <math>LE_i</math>: flux de chaleur latente (<math>Wm^{-2}</math>)  <math>G_i</math>: flux conductif dans le sol (<math>Wm^{-2}</math>)</p> <p><math>T_{sj}</math>: température de surface  <math>T_{aj}</math>: température de l'air à 2 m  <math>q_a</math>: pression de vapeur de l'air</p>
<p><b>Forçage Climatique</b></p> <p><math>u_a</math>: vitesse moyenne journalière du vent (constante/jour)  <math>R_g</math>: rayonnement global à chaque pas de temps (20 mn ici)</p>	

Le modèle PBLs permet de simuler pour chaque pixel les flux de surface et températures à l'échelle de la journée (e.g. entre 5h et 20h) avec un pas de temps de 20 mn dans ce travail (pas de temps correspondant au pas de temps d'acquisition des mesures de  $R_g$ ). Le modèle est très rapide en exécution mais en revanche, l'identification des valeurs initiales des points caractéristiques des profils de l'atmosphère est parfois délicate car les profils ne suivent pas tous la forme standard décrite par la figure II.2a. Le modèle PBLs a déjà été appliqué par

Brunet et al. (1991) dans sa version initiale (surface homogène) sur des cultures d'arachide irriguées au Sénégal, de maïs dans le Sud-Ouest de la France et une steppe aride du Sud-Est de la France, et par Lagouarde et Brunet (1994) sur une forêt de pin maritime près de Bordeaux (France), montrant une description adéquate de l'évolution des flux d'énergie de la surface. De notre côté, nous avons également utilisé PBLs en mode direct avec son approche patchée pour la cartographie du microclimat sur la zone Crau - Camargue, dans la basse vallée du Rhône (Courault et al. 2008). Dans cette étude, le modèle a été utilisé pour simuler non seulement le bilan d'énergie de surface et la température de surface, mais aussi la température et l'humidité de l'air à 2 m tout au long de la journée (ce travail est présenté en **annexe I**). Nous avons montré que PBLs (en mode direct) pouvait avoir des performances assez proches de celles de SEBAL (du moins sur les parcelles de validation). La figure II.4 ci-dessous montre le résultat obtenu sur le flux de sensible  $H$  lors de cette expérimentation. Nous avons confronté l'évolution du flux  $H$  au cours d'une journée sur différentes parcelles par rapport aux mesures. On observe également le résultat fourni par SEBAL autour du midi solaire.



**Fig. II.4. Comparaison entre le flux de chaleur sensible  $H$  estimé par PBLs (trait gras) et les mesures terrains (points) acquises avec un anémomètre sur les principales cultures de la zone CRAU – CAMARGUE. (Estimations SEBAL : carré noir)**

### **3. Assimilation variationnelle de données : théorie et application au modèle PBLs dans notre cas d'étude**

#### **3.1. Théorie**

Les méthodes d'assimilation variationnelle de données de type temporel (4D-Var) visent à adapter les solutions du modèle de façon globale à toutes les observations disponibles à l'échelle spatio-temporelle du domaine considéré (Talagrand, 1997). Elles sont basées sur la minimisation de l'écart entre les sorties du modèle et les observations en faisant varier les différentes valeurs d'un vecteur de contrôle composé d'un ensemble de paramètres d'entrée, conditions initiales et/ou aux limites du modèle sélectionné. Cet écart est exprimé par le calcul d'une fonction scalaire, souvent appelé la fonction de coût ou fonction objectif, qui est généralement une somme des carrés des écarts entre les observations et les valeurs du modèle correspondant aux observations : il s'agit d'un problème de moindres carrés non linéaire. Dans la plupart des cas, ce problème est sous-déterminé car les données sont spatialement ou temporellement rares et indirectement liées au vecteur de contrôle. Dans le but d'en faire un problème bien posé (Fletcher, 2000; Sun, 1994), il est nécessaire de le régulariser. Cela est fait classiquement par l'utilisation d'informations a priori (parfois également appelé ébauche) sur le vecteur de contrôle. Un terme d'écart entre le vecteur estimé et son information a priori est alors ajouté à la fonction coût.

Les informations utilisées dans un système d'assimilation variationnel sont donc les observations, l'information a priori et les propriétés physiques connues du système implémentées sous la forme d'un modèle. Toutes ces sources d'informations sont importantes et doivent être prises en compte pour obtenir une bonne analyse. Par ailleurs, toutes ces sources d'information sont entachées d'erreur et il n'est pas possible de leur faire entièrement confiance. Il faut donc réussir un compromis. Pour cela, une représentation probabiliste de ces erreurs est utilisée. En pratique, ces erreurs sont souvent supposées gaussiennes et centrées. Leurs distributions sont alors caractérisées par des matrices de covariance. Plus l'erreur sur une information utilisée dans le système d'assimilation sera supposée importante, c'est-à-dire plus sa variance d'erreur sera forte, moins cette information aura de poids pour le calcul de l'analyse.

### 3.1.1. La fonction coût et son gradient

Le calcul de l'analyse d'un système 4D-Var sur une fenêtre temporelle  $[t_0, t_n]$  (nous ne garderons que les indices 0, n par la suite pour simplifier les équations) est ainsi un problème d'optimisation défini par une fonction coût  $J(x)$  comme suit (Bouttier & Courtier, 1999):

$$J(x) = \frac{1}{2} (x - x^b)^T B^{-1} (x - x^b) + \frac{1}{2} \sum_{i=0}^n (y_i^o - M_{0 \rightarrow i} [x])^T R_i^{-1} (y_i^o - M_{0 \rightarrow i} [x]) \quad (\text{II.23})$$

Où:

- $x$  est le vecteur des paramètres/variables à contrôler ;
- $x^b$  est l'information a priori sur les paramètres/variables à contrôler ;
- $B$  est la matrice de covariance d'erreur de l'information a priori ;
- $y_i^o$  est le vecteur d'observations au temps  $i$  sur tout le domaine considéré ;
- $M_{0 \rightarrow i}$  est l'opérateur qui permet de passer de l'espace du vecteur de contrôle à l'espace des observations au temps  $i$ , c'est en pratique l'opérateur composé du modèle dynamique et d'un opérateur, souvent appelé opérateur d'observation, qui transforme les sorties du modèle en équivalent observé ;
- $R_i$  est la matrice de covariance d'erreur des observations au temps  $i$

L'équation (II.23) peut être décomposée en deux termes : un terme d'ébauche  $J_B$  et un terme d'observation  $J_0$  définis par :

$$J(x) = J_B(x) + J_0(x) \quad (\text{II.24})$$

$$\text{où } J_B(x) = \frac{1}{2} (x - x^b)^T B^{-1} (x - x^b) \quad (\text{II.25})$$

$$\text{et } J_0(x) = \frac{1}{2} \sum_{i=0}^n (y_i^o - M_{0 \rightarrow i} [x])^T R_i^{-1} (y_i^o - M_{0 \rightarrow i} [x]) \quad (\text{II.26})$$

Un grand nombre de méthodes peuvent être envisagées pour minimiser la fonctionnelle  $J$ . Pour les problèmes de grande taille et les modèles coûteux, on préfère de façon générale utiliser des méthodes de descente qui utilisent conjointement la valeur de la fonction et de son gradient pour construire une suite d'estimateurs du vecteur de contrôle qui converge rapidement vers le minimum de la fonctionnelle.

Le gradient de l'équation (II.25) par rapport au vecteur de contrôle  $x$  est donné par :

$$\nabla_x J_B(x) = B^{-1} \cdot (x - x^b) \quad (\text{II.27})$$

Le calcul du gradient du terme d'observation de la fonction coût (II.26) par rapport au vecteur de contrôle  $x$  nécessite la théorie de la différentiation d'une fonction composée.

Soit  $y_i = M_{0 \rightarrow i}(x)$ , et  $J_{0i}(x) = \frac{1}{2} (y_i^o - M_{0 \rightarrow i}[x])^T R_i^{-1} (y_i^o - M_{0 \rightarrow i}[x])$

$$\text{On a, } \nabla_{y_i} J_{0i}(y_i) = -R_i^{-1} \cdot (y_i^o - y_i) \quad (\text{II.28})$$

Or on sait que les dérivées partielles de  $J_{0i}$  par rapport à  $x$  s'écrivent :

$$\frac{\partial J_{0i}}{\partial x_j} = \sum_{l=1}^k \frac{\partial y_{i,l}}{\partial x_j} * \frac{\partial J_{0i}}{\partial y_{i,l}}, \quad \forall i \in \{1, \dots, p\} \quad (\text{II.29})$$

avec  $p$  la taille du vecteur de contrôle  $x$ , et  $k$  celle du vecteur d'observation au temps  $i$ .

$$\text{Ainsi, on a : } \nabla_x J_{0i}(x) = M_{i \rightarrow 0}^T[x] \cdot \nabla_{y_i} J_{0i}(x), \quad (\text{II.30})$$

avec  $M_{i \rightarrow 0}^T[x]$  la transposée de la matrice Jacobienne de l'opérateur  $M_{0 \rightarrow i}[x]$  par rapport au vecteur de contrôle  $x$ , aussi appelée la matrice adjointe :

$$\mathbf{M}_{i \rightarrow 0}^T = \begin{pmatrix} \frac{\partial y_{i,1}}{\partial x_1} & \dots & \frac{\partial y_{i,k}}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_{i,1}}{\partial x_p} & \dots & \frac{\partial y_{i,k}}{\partial x_p} \end{pmatrix} \quad (\text{II.31})$$

Et le gradient de  $J_0(x)$  par rapport à  $x$  s'écrit alors :

$$\nabla_x J_0(x) = - \sum_{i=0}^n \mathbf{M}_{i \rightarrow 0}^T[x] \cdot R_i^{-1} \cdot (y_i^o - y_i) \quad (\text{II.32})$$

Les équations (II.27) et (II.32) définissent le gradient de la fonction coût du problème d'assimilation.

### 3.1.2. A propos de l'opérateur adjoint

Le calcul de l'adjoint du modèle (et de l'opérateur d'observation) nécessaire pour obtenir la valeur du gradient de la fonction coût est une difficulté pratique pour la mise en place d'un système d'assimilation variationnelle. Ce calcul est soit basé sur une approximation des dérivées en utilisant la méthode des différences finies, soit réalisé par un code de calcul qui implémente l'opérateur adjoint. Dans la pratique, la deuxième solution offre deux avantages : elle permet de calculer les dérivées exactes du modèle, et non pas une approximation, et elle est particulièrement efficace lorsque la taille du vecteur de contrôle est grand (Le Dimet & Talagrand, 1986). Néanmoins, l'implémentation de l'opérateur adjoint d'un modèle est une tâche délicate et souvent cruciale lors de l'implémentation d'un système d'assimilation. En effet, dans de nombreux domaines, comme la météorologie et l'océanographie, les modèles sont souvent complexes et non-linéaires. L'opération d'implémentation de l'adjoint d'un modèle peut être effectuée soit manuellement, soit par des méthodes automatiques (e.g. Giering & Kaminski, 1998). L'écriture manuelle est une opération relativement simple mais longue, fastidieuse et source d'erreurs (Elizondo et al., 2000). Les méthodes automatiques ne sont pas

toujours capables de présumer les hypothèses de linéarisation qui sont faites manuellement et les modèles adjoints obtenus sont parfois coûteux (Xu, 1996). Néanmoins, l'utilisation conjointe d'une méthode automatique et d'une écriture manuelle permet d'obtenir assez rapidement des modèles adjoints efficaces. Dans notre cas, nous avons utilisé le logiciel de différentiation automatique TAPENADE (Hascoët & Pascual, 2004) dont le fonctionnement est présenté dans la section suivante (cf. **partie II section 6**).

### 3.1.3. Préconditionnement de la fonction coût

Le coût de calcul de l'analyse est proportionnel au nombre d'évaluations de la fonction de coût et de son gradient nécessaire pour réaliser la minimisation. A chaque itération, l'algorithme de minimisation propose une valeur du vecteur de contrôle  $x$  pour lequel la valeur de la fonction coût est inférieure ou égale à celle de l'itération précédente. Cela nécessite une ou plusieurs évaluations de la fonction et de son gradient par itération.

Le principal aspect de  $J$  qui affecte les performances des minimiseurs classiques (en supposant que  $J$  est quadratique ou presque) est le conditionnement du problème. Ce conditionnement est défini comme le rapport entre la plus grande et la plus petite valeur propre de la seconde dérivée de  $J$  (noté  $J''$ , la matrice Hessienne). Il permet de comprendre que la qualité de l'analyse est proportionnelle à la convexité de la fonction coût. Moins la fonction coût sera convexe, moins bonne sera l'analyse. Une illustration simple de ce problème est donnée dans (Bouttier & Courtier, 1999). Il est donc très important de formuler le problème de manière à obtenir une fonction très convexe.

Dans le cas général,  $J$  est elliptique, mais il est possible de définir un changement de variable appelée preconditionnement qui permet de diminuer le conditionnement (Gilbert & Lemaréchal, 1989 ; Thépaut & Courtier, 1991; Bouttier & Courtier, 1999).

Pour les méthodes variationnelles, une méthode simple et efficace de preconditionnement est d'utiliser la racine carrée de la matrice de covariance  $B : L = B^{1/2}$  avec  $B = LL^T$ . Ce choix se justifie par l'hypothèse qu'il y a peu d'observations par rapport à la taille du vecteur de contrôle. Par changement de variable du vecteur de contrôle  $x$ , on définit la nouvelle fonction coût telle que :

Le nouveau vecteur de contrôle est  $v = L^{-1} \cdot x$

La fonction coût devient alors  $J(v) = \frac{1}{2} (v - v^b)^T \cdot (v - v^b) + J_0(Lv)$ , (II.33)

Et le gradient :  $\nabla_v J(v) = v + L^T \cdot \nabla_x J_0(Lv)$  (II.34)

### 3.1.4. Algorithme de minimisation

Le problème de minimisation à résoudre ici est un problème de moindres carrés non linéaire sans contrainte ou avec contrainte de bornes (le vecteur de contrôle peut contenir des paramètres ou variables bornées, comme c'est le cas dans notre étude). Lorsque la taille du vecteur de contrôle est importante, une méthode de choix pour résoudre ce problème est la méthode de Quasi-Newton (Gilbert & Lemaréchal, 1989). Cette technique efficace repose sur une estimation de la matrice hessienne du problème et, pour les problèmes de taille importante, des algorithmes (BFGS, CGPLUS) proposent des techniques pour limiter la quantité d'information stockée qui peut dans ce cas être très importante (Liu & Nocedal, 1989). Ce type de méthode est ou a été jusqu'à très récemment largement utilisée dans la plupart des centres météorologiques opérationnels notamment (ECMWF, CNRM).

## 3.2. Application au modèle PBLs

Les observations que nous allons considérer sont les températures de surface instantanées ( $T_s$ ) qui peuvent être acquises soit par une caméra thermique aéroportée à midi (cas des données Alpilles présentées plus loin, cf. **section 4.1**), soit par des satellites type ASTER ou Landsat (cf. Courault et al., 2008), et nous estimerons les paramètres de l'humidité du sol normalisée ( $f_2$ ) et la rugosité aérodynamique ( $z0$ ) (cf. **partie III.1**). En effet, ce sont des variables non directement accessibles du modèle et qui varient beaucoup spatialement. De plus, ces paramètres ont un fort impact sur l'estimation de l'évapotranspiration comme illustré dans la partie III de ce document. Par conséquent, la stratégie de retrouver ces paramètres à partir

des données acquises dans le domaine spectral de l'infrarouge thermique est particulièrement intéressante car ces données sont directement corrélées à l'humidité du sol et la variable acquise (température de surface) est directement liée au fonctionnement physique de l'évapotranspiration.

### 3.2.1. Cas d'un seul pixel

Dans ce cas, on considère une seule zone de culture ou un seul pixel. Par conséquent le vecteur de paramètre à contrôler s'écrit alors  $x = {}^T (f_2, z_0)$ .  $y^o_i$  est la température de surface observée à midi notée  $T_{sobs}$ .  $M_{0 \rightarrow i}[x]$  la température de surface à midi simulée par PBLs notée  $T_{sSim}$ .

L'estimation de 2 paramètres avec une seule observation est manifestement un problème sous-déterminé. Donc, l'ajout de valeurs a priori ( $x_b$ ), est nécessaire pour avoir un problème bien posé. On définit l'information a priori par l'expression :  $x_b = x_t + \varepsilon_b$ , où  $x_t$  est l'information exacte et  $\varepsilon_b$  est l'erreur commise.

Ainsi, on a :  $x_b = {}^T (f_b, z_b) = {}^T (f_t + \varepsilon_{fb}, z_t + \varepsilon_{zb})$ .

Les erreurs  $\varepsilon_{fb}$  et  $\varepsilon_{zb}$  sont supposées gaussiennes et centrées. La matrice de covariance d'erreur de l'information a priori  $B$  s'écrit alors :

$$B = \begin{pmatrix} \text{var} (\varepsilon_{fb}) & \text{cov} (\varepsilon_{zb}, \varepsilon_{fb}) \\ \text{cov} (\varepsilon_{fb}, \varepsilon_{zb}) & \text{var} (\varepsilon_{zb}) \end{pmatrix} \quad (\text{II.35})$$

Les variances de  $\varepsilon_{fb}$  et  $\varepsilon_{zb}$  seront par la suite notées  $\sigma_{fb}^2$  et  $\sigma_{zb}^2$ . Par ailleurs, nous supposons également que les erreurs sur l'information a priori des différents paramètres sont indépendantes. Ces hypothèses ainsi que la méthodologie utilisée pour fixer les valeurs de l'information a priori dans les cas d'études réels seront discutées dans la **section 5** de ce chapitre (cf. chapitre II.5).

Il en résulte que  $\text{cov} (\varepsilon_{fb}, \varepsilon_{zb}) = 0$ , et la matrice  $B$  (II.35) devient alors :

$$B = \begin{pmatrix} \sigma_{fb}^2 & 0 \\ 0 & \sigma_{zb}^2 \end{pmatrix} \quad (\text{II.36})$$

La matrice de covariance d'erreur de l'observation  $R_i$  est ici un scalaire car il n'y a qu'une seule observation dans le temps et dans l'espace. L'erreur sur l'observation, noté  $\varepsilon_0$ , représente à la fois l'erreur de mesure et celle du modèle, et elle est donnée par :  $R_i = \text{cov}(\varepsilon_0, \varepsilon_0) = \sigma^2$ .

Etant données les hypothèses faites ci-dessus et avec l'équation (II.36), le terme d'ébauche de la fonction coût (II.25) devient donc :

$$J_B(x) = \frac{1}{2} \left( \frac{f_2 - f_b}{\sigma_{fb}} \right)^2 + \frac{1}{2} \left( \frac{z_0 - z_b}{\sigma_{zb}} \right)^2 \quad (\text{II.37})$$

Le terme d'observation de la fonction coût (II.26) s'écrit :

$$J_0(x) = \frac{1}{2} \left( \frac{T_{sobs} - T_{sSim}}{\sigma} \right)^2 \quad (\text{II.38})$$

Ainsi la fonction coût du problème  $J(x)$  (II.24) est définie par les équations (II.37) + (II.38).

Les gradients des équations (II.37) et (II.38) par rapport à la variable  $x = {}^T(f_2, z_0)$  sont donnés par :

$$\nabla_x J_B(x) = \left( \frac{f_2 - f_b}{\sigma_{fb}^2}, \frac{z_0 - z_b}{\sigma_{zb}^2} \right)^T \quad (\text{II.39})$$

$$\nabla_x J_0(x) = -M^T[x] \cdot \left( \frac{T_{sobs} - T_{sSim}}{\sigma^2} \right) \quad (\text{II.40})$$

où  $M^T$  est l'opérateur adjoint du modèle PBLs intégré jusqu'à 12h.

Les équations (II.39) et (II.40) définissent le gradient de la fonction coût du problème

d'assimilation  $\nabla J(x)$ .

### 3.2.2. Cas à n pixels

On suppose cette fois que la zone agricole modélisée est composée de plusieurs parcelles de végétations différentes. Dans ce cas, le vecteur de contrôle est composé des paramètres d'entrée  $(f_{2i}, z_{0i})$  sur chacun des n pixels du domaine spatial considéré :

$$x = {}^T(f_{21}, f_{22}, \dots, f_{2n}, z_{01}, z_{02}, \dots, z_{0n}) \quad (\text{II.41})$$

Ces paramètres sont estimés à partir des observations de température de surface  $T_{si}$  à midi sur chaque pixel. Comme dans le cas à un seul pixel, l'ajout de valeurs a priori  $(x_b)$ , est nécessaire pour avoir un problème bien posé :

$$x_b = {}^T(f_{b1}, f_{b2}, \dots, f_{bn}, z_{b1}, z_{b2}, \dots, z_{bn}) \quad (\text{II.42})$$

Concernant les erreurs sur l'information a priori, en plus des hypothèses faites dans le cas à un seul pixel, nous supposons que les erreurs des informations a priori sur les différents pixels sont indépendantes. Cette hypothèse sera discutée pour l'application au cas réel dans la **section 5** de ce chapitre.

$$\text{ainsi, } \text{cov}(\epsilon_{f_{bi}}, \epsilon_{f_{bj}}) = \text{cov}(\epsilon_{z_{bi}}, \epsilon_{z_{bj}}) = 0, \text{ pour tout } i, j \text{ dans } \{1, \dots, n\}. \quad (\text{II.43})$$

Sous ces hypothèses la matrice  $B$  est donc diagonale :

$$B = \begin{pmatrix} \sigma_{fb1}^2 & 0 & 0 & & & \\ 0 & . & 0 & & & \\ 0 & 0 & \sigma_{fbn}^2 & & & \\ & & & \sigma_{zb1}^2 & 0 & 0 \\ & 0 & & 0 & . & 0 \\ & & & 0 & 0 & \sigma_{zbn}^2 \end{pmatrix} \quad (\text{II.44})$$

Dans ce cas à n pixels, la matrice de covariance d'erreur des observations théorique a la forme suivante :

$$R_i = \begin{pmatrix} \text{var}(\varepsilon_1) & \text{cov}(\varepsilon_2, \varepsilon_1) & \cdot & \text{cov}(\varepsilon_n, \varepsilon_1) \\ \text{cov}(\varepsilon_1, \varepsilon_2) & \text{var}(\varepsilon_2) & \cdot & \text{cov}(\varepsilon_n, \varepsilon_2) \\ \cdot & \cdot & \cdot & \cdot \\ \text{cov}(\varepsilon_1, \varepsilon_n) & \text{cov}(\varepsilon_2, \varepsilon_n) & \cdot & \text{var}(\varepsilon_n) \end{pmatrix} \quad (\text{II.45})$$

Dans cette étude, et cela sera également discuté dans la **section 5** de ce chapitre, les erreurs d'observation seront considérées comme indépendantes entre les pixels,

$$\text{Soit } \text{cov}(\varepsilon_i, \varepsilon_j) = 0, \text{ pour tout } i, j \text{ dans } \{1, \dots, n\} \quad (\text{II.46})$$

La matrice de covariance d'erreur d'observation (II.45) considérée dans cette étude sera donc de la forme :

$$R_i = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \cdot & 0 \\ 0 & 0 & \sigma_n^2 \end{pmatrix} \quad (\text{II.47})$$

Avec les équations (II.41), (II.42) et (II.44), le terme d'ébauche de la fonction coût (II.25) pour le cas à n pixels s'écrit donc :

$$J_B(x) = \frac{1}{2} \sum_{i=1}^n \frac{(f_{2i} - f_{bi})^2}{\sigma_{fbi}^2} + \frac{1}{2} \sum_{i=1}^n \frac{(z_{0i} - z_{bi})^2}{\sigma_{zbi}^2} \quad (\text{II.48})$$

Et avec l'équation (II.47), le terme d'observation de la fonction coût (II.26) pour le cas à n pixels est défini par :

$$J_0(x) = \frac{1}{2} \sum_{i=1}^n \frac{(Ts_{obs(i)} - Ts_{Sim(i)})^2}{\sigma_i^2} \quad (\text{II.49})$$

avec  $T_{sobs(i)}$  la température de surface observée sur chaque pixel (i),  $T_{sSim(i)}$  la température de surface chaque pixel (i) calculée par le modèle PBLs.

La somme des équations (II.48) et (II.49) définit la fonction coût du problème  $J(x)$  (II.24) dans le cas à n pixels.

Le gradient de l'équation (II.48) par rapport au vecteur de contrôle  $x$  s'écrit :

$$\nabla_x J_B(x) = \left( \frac{f_{21} - f_{b1}}{\sigma_{fb1}^2}, \dots, \frac{f_{2n} - f_{bn}}{\sigma_{fbn}^2}, \frac{z_{01} - z_{b1}}{\sigma_{zb1}^2}, \dots, \frac{z_{0n} - z_{bn}}{\sigma_{zbn}^2} \right)^T \quad (\text{II.50})$$

et celui du terme d'observation :

$$\nabla_x J_0(x) = \mathbf{M}^T[x] \cdot \nabla_{TsSim} J_0(x) \quad (\text{II.51})$$

$$\text{avec } \nabla_{TsSim} J_0(x) = - \left( \frac{Ts_{obs(1)} - Ts_{Sim(1)}}{\sigma_1^2} \dots \frac{Ts_{obs(n)} - Ts_{Sim(n)}}{\sigma_n^2} \right)^T \quad (\text{II.52})$$

et  $\mathbf{M}^T$  est l'opérateur adjoint du modèle PBLs intégré jusqu'à 12h par rapport au vecteur de contrôle  $x$  (II.41)

Les équations (II.50) et (II.51) définissent le gradient de la fonction coût du problème dans le cas à n pixels.

## 4. Données utilisées

Pour tester la méthode développée, nous nous sommes particulièrement appuyés sur les données qui ont été acquises sur la zone Alpilles au cours du programme Alpilles - ReSeDA (cf. Baret, 2002; Prévot et al., 1998 ; Olioso et al., 2002a ; 2002b). Ces données ont été utilisées dans différentes études portant sur l'estimation des paramètres de surface par télédétection, de l'évapotranspiration et du bilan d'énergie de surface, et pour l'étude de méthode d'assimilation de données (cf. ci dessous). Elles vont être utilisées ici pour tester la mise en œuvre de la méthode d'assimilation,

- d'une part en élaborant un jeu test de données synthétiques pour une évaluation détaillée de la méthode,
- et d'autre part pour effectivement évaluer la méthode sur des mesures réelles de flux de surface.

D'autres jeux de données ont été utilisés pour réaliser les études particulières portant sur l'analyse et la paramétrisation des paramètres  $c_0$  et  $z_{0h}$  (ou  $G/R_n$  et  $kB^{-1}$ ). Il s'agit de jeux de données qui ont été acquis sur le site expérimental 'Flux et Télédétection' (familièrement connu sous le nom de 'site Grue') de l'INRA d'Avignon et de données qui ont été simulées par des modèles de transferts plus complexes que PBLs en particulier sur la description du fonctionnement hydrique et thermique du sol (les modèles TEC (Witono & Bruckler, 1989) et SiSPAT). Tous ces jeux de données sont décrits succinctement dans cette section.

### 4.1. Données Alpilles ReSeDA

#### 4.1.1. Généralités sur le programme Alpilles ReSeDA

L'expérience Alpilles – ReSeDA (<http://www.avignon.inra.fr/reseda>) s'est déroulée sur une petite zone agricole de 4 par 5 km près d'Avignon (Sud-Est de la France) dans la vallée du Rhône (N 43 ° 47 ' et E 4 ° 45'). Elle a couvert l'ensemble de la saison de croissance des cultures d'hiver et d'été entre Octobre 1996 et Novembre 1997. Cette expérience a occupé la

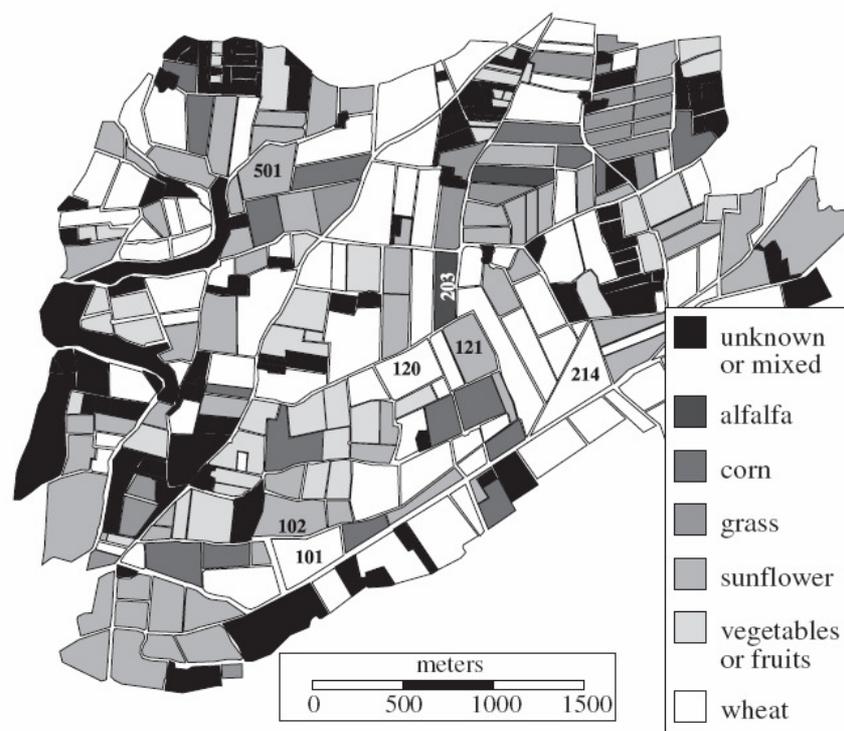
première partie du programme ReSeDA financée par l'Union Européenne, qui visait à améliorer l'utilisation des données de télédétection multitemporelle, multispectrale et multidirectionnelle pour l'assimilation dans des modèles simulant les processus de surface (Baret, 2002; Prévot et al., 1998). La collecte simultanée des données de télédétection et des mesures de terrain permettait d'évaluer des algorithmes d'extraction de l'information sur les processus de surface à partir des données de télédétection à différentes résolutions spatiales. Ces données ont été utilisées au sein de divers programmes, comme par exemple le projet européen WATERMED qui étudiait la résistance à la sécheresse de la végétation sur l'ensemble du bassin méditerranéen (<http://www.avignon.inra.fr/reseda/base/watermed/>; Sobrino et al., 2001), ou par exemple encore le programme GEWEX-Rhône dont l'objectif était la modélisation hydro-météorologique du bassin du Rhône (Ottlé et al., 2002).

Les données Alpilles ont été utilisées dans différentes études portant sur l'estimation des paramètres de surface par télédétection (Jacob et al., 2002a,b ; Weiss et al., 2002 ; Oliosio et al., 2002d ; Hasager et al., 2002a ; Rubio et al., 2003 ; Jacob et al., 2003 ; Jacob & Oliosio, 2005 ; Rivalland et al., 2006) de l'évapotranspiration et du bilan d'énergie de surface (Jacob et al., 2002c ; Oliosio et al., 2002c, Wassenaar et al., 2002, Gomez et al., 2005 ; Oliosio et al., 2006 ; Sobrino et al., 2007 ; Timmermans et al., 2009) et pour l'étude de méthodes d'assimilation de données (Weiss et al., 2001 ; Moulin et al., 2002, Clevers et al., 2002 ; Prévot et al., 2003 ; Demarty et al., 2004, 2005, Oliosio et al., 2005 ; Coudert et al., 2006 ; Hadria et al., 2006 ; Coudert & Ottlé, 2007).

#### **4.1.2. Description (succincte) des données Alpilles**

La zone Alpilles compte une très grande variété de parcelles agricoles d'une taille moyenne de 200 m par 200 m. Les cultures de blé et de tournesol sont les cultures principales représentant respectivement 30% et 20% des surfaces de la zone, ainsi que les vergers qui couvrent environ 15% (pêche, abricot, pomme). D'autres cultures présentes sont le maïs, la luzerne, des cultures légumières, et de la prairie. La carte d'occupation des sols est donnée sur la Figure II.5. L'acquisition des données de terrain a été décrite en détail par Oliosio et al. (2002a). Elles concernent les suivis des flux du bilan d'énergie, de la croissance et de la production de la végétation (*LAI*, biomasse, densité racinaire) et de l'état hydrique du sol sur plusieurs types de cultures. Ces données et en particulier les données de flux ( $R_n$ ,  $H$ ,  $LE$ ,  $G$ ) et

de température de surface ont été acquises en continu sur 7 parcelles (numérotées 101, 102, 120, 121, 203, 214, et 501 ; cf. Fig. II.5). Il faut noter que les parcelles (101, 120, et 214) sont des cultures de blé semées au mois de novembre (pour 101 et 120) tandis que la parcelle 214 a été semée en Février. La parcelle 120 a été irriguée une fois à la fin de mars, les autres parcelles n'ayant pas été irriguées malgré des conditions hydriques défavorables en 1997 (Oliosio et al., 2002a). Le type de sol (argilo-limono-sableux) des 3 parcelles (101, 120, 214) est très similaire (Oliosio et al., 2002a). Le même dispositif expérimental mis en œuvre comporte aussi 3 champs de tournesols (102, 121 et 501) et un champ de luzerne (203) (cf. Tab.II.2). Les principales caractéristiques des sols et des cultures suivies au cours de l'expérimentation sont données dans le tableau II.2.



**Fig. II.5. Carte d'occupation des parcelles du site Alpilles en 1997** issu de Oliosio et al. (2002a). Les 7 parcelles de validation sont numérotées (101: blé, 102: tournesol, 120: blé, 121: tournesol, 203: luzerne, 214: blé et 501: tournesol)

Parallèlement aux données de terrain, de nombreuses données de télédétection ont été acquises tout au long de l'expérimentation à partir de plateforme satellitaires et aéroportées : en particulier capteurs HRV des satellites SPOT (réflectances verte, rouge et proche infrarouge),

capteur POLDER aéroporté (réflectances multidirectionnelles bleu, verte, rouge et proche infrarouge) et caméra aéroportée INFRAMETRICS 760 (luminance infrarouge thermique). L'ensemble de ces données a été acquis avec une résolution spatiale de 20m. Des images ont été ainsi acquises régulièrement tout au long de l'expérimentation, en particulier avec le capteur POLDER et la caméra thermique INFRAMETRICS qui ont opéré tous les 15 à 30 jours (soit une quinzaine de dates entre mars et octobre 1997). Les protocoles d'acquisition aéroportée et le traitement des images ont été détaillés dans Jacob (1999), Jacob et al. (2002a, b, c) et Jacob et al. (2003). Des radiosondages ont été acquis in situ pour certaines de ces journées d'acquisition. Les radiosondages effectués en routine par Météo-France à Nîmes ont été utilisés pour les autres journées.

**Tab. II.2. Caractéristiques des parcelles, issu de Olioso et al. (2002a). (a) Types de sol sont argilo-limoneux pour la classe 10, loam argileux limoneux pour la classe 7 et loam limoneux pour la classe 4 (issu de Clapp & Hornberger, 1978). Ces classes sont déduites de mesures granulométrie dans la couche supérieure du sol de 0 – 30 cm. (b) Méthodes utilisées pour déterminer la capacité au champ, le point de flétrissement et le contenu saturé d'eau peuvent être trouvée dans Braud & Chanzy (2000) et Olioso et al. (2002b).**

DOE représente le jour de l'expérimentation (Day Of Experiment), c'est-à-dire le cumul de jour depuis le 1<sup>er</sup> janvier 2006.

N° parcelle	101	120	214	102	121	501	203
Culture	blé	blé	blé	tournesol	tournesol	tournesol	luzerne
Type soil (a)	10	10	10	7	10	4	7
Capacité au champ ( $m^3 m^{-3}$ ) (b)	0.362	0.368	0.368	0.360	0.368	0.337	0.349
Point de flétrissement ( $m^3 m^{-3}$ ) (b)	0.239	0.239	0.241	0.217	0.229	0.146	0.227
Contenu saturé d'eau ( $m^3 m^{-3}$ ) (b)	0.381	0.380	0.390	0.380	0.400	0.442	0.360
Semis (DOE)	312 (09/11/96)	323 (18/11/96)	409 (12/02/97)	496 (10/05/97)	495 (09/05/97)	447 (12/03/97)	–
Récolte (DOE)	547	548	563	642	631	607	479 529 563 597

### 4.1.3. Données utilisées pour alimenter PBLs

Les différentes données en entrée du modèle PBLs ont été déterminées à partir des informations acquises sur le terrain et par télédétection. Elles ont été fournies au modèle avec une résolution de 20 m sur l'ensemble de la zone. Les paramètres 'sol' ont été prescrits au moyen de la carte des textures élaborées à partir des relevés de terrains et de la carte des sols obtenue auprès de la société du Canal de Provence. Les paramètres de végétation, comme la résistance minimale du couvert, la profondeur racinaire et la rugosité aérodynamique ont été prescrits à partir de valeurs probables obtenues dans la bibliographie ou de relevés de terrain (Oliosio et al., 2002a et b, d) ce qui permet de les cartographier au moyen de la carte d'occupation des sols. la rugosité a été estimée en utilisant la carte d'occupation du sol obtenue à partir d'images SPOT et l'analyse des profils temporels d'indices de végétation pour chaque type de classe d'occupation du sol (Oliosio et al., 2002d, Hasager et al., 2002b). L'albédo a pu être déterminé au moyen des données de réflectances acquises par le capteur Polder aéroporté (Jacob et al., 2002a et b; Jacob & Oliosio, 2005). Grâce au caractère multidirectionnel de ces données, la précision obtenue sur l'albédo est meilleure que 0.02. L'indice foliaire (*LAI*) a été cartographié grâce à la méthode développée par Weiss et al. (2002) qui exploite les mesures multidirectionnelles et multispectrales du capteur aéroporté POLDER au moyen de réseaux de neurones (Rivalland et al., 2006). Les mesures POLDER étant effectuées tous les 15 à 30 jours, ceci a permis d'obtenir facilement les *LAI* journaliers par interpolation linéaire. L'émissivité a été calculée à partir des cartographies d'indice foliaire en se basant sur des relations théoriques de Francois et al. (1997). Les valeurs les plus fortes (au-dessus de 0.98) correspondent aux couverts bien développés, en particulier blé et luzerne, alors que les valeurs de l'ordre de 0.96 correspondent à des sols nus. Les valeurs de l'humidité du sol normalisée  $f_2$  (voir Fig. II.6) sont construites à partir de l'équation (II.22) en utilisant les simulations de la réserve hydrique du sol du modèle ISBA (voir Oliosio et al., 2006 ; Oliosio et al., 2002b)<sup>1</sup>. Par ailleurs, nous avons également testé d'autres formulations proposées pour déterminer l'humidité normalisée (cf. **chapitre II.5.4** et Courault et al., 2008).

---

<sup>1</sup> Les simulations ont été effectuées au moyen du modèle ISBA appliqué à l'ensemble de la zone Alpilles ReSeDA avec une résolution spatiale de 20m du 1<sup>er</sup> janvier au 30 septembre 1997. Les paramètres 'sol' et végétation avaient été prescrit de la même façon que pour l'utilisation actuelle de PBLs, hormis pour le *LAI* et la hauteur du couvert qui ont été prescrits journalièrement en utilisant les estimations POLDER (*LAI*) et les relevés de terrain combinés à l'occupation des sols (hauteur), et l'albédo qui a été calculé journalièrement à partir du *LAI*. Le modèle ISBA requiert également le taux de couverture comme entrée journalière ; celui-ci a été estimé au moyen des données POLDER suivant une méthode similaire à celle utilisée pour le *LAI* (cf. Weiss et al. 2002 et Rivalland et al. 2006) ; les valeurs journalières ont été obtenues par interpolation linéaire entre les dates d'acquisition POLDER.

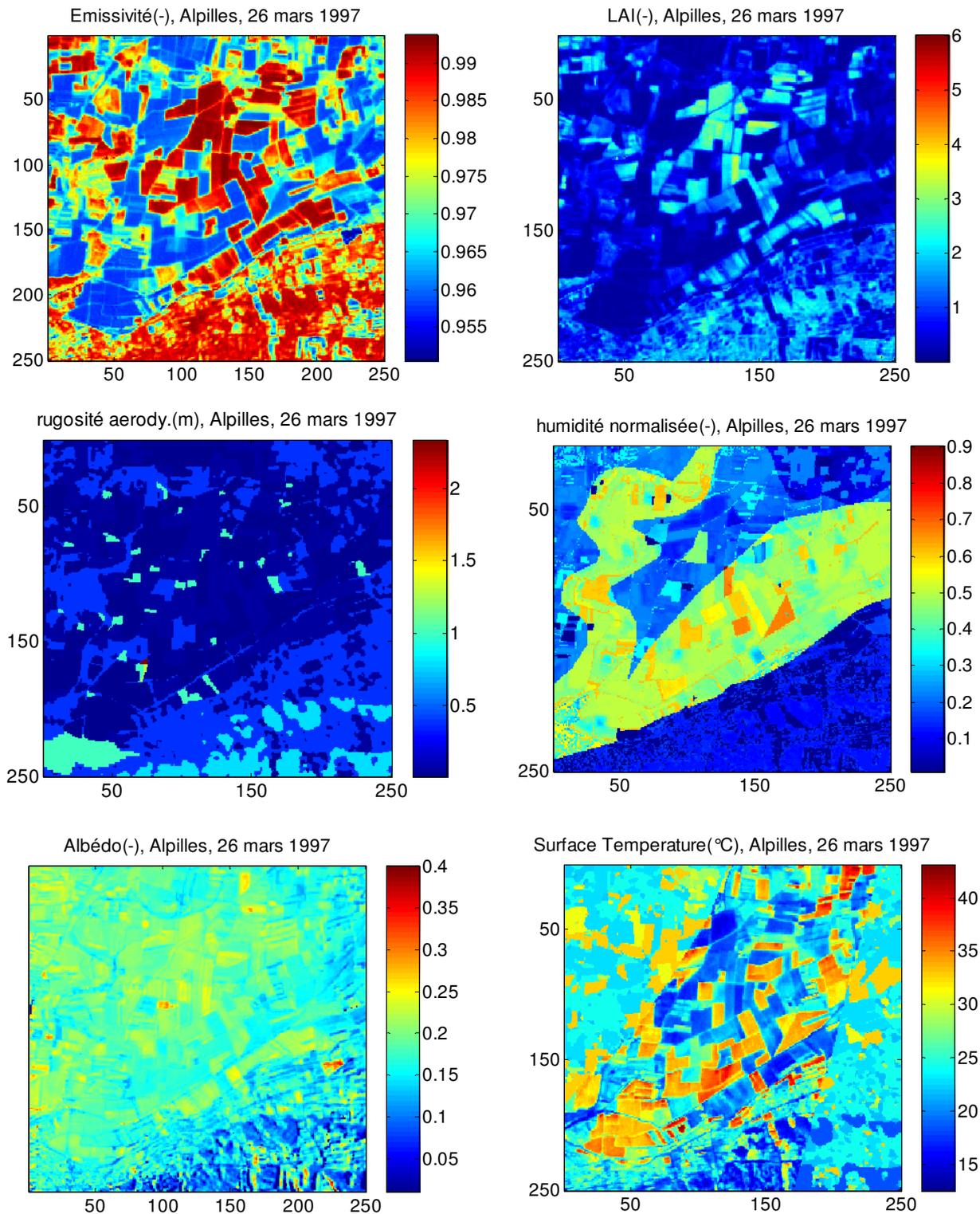
Dans le même temps que les mesures POLDER, les données infrarouge thermique INFRAMETRICS 760 (Jacob, 1999) ont été transformées en température de surface (Jacob et al., 2003), qui seront utilisées plus tard pour alimenter l'algorithme d'assimilation (cf. **section 3.2** ci-dessus).

La figure II.6 propose des cartes de différents paramètres le 26 mars 1997 et permet d'apprécier différents comportements classiques.

- La carte d'émissivité, qui dérive de celle d'indice foliaire, permet de visualiser facilement la présence de couverts végétaux et dans une certaine mesure leur niveau de développement, du bleu pour les sols nus et les couverts peu développés, au rouge pour les plus développés.
- Curieusement, la carte de *LAI* ne permet pas une évaluation aussi aisée, ce qui s'explique en fait par une dynamique beaucoup plus forte du *LAI* que de l'émissivité (l'émissivité saturant assez rapidement lorsque l'indice foliaire augmente, se comporte finalement comme un indice de végétation classique comme le *NDVI* par exemple). Cela dit, il est très intéressant de remarquer que les indices foliaires sont globalement faibles sur la zone, avec seulement quelques parcelles qui atteignent des valeurs de l'ordre de 3 ou plus. Ce type de comportement est assez fréquent dans les zones agricoles qui sont finalement assez peu végétalisées (beaucoup de sols nus, de couverts peu développés ou avec une couverture éparse). Les cartographies de *LAI* réalisées aux autres dates sur la zone montrent également des *LAI* globalement faibles (Weiss et al., 2002, Rivalland et al., 2006).
- En relation avec le type de couvert végétal présent, la carte de rugosité montre également des valeurs souvent faibles dans les zones agricoles, qui sont cependant parsemées de quelques bosquets d'arbres élevés ou de zones construites pour lesquelles des rugosités plus fortes ont été prescrites. Au Sud de la zone, les contreforts des Alpilles montrent la présence plus importante de bois et donc des rugosités plus fortes.
- La carte des températures de surface acquise le 26 mars montre que les sols nus ont en général des températures de surface élevées alors que les zones végétalisées ont des températures plus faibles. La zone étudiée a subi une très forte sécheresse entre fin janvier et mi avril (pas de pluie pendant 3 mois). Les mesures de flux sur le terrain (Oliosio et al., 2002a ; b), ainsi que les cartographies d'évapotranspiration (e.g. Gomez et al., 2005 ; Oliosio & Jacob 2002) ont montré que à cette date les flux de chaleur latente estimés sont en général forts pour les couverts bien développés qui ont la capacité à extraire de l'eau en profondeur, ou qui pour certains, ont pu être irrigués. Les flux de chaleur latente sont par

contre très faibles pour les sols nus qui sont alors très secs (la dernière pluie remonte au 20 janvier, soit plus de deux mois auparavant).

- La carte des humidités normalisées (paramètre  $f_2$ ) montre effectivement que pour certaines parcelles ayant des profondeurs racinaires importantes (cas de la luzerne (parcelle 203)) l'humidité disponible reste importante. La structuration spatiale de la carte du paramètre  $f_2$  est aussi fortement dictée par la structure de la carte des sols qui fait ressortir clairement 3 grands types de zones correspondant à 3 types de sol différents.
- La carte d'albédo montre la faible dynamique de ce paramètre avec des valeurs entre 0.20 et 0.25 pour les cultures et des valeurs plus faibles pour les sols nus et encore plus faible pour les couverts forestiers du sud de la zone. Il a également été remarqué que, bien qu'ayant des albédos plus faibles que les couverts développés, ces sols nus présentent des rayonnements nets plus faibles en raison de leur forte température de surface (non montré ici).



**Fig. II.6.** Cartographies (5 km par 5 km avec une résolution de 20) d'émissivité, de LAI, de rugosité aérodynamique, d'humidité normalisée, d'albédo, et de température de surface, sur le site Alpilles le 26 mars 1997 aux environs de midi solaire.

## 4.2. Les autres jeux de données

En plus des données d'Alpilles ReSeDA décrites ci-dessus, nous avons eu recours à trois autres jeux de données pour explorer nos pistes de recherche sur les modifications apportées au modèle PBLs dans sa nouvelle version (cf. PBLs version 2 **Chapitres III.3** et **V**). Les caractéristiques de ces jeux de données sont présentées dans le Tableau II.3. En nous limitant aux seules données expérimentales d'Alpilles, nous aurions eu accès à une couverture plus partielle des situations. Dans le cadre de ce travail, deux études importantes ont été réalisées pour analyser le comportement de deux paramètres qui sont assez mal connus dans la littérature, ou plutôt pour lesquels il n'existe pas de consensus malgré le nombre important d'études qui leur ont été consacrées. Il s'agit d'une part de la rugosité thermique ( $z0_h$ ) ou du paramètre  $kB^{-1}$  qui sera étudié avec le jeu de données du site 'Flux et Télédétection' d'Avignon ou site grue, et d'autre part du rapport  $G/R_n$  qui correspond au paramètre  $c_{0i}$  dans PBLs. Ce dernier est utilisé pour calculer le flux de chaleur dans le sol par conduction  $G$  à partir de l'estimation du rayonnement net. Nous l'avons étudié grâce à des données acquises sur le terrain (Site grue Avignon et données Alpilles), ainsi qu'à des données simulées dans les configurations d'un sol nu et d'un couvert peu développé au moyen des modèles SiSPAT (Braud et al., 1995 ; Olioso et al., 2002b ; Braud et al., 2005) et TEC (Chanzy & Bruckler, 1993 ; Mumen, 2006 ; Chanzy et al., 2008). Ces modèles sont considérés comme des modèles très aboutis au niveau de la description des transferts au voisinage de la surface du sol (en particulier en considérant les processus de transferts d'eau et de chaleur de manière couplés et en considérant également les transferts convectifs de vapeur dans les couches sèches du sol). Ils sont ainsi des outils puissants pour explorer différentes situations environnementales en détail. Ces différents jeux de données ont servi d'une part à développer la base d'apprentissage des différentes équations mises en œuvre et d'autre part à la validation de ces méthodes. En nous limitant aux seules données expérimentales, nous aurions eu accès à une couverture plus partielle des situations ce qui aurait très vite limité la validité des équations utilisées ou élaborées. C'est pourquoi, les données simulées sont un bon outil d'apprentissage.

**Tab. II.3. Exemple de gamme de valeurs des différentes données, types de culture et textures du sol**

	Sols/culture	Texture (%)	albédo	Humidité (0-5 cm)
TEC Mumen (2006)	AL-SiCL	limon (56), argile (39)	0.12 - 0.20	0.12 - 0.20
	MAR	limon (26), argile (63)	0.15 - 0.22	0.30 - 0.35
	MX-SL	limon (24), sable (59)	0.065 - 0.22	0.15 - 0.22
SiSPAT (Alpilles) Oliosio et al. (2002)	101 (blé)	limon (54), argile (42)	0.23	0.453
	120 (blé irrigué)	limon (54), argile (42)	0.23	0.473
	214 (blé printemps)	limon (52), argile (46)	0.23	0.473
GRUE  Les différentes cultures sur le site grue	2002 (blé)	limon (51), argile (33) sable (16)	0.09 - 0.34	0.12 - 0.31
	2003 (tournesol)		0.08 - 0.40	0.15 - 0.51
	2004 (blé)		0.15 - 0.30	0.07 - 0.40
	2005 (pois)		0.23 - 0.32	0.12 - 0.27
	2006 (blé)		0.13 - 0.25	0.13 - 0.35
	2007 (sorgho)		0.10 - 0.37	0.08 - 0.35

#### **4.2.1. Description (succincte) des données du site 'Flux et Télédétection' d'Avignon (Site GRUE)**

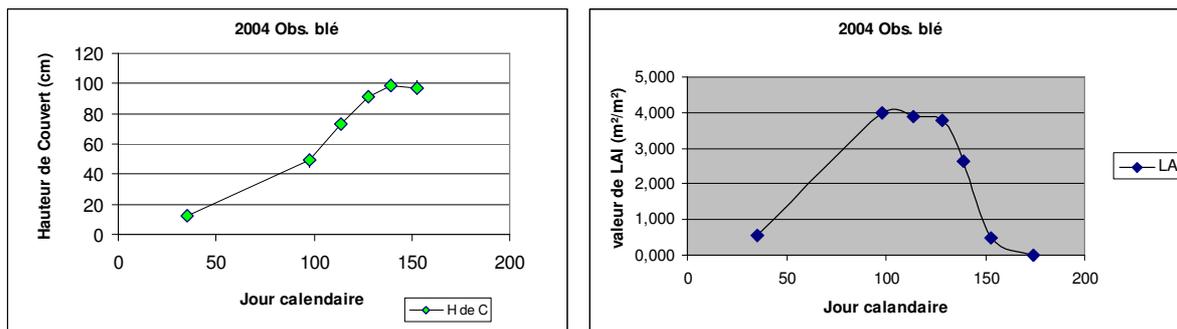
Le site GRUE est situé à proximité du bâtiment Climat de l'unité EMMAH (Environnement Méditerranéen et Modélisation des Agro-Hydro systèmes) localisé sur le domaine Saint Paul de l'INRA d'Avignon (N 43 ° 50 ' et E 4 ° 52'). Il est composé de deux parcelles séparées par une voie de roulement d'une grue à tour de 20 m de haut servant de support aux instruments de télédétection (Fig. II.7a). La parcelle "ouest" de 1.8 ha, adaptée aux mesures de flux est équipée d'une rampe d'irrigation à déplacement frontal permettant de faire des apports d'eau sur l'ensemble de la culture (Fig. II.7b). La parcelle "Est" de 1 ha, destinée à des expérimentations sur de petites surfaces est équipée d'une serre mobile permettant de conduire une culture dans des conditions hydriques variées. Ces deux parcelles servent de terrains d'expérimentations à l'unité.

L'observatoire a pour but de rassembler des données et des informations tout au long de l'année de manière à obtenir un suivi continu sur : les itinéraires techniques, les cultures, les conditions climatiques, et l'état du sol.



**Fig. II.7. Dispositif expérimental du site GRUE : a) GRUE à tour ; b) rampe d'irrigation**

Nous nous sommes intéressés aux données acquises sur la période de 2002 à 2007 dont une partie a été acquise dans le cadre du projet européen CARBOEUROPE-IP (<http://www.carboeurope.org/>). Les flux convectifs sont acquis suivant la méthode des corrélations turbulentes au moyen d'un anémomètre sonique 3D et d'un analyseur de gaz rapide placé au dessus du couvert. L'ensemble des données météorologiques, la température de surface et l'albédo sont acquis en continu fournissant des données moyennes toutes les demi-heures. L'indice foliaire de surface (*LAI*) et la hauteur de la végétation sont mesurés régulièrement sur le terrain (ex. sur la Fig. II.8) permettant d'obtenir par une fonction d'interpolation des estimations en continu. Des mesures d'humidité du sol sont également réalisées en continu près de la surface (0 - 5 cm) au moyen de sondes capacitatives et toutes les semaines jusqu'à 1m50 de profondeur au moyen de sondages neutroniques. Une fiche descriptive des interventions (irrigation, engrais) et sur des itinéraires (date de semis, date de récolte) des cultures est renseignée pour compléter les informations acquises sur la parcelle. Les caractéristiques des cultures réalisées au cours de la période considérée sont données dans le Tableau II.3.



**Fig. II.8. Mesures ponctuelles de la hauteur de la végétation et de LAI du site GRUE en 2004**

#### 4.2.2. Description (succincte) des données simulées par SiSPAT

Le modèle SiSPAT a été développé par Braud et al. (1995) pour étudier les transferts sol-végétation-atmosphère. Comme dit précédemment (**Chapitre I**) ce modèle se caractérise par un traitement fin des transferts hydriques et thermiques dans le sol et une description du bilan d'énergie de surface en deux couches (surface du sol et végétation). Les données générées par le modèle SiSPAT ont été utilisées dans notre étude pour l'étude du rapport  $G/R_n$ . Ce sont des simulations qui ont été réalisées dans le cadre du programme Alpilles suivant trois scénarios répondant à des objectifs tels que : (i) tester la capacité des modèles à décrire les processus en utilisant les paramètres d'entrée des modèles tels qu'ils ont été mesurés dans les champs ; (ii) tester la portabilité des modèles en utilisant des informations a priori en entrée (comme des fonctions de pédotransferts), (iii) et tester la robustesse des modèles par une procédure de calibration - validation. Ces simulations ont fait l'objet de l'étude de Oliosio et al. (2002b) sur la comparaison des modèles TSVA. Elles ont été réalisées sur les parcelles 101, 120, 121, 214, et 501 de la zone Alpilles sur l'ensemble de la période 1996-1997 (cf. Fig. II.5). Ce jeu de données compte les flux de surfaces ( $R_n$ ,  $G$ ,  $H$ ,  $LE$ ) ; les températures de surface et les profils d'humidité du sol simulés par le modèle à un pas de temps fin (30 minutes), ainsi que les entrées du modèle telles que l'albédo, la hauteur de la végétation (soit la rugosité de surface  $z_0$ ), et l'indice foliaire ( $LAI$ ) représentant les parcelles Alpilles.

Contrairement au modèle PBLs, SiSPAT est un modèle dynamique qui peut simuler chaque parcelle de façon continue dans le temps. Il n'a pas besoin d'être réinitialisé et considère un

forçage climatique fin toutes les demi-heures (cf. chapitre I.1.3). Nous avons utilisé l'ensemble du jeu de données en ne retenant que les données sous l'hypothèse d'un *LAI* supérieur à 0.25 (afin de représenter une configuration de couverts peu ou développés, cf. chapitre III.2). Ainsi donc l'ensemble des données SiSPAT (parcelles 101, 120, 121, 214 et 501) constitue 393 points sur l'ensemble des jours simulés autour de midi.

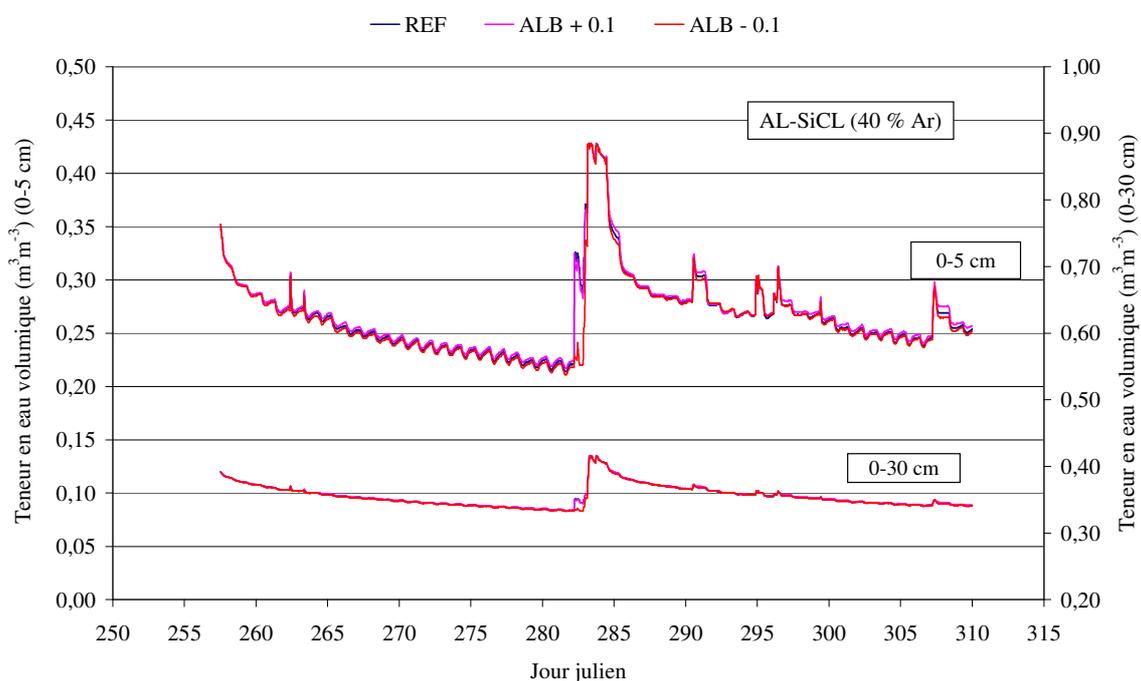
#### **4.2.3. Description (succincte) des données simulées par TEC**

Le modèle TEC est un modèle développé par Witono et Bruckler (1989) (voir aussi Chanzy & Bruckler, 1993) pour l'étude des transferts hydriques et de chaleur dans les sols (il n'inclut pas de végétation). Il a donc été utilisé ici pour étudier le comportement de PBLs pour des cas de sols nus soumis à différentes conditions d'environnement.

En termes théoriques le modèle TEC est très proche de la partie sol du modèle SiSPAT (voir **chapitre I**). Les données du modèle TEC sont utilisées pour approfondir l'étude du rapport ( $G/R_n$ ). Ce modèle permet d'étudier les transferts d'eau et de chaleur dans le sol et permet un bon couplage avec l'atmosphère. La surface du sol dans le modèle TEC est caractérisée par son albédo, sa rugosité aérodynamique et son émissivité de surface. Le modèle nécessite comme conditions initiales un profil de potentiel matriciel et un profil de température. Les conditions aux limites supérieures peuvent être soit un flux d'eau et de chaleur à la surface, soit un ensemble de données climatiques (le rayonnement global, le rayonnement atmosphérique, la température de l'air, la pression de vapeur d'eau de l'air) qui vont permettre, à l'aide du bilan d'énergie, d'estimer le potentiel hydrique et la température de surface (Fig. II.9). Les données TEC que nous avons utilisées ont été simulées par Mumen (2006 ; voir aussi Chanzy et al., 2008) dont l'étude portait sur l'évolution de la teneur en eau dans la couche labourée, considérée comme déterminante pour prendre la décision d'intervenir dans les parcelles ou pas (pénétration d'engin agricole). L'expérience (Mumen, 2006) avait pour objectif d'analyser les performances obtenues en considérant différents niveaux d'information, allant de données limitées à des stratégies de calage sophistiquées permettant de tirer le meilleur parti du modèle TEC.

Les données ont été simulées dans la configuration d'un sol nu sur différents types de sols afin d'explorer différentes textures et structures du sol. Dans cette étude, nous avons eu recours aux données simulées sur les sols nommés AL-SiCL correspondant au sol de la parcelle 101 de

la zone Alpilles (Oliosio et al., 2002b), au sol MAR (Aboudrare, 2002) situé au Maroc et au sol MX-SL (Findeling, 2001) localisé au Mexique. Les simulations ont été réalisées avec un climat identique (durant les mois de février à avril 2004 à Avignon). Nous disposons des données simulées des flux de surface ( $R_n$ ,  $G$ ,  $H$ ,  $LE$ ), de la température de surface, et de la teneur en eau de différentes couches dont la surface (0-5 cm) en continu sur la période de simulation avec un pas de temps d'une heure (voir l'exemple donné en Figure II.9). Nous avons constitué une base de données de 35 jours correspondant à des périodes de parfait rayonnement solaire.



**Fig. II.9. Evolution de la teneur en eau de surface (horizons 0-5 cm et 0-30 cm) pour un sol argile limoneux en fonction de différentes valeurs d'albédo du sol (ALB : valeur d'albédo perturbée par une erreur de plus ou moins de 0.1 ; REF : valeur vraie d'albédo)**

## 5. Mise en œuvre de la méthode d'assimilation sur les données Alpilles

### 5.1. Introduction

La méthode d'assimilation a été appliquée au jeu de données Alpilles. Nous avons défini précédemment que nous allions utiliser les données de télédétection dans l'infrarouge thermique pour retrouver par assimilation les paramètres les plus difficiles à connaître directement, soit ici la rugosité aérodynamique  $z_0$  et l'humidité normalisée  $f_2$ . Les autres paramètres comme l'albédo ( $\alpha$ ), le  $LAI$ , sont eux faciles à estimer par télédétection dans le domaine solaire. Trois versions de l'algorithme d'assimilation ont été élaborées considérant la connaissance préalable ou pas de l'un des deux paramètres recherchés :

- algorithme A : estimation de  $z_0$  uniquement,  $f_2$  étant connu
- algorithme B : estimation de  $f_2$  uniquement,  $z_0$  étant connu
- algorithme C : estimation des deux paramètres simultanément.

Avant de mettre en œuvre l'assimilation sur les données réelles et pour mieux comprendre le fonctionnement de l'algorithme d'assimilation, nous avons opté pour la réalisation d'expériences synthétiques (également appelées expériences jumelles). En effet, l'évaluation de méthodes de cartographie des flux de surface et du microclimat est toujours difficile à partir des données de terrains qui ne permettent pas de disposer d'un jeu de données de validation très exhaustif. Dans le cas de l'expérimentation Alpilles on dispose par exemple de 3 à 7 sites de mesures des flux au sol, selon les dates (au nombre d'une quinzaine), ce qui est très faible par rapport aux centaines de parcelles de la zone (la zone étudiée compte 351 parcelles). Signalons cependant qu'il s'agit d'un des jeux de données les plus larges qui soient utilisables pour ces validations (cf. la compilation faite par Bastiaanssen et al., 2005). Les expériences synthétiques permettent de disposer d'un grand nombre de données et de maîtriser l'ensemble des conditions de leur « acquisition » (avec cependant le désavantage de ne pas représenter parfaitement la réalité puisque simulées au moyen d'un modèle). Elles permettent donc d'étudier en détail le fonctionnement des méthodes en permettant de tester différentes hypothèses de travail, en particulier portant sur les niveaux d'erreur affectant les données ou le modèle, en conditions optimales (les hypothèses sous-jacentes au système d'assimilation sont parfaitement respectées) ou sous-optimales.

L'ensemble des tests (données synthétiques et réelles) a été appliqué à deux versions du modèle PBLs, la version décrite précédemment (section 2 de ce chapitre) et une nouvelle version qui sera décrite plus bas (chapitre III.3). Cette nouvelle version intégrera de nouvelles paramétrisations de la rugosité thermique ( $z0_h$ ) et du rapport  $G/R_n$ .

## 5.2. Expérience synthétique

### 5.2.1. Généralités

Les expériences synthétiques (ou expériences jumelles) consistent à construire un jeu de données simulées. Le principe peut se résumer en trois points :

- on détermine des valeurs « vraies » pour les paramètres à contrôler, et on simule les états observables vrais correspondants avec le modèle original,
- on crée les observations en bruitant les états observables vrais (simulés à l'étape précédente),
- on cherche à retrouver avec l'algorithme d'assimilation les paramètres vrais en partant des observations qui viennent d'être créées.

Beaucoup de travaux ont eu recours à cette méthode lorsque les observations n'étaient pas suffisantes, les conditions météorologiques ne permettant pas les prises d'images satellite par exemple, ou encore comme dans Mazaurec (2003), les observations de hauteur d'eau qu'il souhaitait assimiler pendant les crues se révèlent impossibles à mesurer. Walker et Houser (2001) ont présenté une étude d'assimilation de l'humidité du sol sur des données synthétiques motivée par la nécessité d'améliorer l'humidité des sols pour l'initialisation des modèles de prévision climatologique et hydrologique. Seuffert et al. (2003) ont assimilé des observations synthétiques de la température de brillance de surface associées aux mesures de la température de l'air et l'humidité relative pour tester le potentiel d'amélioration de l'humidité du sol. D'autres détails intéressants et argumentaires concernant l'utilité des expériences jumelles sont donnés par Pipunic et al. (2008).

La caractéristique intéressante de cette technique est *-i)* que les valeurs vraies des paramètres d'entrée et les variables de sorties du modèle sont connues pour chaque pixel et *-ii)*

que les erreurs d'observation et d'ébauche (matrices  $B$  et  $R_i$  définies par les équations II.44 et II.47) sont connues et contrôlées. Les expériences jumelles présentent ainsi beaucoup d'avantages qui doivent permettre de bien évaluer la qualité de la méthode d'assimilation :

- elles permettent de contrôler et définir l'erreur sur les observations ;
- elles permettent de considérer, à juste titre, le modèle comme "parfait" et évitent donc le problème de l'erreur du modèle qui est généralement très difficile à évaluer ;
- les paramètres et variables d'intérêt estimés après assimilation peuvent être comparés à la solution du problème ;
- la base de données créée peut correspondre à des situations très différentes, en particulier en terme de nombre d'observations disponibles, de fréquence, et de répartition temporelle et spatiale.

### 5.2.2. Mise en pratique

Dans notre cas, l'expérience synthétique est construite suivant quatre étapes.

1. Définition d'un scénario "vrai", en choisissant des valeurs des paramètres d'entrée du modèle PBLs (albedo,  $LAI$ , humidité du sol, rugosité, vitesse du vent) et calcul des sorties "vraies" du modèle (température de surface, température de l'air, flux de surface...) correspondant aux paramètres d'entrée "vrais".
2. Calcul des observations synthétiques de la température de surface  $T_s$  en bruitant les valeurs "vraies" de  $T_s$  (par ajout d'erreur) : ( $T_{sobs}$ ).
3. Estimation des paramètres ( $z_0$ ,  $f_2$ ) sur chaque pixel en assimilant les observations synthétiques ( $T_{sobs}$ )
4. Évaluation de l'impact des erreurs des autres paramètres d'entrée de PBLs ( $LAI$ , albedo, vitesse du vent, rapport  $G/R_n$ ) sur les flux de surface et les paramètres ( $z_0$ ,  $f_2$ ) réestimés.

Le but de l'algorithme d'assimilation est d'améliorer la simulation des variables de sortie de PBLs ( $R_n$ ,  $H$ ,  $LE$ ,  $G$ ,  $T_s$ ,  $T_a$ ) sur un domaine spatial donné, en estimant les valeurs des paramètres ( $z_{0i}$ ,  $f_{2i}$ ) sur chaque pixel  $i$  à l'aide de la carte de température de surface pseudo observée ( $T_{sobs}$ ). Aussi, dans les étapes 3 et 4, les paramètres estimés ( $z_{0i}$ ,  $f_{2i}$ ) et les sorties de PBLs (flux de surface,  $T_s$ ,  $T_a$  à 2m) sont comparés aux valeurs correspondantes du scénario

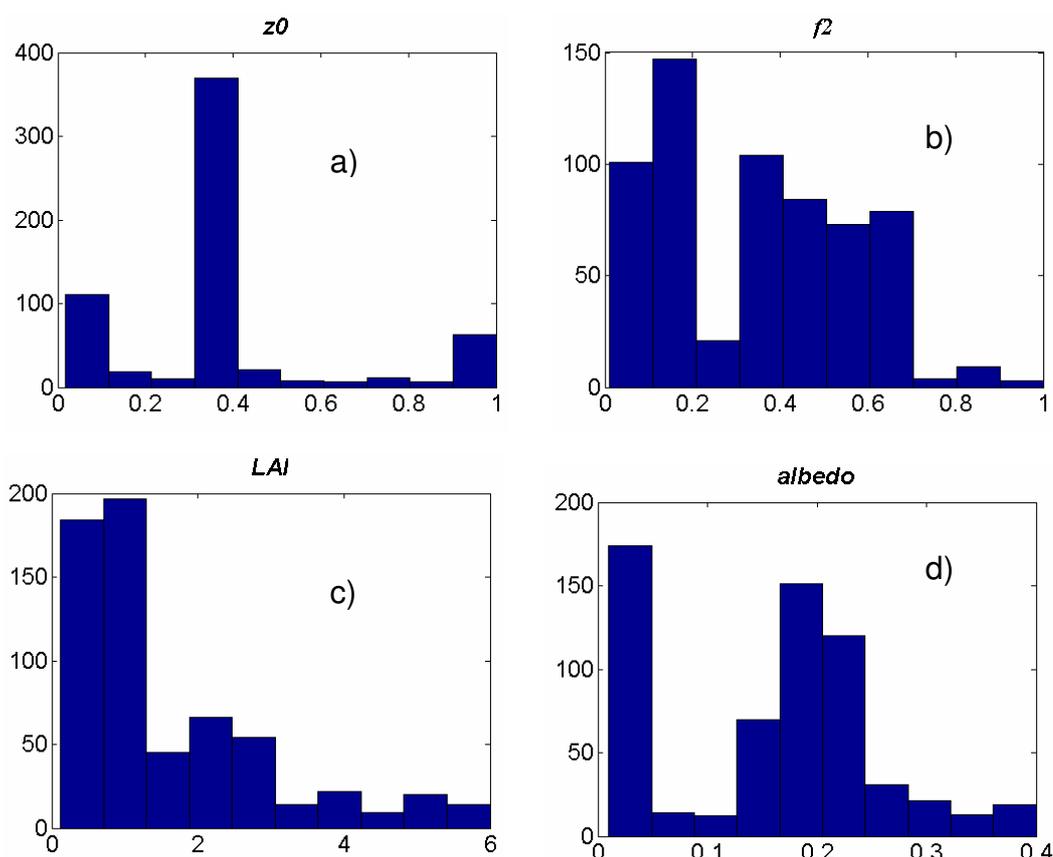
"vrai" défini à l'étape 1. La stratégie d'évaluation des résultats des algorithmes d'assimilation est présentée plus loin dans la **section 5.5**.

Pour être aussi proche que possible d'un contexte réaliste, le scénario "vrai" est construit à partir des données acquises au cours de l'expérience Alpilles – ReSeDA et qui ont été présentées dans la section précédente (cf. section 4.1).

Etape 1.

Les données synthétiques sont construites à partir de 625 points choisis au hasard sur la zone étudiée, de façon à représenter toutes les gammes de valeur des variables et paramètres d'entrées du modèle suivant la structure de la surface et le type de culture en supposant que chacun de ces 625 pixels (ou points) ont le même pourcentage d'occupation de la zone (Fig. II.10). Il y est représenté des sols nus et des couverts développés, des zones humides et sèches. Les données utilisées correspondent à la date du 18 avril 1997 (Day Of Experiment 474) et sont obtenues à partir des cartes et des données météorologiques qui ont été présentées dans la section précédente (*LAI*, albédo, émissivité, rugosité, humidité normalisée....) (cf. section 4.1).

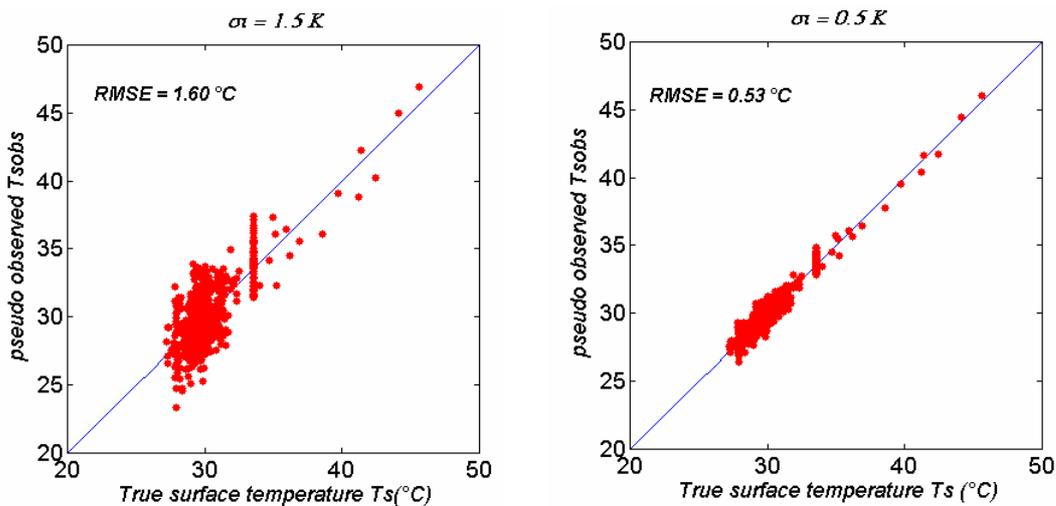
Dans la première étape, le scénario "vrai" est directement obtenu en exécutant le modèle PBLs. L'ensemble des valeurs des variables de sortie de PBLs (température de surface, température de l'air, flux de surface...) et des paramètres d'entrée constitue l'état "vrai".



**Fig. II.10. Distributions de la gamme de valeurs des paramètres utilisés dans le scénario vrai: a)  $z_0$ ; b)  $f_2$ ; c)  $LAI$ ; d) albédo**

Etape 2.

Les observations synthétiques, ou pseudo-observées ( $T_{sobs}$ ), sont générées par l'ajout de perturbations aléatoires aux vraies valeurs de  $T_s$ . Ces perturbations aléatoires sont générées de façon indépendante d'un pixel à un autre. Elles suivent toutes une distribution normale de moyenne nulle et d'écart type  $\sigma_i$ . Une première série d'expériences a été réalisée avec  $\sigma_i = 1.5$  K, et une deuxième expérience avec  $\sigma_i = 0.5$  K. Selon différentes études (cf. Jacob et al., 2008), ces niveaux d'erreurs correspondent respectivement à des erreurs moyennes et à une très bonne précision. La figure II.11 ci-dessous montre la variation des pseudo-observations par rapport aux valeurs vraies suivant l'hypothèse d'erreur.

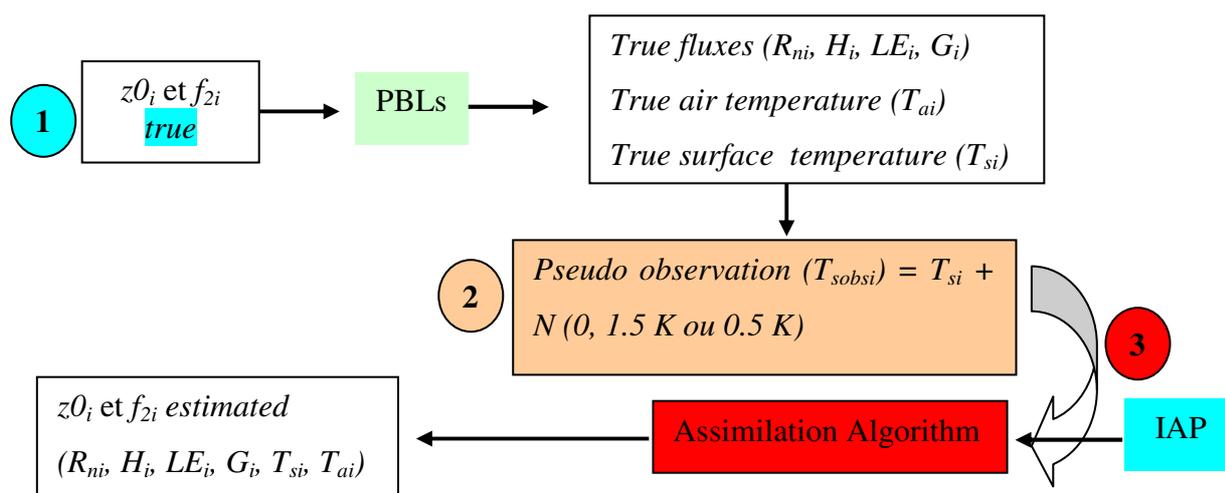


**Fig. II.11. a) Température de surface pseudo-observée ( $T_{sobs}$ ) avec  $\sigma_i = 1.5$  K ; b) Température de surface pseudo-observée ( $T_{sobs}$ ) avec  $\sigma_i = 0.5$  K, par rapport aux vraies valeurs de  $T_s$  (définie à l'étape 1)**

Etape 3.

Cette étape consiste en l'application de l'algorithme d'assimilation défini à la section 3 de ce chapitre. Le schéma II.12 permet d'en illustrer le principe d'application sur les données

synthétiques : à partir d'observations bruitées ( $T_{sobs}$ ) sur l'ensemble des pixels considérés, l'algorithme cherche à retrouver l'état vrai des paramètres estimés (rugosité aérodynamique  $z0_i$  et facteur d'humidité du sol normalisé  $f_{2i}$ ), des flux de surface ( $R_{ni}$ ,  $H_i$ ,  $LE_i$ ,  $G_i$ ) et des températures de surface ( $T_{si}$ ) et de l'air à 2 m ( $T_{ai}$ ) sur chaque pixel. Les paramètres et variables d'entrées du modèle non estimés par l'algorithme d'assimilation sont dans ce cas fixés à leurs valeurs exactes.



**Fig. II.12. Schéma de la procédure d'assimilation à partir des pseudos observations**

#### Etape 4.

En pratique,  $z0$  et  $f_2$  sont les deux principaux paramètres inconnus que nous recherchons par assimilation. Des incertitudes importantes peuvent affecter les autres entrées du modèle qui ne sont pas réestimées par assimilation, tels que  $LAI$ , albédo,  $c_0$  et vitesse moyenne journalière du vent ( $u_a$ ). En mode direct (PBLs seul), ces erreurs sont susceptibles d'affecter les résultats du modèle, en fonction de sa sensibilité à ces paramètres. Toutefois, en fonction de la structure du modèle, l'algorithme d'assimilation peut aussi avoir une influence sur l'impact de ces erreurs. Pour évaluer cette influence, l'algorithme d'assimilation a été appliqué par la suite avec des erreurs introduites sur ces paramètres d'entrée. Nous considérerons dans les analyses faites sur ces données, que les erreurs introduites nous conduisent à utiliser un modèle imparfait permettant de définir ici des « erreurs modèles ». Les valeurs de  $LAI$ , d'albédo ( $\alpha$ ), de  $c_0$  ( $G/R_n$ ) et la vitesse moyenne du vent de la journée ( $u_a$ ) utilisées ici dans le modèle PBLs sont dans ce cas obtenues en ajoutant des erreurs sur les valeurs vraies définies à l'étape 1. Nous avons

ajouté des erreurs aléatoires suivant des distributions gaussiennes pour l'albédo (avec pour écart type 0.02, correspondant à la précision donnée par Jacob et al., 2002a et b ; Jacob et Oliosio, 2005 ; Tasumi et al., 2008) et pour le *LAI* (écart type de 0.5 correspondant à l'évaluation des erreurs par Weiss et al., 2002 et Rivalland et al., 2006). Pour le ratio  $G/R_n$  qui est défini comme constant dans la première version de PBLs ( $c_0 = 0.15$  sur l'ensemble du domaine simulé), des études telles que Norman et al. (1995) et Kustas et al. (1990) ont montré que des variations importantes existaient, entre 0.1 à 0.4. Par conséquent, l'algorithme d'assimilation est exécuté avec la formule décrite par Choudhury et al. (1987) pour calculer  $c_0$  en fonction de *LAI* (correspondant à l'ajout d'erreur dans la procédure d'assimilation). Pour la vitesse moyenne du vent au cours de la journée fixée constante sur l'ensemble de la zone considérée à  $1.8 \text{ ms}^{-1}$ , une erreur de  $\pm 1 \text{ ms}^{-1}$  a été ajoutée à la valeur vraie.

Pour l'étape 3, les trois algorithmes A, B et C définis plus haut ont été appliqués. Pour l'étape 4, seuls les algorithmes B et C sont présentés dans ce mémoire (cf. **Annexe II** pour l'algorithme A). Les valeurs initiales des paramètres pour la minimisation sont, dans la plupart des tests, fixées aux valeurs a priori (sauf pour certains tests de validation présentés dans le **chapitre IV**, pour lesquels il sera alors précisé que celles-ci sont tirées aléatoirement dans une distribution uniforme définies par les bornes des paramètres).

L'application de ces algorithmes nécessite également la connaissance des matrices de covariance d'erreurs ( $B$  (éq. II.44) et  $R_i$  (éq. II.47)) et des valeurs des informations a priori  $x_b$  (éq. II.42). La définition des informations a priori et de leurs erreurs, qui est une étape primordiale de la procédure d'assimilation, va être détaillée dans une section suivante (section 5.4). Pour la plupart des tests sur les expériences synthétiques, les composantes des matrices  $R_i$  et  $B$  seront fixées aux valeurs théoriques utilisées pour générer les erreurs sur les observations et les informations a priori (sauf pour certains tests pour lesquels cela sera précisé). Dans ces cas là, pour l'étape 3, le problème sera donc consistant. Il ne le sera pas dans le cas de l'étape 4 étant donné la prise en compte d'erreurs dans le modèle.

### 5.3. Mise en œuvre sur les données réelles

Pour l'application sur les données Alpilles, nous avons considéré l'ensemble des 62500 pixels de la zone de 5 km par 5 km pour laquelle nous disposons des images aéroportées de température de surface présentées plus haut (caméra INFRAMETRICS, résolution spatiale de 20 m), ainsi que des cartographies d'albédo et d'indice foliaire obtenues à partir des données POLDER. Les dates du 26 mars 1997 (DOE 451), du 16 avril 1997 (DOE 472) et du 18 avril 1997 (DOE 474) ont été sélectionnées pour ce travail. Ce sont des dates dont les conditions météorologiques sont relativement favorables par rapport aux autres et pour lesquelles on avait des radiosondages.

En conditions réelles, l'application du système d'assimilation pose plusieurs problèmes comme la définition des valeurs d'information a priori et des statistiques d'erreur (sur les observations et sur l'information a priori). Nous verrons sur les résultats des expériences synthétiques que ces valeurs ont un impact assez fort sur la qualité des résultats obtenus. Nous avons appliqué les différents schémas de l'algorithme d'assimilation sur les données en supposant que les erreurs d'observation sur les températures de surface étaient indépendantes entre chaque pixel et suivaient une loi normale centrée et d'écart type  $\sigma_i$  constant sur toute la zone et égal à 1.5 K. Cette valeur nous semble réaliste pour une telle application. Elle reflète à la fois l'erreur de mesure et l'erreur modèle (cf. **chapitre III.3**). Nous n'avons cependant que peu d'éléments pour confirmer ces hypothèses. Le réalisme de l'hypothèse d'indépendance des erreurs entre pixels peut être discuté. En effet, l'influence de différents pixels voisins sur la mesure de  $T_s$  en un pixel, les biais introduits par la correction atmosphérique sur l'ensemble des mesures et la corrélation probable des erreurs du modèle sur les différents pixels d'une même parcelle ou ayant le même type de végétation sont autant de sources potentielles de corrélations entre les erreurs d'observations sur les différents pixels. Cependant, nous ne disposons pas de données suffisantes pour pouvoir estimer ces corrélations et nous les négligerons donc.

La définition des valeurs d'information a priori (*IAP*) des paramètres ( $z_0$  et  $f_2$ ) et la matrice de covariance d'erreur de l'information à priori (*B*) sont décrites dans ce chapitre à la section suivante (section 5.4.1).

L'application étant réalisée sur un grand nombre de pixels, il s'agit d'une opération très coûteuse dans notre cas en raison des moyens informatiques dont nous disposons. Les résultats des expériences jumelles, comme on le verra plus tard dans le **chapitre V**, ont montré que l'algorithme B (rechercher seulement le facteur d'humidité du sol  $f_2$ ) apporte une meilleure réponse ou une réponse similaire sur l'estimation de l'évapotranspiration par rapport aux deux autres algorithmes (A et C). Par conséquent une grande partie des mises en œuvre dans le cas réel concerneront l'algorithme B qui a l'avantage d'un coût de calcul moins élevé que l'algorithme C (recherche simultanée de  $z_0$  et  $f_2$ ). Quelques analyses ont cependant été effectuées avec l'algorithme C.

La stratégie d'évaluation des résultats des algorithmes d'assimilation est présentée plus loin dans la **section 5.5**.

#### **5.4. Définition des informations a priori**

Nous commencerons ici par décrire les informations a priori utilisées dans le cas de l'application aux données réelles, puis nous poursuivrons par la description de celles utilisées dans le cas des expériences synthétiques qui découlent de celles utilisées dans le cas réel.

##### **5.4.1. Cas des données réelles**

L'information a priori utilisée pour la rugosité  $z_0$  dans le cas de l'application aux données réelles provient de mesures de terrain spatialisées par classe d'occupation des sols. Le lecteur est invité à consulter la section 4 de cette partie pour plus de détails sur ce point.

L'erreur sur l'information a priori sur  $z_0$  est supposée suivre une loi normale centrée. Afin de fixer une variance réaliste pour cette erreur une analyse des comportements classiques des erreurs sur  $z_0$  a été faite en se basant sur l'expérience accumulée des mesures de terrain ou d'estimation de ces paramètres au sein du laboratoire. Nous avons supposé que des erreurs de l'ordre de 100 % pouvaient facilement être commises lorsque  $z_0$  était faible (cas des sols nus, des cultures basses, des prairies...). Pour les couverts plus hauts des erreurs de 20 % apparaissent plus réalistes. Pour les valeurs intermédiaires, nous avons décrit les erreurs par

une fonction d'interpolation linéaire reliant les deux extrêmes définissant l'écart type pour chaque pixel (i) par :

$$\sigma_{zbi} = \frac{z0_i \times (\max(z0_i) - z0_i) - 0.2 \times z0_i \times (\min(z0_i) - z0_i)}{\max(z0_i) - \min(z0_i)}, \quad (\text{II.53})$$

où  $\max(z0_i)$  et  $\min(z0_i)$  sont les valeurs maximale et minimale de  $z0$  sur l'ensemble de la zone considérée.

Les erreurs de l'information a priori sur  $z0$  sur les différents pixels sont supposées indépendantes. Cette hypothèse est probablement réaliste si l'on considère des pixels pour lesquels la végétation est différente. Par contre, étant donné que la valeur de l'information a priori sur  $z0$  est identique pour les pixels ayant la même végétation, les erreurs sur les pixels d'une même parcelle ou d'une même végétation sont forcément corrélées. Nous négligerons ces corrélations dans cette étude.

Les valeurs de l'information a priori pour le paramètre d'humidité  $f_2$ , sont plus difficiles à déterminer car nous ne disposons pas de mesures suffisamment nombreuses de ce paramètre. Comme décrit dans la section 4 de cette partie, une possibilité est d'utiliser les valeurs simulées au moyen du modèle ISBA. Cependant, il apparaît que ces valeurs sont parfois assez éloignées de la réalité et qu'elles ne sont pas toujours spatialement cohérentes. Deux exemples en donnent une illustration. D'une part, les caractéristiques de sol ont été considérées comme homogènes à l'échelle des grands types de sol définis par la carte des sols. Elles ne tiennent donc pas compte des variations des propriétés des sols à des échelles plus fines qui sont responsables d'une partie de la variabilité du fonctionnement hydrique des couverts et qui conditionnent en partie la variabilité spatiale des températures de surface observées. Ainsi, les valeurs de  $f_2$  simulées par ISBA dépendent essentiellement d'un forçage climatique appliqué de manière uniforme à la zone d'étude, de caractéristiques des sols appliquées de manière homogène à l'échelle de chacun des types de sol de la zone, et d'un forçage d'indice foliaire pour chaque pixel. C'est ce dernier forçage qui seul détermine la variabilité spatiale des simulations pour les pixels ayant le même type de sol. D'autre part, les forçages utilisés pour les simulations n'incluent pas les irrigations, dont les occurrences n'étaient pas connues et qui si elles l'étaient devraient nous amener à modifier fortement les valeurs d'information a priori dans certaines situations (cf. Olioso et al., 2005 ; 2006).

Comme pour l'information a priori sur  $z_0$ , les erreurs de l'information a priori de  $f_2$  sur chaque pixel sont supposées suivre des lois normales centrées. L'écart type de cette erreur ( $\sigma_{f_{2bi}}$ ) relié au paramètre  $f_{2i}$ , est difficile à fixer. Une piste est de le fixer en fonction du type de culture et de sol ou en fonction des conditions de la parcelle. Nous avons ainsi fixé les  $\sigma_{f_{2bi}}$  à des valeurs égales sur les pixels composants une même parcelle, et cette valeur a été calculée en prenant l'écart type du paramètre  $f_2$  sur les différents pixels de la parcelle (cf. Fig. II.5).

Etant données les limites décrites ci-dessus concernant les valeurs de  $f_2$  fournies par le modèle ISBA, il a été décidé d'effectuer des tests supplémentaires avec une autre manière de calculer cette information a priori. Pour cette solution alternative, les valeurs a priori des humidités normalisées  $f_2$  ont été déterminées en utilisant la relation suivante décrite dans Courault et al. (2008) :

$$f_{2i} = \frac{T_{s_i} - T_{s \min_i}}{T_{s \max_i} - T_{s \min_i}} \quad (\text{II.54})$$

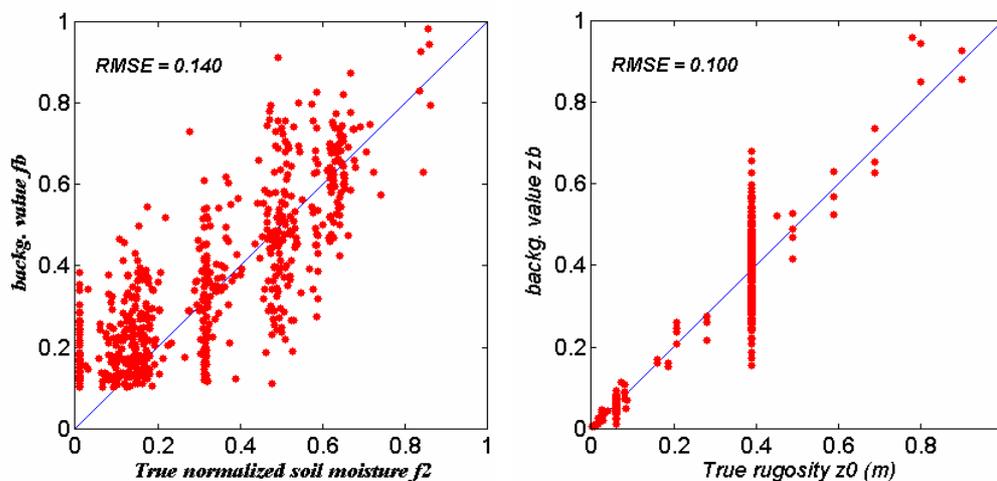
avec  $T_{s \min_i}$  et  $T_{s \max_i}$  les températures de surface minimale et maximale déterminées pour chaque pixel suivant la théorie développée pour le modèle S-SEBI dans l'espace 'température de surface – albédo' et présentée par Roerink et al. (2000) et Gomez et al. (2005) (cf. **Chapitre I, section 1.2.2**). D'autres définitions sont possibles, mais n'ont pas été utilisées ici, en travaillant dans l'espace 'température de surface - indice de végétation' connues sous les noms de méthode du trapèze ou du triangle  $T_s$ -VI ( $T_s$  - vegetation index) et décrites en détail dans les synthèses récentes de Li et al. (2009) et Petropoulos et al. (2009). Nous savons qu'en répétant un grand nombre de fois l'approche mais avec des erreurs différentes dues à des causes aléatoires, il est possible d'estimer des estimateurs statistiques tels que la moyenne, la variance ne dépendant que des processus physiques responsables des erreurs. Par conséquent, nous avons répété l'équation (II.54) dans une deuxième fois avec la température de surface obtenue par la simulation du modèle PBLs avec les valeurs de  $f_2$  déterminées la première fois. Nous avons réitéré plusieurs fois le processus et nous avons calculé l'écart type entre les différentes valeurs de  $f_2$  déterminées par l'équation (II.54). Les écarts type estimés constituent alors les composantes ( $\sigma_{f_{bi}}$ ) de la matrice de covariance d'erreur de l'information à priori ( $B$ ).

Comme pour  $z_0$ , les erreurs des informations a priori sur  $f_2$  sur les différents pixels sont supposées indépendantes. Là encore, des doutes peuvent être émis sur cette hypothèse, en particulier pour les pixels d'une même parcelle, mais encore une fois nous ne disposons pas des données nécessaires pour évaluer ces possibles corrélations et nous les négligerons donc dans cette étude. Enfin, les erreurs des informations a priori sur  $z_0$  et  $f_2$  sont également supposées indépendantes. Cela nous semble réaliste vu les différences entre les techniques utilisées pour fixer les valeurs des informations a priori pour ces deux paramètres.

### 5.4.2. Cas des expériences jumelles

Les valeurs des informations a priori sur les paramètres  $z_0$  et  $f_2$  définis pour l'application au cas réel sont les états vrais des expériences jumelles (pour  $f_2$  il s'agit des valeurs calculées avec la première méthode, i.e. les sorties du modèle ISBA). Les informations a priori dans le cas des expériences jumelles sont alors définies en ajoutant à ces états vrais des erreurs aléatoires indépendantes suivant des lois normales centrées de variance, les variances d'erreur d'information a priori fixées pour le cas réel (cf. section précédente). Les composantes de la matrice  $B$  sont fixées à ces valeurs théoriques (sauf dans quelques cas pour lesquels cela sera précisé).

La figure II.13 ci-dessous illustre la gamme de variation des valeurs a priori obtenue par rapport aux valeurs vraies des paramètres.



**Fig. II.13. a) Valeurs a priori de l'humidité normalisée  $f_2$  générées; b) Valeurs a priori de la rugosité aérodynamique  $z_0$  générées, par rapport aux valeurs vraies (définie à l'étape 1)**

### 5.5. Evaluations des résultats

Dans les différentes configurations et les différents tests effectués dans les expériences jumelles, les performances du système d'assimilation sont évaluées de différentes manières :

- en calculant les erreurs sur les paramètres estimés et variables de sortie du modèle PBLs obtenues après assimilation, par rapport aux états vrais correspondants (défini à l'étape 1), et en comparant ces erreurs avec celles des paramètres et variables disponibles avant assimilation (i.e. calculées à partir des valeurs de l'information a priori).
- en comparant les flux de surfaces ( $R_{ni}$ ,  $H_i$ ,  $LE_i$ ,  $G_i$ ) obtenus après assimilation à ceux obtenus avec la méthode résiduelle définie lors de l'analyse bibliographique dans la partie I de ce document (et décrite dans la suite de ce paragraphe) ; nous avons considérée ici cette méthode comme une méthode de référence car très utilisée dans le cadre des différents algorithmes de cartographie de l'évapotranspiration.

Dans le cas des données réelles, nous ne disposons bien sûr pas des états vrais des différents paramètres et variables. Nous avons cependant à disposition des mesures des flux de surfaces et températures de surface et de l'air à 2m pour 7 parcelles sur lesquelles avaient été disposées des stations de mesure, et ce pour différentes dates. Il s'agit de fait d'un jeu de données très réduit par rapport au nombre total de pixels traités mais qui permet tout de même une première évaluation des résultats de l'algorithme en conditions réelles. Nous avons dans ce cas, comme pour les tests sur les données synthétiques, comparé les erreurs (cette fois par rapport à des mesures, indépendantes de celles utilisées dans la procédure d'assimilation) avant et après assimilation afin d'évaluer le gain apporté par la procédure et donc son intérêt. Afin de compléter cette évaluation, nous avons également représenté et analysé les cartographies des flux de surface, notamment de l'évapotranspiration (flux  $LE$ ), sur l'ensemble de la zone considérée, avant et après assimilation.

La quantification des erreurs, que ce soit par rapport aux états vrais dans le cas des expériences synthétiques ou par rapport à des mesures indépendantes dans le cas de l'application aux données réelles, a été réalisée en calculant les erreurs quadratiques moyennes absolues et relatives (*RMSE* (Root Mean Square Error) et *RRMSE*) calculées selon les formules (II.55 et II.56) et la pente (II.56b) (Wallach, 2006 p17-p22):

$$RMSE(s, s_{true}) = \sqrt{\left( \frac{1}{n} \sum_{i=1}^n (s_i - s_{true_i})^2 \right)} \quad (II.55)$$

$$RRMSE(s, s_{true}) = \frac{RMSE(s, s_{true})}{s_{true}} \quad (II.56)$$

$$pente = \frac{\bar{s}}{s_{true}} \quad (II.56b)$$

avec :

- $n$  le nombre de pixels considérés,
- $s = (s_1, \dots, s_n)$  le vecteur contenant les quantités estimées sur les différents pixels,
- $s_{true} = (s_{true1}, \dots, s_{true_n})$  le vecteur contenant les valeurs vraies (expériences synthétiques) ou mesurées (cas des données réelles) correspondantes.
- $\bar{s}$  = la moyenne des composantes du vecteur

#### Modèle Résiduel

Pour évaluer les performances du modèle PBLs, nous avons également choisi de comparer ses résultats à ceux de la méthode résiduelle classique. Dans cette méthode, le flux de chaleur latente (*LE*) est déterminé à partir de l'équation du bilan énergétique (I.1) rappelée ici par l'équation (II.57).

$$LE = R_n - G - H \quad (II.57)$$

Les flux de surface  $R_n$ ,  $H$  et  $G$  sont obtenus en utilisant les équations (II.3) – (II.5) – (II.8) en considérant des valeurs de paramètre en entrée similaires à ceux qui sont utilisés dans les algorithmes d’assimilation. Dans le cas du modèle modifié (PBLs version 2), les équations équivalentes ont été introduites dans la méthode résiduelle (cf. **chapitre III section 2.1**).

Dans l’analyse sur les données synthétiques, le modèle résiduel a été appliqué de sorte à prendre en compte l’erreur d’observation supposée sur la température de surface ( $T_{sobs}$ :  $T_s$  vrai + erreur) dans le processus d’assimilation. Par rapport au modèle PBLs, la température de l’air atmosphérique ( $T$ ) et la hauteur de la couche limite de surface ( $h_{CLS}$ ) doivent être connues à l’heure du calcul (variables intervenant dans le calcul flux  $H$ ). Ici nous avons utilisé les sorties du modèle PBLs à midi solaire. Le Tableau II.4 ci-dessous montre les différences qui existent au niveau des variables d’entrée et sur la valeur de référence de certaines variables entre le modèle PBLs et le modèle résiduel. Les différences principales du modèle résiduel par rapport au modèle PBLs sont l’absence de la résistance de surface faisant intervenir les variables ( $f_2$ ,  $LAI$ ,  $R_{smin}$ ,  $R_{smax}$ ). Par ailleurs, notre utilisation de la méthode résiduelle cherchant surtout à apporter des éléments d’évaluation de l’assimilation de données dans PBLs, diffère de l’utilisation classique de la méthode. Dans la pratique générale, la méthode résiduelle ne tient pas compte des erreurs de mesure de température de surface. De plus, nous utilisons une connaissance précise de la hauteur de référence (prise à la hauteur de la couche limite de surface) et de la température de l’air à cette hauteur au moment du calcul. Dans la pratique, ces informations ne sont pas faciles à connaître précisément (dans le cas de SEBAL, elles sont estimées à partir de l’analyse de la variabilité spatiale des images).

**Tab. II.4. Comparaison des paramètres d’entrées des modèles PBLs et Résiduel**

	$T_{sobs}$	$z0$	$f_2$	$LAI$	$\alpha$	$R_{smin}/R_{smax}$	$u_a$	$R_g$	$h_{CLS}$	$T$	$q$
PBLs	non	oui	oui	oui	oui	oui	oui (constant)	oui (horaire)	oui (profil initial)		
Résiduel	oui	oui	non	non	oui	non	oui	oui (valeurs à l’instant du calcul)		non	

Les résultats obtenus avec le modèle résiduel décrit ci-dessus sont confrontés au processus d’assimilation, mettant en évidence l’analyse des erreurs des variables et la qualité des observations de télédétection sur les données d’expérience jumelle.

## 6. Outils TAPENADE et PALM

Comme vu dans la **section 3** ci-dessus, le calcul de l'opérateur adjoint du modèle est un point clé dans la mise en place d'un système d'assimilation. Nous avons également discuté de l'intérêt d'utiliser un logiciel de différentiation automatique pour générer cet adjoint. À ce jour, plusieurs logiciels sont développés pour différencier de façon automatique le code de calcul d'un modèle (implémenté soit en FORTRAN, en C, en C++, ou en MATLAB). Parmi eux, Nagware F95 (<http://www.nag.co.uk/nagware>), Adifor (Bischof et al., 1995), TAF, TAPENADE, et ADOL sont parmi les plus connus, mais il en existe maintenant bien d'autres (cf. <http://www.autodiff.org> pour une liste quasi-exhaustive). Nous n'avons pas dans ce travail réalisé de comparaison entre ces outils. Notre choix s'est porté sur le logiciel TAPENADE (Hascoët & Pascual, 2004) dont l'expérience d'utilisation sur plusieurs modèles au sein du laboratoire (Lauvernet, 2005 ; Lauvernet et al., 2008) a permis de montrer de bonnes propriétés de robustesse.

Un aspect important de la mise en place d'un système d'assimilation est son implémentation informatique. En effet, celle-ci consiste à assembler des éléments informatiques (modèle, minimiseur, modèle adjoint, ...) souvent pré-existant et développés de façon indépendante afin de constituer une application complexe. De la façon de réaliser cette implémentation dépendra fortement la souplesse et la durabilité du système d'assimilation. Nous avons pour cela choisi d'utiliser le logiciel PALM (Buis et al., 2006) qui a été développé spécifiquement pour l'implémentation de système d'assimilation.

Nous allons brièvement expliquer dans cette section le mode de fonctionnement de chacun de ces logiciels et nous en profiterons pour aborder la question de la validation du modèle adjoint.

## 6.1. Le logiciel de différenciation automatique TAPENADE

### 6.1.1. Principe de fonctionnement

TAPENADE est un outil logiciel d'analyse statique et de transformation de programmes qui, à partir d'un programme informatique donné calculant le résultat de l'application d'une fonction  $F$ , construit un nouveau programme qui calcule, pour un point donné de l'espace des entrées de la fonction, le produit entre la Jacobienne de cette fonction  $F$  (c'est le mode direct ou mode tangent), ou sa transposée (c'est le mode inverse ou mode adjoint), avec un vecteur donné.

Soit  $P$  un programme informatique composé d'une suite de  $k$  instructions  $\{I_1, I_2, \dots, I_k\}$  prenant en entrée un vecteur  $X$ . Chacune de ces instructions implémente une fonction élémentaire  $f_i$  et le programme  $P$  implémente alors la fonction composée  $F = f_k \circ f_{k-1} \circ \dots \circ f_1$ . Le programme tangent  $P'$  calcule alors  $F'(X).dX = f'_k(X_{k-1}) \times \dots \times f'_2(X_1) \times f'_1(X_0) \times dX$ , où les  $f'_i$  sont les opérateurs jacobiens des fonctions  $f_i$  (qui sont supposées différentiables et telles que  $f_i(X_{i-1}) = X_i$ , avec  $X_0 = X$ ),  $F'(X)$  est la matrice Jacobienne de  $F(X)$ , et  $dX$  la direction de dérivation. Ainsi, lors de la différenciation d'un programme, celui-ci est essentiellement considéré comme une suite d'instructions élémentaires (affectations, opérations et fonctions élémentaires) et de structures de contrôle. En mode tangent, ces instructions élémentaires sont différenciées une à une et organisées dans le même ordre et avec les mêmes structures de contrôle (sauf si celles-ci dépendent de variables différentiables) que dans le programme original. Les instructions élémentaires du programme original sont également incluses dans le programme tangent afin de calculer les états intermédiaires  $X_1, \dots, X_{k-1}$ .

Ainsi, les lignes de code :  $c = a * b$

$$d = 2 * c$$

deviennent dans le code tangent :  $dc = a * db + da * b$

$$c = a * b$$

$$dd = 2 * dc$$

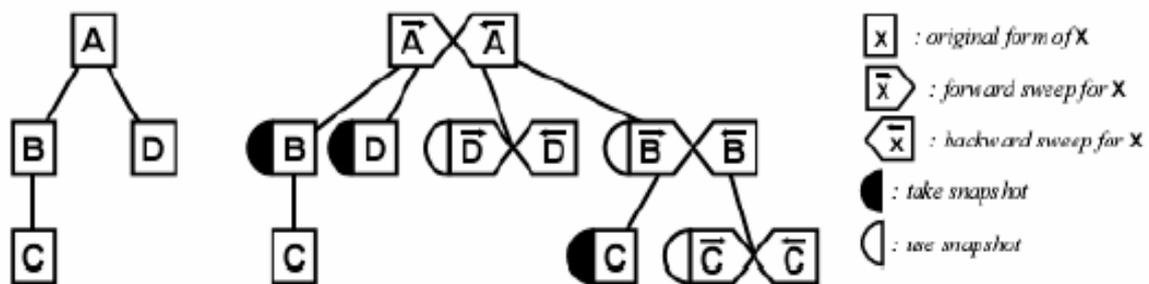
$$d = 2 * c$$

où  $da, db, dc$  et  $dd$  sont les variables différenciées de  $a, b, c$  et  $d$ .

Le programme adjoint  $P^T$  calcule  $F^T(X).dY = f_1^T(X_0) \times \dots \times f_{k-1}^T(X_{k-2}) \times f_k^T(X_{k-1}) \times dY$ . Ainsi, comme la transposition de fonctions composées inverse l'ordre de composition, l'ordre d'exécution des instructions dans un programme adjoint est inversé par rapport au programme original. De plus, l'exécution de ce programme adjoint nécessite les états intermédiaires  $X_k$  calculés par le programme à différencier, comme pour le programme tangent, mais dans un ordre inverse. Deux stratégies sont alors possibles pour réaliser cela :

- soit le programme à différencier est exécuté au préalable et l'ensemble des états intermédiaires nécessaires pour le programme adjoint sont stockés, puis récupérés un par un en ordre inverse par celui-ci,
- soit avant chaque instruction du programme adjoint l'ensemble du programme à différencier est exécuté jusqu'à obtenir l'état intermédiaire dont celle-ci a besoin.

Pour éviter une trop grande quantité de stockage ou un temps de calcul trop important, TAPENADE utilise un compromis entre stockage et recalcul des états intermédiaires appelé checkpointing. Il est appliqué par défaut à chaque routine ou fonction. Son principe est illustré sur la figure suivante tirée de la documentation de TAPENADE.



Cette figure présente à gauche un programme original composé de 4 fonctions ou routines. La fonction A appelle les fonctions B et D, et la fonction B appelle la fonction C. A droite, on peut voir l'arbre d'appel du code adjoint de la fonction A. Le terme « forward sweep » ou « passe avant » en français désigne l'exécution des fonctions du programme original auxquelles TAPENADE a simplement ajouté des points de stockages (snapshot) des variables intermédiaires avant l'appel aux sous-fonctions. Le terme « backward sweep » ou « passe

arrière » en français désigne l'exécution des instructions différenciées en mode adjoint (aussi appelé mode inverse). Le code adjoint de la fonction A est constitué de la passe avant de A puis de sa passe arrière. La passe avant de A effectue un stockage avant chaque sous-fonction appelée, mais ces sous-fonctions ne sont pas des passes avant, ce sont les fonctions du code original. Puis, lors de la passe arrière de A, les adjoints des fonctions D et B sont appelés. Ils sont eux même constitués d'une passe avant (qui nécessite la récupération de l'état intermédiaire stocké précédemment) et d'une passe arrière, ... et ainsi de suite. Au final, cette technique implique des phases de stockage de variables intermédiaires, mais la quantité maximale de mémoire utilisée est toujours inférieure à celle nécessaire à une stratégie de stockage total, et des phases de recalcul (on peut voir que chaque fonction du code original est exécutée plusieurs fois), mais moins qu'avec une stratégie basée sur un recalcul systématique.

Nous n'avons abordé ici qu'une petite partie du principe de fonctionnement du logiciel TAPENADE, celle qui était nécessaire pour la compréhension du chapitre IV notamment. Le lecteur intéressé par une description plus complète est invité à se référer à (Hascoët et Pascual, 2004).

### **6.1.2. Principe d'utilisation**

Pour différencier un programme, TAPENADE requiert :

- la liste des variables dépendantes, i.e. les variables de sortie du programme que l'on souhaite dériver,
- la liste des variables indépendantes, i.e. les variables d'entrée du programme par rapport auxquelles les dérivées seront calculées,
- le nom de la fonction principale à différencier et la liste des fichiers contenant l'ensemble du code à différencier,
- la sélection du mode de différentiation (tangent ou adjoint).

Par ailleurs, TAPENADE requiert que le programme à différencier soit mis sous la forme d'une fonction (ou d'une sous-routine si le programme est implémenté en FORTRAN) et que les variables dépendantes et indépendantes soient des arguments de cette fonction. Bien sûr, les variables indépendantes ne doivent pas être initialisées au sein de cette fonction (sans quoi elles

seraient équivalentes à des constantes). Bien souvent les modèles sont implémentés dans une fonction principale sans argument qui contient (de façon plus ou moins structurée et regroupée selon les cas) les initialisations des différents paramètres et variables d'entrée (leurs valeurs étant soit fixées directement dans le code, soit lues dans un ou plusieurs fichiers d'entrées). Il est donc souvent nécessaire d'effectuer une restructuration du programme principal du code à différencier. Classiquement, on sépare alors le programme principal en trois sections : initialisations, calcul, gestion des sorties. La phase de calcul est placée dans une fonction ou sous-routine dont les arguments contiennent les variables pouvant être choisies comme dépendantes et indépendantes pour la différentiation. C'est cette fonction ou sous-routine qui sera différenciée.

Il n'est pas rare que TAPENADE n'arrive pas à générer le code différencié du premier coup, pour diverses raisons souvent liées à une programmation incorrecte ou peu robuste dans le programme à différencier. Il fournit, en sortie de la différentiation, une liste de messages d'avertissement et d'erreur assez détaillée et par ailleurs bien documentée. L'analyse de ces messages (qui peuvent concerner aussi bien la déclaration des variables que leur initialisation ou les structures de contrôle, etc. ...) permet alors d'effectuer les corrections adaptées et de remédier à ce type de problème.

Cependant, étant donné la complexité du problème de la génération automatique d'un code adjoint, il est possible que le code adjoint généré donne des résultats incorrects. Il est donc indispensable de vérifier son bon fonctionnement, mais cela nécessite un test particulier.

### **6.1.3. Validation du code de l'adjoint**

La validation d'un code adjoint est classiquement réalisée en 2 phases : on teste d'abord l'adéquation entre les résultats du modèle tangent et le calcul des dérivées par différences finies (test de Taylor), puis l'adéquation entre le modèle tangent et le modèle adjoint (test du produit scalaire).

Soit  $f$  une fonction continue et dérivable. La méthode des différences finies centrées permet d'approximer le calcul de la dérivée directionnelle  $f'(x).dx$  par la formule :

$$f'(x).dx \approx \frac{f(x + \varepsilon * dx) - f(x - \varepsilon * dx)}{2\varepsilon} \quad (\text{II.58})$$

avec  $x$  la variable de la fonction,  $\varepsilon$  le pas de calcul de la dérivée (qui doit tendre vers 0) et  $dx$  une direction de dérivation.

Le programme tangent de la fonction  $f$  est censé calculer de façon exacte cette quantité  $f'(x).dx$ .

Par ailleurs, la définition mathématique d'un opérateur adjoint est la suivante : soit un opérateur linéaire  $A$  de l'espace  $E$  vers  $F$ , l'adjoint de  $A$  est l'opérateur  $A^*$  tel que pour tout vecteur  $(x, y)$  dans  $E \times F$ , on a le produit scalaire :

$$\langle Ax, y \rangle_F = \langle x, A^* y \rangle_E \quad (\text{II.59})$$

Ainsi, les modèles tangent et adjoint sont vérifiés en testant les égalités :

$$\left\langle \frac{f(x + \varepsilon * dx) - f(x - \varepsilon * dx)}{2\varepsilon}, \frac{f(x + \varepsilon * dx) - f(x - \varepsilon * dx)}{2\varepsilon} \right\rangle \approx \langle f'(x).dx, f'(x).dx \rangle = \langle dx, f^T(x).f'(x).dx \rangle$$

où  $f'(x).dx$  est le résultat du programme tangent au point  $x$  et pour une direction  $dx$ , et  $f^T(x).f'(x).dx$  est le résultat du programme adjoint au point  $x$  et pour une direction  $f'(x).dx$ .

La première égalité permet de vérifier la validité du code tangent, et la deuxième celle du code adjoint par rapport au code tangent. Il faut noter que cette vérification est bien sûr totalement dépendante du point  $x$  pour lequel les calculs sont effectués et qu'elle devrait donc en théorie être effectuée en chaque point de l'espace du vecteur  $x$  (ce qui est impossible en pratique).

#### 6.1.4. Stratégie et outils de débogage avec TAPENADE

Si les tests de validation de l'adjoint ne donnent pas des résultats satisfaisants, ce qui n'est pas rare en pratique, il est nécessaire de suivre une stratégie adaptée pour trouver la ou les

sources d'erreur. Sans cela, le débogage d'un code adjoint peut se révéler très long et complexe.

Voici quelques pistes pour organiser cette stratégie, tirées des diverses sources de documentations sur TAPENADE et de l'expérience de son utilisation accumulée au sein du laboratoire depuis quelques années :

- 1) Tout d'abord, il est bien sûr indispensable d'avoir analysé l'ensemble des messages d'avertissement donnés par TAPENADE lors de la génération des modèles tangent et adjoint, et de s'être assuré qu'ils ne pointent pas une source d'erreur.
- 2) Le programme de test `testMemSize` dans le package `ADFirstAidKit` de TAPENADE permet de vérifier que les types informatiques utilisés ont bien des tailles correspondantes à celles considérées par TAPENADE (4 pour les entiers, 4 pour les réels et 8 pour les double par défaut).
- 3) Le modèle direct doit ensuite être vérifié : est-ce que son exécution répétée au sein d'un même exécutable donne des résultats identiques ? Est-ce que le code direct généré avec TAPENADE (option *preprocess*) donne des résultats identiques à ceux du programme original ?
- 4) Les options d'optimisation par défaut de TAPENADE voire celles du compilateur peuvent générer des transformations de programme qui conduisent à des erreurs : on peut les inhiber avec les options *-nooptim* (pour TAPENADE) et *-g* (pour les compilateurs).
- 5) Les résultats obtenus par différences finies sont-ils fiables ? Il faut noter que les différences entre les dérivées obtenues par différences finies et les dérivées obtenues avec le code tangent peuvent très bien venir non pas d'un bug (défaut) dans le code tangent, mais de différences finies instables (comportement chaotique du système à très fine échelle en raison de discontinuité, d'arrondis, ...) ou de seuils dans le modèle. Il est donc bon de vérifier la stabilité des résultats obtenus par différences finies autour de la zone d'intérêt (on peut comparer différences finies à droite, à gauche et centrée, ou regarder la variation des sorties du modèle ou des dérivées obtenues par différences finies en chaque point d'une petite plage de variation du paramètre d'entrée concerné).

Si le problème persiste et si le modèle tangent ne donne pas des résultats satisfaisants, alors il est préférable de commencer par chercher les erreurs dans ce code plutôt que dans le code

adjoint, car il est beaucoup plus simple à comprendre et à débogger. Pour cela, TAPENADE fournit des outils de débogage qui permettent de réaliser le test de Taylor et de l'adjoint sur des morceaux choisis du programme (instructions, fonctions, ...). Pour la version de TAPENADE que nous avons utilisé (v2.2) ces outils n'étaient pas complètement finalisés, mais leur utilisation sur des programmes pas trop complexes, tels que PBLs, était possible, et il furent d'ailleurs d'un précieux secours (cf. **section IV.1**).

### **6.1.5. Performances d'un code adjoint**

En théorie, le nombre d'opérations à réaliser dans un code adjoint est au maximum de quatre fois supérieur au nombre d'opérations à réaliser dans le code direct correspondant (Talagrand, 1997). En effet, la dérivation de termes non-linéaires multiplie le nombre d'opérations à réaliser. Cependant, le coût réel peut-être encore supérieur étant donné le stockage des états intermédiaires du modèle direct qui doit être réalisé. De plus, si la différenciation manuelle permet à l'auteur de réaliser des optimisations judicieuses en fonction des cas (stockage ou non de variables intermédiaires, élimination de code inutile, optimisation du calcul en cas de fonction linéaire, ...), ce genre d'opérations est plus difficile à réaliser pour un différentiateur automatique (même si de gros efforts sont fait dans ce sens). Ainsi, Hascoët et Pascual (2004) ont montré les performances des codes différenciés par TAPENADE sur plusieurs exemples provenant de différents domaines d'applications réelles. Pour les modèles testés, ce ratio varie entre 2 pour le modèle LIDAR (optique) et 10 pour le modèle MARGARET (nucléaire) (le ratio de 19 donné pour le modèle de culture STICS a été ramené à 8 récemment).

## 6.2. PALM : un coupleur de code de calculs pour l'assimilation de données

Le logiciel PALM (Projet d'Assimilation par Logiciel Multi - méthodes) est développé par le CERFACS (Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique). C'est un logiciel qui permet d'assembler des applications complexes de façon modulaire et parallèle à partir de composants logiciels indépendants (Buis et al. , 2006). Le principe sous-jacent est proche de celui utilisé pour le couplage de modèles. Toutefois, l'approche de couplage a été étendue pour pouvoir gérer n'importe quel type d'application. En effet, la plupart des coupleurs de codes sont statiques : ils lancent les modèles à exécuter au début de l'application et pour toute sa durée. PALM est un coupleur dynamique dans le sens où il permet de lancer les composants à coupler de façon concurrente ou successive, dans des boucles ou de façon conditionnelle. En effet, le schéma de couplage est créé avec toute la souplesse d'un langage de programmation classique. Cela permet de pouvoir composer n'importe quel algorithme par couplage de ses opérations élémentaires. Le logiciel gère l'exécution de l'algorithme et l'échange de données entre les composants.

Ce logiciel a d'abord été créé pour des applications d'assimilation de données (Lellouche et al., 2002 , Massart et al., 2007). En effet, les différentes méthodes d'assimilation utilisent plus ou moins les mêmes opérateurs ou programmes (modèle, adjoint, matrices de covariance d'erreurs, minimiseur, etc....) mais dans des combinaisons différentes (Lagarde et al., 2001). Le passage d'une méthode à une autre est alors fortement facilité par une implémentation modulaire qui considère chaque opérateur comme étant un programme indépendant des autres (El Akkraoui et al., 2008). D'autre part, les systèmes d'assimilation sont bâtis autour de modèles qui peuvent être complexes, parallèles, difficiles à intégrer dans un autre programme et qui sont souvent développés indépendamment des systèmes d'assimilation. Le concept de couplage permet de les intégrer et de les faire évoluer facilement même au sein d'un système complexe tel qu'une application d'assimilation de données tout en gérant au mieux le parallélisme de l'application.

Dans le langage PALM, les composants à coupler sont appelés *unités*. Ces unités sont assemblées en séquences appelées *branches* qui représentent l'algorithme de l'application. Les branches et les unités peuvent échanger des données, appelées *objets*, via des *communications*.

La création du schéma de couplage (branches et communications) est effectuée grâce à l'interface graphique PrePALM (Fig. II.14).

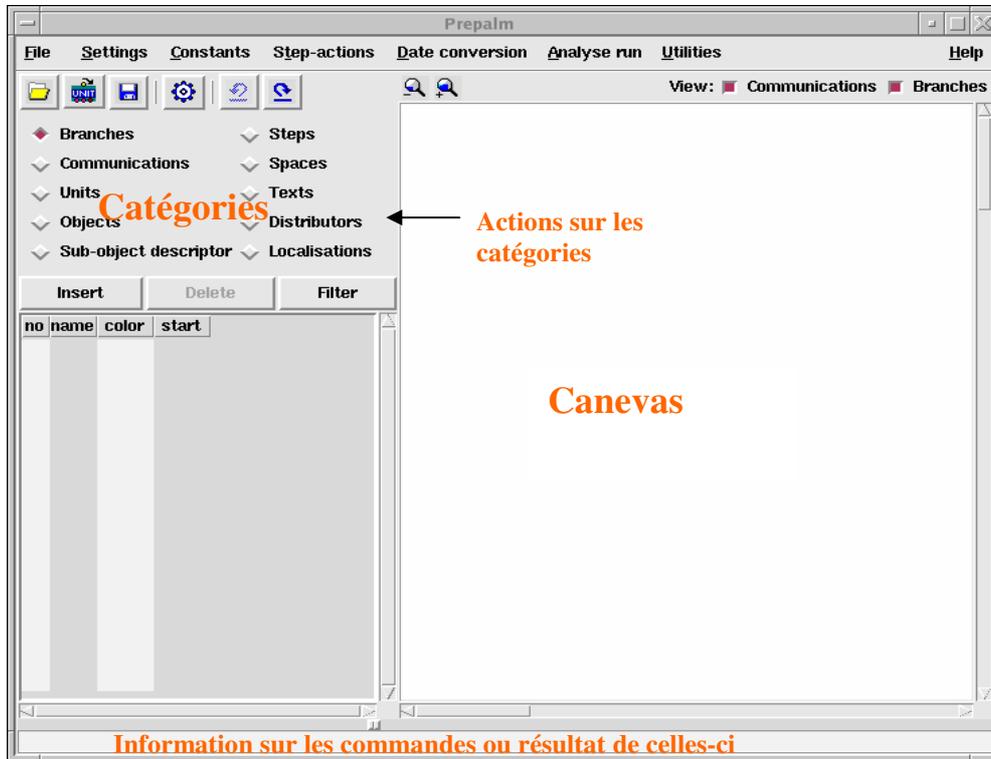
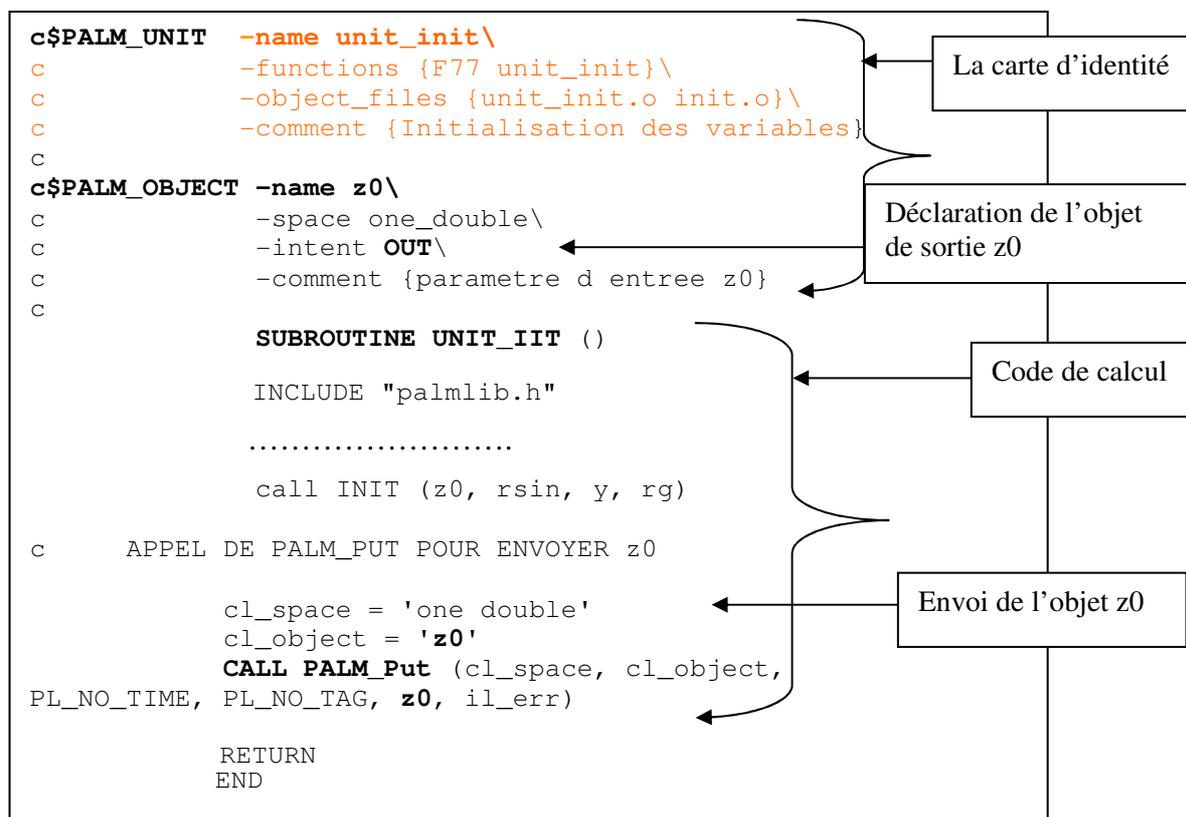


Fig. II.14. Interface graphique PrePALM

### 6.2.1. Unité

Une unité est une fonction ou sous-routine sans argument, implémentée en C, C++ ou FORTRAN, ou un exécutable. Une unité doit être vue comme une tâche élémentaire et indépendante des autres, qui consomme, transforme et produit des données. Toutefois, la granularité du couplage est libre, et chaque unité peut correspondre à un code de calcul complet qui peut même être parallélisé. Les données échangées par une unité sont appelées *objets* et sont décrits par des métadonnées. Ces métadonnées ainsi que diverses informations (langage de programmation, nom de la routine principale, ...) sont décrites dans une carte d'identité associée à chaque unité. L'échange d'objets est effectué via l'appel à des fonctions PALM (PALM\_Put, PALM\_Get) lorsque ces données sont nécessaires ou disponibles au sein d'une

unité. Les codes à coupler doivent donc être instrumentés. Cependant, afin d'assurer l'indépendance entre unités, la source d'un objet consommé et la cible d'un objet produit ne sont pas précisés dans le code de l'unité, mais dans le schéma de couplage propre à chaque application. La figure II.15 présente un exemple d'unité implémentée en FORTRAN.

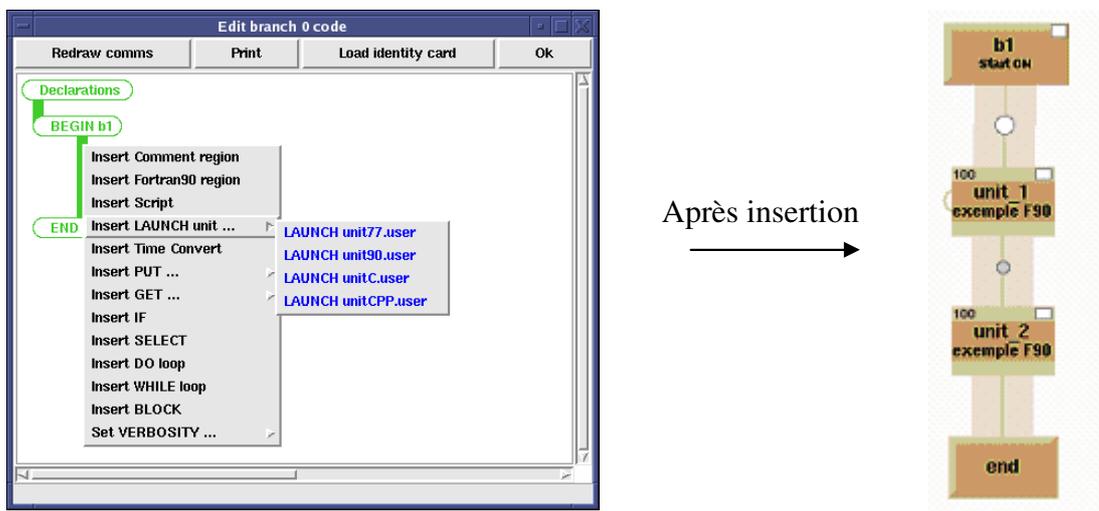


**Fig. II.15. Exemple du code d'une unité PALM**

Par ailleurs, PALM propose des unités prédéfinies pour réaliser des opérations d'algèbre sur les objets échangés par les unités au sein d'une application. Ces opérations vont de simples sommes et produits de vecteurs ou matrices jusqu'à des solveurs de problèmes linéaires directs ou itératifs, de problèmes de valeurs propres ou des minimiseurs. Ces unités sont en fait des interfaces à des bibliothèques d'algèbre (BLAS, LAPACK, ...), ou des codes disponibles gratuitement et largement utilisés dans la communauté scientifique.

### 6.2.2. Branche

Une branche est une séquence pouvant contenir des lancements d'unités et d'autres branches, des structures de contrôle et des échanges d'objets, le tout avec la souplesse d'un langage de programmation classique. Les branches sont créées dans l'interface graphique PrePALM. Une application peut contenir plusieurs branches et l'ensemble des branches représente l'algorithme de l'application. Des branches peuvent s'exécuter en parallèle et contenir des unités elle-même parallèles (et ce indépendamment les unes des autres). Ainsi, ce double niveau de parallélisme permet d'exploiter au maximum le parallélisme intrinsèque de l'application. Un exemple de création d'une branche sous PrePALM est illustré figure II.16.

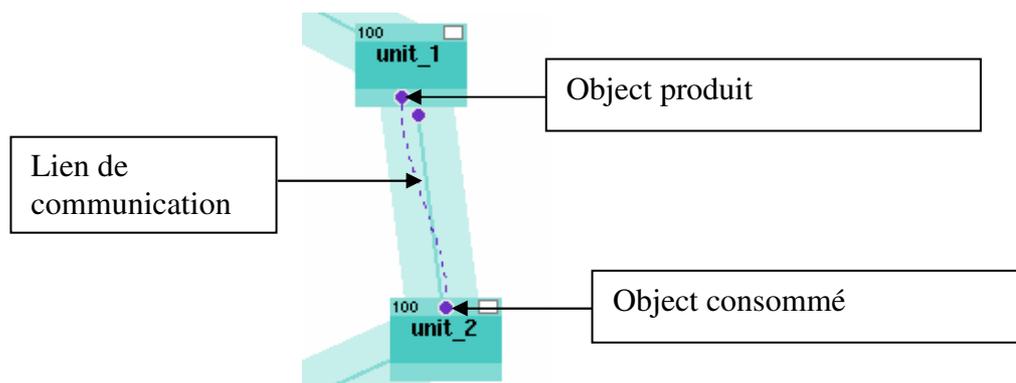


**Fig. II.16. Exemple d'insertion d'unités dans une branche :** à gauche une image de l'éditeur du code de branche dans PrePALM, qui permet de créer et modifier les branches, et à droite, une image d'une branche contenant deux unités (la première étant exécutée au sein d'une boucle), dans le canevas de PrePALM

### 6.2.3. Communication

Une communication matérialise l'échange de données entre deux ou plusieurs unités. Elle est créée dans PrePALM lors de la conception du schéma de couplage, et symbolisée par un fil.

Ce fil permet de relier un objet envoyé par une unité vers une ou plusieurs unités qui attendent de recevoir un objet compatible. Lors de l'exécution de l'application, une communication n'est effectuée que si l'unité source produit l'objet par appel à la primitive PALM\_Put, l'unité cible consomme l'objet par appel à la primitive PALM\_Get et le lien entre les deux a été matérialisé dans PrePALM. Bien sûr, l'appel au PALM\_Put n'est pas bloquant : si l'unité réceptrice n'a pas effectué le PALM\_Get correspondant, l'objet est envoyé et stocké temporairement. Les unités peuvent être parallèles, et ce indépendamment les unes de autres. Malgré cela, des données distribuées différemment au sein des unités source et cible peuvent être échangées. La redistribution des données est effectuée automatiquement sur la base de la description de leur distribution dans les cartes d'identité des unités. La figure II.17 ci-dessous montre une image d'une communication dans le canevas de PrePALM.



**Fig. II.17. Exemple de communication :** Les disques colorés en haut et en bas des unités représentent les objets en entrée et en sortie de ces unités. Leur couleur dépend du type informatique de l'objet associé. Lorsqu' aucune communication n'est associée à un objet, celui-ci est dessiné en dehors du rectangle qui représente l'unité.

#### 6.2.4. Compilation et exécution d'une application PALM

Une fois le schéma de couplage créé, l'interface graphique PrePALM permet de générer les fichiers et codes sources nécessaires pour l'exécution de l'application. Ces codes sources doivent être compilés et liés avec la librairie PALM, les sources compilées des unités et la

bibliothèque MPI qui permet de gérer les échanges de données. L'outil PrePALM permet de suivre en temps réel l'exécution de l'application et d'analyser ses performances a posteriori.

Pour plus de détails sur le fonctionnement et l'utilisation du logiciel PALM, le lecteur est invité à se référer à (Morel and the PALM group, 2006).



***TROISIEME PARTIE :***

***APPLICATIONS, ANALYSES ET MODIFICATIONS DU  
MODELES PBLs***



## 1. Analyse de sensibilité du modèle PBLs

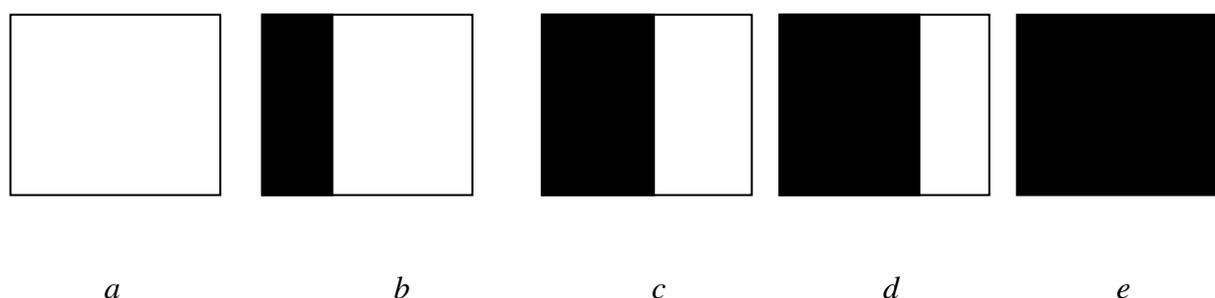
L'approche patchée ou 'tiled approach' consiste à prendre en compte les interactions entre la surface hétérogène, et l'atmosphère. En effet, les flux de quantité de mouvement et de chaleur à l'interface sol - atmosphère déterminent la hauteur de la couche atmosphérique. En termes d'interactions sol – atmosphère, la couche limite atmosphérique a un effet rétroactif sur les flux en surface, la température ou la concentration en vapeur d'eau de la couche déterminant les gradients, donc les flux. L'un des intérêts de PBLs avec l'approche patchée est de représenter au sein du modèle la composition de la zone agricole (l'ensemble des parcelles cultivées, et des zones non cultivées) contribuant à l'évolution de la couche limite atmosphérique. Dans cette section, nous avons analysé l'influence de l'hétérogénéité du domaine agricole sur l'évolution de la hauteur de la couche limite, ainsi que sur le calcul des flux et des températures de surface. Cette analyse est confrontée aux résultats du modèle PBLs dans sa version initiale qui considérait une surface homogène (~ 1 patch).

Par ailleurs, le modèle PBLs devrait être en partie destiné au suivi de grandes surfaces, avec des mesures de température ou des caractéristiques de surface faites à partir de plateformes aéroportées ou satellitaires. Une mise en œuvre opérationnelle s'accompagnerait donc d'un certain nombre de dégradations des conditions d'utilisations liées à la mesure de télédétection (erreur due aux instruments, mauvaise connaissance des caractéristiques des parcelles). Nous traitons de manière globale ces dégradations, en supposant que les variables d'entrée sont assorties d'erreurs de mesure et que les paramètres présentent une certaine incertitude. Nous ne mènerons pas une étude exhaustive, mais nous présenterons des illustrations des répercussions que la dégradation des conditions d'utilisation du modèle (sans assimilation) induit sur l'estimation de l'évapotranspiration.

## 1.1. Effet du ‘tiled approach’ et de la prise en compte des rétroactions sur surface hétérogène – atmosphère à l’échelle du paysage agricole

### 1.1.1. Illustration sur un paysage simplifié

Pour illustrer l’interaction entre les différents types de cultures, nous avons simulé une zone agricole constituée de deux patchs (pixels), dont nous ferons varier le pourcentage d’occupation de la surface suivant le schéma ci-dessous (Fig. III.1).



**Fig. III.1. Différentes configurations de surface étudiées**

Nous avons représenté dans cette configuration une culture de blé (wheat) sèche et une culture de luzerne (alfalfa) irriguée. Le tableau III.1 montre les caractéristiques des deux pixels choisis. Ces caractéristiques des cultures sont quasi similaires aux jeux de données Alpilles décrits dans le chapitre précédent (cf. partie II Fig. II.5)

**Tab. III.1. Caractéristiques des pixels de la surface étudiées**

Cultures	$z0_m$	$LAI$	$f_2$	
Blé	0,07	0,93	0,18	sèche
Luzerne	0,06	3,03	0,62	humide

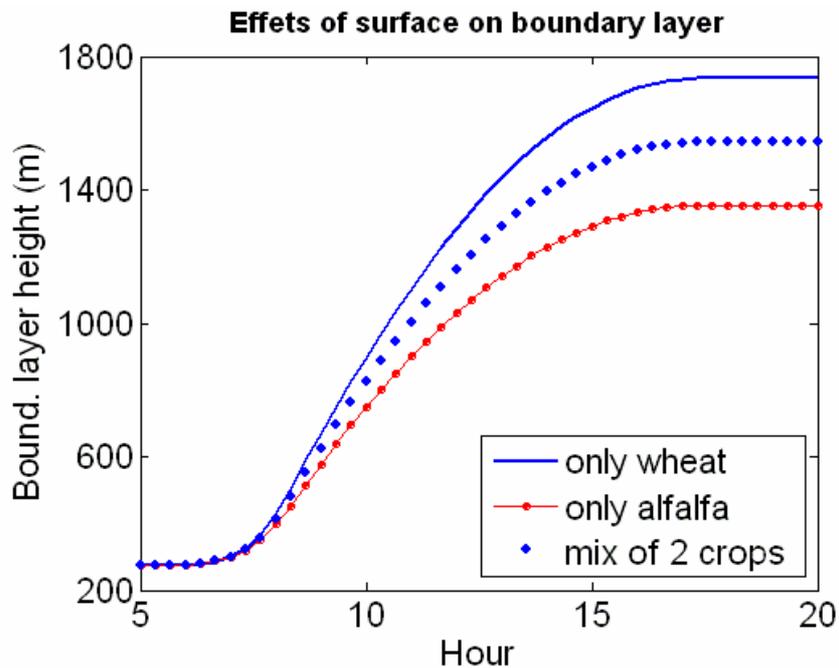
La prise en compte des effets rétroactifs de la couche limite soumise à l’action de la zone hétérogène avec l’approche patchée est mise en évidence en comparant les résultats à ceux d’une parcelle homogène. Le tableau III.2 ci-dessous montre le flux  $LE$  total sur la zone

suivant la configuration de l'occupation des surfaces (cf. Fig. III.1) avec le modèle PBLs patché. Ce résultat est confronté au modèle PBLs dans sa version initiale (PBLs-homogène ou PBLs-1 patch) montrant l'influence des interactions des pixels voisins. Le flux  $LE$  total est obtenu avec le modèle PBLs-homogène en pondérant, suivant le pourcentage d'occupation de surface, les flux  $LE$  obtenus indépendamment sur chacune des cultures homogènes sèche et humide.

**Tab. III.2. Comparaison de PBLs patché et PBLs-homogène sur le flux  $LE$  total à 12h suivant les proportions de zone sèche et humide.**

Proportion sèche - humide	75% 25%	50% 50%	25% 75%
PBLs-homogène ( $Wm^{-2}$ )	143	185	226
PBLs patché ( $Wm^{-2}$ )	157	191	241
Ecart ( $Wm^{-2}$ )	+ 14 (9%)	+ 6 (3%)	+15 (6%)

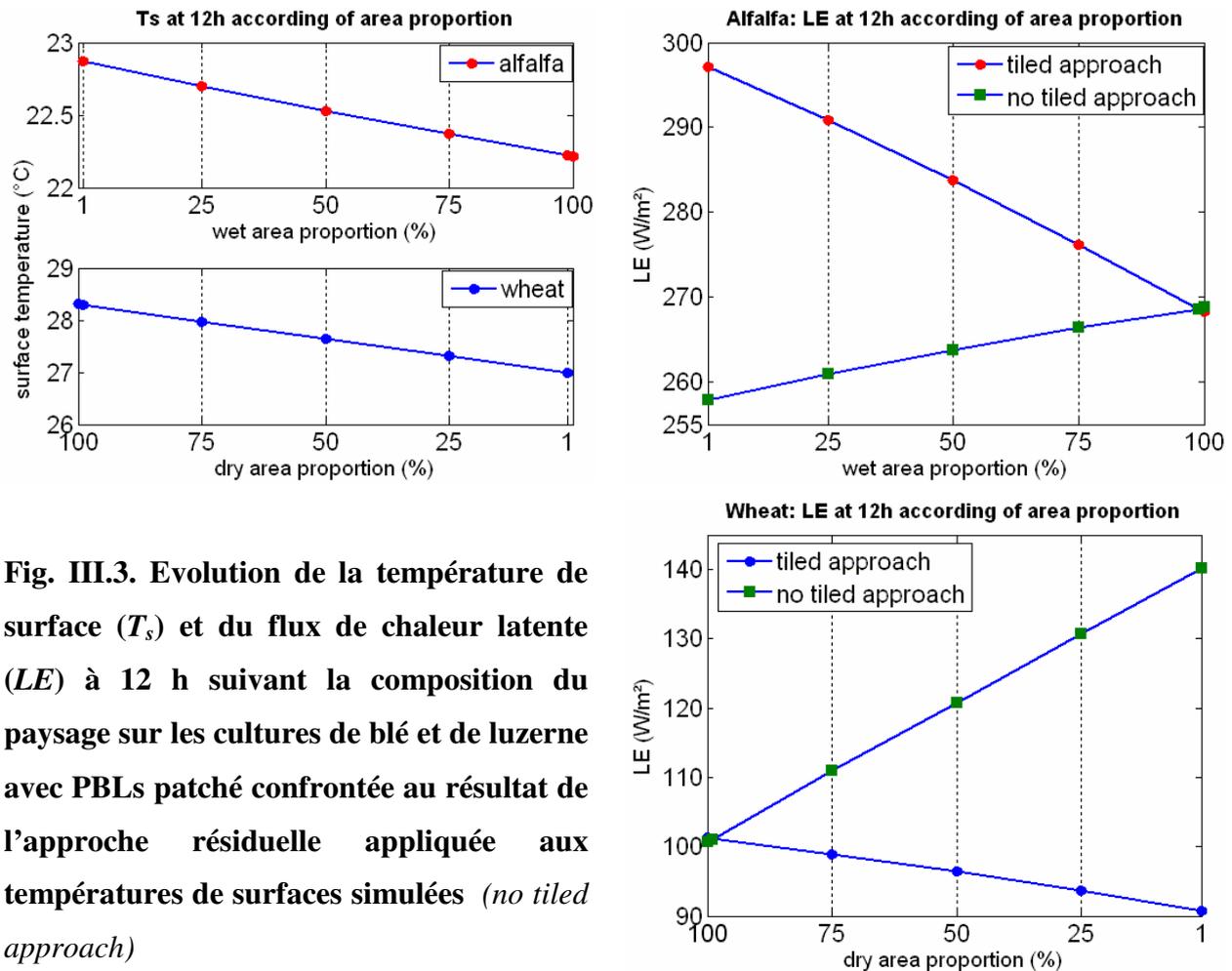
Les résultats du tableau ci-dessus montrent bien qu'il existe une interaction entre les deux zones de culture lorsqu'elles sont considérées simultanément dans la version patchée. La prise ou la non prise en compte de ces effets rétroactifs de la couche limite atmosphérique dans le calcul des flux engendre une différence qui peut atteindre jusqu'à  $15 Wm^{-2}$  dans ce cas étudié. Ces effets de rétroactions entre la surface et l'atmosphère sont illustrés par la figure III.2 montrant l'évolution de la couche limite atmosphérique d'une part par l'action indépendante d'une seule zone et d'autre part par l'action combinée des deux pixels dans la configuration 50% - 50%.



**Fig. III.2. Evolution de la hauteur de la CLA au cours de la journée suivant la configuration de la zone agricole à 100% pour une seule culture simulée et 50 % - 50 % pour les deux cultures présentée en même temps dans la simulation**

On peut remarquer que la hauteur de la couche limite atmosphérique varie en fonction de la composition du paysage et que l'hétérogénéité du paysage a des effets sur la simulation des flux de surface et sur les températures. Ces résultats permettent d'illustrer les raisons du choix de l'approche patchée pour notre travail (tiled approach).

Observons maintenant sur la Figure III.3 comment évolue le flux  $LE$  et la température de surface  $T_s$  lorsque l'on modifie les proportions des deux surfaces en suivant les configurations de la Figure III.1.



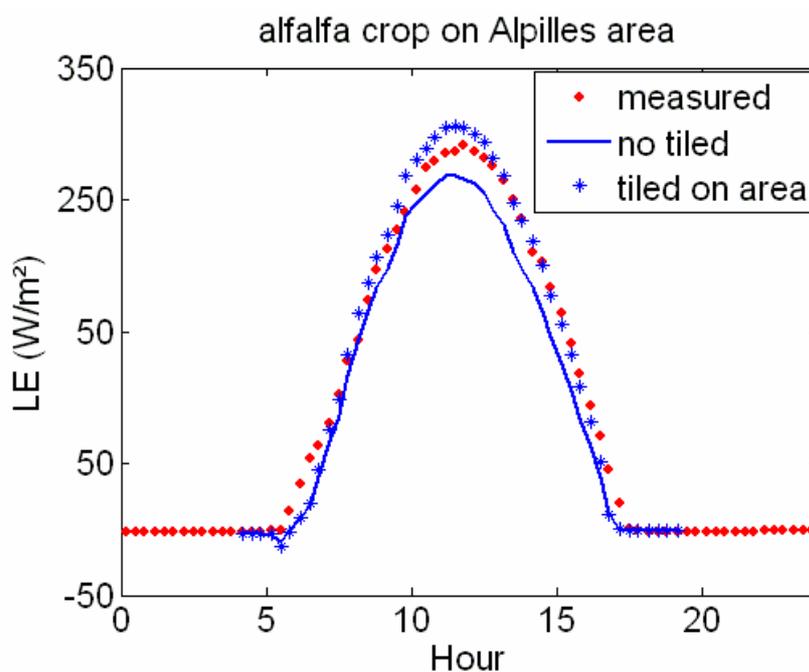
**Fig. III.3. Evolution de la température de surface ( $T_s$ ) et du flux de chaleur latente ( $LE$ ) à 12 h suivant la composition du paysage sur les cultures de blé et de luzerne avec PBLs patché confrontée au résultat de l'approche résiduelle appliquée aux températures de surfaces simulées (*no tiled approach*)**

Une première constatation concerne les variations de la température de surface de chacune des cultures (sèche et humide) en fonction de leur proportion. Les deux températures de surface diminuent quand la proportion de zone humide augmente (d'environ 0,7 K pour la luzerne (humide) et de 1,3 K pour le blé (sec)). Pour le flux de chaleur latente, on observe des variations opposées pour les deux cultures. Quand la proportion de culture humide augmente, le flux  $LE$  de la culture sèche augmente et celui de la culture humide diminue. La variation totale est plus importante sur la zone humide ( $30 Wm^{-2}$ ) que sur la zone sèche ( $10 Wm^{-2}$ ). Les états de surface de chaque zone restant les mêmes, toutes ces évolutions résultent des rétroactions surface – atmosphère. On peut supposer ici que l'augmentation de la proportion de culture humide tend au niveau de l'atmosphère à augmenter l'humidité de l'air et à diminuer la température de l'air, ce qui, par l'intermédiaire des processus turbulents, des variations de

déficit de saturation de l'air, de la régulation stomatique et du bilan d'énergie, conduit à modifier les résultats des calculs de flux de surface.

Il est également intéressant de constater que dans le cas de l'approche patchée et sur la luzerne, la température de surface et le flux de chaleur latente  $LE$  augmentent simultanément ce qui contredit le sens commun d'évolutions opposées de ces deux variables. Pour illustrer le possible impact de ces variations de température de surface sur le calcul des flux à partir de données de télédétection, nous avons calculé le flux  $LE$  à midi au moyen du modèle résiduel et en utilisant les températures de surface et de l'air simulées par PBLs pour chacune des cultures. Les résultats obtenus sont différents des simulations du modèle patché et varient en sens inverses (Figure III.3).

Pour illustrer la différence de quelques  $Wm^{-2}$  obtenue dans le calcul des flux de surface avec la prise ou la non prise en compte des rétroactions de l'atmosphère, nous avons comparé l'évolution horaire du flux latente  $LE$  de la culture de luzerne aux mesures acquises sur la parcelle 203 de la zone Alpilles du 18 avril 1997 suivant les deux approches (tiled approach & PBLs-homogène) (Fig. III.4). Dans le cas patché, le flux de chaleur  $LE$  est obtenu en simulant l'ensemble de la zone Alpilles (62500 pixels). On peut remarquer avec l'approche sans patch une sous estimation du flux  $LE$  autour de 12h même si l'allure générale du flux estimé est cohérente aux mesures.



**Fig. III.4. Flux de chaleur  $LE$  estimé par PBLs patché (\*) et PBLs 1 patch (-) comparé aux mesures de terrain (.) sur la parcelle 203 du 18/04/1997** (Dans le cas PBLs-1patch la proportion de surface est 100% et dans le cas PBLs patché l'ensemble de la zone Alpilles (62500 pixels) est représenté)

### 1.1.2. Illustration sur un vrai paysage (données Alpilles)

Avec l'approche patchée, on suppose que les variables de sortie du modèle PBLs sur un pixel ( $i$ ) sont influencées par les valeurs des paramètres sur les autres pixels (en considérant l'utilisation d'images de télédétection, nous attribuons un patch à chaque pixel). Une façon d'illustrer, voire de quantifier ces effets est de calculer les dérivées d'une variable de sortie du modèle en un pixel par rapport aux paramètres sur les autres pixels. Si ces effets existent alors on devrait avoir  $\frac{\partial Y_i}{\partial x_j} \neq 0$ , avec  $Y_i$  la variable de sortie considérée sur le pixel ( $i$ ) et  $x_j$  la variable d'entrée considérée sur le pixel  $j$ , avec  $j \neq i$ .

Les dérivées de la température de surface par rapport aux paramètres  $z0_m$  et  $f_2$  sont facilement obtenues dans notre cas à l'aide du modèle adjoint que nous avons généré. Les Figures II.5a et II.5b présentent ces dérivées sur un exemple simple à 10 pixels choisis parmi les données d'Alpilles. Nous pouvons constater sur cet exemple que la température de surface en un pixel donné (le 5 en l'occurrence) est influencée par les valeurs du paramètre  $f_2$ , et dans une moindre mesure du paramètre  $z0_m$ , sur tous les pixels. L'influence de chacun des pixels est certes faible dans ce cas par rapport à celle de  $f_2$  et de  $z0_m$  sur le pixel considéré, mais elle n'est pas négligeable. La somme des effets des paramètres  $f_2$  (resp.  $z0_m$ ) sur les pixels « voisins » représente dans ce cas plus de 27% (resp. 7%) du cumul total des effets de  $f_2$  (resp.  $z0_m$ ) sur la température de surface du pixel 5. Les valeurs négatives des dérivées par rapport à  $f_2$  expriment,

- sur le pixel considéré la baisse de température de surface lorsque l'humidité du sol augmente,
- et sur les autres pixels l'impact de l'augmentation d'évapotranspiration qui refroidit l'atmosphère, ce qui entraîne une baisse des températures de surface

Pour  $z0_m$ , des mécanismes différents sont impliqués puisque les dérivées des deux catégories de modèle sont de signe opposé :

- quand la rugosité augmente, localement la température de surface diminue en raison d'une plus forte turbulence qui favorise les échanges surface atmosphère
- l'augmentation de la rugosité sur les autres pixels entraîne une augmentation de la température de surface sur le pixel considéré par la modification des profils atmosphériques de température au dessus de la zone modélisée.

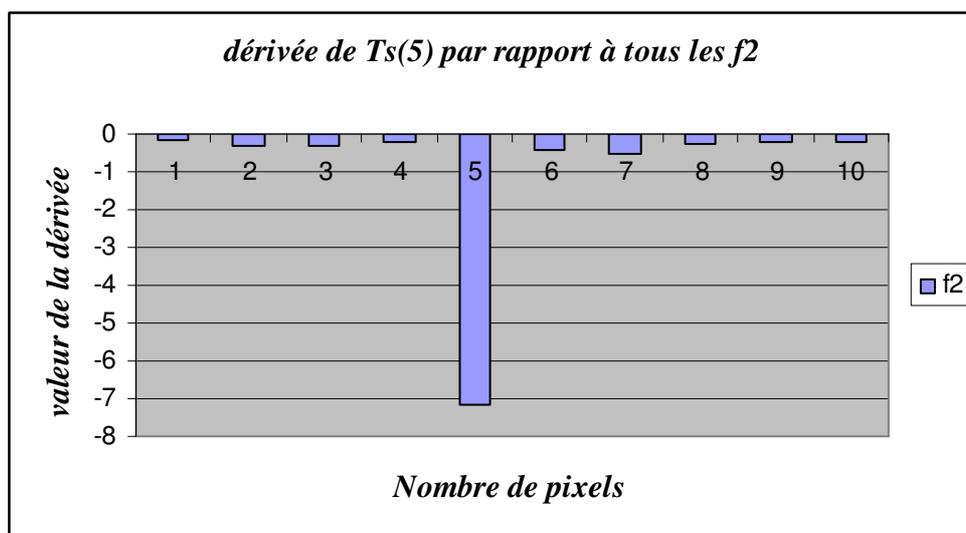
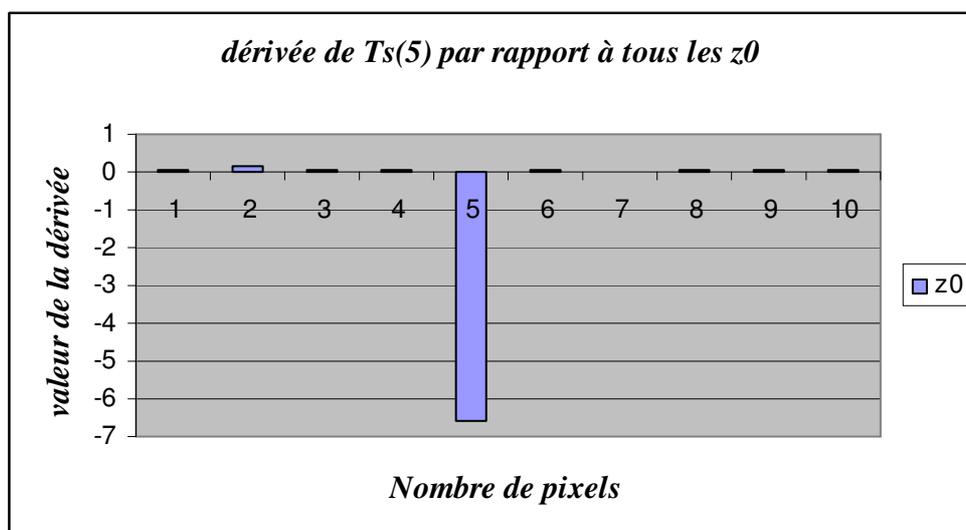


Fig. II.5a. Dérivée de la température de surface (à 12h) sur le pixel 5 par rapport au paramètre  $f_2$  sur les différents pixels, sur un cas à 10 pixels.



**Fig. II.5b. Dérivée de la température de surface (à 12h) sur le pixel 5 par rapport au paramètre  $z0_m$  sur les différents pixels, sur un cas à 10 pixels.**

Nous avons calculé ces dérivées sur toute la zone Alpilles pour la date du 18 avril 1997 pour la température de surface à 12h située sur un pixel au centre de la zone. Les figures II.6a et II.6b représentent les valeurs de ces dérivées. L'influence sur la température de surface associée au pixel (129,123), des paramètres  $f_2$  et  $z0_m$  associés aux autres pixels est très faible par rapport à celle des paramètres  $f_2$  et  $z0_m$  associés à ce pixel lui même. Cependant, la somme des effets des paramètres  $f_2$  (resp.  $z0_m$ ) sur ces pixels représente dans ce cas plus de 20% (resp. 30%) du cumul total des effets (en valeur absolue) de  $f_2$  (resp.  $z0_m$ ) sur la température de surface du pixel (129,123). Par ailleurs, on peut noter que les structures spatiales des effets de  $f_2$  et  $z0_m$  sont assez différentes.

Sur les deux cartes, la structuration spatiale liée aux parcelles agricoles est bien apparente. Sur la carte des dérivées par rapport à  $f_2$  s'y surimpose la structure spatiale de la carte des sols, qui a servi pour élaborer les informations nécessaires au calcul de  $f_2$  (point de flétrissement et capacité au champ). Dans le cas de  $z0_m$  il est possible au premier ordre de relier le niveau des dérivées aux valeurs de  $z0_m$ , les rugosités les plus élevées correspondant aux plus faibles dérivées (cas des bois et des vergers). Les parcelles en sol nu apparaissent pour la plupart dans les verts – verts-jaunes.

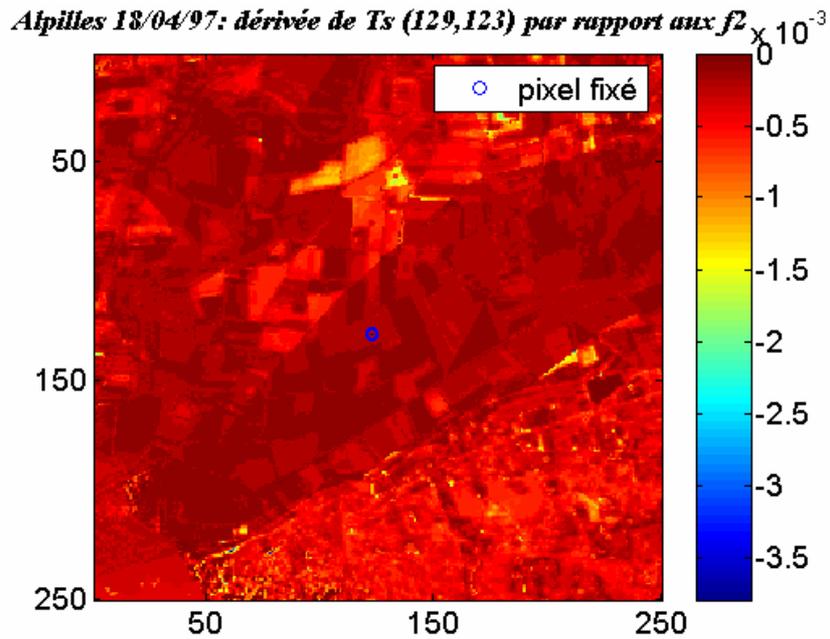


Fig. II.6a. Dérivée de la température de surface (à 12h) du pixel (129,123) par rapport à  $f_2$ . La dérivée de  $T_s$  (129, 123) par rapport à  $f_2$  (129, 123), qui vaut -6,73, n'est pas représentée sur cette carte afin de faire ressortir les structures spatiales des effets de  $f_2$  sur les autres pixels.

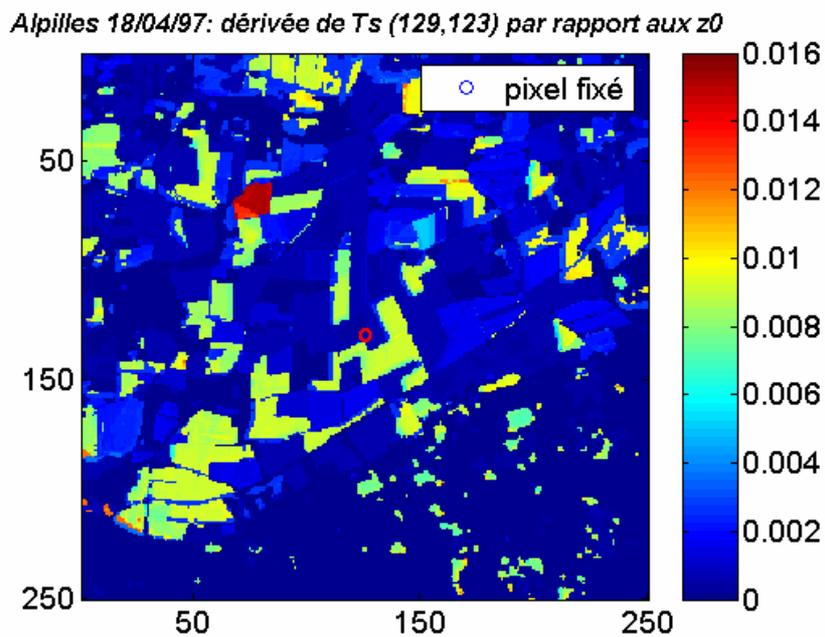


Fig. II.6b. Dérivée de la température de surface (à 12h) du pixel (129,123) par rapport à  $z_0$ . La dérivée de  $T_s$  (129, 123) par rapport à  $z_0$  (129, 123), qui vaut -23,19, n'est pas représentée sur cette carte afin de faire ressortir les structures spatiales des effets de  $z_0$  sur les autres pixels.

## 1.2. Sensibilité des variables d'intérêt du modèle PBLs aux erreurs sur ses variables d'entrée

L'étude de sensibilité est menée sur les données des expériences synthétiques définies dans la deuxième partie (cf. **section 5.2.2**). Pour évaluer l'impact de la dégradation des variables d'entrée du modèle PBLs sur ses sorties, nous avons quantifié les erreurs induites par ces dégradations sur ces sorties par rapport au scénario vrai. Comme défini à l'**étape 4** de la **section 5.2.2** de la **partie II**, nous avons évalué indépendamment l'impact des différentes erreurs des paramètres d'entrée sur les sorties de flux et de températures du modèle (Tab. III.3). Rappelons que les valeurs de ces erreurs sont tirées aléatoirement dans une distribution normale centrée et d'écart type donné, et ce de façon indépendante entre les différents pixels, sauf pour la vitesse moyenne journalière du vent  $u_a$  qui prend une unique valeur sur l'ensemble de la zone. 625 points sont utilisés dans l'analyse afin d'explorer toutes les gammes des valeurs des différentes variables d'entrée. Le modèle PBLs est exécuté à chaque fois supposant que les autres variables d'entrée du modèle sont vraies exceptée la variable dont on veut quantifier l'impact de son erreur sur les sorties de flux de surface, de température de surface et de température de l'air. Les erreurs sur le paramètre de l'humidité normalisée ( $f_2$ ) et de la rugosité aérodynamique ( $z_{0m}$ ) correspondent à celles des informations a priori utilisées dans les expériences d'assimilation de données.

**Tab. III.3. RMSE des écarts entre les sorties simulées par PBLs en prenant en compte des erreurs sur ses variables d'entrées et celles obtenues sans prendre en compte ces erreurs.** En colonne, les différentes variables d'entrée auxquelles des erreurs ont été ajoutées (avec l'écart type de ces erreurs), et en ligne les différentes variables de sorties considérées.

	$f_2$ $\pm 42\%$	$LAI$ $\pm 44\%$	$z_{0m}$ $\pm 25\%$	$\alpha$ $\pm 8\%$	$u_a = 0,8$ $-1 \text{ ms}^{-1}$	$u_a = 2,8$ $+ 1 \text{ ms}^{-1}$	$c_0$ $36\%$	$Tous \varepsilon$ $u_a = 0,8$	$Tous \varepsilon$ $u_a = 2,8$
$R_n \text{ (Wm}^{-2}\text{)}$	4,0	5,0	4,2	10,0	4,0	4,0	3,0	13,4	9,0
$H \text{ (Wm}^{-2}\text{)}$	36	42	8	7	6	7	28	71	67
$LE \text{ (Wm}^{-2}\text{)}$	39	47	5	3	3	4	6	60	59
$G \text{ (Wm}^{-2}\text{)}$	0,5	0,8	0,6	1,5	0,6	0,7	35	34	35
$T_s \text{ (K)}$	0,61	0,82	0,70	0,20	0,60	0,61	0,65	1,50	1,71
$T_a \text{ (K)}$	0,30	0,30	0,40	0,11	0,31	0,10	0,30	0,84	0,63

L'analyse de sensibilité présentée dans le tableau III.3, montre que l'ajout d'erreur sur le  $LAI$ , et sur le paramètre d'humidité  $f_2$  augmente de façon significative l'erreur du modèle sur les flux convectifs  $H$  et  $LE$ . Dans ce scénario, la plupart des valeurs de  $LAI$  sont faibles (cf. Fig. II.10c) donc l'ajout de l'erreur sur ces valeurs a été très forte. Ce qui traduit notre incapacité à fixer avec précision les faibles valeurs de  $LAI$ . De plus, d'une part, le principal terme  $r_s$  dans l'expression du flux  $LE$  est fonction de  $LAI$  et de  $f_2$  et d'autre part, les deux flux ( $H$  et  $LE$ ) se compensent, ainsi donc l'influence des erreurs sur le flux  $LE$  est observable en même temps sur le flux  $H$ .

L'impact de l'erreur sur le terme  $c_0$  est observé sur les flux  $H$  et  $G$ . Cela pourrait être expliqué par l'étude de Cellier et al. (1996) décrivant la partition de la chaleur sensible  $H$  entre le sol et l'atmosphère. Il est supposé que le flux de chaleur sensible ( $H$ ) et le flux conductif du sol ( $G$ ) se comportent de la même manière, car ils sont étroitement couplés à travers la température de surface.

Un autre point de remarque de l'analyse est que seul le flux  $R_n$  est très perturbé par l'erreur sur l'albédo ( $\alpha$ ) ce qui est normal car la valeur de l'albédo est liée directement au calcul de  $R_n$ . Cette observation est confortée par l'étude de Jacob et al. (2002b) qui ont montré que certaines sous-estimations de l'albédo peuvent produire une légère surestimation de rayonnement net.

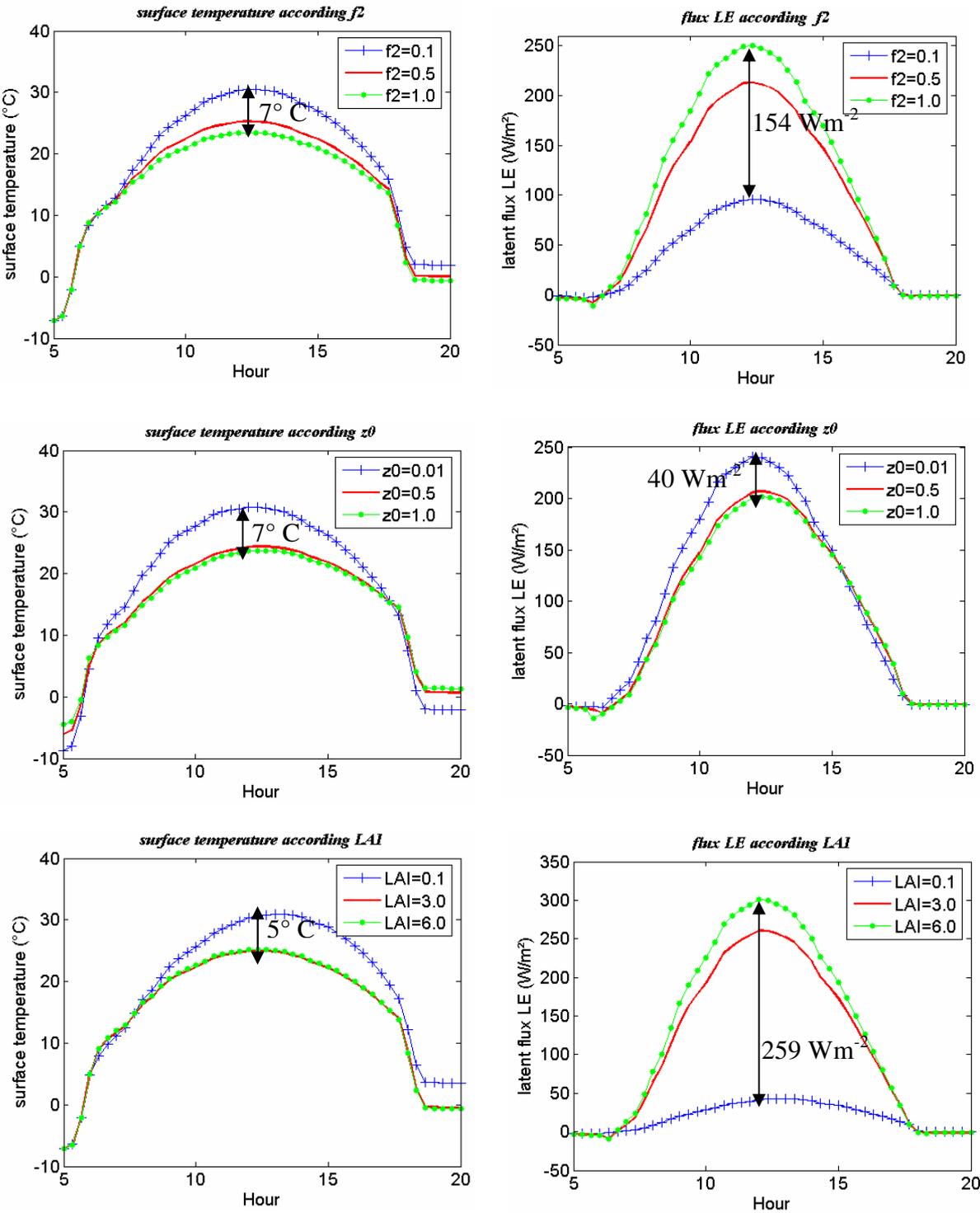
L'ajout d'erreur sur de la vitesse moyenne journalière du vent ( $u_a$ ) ainsi que la rugosité aérodynamique ( $z_0$  ou  $z_{0m}$ ) ont moins d'influence sur les flux par rapport aux autres erreurs. La combinaison de toutes ces erreurs engendre une dégradation importante des flux (particulièrement sur  $H$  et  $LE$ ) et de la température de surface. C'est pourquoi la mise en œuvre des procédures d'assimilation pour réduire l'impact de ces erreurs est nécessaire.

A l'analyse du tableau III.3, on peut observer que les erreurs de  $LAI$  et du paramètre  $f_2$  ont un impact très important sur la simulation de l'évapotranspiration. Dans la partie précédente du mémoire, nous avons justifié le choix des paramètres  $f_2$  et  $z_{0m}$  à rechercher par assimilation par la non possibilité de les mesurer aux échelles spatiales considérées, alors que pour le  $LAI$  il existe maintenant de bonnes connaissances permettant de l'estimer par télédétection avec une faible incertitude (e.g. Bsaïbes et al., 2009 ; CYCLOPES in Baret et al., 2007 et Weiss et al., 2007 ; Baccour et al., 2006 ; Soudani et al., 2006 ; Duchemin et al., 2006 ; Weiss et al., 2002 ; Weiss & Baret, 1999).

L'analyse de sensibilité ci-dessus montre cependant que la sensibilité des flux au paramètre de rugosité est bien moins forte que la sensibilité au  $LAI$  (par contre leur impact sur la température

de surface est fort pour tous les deux). Au vu de cette sensibilité, explorer l'idée d'assimiler le *LAI* pour améliorer l'estimation de l'évapotranspiration serait très pertinente. Cette piste ne sera cependant pas entreprise dans le cadre de ce travail.

La figure III.7 ci-dessous illustre l'évolution de la température de surface et du flux *LE* pour différentes valeurs des 3 paramètres  $f_2$ ,  $z_0$  et *LAI*.



**Fig. III.7. Evolution diurne de la température de surface (à gauche) et du flux de chaleur latente  $LE$  (à droite) pour différentes valeurs des paramètres  $f_2$ ,  $z0_m$  et  $LAI$ .**

On peut noter que à l'heure (à midi) où les données d'observation (température de surface) sont disponibles pour l'assimilation de données, l'écart de température entre les deux extrêmes (d'un état sec à un état humide) du paramètre  $f_2$  lié à l'état hydrique du sol est de  $7^\circ \text{C}$  avec une variation de plus de  $154 \text{ Wm}^{-2}$  sur le flux  $LE$ . Un écart similaire sur la température de surface est obtenu avec la variation de la rugosité aérodynamique ( $z0_m$ ) contre  $5^\circ \text{C}$  avec le  $LAI$ . Par contre l'impact de ces extrêmes (sol sec à sol humide ou sol nu à couvert développé) sur l'évapotranspiration est important avec le  $LAI$ . Il est de l'ordre de  $259 \text{ Wm}^{-2}$ . Notons également que les variations de flux sont en sens opposé des variations de température de surface pour  $f_2$  et  $LAI$ , alors qu'elles sont dans le même sens pour  $z0_m$ .

## 2. Modifications du modèle PBLs

L'analyse du tableau III.3 ci-dessus a montré qu'une dégradation de l'ensemble des paramètres d'entrée du modèle PBLs augmente considérablement l'erreur modèle sur l'estimation des flux de surface et du microclimat. Une partie de l'erreur modèle provient des hypothèses fortes faites dans la paramétrisation des équations. L'étude ci-dessous porte particulièrement sur deux équations qui, à la lumière de la bibliographie, nous apparaissent limitantes pour représenter correctement les flux de surface sur différents types de zones agricoles et pour une représentation correcte de la température de surface nécessaire aux procédures d'assimilation :

- le calcul du flux de chaleur par conduction dans le sol ( $G$ ) qui n'est représenté dans PBLs que par une proportionnalité simple au rayonnement net, alors qu'il dépend en théorie des caractéristiques thermiques du sol et de l'évolution temporelle de la température de surface ;
- le calcul de la rugosité thermique  $z\theta_h$  (intervenant dans le calcul de la résistance aérodynamique  $r_a$ ), définie dans le modèle PBLs comme une simple proportion de la rugosité aérodynamique  $z\theta$ , alors que des variations de plusieurs ordres de grandeur ont été constatés dans la bibliographie en fonction de différents facteurs (abordées en général par l'étude du « paramètre »  $kB^{-1}$  qui sera décrit plus bas).

Plusieurs pistes sont explorées dont l'analyse est décrite dans les sections suivantes.

## 2.1. Etude du rapport $G/R_n$

### 2.1.1. Positionnement du problème et introduction de l'analyse

Le flux de chaleur dans le sol est une variable difficile à quantifier et à spatialiser (Bastiaanssen et al., 2001). Il dépend fortement des propriétés du couvert et du sol, les premières étant généralement accessibles par télédétection, les secondes étant beaucoup plus complexes à connaître, en particulier dans leurs variations spatiales. L'approche utilisée pour estimer le flux  $G$  dans beaucoup de modèles (e.g. Choudhury et al., 1987 ; Norman et al., 1995 ; 2000) est d'estimer le flux comme une proportion ( $c_0$ ) du flux  $R_n$  en considérant cette proportion constante durant la journée. D'après la bibliographie elle varie entre 0,15 et 0,5 selon le développement du couvert végétal (e.g. Brutsaert, 1982 ; Clothier et al., 1986 ; Choudhury et al., 1987 ; Kustas & Goodrich, 1994). Toutefois, plusieurs études ont montré que le rapport  $G/R_n$  pouvait varier sensiblement au cours de la journée, et que considérer  $c_0$  comme une constante pouvait entraîner des sous-estimations du flux  $G$  dans la matinée et des surestimations dans l'après midi jusqu'à 50% (e.g. Santanello & Friedl, 2003).

Différentes approches ont été élaborées pour exprimer le flux  $G$  en fonction du rayonnement net et analyser les causes de variation du rapport  $G/R_n$ . Un récapitulatif de certaines expressions de  $c_0$  est présenté dans le tableau III.4 ci-dessous. Notons que certaines de ces expressions se valent plus ou moins et que par de simples manipulations mathématiques des équivalences peuvent être établies. C'est le cas des expressions considérant simplement le développement du couvert, par le  $LAI$ , la hauteur, le taux de couverture, ou des indices de végétation acquis par télédétection, qui sont souvent eux même bien reliés au  $LAI$  (**expressions 1-6**). Par exemple, Moran et al. (1994) ont exprimé le rapport  $G/R_n$  à l'aide du  $NDVI$  suivant une relation exponentielle ; Kustas et Daughtry (1990) ont proposé une relation linéaire entre les rapports  $G/R_n$  et  $\rho_{NIR} / \rho_{RED}$  (proche infrarouge, rouge) où  $\rho$  est la réflectance mesurée. Cherchant à améliorer la prise en compte de la variabilité spatiale, Bastiaanssen et al. (1998a) ont étendu l'utilisation de la télédétection en introduisant des mesures de température de surface et d'albédo dans l'expression de  $c_0$  (**expression 7**). D'autres approches très différentes ont été élaborées tout d'abord dans le cas des sols nus, en reliant par exemple  $c_0$  à la vitesse moyenne du vent qui a effectivement un impact sur la partition des flux de chaleur à la surface du sol (**expression 9**, Cellier et al., 1996). Par ailleurs, Santanello et Friedl (2003) et Murray et

Verhoef (2007) ont plutôt décrit le calcul en se basant sur une fonction sinusoïdale dépendant du temps dans laquelle la prise en compte de la couverture végétale est plus ou moins détaillée (**expression 8**). Pour l'ensemble de ces formulations, les coefficients sont déterminés empiriquement. Les résultats sont en général satisfaisants, mais de fortes dispersions peuvent être observées en particulier lorsque l'on s'éloigne des conditions d'établissement de la formulation.

#### Stratégie de l'analyse

Dans ce travail, nous recherchons une expression permettant de déterminer le flux  $G$  au sein de PBLs de manière moins approximative qu'une simple proportion constante de  $R_n$ . Nous souhaitons par ailleurs qu'elle permette une spatialisation facile du flux  $G$ . La possibilité d'accès aux données d'entrée et la possibilité d'intégrer l'expression dans PBLs sont donc deux aspects déterminants. L'analyse des expressions réunies dans le Tableau III.4 nous a conduit à séparer le cas des couverts végétaux plus ou moins développés de celui des sols nus (les deux cas ont été traités de manière indépendante dans la suite). Dans le premier cas, les **expressions 1 à 7** sont concernées, mais nous avons déjà signalé que les 6 premières étaient globalement similaires et nous nous sommes simplement intéressés aux **expressions 6 et 7**. Dans le cas des sols nus, nous nous sommes intéressés aux **expressions 8 et 9**. Pour les deux groupes d'expressions, une démarche similaire a été effectuée :

- les deux groupes d'expressions ont d'abord été adaptés au problème posé et en particulier à la disponibilité des informations nécessaires à leur fonctionnement (impliquant certaines modifications de la forme des équations)
- la calibration des expressions a été effectuée sur des données simulées par des modèles détaillés des transferts dans les sols (TEC et SiSPAT).
- les expressions calibrées ont été évaluées sur des données réelles mesurées dans le cadre du programme Alpilles-ReSeDA et sur le site Flux et Télédétection de l'INRA à Avignon.

**Tab. III.4. Différentes expressions du rapport  $G/R_n$**

Equations de $c_0$
<p><b>1.</b> <math>c_0 = \Gamma_c + (1 - f_c) \cdot (\Gamma_s - \Gamma_c)</math> (Su, 2002 ; Colin, 2006)  <math>\Gamma_c = 0,05</math> (surface couverte de végétation : Monteith &amp; Unsworth, 1990)  <math>\Gamma_s = 0,315</math> (surface de sol nu : Kustas &amp; Daughtry, 1990)  <math>f_c</math> : fraction du couvert végétal (Baret et al., 1995 ; Norman et al., 1995 ; Weiss et al., 2002)</p>
<p><b>2.</b> <math>c_0 = n_1 - n_2 NDVI</math> (Kustas et al., 1993). Avec <math>n_1 = 0,4</math> et <math>n_2 = 0,33</math></p>
<p><b>3.</b> <math>c_0 = r_1 - r_2 \frac{\rho_{NIR}}{\rho_{RED}}</math> (Kustas et al., 1994). Avec <math>r_1 = 0,36</math> et <math>r_2 = 0,02</math></p>
<p><b>4.</b> <math>c_0 = m_1 - m_2 h_c</math> (Reginato et al., 1985). Avec <math>m_1 = 0,1</math>, <math>m_2 = 0,042</math> et <math>h_c</math> la hauteur du couvert végétal</p>
<p><b>5.</b> <math>c_0 = c_1 \exp(-c_2 MSAVI)</math> (Sobrino et al., 2005 ; Jackson et al., 1987).  <math>c_1 = 0,5</math>, <math>c_2 = 2,13</math> et <math>MSAVI = \frac{2\rho_{NIR} + 1 - \sqrt{(2\rho_{NIR} + 1)^2 - 8(\rho_{NIR} - \rho_{RED})}}{2}</math></p>
<p><b>6.</b> <math>c_0 = \beta \exp(-\alpha_C LAI)</math> (Choudhury et al., 1987).  Avec <math>\beta</math> le terme sur sol nu qui varie entre 0,3 et 0,4 et <math>\alpha_C</math> est le coefficient d'extinction, qui est de l'ordre de 0,5 (Gomez et al., 2005) mais varie avec le type de végétation et l'angle zénithal solaire (Monteith, 1973 ; Ross, 1981).</p>
<p><b>7.</b> <math>c_0 = Coef_1 * Coef_2</math> (Bastiaanssen et al., 1998a)  <math>Coef_1 = \frac{T_s}{\alpha} \left( \bar{\eta} \bar{\alpha} + \delta \bar{\alpha}^{-2} \right)</math> avec <math>\bar{\alpha}</math> l'albédo moyen sur la journée, <math>\eta = 0,0032</math> et <math>\delta = 0,0062</math>  <math>Coef_2 = 1 - \mu NDVI^4</math> avec <math>\mu = 0,978</math></p>
<p><b>8.</b> <math>c_0 = A * \cos\left(\frac{2\pi(t+10800)}{B}\right)</math> (Santanello &amp; Friedl, 2003).  Avec <math>t</math> le temps en seconde, <math>A = 0,31 - 0,35</math> et <math>B = 74\ 000 - 100\ 000\ s</math></p>
<p><b>9.</b> <math>c_0 = \alpha_1 + \beta_1 u_a</math> (Cellier et al., 1996). <math>\alpha_1 = 0,57</math> et <math>\beta_1 = -0,054</math></p>

### 2.1.2. Cas d'un couvert végétal peu développé ou développé

Dans cette section, nous nous sommes intéressés aux **expressions 6 et 7** du Tableau III.4 renommées ci-dessous  $c_{01}$  et  $c_{02}$ .

✚ Equation de Choudhury et al. (1987) :

$$c_{01} = \beta \exp(-\alpha_C LAI) \quad (\text{III.1})$$

Cette expression est bien représentative de l'ensemble des expressions représentant le rôle du développement du couvert végétal sur le rapport  $G/R_n$  (e.g. **expressions 1-6**).

✚ Equation de Bastiaanssen et al. (1998a) :

$$c_{02} = Coef_1 \cdot Coef_2 \quad (\text{III.2})$$

Cette expression est celle qui est utilisée dans le modèle SEBAL (Bastiaanssen et al. 1998a). Elle a été développée pour essayer de mieux prendre en compte les différents facteurs qui influent sur le flux de chaleur dans le sol.  $Coef_2$  représente l'impact de la couverture végétale sur le rapport  $G/R_n$  suivant une expression équivalente aux **expressions 1-6** présentées ci-dessus. Ce facteur est présenté comme rendant compte de l'extinction du rayonnement par le couvert (valant 1 pour un sol nu).  $Coef_1$  représente le rapport entre le flux de chaleur dans le sol et le rayonnement net pour un sol nu. Sur la base d'une expression théorique, Bastiaanssen et al. (1998a) montrent que ce facteur augmente avec la température de surface et l'albédo du sol nu. Sur la base d'une analyse de données expérimentales, ils montrent également que la prise en compte de l'albédo journalier moyen permet de tenir compte de la teinte du sol et de l'humidité superficielle par l'intermédiaire de la brillance du sol. Toujours selon Bastiaanssen et al. (1998a), le produit des deux coefficients devrait permettre de représenter le rapport  $G/R_n$  dans un grand nombre de situations de développement de couverts végétaux.

#### ✚ Notre approche dans cette étude :

Les deux expressions ci-dessus ont déjà été évaluées dans leurs configurations d'origine (sans recalage des coefficients) à partir des images aéroportées du jeu de données Alpilles par Gómez et al. (2005) [**expression 6**] et Jacob et al. (2002) [**expression 7**] conduisant dans les deux cas à des résultats peu satisfaisants avec pour le premier une forte surestimation ( $RMSE$  de  $48 \text{ Wm}^{-2}$ ) et pour le second une forte sous-estimation ( $RMSE$  de  $44 \text{ Wm}^{-2}$ ). Leur application dans le cadre de cette étude devrait donc nécessiter au minimum un nouvel étalonnage. Par ailleurs, la détermination de l'albédo journalier moyen n'étant pas directe par télédétection, et qui plus est fortement déterminée par l'estimation de l'albédo instantané (Jacob & Olioso, 2005), une simplification de la forme de  $Coef_1$  est proposée. De la même façon, pour être homogène avec l'**expression 6** la formulation de  $Coef_2$  a également été modifiée (en considérant la liaison entre  $NDVI$  et  $LAI$  telle que proposée par Weiss et al. 2002).

Ainsi :

$$Coef_1 = T_s (\eta + \delta \cdot \alpha) \quad (III.3)$$

$$Coef_2 = \exp(-\alpha_{C_2} LAI) \quad (III.4)$$

Nous avons calculé les coefficients empiriques de toutes ces expressions par la méthode des moindres carrés. Cela consistait à réduire l'écart de différence entre les observations de  $G/R_n$  et les résultats donnés par l'équation. L'implémentation de cette application est très simple à réaliser sous MATLAB avec la fonction *fminsearch*. Cet apprentissage a été effectué sur les simulations SiSPAT décrites dans la **section 4.2.2** de la **partie II** en considérant les valeurs de  $G/R_n$  à midi solaire et les couverts peu développées ou développés en ne retenant que les données avec un  $LAI$  supérieur à 0,25 (le  $LAI$  dépassant tout juste 2 dans les données). Ainsi donc l'ensemble des données SiSPAT (parcelles 101, 120, 121, 214 et 501) constitue 393 points sur l'ensemble des jours simulés. Cette partie d'étude est ensuite validée sur les données réelles des sept parcelles d'Alpilles (101,102, 120, 121, 203, 214, et 501) acquises lors de la campagne ReSeDA (662 points) et sur les données des campagnes 2002 (blé, 180 points) et 2005 (pois, 241 points) sur le site 'Flux et Télédétection' à Avignon.

### 2.1.2.1. Apprentissage sur données SiSPAT

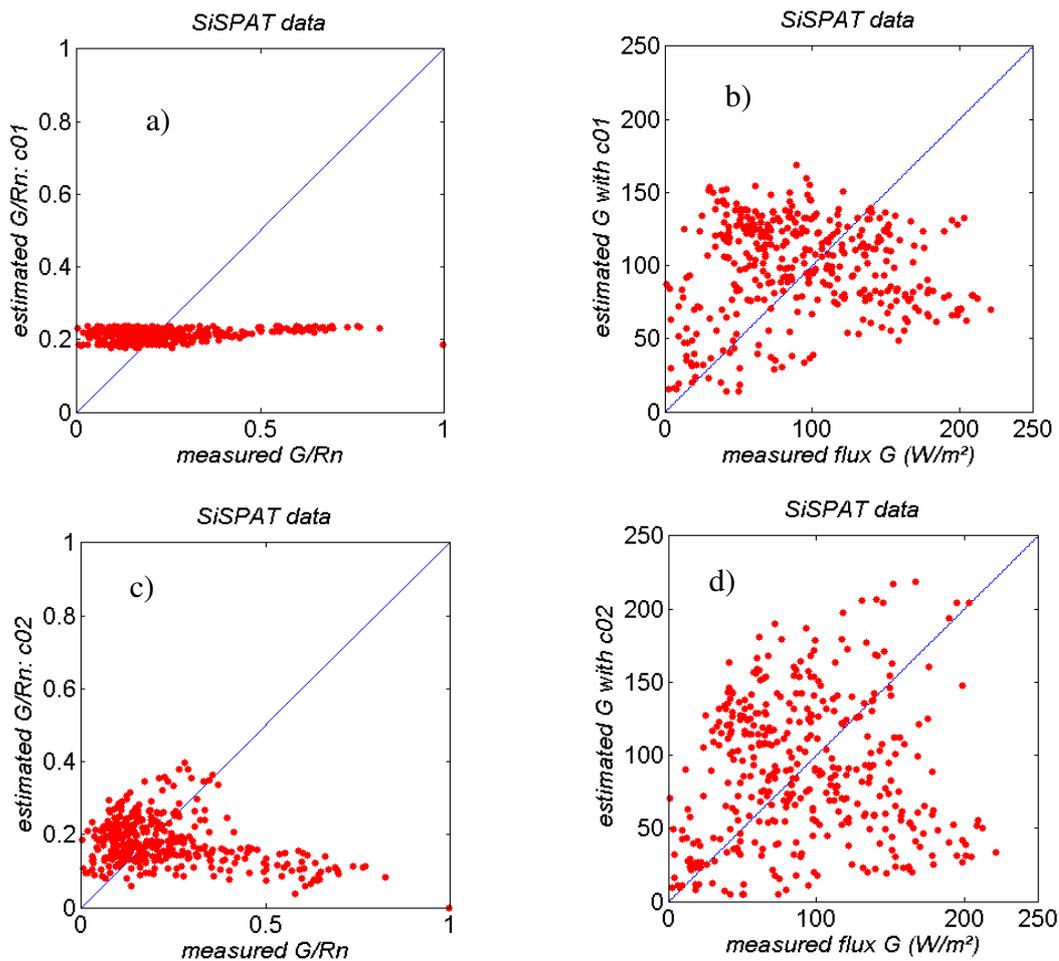
L'application des équations (III.1 et III.2) sur les données SiSPAT donne les résultats suivants (Fig. III.8 et Fig.III.9). Les coefficients empiriques déterminés par la méthode des moindres carrés sont présentés dans le tableau III.5 ci-dessous. Pour comparaison, l'application des coefficients de l'expression 6 donnés par Choudhury et al. (1987) et Kustas et al. (1993), ainsi que le résultat pour un coefficient  $c_0$  constant sont aussi donnés (la valeur de  $c_0$  constant ayant été obtenue par une simple moyenne des valeurs du rapport  $G/R_n$ ).

**Table III.5a. Résultats statistiques du rapport  $G/R_n$**

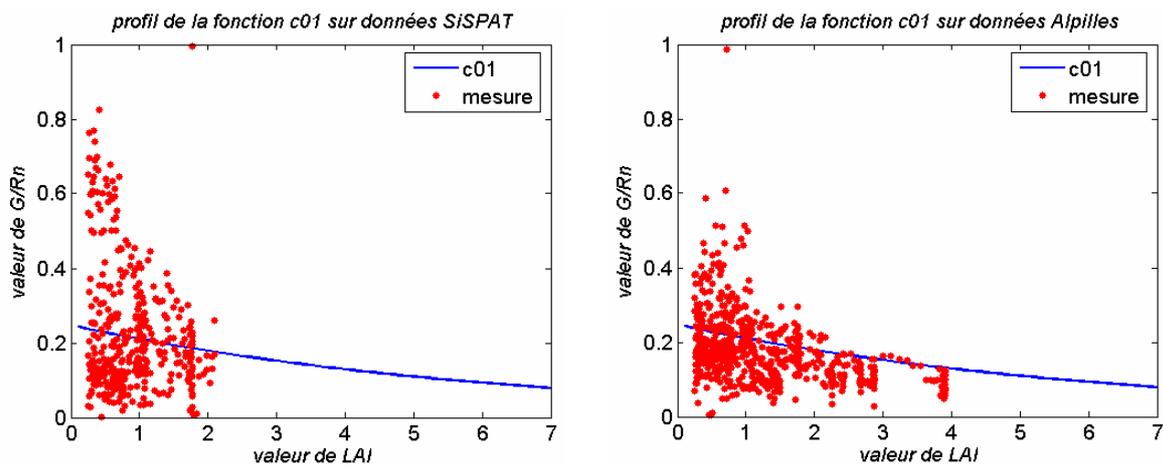
393 points	RMSE	RRMSE	$\eta$	$\delta$	$\beta$	$\alpha_c$
$c_{01}$ (Eq. III.1)	0,17	63%	-	-	0,25	0,16
Choudhury et al. (1987)	-	-	-	-	0,4	0,5
Kustas et al. (1993)	-	-	-	-	0,29	0,26
$c_{02}$ (Eq. III.2)	0,20	71%	0,011	-0,024	-	-0,052

**Table III.5b. Résultats statistiques du flux  $G$  calculé à 12h sur les données SiSPAT**

393 points	RMSE ( $Wm^{-2}$ )	RRMSE (%)
$G$ avec $c_{01}$ (Eq. III.1)	58,3	54,6
$c_0 = 0,214$	58,9	55,2
Choudhury et al. (1987)	70,5	68,1
Kustas et al. (1993)	61,4	58,0
$G$ avec $c_{02}$ (III.2)	67,0	64,2



**Fig. III.8.** Rapport  $G/R_n$  et flux  $G$  à 12h estimés à partir de l'Equation III.1 (en haut  $c_{01}$ ) et de l'Equation III.2 (en bas  $c_{02}$ ).



**Fig. III.9.** Profil de la fonction  $c_{01}$  (Eq. III.1) en fonction du LAI par rapport aux rapports des flux  $G$  et  $R_n$  à 12h, sur les simulations SiSPAT à gauche (apprentissage) et sur les données réelles Alpiilles à droite (validation).

Le calage des deux modèles sur les données simulées ne fournit pas des résultats réellement satisfaisants. Les *RMSE* sont forts et similaires à l'utilisation d'un coefficient  $c_0$  constant. On obtient avec  $c_{01}$  une erreur quadratique (*RMSE*) de  $58 \text{ Wm}^{-2}$  sur l'estimation du flux conductif dans le sol ( $G$ ) correspondant à une dispersion d'environ 54% (Fig. III.8b). Quant à l'expression  $c_{02}$  (III.2), l'erreur quadratique est de l'ordre de  $67 \text{ Wm}^{-2}$ , avec une dispersion de 64% (Fig. III.8d).

Pour la gamme de données explorée les deux modèles ne permettent pas de représenter la variabilité des  $c_0$  représentés, ce qui est bien illustré par la Figure III.8a et par la Figure III.9a (à gauche) qui présente l'**expression 6** et les données utilisées pour la calibration. Si le deuxième modèle, plus complexe, semble pouvoir mieux représenter la variabilité aux faibles  $c_0$ , les forts  $c_0$  restent non simulables (Fig. III.8c). Ces résultats montrent tout d'abord que la seule prise en compte du *LAI* ne permet pas de rendre compte de la variabilité du rapport  $G/R_n$  qui dépend fortement dans cette gamme de *LAI* du fonctionnement physique du sol. La variabilité de ce fonctionnement n'est pas non plus correctement décrite par le deuxième modèle inspiré de l'**expression 7** (Bastiaanssen et al., 1998a). Cela peut être expliqué, au moins partiellement, par la sensibilité du rapport  $G/R_n$  à l'effet de l'humidité et de la conductivité thermique qui ne sont pas pris en compte par les entrées utilisées dans le modèle. En effet, en ajoutant un terme traduisant l'effet de l'humidité du sol dans l'expression de l'**Equation III.2**, on peut réduire la dispersion des résultats de 20% (le test a été réalisé en multipliant  $Coef_2$  par une fonction linéaire de l'humidité de surface). Cependant, l'approche d'introduire l'humidité dans l'équation est difficile à mettre en œuvre pour des raisons d'accessibilité de cette humidité à l'échelle régionale.

#### 2.1.2.2. Validation sur données Alpilles et Site 'Flux et Télédétection'

L'application de ces approches ( $c_{01}$ ,  $c_{02}$ ) sur les données réelles de l'expérimentation Alpilles et du site 'Flux et Télédétection' présente des résultats plus satisfaisants que les résultats obtenus dans la phase de calibration précédente (Fig. III.9b droite, III.10, Fig. III.11 et Tableau III.6). Les estimations de  $G$  à midi solaire sont même d'une assez bonne qualité par comparaison aux résultats généralement présentés dans la littérature (les *RMSE* de l'ordre de  $30 \text{ Wm}^{-2}$  sont bien inférieurs à ceux obtenus dans le même contexte expérimental par Jacob et al., 2002a et Gómez et al., 2005). Les erreurs relatives sont de l'ordre de 40 % ce qui est également satisfaisant. Ces

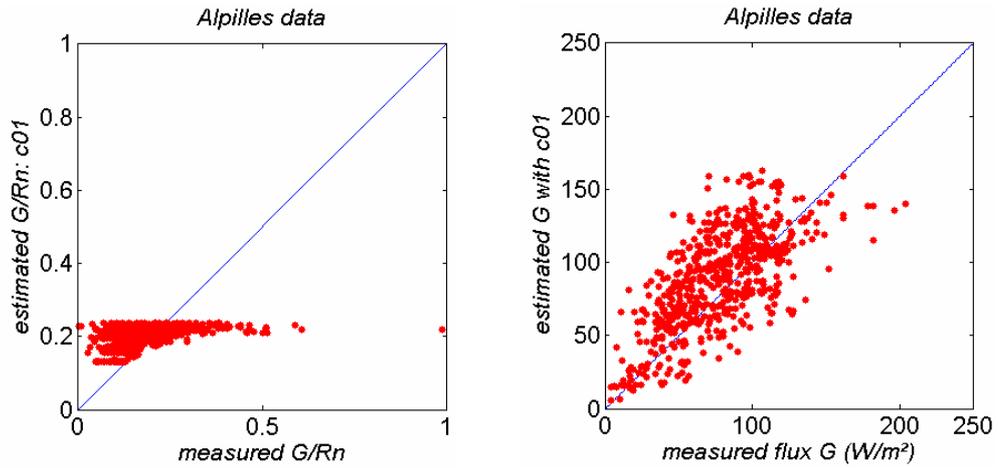
résultats semblent devoir s'expliquer par une moindre variabilité des conditions de sols sur les données réelles que pour les données d'apprentissage. Il est cependant possible de noter sur la figure III.7 que le modèle développé peut laisser penser à des surestimations possibles de  $c_{01}$  du rapport aux plus fortes valeurs de  $LAI$ . Par ailleurs, la supériorité de l'expression la plus simple ( $c_{01}$ ) déjà observée lors de l'apprentissage, est confirmée ici.

**Tab. III.6a. Résultats statistiques du rapport  $G/R_n$**  (entre parenthèse le nombre de points utilisée pour les différentes données)

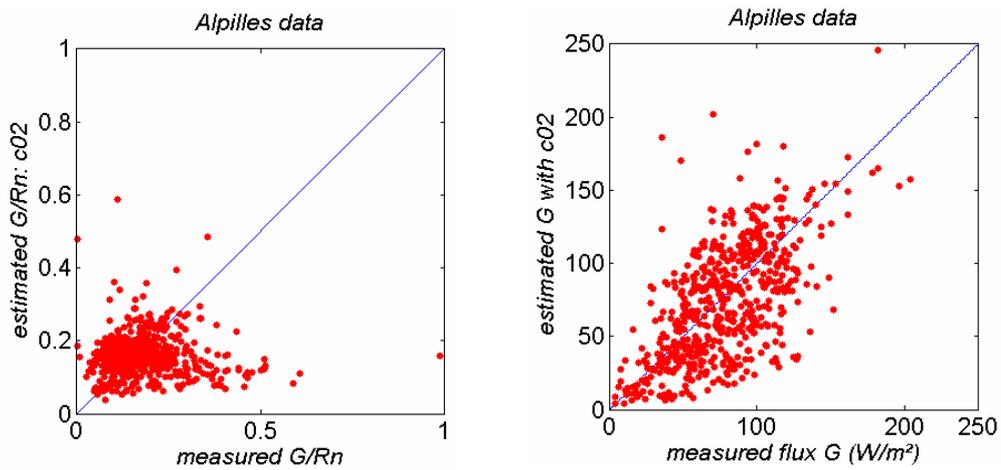
<i>Equations</i>	<i>Données</i>	<i>Alpilles (662)</i>		<i>Grue 2002 (180)</i>		<i>Grue 2005 (241)</i>	
		<i>RMSE Wm<sup>-2</sup></i>	<i>RRMSE %</i>	<i>RMSE Wm<sup>-2</sup></i>	<i>RRMSE %</i>	<i>RMSE Wm<sup>-2</sup></i>	<i>RRMSE %</i>
$c_{01}$ (III.1)		0.095	52.6	0.069	53.4	0.086	45.7
$c_0 = \text{moyenne de } c_{01}$		-	-	-	-	-	-
Choudhury et al. 1987		0.114	61.7	0.101	78.0	0.140	73.6
Kustas et al. (1993)		0.097	53.3	0.073	56.3	0.098	51.7
$c_{02}$ (III.2)		0.083	29.2	0.119	92.6	0.112	59.1

**Tab. III.6b. Résultats statistiques du flux  $G$  calculé à 12h sur les données Alpilles (662 points), Grue 2002 (180 points) et Grue 2005 (241 points).**  $c_0 = (0,201 ; 0,150 ; 0,224)$  : valeurs de  $c_0$  pour les différentes données.

<i>Equations</i>	<i>Données</i>	<i>Alpilles (662)</i>		<i>Grue 2002 (180)</i>		<i>Grue 2005 (241)</i>	
		<i>RMSE Wm<sup>-2</sup></i>	<i>RRMSE %</i>	<i>RMSE Wm<sup>-2</sup></i>	<i>RRMSE %</i>	<i>RMSE Wm<sup>-2</sup></i>	<i>RRMSE %</i>
$G$ avec $c_{01}$ (III.1)		29,7	38,8	29,2	69,8	32,3	45,5
$c_0 = (0,201 ; 0,150 ; 0,224)$		34,8	45,5	30,6	73,4	37,1	52,2
Choudhury et al. 1987		42,5	55,6	45,3	108	55,4	72,0
Kustas et al. (1993)		32,1	42,0	30,7	73,6	37,8	53,2
$G$ avec $c_{02}$ (III.2)		36,8	48,1	53,1	127	37,0	52,2



**Fig. III.10. Rapport  $G/R_n$  (à gauche  $c_{01}$ ) et flux  $G$  à 12h estimés à partir de l'équation (III.1) sur les données Alpillles**



**Fig. III.11. Rapport  $G/R_n$  (à gauche  $c_{02}$ ) et flux  $G$  à 12h estimés à partir de l'équation (III.2) sur les données Alpillles**

### 2.1.3. Cas du sol nu

Les études de Cellier et al., (1996) et Santanello et Friedl (2003) ont apportées certains éléments permettant d'analyser le comportement du rapport  $G/R_n$  sur sol nu. Des études récentes montrent qu'il est possible de le modéliser à partir d'une fonction sinusoïdale dépendant du temps. Murray et Verhoef (2007) se sont appuyés sur l'étude de Santanello et Friedl (2003) pour exprimer une méthode analytique basée sur une analyse harmonique de la température de surface (Van Wijk & De Vries, 1963).

Nous avons analysé les équations de Santanello et Friedl (2003) et de Cellier et al. (1996) afin de déduire une approche pour estimer le flux  $G$  [expressions 8 et 9 du Tableau III.4]. Seule la période de la journée où la valeur du flux  $G$  est positive (soit entre 6h et 16 h) est utilisée.

✚ Equation de Santanello et Friedl (2003) :

Les résultats de simulation avec un modèle détaillé de transfert (SHAW, Flerchinger et al., 1998) sur les données Monsoon'90 leur ont permis d'examiner certaines relations entre le rapport  $G/R_n$ , son évolution dans le temps, et le type de sol, l'humidité du sol et la végétation. L'étude a montrée que le maximum du rapport  $G/R_n$  apparaît vers 9h -10h, soit 2 à 3 heures avant que le maximum du rayonnement net  $R_n$  soit atteint. Ces analyses ont conduit à définir une fonction de la forme sinusoïdale suivante pour modéliser le rapport  $G/R_n$ :

$$c_0 = A * \cos\left(\frac{2\pi(t+10800)}{B}\right) \quad (\text{III.5})$$

où  $A$  représente la valeur maximum de  $G/R_n$  au cours de la journée,  $B$  est choisi pour minimiser l'écart entre  $G/R_n$  de l'équation (III.5) et les simulations du modèle SHAW, et  $t$  est le temps en seconde. La constante 10800 traduit en seconde le décalage de temps de 3h observé entre le flux  $R_n$  et le rapport  $G/R_n$ . L'étude préconise certaines valeurs de  $A$  et  $B$  selon la teneur en eau du sol résumé dans le tableau III.7 ci-dessous :

**Tab. III.7. Valeurs des coefficients de l'équation de Santanello et Friedl (2003)**

	<i>A</i>	<i>B</i>
moist soils ( $\theta = 0,4$ )	0,31	74 000
intermediate moisture ( $\theta = 0,25$ )	0,33	85 000
dry soils ( $\theta = 0,05$ )	0,35	100 000

Santanello et Friedl (2003) montrent également que les coefficients *A* et *B* peuvent être reliés, par une relation empirique, à l'amplitude de la température de surface. Une régression linéaire entre le minimum et le maximum de la température de surface ( $\Delta T$ ) de la journée permet d'approximer les coefficients *A* et *B* (e.g.  $A = 0.0074 \Delta T + 0.088$ ). L'écart de variation observé par cette approximation est de  $\pm 0,05$  pour *A* et de  $\pm 15\ 000$  s pour *B*.

✚ Equation de Cellier et al. (1996) :

L'étude de Cellier et al. (1996) montre que les échanges d'énergie entre un sol nu et l'atmosphère dépendant à la fois des paramètres du sol et des conditions météorologiques, il est possible de mettre en place une méthode simple pour calculer le flux *G* à partir des propriétés du sol et des données météorologiques. Le jeu de données acquis sur trois types de sol (sablo limoneux, limoneux et sols crayeux) dans le nord de la France près de Laon entre Avril – Mai 1991 a servi de base d'analyse. L'étude montre que les relations entre le rayonnement net et le flux conductif dans le sol sont presque linéaires sur tous les sols avec des pentes très similaires. La corrélation entre la moyenne journalière du flux *G* et la conductivité thermique confirme l'influence des propriétés thermiques sur le flux du sol *G*. Il montre que, même avec des valeurs quotidiennes, le flux *G* ne peut pas être considéré comme proportionnel au rayonnement net. Le ratio  $G/R_n$  est plus grand (40 - 50%) à faible rayonnement net (moins de  $250\ \text{Wm}^{-2}$ ), que à fort rayonnement net (30 - 40%). A faible rayonnement net, les pentes de la relation entre le flux sensible *H* et le flux  $R_n$  sont similaires à celles du flux de chaleur *G*. Toutefois, les pentes augmentent et les données deviennent de plus en plus dispersées à fort rayonnement net. Ceci a amené à observer les rapports  $G/R_n$  et  $H/R_n$  par rapport à la vitesse du vent  $u_a$ . L'étude a révélé que  $G/R_n$  diminue progressivement en fonction de la vitesse moyenne du vent sur la même période. Ainsi donc, on obtient l'expression suivante pour calculer le rapport  $G/R_n$ :

$$c_0 = \alpha_1 + \beta_1 u_a \quad (\text{III.6})$$

où  $u_a$  est la vitesse moyenne du vent de la journée. Cependant, contrairement à l'équation III.5 cette approche ne tient pas compte du déphasage des flux. La prise en compte de ce déphasage devrait permettre d'améliorer l'estimation du flux  $G$ .

#### Notre approche dans cette étude

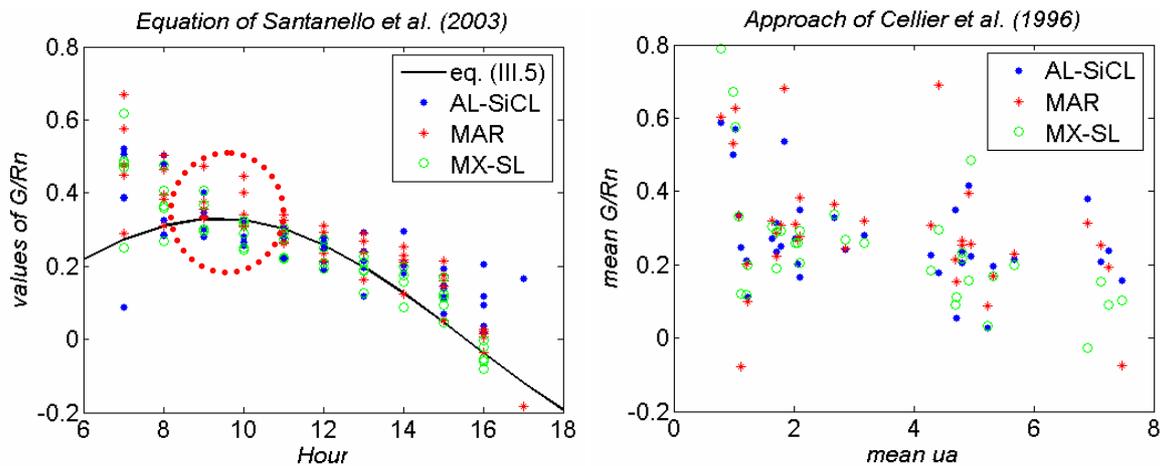
Nous avons cherché à tenir compte des différents facteurs analysés dans les approches de Cellier et al. (1996) et de Santanello et Friedl (2003) (déphasage, humidité du sol, température et vitesse du vent). Les données TEC décrites dans la section 4.2.3 de la partie II conçues dans la configuration d'un sol nu, nous ont permis d'analyser les variations du rapport  $G/R_n$  et de déterminer des expressions potentielles de  $c_0$  et d'estimer leurs coefficients.

L'observation de l'évolution journalière de  $G/R_n$  sur les données TEC montre que le maximum de  $G/R_n$  se situe autour de 10 h (Fig. III.12a) ce qui est en bon accord avec les travaux originaux et justifie l'utilisation de l'équation (III.5), nécessitant la calibration des paramètres  $A$  et  $B$ . Ce dernier est déterminé en analysant le déphasage des simulations de  $G$  et de  $R_n$ . En ce qui concerne le coefficient  $A$  qui représente la valeur maximale du rapport  $G/R_n$  sur la journée, l'idée est d'exploiter l'expression de l'équation (III.3) qui selon Bastiaanssen et al. (1998a) traduit le rapport  $G/R_n$  sur sol nu. Elle peut donc permettre d'exprimer ce coefficient en considérant donc que  $\max(G/R_n)$  est fonction de la température de surface ( $T_s$ ) et de l'albédo ( $\alpha$ ) conduisant à l'expression  $A_1$  donnée ci-dessous. Certes, le terme  $A$  est paramétré en fonction de l'humidité du sol et une relation empirique entre les températures de surface maximale et minimale ( $\Delta T$ ). En général, nous ne disposons d'aucune connaissance sur l'humidité de surface du sol à l'échelle régionale étudiée. De plus, nous n'avons qu'une seule donnée de température de surface autour de midi par télédétection. Ainsi donc, le raisonnement est de mettre en place une démarche dont l'estimation du terme  $A$  ne dépend pas absolument de la connaissance de l'humidité du sol et ne faire intervenir qu'une seule valeur de température de surface :

$$A_1 = T_s (\eta + \delta \cdot \alpha) \quad (\text{III.7})$$

La Figure III.12b montre que le comportement du rapport  $G/R_n$  simulé par TEC se comporte de manière similaire vis-à-vis de la vitesse du vent que dans l'étude de Cellier et al. (1996). Ainsi, nous avons testé une deuxième approche  $A_2$  dans laquelle est introduite la vitesse du vent par l'intermédiaire d'une expression similaire à l'équation (III.6) :

$$A_2 = T_s (\eta + \delta \cdot \alpha) \cdot (\alpha_1 + \beta_1 u_a) \quad (III.8)$$



**Fig. III.12. Evolution du rapport  $G/R_n$  calculé à partir des simulations TEC vis-à-vis du temps (ég. III.5) et de la vitesse du vent (ég. III.6)**

Quant à l'analyse des processus sur les sols nus, les données du modèle TEC sont constituées dans cette optique. Elles se composent de ces 3 types de sols (AL-SiCL, MAR, MX-SL) sur 35 jours correspondant à des périodes de parfait rayonnement global (soit un ensemble de 105 points). Ces résultats préliminaires sont ensuite validés sur les données du site GRUE.

### 2.1.3.1. Apprentissage sur les données TEC

L'équation (III.5) présentée plus haut est formulée afin de tenir compte du décalage observé entre les flux. Les premières simulations montrent qu'avec une valeur de  $B = 89\ 000\ s$ , nous avons un positionnement correct du déphasage par rapport aux données. En effet, une comparaison avec la valeur moyenne originelle  $B = 85\ 000\ s$  montre une amélioration nette des estimations du flux  $G$  (Tab. III.8).

**Tab. III.8. Evaluation des Calculs du flux  $G$  ( $Wm^{-2}$ ) à 12h sur les données TEC suivant la valeur de  $B$**

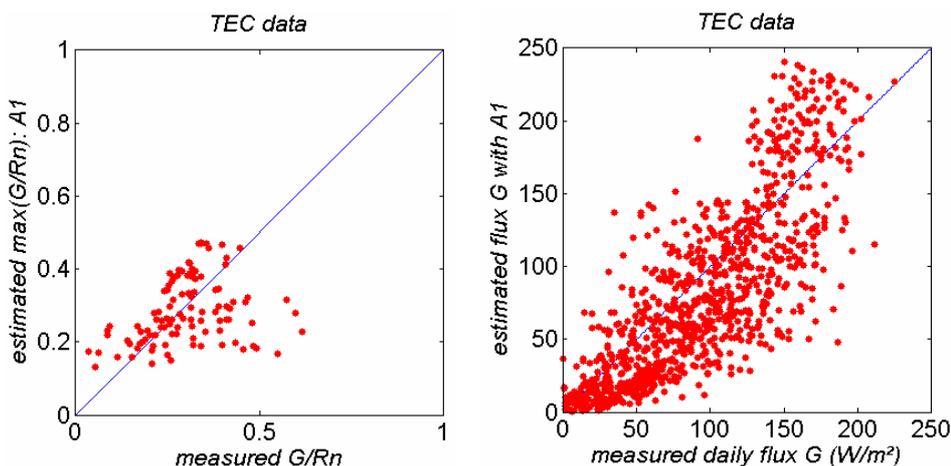
	$B = 89\ 000\ s$			$B = 85\ 000\ s$		
	AL-SiCL	MAR	MX-SL	AL-SiCL	MAR	MX-SL
$RMSE$ ( $Wm^{-2}$ )	35	28	31	46	39	39
$RRMSE$ (%)	34	28	32	44	38	41

Dans un second temps, nous avons évalué les termes  $A_1$  (III.7) et  $A_2$  (III.8) [ou  $\max(G/R_n)$ ] en utilisant la méthode des moindres carrés ce qui permet d'obtenir les coefficients  $\eta$ ,  $\delta$ ,  $\alpha_1$ , et  $\beta_1$  présentés dans le Tableau III.9. L'estimation des termes  $A_1$  et  $A_2$  montre une certaine dispersion (Fig. III.13a) qui peut être rapprochée des résultats obtenus dans l'exercice précédent sur les couverts peu développés, et qui pourrait donc s'expliquer par la non prise en compte du rôle de l'humidité superficielle du sol (notons que cet effet a été signalé dans le travail original de Santanello et Friedl 2003).

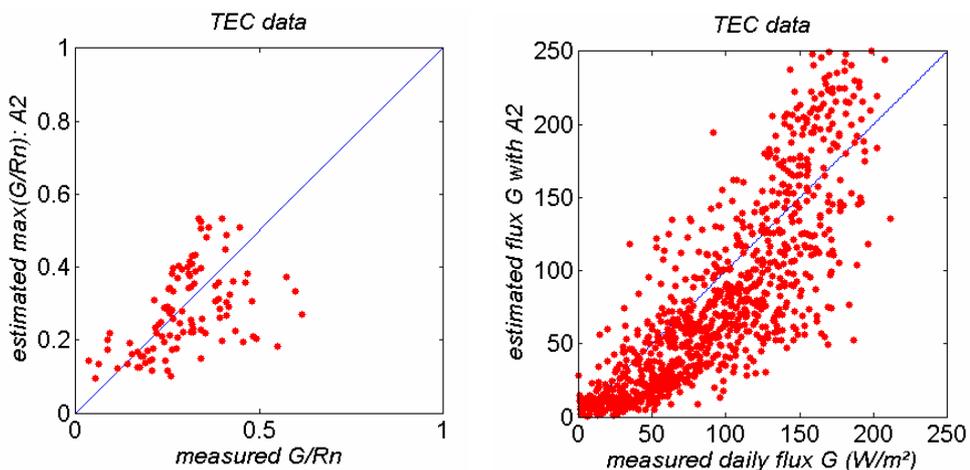
**Table III.9. Calculs des coefficients empiriques du terme  $A$**

$105\ points$	$RMSE$ ( $Wm^{-2}$ )	$RRMSE$ (%)	$\eta$	$\delta$	$\alpha_1$	$\beta_1$
$A_1$ (III.7)	0,12	40	0,0058	0,0036	-	-
$A_2$ (III.8)	0,11	39	0,053	0,631	0,092	-0,005

La Figure III.13b présente l'estimation basée sur la première approche ( $A_1$ ) du flux  $G$  horaire sur l'ensemble des données TEC utilisées. Une dispersion de 45%, qui correspond à une erreur quadratique de  $35 \text{ Wm}^{-2}$ , est observée (Tableau III.10), ce qui représente un résultat très satisfaisant.



**Fig. III.13. Evaluation de l'expression  $A_1$  (III.7) et du flux  $G$  horaire estimé entre 6h – 16h (III.5)**



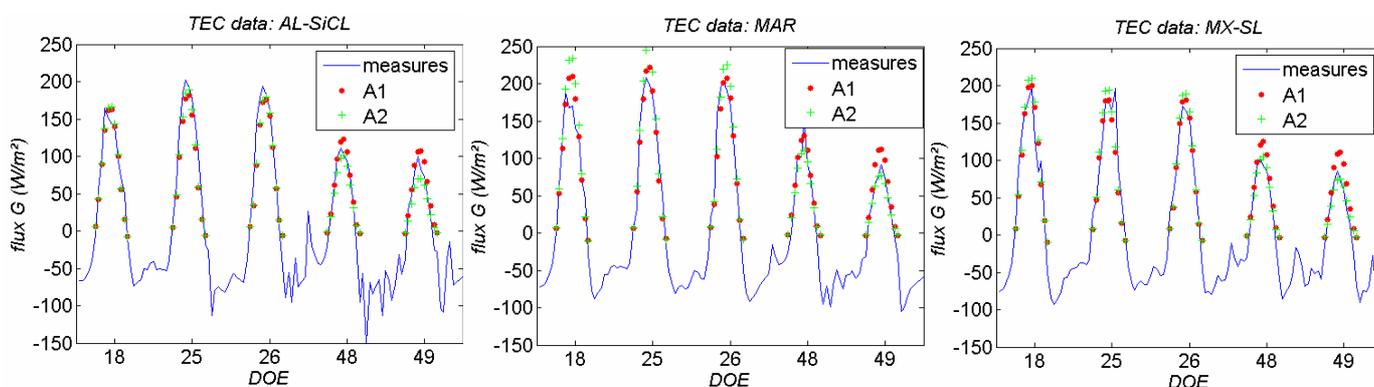
**Fig. III.14. Evaluation de l'expression  $A_2$  (III.8) et du flux  $G$  horaire estimé entre 6h – 16h (III.5)**

**Tab. III.10. Evaluation de l'estimation des flux  $G$  horaires estimés par les deux approches proposées  $A_1$  et  $A_2$  par rapport aux données TEC (cf. Fig. II13b et Fig.III.14b)**

$flux\ G\ horaire\ (Wm^{-2})$	$RMSE\ (Wm^{-2})$	$RRMSE\ (\%)$
$A_1\ (III.7)$	35	45
$A_2\ (III.8)$	35	45

Les résultats obtenus par la deuxième approche ( $A_2$ ) ne montrent aucune amélioration significative par rapport à la première approche ( $A_1$ ) (Tableaux III.9 et III.10, Figure III.14). Certes la dispersion sur l'estimation du terme  $A_2$  se réduit de 1%, mais l'impact sur le flux  $G$  est totalement négligeable. L'apport de la vitesse du vent dans la prédiction du flux  $G$  n'apparaît donc pas sur les données TEC.

Par ailleurs, l'évaluation de la calibration des deux expressions est présentée en fonction des différents sols du jeu de données (Fig. III.15 et Table III.11), montrant des résultats à peu près équivalents (un peu meilleur pour le sol MAR).



**Fig. III.15. Evolution horaire du flux  $G$  sur certains jours des données TEC**

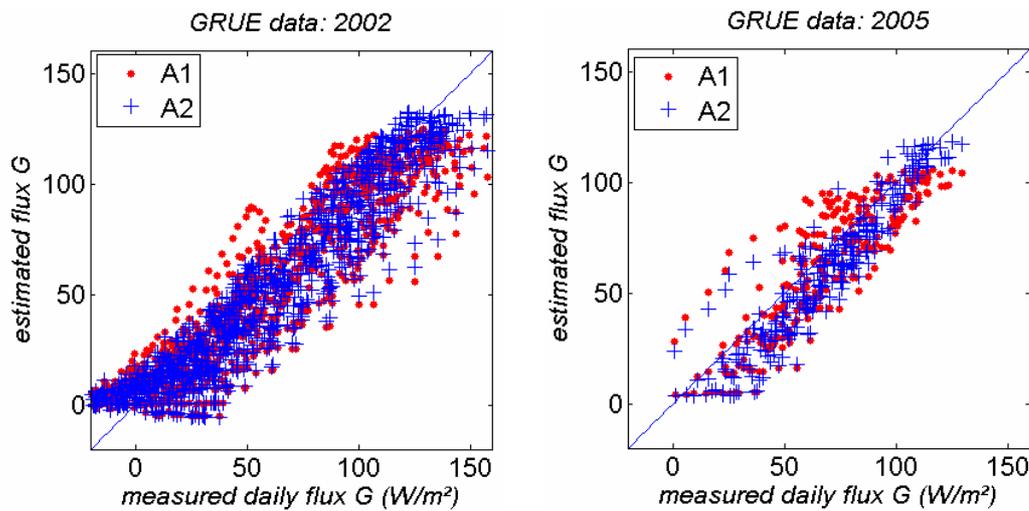
**Tab. III.11. Erreur sur l'estimation du flux  $G$  sur certains jours des données TEC (Fig. III.15)**

	$A_1\ (Eq.\ III.7)$			$A_2\ (Eq.\ III.8)$		
	AL-SiCL	MAR	MX-SL	AL-SiCL	MAR	MX-SL
$RMSE$	25	23	26	24	26	25
$RRMSE$	33	29	38	32	33	35

Nous avons réalisé d'autres analyses lors de cet apprentissage dont nous allons simplement évoquer les résultats. Dans les approches ci-dessus, les données de télédétection de la température de surface sont obtenues autour du midi solaire or l'étude montre un maximum du rapport  $G/R_n$  vers 10h. Nous avons essayé un apprentissage utilisant des températures mesurées à proximité de ce maximum pour évaluer l'impact du décalage temporel. Les résultats n'ont pas montré une amélioration significative (variation de  $RMSE$  inférieure à  $2 \text{ Wm}^{-2}$  sur l'estimation du flux  $G$ ). Ce qui montre qu'il est tout à fait possible d'exprimer le rapport maximum  $\max(G/R_n)$  avec la température de surface de midi. Une autre remarque sur l'étude concerne la stabilité de l'estimation des coefficients empiriques ( $\alpha, \beta, \eta, \delta$ ) lors de la calibration du rapport  $G/R_n$ . Celle-ci a été étudiée en réalisant plusieurs calibrations en partant de valeurs initiales différentes des paramètres dans le processus de minimisation. La stabilité des coefficients déterminés était très liée au nombre de coefficients. Elle est la meilleure pour les équations III.1 et III.7, tandis qu'avec les autres approches (4 coefficients à déterminer), nous sommes confrontés à une certaine instabilité des coefficients. Par exemple, on a observé que  $\eta$  pouvait varier entre 0,0050 et 0,0060 et  $\delta$  entre 0,003 et 0,006. Cependant ces instabilités affectent peu les estimations du flux  $G$ .

### 2.1.3.2. Validation du terme A sur les données Grue

Après leur calibration sur les données simulées par TEC, nous avons appliqué les deux approches du calcul du terme A sur les données du site GRUE. Nous nous sommes intéressés aux périodes de sol nu des années 2002 et 2005 (périodes estivales après les récoltes de blé et de pois). Nous comptons 36 jours en 2002 correspondant à des journées parfaitement dégagées (sans nuages) et seulement 14 jours en 2005. Les résultats des évaluations sont présentés dans la Figure III.16 et le Tableau III.12. Ces résultats sont très satisfaisants puisque les  $RMSE$  sont bas ( $< 20 \text{ Wm}^{-2}$ ). Les résultats des deux approches sont proches, la prise en compte de la vitesse du vent ( $A_2$ ) apportant cependant un léger mieux en particulier en 2005 (baisse de 4 % de la dispersion des données). Ces différences entre les deux approches sont essentiellement dues à une meilleure adéquation de la deuxième expression dans certaines situations de fort vent (mais ce n'est pas systématique comme le montre les Figures III.17 et III.18).

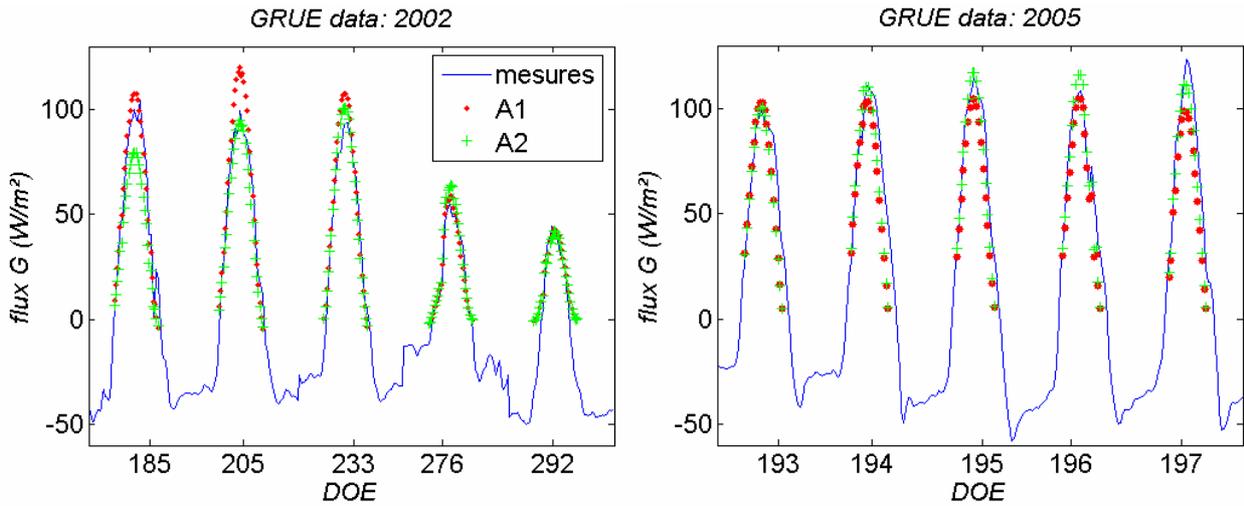


**Fig. III.16.** Evaluation du flux  $G$  horaire estimé entre 6h – 16h avec les termes ( $A_1$  et  $A_2$ ) sur les données GRUE

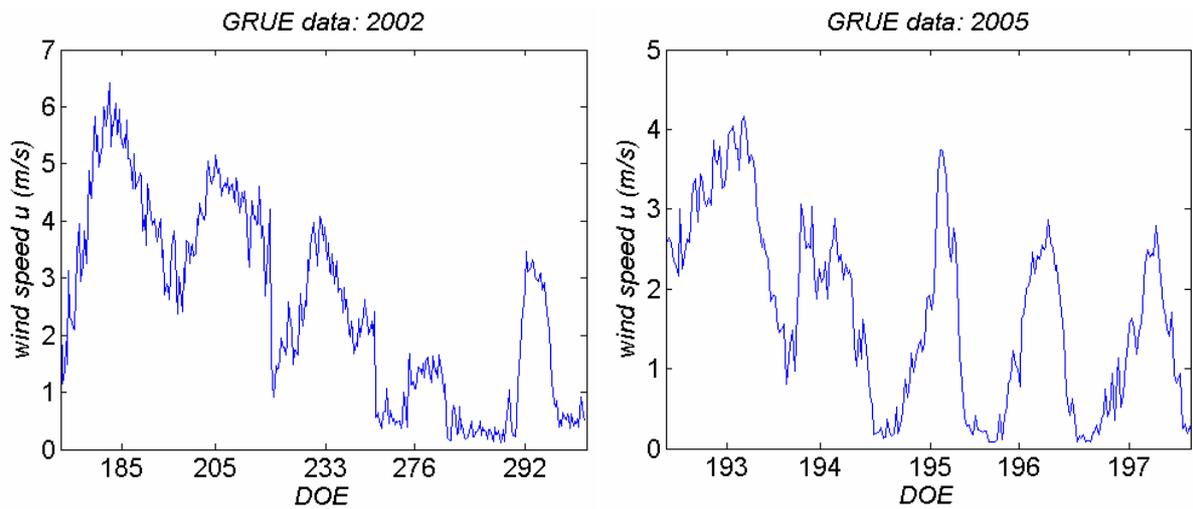
**Tab. III.12.** Evaluation de l'estimation des flux  $G$  horaires estimés entre 6h – 16h avec le terme  $A$  sur les données GRUE (cf. Fig. II.16)

	2002		2005	
	$A_1$ (III.7)	$A_2$ (III.8)	$A_1$ (III.7)	$A_2$ (III.8)
$RMSE (Wm^{-2})$	17	16	16	13
$RRMSE (\%)$	32	29	24	20

La difficulté dans l'implémentation de ces deux approches est liée aux coefficients du terme  $A$ . En effet, d'une part nous avons 2 coefficients à déterminer tandis que d'un autre côté 4 coefficients sont nécessaires. Par ailleurs, l'analyse des tableaux (III.10, III.11, III.12, et III.13) nous montre que l'impact de la vitesse du vent dans le calcul du terme  $A$  sur l'estimation du flux  $G$  est relativement important. L'écart entre les deux approches sur l'estimation du flux  $G$  par rapport aux mesures n'est que de  $2 \text{ Wm}^{-2}$  en moyenne. Par conséquent, privilégier la première approche ( $A_1$ ) se révèle judicieux, d'autant plus qu'elle permet d'éviter d'utiliser une approche plus complexe semblant pouvoir poser des problèmes quant à la stabilité des estimations des coefficients empiriques à déterminer.



**Fig. III.17. Evolution horaire du flux  $G$  sur certains jours des données GRUE**



**Fig. III.18. Evolution de la vitesse du vent ( $u_a$ ) sur certains jours des données GRUE**

**Table III.13. Erreur sur l'estimation du flux  $G$  sur certains jours des données GRUE (Fig. III.17)**

	2002		2005	
	$A_1$ (III.7)	$A_2$ (III.8)	$A_1$ (III.7)	$A_2$ (III.8)
$RMSE (Wm^{-2})$	9	10	15	13
$RRMSE (\%)$	21	24	20	16

#### 2.1.4. Conclusion sur le rapport $G/R_n$

Pour conclure cette étude du rapport  $G/R_n$ , il ressort tout d'abord d'après l'étude bibliographique que l'approche d'estimer le flux conductif du sol  $G$  comme une simple proportion du rayonnement net ( $R_n$ ) a fait émerger un grand nombre d'approches différentes. Ces différentes formulations relient le rapport des flux aux propriétés caractéristiques de la surface ( $LAI$ ,  $NDVI$ , réflectance, albédo) et à des données météorologiques (température, vitesse du vent). Leur grande variabilité montre qu'elles ne proposent un modèle général et ne permettent pas d'obtenir systématiquement des résultats corrects. Par exemple, si le rapport  $G/R_n$  est considéré constant au cours de la journée, elles peuvent conduire à une sous-estimation du flux  $G$  dans la matinée et une surestimation du flux  $G$  dans l'après midi, en particulier pour les sols nus. Cette observation est appuyée par plusieurs études empiriques montrant que le rapport entre les flux  $G$  et  $R_n$  n'est pas constant au cours de la journée. Elles ne semblent pas non plus pouvoir rendre compte de la variabilité des situations constatées, ce qui s'exprime par une grande variabilité des rapports, essentiellement en conditions de sol nu. Par conséquent, cette méthode n'est pas toujours appropriée dans plusieurs situations (particulièrement lorsque l'hétérogénéité régionale de la surface doit être prise en compte).

Dans un premier abord, nous avons analysé différentes expressions dérivées de formules décrites dans la bibliographie et évalué leur impact sur l'estimation du flux conductif dans le sol. L'étude révèle des résultats contrastés selon les exercices réalisés.

Dans le cas de la présence de couvert végétaux, l'expression du  $G/R_n$  en fonction simplement de l'indice foliaire  $LAI$  (Eq. (III.1)), bien que difficile à calibrer sur les données SiSPAT a permis d'obtenir des résultats corrects sur les mesures de terrain de l'expérimentation Alpilles et du site 'Flux et Télédétection' d'Avignon en considérant les estimations à midi solaire. L'expression dérivée du modèle SEBAL tenant compte également de la température de surface et de l'albédo a systématiquement été moins performante.

Le comportement du rapport  $G/R_n$  a été examiné dans la configuration d'un sol nu avec des expressions tenant compte de ses évolutions horaires (III.5) en décomposant l'équation en deux termes ( $A$  et  $B$ ). Si nous avons réussi à nous affranchir du terme  $B$  exprimant le déphasage entre les flux  $R_n$  et  $G$  pour obtenir des résultats cohérents, le terme  $A$  correspondant au maximum du rapport  $G/R_n$  de la journée est plus difficile à paramétrer. Nous avons introduit deux approches (III.7 et III.8) pour l'exprimer, qui toutes deux donnent des résultats de flux  $G$  de très bonne

qualité, en particulier sur les données réelles. La comparaison de ces deux approches du calcul du terme  $A$  a montré qu'il ne semblait pas pertinent de complexifier l'approche en tenant compte de l'impact de la vitesse du vent sur l'estimation du flux  $G$ , bien que pour certaines journées particulières elle pouvait apporter une amélioration sensible du résultat.

Les deux études que nous avons réalisées sur l'estimation du flux  $G$  à partir du rayonnement net, si elles ont permis d'obtenir des flux  $G$  de bonne qualité sur les données d'évaluation, n'ont pas permis de dégager une modélisation permettant de décrire correctement l'ensemble des situations abordées, en particulier lors des calibrations. Un travail plus approfondi devrait être mené, en particulier en élaborant un plan de simulation au moyen de modèles de transfert détaillés spécifique au problème posé (plutôt que d'utiliser des jeux de données déjà existants), et en testant les expressions élaborées sur des situations expérimentales plus nombreuses. Enfin, il conviendrait d'analyser les nouvelles expressions, comme celles qui ont été proposées par Murray et Verhoef (2007) qui ne nous étaient pas connues au moment de réaliser cette étude.

## 2.2. Etude du rapport $kB^{-1}$

### 2.2.1. Positionnement du problème et introduction de l'analyse

Estimer le flux de chaleur sensible  $H$  entre la surface et l'air au dessus requiert une estimation de la résistance aérodynamique  $r_a$  (voir la description du modèle PBLs dans la partie II de ce mémoire). Cette résistance est souvent exprimée comme une fonction d'un paramètre sans dimension donnée par (Owen & Thompson, 1963; Chamberlain, 1968)  $kB^{-1}$ . L'expression de  $kB^{-1}$  (III.9) est le rapport logarithmique entre la rugosité aérodynamique ( $z0$  ou  $z0_m$ ) et de la rugosité thermique ( $z0_h$ ).

$$kB^{-1} = \ln\left(\frac{z0_m}{z0_h}\right) \quad \text{(III.9)}$$

L'analyse expérimentale de ce paramètre sur diverses végétations peu denses a montré une large gamme de variation (Kustas et al., 1989 ; Troufleau et al., 1997 ; Verhoef et al., 1997) avec des valeurs fortes dans toutes sortes de situation, indiquant des différences de plusieurs ordres de grandeur entre les deux rugosités. Pour des surfaces homogènes avec une végétation développée une valeur de l'ordre de 2,3 est souvent supposée valide (Hasager, 1997). Cela correspond à une rugosité thermique du  $10^{\text{ème}}$  de la rugosité aérodynamique.

Les différences entre les deux rugosités sont causés par des différences de processus physiques entre les transports de quantité de mouvement et les transports de chaleur à proximité des surfaces d'échange, ainsi que par des différences de répartition des sources et puits de chacun de ces flux au sein des couverts (en particulier en lien avec la répartition des températures de surfaces dans le feuillage et au sol).

En pratique, une question fondamentale se pose du fait que les valeurs de  $kB^{-1}$  ne peuvent pas être mesurées directement, mais sont issues de la formulation des transferts de masse utilisant les mesures d'autres variables (selon les méthodes, les profils de vent et de température dans les plus basses couches de l'atmosphère, la température de surface, la vitesse de frottement, le flux  $H...$ ). Toute incertitude associée à ces mesures se reportera sur la précision des estimations des valeurs de  $kB^{-1}$ . Par exemple, la difficulté de connaître de façon précise l'émissivité de surface,

qui est nécessaire à la conversion de la mesure infrarouge thermique en température de surface, affectera l'estimation de  $kB^{-1}$ . (À titre indicatif, une différence de 1% sur l'émissivité de surface peut se traduire par une différence de 0,6 K sur la température de surface dérivée). Ces incertitudes sont certainement des facteurs importants qui contribuent aux incertitudes observées sur les valeurs de  $kB^{-1}$ . Comme conséquence de ces incertitudes, le sens physique du paramètre  $kB^{-1}$  ainsi que la précision associée à attendre ne sont pas clairement définies (Su et al., 2001).

La plupart des modèles du bilan d'énergie de surface ont en entrée la rugosité aérodynamique ( $z0_m$ ) et, essentiellement pour les modèles monocouches, la rugosité thermique ( $z0_h$ ). Celle-ci est en général décrite comme une simple fraction de la rugosité aérodynamique (tel est le cas du modèle PBLs avec un rapport de 1/10). Lorsque ces modèles utilisent la température de surface en entrée, la précision du calcul du flux  $H$  est fortement liée à celle de la rugosité thermique.

Bien que de nombreuses formulations empiriques de  $kB^{-1}$  aient été proposées au cours des dernières décennies, les incertitudes associées à ces formulations sont grandes, en particulier pour les couverts peu denses. Par exemple, le paramétrage proposé empiriquement par Kustas et al. (1989) comme une relation linéaire du produit  $u_a (T_s - T_a)$  ne semble valable que pour les valeurs de  $H$  élevées ; pour des valeurs inférieures, la pente de la relation doit être ajustée en fonction du niveau de  $H$ . Plus récemment, des modèles plus complexes, élaborés à partir de considérations théoriques ou de modèles détaillés des transferts surface - atmosphère ont été proposés (e.g. Jensen et Hummelshøj, 1995 ; Massman, 1999a ; Blümel, 1999). Par exemple, l'expression de  $kB^{-1}$  développée par Massman (1999a) est dérivée d'un modèle physique complexe fondé sur la théorie lagrangienne. Ces nouvelles expressions se basent généralement sur des relations avec la vitesse de frottement ( $u^*$ ), l'indice de surface foliaire ( $LAI$ ) ou l'indice de surface de l'ensemble des éléments végétaux, feuilles et tiges ( $PAI$ : Plant Area Index), et la rugosité aérodynamique ( $z0_m$ ).

#### Stratégie de l'analyse

L'objectif principal de notre étude est de rechercher une méthode permettant de déterminer de façon indépendante la rugosité pour le transfert de chaleur de façon à l'intégrer au sein du modèle PBLs. Ce modèle ne considère pour l'instant qu'un rapport constant entre les rugosités

thermique et aérodynamique (correspondant à une valeur de 2,3 pour le paramètre  $kB^{-1}$ ). Trois modèles de  $kB^{-1}$  récents ont été retenus pour cette étude, le modèle utilisé par Hasager et al. (2002a), dérivés des travaux de Jensen et Hummelshøj, 1995), le modèle développé par Blümel (1999) et le modèle développé par Massman (1999a) dans sa version modifiée par Su et al. (2001). Ces trois modèles seront dénommés modèles de Hasager, de Blümel et de Massman dans la suite du texte.

Ces trois modèles sont évalués au moyen de trois jeux de données issus de l'expérimentation du site 'Flux et Télédétection' (ou site GRUE) d'Avignon (les années 2004, 2005 et 2007. cf. la **section 4.2.1** de la **partie II**). L'ensemble de données est formé de 152 jours pour l'année 2004, 96 jours pour 2005, et 149 jours pour l'année 2007. Les gammes de variation des variables utilisées sont décrites dans le tableau III.14 ci-dessous. Les données acquises au cours de l'expérimentation Alpilles sont également utilisées pour les jours correspondant aux acquisitions de données de télédétection aéroportée.

Pour réaliser l'évaluation des modèles par rapport aux données expérimentales, le paramètre  $kB^{-1}$  a été calculé à partir de l'équation du flux de chaleur sensible  $H$  (II.8, III.33). Nous avons dans un premier temps analysé la sensibilité du calcul par rapport aux erreurs sur les variables intervenant dans ce calcul. Dans un second temps, les modèles ont été évalués par rapport à ces estimations de  $kB^{-1}$ . Dans un troisième temps les modèles ont été évalués par rapport à leur impact sur les estimations de flux de chaleur sensible.

**Tab. III.14. Gamme de variation des valeurs à 12 h**

	$z0_m$ (m)	$PAI$ (-)	$u^*$ ( $ms^{-1}$ )	$T_s$ ( $^{\circ}C$ )	$T_a$ ( $^{\circ}C$ )	$H$ ( $Wm^{-2}$ )
2004	0,01 - 0,08	0,56 - 7	0,04 - 0,97	5 - 41	4 - 33	-123 - 415
2005	0,01 - 0,05	0,10 - 4	0,12 - 0,69	17 - 56	13 - 34	-35 - 367
2007	0,01 - 0,11	0,40 - 6	0,03 - 0,93	14 - 45	13 - 34	-11 - 449

## 2.2.2. Description des modèles étudiés

### 2.2.2.1. Modèle de Hasager

Hasager et al. (2002a) ont appliqué aux données aéroportées Alpilles, les équations dérivées sur des considérations théoriques pour tout type de surface et en particulier les couverts végétaux par Jensen et Hummelshøj (1995 ; cf. aussi Jensen et Hummelshøj, 1997). Elles avaient alors donné des résultats encourageants par rapport à l'utilisation de  $kB^{-1}$  constants (qui posait le problème de comment bien définir la valeur de ce paramètre). Le modèle est défini à la fois pour des couverts peu développés par l'expression suivante :

$$z0_h = \frac{z0_m}{\exp\left(100\sqrt{u^*}\right)} \quad \text{for } z0_m < 0.05 \text{ m} \quad (\text{III.10})$$

et

$$z0_h = \frac{z0_m}{\exp\left(22\sqrt{u^*}\right)} \quad \text{for } z0_m \geq 0.05 \text{ m} \quad (\text{III.11})$$

Pour les couverts végétaux denses, l'expression donnée par Jensen et Hummelshøj (1997) est utilisée :

$$r_b u^* = c \frac{Pr}{LAI^{2/3}} \left( \frac{l u^*}{\nu} \right)^{1/3} \quad (\text{III.12})$$

avec  $Pr$  le nombre de Prandtl,  $c$  une constante égale environ à 5 ici,  $\nu$  la viscosité cinématique de l'air et  $l$  une dimension caractéristique de la taille des éléments végétaux d'environ  $3 \cdot 10^{-3}$  m pour l'herbe, les céréales et les forêts de conifères et d'environ  $3 \cdot 10^{-2}$  m pour les forêts à feuilles caduques. L'expression de  $z0_h$  pour les couverts dense est donc dans notre cas :

$$z0_h = \frac{z0_m}{\exp\left(\frac{5.85}{LAI^{2/3}} u^{*1/3}\right)} \quad (\text{III.13})$$

### 2.2.2.2. Modèle de Blümel

Le modèle de  $kB^{-1}$  de Blümel (1999) a été obtenu par ajustement des résultats de simulation d'un modèle de transfert de masse multi-sources. Tout d'abord, Blümel (1999) a développé un modèle de transfert, qui nécessite les températures du sol, de la végétation, et de l'air à un niveau de référence pour simuler le flux de chaleur sensible et le flux de quantité de mouvement des différents types de surface sans une connaissance explicite des longueurs de rugosité thermique et aérodynamique. Les résultats des simulations ont alors été utilisés pour établir une relation fonctionnelle permettant de déterminer la valeur de  $kB^{-1}$  d'un couvert végétal donné en utilisant les valeurs limites de sol nu ( $kB_s^{-1}$ ) et du couvert végétal dense ( $kB_c^{-1}$ ). Le terme de  $kB^{-1}$  est exprimé comme suit (pour la description de ce modèle voir aussi Su et al. 2001) :

$$kB^{-1} = \frac{C(f_c)}{\ln(z_{eff} / z_{0meff})} - \ln\left(\frac{z_{eff}}{z_{0meff}}\right) \quad (\text{III.14})$$

avec la fonction  $C(f_c)$  définie par :

$$C(f_c) = \ln(z_{eff} / z_{0meff}) \ln(z_{eff} / z_{0heff}) \quad (\text{III.15})$$

L'indice *eff* fait référence à une valeur effective mesurée ou estimée en tenant compte de la fraction végétale du couvert  $f_c$ . Ainsi donc, la hauteur de référence effective est définie comme :

$$z_{eff} = z - f_c * d \quad (\text{III.16})$$

et la rugosité aérodynamique effective par :

$$z_{0meff} = z_{eff} \exp(-k / C_{DMeff}^{1/2}) \quad (\text{III.17})$$

$$\text{avec } C_{DMeff} = g(f_c)C_{DMC} + [1 - g(f_c)]C_{DMS} . \quad (\text{III.18})$$

Les coefficients de transfert (en conditions neutres) sont donnés pour le sol nu comme :

$$C_{DMS} = [k / \ln(z / z_{0ms})]^2 \quad (\text{III.19})$$

et pour le couvert développé comme :

$$C_{DMC} = [k / \ln(z / z_{0mc})]^2 \quad (\text{III.20})$$

où  $z_{0ms}$  correspond à la hauteur de la rugosité aérodynamique sur sol nu, fixée à 0.01 m si aucune information n'est disponible, et  $z_{0mc}$  est équivalent à  $z_{0m}$  pour le couvert développé.

La fonction  $g(f_c)$  est une fonction empirique donnée par :

$$g(f_c) = f_c^{\gamma} + f_c(1 - f_c)\zeta f \quad (\text{III.21})$$

où  $\gamma$  et  $\zeta f$  sont prescrits comme  $0.5 \leq \gamma \leq 1.0$  et  $0.0 \leq \zeta f \leq 1.0$  afin de déterminer l'influence de la géométrie de la végétation sur les flux de quantité de mouvement. Les détails sur ces paramètres peuvent être trouvés dans Blümel (1999).

Le calcul de  $kB^{-1}$  dans les conditions limites pour le sol nu et le couvert dense nécessite d'exprimer la fonction  $C(f_c)$  donnée par l'Equation III.15 pour ces conditions ( $C_s$  et  $C_c$ ) :

- pour un sol nu :

$$C_s = \ln\left(\frac{z}{z_{0ms}}\right) \ln\left(\frac{z}{z_{0ms}}\right) = \ln\left(\frac{z}{z_{0ms}}\right) \left[ \ln\left(\frac{z}{z_{0ms}}\right) + kB_s^{-1} \right] \quad (\text{III.22})$$

avec l'expression de  $kB^{-1}$  sur sol nu donnée selon Brutsaert (1982) par :

$$kB_s^{-1} = 2.46(\text{Re}^*)^{1/4} - \ln(7.4) \quad (\text{III.23})$$

où  $\text{Re}^*$  est le nombre de Reynolds.

- pour un couvert végétal dense :

$$C_c = \ln\left(\frac{z-d}{z_{0mc}}\right) \left[ \ln\left(\frac{z-d}{z_{0mc}}\right) + kB_c^{-1} \right] \quad (\text{III.24})$$

$$kB_c^{-1} = C_k (\sigma_a LSAI^3)^{-1/4} \left[ D_l u / \ln\left(\frac{z-d}{z_{0mc}}\right) \right]^{1/2} \quad (\text{II.25})$$

avec la constante  $C_k$  fixée à  $16.4 \text{ m}^{-1}\text{s}^{1/2}$ , et  $\sigma_a$  un facteur pour partitionner la quantité de mouvement définie par :

$$\sigma_a = 1 - \frac{0.5}{0.5 + PAI} \exp\left(-\frac{PAI^2}{8}\right) \quad (\text{III.26})$$

où  $PAI$  (Plant Area Index) représente l'indice de surface de l'ensemble des éléments végétaux, feuilles et tiges (équivalent du  $LAI$  qui est lui défini pour les feuilles seulement), et  $D_l$  est la dimension d'une feuille type fixée entre 0,005 – 0,05 (Su et al., 2001).

De manière à « fermer » le système d'équation ( $z_{0heff}$  est toujours inconnu), une deuxième expression de  $C$  ( $f_c$ ) est obtenue par l'ajustement d'une relation empirique sur les résultats des simulations du modèle multi-sources comme suit :

$$C(f_c) = A \exp(-a_1 f_c) + B \quad (\text{III.27})$$

avec

$$A = \frac{C_s - C_c}{1 - \exp(-a_1)} \quad (\text{III.27a})$$

$$B = C_s - A \quad (\text{III.27b})$$

$$a_1 = 2.6 \left( \frac{10h_c}{z} \right)^{0.355} \quad (\text{III.27c})$$

En résumé, le modèle de Blümel permet de déterminer la valeur de  $kB^{-1}$  pour n'importe quel type de surface végétalisée ou en sol nu à partir de l'équation (III.14) et en utilisant l'équation (III.27) pour interpoler en fonction du taux de couverture entre les valeurs de sol nu données par les équations (III.22 et III.23) et de couvert développé par les équations (III.24 et III.25).

### 2.2.2.3. Modèle de Massman

Comme dit plus haut, Massman (1999a) a dérivé une expression de  $kB^{-1}$  à partir d'un modèle physique complexe des transferts turbulents dans un couvert végétal basé sur la théorie lagrangienne. Cette expression a été modifiée par Su et al. (2001) pour simplifier les informations nécessaires à sa mise en oeuvre. Elle est exprimée comme la somme de trois termes :

$$kB^{-1} = \frac{kC_d}{4C_t \frac{u_*}{u(h)} (1 - e^{-n/2})} f_c^2 + 2f_c f_s \frac{k \frac{u_*}{u(h)} \frac{z_{0m}}{h}}{C_t^*} + kB_s^{-1} f_s^2 \quad (\text{III.28})$$

où :

- $f_c$  est le taux de couverture (ou fraction de couvert) et  $f_s$  son complément ( $f_c = 1 - f_s$ )
- $h$ , la hauteur du couvert
- $C_d$  le coefficient de traînée des éléments foliaires supposé prendre la valeur de 0,2
- $U(h)$  la vitesse du vent horizontal au sommet du couvert
- $C_t$  le coefficient de transfert thermique de la feuille (pour la plupart des couverts et des conditions environnementales,  $C_t$  est délimité par  $0,005 N \leq C_t \leq 0,075 N$ ,  $N$  étant le nombre de faces d'une feuille qui participe à l'échange de chaleur)
- $C_t^*$  le coefficient de transfert thermique du sol, donnée par  $C_t^* = Pr^{-2/3} Re^{*-1/2}$ , où  $Pr$  est le nombre de Prandtl (fixé à 0.71; Massman 1999b) et  $Re^*$  le nombre de Reynolds donné par  $Re^* = h_s u^*/\nu$ , avec  $h_s$  la hauteur de rugosité du sol nu et  $\nu$  la viscosité cinématique de l'air
- $n$  le coefficient d'extinction du profil de la vitesse du vent dans le couvert tel que

$$n = \frac{C_d LAI}{2u_*^2 / u(h)^2} \quad (\text{III.29})$$

Le rapport  $u^*/u(h)$  est paramétré comme suit :  $\frac{u_*}{u(h)} = c_1 - c_2 \exp(-c_3 C_d LAI)$  (III.30)

avec  $c_1 = 0,320$ ,  $c_2 = 0,264$  et  $c_3 = 15,10$  des constantes du modèle qui sont décrites par Massman (1997).

Enfin, la vitesse horizontale du vent au sommet du couvert  $u(h)$  est définie à partir du profil logarithmique au-dessus du couvert par :

$$u(h) = u_a \frac{\ln(h - d) - \ln(z_{0m})}{\ln(z - d) - \ln(z_{0m})} \quad (\text{III.31})$$

Le premier terme de l'équation (III.28) suit l'expression du modèle de Choudhury et Monteith (1988) pour un couvert dense, le troisième terme celle de Brutsaert (1982) pour une surface de sol nu (cf. Eq. III.23), tandis que le second terme décrit l'interaction entre la végétation et une surface de sol nu. La pondération quadratique basée sur le taux de couverture ( $f_c$  et son complément  $f_s$ ) permet de décrire les conditions de toutes situations entre la végétation dense et le sol nu.

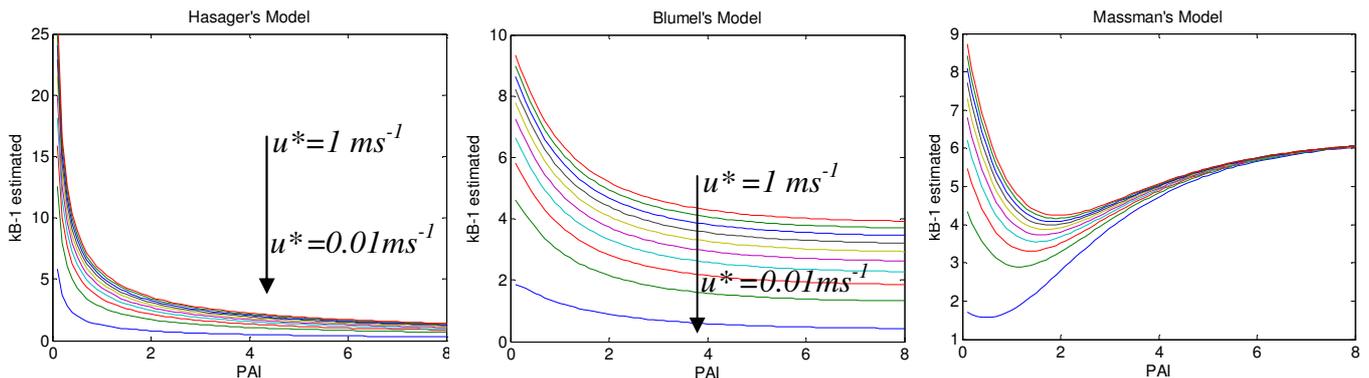
Comme pour les deux autres modèles, la hauteur de rugosité ( $z_{0h}$ ) pour le transfert de chaleur peut être déterminé par :

$$z_{0h} = \frac{z_{0m}}{\exp(kB^{-1})} \quad (\text{III.32})$$

### 2.2.2.4. Illustration du fonctionnement des modèles

La Figure III.19 illustre le comportement des trois modèles en fonction de l'indice  $PAI$  pour différentes vitesses de frottement  $u^*$  donnée. Le  $PAI$  est utilisé ici à la place du  $LAI$  pour tenir compte de l'ensemble des surfaces qui jouent sur les interactions concernées. Les trois modèles ont des comportements différents en fonction des deux variables présentées, et en particulier du  $PAI$ . Les  $kB^{-1}$  décrits par les modèles de Hasager et Blümel décroissent lorsque le  $PAI$  augmente. Le modèle de Massman présente un profil différent avec tout d'abord une baisse, puis une augmentation après un niveau de  $PAI$  relativement bas (entre 1 et 2). Aux fortes valeurs de  $PAI$  les trois modèles présentent une saturation. Les valeurs de  $kB^{-1}$  aux faibles  $PAI$  sont similaires pour les modèles de Massman et de Blümel (2 à 9 selon  $u^*$ ) et beaucoup plus forte pour le modèle d'Hasager (5-25). Les valeurs pour les forts  $PAI$  sont différentes entre les 3 modèles (Hasager (0,5-2) < Blümel (0,5-4) < Massman (6)). Les variations des  $kB^{-1}$  à faibles  $PAI$  d'un coté et à forts  $PAI$  d'un autre coté en fonction de la vitesse de frottement déterminent les variations de  $kB^{-1}$  à un  $PAI$  intermédiaire donné montrant des différences de comportement notable entre les trois modèles (mais toujours une augmentation de  $kB^{-1}$  avec  $u^*$  et donc le niveau de turbulence).

Par ailleurs, tous les modèles sont très dépendants de la rugosité mécanique  $zO_m$ .



**Fig. III.19. Profils des modèles de  $kB^{-1}$  suivant les valeurs de  $PAI$  présenté pour toutes les vitesses de frottement entre  $0.01 \text{ ms}^{-1}$  et  $1 \text{ ms}^{-1}$**

### 2.2.3. Détermination de $kB^{-1}$ sur les données expérimentales

Le test des modèles de rugosité thermique présentés ci-dessus nécessite de disposer de données de  $kB^{-1}$  dans des situations réelles. La difficulté dans la détermination de  $kB^{-1}$  découle du fait qu'il ne peut pas être mesuré directement et que sa valeur doit être dérivée à partir de la mesure de différentes variables à la fois aérodynamiques et thermiques. Toute erreur sur les mesures de ces variables contribue aux incertitudes des valeurs de  $z0_h$  et de  $kB^{-1}$ .

Dans cette section, nous évaluons la possibilité de déterminer le paramètre  $kB^{-1}$  sur les données acquises sur le site 'Flux et Télédétection' d'Avignon (site Grue) en analysant plus particulièrement les incertitudes sur le calcul.

#### 2.2.3.1. Méthode de calcul du paramètre $kB^{-1}$

$kB^{-1}$  peut être calculé à partir des équations du flux de chaleur sensible et de la mesure des quantités qui interviennent dans ces équations. Pour rappel, l'expression classique pour exprimer le flux de chaleur sensible est la suivante :

$$H = \rho c_p \frac{T_s - T_a}{r_a} \quad (\text{III.33})$$

avec la rugosité aérodynamique  $r_a$ ,  $T_s$  la température de surface calculée à partir des températures de brillance et  $T_a$  la température de l'air à 2 m.

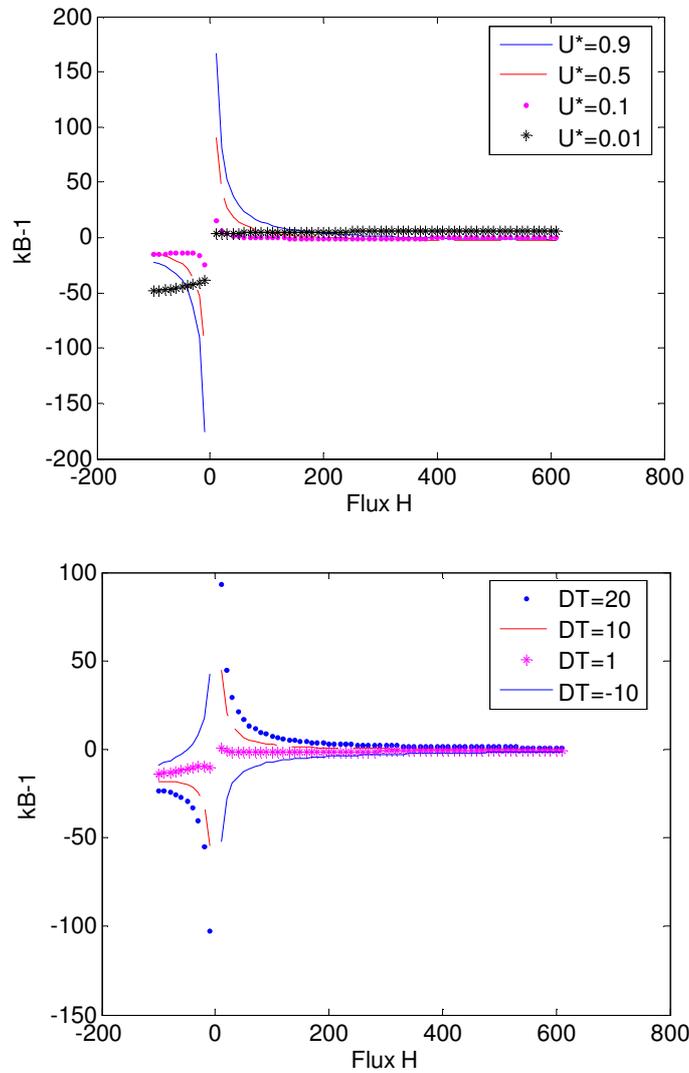
$$r_a = \frac{1}{ku^*} \left( \ln\left(\frac{z}{z0_h}\right) - \Psi_h \right) \quad (\text{III.34})$$

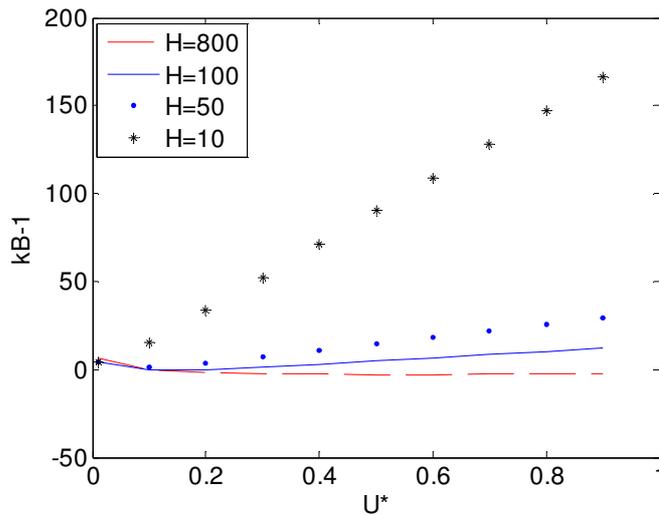
A partir des deux expressions (III.33 et III.34), on peut obtenir une expression permettant le calcul direct de la rugosité thermique ( $z0_h$ ) avec l'équation suivante et du paramètre  $kB^{-1}$  au travers de l'**Equation III.9**.

$$z0_h = \frac{z}{\exp\left(k\rho c_p \frac{u^* (T_s - T_a)}{H} + \Psi_h\right)} \quad (\text{III.35})$$

Ainsi  $kB^{-1}$  peut être déterminé à partir de la mesure de  $u^*$ ,  $H$ , et de la différence entre les températures de surface et de l'air ( $T_s - T_a$ ).

Les graphiques présentés sur la Figure III.20 ci-dessous illustrent le comportement du paramètre  $kB^{-1}$  en fonction des variations des différentes variables utilisées dans son calcul ( $H$ ,  $DT = (T_s - T_a)$  et  $u^*$ ).





**Fig. III.20. Comportement du paramètre  $kB^{-1}$  calculé à partir du flux  $H$  : a)  $kB^{-1}$  calculé pour  $DT = 3 \text{ K}$  fixé ; b)  $kB^{-1}$  calculé pour  $u^* = 0,5 \text{ ms}^{-1}$  fixé; c)  $kB^{-1}$  calculé pour  $DT = 3 \text{ K}$  fixé**

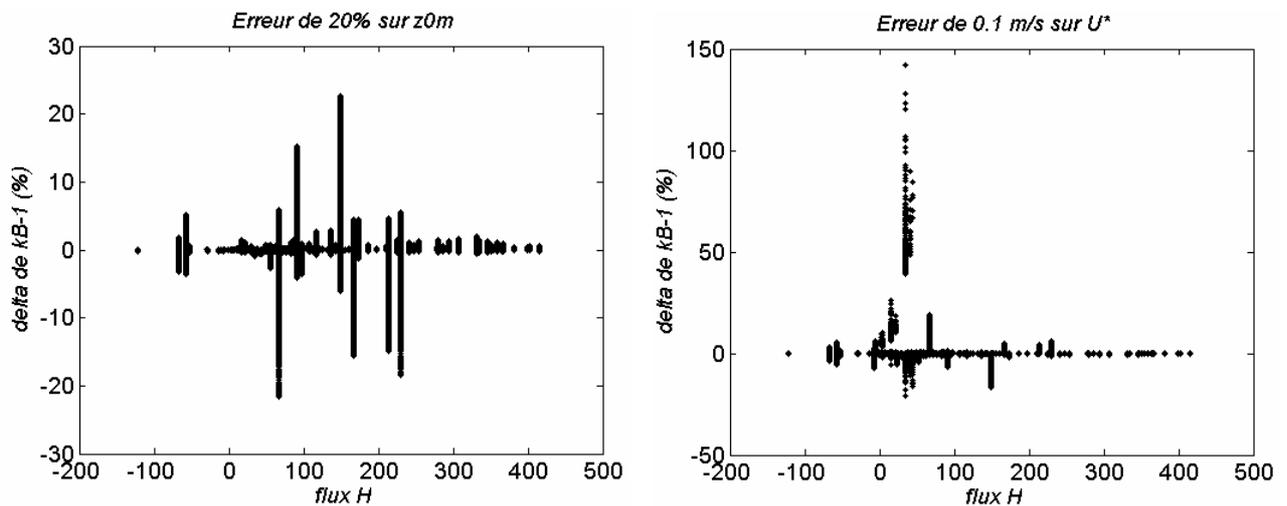
Les valeurs de  $kB^{-1}$  apparaissent positives ou négatives selon les conditions. Les valeurs négatives sont très souvent observées lorsque  $H$  est négatif (mais pas seulement). Cela signifie dans ce cas que  $z0_h$  est supérieur à  $z0_m$ , ce qui bien qu'ayant été rapporté dans certaines études est peu fréquent. Par exemple pour une surface homogène de sol nu, des valeurs de  $kB^{-1}$  de -7.0 à 7.0 ont été obtenues par Verhoef et al. (1997). Par ailleurs, la Figure III.20 montre que les valeurs de  $kB^{-1}$  peuvent être très élevées (en valeur absolue) lorsque  $H$  se rapproche de zéro, et ce d'autant plus que  $DT$  et/ou  $u^*$  sont forts. Les valeurs obtenues sont alors très fortes ( $> 50$ ) ce qui excède la gamme des valeurs obtenues dans la littérature, qui se situent plutôt entre 1 et 20 comme par exemple obtenus par Kustas et al. (1989) ou Troufleau et al. (1997). Il importe de signaler cependant que certaines des valeurs présentés sur les graphes correspondent à des situations physiquement impossibles, comme par exemple des flux  $H$  positifs avec des différences de températures négatives (hors erreur de mesure, cet effet ne doit pas apparaître sur les données réelles).

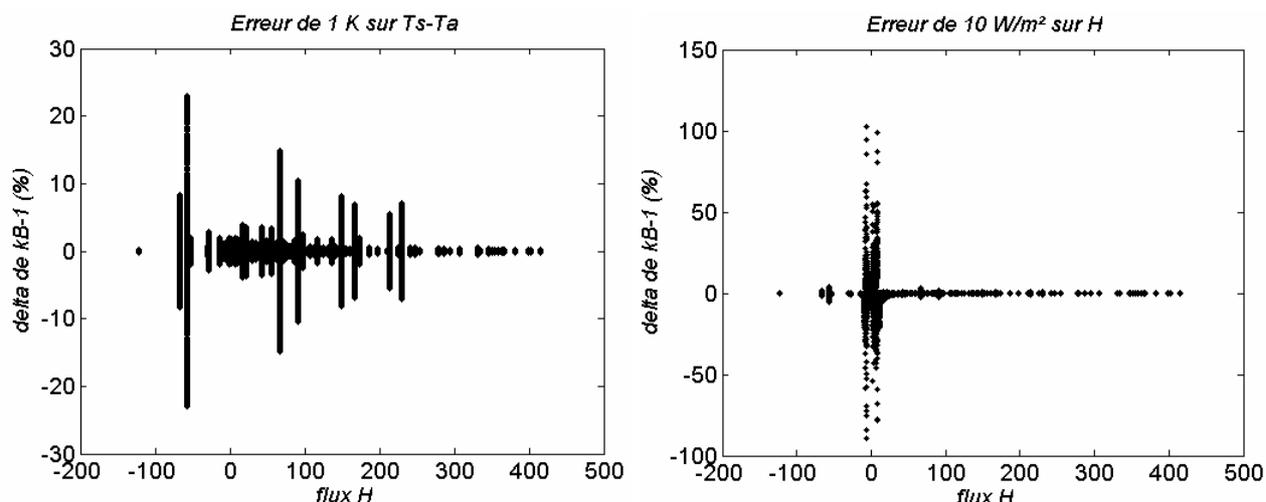
L'examen des courbes de la Figure III.20 montre également que dans certaines conditions, de faibles variations des données d'entrée peuvent entraîner de fortes variations de  $kB^{-1}$ , en particulier aux faibles valeurs de flux  $H$ . Avant de pouvoir utiliser les estimations de  $kB^{-1}$  pour le test des formulations proposées plus haut, il importe donc d'en vérifier la précision en

fonction de la précision des mesures des différentes variables nécessaires au calcul ( $H$ ,  $DT$  et  $u^*$ ).

### 2.2.3.2. Analyse de la sensibilité du calcul de $kB^{-1}$ aux mesures

Nous avons évalué la sensibilité de  $kB^{-1}$  aux erreurs sur les variables mesurées au moyen des données acquises sur le site ‘Flux et Télédétection’ (données à 12 h en 2004 sur une culture de blé). Pour chaque jour de données, nous avons généré des erreurs autour des valeurs mesurées en tirant 1000 points dans des lois uniformes de  $\pm 10 \text{ Wm}^{-2}$  pour le flux  $H$ ,  $\pm 0.1 \text{ ms}^{-1}$  pour la vitesse de frottement  $u^*$  et de  $\pm 1 \text{ K}$  pour la différence de températures  $DT$ . Une perturbation aléatoire de  $\pm 20\%$  a été également ajoutée aux valeurs de rugosité mécanique  $z_{0m}$  pour en tester l’impact au travers de l’équation III.9. Ces niveaux d’erreurs correspondent bien aux niveaux d’erreurs constatés dans les mesures du site d’Avignon. La figure III.21 présente les variations de  $kB^{-1}$ , calculé au moyen de l’équation III.35, relativement au calcul sans introduction d’erreur.

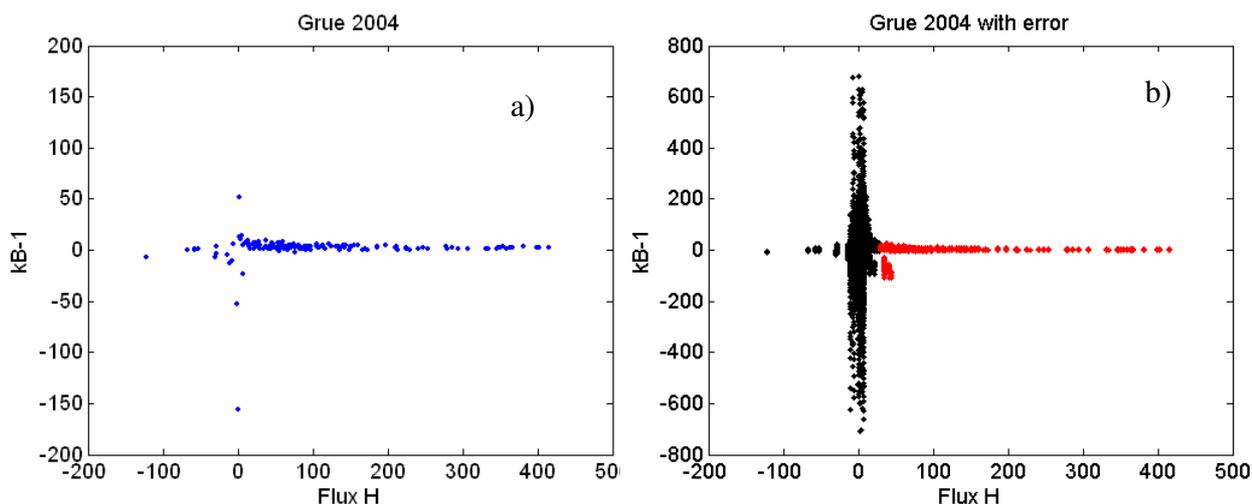




**Fig. III.21. Pourcentages de variation de  $kB^{-1}$  en fonction des différentes erreurs introduites dans les variables permettant le calcul de  $kB^{-1}$  ( $H$ ,  $DT$  et  $u^*$  et  $z0_m$ )**

On peut constater que les variations de  $kB^{-1}$  peuvent être très importantes (jusqu'à 150%) pour les deux variables  $u^*$  et  $H$ , et que les deux autres variables  $z0_m$  et  $DT$  ont un impact plus limité (au plus 25%). Comme le laissait supposer les résultats de la Figure III.20 dans la section précédente, les plus fortes erreurs sont obtenues aux alentours des flux de chaleur sensible les plus faibles, soit au sein d'une fenêtre de  $\pm 20$  à  $\pm 30 \text{ Wm}^{-2}$ . Les fortes variations liés à  $u^*$  semblent également se produire en conditions de faible flux de quantité de mouvement (voir également les résultats sur la Figure III.20). La Figure III.22 permet de comparer les valeurs de  $kB^{-1}$  obtenues sans considérer d'erreur (à gauche) aux valeurs obtenues en considérant toutes les erreurs simultanément (à droite). Elle confirme que les valeurs du flux  $H$  près de 0 présentent une variabilité importante associée à des risques d'erreur important. Ceci dit, dans le cas de ce jeu de données, les situations dans lesquelles le flux  $H$  est faible (dans notre cas à midi solaire) sont peu fréquentes.

Pour conclure, on peut constater que l'équation permettant de calculer  $kB^{-1}$  à partir des mesures de terrain (au midi solaire) donne des résultats fiables dans des conditions de flux  $H$  supérieur à  $30 \text{ Wm}^{-2}$ , les autres facteurs d'erreurs possibles ayant un impact relativement faible. Dans le cas ici d'une culture sèche, ces conditions de validité du calcul sont les plus fréquentes. Pour d'autres situations, on peut penser en particulier à des cultures d'été irriguées, les situations pour lesquelles l'application du calcul serait risquée, pourraient être beaucoup plus fréquentes. Dans la suite de ce travail, les calculs effectués pour les situations à faible flux  $H$  ont été supprimés des analyses.

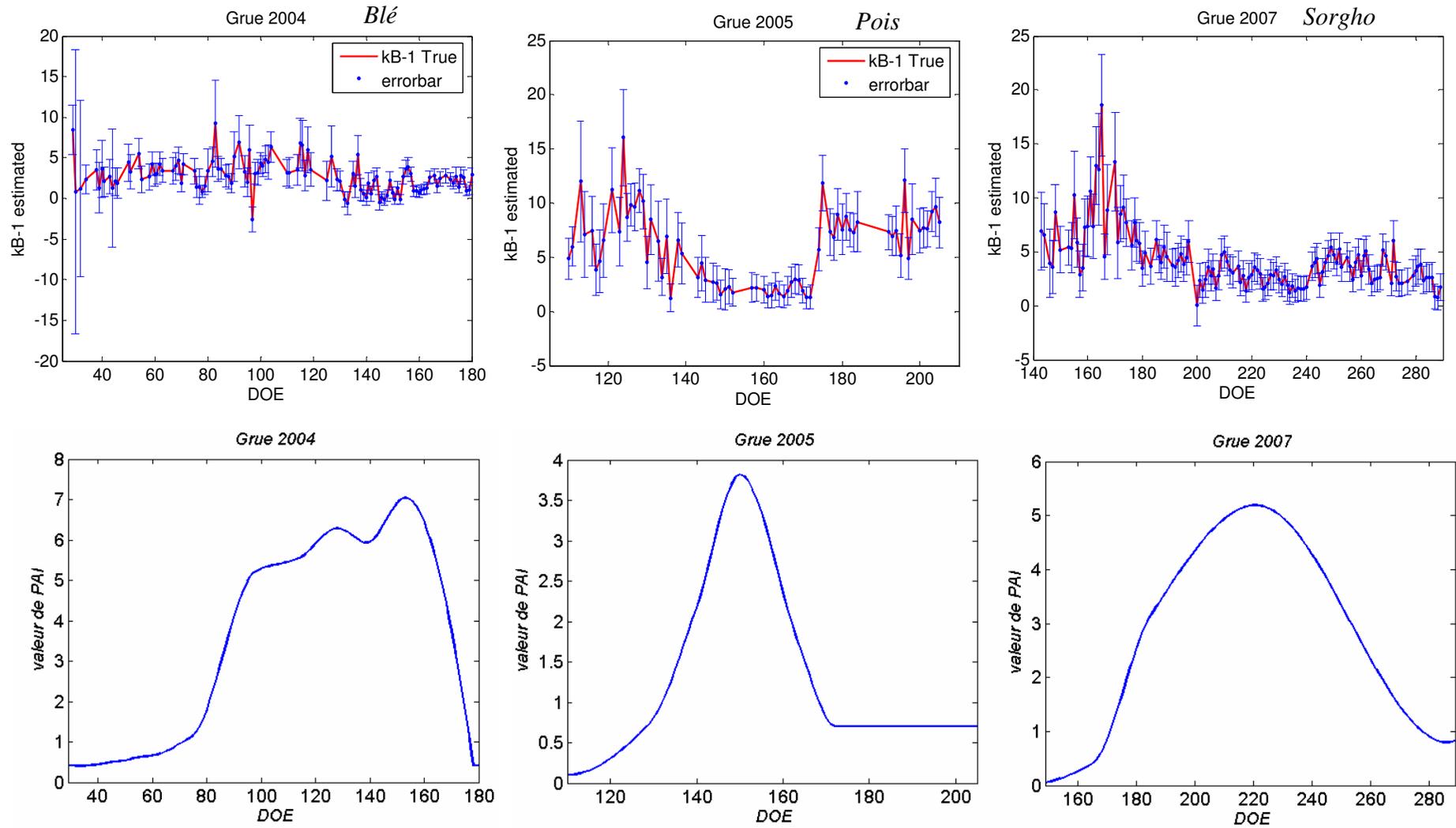


**Fig. III.22. a) Estimation du paramètre  $kB^{-1}$  à partir du flux  $H$  sur les données GRUE 2004 et b) Estimation de  $kB^{-1}$  avec ajout d'erreur sur toutes les variables permettant de le calculé (en rouge les résultats pour flux  $H > 30 \text{ Wm}^{-2}$ )**

### 2.2.3.3. Evolution des $kB^{-1}$ calculés sur les données su site 'Flux et Télédétection' d'Avignon

La Figure III.23 présente l'évolution des  $kB^{-1}$  au cours des trois saisons de culture (avec leurs barres d'erreur associées par rapport à l'écart type). La représentation des barres d'erreur, après avoir éliminé les données avec des flux  $H$  inférieurs à  $30 \text{ Wm}^{-2}$ , met en évidence les jours où la sensibilité du calcul du paramètre  $kB^{-1}$  est importante, ce qui pourrait souligner des incertitudes dans l'analyse des estimations ces jours ci. Les valeurs de  $kB^{-1}$  obtenues sont dans la gamme des valeurs présentées dans la littérature. On peut noter que certaines valeurs de  $kB^{-1}$  vraies sont négatives (e.g. en 2004 pour DOE 97, 132, 133) ce qui semble être lié ici à des conditions atmosphériques particulières conduisant à des erreurs sur la détermination de la différence de températures ( $T_s - T_a$ ). Liu et al. (2007) ont montré que les valeurs négatives de  $kB^{-1}$  sont obtenues sur une culture de maïs lorsque la température de surface est très proche de la température de l'air.

La Figure III.23 propose en regard des évaluations de  $kB^{-1}$  et les évolutions de  $PAI$ , ce qui permet de bien noter la baisse de  $kB^{-1}$  alors que la densité de la végétation augmente (tel que le montre aussi les modèles).



**Fig. III.23.** En haut: évolution de  $kB^{-1}$  (+/- barres d'erreur par rapport à l'écart type) au cours des trois saisons de culture étudiées (2004: blé, 2005: pois; 2007: sorgho). En bas : évolutions des PAI. DOE représente le nombre de jour depuis le début de l'année.

Le tableau III.15 présente les erreurs d'incertitudes sur l'estimation de  $kB^{-1}$  obtenues pour les différentes saisons de cultures sur le site 'Flux et Télédétection' d'Avignon. Comparés aux résultats de Liu et al. (2007) et Troufleau et al. (1997) les résultats obtenus sont dans les mêmes gammes d'erreur. Troufleau et al. (1997) ont relevé certaines incertitudes sur le paramètre  $kB^{-1}$  résultant de l'ajout des erreurs sur la rugosité  $z0_m$  et le flux de chaleur sensible  $H$ . Leur étude sur une surface de savane en jachère montre que des erreurs de 20% sur la rugosité mécanique ( $z0_m$  égale à 0.25 m) et respectivement sur le flux  $H$  engendrent des erreurs de 13% et 35% sur le calcul de  $kB^{-1}$  lorsque celui-ci vaut 5, et des erreurs de 8% et 25% lorsque  $kB^{-1}$  vaut 15. L'analyse de sensibilité de Liu et al. (2007) a aussi montré qu'une sous-estimation de 50% de la rugosité mécanique  $z0_m$  (valeur de référence de 0.016 m) peut entraîner une erreur de 25% en moyenne sur l'estimation de  $kB^{-1}$ .

**Tab. III.15. RMSE et RRMSE (en parenthèse) sur  $kB^{-1}$  estimé à partir du flux  $H$  avec les incertitudes d'erreur sur les données GRUE (après suppression des  $H < 30 \text{ Wm}^{-2}$ ; all  $\varepsilon$  = représente toutes les erreurs introduites)**

2004					2005	2007
$z0_m \pm 20\%$	$u^* \pm 0,1 \text{ ms}^{-1}$	$(T_s - T_a) \pm 1 \text{ K}$	$H \pm 10 \text{ Wm}^{-2}$	all $\varepsilon$	all $\varepsilon$	all $\varepsilon$
0,57 (21%)	0,77 (29%)	1,13 (43%)	0,4 (15%)	1,56 (59%)	2,28 (40%)	1,96 (48%)

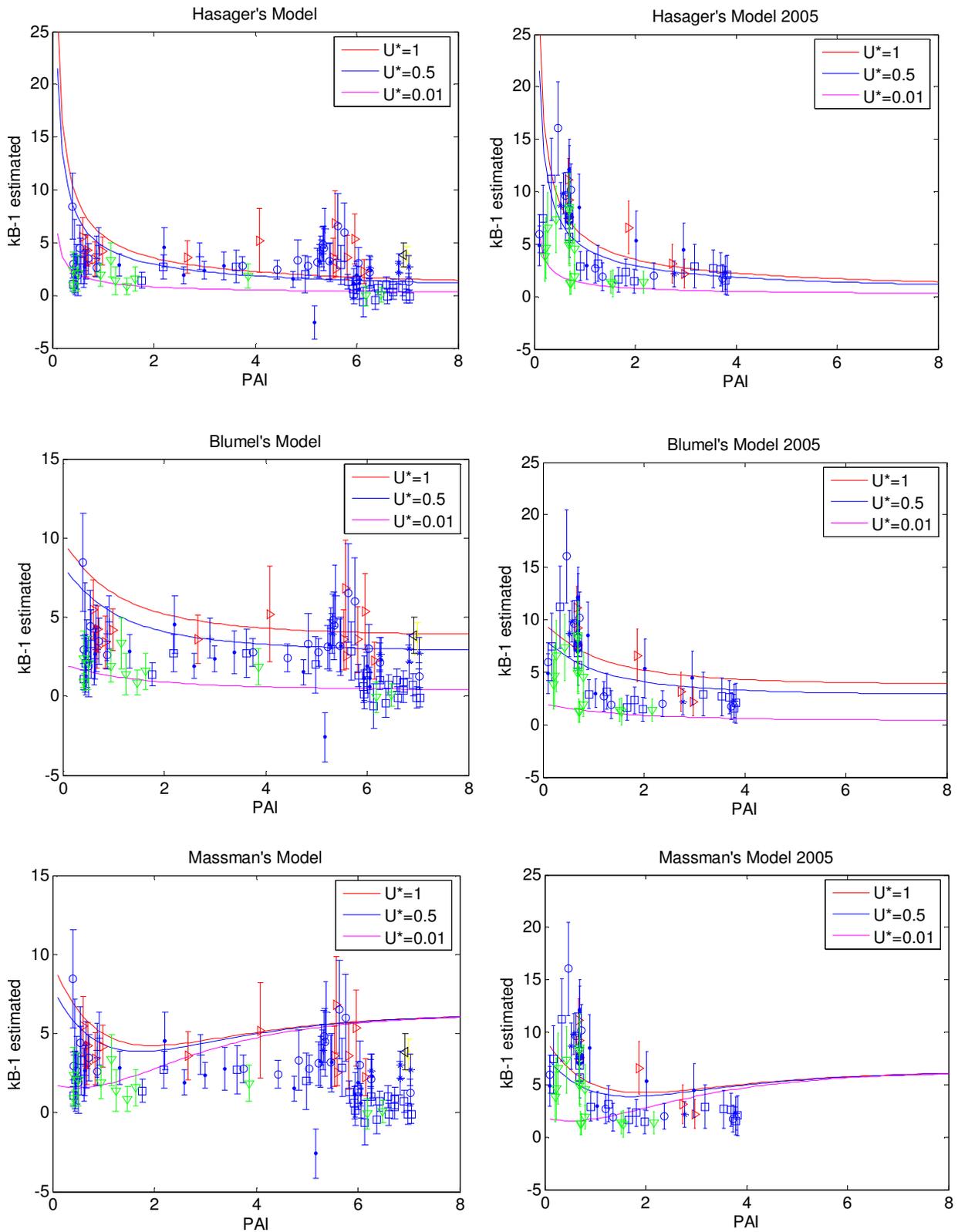
## 2.2.4. Application des modèles de $kB^{-1}$ sur les données expérimentales

Nous avons appliqué les trois modèles de  $kB^{-1}$  décrits dans les paragraphes 2.2.2.1, 2.2.2.2 et 2.2.2.3 ci-dessus aux données expérimentales, d'une part pour les comparer aux estimations de  $kB^{-1}$  analysées ci-dessus, et d'autre part, inclus au sein du calcul du flux de chaleur sensible  $H$ , pour en évaluer l'impact sur les flux (le flux de chaleur sensible  $H$  est ici calculé au moyen de l'équation III.33 avec un forçage des températures de surface et de l'air).

### 2.2.4.1. Evaluation des modèles sur les données 'Flux et Télédétection' d'Avignon

Les comparaisons des modèles de  $kB^{-1}$  aux estimations sur les données expérimentales à midi solaire sont présentées dans la Figure III.24 (pour les années 2004 et 2005). Les résultats apparaissent difficiles à interpréter. D'une part, les estimations de  $kB^{-1}$  sont entachées d'erreurs souvent larges par rapport aux écarts qui existent entre les différentes expressions des modèles. D'autre part, les trois modèles présentent des résultats très différents qui ne semblent pas pouvoir être discriminés à partir des observations. Les données expérimentales présentent un accord global avec les tendances générales des modèles : diminution de  $kB^{-1}$  alors que la densité de la végétation ( $PAI$ ) augmente (avec dans certains cas des ré-augmentations comme pour le modèle de Massman), et valeurs plus fortes lorsque la vitesse de frottement  $u^*$  est plus forte. Cependant les valeurs expérimentales apparaissent souvent plus élevées que les modèles, hormis dans le cas des  $PAI$  au-delà de 2 ou 3 avec le modèle de Massman. L'évaluation des modèles de  $kB^{-1}$  au travers de l'évaluation des calculs de flux  $H$  semble susceptible de nous apporter plus d'information sur la qualité des modèles de  $kB^{-1}$ .

La figure III.25 présente l'évaluation, par rapport aux mesures de flux turbulent, des estimations de flux de chaleur sensible au moyen des équations III.33 et III.34 en y intégrant les estimations de  $kB^{-1}$  par les différents modèles étudiés (données du site 'Flux et Télédétection' d'Avignon à midi solaire pour les cycles de cultures de 2004, 2005 et 2007). Les évaluations numériques des résultats sont données dans le Tableau III.16.



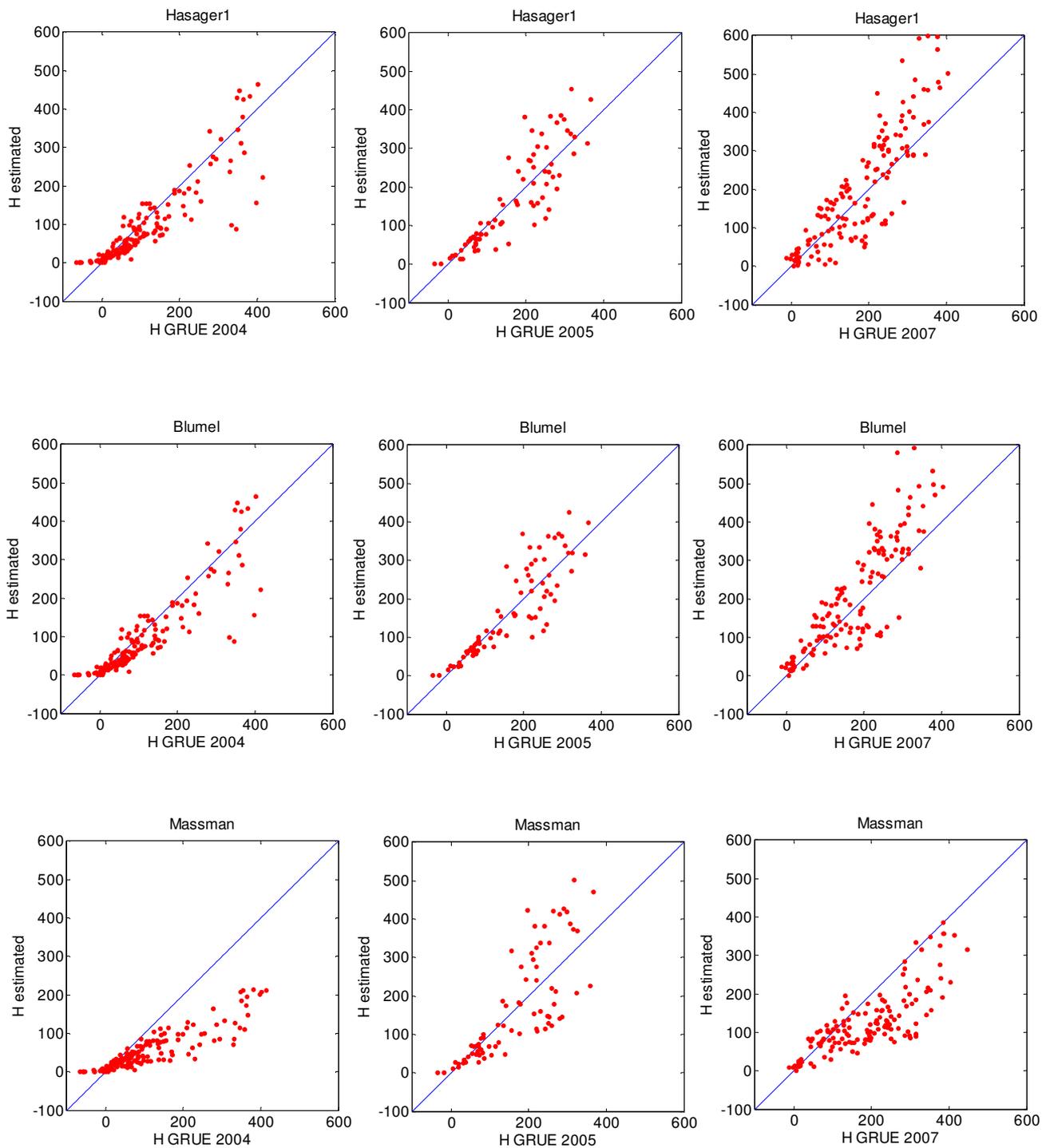
**Figure III.24 : comparaison des modèles de  $kB^{-1}$  aux estimations expérimentales sur les saisons 2004 (à gauche) et 2005 (à droite). Les modèles sont présentés pour trois valeurs de  $u^*$  (1, 0,5 et 0,01  $ms^{-1}$ ) et les données selon des classes de 0,1  $ms^{-1}$  (voir légende à la page suivante)**

◇	0.9<U*<1.0
◁	0.8<U*<0.9
+	0.7<U*<0.8
*	0.6<U*<0.7
▷	0.5<U*<0.6
○	0.4<U*<0.5
•	0.3<U*<0.4
□	0.2<U*<0.3
▽	0.1<U*<0.2

**Tab. III. 16.** *RMSE* et *RRMSE* du flux *H* estimés par les modèles de  $kB^{-1}$  par rapport aux données GRUE (cf. Fig. II.25)

	2004			2005			2007		
	<i>RMSE</i> ( $Wm^{-2}$ )	<i>RRMSE</i> (%)	<i>Pente</i> (-)	<i>RMSE</i> ( $Wm^{-2}$ )	<i>RRMSE</i> (%)	<i>Pente</i> (-)	<i>RMSE</i> ( $Wm^{-2}$ )	<i>RRMSE</i> (%)	<i>Pente</i> (-)
<i>Hasager</i>	56	50	0,92	59	36	1,05	96	52	1,17
<i>Blümel</i>	55	46	0,91	56	33	1,05	99	54	1,25
<i>Massman</i>	92	82	0,52	84	51	1,06	92	49	0,67

On peut noter dans un premier temps que l'allure générale des résultats avec le modèle Hasager est très similaire à ceux de Blümel. Par contre les résultats du modèle de Massman présentent des dispersions plus importantes que les deux autres (forts *RMSE* en 2004 et 2005, années pour lesquelles on note une forte sous estimation de *H* par ce modèle comme indiquée par l'information sur la pente dans le Tableau III.16). Ces différences sont à mettre en relation avec les différences entre le modèle et les estimations de  $kB^{-1}$  apparaissant sur les Figures III.24. Les niveaux d'erreurs constatés sur *H* sont en accord, quoique dans la gamme haute des valeurs, avec les niveaux d'erreurs obtenus dans la littérature lorsque l'on s'intéresse à l'évaluation des flux turbulents à partir de mesures acquises au sol. Signalons également que le test du modèle de Hasager sur les données Alpilles à partir des mesures aéroportées avait donné des résultats similaires avec des *RMSE* se situant entre  $70 Wm^{-2}$  et  $85 Wm^{-2}$  (Hasager et al., 2002a).



**Fig. III.25. Comparaison des calculs de flux de chaleur sensible  $H$  incorporant les trois modèles de  $kB^{-1}$  avec les mesures expérimentales à midi solaire sur le site ‘Flux et Télédétection’ d’Avignon (Site GRUE). Les données de flux sont en  $Wm^{-2}$ .**

### 2.2.4.2. Analyse de l'impact de la détermination de la rugosité aérodynamique $z0_m$ sur les modèles de $kB^{-1}$ et les calculs de flux $H$

La rugosité aérodynamique ou mécanique ( $z0_m$ ) est un paramètre essentiel dans les modèles de  $kB^{-1}$  ci-dessus. Par conséquent, une erreur sur les mesures de  $z0_m$  pourrait probablement modifier la précision des résultats des modèles de  $kB^{-1}$ . L'analyse de l'impact de la rugosité mécanique dans le processus d'estimation du flux de chaleur sensible  $H$  et notamment du paramètre  $kB^{-1}$  permet d'évaluer la précision des modèles. De ce fait, nous avons évalué deux approches permettant de fixer les valeurs de la rugosité mécanique ( $z0_m$ ). Il s'agit d'une part la formulation standard de Brutsaert (1982) rappelé par l'équation (III.36) et d'autre part une formule plus complexe (Raupach, 1992 ; 1994 ; Olioso, 1992 ; Troufleau et al., 1997) définie par les équations III.37 - III.38.

$$z0_{m1} = 0.136h \quad (III.36)$$

avec  $h$  la hauteur du couvert.

$$z0_{m2} = z0_s \frac{0.2 - C_d LAI}{0.2} + 0.28h(C_d LAI)^{1/2} \quad 0 \leq C_d LAI \leq 0.2 \quad (III.37)$$

et

$$z0_{m2} = 0.30h \left(1 - \frac{d}{h}\right) \quad 0 \leq C_d LAI \leq 0.2 \quad (III.38)$$

avec

$$d = h \left[ \ln \left( 1 + (C_d LAI)^{1/6} \right) + 0.03 \ln \left( 1 + (C_d LAI)^6 \right) \right] \quad (III.39)$$

où  $z0_s$  est la longueur de rugosité sur le sol nu (e.g. 0.01 m),  $C_d$  le coefficient de traînée des feuilles et pour rappel  $d$  la hauteur de déplacement.

Ces deux calculs aboutissent à des rugosités qui peuvent être très différentes en fonction du niveau de développement du couvert en particulier lorsque le  $LAI$  est important. Par exemple pour une hauteur de 1m, un  $LAI$  de 6 et un coefficient de traînée de 0,2,  $z0_{m1}$  vaut 0,136 et  $z0_{m2}$

juste un peu plus de la moitié (0,075). Cette différence augmente avec la densité de la végétation. Les différences peuvent également être importantes pour les couverts peu développés ( $C_d LAI < 0,1$ ). Des valeurs équivalentes sont obtenues pour  $C_d LAI \sim 0,2$ . Les deux approches sont appliquées sur les données du site 'Flux et Télédétection'. Les statistiques des différences entre les deux méthodes sont récapitulées par le tableau III.17 ci-dessous.

**Tab. III.17. Quantification des différences entre les deux estimations de rugosité aérodynamique :  $RMSE$  et  $RRMSE$  entre les deux  $z_{0m1}$  et  $z_{0m2}$  calculés sur les données GRUE**

	$RMSE$ (m)	$RRMSE$ (%)	$Pente$ (-)
2004	0,04	49	0,62
2005	0,01	20	0,84
2007	0,03	33	0,73

On note une erreur moyenne de 34% entre les deux méthodes de calcul de la rugosité mécanique (Tab. III.17). Cette erreur correspond à une erreur de 7% sur le paramètre  $kB^{-1}$  (Tab. III.18) et respectivement de 37% sur l'estimation du flux de chaleur sensible  $H$  (Tab. III.19) avec le modèle Hasager. Cet impact correspond respectivement à 5% et 35% avec le modèle de Blümel sur le paramètre  $kB^{-1}$  et sur le flux  $H$  contre 1% et 30% avec le modèle de Massman. Soulignons ici l'impact faible sur le modèle de  $kB^{-1}$  de Massman (1%) comparé aux modèles de Hasager et de Blümel. Par contre cette erreur a un impact important sur l'estimation du flux de chaleur sensible. Cela permet de justifier notre choix d'assimiler ce paramètre ( $z_{0m}$ ) dans le processus d'assimilation de données pour améliorer l'estimation des flux énergétiques de surface.

**Tab. III.18.  $RMSE$  et  $RRMSE$  sur le paramètre  $kB^{-1}$  calculé avec  $z_{0m1}$  et  $z_{0m2}$  sur les données GRUE par les modèles**

	2004			2005			2007		
	$RMSE$ (-)	$RRMSE$ (%)	$Pente$ (-)	$RMSE$ (-)	$RRMSE$ (%)	$Pente$ (-)	$RMSE$ (-)	$RRMSE$ (%)	$Pente$ (-)
<i>Hasager</i>	0,10	4	0,99	0,10	3	0,98	1,9	14	1
<i>Blümel</i>	0,15	6	0,96	0,11	3	0,98	0,14	5	0,96
<i>Massman</i>	0,04	0,78	0,99	0,03	0,68	0,99	0,08	2	0,99

**Tab. III.19. RMSE et RRMSE sur le flux  $H$  calculé avec  $z_{0m1}$  et  $z_{0m2}$  sur les données GRUE par les modèles**

	2004			2005			2007		
	<i>RMSE</i> ( $Wm^{-2}$ )	<i>RRMSE</i> (%)	<i>Pente</i> (-)	<i>RMSE</i> ( $Wm^{-2}$ )	<i>RRMSE</i> (%)	<i>Pente</i> (-)	<i>RMSE</i> ( $Wm^{-2}$ )	<i>RRMSE</i> (%)	<i>Pente</i> (-)
<i>Hasager</i>	87	58	0,69	16	9	0,94	139	45	0,71
<i>Blümel</i>	77	53	0,70	15	8	0,94	142	45	0,70
<i>Massman</i>	32	45	0,76	13	8	0,94	60	38	0,75

Un autre point qui ressort de l'analyse est que l'impact des différences entre les rugosités mécaniques diffère fortement d'une année à l'autre. Cela peut s'expliquer par des niveaux et des répartitions différentes au cours de la saison des mesures de  $PAI$  et de hauteur de couvert en entrée des calculs de rugosité mécanique. En particulier, il importe de noter les différences de hauteurs de couverts entre les 3 cultures (~50 cm pour le pois en 2005, ~80 cm pour le blé en 2004 et ~1m40 pour le sorgho en 2007) et les différences dans les profils temporels de  $LAI$  (voir Figure III.21). Rappelons aussi ici que les modèles de  $kB^{-1}$  sont sensibles aux variations de  $LAI$  essentiellement aux faibles niveaux de  $LAI$  (cf. Figure III.19).

#### 2.2.4.3. Evaluation des modèles sur les données Alpilles

Les modèles de  $kB^{-1}$  ont été évalués au travers des estimations de flux de chaleur sensible sur les 16 dates des mesures aéroportées de la température de surface au cours de l'expérimentation Alpilles (Tab. III.20). Les données d'évaluation concernent le flux de chaleur sensible  $H$  à midi solaire sur les 7 parcelles instrumentées (101, 102, 120, 121, 203, 214, et 501). Les niveaux d'erreurs se situent dans la même gamme que les erreurs constatées sur les données du site d'Avignon. On peut remarquer que le modèle de Massman donne un résultat plus cohérent que les modèles de Hasager et de Blümel. On peut aussi noter une dispersion plus forte des résultats en utilisant le rapport des deux rugosités, aérodynamique et thermique, constant et correspondant à une valeur de  $kB^{-1}$  de 2,3 (comme fait dans PBLs), montrant bien l'inadaptation de ce choix.

**Tab. III.20. *RMSE* et *RRMSE* sur le flux *H* estimé à 12 h sur les données Alpilles avec les modèles de  $kB^{-1}$  (sur les mesures des 7 parcelles durant les 16 jours des mesures aéroportées de la température de surface)**

	Données Alpilles		
	<i>RMSE</i> ( <i>Wm-2</i> )	<i>RRMSE</i> (%)	<i>Pente</i> (-)
<i>Hasager</i>	86	52	0,67
<i>Blümel</i>	77	47	0,63
<i>Massman</i>	74	45	0,73
$z0_m/z0_h=10$	99	57	1,14

### 2.2.5. Conclusion sur le paramètre $kB^{-1}$

On peut retenir de cette étude que le paramètre  $kB^{-1}$  peut prendre des valeurs très variables en fonction des conditions. Dans beaucoup de modèle de transfert, comme dans PBLs, l'utilisation d'une valeur constante et relativement faible est présentée comme adaptée à la simulation du fonctionnement de couverts végétaux développés et homogènes. Elle ne semble pas adaptée au problème que nous cherchons à traiter, la simulation du fonctionnement des couverts à l'échelle d'un paysage agricole au sein duquel on trouve une grande variabilité de couvert. Les modèles de  $kB^{-1}$  analysés ci-dessus donnent une première réponse de modélisation de ce paramètre en fonction de paramètre pouvant être reliés à cette variabilité (comme le  $LAI$ ). Nous avons analysé les performances de trois paramétrisations de  $kB^{-1}$  à partir de données expérimentales. Les comparaisons par rapport à des estimations de  $kB^{-1}$  faites à partir des mesures de terrains montrent que les 3 modèles étudiés ont des comportements différents. Il est cependant difficile de les évaluer précisément en raison, d'une part de la variabilité des estimations, et d'autre part du possible impact des erreurs de mesures qui entraînent des risques d'imprécision importants sur les valeurs du paramètre (nous avons aussi montré que les estimations ne pouvaient pas du tout être considérées comme fiables pour des flux de chaleur sensible inférieurs à  $30 \text{ Wm}^{-2}$ ). Les valeurs estimées à partir des mesures de terrain et les valeurs simulées sont bien dans la gamme des valeurs rapportée dans la bibliographie. Intégrées au calcul du flux de chaleur sensible, les paramétrisations de  $kB^{-1}$  ont également été évaluées directement par rapport aux mesures de flux de chaleur sensible, montrant de fortes similitudes entre les modèles de Blümel et de Hasager. Le modèle de Massman présente des performances moindres sur les données du site 'Flux et Télédétection' d'Avignon, mais apparaît plus cohérent sur les données Alpilles. De plus le modèle de Massman montre une stabilité aux erreurs (émissivité,  $LAI$ ,  $zO_m$ ) plus intéressante par rapport aux deux autres.

### 3. Applications : étude comparative des versions 1 & 2 du modèle PBLs par rapport aux données Alpilles

Dans cette section, nous avons intégré les deux modifications majeures ( $G/R_n$  et  $kB^{-1}$ ) dans le modèle PBLs (nommé version 2, PBLs2). Le modèle PBLs décrit dans la section 2 de la partie II (noté ici version 1, PBLs1) a été utilisé dans plusieurs applications telle que Courault et al. (2008) et Courault et al. (2009b). Les résultats de toutes ces applications et les études des deux paramètres étudiés ci-dessus ont montré que la paramétrisation de ces rapports n'est pas forcément appropriée dans certaines conditions. En effet, prescrire une valeur constante pour une étude sur une zone hétérogène souligne des sous-estimations ou surestimations des flux énergétiques de surface par endroit. La différence notoire entre les deux versions du modèle PBLs réside sur les paramètres  $c_0$  ( $G/R_n$ ) et  $z0_h$  ( $kB^{-1}$ ).

- **Dans la version 1**, le flux conductif du sol ( $G$ ) est estimé comme une simple proportion du flux du rayonnement net ( $R_n$ ) sur toute la zone, ainsi que la rugosité thermique ( $z0_h$ ) définie par un rapport à la rugosité aérodynamique ( $z0_m$ ).
- **Dans la version 2**, suite aux travaux présentés au dessus, nous avons intégré l'équation (III.1) de Choudhury et al. (1987) et les équations (III.5 et III.7) de Santanello et Friedl (2003) pour représenter le rapport  $G/R_n$  selon les configurations couvert végétal et sol nu. Par ailleurs, nous avons opté pour le modèle de Massman pour représenter le  $kB^{-1}$ .

Pour étayer les modifications apportées par les nouvelles formulations introduites dans le modèle PBLs, nous avons effectué une étude comparative des deux versions sur les parcelles de la zone Alpilles (le modèle étant appliqué à l'ensemble de la zone Alpilles à partir des informations utilisées comme valeurs a priori dans les tests d'assimilation présentés dans la cinquième partie de ce mémoire ; voir aussi les **sections 5.3** et **5.4** de la **partie II** pour la description des données).

Le Tableau III.21 ci-dessous présentent les résultats comparés des deux versions de PBLs par rapport aux mesures des flux de surface et des températures de surface et de l'air mesurées sur les quatre parcelles 101, 120, 203 et 214 dans la zone Alpilles tout au long des journées du 18 avril 1997 (DOE 474 de 5h à 20h) et du 26 mars 1997 (DOE 451). La Figure III.26 présente les comparaisons des flux de chaleur  $H$  et  $G$ . D'autres figures présentent d'autres résultats dans l'**Annexe III**.

**Tab. III.21. Comparaison des évaluations des flux et des températures simulés par les deux versions du modèle PBLs (PBLs1 et PBLs2) par rapport aux mesures effectuées sur les parcelles 101, 120, 203 et 214 entre 5 et 20 h pour les jours (DOE) 474 et 451 de l'expérimentation Alpilles.**

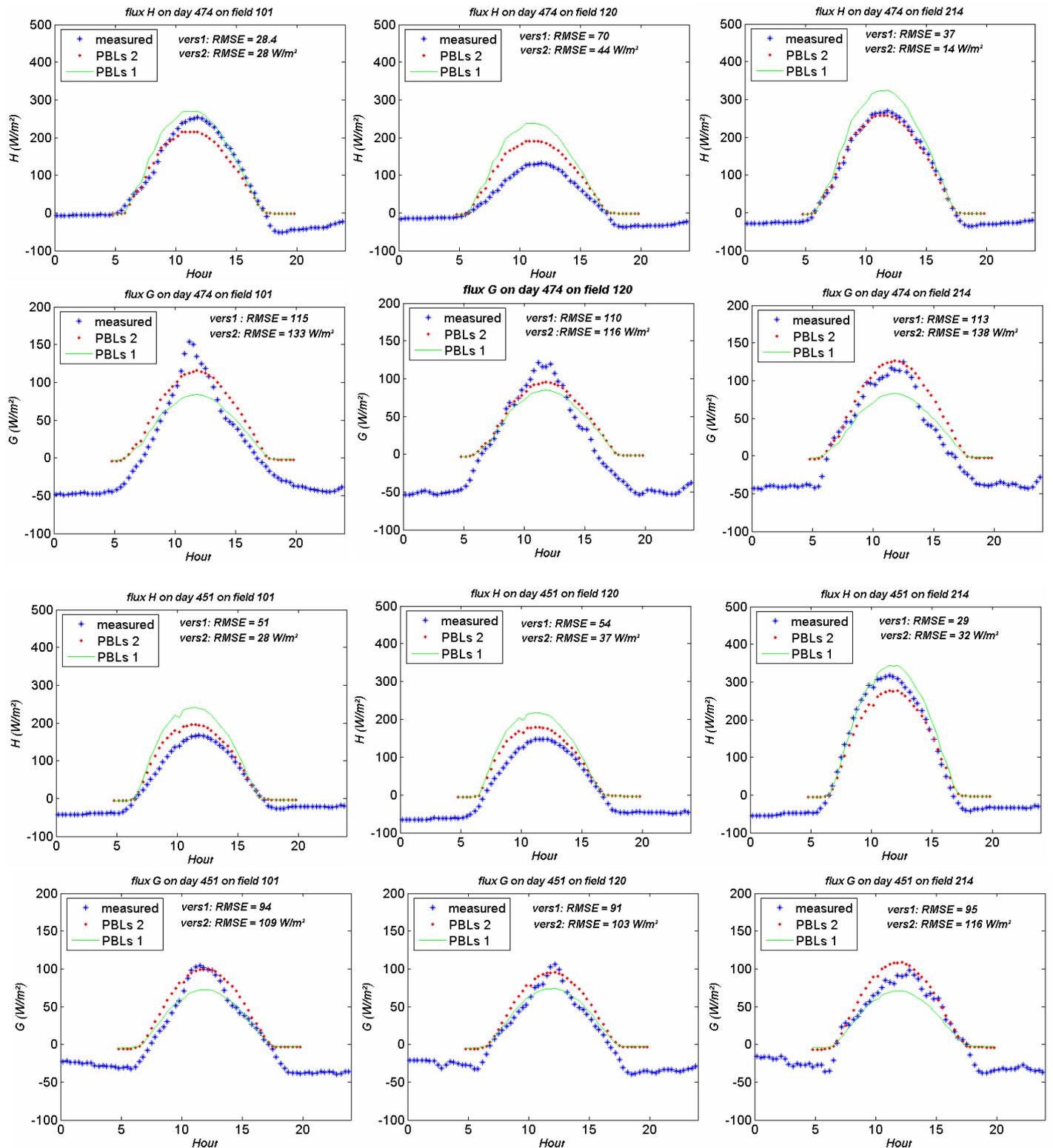
DOE	$R_n (Wm^{-2})$		$H (Wm^{-2})$		$LE (Wm^{-2})$		$G (Wm^{-2})$		$T_s (K)$		$T_a (K)$	
	<i>PBLs1</i>	<i>PBLs2</i>	<i>PBLs1</i>	<i>PBLs2</i>	<i>PBLs1</i>	<i>PBLs2</i>	<i>PBLs1</i>	<i>PBLs2</i>	<i>PBLs1</i>	<i>PBLs2</i>	<i>PBLs1</i>	<i>PBLs2</i>
101	33	31	28	28	26	23	115	133	4,2	5,0	3,7	3,9
120	45	48	70	44	70	62	110	116	5,0	6,3	3,3	3,8
203	30	32	24	29	31	24	86	83	4,9	6,1	3,2	3,7
214	38	35	37	14	22	18	113	138	3,8	4,4	3,8	3,8

DOE	$R_n (Wm^{-2})$		$H (Wm^{-2})$		$LE (Wm^{-2})$		$G (Wm^{-2})$		$T_s (K)$		$T_a (K)$	
	<i>PBLs1</i>	<i>PBLs2</i>	<i>PBLs1</i>	<i>PBLs2</i>	<i>PBLs1</i>	<i>PBLs2</i>	<i>PBLs1</i>	<i>PBLs2</i>	<i>PBLs1</i>	<i>PBLs2</i>	<i>PBLs1</i>	<i>PBLs2</i>
101	23	24	51	28	41	38	94	109	4,6	5,4	4,6	4,8
120	40	40	54	37	53	51	91	103	4,6	5,4	4,5	4,8
203	24	29	37	31	52	53	67	63	4,7	6,0	4,4	4,8
214	43	42	29	32	12	15	95	116	3,8	4,9	4,7	5,0

De façon générale, les résultats sur le rayonnement net et le flux de chaleur latente sont très proches entre les deux versions du modèle. Les différences notoires se situent sur les estimations des variables les plus reliées aux deux paramètres modifiés : les flux de chaleur sensible  $H$ , de chaleur dans le sol  $G$  et les températures. Pour le flux  $H$  les simulations avec la nouvelle version se rapprochent sensiblement des mesures, en particulier le jour 451. Par contre les erreurs quadratiques ( $RMSE$ ) calculées sur les évolutions journalières du flux  $G$  et des températures montrent en général un éloignement des mesures avec la nouvelle version. Les flux  $G$  et  $H$  sont présentés sur la Figure III.26 montrant une forte augmentation de  $G$  et une forte baisse de  $H$ . Les évaluations numériques des modèles sont à prendre avec prudence ici. En effet, les conditions de simulations, et en particulier les humidités du sol, ont été fixées à des valeurs a priori pas forcément adaptée à la situation réelle des parcelles (l'assimilation aura pour objectif de retrouver des valeurs plus adaptées : cf. chapitre V). On remarque aussi pour le flux  $G$  que la dégradation des  $RMSE$  correspond à des surestimations par rapport aux mesures, excepté pour les simulations autour du midi solaire qui sont en meilleurs agréments (sur l'ensemble des données utilisées ici PBLs1 présente à midi solaire un  $RMSE$  de  $35 Wm^{-2}$  et PBLs2 un  $RMSE$  de  $18 Wm^{-2}$  pour la date 474. Notons que le flux  $G$ , étant peu sensible à

l'humidité du sol, sera peu modifié par l'assimilation de la température de surface dans PBLs, et que les résultats présentés ici resteront valables lors des tests d'assimilation, facilitant l'amélioration des estimations au midi solaire (cinquième partie de ce mémoire).

Une problématique posée par la deuxième version du modèle PBLs concerne l'augmentation de la température de surface (cf. **Annexe III**). Celle-ci s'explique par l'introduction d'une rugosité thermique ( $z0_h$ ) plus faible par le modèle de Massman que dans la version initiale de PBLs. Kalma et Jupp (1990) ou Braud et al. (1993) par exemple ont montré que la baisse de rugosité thermique entraînait des modifications de la simulation de la température de surface jusqu'à quelques degrés. L'examen des simulations des flux (Figure II.26 et Annexe III) montre des flux  $H$  trop forts, des flux  $LE$  trop faibles et des températures de surface trop fortes au midi solaire en cohérence avec la possibilité de corriger les estimations par assimilation de la température. En effet, ce sont ces sous-estimations ou surestimations que l'on cherche à corriger via le processus d'assimilation variationnelle de données.



**Fig. III.26. Comparaison des simulations des flux  $H$  et  $G$  par les deux versions de PBLs (vers 1 et vers 2) aux flux mesurés sur les parcelles 101, 120, et 214 les DOE 451 et 474.**



***QUATRIEME PARTIE :***

***IMPLEMENTATION ET VERIFICATION DU SYSTEME  
D'ASSIMILATION DE DONNEES***



## 1. Introduction

Trois aspects importants de la mise en oeuvre d'un système d'assimilation seront abordés dans cette partie : l'implémentation, la vérification et le coût en temps de calcul.

L'implémentation d'un système d'assimilation variationnel pose le problème du développement d'un modèle adjoint. La solution que nous avons choisie ici est d'utiliser le différentiateur automatique TAPENADE. Si cette solution semble être adaptée à notre cas, il n'en reste pas moins que l'adjointisation automatique d'un code de calcul est une tâche difficile. Ce type de logiciel fait encore l'objet de recherches actives et leur application est parfois délicate : elle peut nécessiter des modifications du modèle considéré, l'adjoint généré peut donner des résultats incorrects pour diverses raisons et se pose alors le problème de l'identification des sources d'erreurs, le coût en temps de calcul de l'adjoint généré peut être important, etc ... Nous présenterons dans ce chapitre le détail des opérations qui ont été nécessaires afin d'obtenir un code adjoint valide et performant pour le modèle PBLs. Nous détaillerons également les résultats des tests de validation de cet adjoint et ses performances en termes de temps de calcul et de mémoire utilisée.

L'utilisation du logiciel PALM pour l'assemblage du système d'assimilation doit permettre de le rendre flexible grâce à une implémentation modulaire. Cela est particulièrement d'intérêt pour l'évolution future de ce travail et notamment dans le cadre de l'étude de la modélisation du paysage dans le laboratoire. Nous présenterons ici comment le système d'assimilation a été implémenté avec ce logiciel et quelles modifications ont dûes être réalisées dans les codes existants. Le nombre de sources d'erreurs possibles dans l'implémentation d'un système d'assimilation variationnel est très important. Ces erreurs peuvent se situer dans ses composants (en particulier le modèle adjoint, mais aussi le modèle direct qui est utilisé de façon différente dans un algorithme d'assimilation par rapport à une utilisation classique) ou leur assemblage. Il est ainsi essentiel de mettre en oeuvre une série de vérifications du comportement du système développé, dans des conditions d'information maîtrisées, afin de s'assurer qu'il fonctionne correctement. Les expériences synthétiques sont particulièrement précieuses dans ce cadre. La fin de cette partie sera consacrée à la description des expériences de vérifications qui ont été effectuées et de leurs résultats. Enfin, des éléments concernant le coût en temps de calcul du système d'assimilation seront donnés en conclusion.

## 2. Génération, validation et performances de l'adjoint

Comme décrit dans le deuxième chapitre, le système d'assimilation variationnel repose sur l'utilisation d'un algorithme de minimisation (méthode quasi-newton) nécessitant le calcul du gradient de la fonction coût du problème (II.24). Nous avons noté que la méthode adjointe est la plus adaptée pour calculer ce gradient de façon efficace pour des problèmes de grande taille et avons choisi d'utiliser le logiciel de différentiation automatique TAPENADE pour générer l'adjoint de PBLs. Nous présentons dans cette section l'application du logiciel TAPENADE au modèle PBLs, la validation de l'adjoint obtenu ainsi que ses performances.

### 2.1. Génération de l'adjoint du modèle PBLs

Dans notre application, la liste des variables dépendantes et indépendantes par rapport auxquelles différencier PBLs sont respectivement le tableau qui contient les valeurs des températures de surface sur l'ensemble des pixels et les tableaux qui contiennent les valeurs des paramètres  $f_2$  et  $z_0$  sur l'ensemble des pixels. Il faut noter ici que le même code adjoint est utilisé dans les trois algorithmes A, B et C (qui diffèrent de par la liste des paramètres estimés : uniquement  $z_0$  pour A, uniquement  $f_2$  pour B, et  $z_0$  et  $f_2$  pour C). Lorsqu'une seule des deux variables indépendantes est estimée dans le processus d'assimilation, le gradient de la fonction coût  $J$  ne doit être calculé que par rapport à cette variable. Il suffit alors de ne sélectionner que les valeurs correspondant à cette variable en sortie du modèle adjoint.

Différentes modifications du code du modèle PBLs ont dû être réalisées afin d'obtenir son code adjoint. Tout d'abord, afin d'isoler la partie à différencier du modèle, nous avons restructuré le programme principal en trois parties (initialisation, calcul, écriture des résultats) selon les principes détaillés dans le **chapitre II**. Cette opération s'est avérée relativement légère vu la taille du code et sa relative bonne structuration originale.

L'application de TAPENADE sur la subroutine contenant le code à différencier n'a pas permis de générer l'adjoint sans modifications supplémentaires. L'analyse des messages d'erreur a révélé certaines limites de TAPENADE mais aussi des faiblesses dans l'implémentation de PBLs. Ainsi, les passages de fonctions par argument ont dû être enlevés et certaines parties du

code contenant des GOTO ont dû être réécrites car TAPENADE ne réussissait pas à traiter ces instructions dans certains cas. Par ailleurs, l'utilisation de variables de type entier comme variables intermédiaires dans le calcul de variables réelles, qui n'est pas à conseiller lors de l'implémentation d'un programme, pose problème à TAPENADE. En effet, cela coupe la chaîne de dépendance des variables à dériver (on ne peut dériver par rapport à une variable entière), et cela a donc dû être corrigé à différents endroits du code de PBLs.

Enfin, certains messages d'avertissement délivrés par TAPENADE ont également conduit à des modifications. Ces problèmes n'étaient pas bloquants pour la génération de l'adjoint mais auraient induits des erreurs dans ses résultats. Ces messages (AD06 et AD07) suggèrent la déclaration de variables (souvent via l'inclusion de COMMON) dans certaines fonctions du code direct. Ces variables ne sont pas utilisées dans ces fonctions, c'est pourquoi elles n'ont pas été déclarées originellement dans le code direct. Toutefois, dans le code adjoint généré, des appels aux diverses fonctions du code direct sont réalisés afin d'effectuer une simulation du modèle direct et de stocker l'état de variables intermédiaires à différentes étapes (cf. mécanisme de checkpointing détaillé dans le **chapitre II section 6**).

Or ces stockages de variables sont parfois réalisés dans des fonctions du code direct dans lesquelles ces variables ne sont pas déclarées. Typiquement, une variable doit parfois être stockée avant l'appel à une fonction qui la modifie. Si cette variable est déclarée dans un COMMON, (et n'est donc pas passée en argument de cette fonction, que nous noterons **fonction1**.) il est tout à fait possible que dans le code direct, cette variable ne soit pas déclarée dans la fonction **fonction2** qui appelle **fonction1**. Si le code du modèle considéré (comme c'était le cas de PBLs) a été implémenté avec le système des déclarations implicites, système qui est malheureusement celui par défaut en FORTRAN, alors ce problème passera inaperçu lors de la compilation du code adjoint. Dans ce cas, les résultats du code adjoint seront incorrects et la source de ces erreurs sera particulièrement difficile à détecter. Ces messages d'avertissement donnés par TAPENADE ont donc été particulièrement précieux pour éviter ce type de problème avec PBLs.

Une fois le code adjoint généré, les premiers tests de validation ont montré un écart important entre les résultats obtenus avec le code adjoint et ceux obtenus par différences finies. Après diverses vérifications (cf. chapitre II pour les vérifications d'usages), nous avons testé le code tangent qui s'est révélé lui aussi incorrect. Un code tangent étant beaucoup plus facile à déboguer (corriger l'erreur) qu'un code adjoint, nous avons recherché la source de l'erreur du

code tangent grâce aux outils de débogage fournis avec TAPENADE. La manipulation de ces outils est fastidieuse mais relativement simple, et nous avons pu détecter le problème qui se situait en fait dans le code de PBLs. Il s'agissait d'un bug (défaut) de PBLs qui ne s'exprimait pas lors de la simulation du code direct, mais que la transformation de programme effectuée par TAPENADE dans les codes tangent et adjoint activait. La correction de ce bug (défaut) dans le code de PBLs (problème d'initialisation de variable dans la routine nb02a), nous a permis d'obtenir un code adjoint ayant des résultats corrects comme nous le verrons dans la section suivante.

Enfin, un dernier type de modification a été effectué dans le modèle PBLs afin d'optimiser les performances du modèle adjoint en terme de temps de calcul. Ce point est détaillé dans la **section 2.3** de ce chapitre.

## 2.2. Validation de l'adjoint du modèle PBLs

Le modèle adjoint est jugé correct si les résultats du test de Taylor et du test du produit scalaire de l'adjoint (cf. **section II.6**) sont satisfaisants.

Nous avons validé le modèle adjoint de PBLs pour différents pixels (tirés au hasard dans la carte Alpilles - ReSeDA), et avons testé différents pas ( $\epsilon$ ) pour le calcul des dérivées par différences finies. Nous avons réalisé une simulation entre 5h et 12h et utilisé les sorties des températures de surface à 12h pour procéder aux tests de validation de l'adjoint. Le tableau IV.1 ci-dessous résume les différents résultats obtenus.

**Tab. IV.1. Norme de la dérivée de la température de surface  $T_{si}$  :**

*Colonne 1* : nombre de pixels utilisé dans la validation ; *Colonne 2* : pas ( $\epsilon$ ) pour le calcul des dérivées par différences finies. *Colonnes 3 à 4* : valeurs de la norme : *DD* norme par différences finies, *NormTgt* celle du code tangent et *NormAdj* celle de l'adjoint

	$\epsilon$	DD	NormTgt	NormAdj
$T_s$ à 12h	$10^{-6}$	999746,2065465		
1 pixel	$10^{-7}$	999746,1689571	999746,1439858	999746,1439861
	$10^{-8}$	999746,2867177		
$T_s$ à 12h	$10^{-6}$	1297947,247577		
2 pixels	$10^{-7}$	1301922,079753	1301922,061621	1301922,061626
	$10^{-8}$	1301921,763272		
$T_s$ à 12h				
100 pixels	$10^{-7}$	14016718,849485	14016718,437711	14016718,437732

Les différents résultats du tableau IV.1 montrent bien l'égalité entre les différentes normes permettant de considérer comme validé le modèle adjoint de PBLs. En effet, on note une égalité avec 7 à 8 chiffres significatifs comme le montre les résultats avec 1 pixel et 2 pixels avec un pas de calcul à  $10^{-7}$  pour le calcul par différences finies. Le nombre de chiffres significatifs obtenu est dans la gamme de ceux classiquement rencontrés lors de ces processus de validation de l'adjoint (Hascoët & Pascual, 2004). Par ailleurs, on peut constater que le pas

de calcul des différences finies influe sur la précision du calcul des dérivées (notamment pour le cas à 2 pixels). Pour un pas de calcul trop grand ( $10^{-4}$ ) ou surtout très petit ( $10^{-9}$ ), on obtient une faible précision du résultat avec la formule des différences finies (résultats non montrés ici). Ces types d'imprécision ne remettent pas en cause la validité de l'adjoint mais plutôt traduisent l'imprécision des différences finies : pour un pas trop grand, l'hypothèse de linéarité sous-jacente à l'utilisation de la formule des différences finies n'est plus valide, et pour un pas trop petit, les erreurs d'arrondi dominant le calcul.

La différence entre les produits scalaires issus du modèle linéaire tangent et du modèle adjoint se limite au dernier ou à l'avant dernier digit. C'est également un résultat attendu (Hascoët & Pascual, 2004).

Les résultats de ces tests sont donc satisfaisants. Toutefois un autre aspect à analyser est le coût de l'adjoint en termes de mémoire et de temps d'exécution.

### **2.3. Performances et optimisation de l'adjoint**

Comme rappelé dans la **section 6** du **chapitre II**, le temps d'exécution minimal d'un modèle adjoint est généralement de l'ordre de quelques fois celui du modèle direct considéré. Les premiers tests de performance réalisés avec PBLs ont montré un ratio temps de calcul du modèle adjoint sur temps de calcul du modèle direct de l'ordre de 30 (pour 1000 pixels), ce qui est assez élevé. L'application d'un outil d'analyse de performance du code adjoint a permis de localiser les lignes de codes particulièrement coûteuses. Il s'agissait du stockage de tableaux (par le mécanisme de checkpointing utilisé en programmation) dans le calcul des flux de surface. Dans PBLs, le calcul des flux de surface est effectué de façon indépendante entre pixels au sein d'une itération de couplage avec le modèle de couche limite atmosphérique. Or les routines responsables de ce calcul prenaient en argument l'adresse des différents tableaux contenant les données pour l'ensemble des pixels. Cela ne pose aucun problème dans le cas du modèle direct. Cependant, TAPENADE ne peut détecter qu'à chaque itération les routines n'utilisaient qu'un seul élément du tableau, et stocke donc l'ensemble de ces tableaux à chaque étape de checkpointing au lieu d'un seul élément. En conséquence, le coût de l'adjoint devient vite important avec l'augmentation du nombre de pixels traités. Le modèle PBLs a donc été modifié afin qu'au sein de la boucle sur les pixels, seules les valeurs des tableaux

correspondant au pixel traité soient passées en argument des différentes routines impliquées dans le calcul des flux de surface. Le tableau IV.2 ci-dessous montre les performances du modèle adjoint, après cette modification, pour différents nombres de pixels traités. Nous pouvons constater que cette simple modification a permis de réduire drastiquement le coût du modèle adjoint et de rendre son ratio avec le temps de calcul du modèle direct quasiment indépendant du nombre de pixels traités.

**Tab. IV.2. Temps d'exécution du modèle adjoint de PBLs entre 5h – 12h et son stockage mémoire**

Nombre de pixels	Temps CPU modèle direct (sec.)	Temps(adjoint)/Temps(direct)	mémoire adj (MB)
1	1.000E-02	8	0.01
10	9.999E-02	6.70001	0.031250
100	1.07000	6.81309	0.343750
1000	10.37000	7.33655	3.437500
10000	106.460	7.27682	34.468750

Au final, le rapport entre le temps d'exécution du modèle adjoint et celui du modèle direct se situe autour de 7, ce qui est tout à fait acceptable (cf. chapitre II section 6 pour des exemples sur d'autres modèles). Le stockage mémoire dû au mécanisme de checkpointing est linéaire au nombre de pixels traité et reste également raisonnable.

Ces indications nous sont fort utiles car elles nous permettent d'établir des prévisions du coût de la méthode en fonction de la taille du problème considéré. A partir de la valeur du temps d'exécution pour 100 pixels, nous pouvons évaluer le coût nécessaire de l'application de la méthode sur toute la zone d'Alpilles subdivisée en 250 par 250 pixels (soit 62500 pixels). Il nous faudrait 11 minutes CPU pour une simulation du modèle PBLs et 1h18 mn pour une simulation de son adjoint avec une mémoire de stockage autour de 223 MB.

### 3. Assemblage et vérification du système d'assimilation

#### 3.1. Assemblage du système d'assimilation avec le logiciel PALM

##### 3.1.1. Unités

La première étape de la construction d'une application avec le logiciel PALM est le développement des unités. Un choix naturel pour la granularité du couplage dans le cas d'un système d'assimilation variationnel est de faire correspondre les unités aux principaux opérateurs mathématiques qui composent l'algorithme (Lagarde et al., 2001). Ainsi, pour implémenter le système d'assimilation dans notre cas, les principales unités seront :

- le modèle PBLs,
- l'adjoint du modèle PBLs,
- les opérateurs de multiplication d'un vecteur par l'inverse des matrices de covariance d'erreur  $R_i$  et  $B$ ,
- le minimiseur,
- des unités permettant de réaliser les différentes opérations algébriques nécessaires pour le calcul de la fonction et de son gradient (produit scalaire et somme de vecteurs),
- des unités permettant la lecture et l'envoi de différentes variables (information a priori, observations, initialisation des variables du modèle, ...) et le traitement des résultats.

Le modèle PBLs est découpé en trois unités correspondant à la réorganisation effectuée pour son adjointisation (cf. section précédente) : une unité qui initialise ses variables d'entrée, une unité qui réalise les calculs et une unité qui gère l'écriture des résultats dans les fichiers de sortie. Ce découpage permet une manipulation plus souple du modèle. Le modèle adjoint, qui rappelons-le, n'est l'opérateur adjoint que de la partie calcul du modèle PBLs, est donc représenté par une seule unité. L'unité d'initialisation des variables d'entrées du modèle est utilisée à la fois pour le modèle direct et pour le modèle adjoint. La transformation des codes existant en unités PALM est une opération relativement légère : il s'agit d'y insérer les appels aux primitives PALM\_Put et PALM\_Get lorsque des objets sont produits ou consommés. Dans notre cas, les objets manipulés ne concernent que les vecteurs contenant les valeurs de  $z_0$ ,  $f_2$  et

$T_s$  à 12h sur l'ensemble des pixels. Les autres variables échangées entre les unités concernant le modèle et son adjoint sont passées par COMMON, comme dans le code original, car elles ne concernent pas les autres unités de l'algorithme. Les modifications apportées à ces codes sont donc mineures. Les figures IV.1 et IV.2 présentent respectivement la routine principale du modèle PBLs instrumenté, avec sa carte d'identité, et une illustration d'une application du code direct de PBLs dans PrePALM.

```

SUBROUTINE UNIT_PBL ()
  IMPLICIT NONE
  interface PALM
    include 'palmLib.h'
    include 'palm_user_param.h'

    CHARACTER(LEN=PL_LNAME) :: cl_object, cl_space
    integer il_err
    integer nb_zones, ne, maxi

    real y(ne), z0(nb_pixels), f2(nb_pixels), Ts(nb_pixels),
+     z0h(nb_pixels)

    common/parametre_calcul/y, nb_zones, z0h

  end interface

  c Appel de Palm_Get pour recevoir z0
  c cl_space = 'vect1d'
  c cl_object = 'z0'
  CALL PALM_Get(cl_space, cl_object, PL_NO_TIME,
+   PL_NO_TAG, z0, il_err)

  c Appel de Palm_Get pour recevoir f2
  c cl_space = 'vect1d'
  c cl_object = 'f2'
  CALL PALM_Get(cl_space, cl_object, PL_NO_TIME,
+   PL_NO_TAG, f2, il_err)

  c Appel de la routine calcul
  call CALCUL(nb_zones, z0, z0h, f2, y, Ts)

  c Envoi du Ts a 12h:
  c cl_space = 'vect1d'
  c cl_object = 'Ts'
  CALL PALM_Put(cl_space, cl_object, PL_NO_TIME,
+   PL_NO_TAG, Ts, il_err)

  END

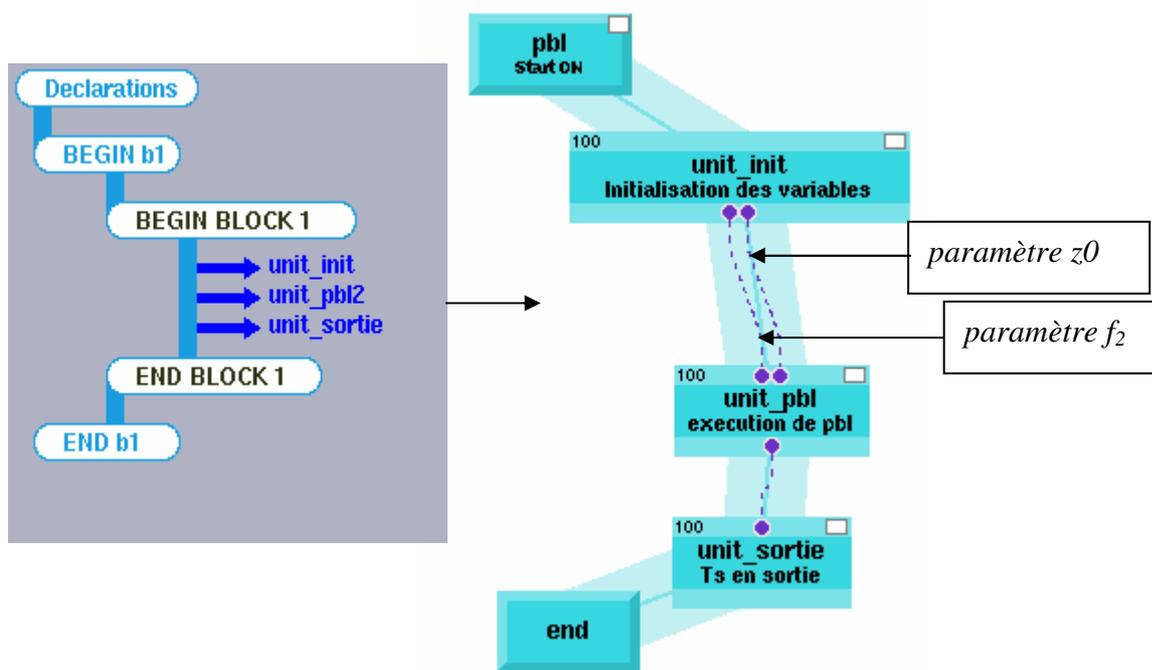
```

```

c$PALM_UNIT -name unit_pbl\
c -functions {F77 unit_pbl}\
c -object_files { pbl.a}\
c -comment {Execution de PBLs}
c
c$PALM_SPACE -name vect1d\
c -shape (nb_pixels)\
c -element_size PL_DOUBLE_PRECISION\
c -comment {Vecteur de taille le nombre de pixels}
c
c$PALM_OBJECT -name z0\
c -space vect1d\
c -intent IN\
c -comment {parametre z0}
c
c$PALM_OBJECT -name f2\
c -space vect1d\
c -intent IN\
c -comment {parametre f2}
c
c$PALM_OBJECT -name Ts\
c -space vect1d\
c -intent OUT\
c -comment {Ts a 12h}

```

**Fig. IV.1. Carte d'identité et routine principale de l'unité de calcul de PBLs. Il faut noter que seule la routine principale de PBLs a été modifiée pour son interfaçage dans PALM.**

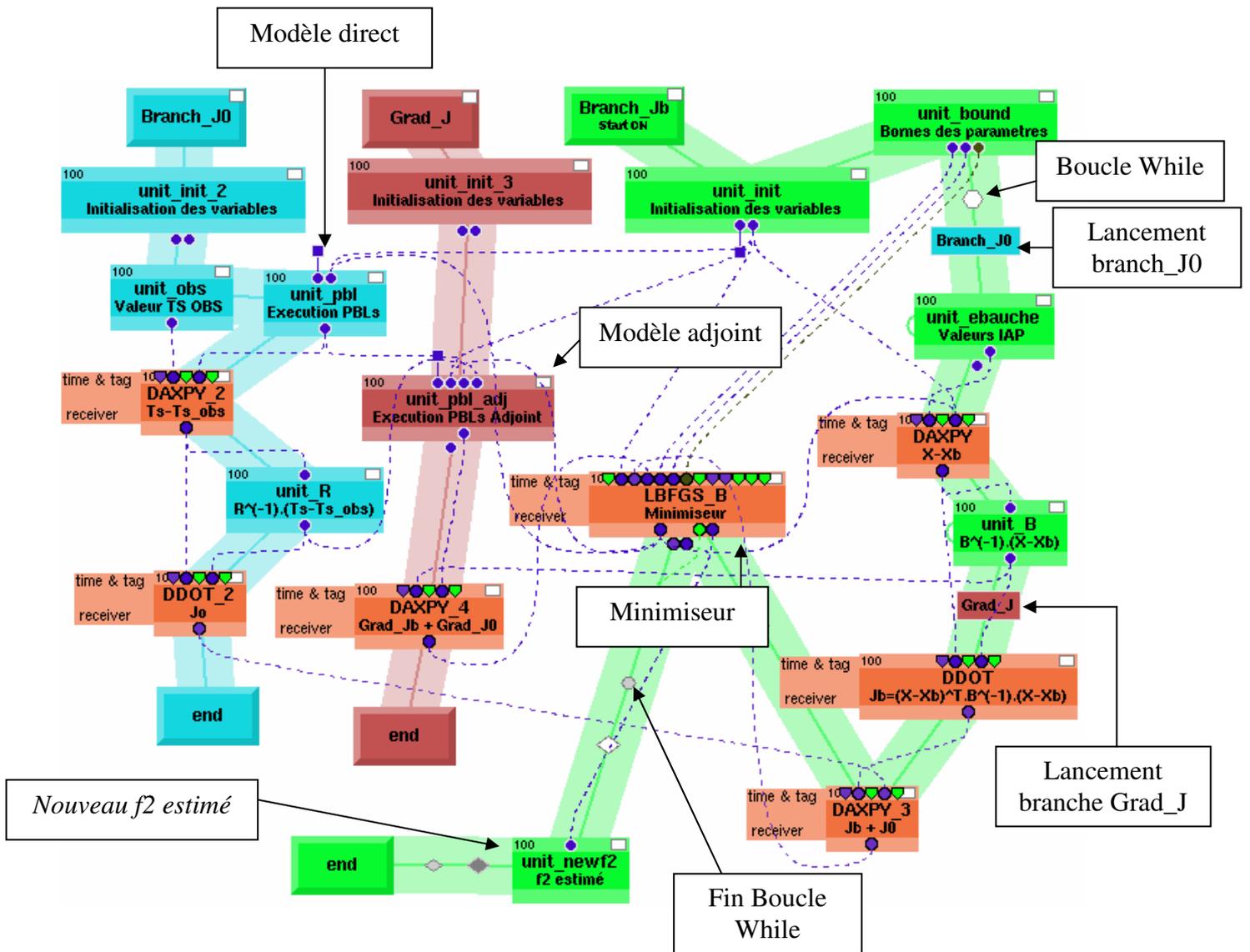


**Fig. IV.2. Illustration d’une application PALM qui réalise une exécution du modèle PBLs : à gauche le code de branche, et à droite la représentation de la branche dans le canvas de l’interface graphique PrePALM.**

Les unités concernant les opérations algébriques de base et le minimiseur proviennent de la boîte à outil d’algèbre de PALM : nous n’avons donc pas à les développer ou à les modifier. L’unité du minimiseur est une interface PALM au code L-BFGS-B disponible sous licence GPL à l’adresse <http://www.ece.northwestern.edu/~nocedal/lbfgs.html> et qui implémente une version avec contraintes de bornes de l’algorithme de Quasi-Newton à mémoire limitée, tel que décrit dans (Liu & Nocedal, 1989).

### 3.1.2. Construction schéma de couplage

Nous avons choisi d’implémenter le schéma de couplage avec trois branches : une branche appelée *Branch\_J0* pour le modèle direct et le calcul de la partie « comparaison aux observations » ( $J_0$ ) de la fonction coût, une branche appelée *Grad\_J* pour le modèle adjoint et le calcul de  $\nabla J_o$ , et une branche appelée *Branch\_Jb* qui pilote la minimisation et réalise les calculs des termes d’ébauche de la fonction coût et de son gradient. La figure IV.3 est une illustration de la représentation de l’application dans PrePALM.



**Fig. IV.3. Schéma final de l’algorithme d’assimilation B dans PrePALM (sans préconditionnement). A chaque branche est associée une couleur. Les unités provenant de la boîte à outils d’algèbre de PALM sont colorées en orange.**

L’algorithme est organisé au sein d’une boucle « while » : tant que le minimiseur ne donne pas le signal de la fin de la minimisation (lorsqu’un des critères d’arrêt est rempli), la fonction coût et son gradient doivent être calculés pour la valeur des paramètres qu’il fournit, et les résultats doivent lui être envoyés. Ainsi, les branches et unités qui permettent le calcul de la fonction coût et de son gradient ainsi que l’unité du minimiseur sont lancés au sein de cette boucle. Le

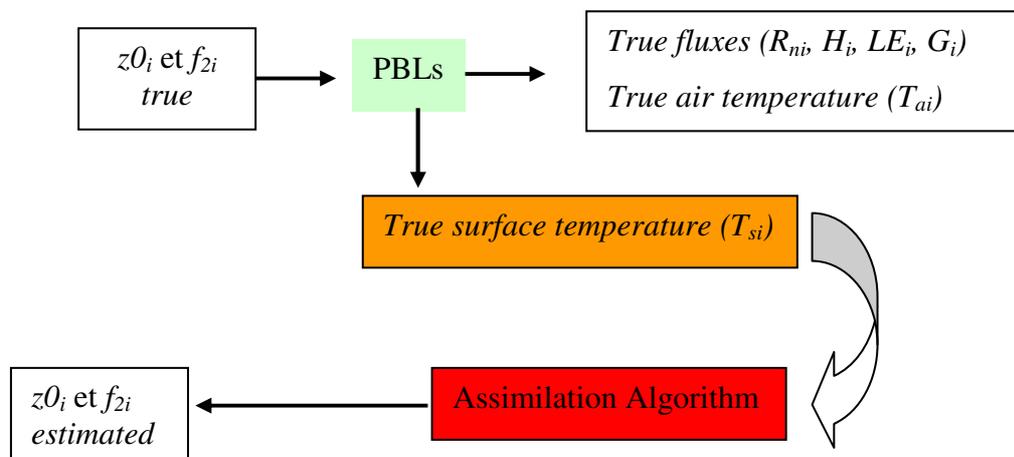
processus est initialisé par l'unité `unit_init` qui envoie les valeurs initiales des paramètres. Lorsque la minimisation se termine, les valeurs estimées des paramètres sont récupérées (dans cet exemple par l'unité `unit_newf2`) et écrites dans un fichier. Nous ne décrivons pas plus en détail l'implémentation de cet algorithme. Cependant, au-delà des aspects de flexibilité de l'application, nous souhaitons insister ici sur un autre intérêt de l'utilisation de PALM : la représentation de l'application dans PrePALM. En effet, pour un utilisateur ayant une connaissance minimale des symboles utilisés dans PrePALM, la représentation de l'application en donne une vue à la fois globale et détaillée : on peut voir sur la figure IV.3 les différents composants de l'application, l'algorithme (les branches et les structures de contrôle, qui sont représentées graphiquement par différents symboles), les flux de données entre les composants, les opérations d'algèbre, le parallélisme de l'application, ... D'autre part, l'utilisation de PrePALM donne en plus accès aux commentaires associés aux entités pointées par la souris (unités, objets, communications, ...).

### 3.2. Vérification

Dans cette section, nous détaillons les tests de vérification que nous avons effectués pour valider l'implémentation du système d'assimilation. Nous avons mené différentes analyses sur les données synthétiques décrites dans la **section 5** du **chapitre II**.

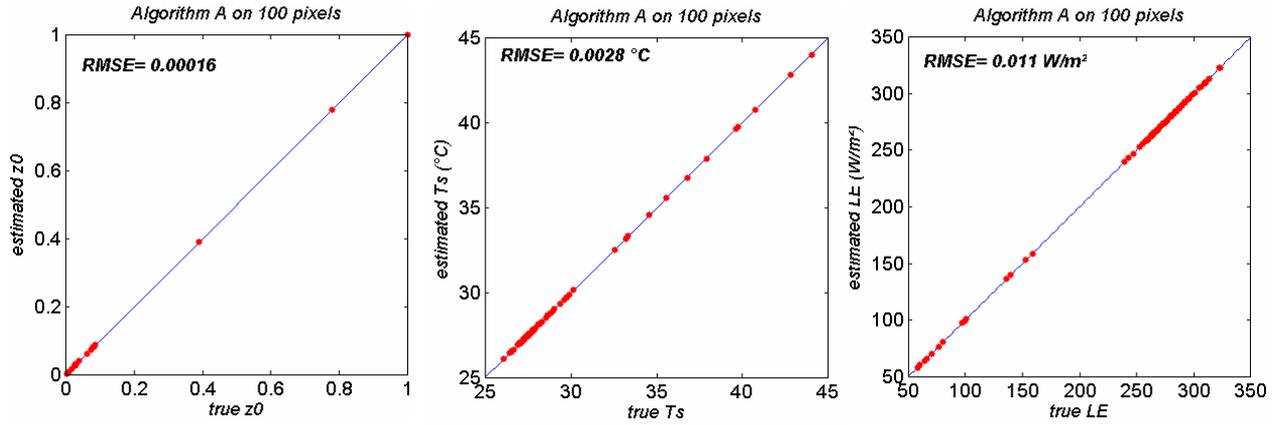
#### 3.2.1. Estimation des états vrais des paramètres $z0_i$ et $f2_i$

Une méthode pour valider l'implémentation de l'algorithme d'assimilation est de vérifier si le système d'assimilation retrouve les paramètres vrais définis dans le cadre d'une expérience synthétique, à partir d'observations exactes sans utiliser de terme d'ébauche. Dans notre cas, on assimile les températures de surface vraies ( $T_{si}$  *true*) dans l'algorithme et on cherche à retrouver les paramètres ( $z0_i$  et  $f2_i$ ) vrais et les flux vrais correspondants comme indiqués par la figure IV.4 ci-dessous.

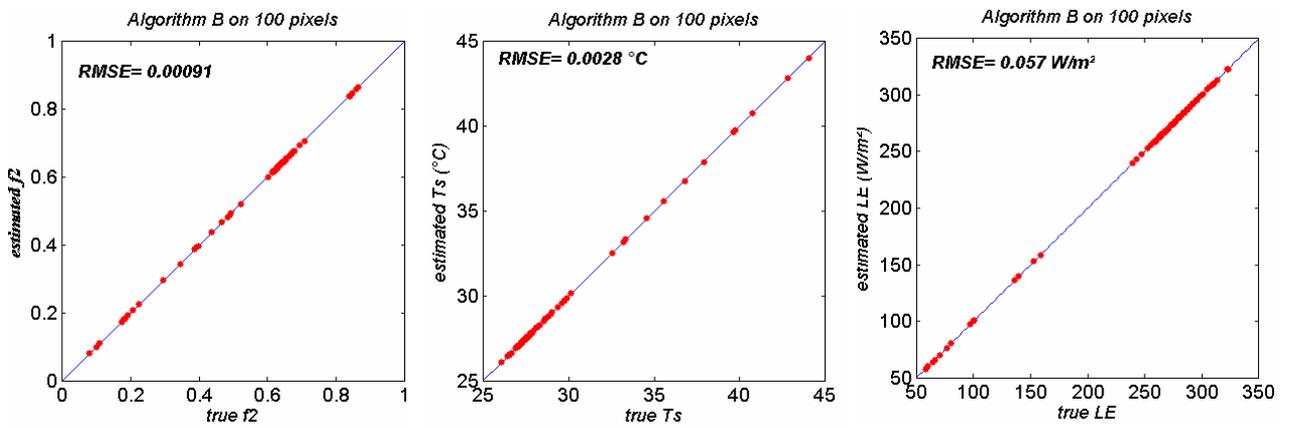


**Fig. IV.4. Schéma de validation de l'algorithme pour retrouver l'état vrai des variables définies**

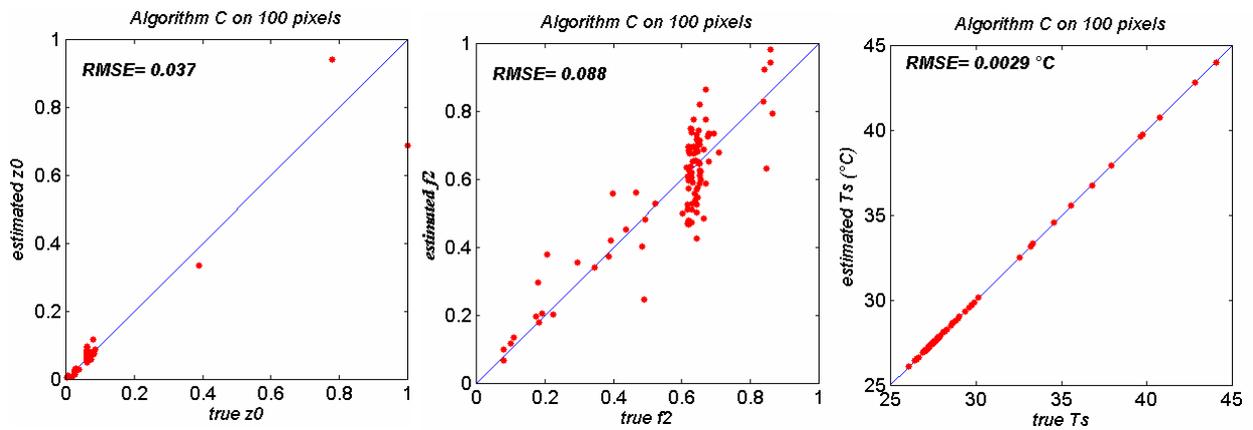
Nous avons réalisé ce type de validation sur 100 pixels tirés au hasard sur la carte Alpilles - ReSeDA et pour les différents algorithmes implémentés (A, B et C). Les conditions initiales sur les paramètres sont fixées aux valeurs a priori (cf. section 5.4 de la partie II) dont les erreurs sont de 0,23 pour  $z0$  et de 0,11 pour  $f2$ . Les résultats sont présentés ci-dessous (Fig. IV.5, IV.6 et IV.7).



**Fig. IV.5.** Les états vrais de la rugosité ( $z_{0i}$ ), de la température de surface ( $T_{si}$ ) et du flux ( $LE_i$ ) recherchés avec l'algorithme A



**Fig. IV.6.** Les états vrais de l'humidité normalisée ( $f_{2i}$ ), de la température de surface ( $T_{si}$ ) et du flux ( $LE_i$ ) recherchés avec l'algorithme B



**Fig. IV.7.** Les états vrais de la rugosité ( $z_{0i}$ ), de l'humidité normalisée ( $f_{2i}$ ) et de la température de surface ( $T_{si}$ ) recherchés avec l'algorithme C

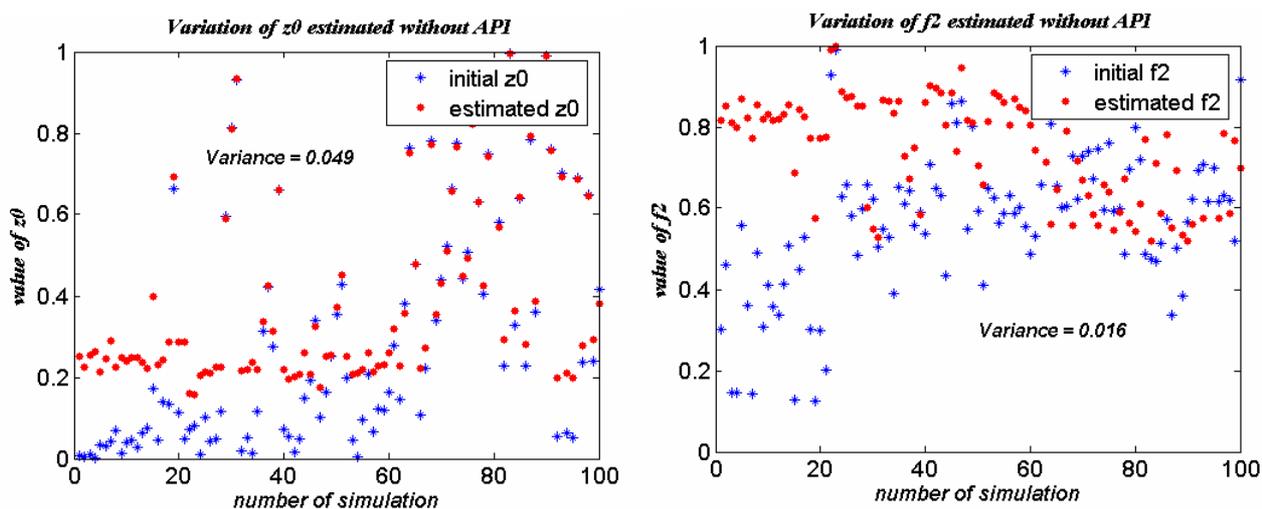
On peut noter qu'avec les algorithmes *A* (Fig. IV.5) et *B* (Fig. IV.6), les états vrais sont très bien retrouvés. Dans ces deux cas, les problèmes sont bien posés : il n'y a pas de problèmes de minima locaux ou d'effets de compensation entre les valeurs des paramètres sur les différents pixels. Sur les résultats obtenus à partir de l'algorithme *C* (Fig. IV.7), on note une dispersion de l'ordre de 15% à 30% sur l'estimation des paramètres ( $z_{0i}$  et  $f_{2i}$ ) alors même que les températures de surface sont estimées avec la même qualité que dans les cas *A* et *B*. Cela est dû à la sous-détermination du problème qui, du fait du nombre insuffisant d'information pour estimer tous les paramètres, devient un problème mal posé. Dans ce cas, l'estimation des paramètres ( $z_{0i}$  et  $f_{2i}$ ) dépend beaucoup des valeurs initiales de  $z_0$  et  $f_2$  choisies pour débiter le processus itératif de l'algorithme *C* ce qui n'est pas le cas des algorithmes *A* et *B*. C'est pourquoi l'ajout du terme d'information a priori dans le processus d'assimilation est important. Il permet de régulariser le problème mal posé. On peut quand même noter que malgré ce phénomène de compensation entre paramètres, les observations ont apporté de l'information pour estimer ces paramètres : les valeurs estimées pour  $z_{0i}$  et  $f_{2i}$  sont nettement plus précises que les valeurs initiales : les *RMSE* passent respectivement de 0,23 à 0,037 et de 0,11 à 0,088.

### 3.2.2. Indépendances des résultats par rapport aux valeurs initiales

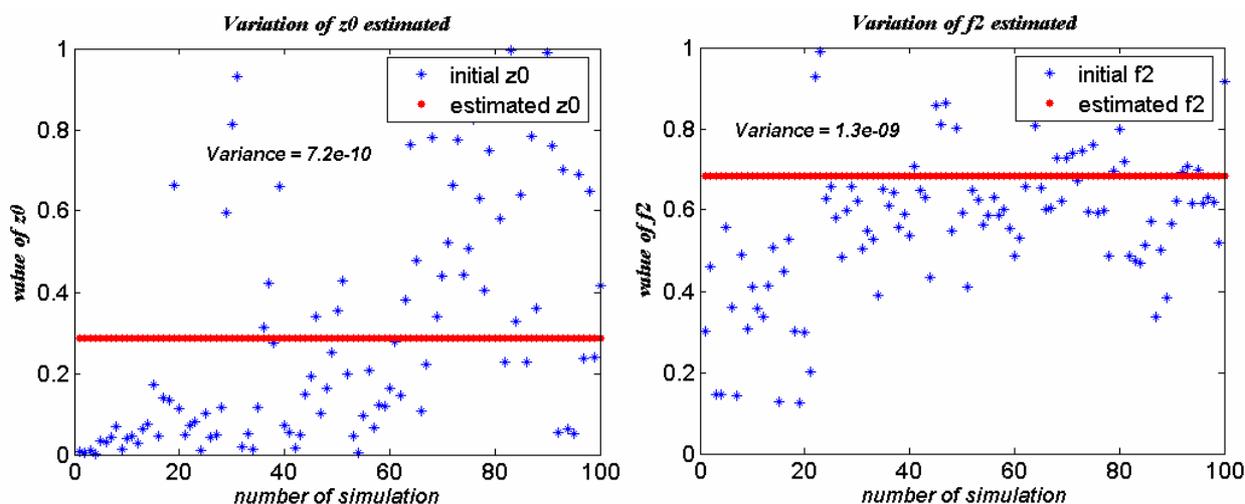
En pratique, les valeurs initiales des paramètres pour le processus de minimisation sont prises comme égale à l'information a priori (ce n'est pas obligatoire). Toutefois, il est important de distinguer clairement le vecteur d'information a priori qui est utilisé dans la définition de la fonction coût, du vecteur initial qui est utilisé pour initialiser la procédure de minimisation. Dans un processus d'assimilation satisfaisant, l'analyse ne dépend pas de façon significative du choix de l'initialisation des paramètres à contrôler (Bouttier & Courtier, 1999). En revanche, l'ajout du terme d'information a priori, considéré comme une information supplémentaire, doit permettre d'obtenir une analyse indépendante de l'initialisation des paramètres à contrôler en régularisant le problème d'optimisation. C'est cet effet de dépendance aux valeurs initiales que nous avons mis en évidence d'une part dans une configuration sans information a priori (Fig. IV.8) puis avec (Fig. IV.9). Dans ces tests, nous avons répété 100 fois le processus d'assimilation avec l'algorithme *C*, appliqué sur un seul pixel, avec différentes valeurs initiales des paramètres  $z_0$  et  $f_2$  tirées aléatoirement selon une loi uniforme définie par

les bornes des paramètres. Il faut noter que dans ce test, l'observation de température de surface utilisée, est différente de la valeur vraie.

Dans le cas où l'on n'utilise pas d'information a priori, on obtient des solutions assez dépendantes des états initiaux des paramètres de la rugosité aérodynamique et de l'humidité normalisée (Fig. IV.8).



**Fig. IV.8. Estimation de la rugosité aérodynamique  $z_0$  et de l'humidité normalisée  $f_2$  sans information a priori (sans IAP) suivant différentes valeurs initiales : valeurs initiales (en étoile) ; valeurs estimées (en point)**



**Fig. IV.9. Estimation de la rugosité aérodynamique  $z_0$  et de l'humidité normalisée  $f_2$  avec information a priori (avec IAP) suivant différentes valeurs initiales : valeurs initiales (en étoile) ; valeurs estimées (en point)**

Concernant le test avec information a priori, on peut constater la stabilité des résultats obtenus (Fig. IV.9). La variance entre les différentes estimations des paramètres est de l'ordre de  $10^{-10}$  pour la rugosité aérodynamique et  $10^{-9}$  pour l'humidité du sol normalisée. Les solutions obtenues ( $z_0 = 0,29$  ;  $f_2 = 0,68$ ) sont proches mais différentes des valeurs a priori (0,28 ; 0,65). Si les valeurs initiales des paramètres  $z_0$  et  $f_2$  n'ont pas d'influence sur l'estimation des nouveaux paramètres, cela est dû à l'ajout de l'information a priori qui régularise le problème mal posé par définition. Le terme d'ébauche de la fonction coût ( $J_B$ ) permet, en fonction du poids qui lui est donné via la matrice de covariance d'erreur de l'information à priori ( $B$ ), de rendre plus convexe la fonction coût et d'éviter ainsi la présence de minima locaux.

### 3.2.3. Poids des matrices de covariances d'erreur de l'ébauche et d'observation dans l'estimation

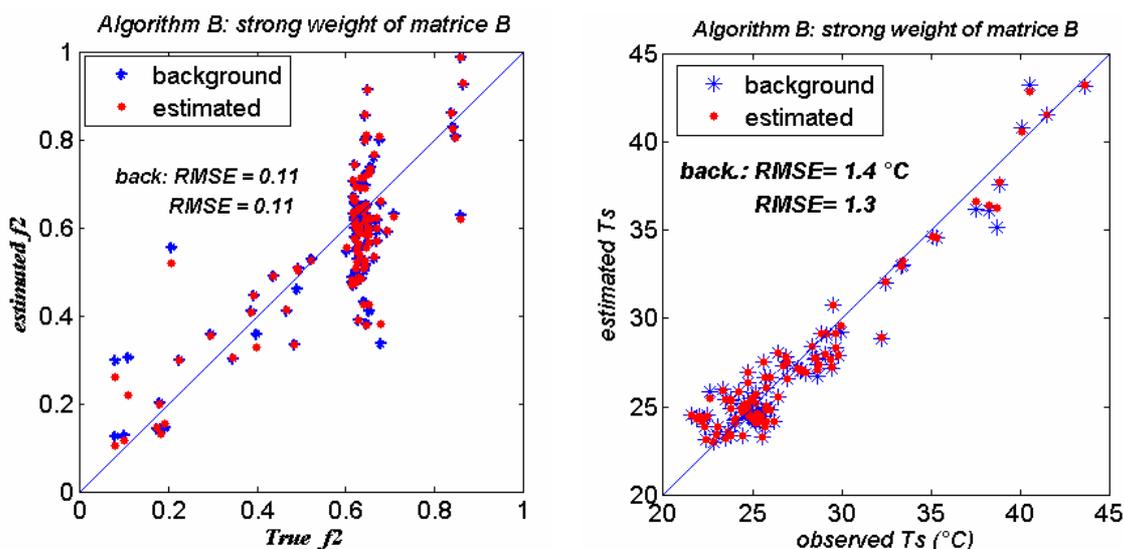
Les matrices de covariance d'erreur de l'information a priori ( $B$ ) et des observations ( $R_i$ ) jouent un rôle très important dans l'algorithme d'assimilation. En effet, l'algorithme réagit en fonction des valeurs fixées pour les composantes ( $\sigma_{bi}$  et  $\sigma_i$ ) des matrices. Pour une faible valeur de  $\sigma_i$  (composante de la matrice  $R_i$ ), l'observation sur le pixel  $i$  a un poids élevé dans le processus d'assimilation. Cela traduit physiquement que la valeur de cette observation est peu entachée d'erreur. Par conséquent, l'algorithme aura tendance à approcher plus cette observation que les autres. De même, si le poids de l'information a priori est élevé, les paramètres estimés auront tendance à être proches des valeurs de cette information a priori.

La qualité de l'analyse passe par une spécification adaptée des covariances d'erreur de l'observation (matrice  $R_i$ ) et de l'information a priori (matrice  $B$ ). Nous avons vérifié le comportement du système d'assimilation par rapport aux modifications de ces matrices sur une application de l'algorithme B sur 100 pixels de l'expérience synthétique. Dans cette expérience synthétique, les observations ont été générées avec une erreur d'écart type de 1,5 K et l'information a priori avec une erreur d'écart type de 0,11 en moyenne.

Nous avons réalisé deux applications du système d'assimilation : une avec une information a priori ayant plus de poids que défini dans l'expérience synthétique ( $\sigma_{fbi}$  fixé à la moitié de sa valeur réelle, soit 0,051 en moyenne, et  $\sigma_i$  fixé à 1,5 K comme dans l'expérience synthétique) et l'autre avec des observations ayant plus de poids que défini dans l'expérience synthétique ( $\sigma_i$  fixé à 0,5 K au lieu de 1,5 K, et  $\sigma_{fbi}$  fixé au double de sa valeur réelle, soit 0,22 en

moyenne). Les résultats sont illustrés sur les figures IV.10 et IV.11. On y compare l'humidité du sol normalisée  $f_2$  estimée par l'algorithme d'assimilation ainsi que les valeurs de l'information a priori, à l'état vrai sur l'ensemble des pixels. Par ailleurs, on trace également les températures de surface ( $T_{si}$ ) estimées par l'algorithme ainsi que celles simulées à partir de l'information a priori par rapport aux observations  $T_{sobs}$  (température de surface vraie à laquelle on a ajouté un bruit gaussien) afin de caractériser l'apport de l'observation dans le processus d'assimilation.

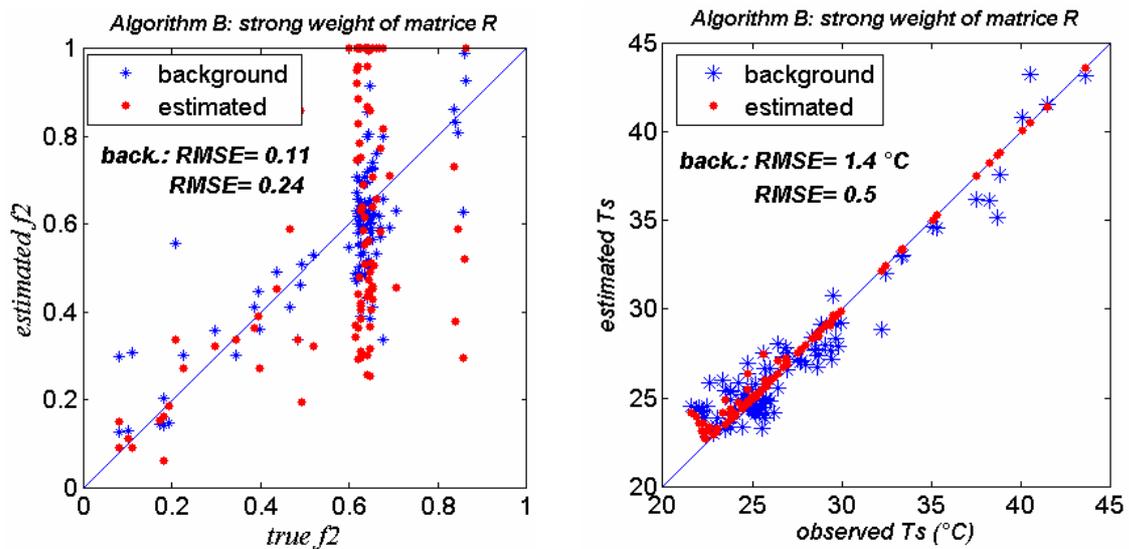
Comme attendu, on peut voir sur le premier graphique de la figure IV.10 que le vecteur de paramètre de l'humidité du sol normalisée ( $f_{2i}$ ) estimée reste proche des valeurs a priori. Par conséquent, on peut observer sur la figure IV.10b que le vecteur de température de surface estimé par l'algorithme d'assimilation est quasi similaire au vecteur de température déterminé avec le vecteur d'information a priori (background).



**Fig. IV.10. Poids important de la matrice de covariance  $B$  ( $f_2$ ,  $\sigma_{f_{bi}}$  fixé à 0,051) par rapport à la matrice de covariance  $R_i$  (observation,  $\sigma_i$  fixé à 1,5 K) dans le processus d'assimilation : l'information a priori (en étoile) et l'estimation de l'algorithme (en point)**

Au contraire, sur la figure IV.11, on observe un vecteur de température de surface estimé très proche des observations (avec un  $RMSE$  consistant avec l'erreur d'observation fixée dans le schéma d'assimilation) et un vecteur de paramètres  $f_2$  estimé relativement loin de

l'information a priori. On peut noter également que les valeurs estimées des paramètres  $f_{2i}$  sont nettement moins proches de l'état vrai que ne l'est l'information a priori : dans ce cas les observations ont dégradé notre niveau d'information sur les paramètres  $f_{2i}$ . Cela est dû à l'inconsistance du schéma d'assimilation : les statistiques d'erreur fixées dans le schéma d'assimilation ne sont pas cohérentes avec la réalité. Il en résulte que l'algorithme d'assimilation corrige abusivement les paramètres  $f_{2i}$  afin d'approcher les observations de manière trop importante.

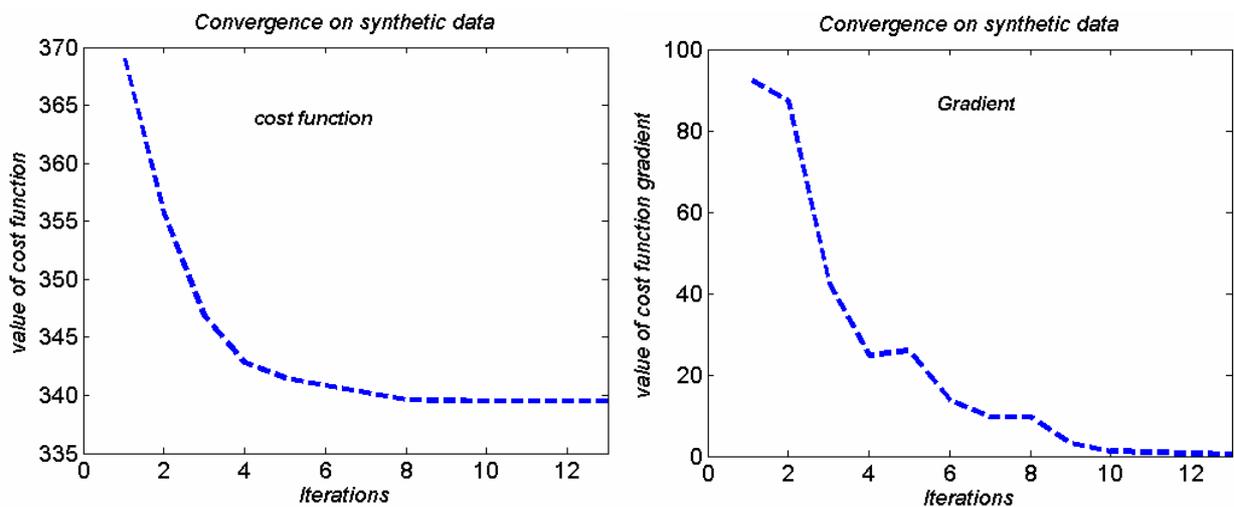


**Fig. IV.11. Poids important de la matrice de covariance  $R_i$  (observation  $\sigma_i$  fixé à 0,5 K) par rapport à la matrice de covariance  $B$  ( $f_2$ ,  $\sigma_{f_{2i}}$  fixé à 0,051) dans le processus d'assimilation : l'information a priori (en étoile) et l'estimation de l'algorithme (en point)**

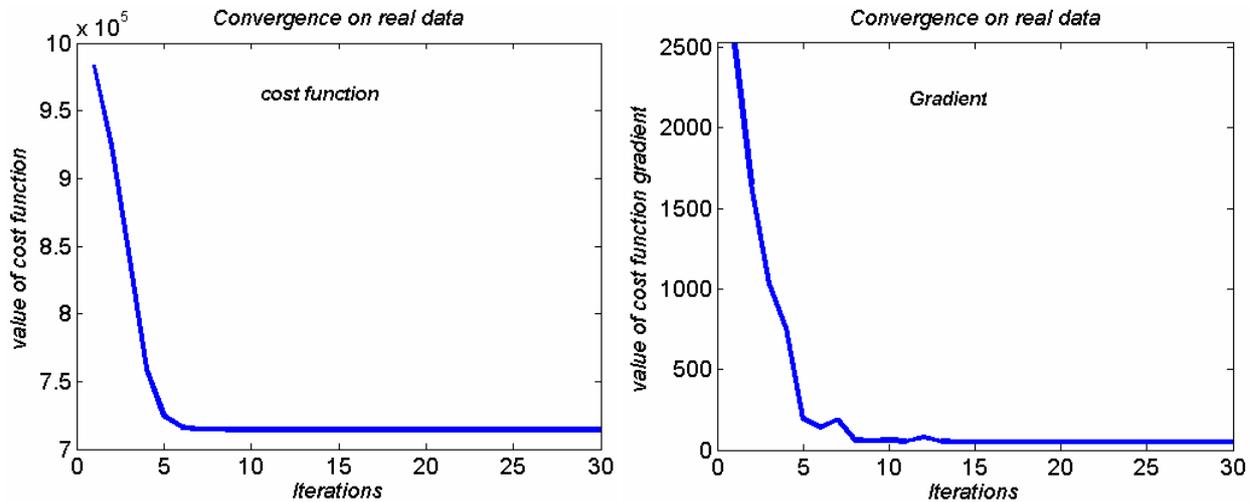
### 3.2.4. Convergence de la fonction coût

Pour conclure sur la validation du schéma de l'algorithme d'assimilation, nous avons analysé la convergence de la fonction coût ( $J$ ) et de son gradient ( $\nabla J$ ) d'une part sur 625 pixels des données d'expérience synthétique (Fig. IV.12) et d'autre part sur l'application au cas réel de la zone Alpilles dans son ensemble, soit sur 62500 pixels (Fig. IV.13).

On peut observer sur les différentes figures que la fonction coût atteint assez rapidement une valeur stable (cela nécessite une dizaine d'itérations sur ces deux exemples), tout comme la norme de son gradient dont la valeur diminue nettement. En théorie, si les hypothèses sous-jacentes au problème d'assimilation sont respectées (et en particulier si les statistiques d'erreur fixées dans les matrices  $B$  et  $R_i$  sont correctes) et que le problème d'optimisation est linéaire ou quasi-linéaire, alors la valeur finale de la fonction devrait être en moyenne égale à la moitié du nombre d'observations (Tarantola, 1987 ; 2005). Pour le cas sur l'expérience synthétique, qui respecte les hypothèses effectuées concernant les erreurs d'observation, de modèle et d'information a priori, on peut voir que la valeur finale de la fonction est bien d'un ordre de grandeur correct. En effet, la valeur finale de la fonction est environ 340 pour une valeur moyenne théorique de  $625/2 = 312,5$  pour les 625 observations. Ce n'est pas le cas de l'application sur le cas réel : la valeur finale est plus de 20 fois plus élevée que la valeur moyenne théorique ce qui montre que les statistiques d'erreurs fixées dans ce cas ne sont pas cohérentes avec la réalité.



**Fig. IV.12. Convergence de la fonction coût ( $J$ ) et du gradient ( $\nabla J$ ) sur données synthétiques**



**Fig. IV.13. Convergence de la fonction coût ( $J$ ) et du gradient ( $\nabla J$ ) sur donnée réelle**

L'arrêt de l'algorithme de minimisation est ici conditionné à deux critères (dont un seul doit être rempli). Le premier concerne la différence normalisée entre la valeur de la fonction à l'itération courante et celle à l'itération précédente. La valeur du seuil en dessous duquel l'algorithme s'arrête a ici été fixée à  $10^6 \cdot \text{eps}$ , où  $\text{eps}$  est la précision machine. Le deuxième critère concerne la norme infinie du gradient. La valeur du seuil en dessous duquel l'algorithme s'arrête a ici été fixée à  $10^{-6}$ . Il semble, sur l'exemple avec les données réelles, que les valeurs fixées pour ces critères sont peut-être un peu faibles puisque l'algorithme effectue encore une quinzaine d'itérations après le début des zones de stabilité de la fonction et du gradient avant de s'arrêter.

## 4. Conclusion

Nous avons consacré cette partie à l'application des outils choisis pour implémenter le système d'assimilation et à la validation du système implémenté.

Le logiciel TAPENADE nous a permis d'obtenir assez rapidement un adjoint valide et performant pour le modèle PBLs. Le nombre de modifications apportées au code de PBLs a été faible, et il faut souligner ici la robustesse de l'outil TAPENADE : les quelques problèmes rencontrés ont été dus à l'implémentation parfois peu robuste, voire incorrecte de PBLs, et lorsque TAPENADE n'arrivait pas à interpréter quelques passages, ceux-ci ont été signalés dans les messages d'avertissement ou d'erreur générés lors de son application. Les outils de debuggage fournis avec TAPENADE ont été précieux même s'il faut noter que nous n'avons eu à les utiliser que sur le code tangent et que leur application sur des codes plus complexes n'est probablement pas aussi aisée. L'utilisation de PALM n'a nécessité que des modifications mineures des codes existants. Si l'intérêt de son utilisation n'est pas flagrant sur cette application, il faut cependant considérer les évolutions futures de ce travail. En effet, il s'agira dans le cadre des travaux sur la modélisation du paysage, dont cette étude ne représente que les prémisses, de coupler des modèles complexes et parfois très coûteux et de les inclure au sein d'un ou plusieurs système(s) d'assimilation. Les potentialités de l'outil PALM seront alors pleinement exploitées.

Au vu des résultats des expériences de vérification effectuées, on peut conclure que l'algorithme d'assimilation réagit correctement en fonction de l'information qu'on lui fournit, et on considèrera donc son implémentation comme vérifiée. On peut noter également que les statistiques d'erreur influent fortement sur les résultats de la procédure d'assimilation, et qu'il est nécessaire que ces statistiques soient consistantes avec la réalité pour utiliser correctement les informations disponibles. Nous avons aussi vu que les valeurs initiales des paramètres  $z0_i$  et  $f_{2i}$  n'influencent pas l'estimation des paramètres grâce à l'ajout de l'information a priori. Ces expériences nous permettent également d'évaluer le coût de la procédure d'assimilation. Ce coût est proportionnel à la somme du temps d'exécution du modèle PBLs et de celui de son adjoint qui doit être multiplié par le nombre de leur évaluation au sein du processus de minimisation. Par ailleurs, ce coût est presque linéaire au nombre de pixels traités. Le tableau IV.3 ci-dessous donne le coût estimé de l'application de l'algorithme sur toute la zone Alpilles

pour une simulation de 5h à 12h. Ce temps est assez important, même s'il a été réduit d'un tiers grâce à l'approche du préconditionnement de la fonction coût. On peut remarquer qu'il est deux fois plus important dans le cas où l'on cherche à estimer les deux paramètres à la fois (algorithme C) que dans le cas des algorithmes A et B. Il faut cependant noter que ces temps de calcul ont été mesurés sur une machine relativement ancienne (processeur: UltraSPARC-IIIi, 1.25GHz, 2Go de mémoire), et que plusieurs optimisations pourraient être effectuées pour réduire ce temps de calcul (fixer les critères d'arrêt à des valeurs moins contraignantes, optimiser voire paralléliser les modèles directs et adjoints).

**Tab. IV.3. Temps CPU d'exécution de l'algorithme d'assimilation entre 5h – 12h**

<i>Nombre de pixels</i>	<i>PBLs (une exécution)</i>	<i>PBLs adjoint (une exécution)</i>	<i>Système d'assimilation (algorithme A ou B)</i>	<i>Système d'assimilation (algorithme C)</i>
62500 (250 x 250)	11 mn	1h18	5 jours	10 jours

La partie suivante sera consacrée à l'analyse des performances du système d'assimilation en termes de qualité des résultats obtenus.



***CINQUIEME PARTIE :***

***PERFORMANCES DE L'ALGORITHME D'ASSIMILATION SUR  
LES DONNEES SYNTHETIQUES ET SUR LES DONNEES  
REELLES DE L'EXPERIENCE ALPILLES – ReSeDA***

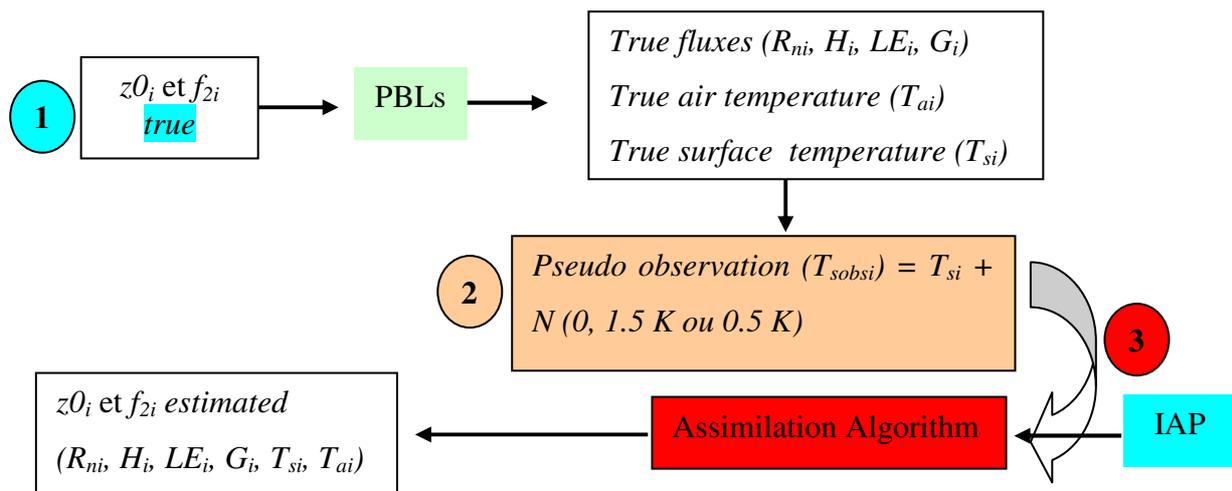


## ***Introduction***

Dans cette partie, nous avons évalué la performance de l'algorithme d'assimilation d'une part sur les données des expériences synthétiques et d'autre part sur les données Alpilles ReSeDA (cf. **Chapitre II Section 4**).

Le schéma V.1 rappelle le principe de l'application de l'algorithme d'assimilation sur les données synthétiques : à partir d'observations bruitées ( $T_{sobs}$ ) sur l'ensemble des pixels considérés, l'algorithme cherche à retrouver l'état vrai des paramètres estimés (rugosité aérodynamique  $z0_i$  et facteur d'humidité du sol normalisé  $f_{2i}$ ), des flux de surface ( $R_{ni}$ ,  $H_i$ ,  $LE_i$ ,  $G_i$ ) et des températures de surface ( $T_{si}$ ) et de l'air à 2 m ( $T_{ai}$ ) sur chaque pixel. Nous avons essayé pour ces expériences synthétiques de se rapprocher de conditions réalistes. En effet ces données sont construites à partir des variables mesurées sur la zone Alpilles (cf. section 4.1 du chapitre II). 625 pixels choisis au hasard de la carte d'occupation d'Alpilles ont constitué la base de données de l'expérience synthétique. Il y est représenté des sols nus et des couverts développés, des zones humides et sèches. Ces expériences synthétiques nous ont permis de tester différentes conditions d'information pour l'application du système d'assimilation : des observations de température de surface plus ou moins précises, et l'hypothèse de connaissance ou non de l'un des deux paramètres  $z0$  et  $f_2$  (Algorithme A : estimer uniquement  $z0$ ,  $f_2$  est connu ; Algorithme B : estimer uniquement  $f_2$ ,  $z0$  est connu ; Algorithme C : estimer les deux paramètres à la fois).

Enfin, nous avons réalisé ces tests soit en supposant que le modèle était parfait, soit en prenant en compte une forme d'erreur de modèle, par l'introduction d'erreur de mesure sur ses variables d'entrée non estimées par l'algorithme (albédo,  $LAI$ , vitesse du vent ...). Les performances du système d'assimilation dans ses différentes configurations et pour les différents tests sont alors évaluées en comparant notamment les variables de sortie d'intérêt du modèle PBLs (flux de surface ( $R_{ni}$ ,  $H_i$ ,  $LE_i$ ,  $G_i$ ), températures de surface ( $T_{si}$ ) et de l'air à 2 m ( $T_{ai}$ )), avant l'assimilation (i.e calculées à partir des valeurs de l'information a priori) et après l'assimilation, à l'état vrai choisi. L'ensemble des tests (données synthétiques) a été appliqué sur les deux versions du modèle PBLs.



**Fig. V.1. Schéma de la procédure d'assimilation à partir des pseudos observations**

En conditions réelles, l'application d'un système d'assimilation pose un certain nombre de problèmes qui peuvent nuire à la qualité de ses résultats. Pour notre application, ces problèmes sont essentiellement liés à la qualité des informations à disposition (modèle, observations, information a priori) et au non respect des hypothèses sous-jacentes au système d'assimilation (erreurs gaussiennes, centrées, non corrélées, ayant des variances similaires à celles fixées dans le système, ...). Si l'impact de certains de ces problèmes peut être évalué sur des expériences synthétiques, rien ne remplace une application à des données réelles. Nous avons donc réalisé une application du système d'assimilation avec la deuxième version de PBLs sur les données d'Alpilles, en considérant l'ensemble des 62500 pixels pour lesquels nous disposons de mesures de température de surface. Deux façons d'estimer l'information a priori ont été testées. Pour l'évaluation des résultats, nous avons considéré en particulier les 7 parcelles sur lesquelles avaient été disposées des stations de mesure et pour lesquelles nous disposons donc de mesures des variables d'intérêt du modèle. Nous avons également représenté et analysé les cartographies des flux de surface, notamment de l'évapotranspiration (flux  $LE$ ), sur l'ensemble de la zone considérée, avant et après assimilation.

D'autre part, les résultats obtenus après assimilation sont également comparés à ceux obtenus avec la méthode résiduelle définie dans la **section 5** du **chapitre II**.

## 1. Performances sur les données synthétiques

### 1.1. Résultats de l'algorithme d'assimilation avec la version 1 de PBLs

Il faut rappeler que la version 1 du modèle PBLs est caractérisée par le rapport  $G/R_n$  défini constant ( $c_0$ ) pour toute la zone et par le paramètre  $kB^{-1}$  (la rugosité thermique  $z0_h$ ) déterminé à partir de la rugosité aérodynamique ( $z0$ ) par un simple ratio. Ces caractéristiques sont importantes pour l'analyse de l'erreur modèle et particulièrement pour la comparaison des résultats du système d'assimilation avec ceux obtenus par la méthode résiduelle.

#### 1.1.1. Cas sans erreur modèle

Les résultats obtenus par les trois algorithmes A (on n'estime que  $z0$ ), B (on n'estime que  $f_2$ ) et C (on estime  $z0$  et  $f_2$ ) sont présentés dans le tableau V.1 ci dessous.

Le premier groupe de colonnes montre les performances du modèle sans assimilation : les valeurs des paramètres sont fixées aux valeurs a priori ; dans le cas A, les  $f_2$  sont fixés à leurs valeurs vraies, et dans le cas B, ce sont les  $z0$  qui sont fixés à leurs valeurs vraies. La comparaison des performances du modèle sur les cas A et B montre que les flux turbulents  $H$  et  $LE$  sont particulièrement sensibles aux erreurs sur le paramètre  $f_2$  alors que les flux  $R_n$  et  $G$  et les températures  $T_s$  et  $T_a$  sont légèrement plus sensibles aux erreurs sur  $z0$  qu'à celles sur  $f_2$ . Ce résultat exprime les différences de sensibilité des sorties de PBLs à ses paramètres d'entrées  $z0$  et  $f_2$ . Le cas C montre des résultats légèrement dégradés par rapport au cas B (flux turbulents) ou par rapport au cas A ( $R_n$ ,  $G$  et températures).

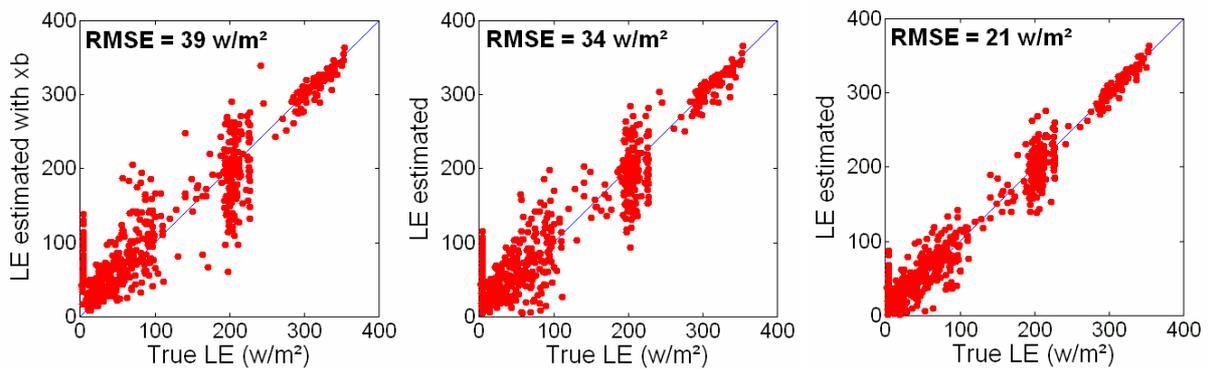
**Tab. V.1. RMSE entre les résultats de l'assimilation avec l'erreur d'observation et les états vrais pour les flux d'énergie (à midi), la température de surface et la température de l'air à 2 m (à midi) et les paramètres de  $z_0$  et  $f_2$  retrouvés.**

	PBLs (sans Assimilation)			Assimilation $\sigma_i = 1,5$ K			Assimilation $\sigma_i = 0,5$ K		
	A	B	C	A	B	C	A	B	C
$R_n$ ( $Wm^{-2}$ )	4,2	4,0	5,2	3,6	3,1	4,0	2,3	2,0	2,5
$H$ ( $Wm^{-2}$ )	8	36	38	7	31	35	4	19	32
$LE$ ( $Wm^{-2}$ )	5	39	39	4	34	37	3	21	32
$G$ ( $Wm^{-2}$ )	0,6	0,5	0,8	0,5	0,4	0,6	0,3	0,3	0,4
$T_s$ (K)	0,70	0,61	0,80	0,59	0,52	0,71	0,37	0,33	0,41
$T_a$ (K)	0,40	0,30	0,40	0,36	0,20	0,38	0,23	0,12	0,25
$z_0$ (m)	0,10	x	0,10	0,09	X	0,09	0,06	x	0,09
$f_2$ (-)	x	0,140	0,140	x	0,121	0,124	x	0,080	0,109

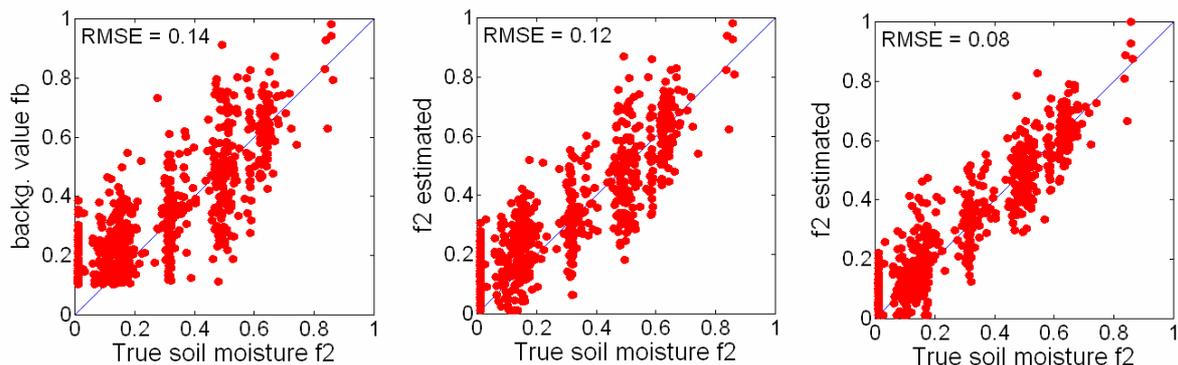
Les deux autres groupes de colonnes présentent les résultats des assimilations au moyen des 3 algorithmes et pour les deux niveaux d'erreur d'observation fixés. Pour chacun des niveaux d'erreur d'observation, la comparaison des résultats obtenus par les algorithmes A et B reflète les résultats obtenus sans assimilation : les résultats sur les flux  $H$  et  $LE$  sont nettement meilleurs lorsque  $f_2$  est fixé à sa vraie valeur (algorithme A) et ceux sur les variables  $R_n$ ,  $G$ ,  $T_s$  et  $T_a$  sont meilleurs lorsque  $z_0$  est fixé à sa vraie valeur (algorithme B). La comparaison des résultats obtenus avec les algorithmes A, B et C dans les cas avec assimilation montre que les résultats du cas C sont systématiquement moins bons que ceux des cas A et B, et ce de façon parfois assez marquée lorsque les observations sont plus précises (cas des flux  $H$  et  $LE$  pour une erreur d'observation de 0,5 K). Cela s'explique par le fait que dans les cas A et B un seul paramètre par pixel parmi  $f_2$  et  $z_0$  est estimé et l'autre est fixé à sa valeur vraie. Dans le cas C, l'information à disposition est moins précise : les deux paramètres doivent être estimés sur chaque pixel à partir du même jeu d'observation que dans les cas A et B, et d'une information a priori entachée d'erreur pour les deux paramètres.

Une comparaison globale des résultats avec et sans assimilation montre que les résultats sont, en moyenne sur l'ensemble des pixels, systématiquement meilleurs avec assimilation que sans assimilation. Ceci est vrai quels que soient les variables ou paramètres estimés considérés, même si parfois l'amélioration peut être faible. Les résultats obtenus avec une erreur d'observation de 0,5 K sont très nettement meilleurs que ceux obtenus avec une erreur d'observation de 1,5 K (sauf pour les flux  $H$  et  $LE$  pour l'algorithme C pour lequel

l'amélioration est moins sensible). Cela est particulièrement mis en évidence sur les figures V.2 et V.3 pour l'estimation du flux  $LE$  et du paramètre  $f_2$ ; pour  $LE$ , dans le cas de l'algorithme B, l'assimilation des observations avec une erreur de 1,5 K permet une légère amélioration de  $RMSE$  de  $39 \text{ Wm}^{-2}$  à  $34 \text{ Wm}^{-2}$  par rapport au cas sans assimilation; l'assimilation des observations avec une erreur de 0,5 K permet, elle, une amélioration beaucoup plus nette portant le  $RMSE$  à  $21 \text{ Wm}^{-2}$ .



**Fig. V.2.** a)  $LE$  simulé avec le modèle PBLs à partir de l'information a priori à 12h; b)  $LE$  estimé en utilisant l'algorithme B avec l'erreur d'observation de 1,5 K à 12h; c)  $LE$  estimé en utilisant l'algorithme B avec l'erreur d'observation de 0,5 K à 12h



**Fig. V.3.** a) IAP de l'humidité du sol normalisée ( $f_2$ ); b)  $f_2$  estimé par l'algorithme B avec erreur d'observation de 1,5 K; c)  $f_2$  estimé par l'algorithme B avec erreur d'observation de 0,5 K.

En général, l'intérêt du système d'assimilation peut paraître limité pour certaines variables dans le cas avec une erreur d'observation de 1,5 K : ces observations n'apportent pas beaucoup d'information en plus de l'information a priori. Par contre, le cas avec une erreur de 0,5 K montre tout l'intérêt d'avoir des observations précises à disposition. Nous pouvons cependant constater dans le tableau V.1 que les résultats obtenus avec l'algorithme C pour une erreur d'observation de 0,5 K ne sont que peu améliorés par rapport à ceux obtenus pour une erreur de 1,5 K. Dans ce cas (Algorithme C, erreur 0,5 K), l'information apportée par des observations précises n'est pas suffisante pour compenser fortement la mauvaise connaissance simultanée des deux paramètres  $z_0$  et  $f_2$ . On peut observer de plus que les résultats obtenus avec l'algorithme C pour une erreur de 0,5 K sont d'une qualité à peu près équivalente à ceux obtenus avec l'algorithme B pour une erreur d'observation de 1,5 K. Ainsi, la connaissance parfaite de  $z_0$  compense la perte de précision des observations de température de surface.

Le tableau V.1 présente également les résultats d'estimation de l'humidité du sol normalisée ( $f_2$ ) et de la rugosité aérodynamique ( $z_0$ ) par assimilation (voir aussi la Figure V.3). Il apparaît que la procédure d'assimilation améliore toujours les valeurs de ces paramètres, en particulier lorsque les données observées sont les plus précises. Les estimations sont également plus précises pour les algorithmes A et B, pour lesquels un seul paramètre est recherché à la fois.

Ces erreurs observées s'inscrivent dans le même ordre de grandeur que celles données par d'autres auteurs qui ont travaillé avec des données synthétiques. L'étude de Pipunic et al. (2008) a relevé des *RMSE* sur le flux *LE* par rapport à l'état vrai qui peuvent atteindre jusqu'à  $74 \text{ Wm}^{-2}$  (notamment  $48 \text{ Wm}^{-2}$  sur le flux *H*).

Tous ces premiers résultats montrent donc l'intérêt de la méthode d'assimilation et le rôle important de la sensibilité au paramètre d'humidité normalisée  $f_2$ . Ils mettent également bien en évidence le rôle de la précision des mesures utilisées pour réaliser l'assimilation.

Remarquons cependant que, si nous nous plaçons ici dans un cas avec erreur d'observation, il n'en reste pas moins que les conditions d'application sont idéales : le modèle est parfait et les hypothèses effectuées sur les erreurs d'information a priori et sur les observations sont parfaitement respectées. Le système d'assimilation réussit dans ces conditions à combiner de façon optimale l'ensemble des informations mises à sa disposition. Ainsi, on peut par exemple voir que l'erreur sur la température de surface  $T_s$  après assimilation est à la fois inférieure à

celle obtenue par le modèle direct en propageant l'information a priori et à l'erreur sur les observations assimilées. Dans les sections suivantes, le cas du modèle « non parfait » et le cas des données réelles vont être analysés pouvant nous amener à nuancer nos premières affirmations.

### 1.1.2. Cas avec erreur modèle

L'assimilation de la température de surface dans PBLs a été utilisée pour retrouver les paramètres  $f_2$  et  $z_0$  de PBLs qui sont difficiles à connaître à l'échelle d'une petite région, en particulier parce qu'il n'existe pas de méthode permettant de les mesurer directement à cette échelle et d'accéder facilement à leur variabilité spatiale. Les autres paramètres sont supposés plus accessibles et ont donc jusqu'à présent été directement introduits en entrée du modèle. Ce sont l'albédo ( $\alpha$ ), le  $LAI$ , la vitesse moyenne journalière du vent  $u_a$ , le rapport  $G/R_n$  ou  $c_0$ , et la rugosité  $z_0$  (pour cette dernière uniquement dans le cas de l'algorithme B). L'obtention de ces paramètres a été détaillée dans les parties précédentes (en particulier **Partie II - section 4**, ainsi que **Partie III** pour  $c_0$ ). Leur connaissance est bien sûr entachée d'erreurs qui ont été décrites dans la **section 5** de la **Partie II**. Ces erreurs correspondent à celles que nous avons pu évaluer lors des précédents travaux portant sur les données Alpilles. Elles sont cependant représentatives des niveaux d'erreur constatés de nos jours dans la bibliographie.

Dans cette partie, nous présenterons les résultats des expériences d'assimilation sur les données synthétiques lorsque ces erreurs sont introduites en entrée du modèle. Les erreurs sur les sorties de PBLs induites par celles sur les entrées non estimées dans le système d'assimilation peuvent être considérées comme des erreurs de modèle. Contrairement au cas précédent, nous nous placerons ici dans des conditions sous-optimales : d'une part les erreurs de modèle sur la température de surface n'ont pas une distribution normale centrée et ne sont probablement pas indépendantes entre pixels, et d'autre part elles ne seront pas prises en compte dans la matrice  $R_i$ , qui restera identique à celle utilisée dans le cas précédent (i.e. diagonale et de variance, la variance d'erreur des observations).

Pour rappel, les erreurs sont introduites dans l'analyse pour le  $LAI$ , l'albédo ( $\alpha$ ),  $c_0$  ( $G/R_n$ ) et la vitesse moyenne journalière du vent ( $u_a$ ) en les ajoutant aux valeurs vraies définies dans le scénario vrai. Nous avons ajouté des erreurs aléatoires suivant des distributions gaussiennes pour l'albédo (avec un écart type de 0,02) et pour le  $LAI$  (avec un écart type de 0,5). Les valeurs de  $c_0$ , habituellement considérées constantes pour tout le domaine, sont calculées ici avec la formule décrite par Choudhury et al. (1987) en fonction du  $LAI$  pour chaque pixel. La vitesse moyenne journalière du vent, également fixée constante sur l'ensemble de la zone considérée à  $1,8 \text{ ms}^{-1}$ , a été modifiée en y ajoutant  $\pm 1 \text{ ms}^{-1}$ . Quant à la rugosité  $z_0$  (introduite

dans le cas de l'algorithme B seulement), elle a été fixée égale à l'information a priori qui a été définie pour l'ensemble des tests des algorithmes A et C.

Dans cette section l'impact de ces erreurs n'est présenté que pour les algorithmes B et C (cf. **annexe II** pour les résultats de l'algorithme A).

Le Tableau V.2 présente l'influence des erreurs des données d'entrée du modèle PBLs sur les résultats de l'assimilation ; l'ensemble des *RMSE* est calculé par comparaison à l'état vrai.

Dans le tableau :

- la troisième colonne du tableau correspond à une analyse de sensibilité du modèle direct : impact des erreurs des entrées sur les simulations effectuées dans les conditions de l'état de référence vrai ;
- la quatrième colonne présente l'impact des erreurs des entrées sur les simulations quand l'humidité normalisée du sol  $f_2$  est fixée à sa valeur a priori (valeur qui est utilisée dans les expériences d'assimilation pour l'algorithme B) ; et quand l'humidité normalisée du sol  $f_2$  et la rugosité aérodynamique  $z_0$  sont fixées à leurs valeur a priori (pour l'algorithme C)
- les colonnes suivantes présentent les résultats des algorithmes d'assimilation B et C pour les erreurs d'observation  $\sigma_i$  de 1,5 K et 0,5 K (toujours par rapports aux états de référence vrai) ;
- il importe de noter également que la première ligne du tableau (cas sans erreur), similaire aux résultats d'assimilation pour le modèle parfait présentés dans la section précédente, servira de référence pour évaluer l'impact des erreurs sur les performances de l'assimilation.

#### Analyse de sensibilité

Les principaux résultats de l'analyse de sensibilité présentée dans la troisième colonne (voir aussi le tableau III.3 du **chapitre III**) sont les suivants :

- si l'on considère les flux de chaleur, les paramètres se classent du moins sensibles au plus sensibles ainsi :  $\alpha$ ,  $u_a$ ,  $z_0$ ,  $c_0$ , et  $LAI$  (en observant la première ligne de la 4<sup>ème</sup> colonne, il est possible d'ajouter  $f_2$  aux paramètres les plus sensibles)
- l'impact du  $LAI$  est très fort sur les flux de chaleur sensible  $H$  et latente  $LE$  (avec des *RMSE*  $> 40 \text{ Wm}^{-2}$ ), exprimant le rôle central du  $LAI$  sur la représentation des échanges d'énergie

dans PBLs (en particulier en définissant la taille de la surface d'échange de vapeur entre les plante et l'atmosphère)

- le paramètre  $c_0$  affecte essentiellement les flux de  $H$  et  $G$ , exprimant une compensation classiquement constatée dans les modèles TSVA entre ces deux flux lorsque les propriétés de surface du sol sont modifiées (cf. par exemple Olioso et al., 2002b)
- l'albédo affecte essentiellement le flux de rayonnement net ( $R_n$ ) dont il est un déterminant principal (Equation II.3)
- hormis l'albédo, l'ensemble des paramètres a un fort impact sur la température de surface (avec un  $RMSE$  supérieur à 0,6 K) ; le classement des paramètres du moins au plus sensible est légèrement différent du classement établi pour les flux de chaleur :  $\alpha$ , [ $u_a$   $c_0$   $f_2$ ],  $z_0$  et  $LAI$  ; comme la température de surface constitue l'observation utilisée, ces différences peuvent avoir un impact sur la qualité des résultats d'assimilation
- la température de l'air est surtout sensible aux paramètres qui affectent la température de surface, et à la rugosité qui a un impact direct sur la forme des profils de température dans l'atmosphère.

Notons que la colonne 4 qui a été obtenue également sans assimilation, mais en considérant la carte des valeurs a priori de  $f_2$  à la place des valeurs vraies, montre des résultats de sensibilité équivalents à la colonne 3 (hormis les niveaux de  $RMSE$  plus élevés en raison des différences de  $f_2$ ).

**Tab. V.2.  $RMSE$  à 12h entre les états vrais et les sorties de l'algorithme d'assimilation comparées aux résultats de PBLs (sans assimilation) avec ajout d'erreur sur l'albédo ( $\alpha$ ), sur  $LAI$ , sur la vitesse du vent ( $u_a$ ), sur le rapport  $G/R_n$  ( $c_0$ ) et sur la rugosité  $z_0$  par l'algorithme B. Colonne 1 (ajout d'erreur) : erreur moyenne ajoutée par rapport à l'état vrai sur l'ensemble des 625 pixels. Colonne 3: (analyse de sensibilité) montrant l'impact de chaque variable avec l'erreur sur les résultats de PBLs par rapport au scénario vrai. Colonnes 7 et 8: Résultats de l'algorithme C montrant la similitude avec l'algorithme B**

Ajout erreur		Analyse de sensibilité	PBLs		Algo B	Algo B	Algo C	Algo C
			B	C	$\sigma = 1, 5 \text{ K}$	$\sigma = 0, 5 \text{ K}$	$\sigma = 1, 5 \text{ K}$	$\sigma = 0, 5 \text{ K}$
(a) sans erreur	$R_n \text{ (Wm}^{-2}\text{)}$	0	4,0	5,2	3,1	2,0	4,0	2,5
	$H \text{ (Wm}^{-2}\text{)}$	0	36	38	31	19	35	32
	$LE \text{ (Wm}^{-2}\text{)}$	0	39	39	34	21	37	32
	$G \text{ (Wm}^{-2}\text{)}$	0	0,5	0,8	0,4	0,3	0,6	0,4
	$T_s \text{ (}^\circ\text{C)}$	0	0,61	0,80	0,52	0,33	0,71	0,41
	$T_a \text{ (}^\circ\text{C)}$	0	0,30	0,40	0,20	0,12	0,38	0,25
	$z0 \text{ (m)}$	-	-	0,100	-	-	0,094	0,087
	$f_2 \text{ (-)}$	-	0,140	0,140	0,121	0,080	0,124	0,109
(b) $\alpha$ $\pm 8\%$	$R_n \text{ (Wm}^{-2}\text{)}$	10,0	10,2	12	10,1	9,7	11,0	10,9
	$H \text{ (Wm}^{-2}\text{)}$	7	38	40	32	20	37	33
	$LE \text{ (Wm}^{-2}\text{)}$	3	38	39	34	21	36	32
	$G \text{ (Wm}^{-2}\text{)}$	1,5	1,5	1,6	1,5	1,4	1,6	1,6
	$T_s \text{ (}^\circ\text{C)}$	0,20	0,70	0,90	0,51	0,35	0,73	0,42
	$T_a \text{ (}^\circ\text{C)}$	0,11	0,31	0,50	0,23	0,14	0,41	0,25
	$z0 \text{ (m)}$	-	-	0,100	-	-	0,095	0,090
	$f_2 \text{ (-)}$	-	0,140	0,140	0,120	0,084	0,123	0,107
(c) LAI $\pm 44\%$	$R_n \text{ (Wm}^{-2}\text{)}$	5	6	7	5	3	5	3
	$H \text{ (Wm}^{-2}\text{)}$	42	56	57	49	33	53	47
	$LE \text{ (Wm}^{-2}\text{)}$	47	61	61	53	36	56	49
	$G \text{ (Wm}^{-2}\text{)}$	0,8	0,9	1,1	0,7	0,5	0,8	0,4
	$T_s \text{ (}^\circ\text{C)}$	0,82	1,02	1,21	0,82	0,51	0,92	0,44
	$T_a \text{ (}^\circ\text{C)}$	0,30	0,39	0,52	0,31	0,21	0,42	0,21
	$z0 \text{ (m)}$	-	-	0,100	-	-	0,098	0,096
	$f_2 \text{ (-)}$	-	0,140	0,140	0,125	0,125	0,127	0,124
(d) $u_a = 0,8$ $-1 \text{ ms}^{-1}$	$R_n \text{ (Wm}^{-2}\text{)}$	4	5	7	4	3	5	3
	$H \text{ (Wm}^{-2}\text{)}$	6	36	38	31	21	36	33
	$LE \text{ (Wm}^{-2}\text{)}$	3	39	40	34	22	37	33
	$G \text{ (Wm}^{-2}\text{)}$	0,6	0,8	1,1	0,6	0,4	0,8	0,4

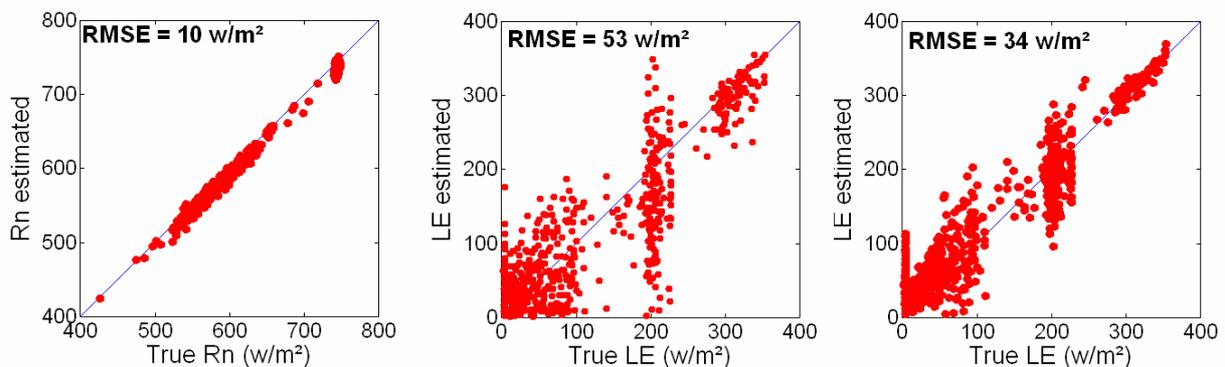
	$T_s$ (°C)	0,60	0,83	1,10	0,72	0,44	0,82	0,45
	$T_a$ (°C)	0,31	0,42	0,54	0,32	0,34	0,43	0,36
	$z0$ (m)	-	-	0,100	-	-	0,102	0,095
	$f_2$ (-)	-	0,140	0,140	0,121	0,097	0,124	0,115
(e) $u_a = 2,8$ $+1 \text{ ms}^{-1}$	$R_n$ ( $\text{Wm}^{-2}$ )	4	6	6	5	3	5	3
	$H$ ( $\text{Wm}^{-2}$ )	7	35	36	33	28	35	34
	$LE$ ( $\text{Wm}^{-2}$ )	4	38	39	35	28	36	34
	$G$ ( $\text{Wm}^{-2}$ )	0,7	0,9	1,0	0,7	0,5	0,7	0,4
	$T_s$ (°C)	0,61	0,95	1,04	0,75	0,44	0,83	0,45
	$T_a$ (°C)	0,10	0,21	0,43	0,20	0,29	0,42	0,40
	$z0$ (m)	-	-	0,100	-	-	0,101	0,094
	$f_2$ (-)	-	0,140	0,140	0,121	0,100	0,123	0,112
(f) $c_0$ $\pm 36\%$	$R_n$ ( $\text{Wm}^{-2}$ )	3	5	6	4	2	5	2
	$H$ ( $\text{Wm}^{-2}$ )	28	48	50	42	27	46	40
	$LE$ ( $\text{Wm}^{-2}$ )	6	38	39	34	28	36	34
	$G$ ( $\text{Wm}^{-2}$ )	35	35	35	34	34	34	34
	$T_s$ (°C)	0,65	0,90	1,10	0,71	0,41	0,84	0,45
	$T_a$ (°C)	0,30	0,42	0,53	0,33	0,20	0,44	0,28
	$z0$ (m)	-	-	0,100	-	-	0,102	0,099
	$f_2$ (-)	-	0,140	0,140	0,120	0,102	0,122	0,111
(g) $z0$ $\pm 25\%$	$R_n$ ( $\text{Wm}^{-2}$ )	4,2	5,5	a)	4	2	a)	a)
	$H$ ( $\text{Wm}^{-2}$ )	8	38	a)	36	35	a)	a)
	$LE$ ( $\text{Wm}^{-2}$ )	5	39	a)	37	35	a)	a)
	$G$ ( $\text{Wm}^{-2}$ )	0,6	0,8	a)	0,6	0,4	a)	a)
	$T_s$ (°C)	0,70	0,80	a)	0,71	0,44	a)	a)
	$T_a$ (°C)	0,40	0,40	a)	0,28	0,31	a)	a)
	$z0$ (m)	-	-	a)	-	-	a)	a)
	$f_2$ (-)	-	-	a)	0,125	0,123	a)	a)

(h) <i>toutes erreurs <math>u_a = 0,8</math></i>	$R_n$ ( $Wm^{-2}$ )	14,0	13,4	13,4	13,2	12,0	13,0	12,0
	$H$ ( $Wm^{-2}$ )	55	71	71	61	49	63	56
	$LE$ ( $Wm^{-2}$ )	45	60	60	53	48	54	49
	$G$ ( $Wm^{-2}$ )	34	34	34	33	33	33	33
	$T_s$ ( $^{\circ}C$ )	1,34	1,50	1,50	1,21	0,82	1,12	0,52
	$T_a$ ( $^{\circ}C$ )	0,72	0,84	0,84	0,71	0,53	0,63	0,45
	$z0$ (m)	-	-	0,100	-	-	0,106	0,116
	$f_2$ (-)	-	0,140	0,140	0,128	0,130	0,126	0,128
(i) <i>toutes erreurs <math>u_a = 2,8</math></i>	$R_n$ ( $Wm^{-2}$ )	9	9	9	8	9	9	9
	$H$ ( $Wm^{-2}$ )	52	67	67	57	48	58	50
	$LE$ ( $Wm^{-2}$ )	44	59	59	54	55	54	52
	$G$ ( $Wm^{-2}$ )	34	35	35	34	33	34	33
	$T_s$ ( $^{\circ}C$ )	1,51	1,71	1,71	1,22	0,74	1,20	0,52
	$T_a$ ( $^{\circ}C$ )	0,60	0,63	0,63	0,40	0,42	0,43	0,41
	$z0$ (m)	-	-	0,100	-	-	0,107	0,119
	$f_2$ (-)	-	0,140	0,140	0,132	0,137	0,131	0,14

## ✚ Analyse générale des résultats

La Figure V.4 est présentée à titre d'illustration des résultats obtenus :

- impact de la dégradation de l'albédo sur l'estimation du flux  $R_n$  avec une forte augmentation du  $RMSE$  par rapport au cas sans erreur (de  $4 \text{ Wm}^{-2}$  à  $10,2 \text{ Wm}^{-2}$  dans le cas de l'algorithme B) ; l'assimilation de la température de surface ne permet pas d'améliorer l'estimation de  $R_n$  ce qui s'explique par **-i**) la dépendance directe de  $R_n$  à l'albédo, et **-ii**) la faible dépendance de  $R_n$  aux faibles variations de température de surface considérées ;
- impact de la dégradation du  $LAI$  sur les flux  $H$  et  $LE$  avec des résultats d'assimilation fortement dégradés par rapport au cas sans erreur (quelque soit l'algorithme et le niveau d'erreur sur la température de surface) ; l'assimilation permet d'améliorer sensiblement les résultats en particulier pour  $LE$  et dans les cas où  $\sigma_i = 0,5 \text{ K}$  ;
- impact de la dégradation de la vitesse du vent sur les flux turbulents avec un comportement très proche de celui du cas sans erreur.



**Fig. V.4.** a)  $R_n$  estimé à 12h par l'algorithme B avec erreur d'observation de 1,5 K et erreur sur l'albédo ; b)  $LE$  à 12h estimé par l'algorithme B avec erreur d'observation de 1,5 K et erreur sur  $LAI$  ; c)  $LE$  estimé à 12h par l'algorithme B avec erreur d'observation de 1,5 K et erreur sur  $u_a$

Dans l'ensemble (Tableau V.2), les différents algorithmes d'assimilation montrent des améliorations sur les estimations des flux et des températures. Pour les flux turbulents  $H$  et  $LE$ , les améliorations par rapport aux états a priori peuvent être du même ordre que celles obtenues pour le modèle sans erreur (cas  $\alpha$ ,  $LAI$ ,  $u_a = 0,8 \text{ ms}^{-1}$ ,  $c_0$ ) : meilleurs résultats pour l'algorithme B que pour l'algorithme C, forte amélioration des estimations de flux avec une température

observée précise dans le cas de l'algorithme B ; amélioration plus légère dans le cas de l'algorithme C. Dans plusieurs autres cas ( $u_a = 2,8 \text{ ms}^{-1}$ ,  $z0$ , et toutes erreurs), les améliorations sont plus faibles, en particulier pour  $LE$ . Dans le cas de  $z0$  aucune amélioration des flux n'est obtenue. Il en est de même pour le cas où toutes les erreurs sont combinées avec  $u_a = 2,8 \text{ ms}^{-1}$ .

Concernant les températures de surface, le processus d'assimilation conduit en général à des améliorations sensibles des simulations (par rapport aux observations), en particulier pour le niveau d'erreur  $\sigma_i = 1,5 \text{ K}$  pour lequel des  $RMSE$  finaux entre  $0,51 \text{ K}$  et  $0,92 \text{ K}$  sont obtenus (mais attention cette amélioration est essentiellement due à la qualité de l'information a priori). Les améliorations sont plus faibles pour  $\sigma_i = 0,5 \text{ K}$  avec des  $RMSE$  finaux de l'ordre de  $0,33$  et  $0,42 \text{ K}$  (plus fortement influencé par contre par la qualité de l'observation). Cependant, lorsque toutes les erreurs sont considérées simultanément, les améliorations sont faibles et on constate même une dégradation des températures pour l'algorithme B et l'erreur de  $0,5 \text{ K}$ .

Lorsque l'on s'intéresse aux paramètres retrouvés, des améliorations sont obtenues dans presque tous les cas pour  $f_2$  (sauf algorithme C avec  $0,5 \text{ K}$  d'erreur sur la température) alors que pour  $z0$  les estimations sont souvent faiblement améliorées, voire dégradées. Lorsque toutes les erreurs sont considérées simultanément, la dégradation de  $z0$  peut même être importante (en particulier dans le cas de la meilleure précision des observations de températures).

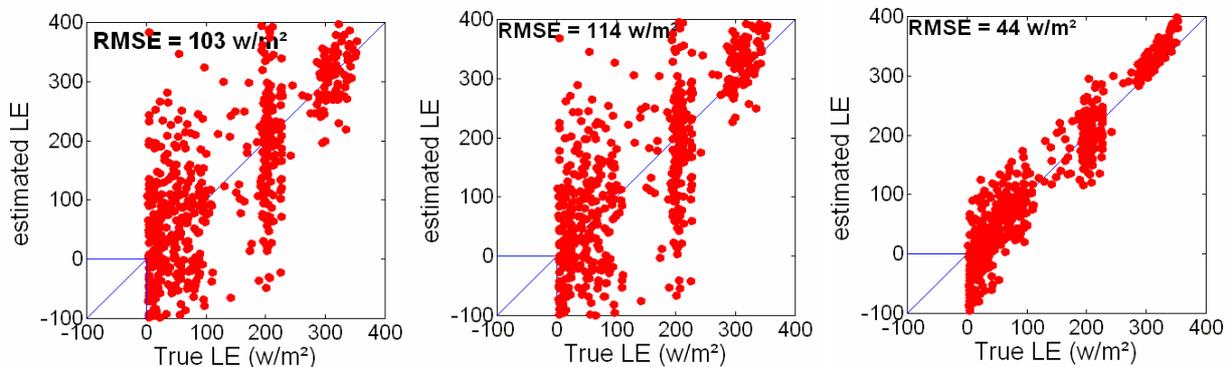
L'ensemble des résultats ci-dessus montre que lorsque la température de surface est peu sensible aux erreurs introduites dans le modèle, les paramètres recherchés, et les flux, sont correctement estimés. Par contre lorsque la température de surface est fortement affectée par les erreurs introduites, les paramètres sont en général mal estimés, voire dégradés par rapports à l'information a priori (IAP). Les paramètres sont dans ce cas « corrigés » afin de compenser les erreurs sur les autres entrées du modèle pour ainsi rapprocher les températures simulées des observations. Leur estimation est donc biaisée (ce biais s'appelle le biais d'omission). Ce biais peut être très fort pour  $z0$  (algorithme C) qui a un lien fort avec la température de surface. Pour autant, les variables de sorties peuvent être améliorées. En particulier, les variables fortement liées à  $T_s$  peuvent être sensiblement améliorées du fait des améliorations de  $T_s$ , alors que ce n'est pas toujours le cas des autres variables. C'est particulièrement visible dans le cas où plusieurs erreurs se combinent :  $H$  et  $T_a$  sont nettement améliorées par rapport aux simulations a priori ; c'est moins net pour  $LE$  et pas du tout le cas pour les autres flux. Les améliorations de

$H$  et  $T_a$  sont d'ailleurs bien plus importantes pour les températures avec la meilleure précision (alors qu'aucune différence n'apparaît pour les autres variables). En ce qui concerne le rayonnement net et le flux de chaleur dans le sol, les impacts de l'assimilation sont en général faibles car la température de surface ne joue qu'un rôle mineur dans leur évolution, ce qui ne permet pas de les corriger en cas de perturbation (cas de l'albédo et de  $c_0$ ). Des améliorations simultanées des flux turbulents sont par contre constatées dans beaucoup de cas correspondant à une amélioration simultanée de la température de surface et des paramètres ré-estimés (ex : cas avec erreurs sur le  $LAI$ ,  $u_a$ , ou  $c_0$ ). Reste le cas particulier de l'algorithme B avec introduction d'erreurs sur  $z_0$  pour lequel l'assimilation n'apporte pas d'amélioration nette alors même que la précision sur la température observée augmente. Cela doit pouvoir s'expliquer par des compensations directes entre la température de surface et la rugosité dans le calcul des flux turbulents, ne permettant pas de transmettre d'information à la détermination de  $f_2$  qui reste assez mal estimée dans ce cas. De la même façon, si la température de l'air semble pouvoir être améliorée par l'assimilation dans un certain nombre de cas, elle ne l'est pas ou peu lorsque les perturbations introduites concernent la vitesse du vent ou la rugosité, des paramètres qui affectent directement les profils atmosphériques dont dépend  $T_a$ . Dans les autres cas,  $T_a$  est directement reliée à  $T_s$  et profite donc directement de ses améliorations.

Un dernier point à noter concerne la prise en compte de  $z_0$  dans la procédure d'assimilation soit comme paramètre à rechercher (algorithme C), soit comme entrée externe avec les erreurs associées (algorithme B). Dans les résultats présentés, l'information a priori concernant  $z_0$  introduite dans l'algorithme C est similaire à l'information introduite dans l'algorithme B (les  $RMSE$  du cas C des simulations a priori sont identiques à ceux du cas B avec  $z_0$  prescrit avec erreur). Le scénario avec ajout d'erreur sur la rugosité  $z_0$  dans l'analyse de l'algorithme B (Tab. V.2.g) présente des erreurs relativement similaires au scénario sans erreur algorithme C (Table V.2.a) avec cependant des améliorations légèrement supérieures des résultats (en particulier pour les flux  $H$  et  $LE$ ) pour l'algorithme C avec la meilleure précision de température de surface.

### 1.1.3. Comparaison méthode résiduelle – assimilation

La Figure V.5 et le Tableau V.3 présentent les résultats obtenus au moyen de la méthode résiduelle sur le jeu de données synthétiques utilisé pour tester les méthodes d'assimilation. Rappelons (cf. **Partie II section 5**) que cette méthode est appliquée en utilisant les mêmes entrées, du moins en ce qui concerne celles qui sont requises, que PBLs dans la configuration de l'algorithme B ( $T_{sobs}$ , rugosité,  $LAI$ , albédo, vitesse du vent, rapport  $G/R_n$ , rayonnements incidents). La température de l'air utilisée correspond à celle qui est simulée par PBLs à la hauteur de la couche limite de surface (dans l'état vrai, ce qui donne des conditions d'utilisation optimiste de la méthode résiduelle). Les résultats de la méthode résiduelle sont évalués par rapport aux flux de chaleur latente vrai et présenté sur la Figure V.5.a. Dans un second temps, l'analyse de l'impact des erreurs dans les différentes entrées (albédo,  $LAI$ , vitesse du vent, rapport  $G/R_n$ , rugosité aérodynamique  $z_0$ ) est effectuée et présenté sur les Figures V.5.b et V.5.c.



**Fig. V.5. Modèle résiduel:** a)  $LE$  estimé à 12h avec une erreur d'observation de 1,5 K ; b)  $LE$  estimé à 12h avec une erreur d'observation de 1,5 K et des erreurs sur  $u_a$ ,  $\alpha$ , et  $LAI$ ; c)  $LE$  à 12h estimé avec une erreur d'observation de 0,5 K et des erreurs sur  $u_a$ ,  $\alpha$ , et  $LAI$

Les premiers résultats obtenus sont illustrés par une comparaison des dispersions obtenues pour des erreurs d'observation  $\sigma_i$  de 1,5 K sur les Figures V.5.a (méthode résiduelle avec un  $RMSE$  de  $103 \text{ Wm}^{-2}$ ) et V.2.b (algorithme B avec un  $RMSE$  de  $34 \text{ Wm}^{-2}$ ). La Figure V.5 montre également que la méthode résiduelle, telle qu'elle a été utilisée ici, conduit à un nombre

important de valeurs négatives de  $LE$ . Cela se produit lorsque  $H$  est supérieur à  $R_n - G$  en lien avec la précision des températures de surface observées. Par la suite, une méthode simple permettant de limiter l'impact de ce problème a été utilisée consistant à considérer des valeurs nulles de  $LE$  dans ce cas. Les résultats dans le Tableau V.3 présentent les  $RMSE$  originaux et les  $RMSE$  tenant compte de ces modifications.

L'analyse d'erreur effectuée en tenant compte des erreurs sur l'albédo ( $\alpha$ ), l'indice foliaire  $LAI$ , la vitesse du vent  $u_a$ , le rapport  $G/R_n$  ( $c_0$ ) et la rugosité aérodynamique  $z_0$  est illustrée par les Figures V.5.b et V.5.c. La première de ces figures présente le flux  $LE$  estimé utilisant le modèle résiduel avec une erreur d'observation de 1,5 K dans la configuration d'une analyse d'erreur considérant les erreurs sur la vitesse du vent, l'albédo et le  $LAI$ . La plupart de la dispersion supplémentaire par rapport au cas sans erreur (Figure V.5.a) est due à l'erreur ajoutée sur la valeur de  $LAI$ . La figure V.5c illustre le flux  $LE$  déterminé dans des conditions similaires mais avec une erreur d'observation de 0,5 K qui explique la très forte réduction de la dispersion des résultats.

Le tableau V.3 présente l'ensemble des résultats obtenus en utilisant le modèle résiduel dans les différents scénarios et pour les erreurs d'observation fixée à 1,5 K ou 0,5 K et l'ajout de différentes erreurs sur les autres paramètres. Ces résultats obtenus en tenant compte des erreurs sur les observations de température de surface montrent dans tous les cas des  $RMSE$  (calculés par rapport aux flux vrais) supérieurs, souvent nettement, à ceux obtenus dans les exercices d'assimilation. Seuls les résultats calculés avec une erreur d'observation de 0,5 K présente des résultats presque équivalents entre les deux méthodes.

En conclusion, il apparaît une très nette supériorité de la méthode d'assimilation par rapport à la méthode résiduelle, en particulier lorsque la précision des mesures de température est limitée. Cela est en grande partie lié à la capacité de la méthode d'assimilation de réduire l'impact des erreurs des observations (comme le montre les baisses de  $RMSE$  sur la température de surface obtenues dans le Tableau V.2) alors que la méthode résiduelle reste très sensible aux erreurs sur la mesure de température de surface.

**Tab. V.3: RMSE sur LE estimés par le modèle résiduel comparé à l'algorithme d'assimilation à 12 h. Entre parenthèse les résultats obtenus avec l'erreur d'observation de 0,5 K. Dans le cas avec l'erreur sur  $z_0$  inclus, ce résultat peut être comparé à l'algorithme C car les deux paramètres ( $z_0$  et  $f_2$ ) sont estimés.**

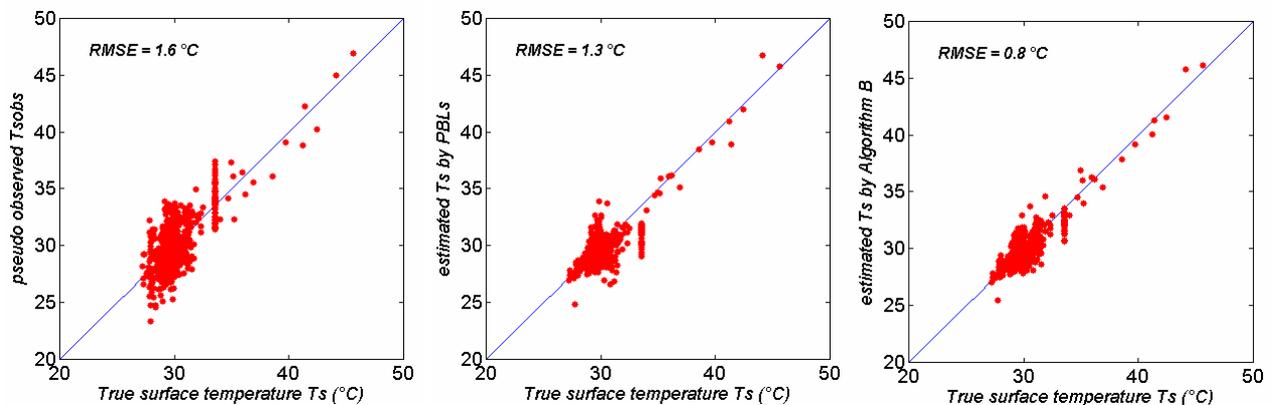
		Modèle résiduel	Algo B	Algo C
$\sigma = 1,5 \text{ K}$ <b>(0,5 K)</b>		103 <b>(33)</b>	34 <b>(21)</b>	x
	$LE < 0 \Rightarrow LE = 0$	78 <b>(27)</b>		
Ajout d'erreur sur $u_a = 0,8, \alpha, LAI$		114 <b>(44)</b>	53 <b>(36)</b>	x
	$LE < 0 \Rightarrow LE = 0$	80 <b>(39)</b>		
Ajout d'erreur sur $u_a, \alpha, LAI, c_0, z_0$		129 <b>(78)</b>	53 <b>(48)</b>	54 <b>(49)</b>
	$LE < 0 \Rightarrow LE = 0$	83 <b>(55)</b>		

## 1.2. Résultats de l'algorithme d'assimilation avec la version 2 de PBLs

Nous rappelons ici que la version 2 du modèle PBLs introduit deux fonctions permettant de modéliser le rapport  $G/R_n$  et la rugosité thermique  $z0_n$  (ou le paramètre  $kB^{-1}$ ) en fonction des caractéristiques de la surface et de l'écoulement atmosphérique, alors qu'ils étaient considérés comme des constantes dans la première version du modèle. Cette deuxième version de PBLs devrait fournir des résultats plus réalistes en particulier en ce qui concerne les flux de chaleur ( $H$ ,  $LE$ ,  $G$ ) et la représentation de leur variabilité spatiale au sein du paysage.

Nous avons réalisé des analyses similaires à celles qui ont été entreprises sur la première version du modèle (détaillées dans la section précédente). A partir des pseudo-observations de température de surface, nous avons recherché les états vrais des paramètres  $z0_m$  et  $f_2$  et des flux en considérant différentes hypothèses sur la qualité des observations et sur l'introduction des erreurs modèles dans l'analyse. Les trois algorithmes A, B et C ont été analysés aux deux niveaux d'erreur d'observation  $\sigma_i$  égale à 1,5 K et 0,5 K.

Nous avons illustré par la figure V.6 ci-dessous la correction de l'erreur d'observation par l'algorithme d'assimilation en présentant les pseudo-observations de température de surface de départ qui correspondent à l'état vrai bruité (Fig. V.6a), les températures de surface simulées par PBLs à partir des informations a priori (Fig. V.6b), et l'estimation de l'algorithme B (Fig. V.6c). Cette dernière montre une amélioration nette de la simulation. Les résultats des différents algorithmes et les améliorations apportées sont détaillés dans les sections ci-dessous.



**Fig. V.6. Comparaisons, aux températures vraies, des températures de surface au cours des différentes étapes de l'assimilation. a) Température de surface  $T_s$  pseudo observée avec  $\sigma_i = 1,5$  K ; b)  $T_s$  simulée par PBLs à partir des informations a priori ; c)  $T_s$  estimée par l'algorithme B avec erreur d'observation de 1,5 K**

### 1.2.1. Cas sans erreur modèle

Les résultats des algorithmes A (recherche de  $z0_m$  seul), B (recherche de  $f_2$  seul) et C (recherche simultanée de  $z0_m$  et de  $f_2$ ) sans erreurs modèles sont présentés dans le Tableau V.4. Tout d'abord, il est intéressant de remarquer que les simulations à partir des données a priori de  $z0_m$  et de  $f_2$  présentent en règle générale des écarts aux données vraies plus grands que les simulations équivalentes avec PBLs1 (sauf pour les flux de chaleur  $H$  et  $LE$  dans la situation de l'algorithme A). Cela traduit d'une part des sensibilités différentes des deux versions du modèle aux deux paramètres rugosité et humidité normalisée en raison de la modification des équations du modèle (on a vu par exemple dans la partie III comment la sensibilité du flux  $H$  à la rugosité aérodynamique pouvait être modifiée par l'introduction de paramétrisations différentes de  $kB^{-1}$ ). D'autre part, nous avons également vu que le niveau même des simulations de PBLs était modifié (eg. Figure III.26), ce qui peut également affecter les écarts par rapport aux états vrais. Une des modifications majeures concerne les simulations de température de surface qui sont fortement affectées par le bruit ajouté aux humidités du sol et qui affectent en conséquence les flux d'énergie et la température de l'air. Cette modification de la sensibilité de  $T_s$  est directement liée à l'introduction de la paramétrisation de  $z0_h$  qui affecte l'ensemble des calculs des transferts turbulents en raison d'une forte modification du calcul des résistances aérodynamiques. Par rapport aux cas de la première version du modèle, la dégradation des simulations a priori illustre une diminution de l'information apportée par l'information a priori introduite dans les algorithmes d'assimilation. Elle illustre aussi une modification de la sensibilité du modèle à ces paramètres (la sensibilité à  $z0_m$  diminue ; la sensibilité à  $f_2$  augmente).

L'examen des résultats des trois algorithmes d'assimilation apporte des conclusions souvent similaires au cas présenté plus haut pour la version 1 de PBLs, mais il apparaît des différences notables que nous détaillerons ci-dessous (cf. Tableau V.4 en comparaison au Tableau V.1).

- Il apparaît immédiatement que les performances obtenues avec la deuxième version du modèle sont bien meilleures pour les flux turbulents  $H$  et  $LE$  et la température de l'air qu'avec la première version (par exemple on obtient un  $RMSE$  de  $13 \text{ Wm}^{-2}$  pour  $LE$  avec l'algorithme B et  $\sigma_i = 0,5 \text{ K}$ ). C'est l'inverse pour  $R_n$  et  $G$ , et surtout pour  $T_s$ , en particulier pour  $\sigma_i = 1,5 \text{ K}$ .

- Pour les algorithmes B et C, les améliorations apportées par les procédures d'assimilation aux estimations des flux et des températures sont bien meilleures que dans le cas de la première version du modèle (ou proches). D'une part comme on vient de le voir, les  $RMSE$  pour les valeurs a priori sont en général supérieurs pour PBLs2 par rapport à PBLs1, et d'autre part après assimilation les  $RMSE$  obtenus avec PBLs2 sont inférieurs ou proches. Cette supériorité de la deuxième version du modèle est bien mise en évidence dans le tableau V.5 qui montre le pourcentage d'amélioration apporté par l'assimilation (en comparant aux simulations a priori).

De son côté, l'algorithme A ne permet de meilleures améliorations que pour  $R_n$ ,  $H$  et  $T_s$  trois variables fortement reliées entre elles (et à la rugosité pour  $H$  et  $T_s$ ).

- Il existe maintenant une claire hiérarchie des résultats en fonction des différents algorithmes et de la précision des données. L'algorithme A donne de meilleurs résultats que B pour l'ensemble des variables (sauf  $T_a$ ). L'algorithme B donne de meilleurs résultats que C (sauf pour  $T_s$  qui ont des niveaux équivalents).

Les résultats obtenus avec la précision des températures observés la plus forte sont toujours bien meilleurs qu'avec les observations moins précises. On notera en particulier que contrairement à PBLs1, les résultats obtenus par l'algorithme C avec  $\sigma_i = 0,5 \text{ K}$  sont nettement améliorés par rapport aux résultats de l'algorithme C avec  $\sigma_i = 1,5 \text{ K}$ .

- Concernant l'estimation des paramètres recherchés, les résultats obtenus avec la deuxième version de PBLs montrent d'une part des améliorations très sensibles des valeurs de l'humidité normalisée  $f_2$  par rapport à la première version, et d'autre part des estimations de moins bonne qualité de la rugosité aérodynamique  $z_{0m}$ . Cette dernière n'est d'ailleurs pas améliorée par l'algorithme C.

L'ensemble des résultats décrits ci-dessus montre que les modifications apportées au modèle PBLs jouent un rôle considérable dans le processus d'assimilation. Ce qui s'explique par les modifications de la sensibilité du modèle aux deux paramètres  $z0_m$  et  $f_2$ .

En particulier, le rôle de la rugosité mécanique  $z0_m$  est profondément modifié via la paramétrisation de la rugosité thermique  $z0_h$  qui entraîne une baisse de la sensibilité du modèle à  $z0_m$ . Cette rugosité devient plus difficile à retrouver par assimilation (il ne l'est plus du tout dans le cas de l'algorithme C). Cela doit nous conduire par la suite à abandonner les deux algorithmes A et C qui n'apportent pas d'informations supplémentaires par rapport à l'algorithme B dans lequel seule l'humidité normalisée doit être retrouvée.

De manière opposée, la sensibilité à l'humidité normalisée  $f_2$  a augmenté ce qui se traduit par une meilleure capacité de la température de surface à contrôler ce paramètre. Ce qui permet en particulier d'améliorer la restitution des flux turbulents.

Cela se traduit également par un renforcement de l'intérêt de la méthode d'assimilation qui devient performante pour l'ensemble des niveaux d'erreurs d'observation considérés ici (1,5 K et 0,5 K).

**Tab. V.4. RMSE entre les résultats de l'assimilation avec l'erreur d'observation et les états vraies (défini à l'étape 1) pour les flux d'énergie (à midi), la température de surface et la température de l'air à 2 m (à midi) et les paramètres de  $z0_m$  et  $f_2$  retrouvés.**

	PBLs2 (sans Assimilation)			Assimilation $\sigma_i = 1,5$ K			Assimilation $\sigma_i = 0,5$ K		
	A	B	C	A	B	C	A	B	C
$R_n$ ( $Wm^{-2}$ )	4	8	8	3	5	5	2,3	2,5	2,9
$H$ ( $Wm^{-2}$ )	6	39	41	5	24	27	4	11	19
$LE$ ( $Wm^{-2}$ )	3,5	45	46	3,1	28	31	2	13	20
$G$ ( $Wm^{-2}$ )	0,8	1,6	1,7	0,7	1,1	1,2	0,52	0,55	0,64
$T_s$ (K)	0,70	1,30	1,31	0,55	0,82	0,93	0,37	0,40	0,46
$T_a$ (K)	0,43	0,42	0,60	0,41	0,20	0,40	0,30	0,13	0,34
$z0$ (m)	0,100	x	0,100	0,097	X	0,100	0,072	x	0,100
$f_2$ (-)	x	0,140	0,140	x	0,107	0,108	x	0,060	0,079

**Tab. V.5. Pourcentage d'amélioration des estimations par la méthode d'assimilation avec l'erreur d'observation de 1,5 K par rapport aux simulations du modèle PBLs avec les valeurs a priori**

	PBLs version 1 (%)			PBLs version 2 (%)		
	A	B	C	A	B	C
$R_n$	14	23	23	25	38	38
$H$	13	14	8	17	39	34
$LE$	20	13	5	11	38	33
$G$	17	20	23	13	31	29
$T_s$	14	17	13	29	39	31
$T_a$	10	33	5	0	52	33
$ZO_m$	10	x	6	3	x	0
$f_2$	x	14	11	x	24	23

### 1.2.2. Cas avec erreur modèle

Une analyse de l'impact de l'erreur modèle a été réalisée comme pour la première version du modèle en ajoutant des erreurs aux paramètres d'entrée qui ne sont pas retrouvés par l'assimilation. Nous avons rajouté les erreurs de mesures liées à l'albédo, au *LAI*, à la vitesse moyenne journalière du vent et à la rugosité (dans le cas de l'algorithme B seulement pour cette dernière). Les impacts des erreurs sont présentés dans le tableau V.6 en fonction des deux niveaux d'erreur d'observation et des deux algorithmes B et C.

Les résultats liés aux erreurs sur l'albédo et la vitesse du vent sont très peu différents de ceux qui ont été obtenus avec PBLs1. Ils ne sont pas présentés.

Le premier résultat apparent concerne comme dans le cas sans erreur l'augmentation générale des *RMSE* par rapport à PBLs1 pour l'ensemble des simulations présentées sans assimilation : colonnes 3 et 4 du tableau correspondant à l'analyse de sensibilité du modèle par rapport à l'état vrai et aux états a priori générés pour les deux algorithmes B et C. Comme pour PBLs1, les erreurs sur le *LAI* génèrent de fortes erreurs sur les résultats du modèle. Notons que par rapport à la première version de PBLs, le *LAI* a ici été introduit dans de nouveaux processus par l'intermédiaire des paramétrisations nouvelles introduites dans PBLs2. Ceci augmente son impact sur les simulations des flux de chaleur et des températures.

Dans les cas avec assimilation on rencontre deux comportements en fonction de la proximité des résultats au cas sans erreur.

- Pour la rugosité, le comportement est proche du cas sans erreur avec une amélioration importante des résultats, d'une part par rapport aux états a priori, et d'autre part par rapport à PBLs1.
- Dans le cas du *LAI* l'assimilation apporte des améliorations sensibles, en particulier pour l'algorithme B (0,5 K). Les améliorations sont cependant moins nettes que dans le cas sans erreur, en particulier pour l'algorithme C. Elles sont aussi moins bonnes que celles qui avaient été obtenues pour PBLs1. Dans le cas de l'algorithme C, les résultats obtenus sont similaires pour les deux niveaux d'erreur de températures observées malgré les fortes différences obtenues sur les températures simulées (ce qui diffère du cas sans erreur et de PBLs1). Ils sont également similaires aux résultats obtenus pour l'algorithme B avec l'erreur 1,5 K.

- Dans le cas où toutes les erreurs sont combinées, les améliorations par l'assimilation sont encore moins nettes que pour le cas du *LAI*. Les deux algorithmes fournissent des résultats très proches quelque soit le niveau des erreurs d'observation. Les niveaux d'erreur après assimilation sont importants, par exemple de l'ordre de 60-65  $Wm^{-2}$  pour les flux de chaleur turbulents (à comparer aux 71-73  $Wm^{-2}$  des simulations a priori).

Lors de l'assimilation dans le modèle PBLs1, de fortes corrections des températures de surface simulées amenaient de fortes corrections des variables auxquelles elles étaient directement liées (en particulier  $H$  et  $T_a$ ). Ce comportement est moins évident ici, particulièrement parce que les erreurs sur les simulations de température de surface sont bien supérieures.

Concernant les paramètres retrouvés, comme dans le cas de PBLs1, la rugosité est dégradée par rapport aux états a priori. Par contre, des améliorations sensibles de l'humidité normalisée  $f_2$  sont constatées.

Les enseignements principaux à tirer de cet exercice concernent

- le trop faible apport d'information par les observations pour contrôler l'ensemble des erreurs introduites dans le modèle
- le rôle important du *LAI* dans la dégradation des résultats
- l'impossibilité de retrouver la rugosité
- le peu de différence existant entre les résultats des algorithmes B et C
- l'apport assez faible de l'augmentation de précision des observations

Concluons enfin que les modifications apportées au modèle PBLs ont affecté significativement le fonctionnement du modèle, aussi bien en ce qui concerne les simulations directes comme on a pu le montrer dans le chapitre III de la thèse ou dans les analyses de sensibilité ci-dessus, que dans son fonctionnement inverse dans les procédures d'assimilation.

**Tab. V.6. RMSE à 12h entre les états vrais et les sorties de l'algorithme d'assimilation comparées aux résultats de PBLs (sans assimilation : colonne 4) avec ajout d'erreur sur l'albédo ( $\alpha$ ), sur LAI, sur la vitesse du vent ( $u_a$ ), et sur la rugosité  $z_0$  par l'algorithme de B. Colonne 1 (ajout d'erreur) : erreur moyenne ajoutée par rapport à l'état vrai sur l'ensemble des 625 pixels. Colonne 3: (analyse de sensibilité) montrant l'impact de chaque variable avec l'erreur sur les résultats de PBLs par rapport au scénario vrai. Colonnes 7 et 8: Résultats de l'algorithme C montrant la similitude avec l'algorithme B**

ajout erreur		Analyse de sensibilité	PBLs		Algo B	Algo B	Algo C	Algo C
			B	C	$\sigma = 1,5 \text{ K}$	$\sigma = 0,5 \text{ K}$	$\sigma = 1,5 \text{ K}$	$\sigma = 0,5 \text{ K}$
(a) sans erreur	$R_n \text{ (Wm}^{-2}\text{)}$	-	8	8	5	2,5	5	2,9
	$H \text{ (Wm}^{-2}\text{)}$	-	39	41	24	11	27	19
	$LE \text{ (Wm}^{-2}\text{)}$	-	45	46	28	13	31	20
	$G \text{ (Wm}^{-2}\text{)}$	-	1,6	1,7	1,1	0,5	1,2	0,6
	$T_s \text{ (}^\circ\text{C)}$	-	1,30	1,31	0,82	0,40	0,93	0,46
	$T_a \text{ (}^\circ\text{C)}$	-	0,42	0,60	0,20	0,13	0,40	0,34
	$z_0 \text{ (m)}$	-	-	0,100	-	-	0,100	0,100
	$f_2 \text{ (-)}$	-	0,140	0,140	0,107	0,060	0,108	0,079
(b) LAI $\pm 44\%$	$R_n \text{ (Wm}^{-2}\text{)}$	7	10	10	7	5	7	5
	$H \text{ (Wm}^{-2}\text{)}$	49	64	65	52	45	55	54
	$LE \text{ (Wm}^{-2}\text{)}$	52	71	71	57	48	59	58
	$G \text{ (Wm}^{-2}\text{)}$	16	16	16	15	14	16	16
	$T_s \text{ (}^\circ\text{C)}$	1,10	1,52	1,60	1,22	0,87	1,16	0,92
	$T_a \text{ (}^\circ\text{C)}$	0,40	0,54	0,64	0,43	0,32	0,54	0,45
	$z_0 \text{ (m)}$	-	-	0,100	-	-	0,112	0,105
	$f_2 \text{ (-)}$	-	0,140	0,140	0,112	0,110	0,114	0,099
(c) $z_0$ $\pm 25\%$	$R_n \text{ (Wm}^{-2}\text{)}$	4	8	a)	5,7	3	a)	a)
	$H \text{ (Wm}^{-2}\text{)}$	6	41	a)	28	20	a)	a)
	$LE \text{ (Wm}^{-2}\text{)}$	3,4	46	a)	31	21	a)	a)

	$G$ ( $Wm^{-2}$ )	0,8	1,7	a)	1,2	0,6	a)	a)
	$T_s$ ( $^{\circ}C$ )	0,60	1,31	a)	0,91	0,47	a)	a)
	$T_a$ ( $^{\circ}C$ )	0,43	0,60	a)	0,45	0,36	a)	a)
	$z0$ (m)	-	-	a)	-	-	a)	a)
	$f2$ (-)	-	0,140	a)	0,109	0,088	a)	a)
(d) toutes erreurs $\alpha$ , LAI $u_a = 0,8$	$Rn$ ( $Wm^{-2}$ )	15	15	16	14	13	14	12
	$H$ ( $Wm^{-2}$ )	56	71	72	62	61	65	64
	$LE$ ( $Wm^{-2}$ )	55	73	73	63	62	65	64
	$G$ ( $Wm^{-2}$ )	18	18	18	17	17	17	17
	$T_s$ ( $^{\circ}C$ )	1,62	1,90	2,0	1,61	1,28	1,43	0,67
	$T_a$ ( $^{\circ}C$ )	0,51	0,80	0,81	0,58	0,48	0,64	0,50
	$z0$ (m)	-	-	0,100	-	-	0,115	0,120
	$f2$ (-)	-	0,140	0,140	0,111	0,113	0,113	0,104

### 1.2.3. Comparaison méthode résiduelle – assimilation

La comparaison entre le modèle résiduel et l’algorithme d’assimilation B est présentée dans le Tableau V.7. Les améliorations apportées au modèle PBLs concernant le rapport  $G/R_n$  et la paramétrisation de la rugosité thermique  $z_{0h}$  par l’intermédiaire du paramètre  $kB^{-1}$  ont également été intégrées au calcul résiduel du flux de chaleur latente. Ce modèle résiduel modifié présente une amélioration des performances très nette par rapport à la version originelle (présentée dans le Tableau V.3).

La comparaison avec l’algorithme B montre toujours que la méthode d’assimilation apporte généralement une meilleure réponse dans l’estimation de l’évapotranspiration (par exemple, avec une erreur d’observation de 1,5 K et pas d’erreur modèle, le *RMSE* sur le flux *LE* est de  $28 \text{ Wm}^{-2}$  avec l’algorithme B contre  $54 \text{ Wm}^{-2}$  avec le modèle résiduel). Cependant, l’écart s’est réduit par rapport au cas précédent, d’une part parce que le modèle résiduel est plus performant, et d’autre part parce que l’algorithme d’assimilation est moins performant (dans le cas avec erreurs). Comme dans le cas précédent, les estimations par le modèle résiduel se rapprochent des résultats de l’assimilation quand la température observée est plus précise (une erreur équivalente autour de  $60 \text{ Wm}^{-2}$  est obtenue dans le cas avec erreur). Nous pouvons en déduire dans cette configuration que la principale source d’erreur est plutôt liée aux paramètres d’entrée qu’au forçage de la température de surface.

**Tab. V.7: *RMSE* sur *LE* estimés par le modèle résiduel comparé à l’algorithme d’assimilation à 12 h. Entre parenthèse les résultats obtenus avec l’erreur d’observation de 0.5 °C**

	Residual Method	Algorithm B
$\sigma = 1.5 \text{ °C (0.5 °C)}$		54 (19)
	$LE < 0 \Rightarrow LE = 0$	44 (16)
adding error on		103 (85)
$u_a = 0.8, \alpha, LAI$	$LE < 0 \Rightarrow LE = 0$	67 (59)

### 1.3. Conclusion sur les données synthétiques

Les expériences synthétiques ont permis de mettre en place et d'analyser les performances des algorithmes d'assimilation variationnelle sur le modèle PBLs et une évolution de ce dernier (PBLs2). Les différentes analyses nous ont permis de mieux comprendre le fonctionnement de la méthode d'assimilation suivant la qualité des observations de température de surface et des informations a priori (humidité normalisée, rugosité) et suivant la qualité du modèle (soit en fonction des erreurs introduites sur les entrées du modèle non recherchées par l'assimilation comme le *LAI*, l'albédo ou la vitesse du vent ; soit en analysant comparativement deux versions différant par la paramétrisation de certains paramètres importants du modèle).

La méthode d'assimilation a été évaluée vis à vis d'un état simulé de référence, ou d'un état vrai, après avoir bruité les différentes informations en entrée (observations, paramètres). Les avantages d'utiliser des données synthétiques sont de permettre l'accès à une exhaustivité des données d'évaluation et la maîtrise des erreurs affectant les différentes entrées du modèle et de l'algorithme (observation, données a priori, autres paramètres). Le jeu de données synthétiques générées a également permis d'évaluer la méthode d'assimilation par rapport aux résultats d'une méthode plus classique consistant en un calcul résiduel du flux de chaleur latente.

Signalons tout d'abord que les niveaux de *RMSE* obtenus dans les différents tests sont tout à fait conformes aux niveaux constatés dans l'ensemble de la littérature portant sur l'évaluation de méthodes d'estimation des flux du bilan d'énergie de surface ou de cartographie de l'évapotranspiration : selon les situations expérimentales, ces niveaux se situent entre quelques dizaines de  $Wm^{-2}$  et un peu plus d'une centaine. Les niveaux d'erreurs en entrée des analyses ayant été choisis de manière à être eux aussi représentatifs des niveaux constatés dans les données réelles, les analyses synthétiques réalisées apparaissent tout à fait cohérentes par rapport à la réalité.

Les exercices d'assimilation de la température de surface dans les deux versions du modèle PBLs ont montré des améliorations souvent assez nettes des résultats, en fonction du niveau d'erreurs des observations, par rapport aux simulations a priori dans le cas optimal d'un modèle sans erreur. Dans le cas d'un modèle avec erreur, nous avons constaté une forte dégradation des résultats d'assimilation, la méthode ayant des difficultés à réduire l'impact de toutes les erreurs introduites. Lorsque toutes les erreurs modèles sont prises en compte, les

améliorations constatées sont tout au plus d'une quinzaine de  $Wm^{-2}$  sur les flux turbulents et les différences entre les niveaux d'erreur d'observation sont faibles. La réduction du niveau d'erreur sur les sorties du modèle s'est en général accompagnée, par mécanisme de compensation, d'une dégradation de l'estimation des paramètres à retrouver ( $z0_m$  et/ou  $f_2$ ) par rapport aux cas sans erreurs. Dans certaines situations la longueur de rugosité aérodynamique  $z0_m$  a même été dégradée par rapport à sa valeur a priori.

La comparaison à la méthode résiduelle montre, en général, une supériorité des algorithmes d'assimilation, en particulier lorsque la précision des mesures de température est limitée. Cela est en grande partie lié à la capacité de la méthode d'assimilation, grâce à l'introduction de l'information a priori, de réduire l'impact des erreurs des observations alors que la méthode résiduelle reste très sensible aux erreurs sur la mesure de température de surface. Finalement dans les cas traités ici, l'écart entre la méthode résiduelle et l'assimilation diminue lorsque la précision des observations augmente. Signalons cependant ici que l'application de la méthode résiduelle est faite dans des conditions assez optimistes car nous n'avons pas tenu compte des erreurs possibles, et même très probables, sur la connaissance de la température de l'air  $T$  utilisée en forçage à la hauteur de la couche limite de surface ( $h_{CLS}$ ) puisque celles ci ont été fixées aux valeurs simulées par PBLs. Dans la méthode résiduelle, la prise en compte de cette erreur serait équivalente à augmenter l'erreur sur l'observation  $T_{sobs}$  puisque c'est la différence entre la température de l'air et la température de surface qui est en fin utilisée dans les calculs. Rappelons aussi que l'enjeu des améliorations récentes de l'utilisation de la méthode résiduelles consistait justement à réduire l'erreur sur cette différence de température (cf. les modèles comme SEBAL).

Les résultats ci-dessus nous font pointer du doigt le rôle de l'information a priori et en particulier des composantes de la matrice de covariance de l'erreur ( $B$ ) dans le processus d'assimilation. Une spécification correcte des erreurs de l'observation et de l'information a priori est cruciale pour la qualité de l'estimation, car ils déterminent dans quelle mesure les valeurs a priori ( $IAP$  ou valeurs d'ébauche) seront corrigées pour correspondre aux observations. En donnant un poids aux matrices de covariance d'erreur ( $B$  et  $R_i$ ) plus ou moins important l'une par rapport à l'autre l'algorithme d'assimilation peut être orienté vers les valeurs d'ébauche ou vers l'observation. La définition de l'information a priori ainsi que les erreurs statistiques associées est un des points les plus critiques de la méthodologie proposée.

Elle dépend des connaissances de l'utilisateur sur le domaine d'application et sa capacité à définir des valeurs d'ébauche aussi probable et réaliste.

Dans les tests réalisés dans cette section, les deux versions du modèle PBLs ont été utilisées. La deuxième version (PBLs2) a apporté des modifications importantes des résultats de l'assimilation. Les nouvelles paramétrisations de  $G/R_n$  et  $kB^{-1}$  ont profondément modifié le lien entre la température de surface et les flux de chaleur et la sensibilité du modèle aux paramètres que l'on cherche ( $z0_m$  ou  $f_2$ ). Ainsi, dans toutes les situations, les assimilations réalisées avec PBLs2 ont permis de mieux retrouver l'humidité normalisée  $f_2$ . Par contre, la rugosité aérodynamique, qui est un paramètre moins sensible dans PBLs2 que dans PBLs1, est devenue quasiment impossible à retrouver. Concernant les variables de sortie, si le modèle PBLs2 est apparu plus performant que la première version dans certains exercices d'assimilation, ses résultats étaient moins bons dans les situations incluant les erreurs d'indice foliaire dans le modèle. Cela peut s'expliquer par le rôle important joué par l'indice foliaire dans les nouvelles paramétrisations introduites et qui se traduit par la plus forte sensibilité de PBLs2 au *LAI* que PBLs1.

Concernant les différents algorithmes testés (A, B, C), il est vite apparu que la sensibilité du modèle étant beaucoup plus forte à  $f_2$  qu'à  $z0_m$ , l'algorithme A qui visait à retrouver  $z0_m$  seul n'était pas adapté au problème posé. Par ailleurs, l'algorithme B visant à estimer  $f_2$  seul montre des performances quasiment similaire à l'algorithme C visant à rechercher simultanément  $z0_m$  et  $f_2$  en particulier dans les situations cumulant l'ensemble des erreurs. Enfin, il semble plus intéressant d'utiliser l'algorithme B en incluant les erreurs sur  $z0_m$  comme erreurs internes au modèle, plutôt que de rechercher simultanément  $z0_m$  et  $f_2$ . Lors du choix des paramètres à retrouver (Chapitre I), nous avons justifié la recherche de  $z0_m$  et  $f_2$  par la difficulté à connaître ces paramètres à l'échelle du paysage agricole. Après cette première analyse, il apparaît que si ce choix était bien justifié pour  $f_2$  il l'est moins pour  $z0_m$ . Notons cependant que le résultat obtenu ici dépend du niveau d'erreur supposé pour la rugosité, qui dans d'autres situations expérimentales pourrait être plus élevé, donnant peut être un poids supplémentaire à ce paramètre dans le processus d'assimilation. Nous avons également supposé dans ce travail que le *LAI* pouvait être facilement connu de manière indépendante, justifiant son introduction dans l'erreur modèle plutôt que dans la procédure d'assimilation. La sensibilité importante du modèle à ce paramètre dans la situation analysée fait poser la question de la recherche de ce paramètre par assimilation. Il est probable dans ce cas que la piste à explorer consisterait à

modéliser les signaux de réflectance simultanément au bilan d'énergie et à la température de surface pour finalement réaliser l'assimilation sur l'ensemble des signaux de télédétection, réflectances et température de surface.

L'étude sur les données synthétiques nous a permis une première analyse des performances des algorithmes d'assimilation, confirmant leur intérêt à notre échelle de travail, montrant le rôle important de la spécification des erreurs et des valeurs a priori, et permettant déjà d'effectuer un premier choix sur les algorithmes et les modèles à utiliser dans la pratique. La section suivante va illustrer leur utilisation sur les données, réelles, de l'expérimentation Alpilles.

## 2. Performances sur les données réelles d'Alpilles – ReSeDA

### 2.1. Introduction

Nous avons appliqué les différents schémas de l'algorithme d'assimilation sur les données réelles d'Alpilles ReSeDA. Le modèle utilisé est la deuxième version de PBLs (PBLs2) qui intègre les modifications portant sur la paramétrisation du rapport  $G/R_n$  et du paramètre  $kB^{-1}$ . Comme l'a montrée l'étude sur les données d'expérience jumelle, une erreur d'observation (erreur de mesures + erreur modèle) de 1,5 K est réaliste pour une telle application. Les valeurs d'ébauche (ou valeurs *IAP*) des paramètres ( $z0_m$  et  $f_2$ ) et la matrice de covariance d'erreur de l'information a priori ( $B$ ) sont telles que décrites dans la **section 5.4** de la **partie II** de ce travail (pour rappel, la carte de la rugosité aérodynamique ( $z0_m$ ) est issue de la carte d'occupation des sols et de l'information sur le développement des couverts et la carte de l'humidité du sol normalisée ( $f_2$ ) provient de simulations au moyen du modèle ISBA).

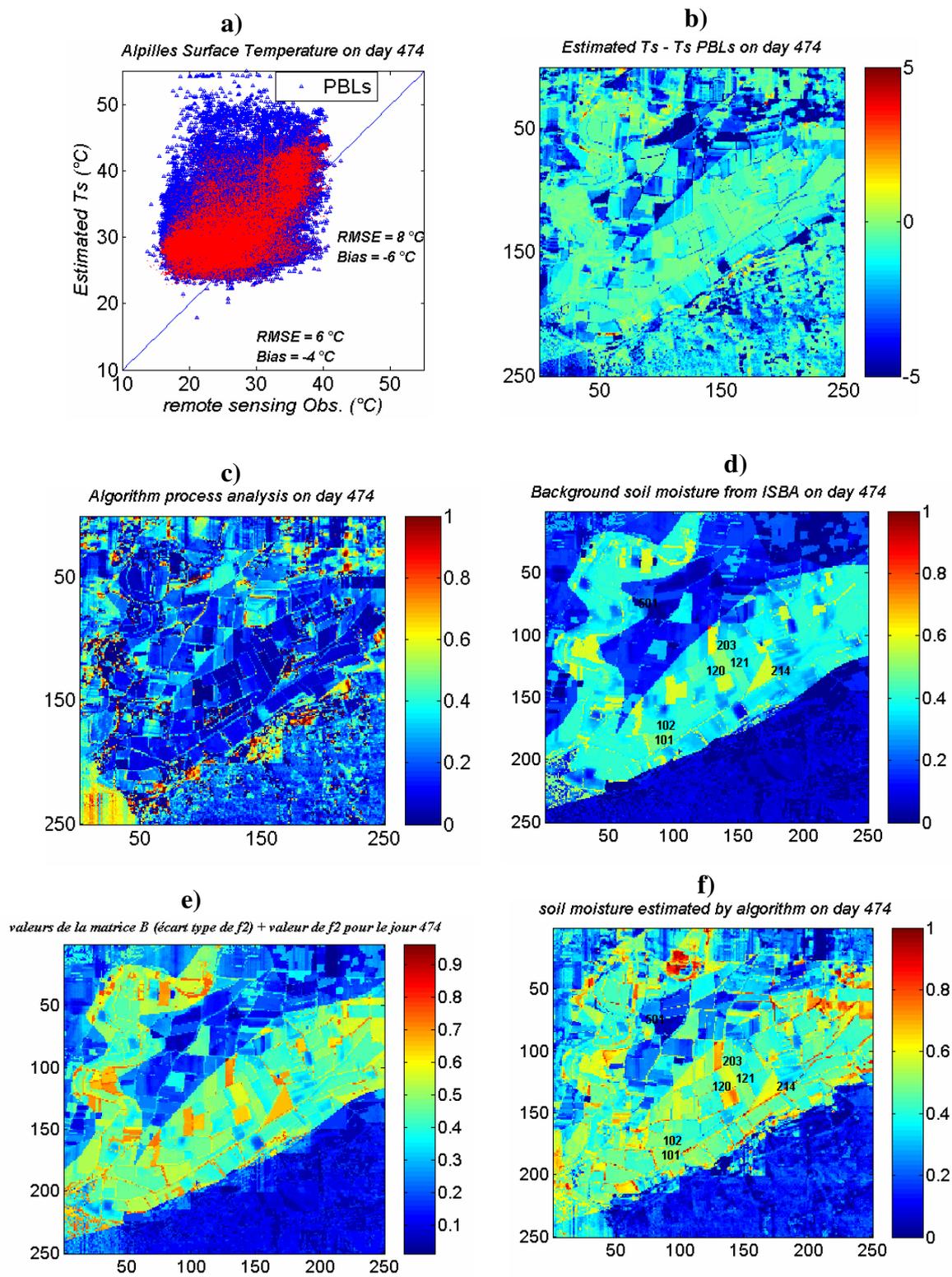
Nous avons assimilé des mesures aéroportées de température de surface pour quelques dates du jeu de données Alpilles. L'application est réalisée sur l'ensemble de la zone de 5 km sur 5 km subdivisée en 62500 pixels (soit 250 x 250 pixels). C'est une opération très coûteuse dans notre cas en raison des moyens informatiques dont nous disposons. Par ailleurs, il ressort aussi de l'analyse sur les données d'expérience jumelle que l'algorithme B (recherche de l'humidité normalisée  $f_2$  seule) apporte une meilleure réponse, ou au moins similaire, pour l'estimation de l'évapotranspiration par rapport aux deux autres schémas. Par conséquent l'essentiel des résultats présentés ci-dessous concerne l'algorithme B. Ce choix est également motivé par le coût de l'application de l'algorithme C qui nécessite plus de temps que l'algorithme B.

Dans cette section les résultats de la méthode d'assimilation sont présentés sous l'hypothèse d'une erreur d'observation de 1,5 K. Dans un premier temps, les résultats de l'algorithme sont analysés par rapport aux observations assimilées afin de déterminer le poids des informations apportées par les observations ou par l'ébauche (= information a priori). Dans un second temps, nous avons validé les estimations sur les mesures des stations météorologiques des 7 parcelles de la zone. Dans un troisième temps, nous avons testé l'impact d'informations a priori générées avec une méthode différente que dans le premier cas.

Ces différentes analyses vont nous permettre de justifier la stratégie de définition de l'information a priori et de valider la méthode ou du moins de la tester sur des données expérimentales réelles. Certes la validation sur les parcelles expérimentales restent partielles (7 parcelles sur plus de 351 parcelles cf. Fig. II.5), mais les résultats de ces assimilations permettent également la comparaison des cartographies des flux de surface avant et après assimilation sur l'ensemble de la zone d'Alpilles.

## 2.2. Analyse de l'assimilation par rapport aux observations

Une première analyse de la performance de l'algorithme d'assimilation sur les données réelles consiste à comparer les températures de surface estimées (après assimilation, notées  $T_s$ ), d'une part aux simulations a priori (ébauche, notées  $T_{sPBLs}$ ) et d'autre part aux observations assimilées ( $T_{sobs}$ ). La Figure V.7 permet d'observer les différences entre ces températures, et en quelque sorte d'observer l'évolution de la température de surface au cours du processus d'assimilation. Elle présente les résultats de l'algorithme B pour la date du 18 avril 1997 (DOE 474). Le premier graphe présente la comparaison des deux températures de surface simulées par PBLs, estimées après assimilation et simulées avec l'information a priori, par rapport aux observations (Fig. V.7a). Le deuxième graphe (Figure V.7b) présente la carte de différence entre la température de surface estimée par l'algorithme d'assimilation et la température simulée avec l'information a priori ( $T_s - T_{sPBLs}$ ). Le troisième graphe (Figure V.7c) montre un rapport ( $\check{R}$ ) permettant d'analyser le poids relatif de l'observation et de l'information a priori dans le processus d'assimilation. Les trois derniers graphes présentent les humidités normalisées  $f_2$  a priori (Figure V.7d), la somme de  $f_2$  a priori et l'écart type de  $f_2$  (matrice  $B$ ) (Figure V.7e), et  $f_2$  après assimilation (Figure V.7f).



**Fig. V.7. Résultats du jour 474 de l'expérimentation Alpilles :** a) Températures de surface estimées par l'algorithme B ( $T_s$ , points rouges) et simulées a priori ( $T_s$  PBLs, étoiles bleues) comparées aux observations aéroportées ; b) Carte de la différence  $T_s$  estimée –  $T_s$  PBLs ; c) Cartographie du rapport  $\check{R}$  permettant un diagnostic du rôle des informations

**introduite dans la procédure d'assimilation** (voir texte et Eq.V.1 ci-dessous) ; d) **information a priori de l'humidité normalisée  $f_2$**  ; e) **somme de l'information a priori de  $f_2$  et de l'écart type de  $f_2$**  (élément de la matrice  $B$ ) ; f) **humidité normalisée  $f_2$  retrouvée par l'assimilation.**

La figure V.7a montre une légère évolution des températures de surfaces simulées de la valeur a priori (bleues) à la valeur estimée par l'algorithme B (rouge). Le résultat de PBLs (avant assimilation) sur l'ensemble de la zone Alpilles pour la date 474 présente un *RMSE* de 8 K contre 6 K après le processus d'assimilation. Nous pouvons déduire de ce résultat que la plupart des estimations sont en fait très liées aux informations a priori choisies. En effet, les estimations par l'algorithme d'assimilation ne présentent pas un grand écart par rapport aux simulations a priori. L'écart entre ces deux températures est illustré sur la figure V.7b pour la zone Alpilles complète. On peut observer que l'écart avoisine zéro dans la majorité des cas. Cela montre que le poids (ou confiance) de la matrice de covariance d'erreur de l'information a priori ( $B$ ) est très important par rapport à celle de l'observation dans le processus d'assimilation. Par conséquent, l'algorithme d'assimilation est plus influencé par l'apport de l'information a priori que par celui de l'observation. Cela pourrait indiquer que le choix des valeurs d'ébauche de l'humidité du sol normalisée ( $f_b$ ) et des erreurs statistiques (matrice  $B$ ) associées n'est pas très adapté. Par ailleurs, on peut observer qu'il existe sur certains pixels des variations importantes entre les deux simulations, ce qui semble indiquer que dans ce cas les observations ont apporté plus d'information dans le processus d'assimilation.

Pour mieux comprendre les résultats, nous avons établi le rapport ( $\check{R}$ ) montrant la position de la température de surface estimée par rapport aux observations d'une part et aux informations a priori d'autre part. Ce diagnostic d'analyse (Fig. V.7c) est défini par l'équation (V.1) ci-dessous :

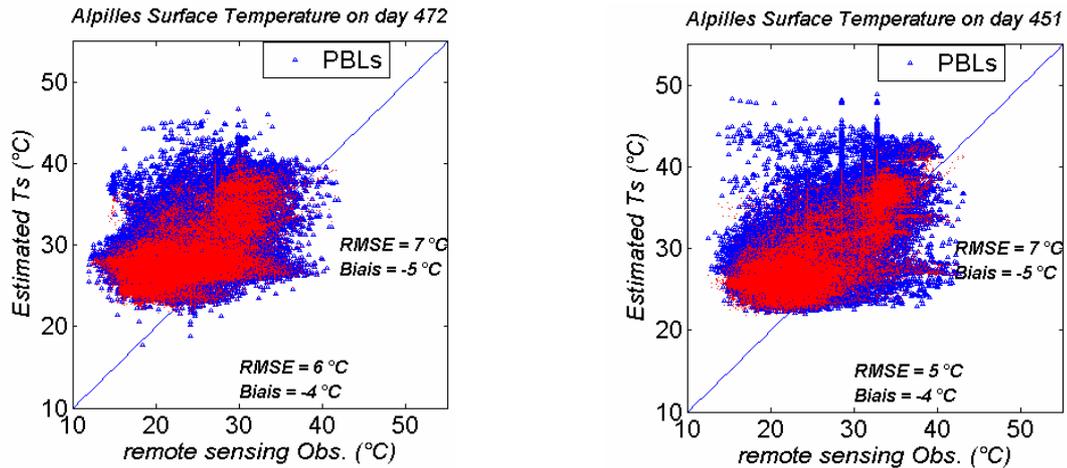
$$\check{R} = \frac{T_s - T_{sPBLs}}{Obs. - T_{sPBLs}} \quad (V.1)$$

Lorsque  $\check{R}$  est proche de 1, la température de surface estimée est proche de l'observation. Quand  $\check{R}$  est proche de 0, l'estimation est proche de la simulation PBLs obtenue avec

l'information a priori. Ce diagnostic nous aide à localiser les pixels dont l'information a priori et l'erreur associée choisies peuvent être inappropriées ou déjà correctes. La Figure V.7c présente la carte du rapport  $\check{R}$  qui est souvent proche de 0, ce qui est en bon accord avec la figure V.7b dans laquelle beaucoup des estimées sont proches de l'ébauche. L'ensemble des résultats montre donc que l'information fournie par les observations reste insuffisante pour apporter plus d'améliorations.

Les Figures V.7d-f montrent comment les humidités normalisées sont modifiées au cours du processus d'assimilation. On note une augmentation générale de l'humidité et une augmentation de la variabilité entre les parcelles (correspondant à une baisse du biais entre les températures de surface simulées et observées sur la Figure V.7a). Cependant, l'augmentation de l'humidité n'a pas permis de corriger complètement ce biais, ce qui montre que les valeurs de  $f_2$  retrouvées sont trop faibles et restent trop proches de l'information a priori. Cela est dû au poids accordé à la matrice de covariance d'erreur  $B$  : la carte des humidités retrouvées est pratiquement similaire à la carte des informations a priori augmentées des valeurs de la matrice  $B$  (Figure V.7e). Plusieurs autres facteurs peuvent participer à l'explication de ce résultat : le fort écart entre les simulations et les températures observées qui indique probablement des erreurs modèle + mesure importantes ; les erreurs prescrites dans l'algorithme (1,5 K) peuvent être trop optimistes ; les erreurs sur les autres paramètres du modèle PBLs ( $LAI$ , vitesse du vent...) dont les impacts ne peuvent pas toujours être corrigés (voir les résultats des expérimentations synthétique dans ce chapitre).

Nous avons réalisé l'assimilation pour deux autres dates, les 16 avril 1997 (DOE 472) et 26 mars 1997 (DOE 451). La figure V.8 présente la comparaison des températures de surface simulées par PBLs avant assimilation (valeurs a priori en bleu) et après assimilation (valeurs estimées par l'algorithme B en rouge) par rapport aux observations (toujours en considérant une erreur d'observation de 1,5 K).



**Fig. V.8. Températures de surface estimées par l’algorithme B ( $T_s$ , points rouges) et simulées avec l’information a priori ( $T_s$  PBLs, étoiles bleues) comparées aux observations aéroportées. a) pour le jour 472 ; b) pour le jour 451.**

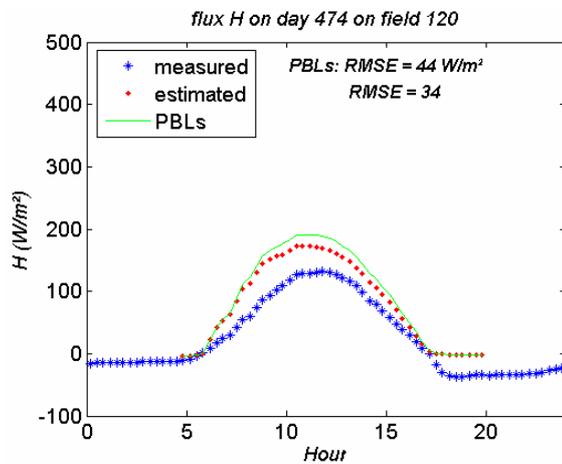
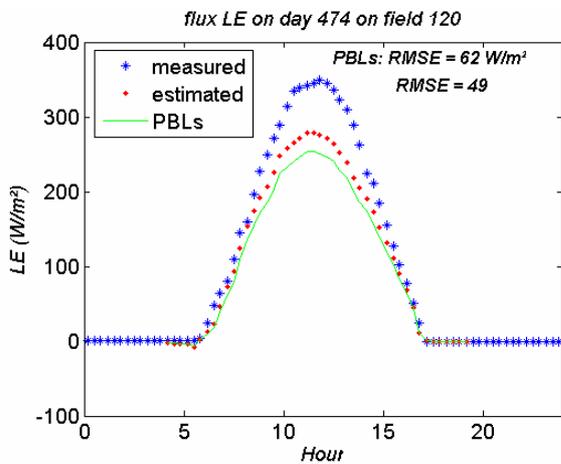
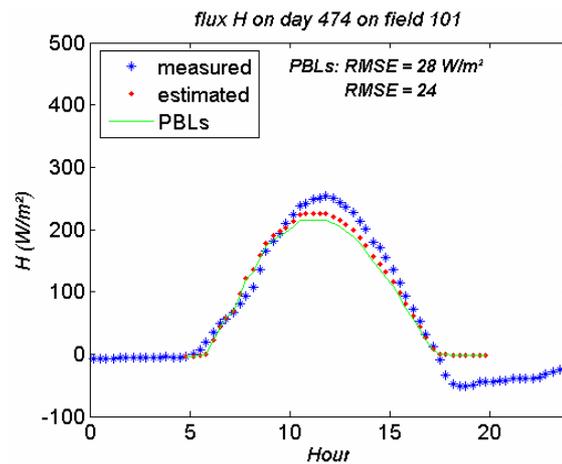
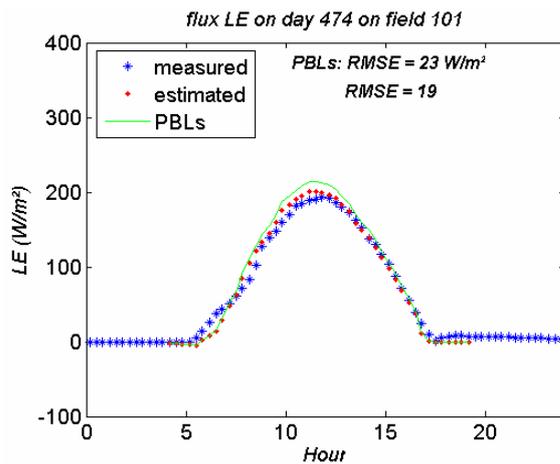
Les résultats obtenus sur la Figure V.8 sont très proches des résultats présentés pour le DOE 474 (figure V.7a) plus haut. Les températures estimées pour les dates 472 et 451 sont également très proches des simulations du modèle PBLs obtenues avec les valeurs d’ébauche (*RMSE* et biais du même ordre de grandeur que pour la première journée). L’écart de *RMSE* avant et après le processus d’assimilation est de 1 K pour le jour 472 et de 2 K pour le 451.

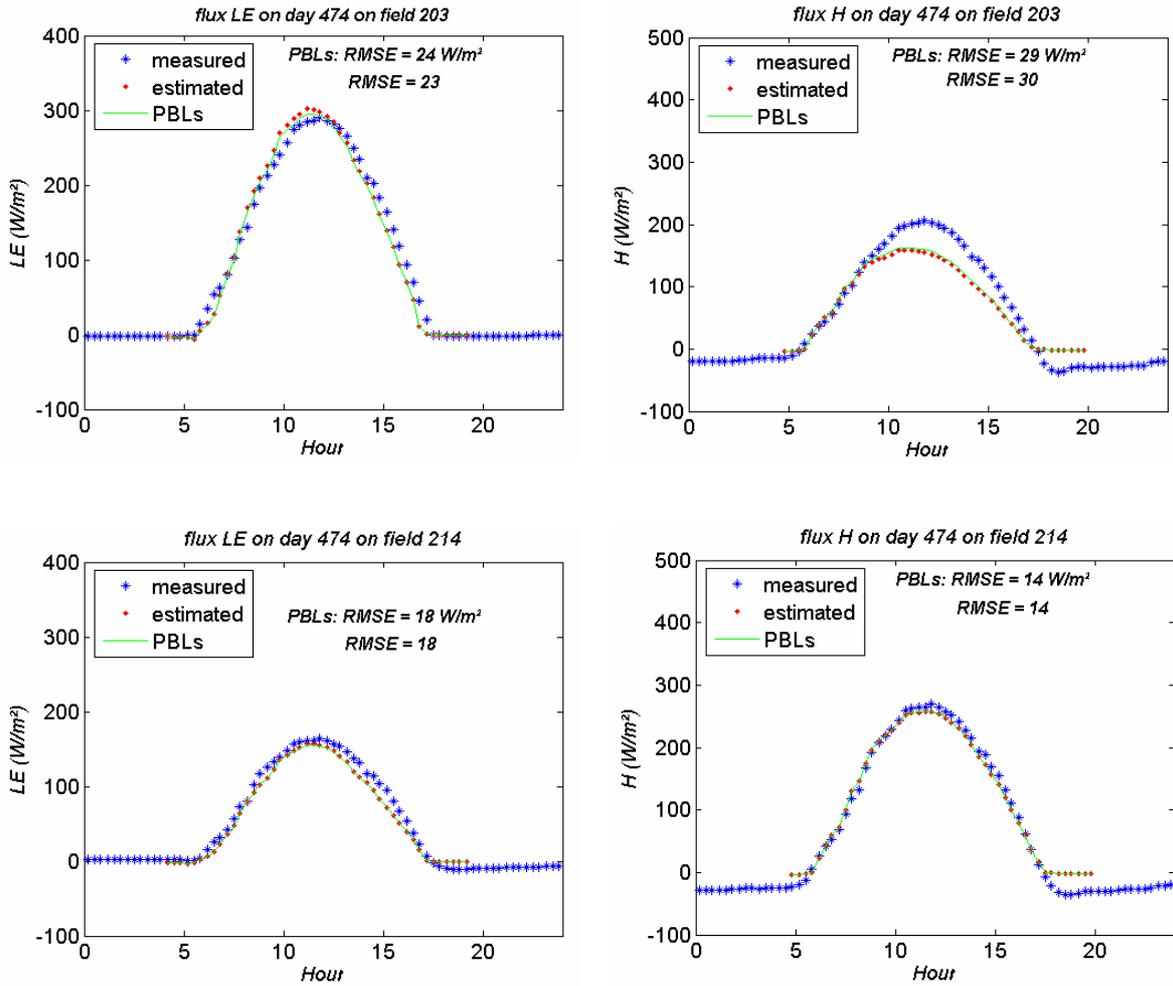
Les figures V.7a et V.8 présentent une forme d’ellipse sur les températures de surface estimées par l’algorithme, ce qui semble montrer que l’algorithme d’assimilation corrige de façon homogène l’erreur sur les observations.

Dans la section suivante, l’analyse des résultats de la méthode d’assimilation est approfondie par la validation sur les mesures de flux de surface des parcelles. La comparaison entre les mesures et les estimations de l’algorithme sur les parcelles nous permettra d’examiner l’apport des observations et le choix de la matrice de covariance d’erreur de l’information a priori ( $B$ ).

### 2.3. Validation sur les parcelles de mesure

Nous avons validé les estimations de l'algorithme d'assimilation sur les mesures de flux de surface et de températures sur les parcelles de la zone Alpilles. Nous avons présenté les résultats pour les DOE (Day Of Experiment) 451, 472, et 474 sur la Figure V.9-V11 et les Tableaux V.8-V.10 ci-dessous montrent la comparaison de l'évolution horaire du flux de surface estimé par l'algorithme d'assimilation et par le modèle PBLs par rapport aux mesures. La figure V.9 et le Tableau V.8 présentent les résultats de la date du 18 avril 1997 (DOE 474) sur les flux convectifs  $LE$  et  $H$  des parcelles 101 (blé), 120 (blé irrigué), 203 (luzerne), et 214 (blé de printemps). Il faut noter qu'à cette date le dispositif expérimental n'est pas en place sur les autres parcelles qui sont en sol nu (102, 121, et 501).





**Fig. V.9. Comparaison des flux  $LE$  et  $H$  simulés a priori (PBLs en vert) et estimés par l’algorithme d’assimilation (en rouge) sur les parcelles 101, 120, 203, et 214 de la zone Alpilles pour la date du 18 avril 1997 (DOE 474)**

**Tab. V.8. RMSE des flux de surface et des températures estimés (5h – 20h) sur les parcelles de mesures de la zone Alpilles pour la date 474: Comparaison entre le modèle PBLs et l’algorithme d’assimilation par rapport aux mesures**

	$R_n (Wm^{-2})$		$H (Wm^{-2})$		$LE (Wm^{-2})$		$G (Wm^{-2})$		$T_s (K)$		$T_a (K)$	
	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo
101	31	30	28	24	23	19	133	133	5,0	5,1	3,9	4,0
120	48	47	44	34	62	49	116	116	6,3	6,0	3,8	3,6
203	32	31	29	30	24	23	83	83	6,1	6,2	3,7	3,7
214	35	34	14	14	18	18	138	138	4,4	4,4	3,8	3,8

A première vue, on peut observer que les résultats de l'algorithme sur les flux de chaleur latente ( $LE$ ) et sensible ( $H$ ), ainsi que les simulations a priori, sont proches des mesures et les statistiques d'erreurs sont dans la gamme haute des résultats de la littérature (sauf pour la parcelle 120). Les résultats d'assimilation sont en général meilleurs que les simulations a priori, mais l'écart est faible (voire très faible). Par exemple, la validation sur la parcelle 101 montre un  $RMSE$  de  $23 \text{ Wm}^{-2}$  sur le flux  $LE$  avec le modèle PBLs contre  $19 \text{ Wm}^{-2}$  après le processus d'assimilation. Pour les parcelles 101, 203 et 214, les informations a priori choisies (facteur de l'humidité du sol  $f_b$ ) permettent de très bonnes simulations des flux. Pour la parcelle 120, ces informations a priori sont moins adaptées et l'assimilation permet une amélioration plus nette des estimations. Cependant, l'apport des observations reste insuffisant pour s'approcher d'avantage des mesures. Comme l'ont montré les études synthétiques le comportement pourrait être lié aux erreurs de mesure sur d'autres paramètres tels que le  $LAI$  ou la vitesse du vent qui ne sont pas contrôlés par la méthode d'assimilation. Cela peut aussi être lié à une mauvaise prescription de la valeur a priori de  $f_2$  et de ses erreurs associées. Cette dernière explication est tout à fait probable car la parcelle 120 a été fortement irriguée 15 jours avant la date étudiée (Oliosio et al., 2002a), irrigation qui n'a pas été intégrée aux simulations ISBA qui ont servi à l'établissement des propriétés a priori (Oliosio et al., 2006). La valeur de  $f_2$  pour cette période est aux alentours de 0,4 ce qui est faible pour une culture irriguée (elle devrait plutôt être de l'ordre de 0,8). On pourrait orienter le processus d'assimilation dans ce cas vers les observations en augmentant l'erreur de la variance de l'information a priori.

Le Tableau V.8 présente tous les résultats de la comparaison du modèle PBLs et de la méthode d'assimilation par rapport aux mesures des différentes parcelles pour l'ensemble des flux de surface et des températures. L'estimation des flux du rayonnement net ( $R_n$ ) par le modèle de PBLs et par l'algorithme d'assimilation montre une bonne cohérence par rapport aux mesures. La comparaison du flux  $R_n$  sur les parcelles présentées en annexe (cf. **Annexe IV**) souligne déjà une bonne modélisation du rayonnement net par le modèle PBLs. La méthode d'assimilation n'apporte qu'une légère correction (Tab. V.8). Concernant le flux de conduction dans le sol ( $G$ ), on ne note pas de changement. Comme l'a montré la comparaison des deux versions du modèle PBLs dans la **section 3** de la **partie III**, l'allure générale du flux  $G$  est très proche des mesures autour du midi solaire (voir **Annexe IV**). Quant aux températures de surface et de l'air à 2 m, on observe d'une part une amplitude plus importante par rapport aux mesures entre 5h et 10h et d'autre part une surestimation du modèle PBLs et de l'algorithme sur certaines parcelles (120, 203) particulièrement sur la température de surface (voir **Annexe**

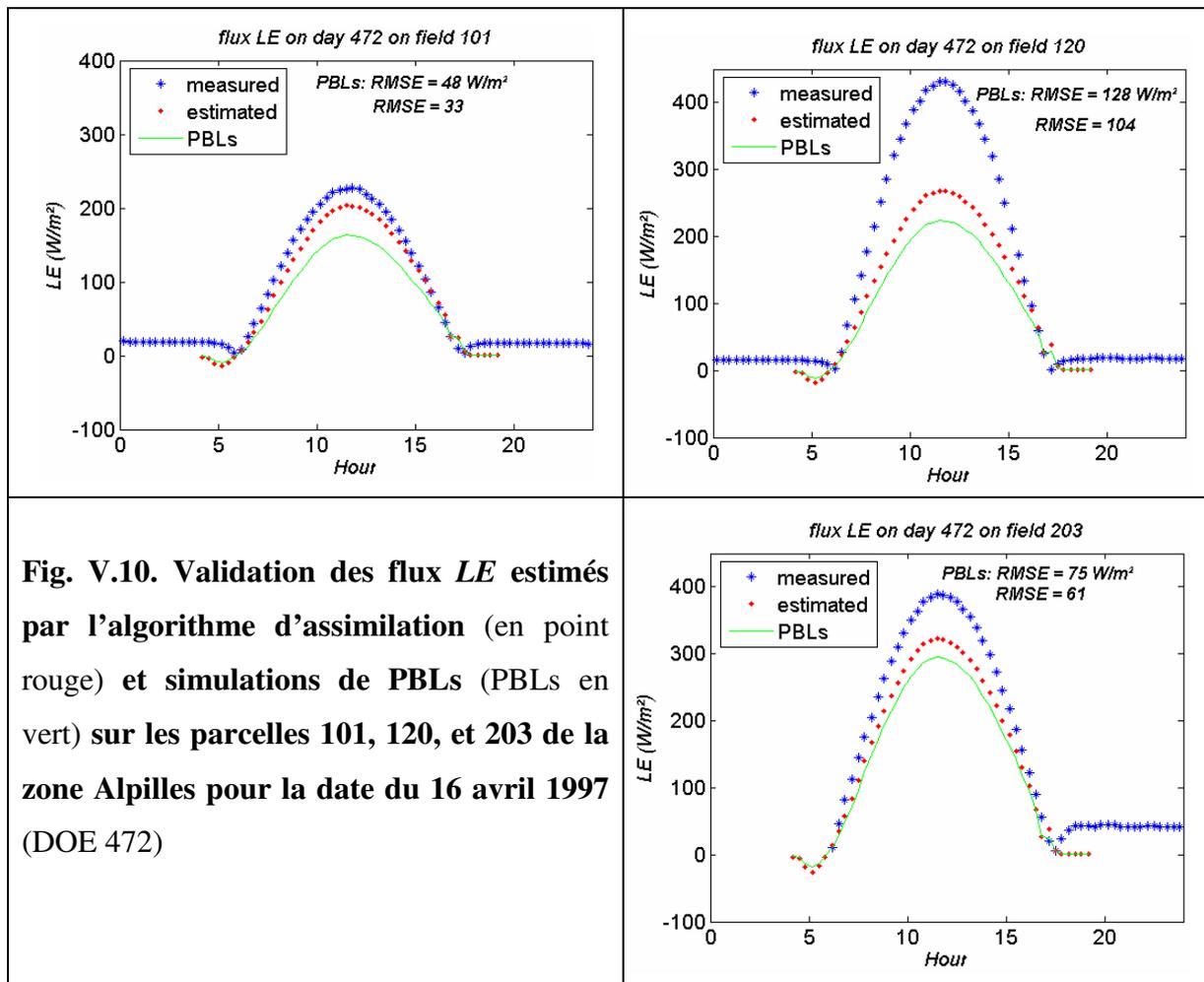
IV). En effet, les mesures de la température de surface sur les parcelles sont indépendantes des observations assimilées par l'algorithme (acquises par avion à midi solaire). On constate une différence de 1,5 K voire de 2 K entre les deux jeux de données. Par conséquent l'hypothèse d'une erreur d'observation de 1,5 K dans le processus d'assimilation est réaliste mais l'information apportée par les observations et l'information a priori reste insuffisante pour le contrôle des erreurs sur la température de surface.

L'application sur les deux autres dates (472 et 451) est présentée dans les tableaux V.9 et V.10.

**Tab. V.9. RMSE des flux de surface et des températures estimés (5h – 20h) sur les parcelles de mesures de la zone Alpilles pour la date 472:** Comparaison entre le modèle PBLs et l'algorithme d'assimilation par rapport aux mesures

	$R_n (Wm^{-2})$		$H (Wm^{-2})$		$LE (Wm^{-2})$		$G (Wm^{-2})$		$T_s (K)$		$T_a (K)$	
	<i>PBLs</i>	<i>Algo</i>	<i>PBLs</i>	<i>Algo</i>	<i>PBLs</i>	<i>Algo</i>	<i>PBLs</i>	<i>Algo</i>	<i>PBLs</i>	<i>Algo</i>	<i>PBLs</i>	<i>Algo</i>
101	28	31	35	30	48	33	126	126	2,4	2,2	3,1	2,8
120	46	46	100	80	128	104	103	104	4,4	3,4	2,7	2,4
203	30	31	61	52	75	61	78	78	4,1	3,5	2,4	2,2
214	37	37	60	54	67	59	133	133	1,4	1,4	3,3	3,1

Les résultats de la date 472 montrent de façon générale un apport de la méthode d'assimilation sur l'estimation des flux de surface. On note une bonne cohérence des résultats sur certaines parcelles telles que les flux convectifs de la parcelle 101 et la température de surface de la parcelle 214. En effet, l'estimation du flux  $LE$  sur la parcelle 101 est très bien différenciée de la simulation de PBLs (Fig. V.10), ce qui met en évidence l'apport de la méthode d'assimilation. Les températures de surface estimées à cette date du jour 472 (Tab. V.9) présentent une faible dégradation par rapport au jour 474 (Tab. V.8). On note un écart d'environ de 2 K par rapport aux mesures contre 5 K sur la date du 474. Par ailleurs, la parcelle 120 affiche un écart beaucoup plus fort d'environ  $20 Wm^{-2}$  sur l'estimation du flux  $LE$  par rapport au modèle PBLs. Cependant, le résultat final reste très éloigné des mesures (Fig. V.10). Nous avons noté que la valeur d'ébauche choisie (l'information a priori) pouvait ne pas être adaptée dans l'analyse du résultat de la parcelle 120 pour la date 474. Cette constatation est également valable pour la date 472. Dans cette situation, l'algorithme a pu extraire de l'information des observations de températures permettant de réduire l'écart entre les simulations de flux et les mesures.

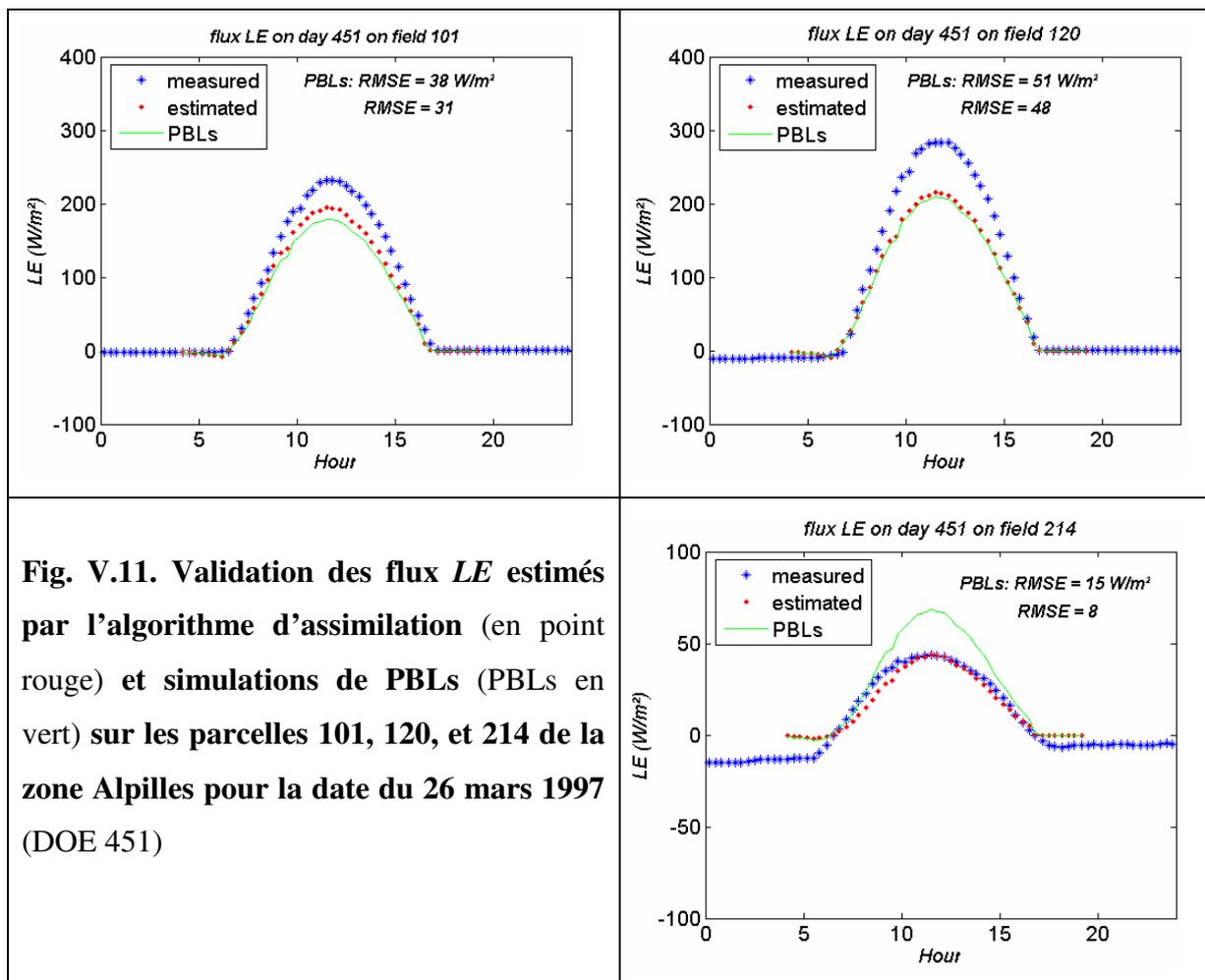


**Tab. V.10. RMSE des flux de surface et des températures estimés (5h – 20h) sur les parcelles de mesures de la zone Alpilles pour la date 451: Comparaison entre le modèle PBLs et l'algorithme d'assimilation par rapport aux mesures**

	$R_n (Wm^{-2})$		$H (Wm^{-2})$		$LE (Wm^{-2})$		$G (Wm^{-2})$		$T_s (K)$		$T_a (K)$	
	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo
101	24	23	28	23	38	31	109	110	5,4	5,2	4,8	4,8
120	40	40	37	35	51	48	103	103	5,4	5,3	4,8	4,8
203	29	29	31	31	53	54	63	62	6,0	6,0	4,8	4,8
214	42	41	32	27	15	8	116	115	4,9	4,9	5,0	5,0

Des résultats identiques à ceux des autres journées sont observés pour la date du 26 mars 1997 (DOE 451) (Tab. V.10 et Fig. V.11). Concernant la parcelle 120, les résultats sont peu modifiés

entre la simulation a priori et les flux estimés par l'assimilation, ce qui reste cohérent avec l'hypothèse de la non prise en compte de l'irrigation car la parcelle a été irriguée après le jour 451 (DOE 457- 460). Globalement, les résultats obtenus pour cette journée montre que la méthode d'assimilation permet des estimations correctes de l'évapotranspiration (~ flux  $LE$ ).



**Fig. V.11. Validation des flux  $LE$  estimés par l'algorithme d'assimilation (en point rouge) et simulations de PBLs (PBLs en vert) sur les parcelles 101, 120, et 214 de la zone Alpilles pour la date du 26 mars 1997 (DOE 451)**

Dans cette section, nous avons réalisé des assimilations en utilisant des informations a priori qui peuvent apparaître inappropriée soit dans leur valeur de base (comme par exemple pour la parcelle 120), soit dans l'erreur statistique associée. Certes, une précédente analyse ou la simulation d'un modèle peut constituer la base pour déterminer les valeurs de l'information a priori à condition que la dynamique du problème soit connue, mais la grande difficulté est liée aux composantes de la matrice de covariance d'erreur  $B$ . En effet, il faut donc être susceptible de minimiser l'impact d'observations très incertaines dans le processus. Ce qui est très complexe pour une application à l'échelle régionale en termes de connaissance de la zone

étudiée, et du coût de production (ici 10 jours de calcul). C'est pourquoi il est nécessaire de tester différentes approches pour déterminer l'information a priori autre que les sorties de modèle. Il est aussi probable d'analyser en amont les simulations du modèle direct (ici qui a un temps d'exécution relativement rapide, cf. Tableau IV.3) pour en tirer des règles d'adaptation des valeurs a priori. L'inconvénient résulte donc de la faible disponibilité des mesures pour une étude comparative avancée. Nous avons examiné dans la section 2.4 ci-dessous, l'équation (II.54) pour déterminer les valeurs d'ébauche de l'humidité normalisée du sol ( $f_b$ ) pour la date du 18 avril 1997 (DOE 474).

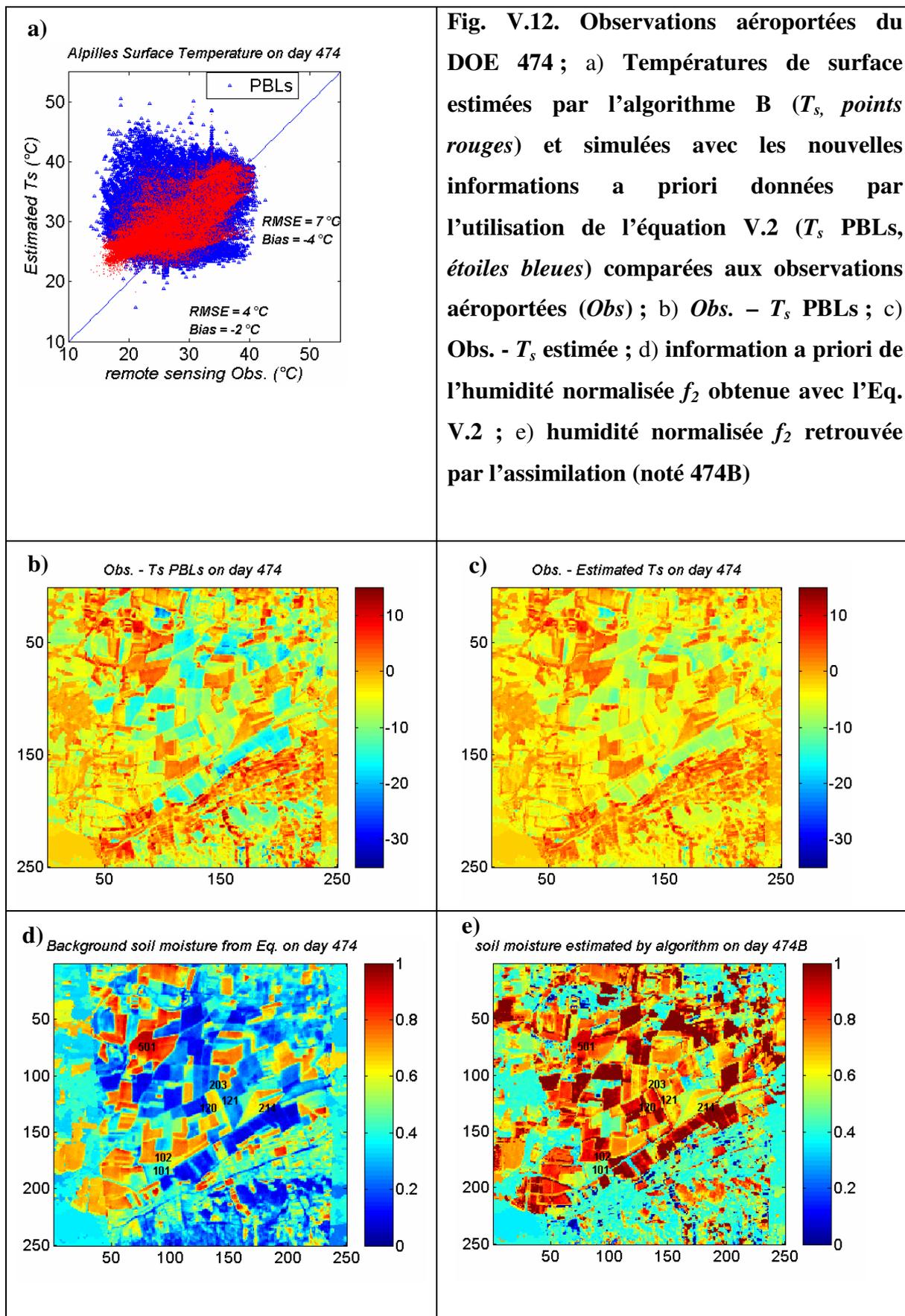
#### 2.4. Impact de l'information a priori de l'humidité du sol

Nous avons testé dans cette section une autre approche pour définir les valeurs de l'information a priori de l'humidité normalisée du sol. L'approche est décrite par l'équation (II.54) dans la **section 5.4** de la **partie II** rappelée ci-dessous par l'expression (V.2) :

$$f_{2i} = \frac{T_{s_i} - T_{s \min i}}{T_{s \max i} - T_{s \min i}} \quad (\text{V.2})$$

Pour la journée 474,  $T_{s_i}$  la température de surface utilisée correspond aux mesures aéroportées sur la zone Alpilles, avec  $T_{s \min i}$  et  $T_{s \max i}$  les températures de surface minimale et maximale déterminées pour chaque pixel suivant la théorie développée pour le modèle S-SEBI dans l'espace 'température de surface – albédo' et présentée par Roerink et al. (2000) et Gomez et al. (2005). Les écart type d'erreurs (matrice B) sont tels décrits dans la section 5.4 de la partie II.

L'impact de cette approche est analysé sur les résultats de la date du 18 avril 1997 (DOE 474). La figure V.12 montre la température de surface estimée par l'algorithme d'assimilation B avec une erreur d'observation de 1,5 K par rapport à l'observation, et l'écart entre les observations et la température de surface simulée par le modèle PBLs ainsi que celle de l'algorithme.



**Fig. V.12. Observations aéroportées du DOE 474 ; a) Températures de surface estimées par l’algorithme B ( $T_s$ , points rouges) et simulées avec les nouvelles informations a priori données par l’utilisation de l’équation V.2 ( $T_s$  PBLs, étoiles bleues) comparées aux observations aéroportées (Obs) ; b) Obs. -  $T_s$  PBLs ; c) Obs. -  $T_s$  estimée ; d) information a priori de l’humidité normalisée  $f_2$  obtenue avec l’Eq. V.2 ; e) humidité normalisée  $f_2$  retrouvée par l’assimilation (noté 474B)**

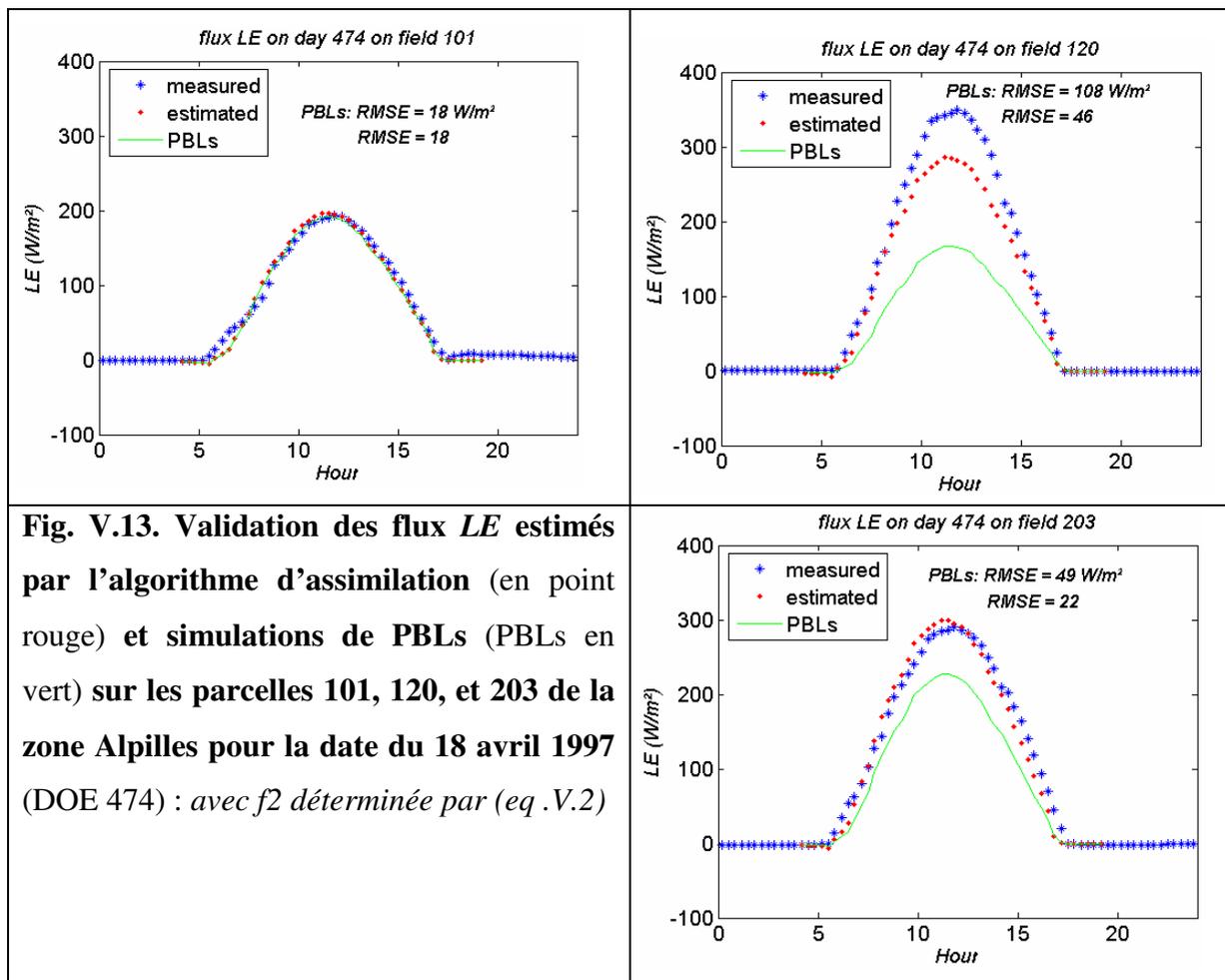
La figure V.12a montre bien l'impact de l'information a priori sur les résultats. En effet, la comparaison des températures de surface estimées par rapport aux observations présente un biais un peu plus faible par rapport à la figure V.7a sur la même date. On note aussi un écart de 3 K sur le *RMSE* entre les simulations du modèle PBLs et les estimations de l'algorithme. Les figures V.12b et V.12c montrent comment la différence entre les températures avant et après l'assimilation se répartissent dans l'espace. Par ailleurs, le résultat sur certains pixels de la cartographie (Fig. V.12b et c) est identique à la simulation de PBLs. Cela pourrait être à la fois dû à l'apport insuffisant des observations dans le processus ou à une valeur d'ébauche déjà correcte. Les Figures V.12d et V.12.e montrent comment les humidités normalisées sont modifiées au cours du processus d'assimilation. Tout d'abord on peut remarquer que les valeurs a priori obtenues avec l'équation (V.2) sont profondément différentes de celles qui ont été utilisées au préalable (Figure V.7d). On note comme précédemment une augmentation quasi générale de l'humidité (correspondant à une baisse du biais entre les températures de surface simulées et observées sur les Figures V.12a-c). Les humidités normalisées retrouvées montrent des valeurs atteignant dans beaucoup de situations la valeur seuil maximale de 1. Dans ces situations il n'est pas possible de mieux corriger les estimations de températures, mettant en évidence une mauvaise prise en compte des erreurs d'observation ou de modèles.

Le tableau V.11 ci-dessous résume la comparaison des estimations des flux et des températures sur les parcelles expérimentales. Il est à comparer aux résultats obtenus plus haut avec la première méthode permettant de fixer les informations a priori présentées dans le tableau V.8.

**Tab. V.11. *RMSE* des flux de surface et des températures estimés (5h – 20h) sur les parcelles de mesures de la zone Alpilles pour la date 474:** Comparaison entre le modèle PBLs et l'algorithme d'assimilation par rapport aux mesures avec  $f_2$  déterminée par (Eq. V.2)

	$R_n (Wm^{-2})$		$H (Wm^{-2})$		$LE (Wm^{-2})$		$G (Wm^{-2})$		$T_s (K)$		$T_a (K)$	
	<i>PBLs</i>	<i>Algo</i>	<i>PBLs</i>	<i>Algo</i>	<i>PBLs</i>	<i>Algo</i>	<i>PBLs</i>	<i>Algo</i>	<i>PBLs</i>	<i>Algo</i>	<i>PBLs</i>	<i>Algo</i>
101	30	29	23	23	18	18	132	132	5,4	5,4	4,1	4,1
120	55	46	82	31	108	46	113	117	8	5	4,4	3,6
203	39	31	17	29	49	22	81	83	8	6	4,2	3,7
214	35	34	15	14	17	17	138	138	4,5	4,6	3,8	3,8

Le tableau V.11 montre des améliorations importantes des estimations par rapport aux simulations a priori. Bien qu'obtenus à partir d'un jeu de valeurs d'ébauche ( $f_b$ ) différent, les résultats de l'assimilation sont numériquement très proches de ceux du tableau V.8. Par contre, les simulations a priori (résultats de PBLs) diffèrent de celles obtenues dans l'expérience précédente, et sont plus éloignées des valeurs mesurées. Cela montre, comme constaté dans les expériences synthétiques, le poids important de la température de surface observée sur la correction des variables simulées : quelque soit l'information a priori, l'algorithme d'assimilation converge vers le même état de températures simulées, le plus proche possible des observations de température de surface. Cela montre également que le poids accordé à la matrice de covariance d'erreur de l'information a priori ( $B$ ) permet de trouver un bon compromis par rapport aux observations dans cette configuration. L'algorithme d'assimilation a su extraire les informations des observations pour corriger l'écart entre les simulations du modèle PBLs et les mesures. La figure V.13 représentant la validation du flux  $LE$  sur les différentes parcelles par rapport aux mesures illustre l'information apportée par l'information a priori et l'observation à l'algorithme pour estimer ce flux.



Au vu de ces résultats, nous pouvons dire que définir l'information a priori de l'humidité du sol (notamment la matrice d'erreur à associer) par l'équation (II.54 ou V.2) autre que les précédentes simulations de modèle permet d'obtenir d'excellents résultats. L'application de la méthode d'assimilation sur les données réelles de la zone Alpilles montre que la méthode est pertinente pour l'estimation de l'évapotranspiration. L'apport de l'information a priori et de l'observation constitue l'élément essentiel de la méthode. Le compromis entre le poids accordé aux deux matrices de covariance d'erreur ( $B$  et  $R_i$ ) oriente le processus d'assimilation vers les observations de confiance.

Les différents résultats présentés ci-dessus montrent que l'algorithme d'assimilation mis en place reste assez positif et surtout prometteur. Nous avons vu qu'il est réalisable et que si il n'est pas trop gourmand en ressource informatique il pourrait être employé de façon régulière et ainsi permettre des suivis de l'évapotranspiration au cours du cycle de développement des cultures.

## **2.5. Récapitulatif des résultats autour du midi solaire et cartographies de l'évapotranspiration**

Le récapitulatif de tous les résultats autour du midi solaire (heure d'acquisition des données aéroportées) montre que la méthode d'assimilation variationnelle de données fournit des résultats très satisfaisants pour l'estimation des flux et du bilan d'énergie. Les résultats sont moins satisfaisants pour les températures. L'apport de la procédure d'assimilation est net pour les flux turbulents ( $H$  et  $LE$ ) par rapport aux simulations a priori (résultats de PBLs). Le tableau V.12 montre que les niveaux d'erreur par rapport aux mesures de terrain sont sensiblement réduits par la méthode d'assimilation autour du midi solaire. Pour le flux de chaleur  $LE$ , on observe l'écart le plus important par rapport à la simulation a priori de PBLs pour la date 474B ( $97 - 32 = 65 \text{ Wm}^{-2}$ ).

La figure V.14 présente la comparaison aux mesures des flux convectifs ( $LE$  et  $H$ ) estimés par la méthode d'assimilation à midi pour les trois dates et sur l'ensemble des parcelles de la zone Alpilles (culture de blé : 101, 120, 214 ; et culture de luzerne : 203). On note une bonne cohérence entre les estimations et les mesures.

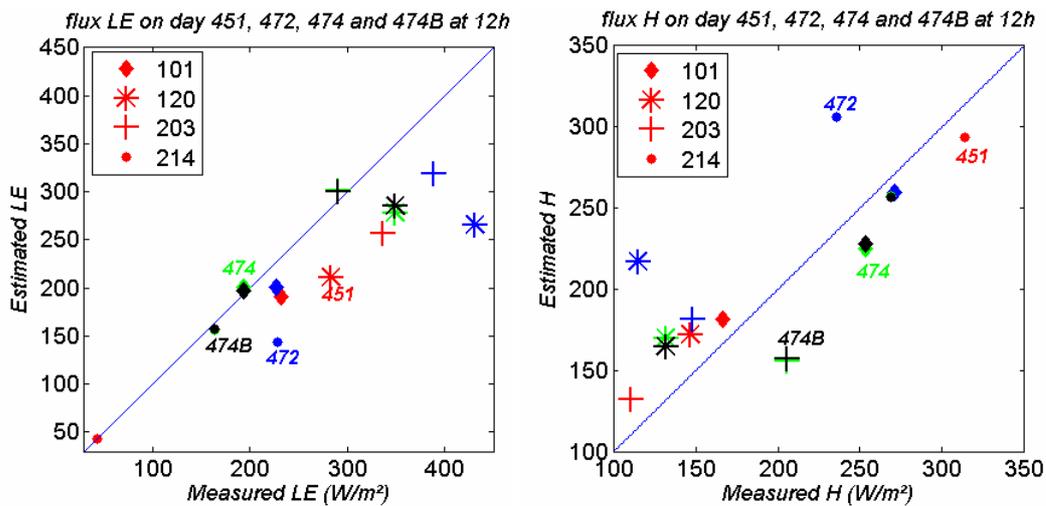
Dans le tableau V.12, on notera également les bons résultats obtenus pour le flux  $G$ , en apparence contradiction avec les résultats des tableaux précédents. Cela montre que la paramétrisation du flux introduite à partir de l'approche de Choudhury et al. (1987) est correcte à cette heure là, mais qu'elle ne peut pas rendre compte correctement des variations horaires du flux. La comparaison de la méthode d'assimilation au modèle résiduel confirme les résultats obtenus sur les données d'expériences jumelles. La réponse apportée par le système d'assimilation dans l'estimation de l'évapotranspiration est nettement supérieure à celle du modèle résiduel. En effet, les flux de chaleur sensible et latente obtenus par la méthode résiduelle montre des écarts importants par rapport aux mesures, seuls les flux  $R_n$  et  $G$  présentent de bons résultats (Tab. V.12a, voir aussi **Annexe IV** pour les figures du flux  $LE$ ).

**Tab. V.12a. RMSE des flux de surface et des températures estimés à 12h sur l'ensemble des parcelles de la zone Alpilles (101, 120, 203, 214) pour les dates 451, 472, 474, et 474B ( $f_2$  obtenue par Eq. V.2):** Comparaison entre le modèle PBLs, l'algorithme d'assimilation (avec  $\sigma_i = 1,5$  K) et le modèle Résiduel (**section II.5**) par rapport aux mesures. Résultats du modèle résiduel obtenus en utilisant en forçage la température de surface des mesures aéroportées et les valeurs du **Tableau V.12b** ci-dessous

Date	451			472			474			474B	
	PBLs	Algo	Rési	PBLs	Algo	Rési	PBLs	Algo	Rési	PBLs	Algo
$R_n (Wm^{-2})$	21	19	18	26	27	55	29	25	24	44	25
$H (Wm^{-2})$	29	20	79	86	64	111	42	34	98	63	32
$LE (Wm^{-2})$	63	57	70	129	100	105	49	36	99	97	32
$G (Wm^{-2})$	15	14	19	37	38	45	18	17	18	19	17
$T_s (K)$	5,4	5,3	-	4,7	3,7	-	6,0	5,0	-	8,2	4,5
$T_a (K)$	3,2	3,0	-	3,3	2,6	-	2,6	2,2	-	3,1	1,9

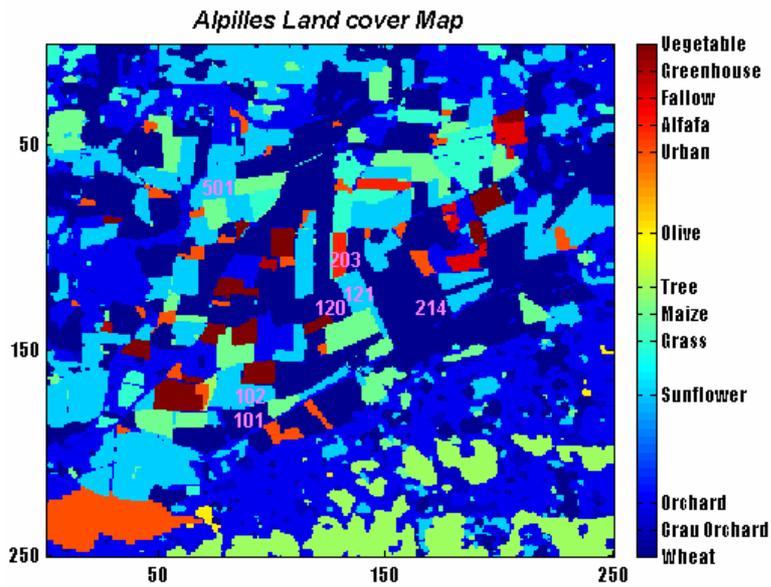
**Tab. V.12b. Valeurs de la hauteur et de la température de l'air CLS utilisées pour le modèle résiduel (mesures radiosondages à Nîmes à 60 m) et les valeurs simulées par le modèle PBLs pour les dates 451, 472, 474, et 474B ( $f_2$  obtenue par Eq. V.2)**

Date	451			472			474			474B	
	PBLs	Algo	Rési	PBLs	Algo	Rési	PBLs	Algo	Rési	PBLs	Algo
$h_{CLS} (m)$	92	89	130	133	130	119	121	117	131	115	111
$T_{CLS} (°C)$	14,1	13,9	15,1	14,2	14,0	17,8	13,4	13,2	13,7	13,1	12,8

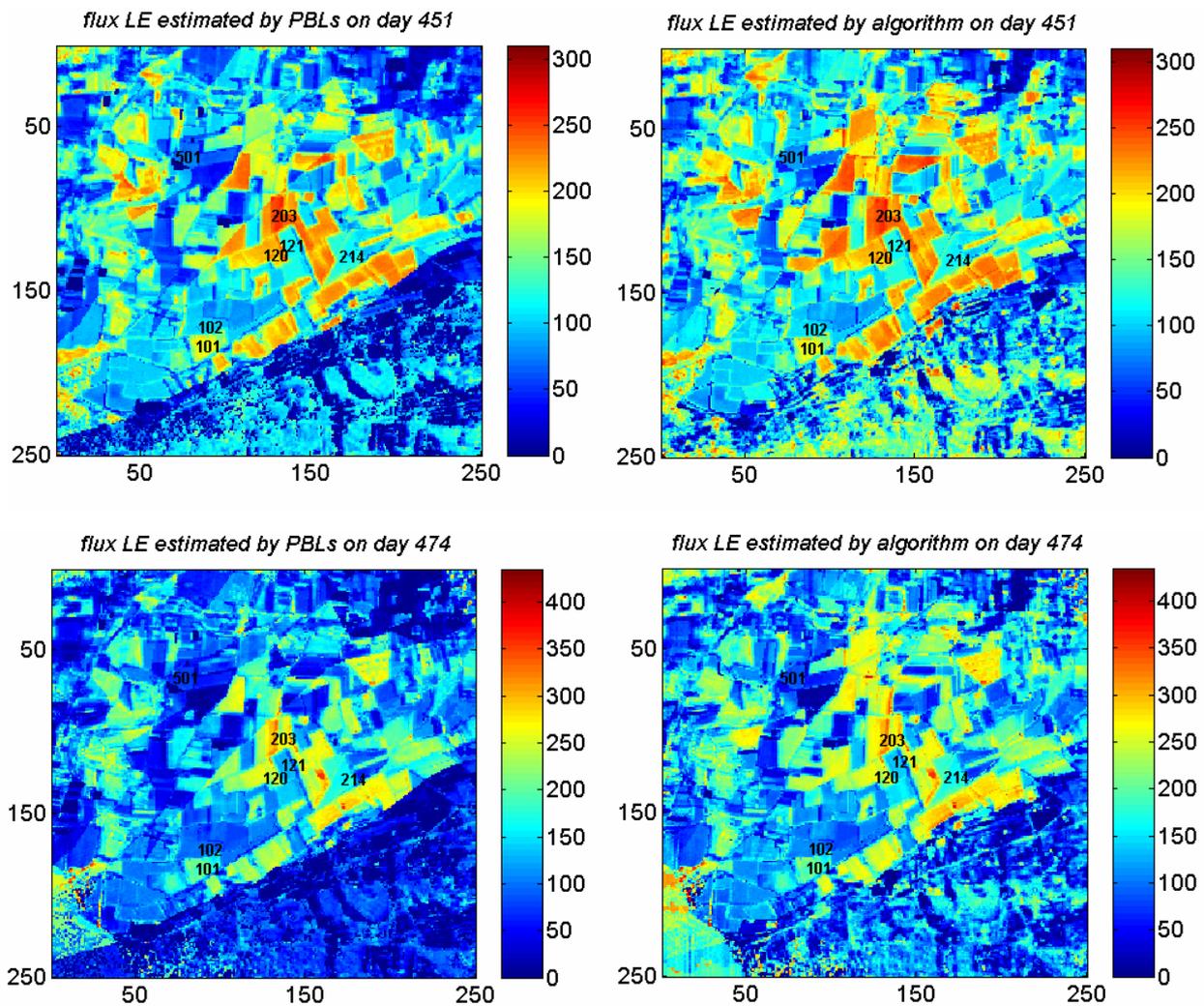


**Fig. V.14. Flux  $LE$  (à gauche) et Flux  $H$  (à droite) estimés par l’algorithme d’assimilation à midi solaire avec  $\sigma_i = 1,5$  K sur l’ensemble des parcelles de la zone Alpilles pour les dates 451(couleur rouge), 472(couleur bleue), 474(couleur verte), et 474B (couleur noire) ( $f_2$  obtenue par Eq. V.2) :**

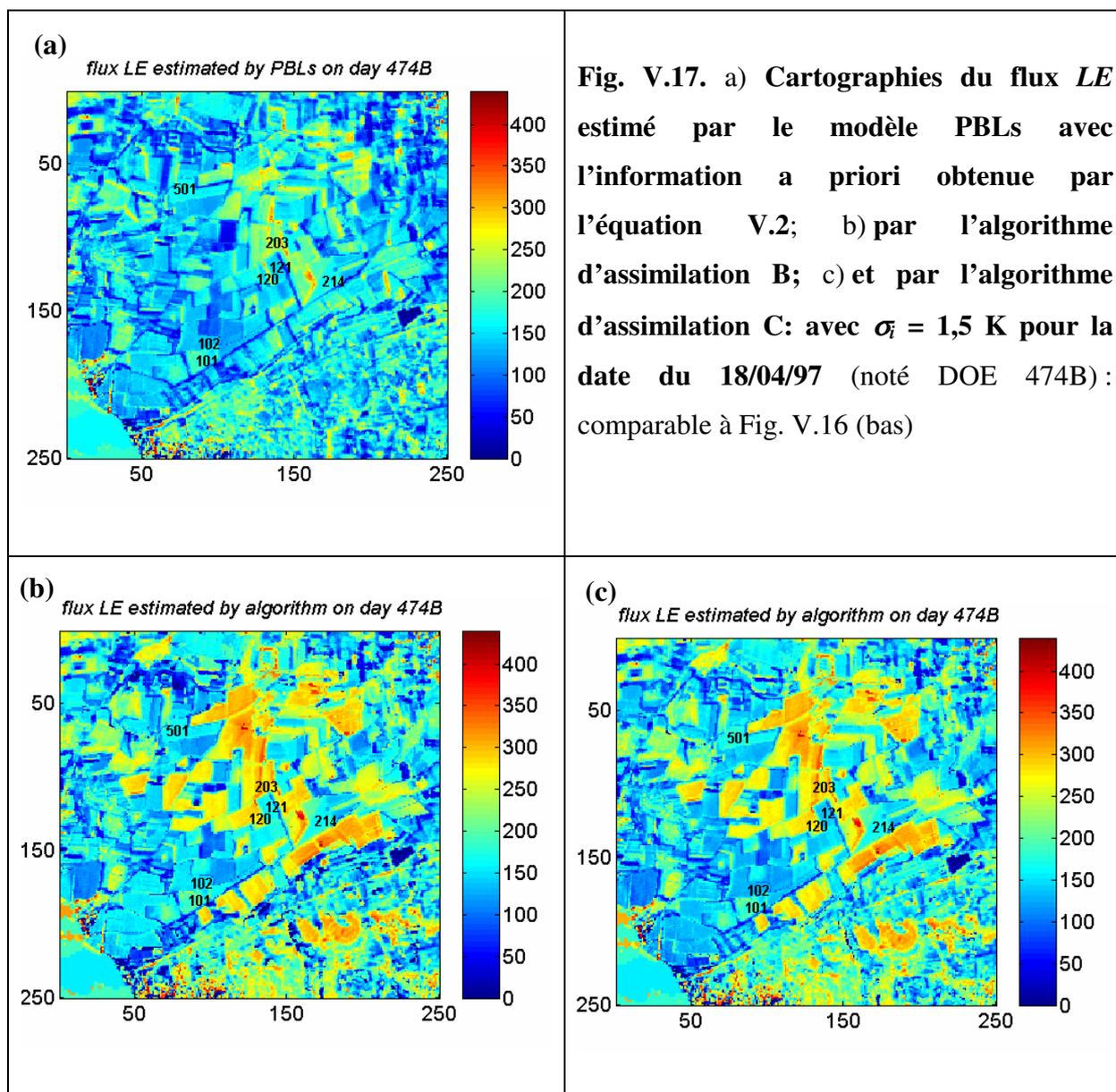
L’analyse des différents résultats sur les parcelles et pour les différentes dates, nous donne une première validation de la méthode d’assimilation variationnelle de données. Nous présentons maintenant ci-dessous (Fig. V.16 et Fig. V.17) les cartes de flux de chaleur latente ( $LE$ ) à 12h sur toute la zone Alpilles pour les trois dates. D’une part on présente la simulation de PBLs obtenue avec les informations a priori, et d’autre part le résultat déterminé avec la méthode d’assimilation variationnelle de données en corrigeant les paramètres de l’humidité du sol normalisée ( $f_2$ ) ou/et la rugosité aérodynamique ( $z0_m$ ). L’analyse des cartes avant et après assimilation nous permettra de remarquer les changements qui s’opèrent en fonction de la connaissance des parcelles de la zone (la figure V.15 montre la carte d’occupation des différentes cultures sur la zone Alpilles).



**Fig. V.15. Carte d'occupation par classe des cultures de la zone Alpilles**



**Fig. V.16. Cartographies du flux  $LE$  estimé par le modèle PBLs avec l'information a priori obtenue au moyen du modèle ISBA (à gauche) et par l'algorithme d'assimilation B avec  $\sigma_i = 1,5$  K (à droite) pour les dates du 26/03/97 (DOE 451 en haut) et du 18/04/97 (DOE 474 en bas)**



La comparaison des cartes du flux  $LE$  simulé par le modèle PBLs avec l'information a priori et l'algorithme d'assimilation montre d'importantes différences sur la Figure V.16. Elles sont

surtout conséquentes sur les zones périphériques et en particulier la zone de vergers et de bois au Sud-Est. Globalement une augmentation des flux  $LE$  est constatée (illustrant une augmentation de l'humidité entre l'information a priori et l'information retrouvée par l'algorithme, traduisant donc des conditions a priori trop sèches par rapport à l'information fournie par les images de température de surface observée).

L'analyse des cartes des dates 474 (Fig. V.16d) et 474B (Fig. V.17b) obtenues avec l'algorithme B montre également des différences importantes sur certaines zones et une différence globale de niveau d'évapotranspiration ( $474B > 474$ ). Bâties à partir d'informations a priori différentes, les assimilations ne convergent pas vers un résultat similaire ce qui a tendance à contredire les résultats obtenus lors des comparaisons aux mesures de terrains (cf. sections précédentes). Par contre, cela illustre le rôle des conditions a priori (valeurs et niveaux d'erreurs) qui dans le premier cas (474) limitent les possibilités de correction des humidités par les observations.

Le flux de chaleur  $LE$  pour la date 474B a été établi également avec l'algorithme C (qui cherche à identifier la rugosité et l'humidité simultanément) (Fig. V.17c). Comparée à la carte obtenue par l'algorithme B (Fig. V.17b), elles présentent des résultats extrêmement proches. Ce qui est en accord avec l'analyse sur les données d'expérience jumelle (Tab. V.6). En effet nous avons observé que l'écart entre les estimations de l'algorithme B (recherche de  $f_2$  seul) et celles de l'algorithme C (recherche simultanée de  $z_{0m}$  et  $f_2$ ) était faible. Par contre, la comparaison de la température de surface estimée par l'algorithme C par rapport à l'observation assimilée montre une correction légèrement au dessus du résultat de l'algorithme B (cf. **Annexe IV**).

## 2.6. Conclusion sur les données réelles

Nous avons testé l'algorithme d'assimilation de la température de surface dans le modèle PBLs sur les données réelles de l'expérimentation Alpilles. Ce test effectué sur 3 journées de printemps, s'est montré très encourageant à plusieurs niveaux.

- La validation des résultats obtenus sur les parcelles de mesures montre un bon accord par rapport aux mesures (Fig. V.14 et Tab. V.12a). Les résultats à 12 h, heure d'acquisition des données de télédétection sont d'assez bonne qualité. Pour le flux de chaleur latente, ils sont meilleurs que la méthode résiduelle. La figure V.17 illustrant la cartographie de l'évapotranspiration (flux  $LE$ ) à l'échelle régionale décrit une toute autre variabilité que la simulation du modèle PBLs à partir des informations a priori. De plus, les estimations de variables difficiles d'accès comme la hauteur de la couche limite et la température de l'air à cette hauteur sont même apparues très correctes (Tableau V.12b). La qualité des variables recherchées laisse penser que la méthode devrait permettre d'observer leur évolution à pas de temps régulier ouvrant la porte à une application plus continue de la méthode pour le suivi des flux de surface et du microclimat.
- L'analyse sur les données réelles a permis d'illustrer l'importance du choix des valeurs a priori des paramètres à retrouver et des erreurs associées d'une manière assez parlante. Pour une journée d'étude, deux processus de définition des valeurs a priori du paramètre d'humidité normalisée  $f_2$  ont été comparés (issues d'une simulation du modèle ISBA ou définies par une simple équation empirique exploitant la variabilité spatiale des températures de surface (Eq. V.2)). Le rôle des termes d'erreurs de la matrice  $B$  est particulièrement bien mis en évidence dans le premier cas (Figures V.7 et V.12).

Par contre, certains points restent encore à mieux maîtriser. Outre, les inconvénients de la méthode d'assimilation liés la définition de façon systématique et plus précise de l'information a priori et de la matrice de covariance d'erreur ( $B$ ) ; Il faut ajouter que des observations (température de surface,  $LAI$ , albédo et vitesse du vent) de meilleures qualités ou fiables sont parfois nécessaires pour trouver un bon compromis par rapport au terme d'ébauche. Par ailleurs, le coût de production en termes de temps calcul de l'évapotranspiration à l'échelle régionale est pour l'instant considérable avec le matériel dont nous disposons. La zone d'étude

de 5 km par 5 km considérée nécessite en moyenne une durée de 10 jours pour déterminer les paramètres optimaux (si aucun autre problème technique n'intervient). Certaines pistes d'amélioration concerne une analyse plus fine des critères d'arrêt et de précision de l'algorithme d'assimilation (nous avons fixé ici une précision de  $10^{-6}$  pour l'estimation des paramètres optimaux, ce qui semble être très élevé pour la gamme de variation des variables recherchées), ou encore la possibilité de travailler par sous-classe de culture en regroupant certains pixels suivant différentes informations pouvant être disponibles comme les propriétés des sols et l'analyse de la phénologie des cultures, ou encore en réalisant une classification non supervisée des informations de télédétection disponible.

## **Conclusion**

La performance de l'algorithme d'assimilation d'une part sur les données synthétiques et d'autre part sur les données réelles d'Alpilles ReSeDA, nous permet d'évaluer l'apport de la méthode et de fixer quelques limites en terme d'impact d'erreur. Les résultats de la méthode d'assimilation montrent de façon générale une meilleure estimation par rapport au modèle PBLs. L'analyse sur les données d'expérience jumelle a montré que les modifications apportées au modèle initial de PBLs ( $kB^{-1}$  et  $G/R_n$ ) peuvent réduire l'erreur modèle dans certains cas, mais ont surtout un impact sur le lien entre les températures de surface, les flux et l'humidité du sol, permettant en particulier une meilleure estimation de ce dernier.

Nous avons vu que l'impact d'erreur des autres paramètres d'entrée (non corrigé par la méthode) tels que l'albédo, la vitesse moyenne journalière du vent et le  $LAI$  ne sont pas négligeables sur la précision que nous pouvons attendre de l'estimation de l'évapotranspiration. En effet, la dégradation de l'indice foliaire  $LAI$  traduit une importante dispersion sur les flux de chaleur latente ( $LE$ ) et sensible ( $H$ ). Certes la méthode d'assimilation contrôle relativement ces erreurs suivant la qualité des observations mais l'information apportée par les observations à la procédure d'assimilation se révèle parfois insuffisante pour réduire l'ensemble de toutes ces erreurs.

Par ailleurs, les stratégies d'estimer un seul paramètre ou deux paramètres à la fois produisent des résultats similaires (à condition que la valeur a priori de la rugosité  $z0$  ne soit pas trop fausse). Cela est dû à la sensibilité conséquente du facteur d'humidité du sol normalisée ( $f_2$ ) par rapport à la rugosité aérodynamique ( $z0$ ). En effet la connaissance actuelle de la rugosité  $z0$  dans notre cas sur les données Alpilles permet d'obtenir des estimations correctes avec un faible impact d'erreur du paramètre. De ce fait, les algorithmes B (ne rechercher que  $f_2$ ) et C (rechercher  $z0$  et  $f_2$ ) donnent des résultats presque similaires dans les différents scénarios testés ci-dessus.

La performance aussi bien sur les données synthétiques que sur les données réelles ont révélé que la méthode d'assimilation apporte une meilleure réponse par rapport au modèle résiduel dans l'estimation de l'évapotranspiration à l'échelle régionale. En effet, l'approche résiduelle ne tient pas compte d'éventuelles erreurs d'observation dans l'analyse. Les interactions dues à

l'hétérogénéité de la surface sont négligées (Tab. II.4). De plus, la connaissance avec précision des variables intervenant dans le modèle (hauteur de couche limite, température de l'air) au temps de calcul est très problématique.

Il faut noter que le bon fonctionnement de l'algorithme d'assimilation est déterminé aussi par le terme d'ébauche (l'information a priori fournie) et surtout par les matrices de covariance d'erreur de l'information a priori ( $B$ ) et de l'observation ( $R_i$ ). Les valeurs de l'information a priori sont considérées comme des observations supplémentaires. L'information a priori apporte donc une information à l'algorithme d'assimilation en indiquant comment et à quel degré le paramètre recherché doit être corrigé. Par conséquent le choix de l'information a priori et l'erreur statistique associée (matrice  $B$ ) constitue un élément essentiel dans la procédure d'assimilation. L'analyse sur les données réelles, comparant deux méthodes de définition des données a priori, a permis de bien illustrer le poids de ce facteur dans les résultats de l'assimilation. La connaissance de ce terme pour une application opérationnelle est déterminante. C'est pourquoi une analyse en amont sur le domaine d'étude est nécessaire afin de définir de façon systématique et aussi précise que possible le paramètre a priori.

## ***CONCLUSION GENERALE***

### ***Rappel des objectifs***

L'objectif de ce travail était d'améliorer l'estimation des différents échanges entre la surface et l'atmosphère, en particulier l'évapotranspiration d'un paysage agricole. Pour ce travail, nous avons choisi d'utiliser un modèle simulant les transferts entre le sol, la plante et l'atmosphère, le modèle PBLs, qui est couplé à une description simplifiée de l'évolution de la couche limite atmosphérique (CLA). Le choix de ce modèle résultait de plusieurs critères: il représentait en effet l'ensemble des processus qui nous intéressaient, en simulant à la fois les principaux flux de surface mais aussi les températures de surface et de l'air, l'humidité de l'air et l'évolution de la CLA, et sa relative simplicité facilitait sa mise en oeuvre pour différentes applications.

Son application à l'échelle régionale nécessite l'utilisation de données de télédétection afin d'avoir accès à la variabilité spatiale des principales caractéristiques de surface. Si certaines variables requises pour le fonctionnement du modèle PBLs telles que l'indice foliaire ou l'albédo sont obtenues avec une précision acceptable à partir d'images satellitaires, d'autres variables en revanche, comme la rugosité de surface ( $z_0$ ) ou encore l'humidité des sols ( $f_2$ ) sont encore difficiles à estimer correctement. Or ce sont des variables qui ont une forte influence sur les flux de surface et les températures et qui présentent une grande variabilité spatiale à l'échelle d'un paysage agricole.

L'originalité de ce travail a été ici de proposer une méthode afin d'estimer ces derniers paramètres plus précisément, en implémentant dans PBLs une méthode d'assimilation variationnelle, basée sur l'introduction de données thermiques observées. Ces méthodes étaient jusqu'à présent peu utilisées dans ces types de modèles (type TSVA).

Les satellites actuels délivrant des températures de surface à fine résolution ne permettent d'acquérir qu'une seule donnée par jour, aussi nous avons donc évalué quel serait l'apport d'une seule température de surface observée par jour pour retrouver correctement  $z_0$  et  $f_2$ , et améliorer ainsi l'estimation des flux de surface et de la température de l'air.

Afin de mieux prendre en compte l'effet de l'hétérogénéité du paysage sur l'évolution de la CLA, nous avons modifié le modèle PBLs en proposant une approche patchée qui permet de représenter les différents constituants du paysage suivant leur pourcentage de surface occupée et les caractéristiques des couverts végétaux. Ainsi, l'évolution de la couche limite dépend de la proportion et des caractéristiques des constituants qui composent le paysage. Nous avons

ensuite introduit la méthode d'assimilation variationnelle de données dans le modèle PBLs. Elle consiste à minimiser une fonction mesurant l'écart entre la température de surface simulée par le modèle PBLs et les températures observées. Un algorithme ajuste les variables du modèle qui nous intéressent ici (la rugosité aérodynamique et l'humidité normalisée du sol) afin d'obtenir une bonne cohérence entre les sorties du modèle et les observations. Le grand nombre d'informations à retrouver, nous a amené à utiliser la méthode adjointe habituellement appliquée pour les problèmes de grande taille. Par conséquent, nous avons développé l'adjoint de PBLs permettant ainsi de réduire le temps de calcul du processus de minimisation. Les logiciels tels que TAPENADE et PALM ont été utilisés pour la mise en place de ce système d'assimilation.

Par ailleurs, la mise en place du modèle PBLs et de la procédure d'assimilation de données infrarouges thermiques, nous a montré que les connaissances imparfaites sur certains autres paramètres pouvaient avoir un impact important sur l'estimation des flux de surface. Nous avons ainsi analysé plus particulièrement le paramètre  $c_0$  représentant le rapport entre le flux conductif dans le sol et le rayonnement net ( $c_0 = G/R_n$ ), ainsi que le paramètre lié à l'estimation de la rugosité thermique  $z_0h$  (paramètre aérodynamique  $kB^{-1}$ ). Ces deux paramètres varient grandement en fonction des surfaces et conditions environnementales. Nous avons comparé différents formalismes proposés dans la littérature pour l'estimation de ces deux paramètres afin de déterminer ceux qui donnent les meilleurs résultats sur les flux de surface et les températures. De nouvelles paramétrisations ont ainsi été introduites dans le modèle PBLs (version 2).

L'étude suivante du travail a été d'évaluer le modèle et l'apport du système d'assimilation à partir de différents tests, notamment des tests d'expériences jumelles. Nous avons ainsi pu tester l'impact de la qualité des observations et confronter les résultats à une approche plus classique d'estimation de l'évapotranspiration.

### ***Principaux Résultats***

Le modèle PBLs permet de simuler les interactions surface - couche limite atmosphérique à l'échelle de la journée. Les travaux réalisés sur le site de Crau – Camargue montrent que l'introduction des patches dans le modèle PBLs permet une amélioration des variables recherchées, en particulier en termes de variabilité spatiale du microclimat (Courault et al.

2008). Nous avons montré que PBLs (en mode direct) pouvait avoir des performances assez proches de celles de SEBAL (du moins sur les parcelles de validation). Dans le travail présenté dans ce mémoire, la comparaison des approches avec ou sans patch a montré que les rétroactions de l'atmosphère sur la surface à travers la température de l'air et la pression de vapeur engendrent des écarts importants sur l'estimation du flux de chaleur latente  $LE$  (notamment dans les cas extrêmes : zones sèche ou humide).

L'analyse des paramètres  $z0_h$  et  $G/R_n$  qui sont assez mal connus de façon théorique a révélé que :

- si le rapport  $G/R_n$  est considéré constant au cours de la journée, cela peut conduire à une sous-estimation du flux  $G$  dans la matinée et une surestimation du flux  $G$  dans l'après midi, en particulier pour les sols nus. Dans le cas d'une surface couverte de végétation plus ou moins dense, l'expression du rapport  $G/R_n$  en fonction de l'indice foliaire (modèle de Choudhury et al. 1987) a permis d'obtenir des résultats corrects sur les mesures de terrain de l'expérimentation Alpilles et du site grue (en considérant les estimations à midi solaire). Dans le cas des sols nus, nous avons évalué deux approches tenant compte des évolutions horaires du rapport  $G/R_n$ . La comparaison de ces deux approches a montré qu'il ne semblait pas pertinent de complexifier l'approche en tenant compte de l'impact de la vitesse du vent sur l'estimation du flux  $G$ . En effet, cette dernière est plus difficile à calibrer en raison d'un plus grand nombre de coefficients à identifier, sans apporter d'améliorations décisives des simulations de  $G$ .

Les deux études que nous avons réalisées sur l'estimation du flux  $G$  à partir du rayonnement net, n'ont pas permis de dégager une modélisation permettant de décrire correctement l'ensemble des situations abordées, mais elles nous ont permis en revanche d'obtenir des flux  $G$  de bonne qualité sur les données d'évaluation.

- On retient également de l'étude sur la rugosité thermique  $z0_h$  (ou le paramètre  $kB^{-1}$ ) que prescrire une valeur constante est plutôt adapté pour une surface totalement homogène. Des valeurs beaucoup plus grandes et variées sont par contre observées sur un paysage agricole hétérogène. Dans cette étude, nous avons évalué des modèles permettant de décrire ce paramètre. On a pu noter que le modèle proposé par Su et al. (2001) (le modèle de Massman) donne un meilleur résultat et affiche une sensibilité faible aux erreurs liées aux variables d'entrée du modèle (émissivité,  $LAI$ ,  $z0$ ) par rapport aux modèles de Hasager et de Blümel.

L'intégration de ces deux résultats (sur  $G/R_n$  et  $z0_h$ ) a conduit à introduire de nouvelles équations dans le modèle PBLs qui a alors montré de meilleurs résultats que la version initiale qui considérait ces paramètres constants sur toute la zone étudiée. La validation sur les parcelles de mesure d'Alpilles a montré une différence significative sur l'estimation des flux convectifs.

Nous avons eu besoin du modèle adjoint de PBLs pour la création du système d'assimilation. Le logiciel TAPENADE nous a permis d'obtenir assez rapidement un adjoint valide et performant après avoir apporté certaines modifications au code de calcul de PBLs. L'utilisation du logiciel PALM pour l'assemblage du système d'assimilation a permis de le rendre flexible grâce à une implémentation modulaire à partir de composants logiciels indépendants (modèle PBLs, adjoint PBLs, minimiseur). C'est un point particulièrement intéressant pour l'évolution future de ce travail si l'on doit ajouter de nouveaux composants au modèle ou prendre en compte d'autres observations et notamment dans le cadre de l'étude de la modélisation du paysage dans le laboratoire.

L'étude sur les données synthétiques nous a donné un premier aperçu des performances du système d'assimilation en vue d'une application opérationnelle. Il ressort de ces tests que la méthode d'assimilation apporte une meilleure réponse sur l'estimation des flux de surface et du microclimat que les simulations directes du modèle PBLs.

Nous avons aussi vu que l'impact d'erreur des autres paramètres d'entrée (non corrigés par le système d'assimilation) n'est pas négligeable sur la précision que nous pouvons attendre de l'estimation de l'évapotranspiration. En particulier, la dégradation de l'indice foliaire *LAI* se traduit par une importante dispersion des flux de chaleur latente et sensible. Certes, la méthode d'assimilation permet de corriger certaines de ces erreurs selon la qualité des observations mais l'information apportée par les observations dans le système d'assimilation se révèle parfois insuffisante pour réduire l'ensemble de ces erreurs.

Il apparaît dans l'analyse sur les données synthétiques, une nette supériorité de la méthode d'assimilation par rapport à la méthode résiduelle, en particulier lorsque la précision des mesures de température est limitée. Cela est en grande partie lié à la capacité de la méthode d'assimilation de combiner l'information provenant du modèle, des observations et de

l'information a priori et d'être ainsi moins sensible aux erreurs des observations que la méthode résiduelle.

L'application du système d'assimilation sur les données réelles d'Alpilles montre que la méthode pourrait être un outil de suivi ou de diagnostic de l'état hydrique des cultures. Nous avons testé dans ce travail deux approches permettant de définir l'information a priori et les statistiques d'erreur associées. Les résultats sur les parcelles de mesure de la zone Alpilles ont montré qu'il est possible d'obtenir des estimations de bonne qualité. Ces résultats obtenus sur les quelques dates de l'expérimentation d'Alpilles ReSeDA constituent à ce stade une première évaluation de l'application opérationnelle.

### *Perspectives*

Le système d'assimilation mis en place a montré des résultats assez positifs et surtout prometteurs. Nous avons vu qu'il est réalisable et pourra être employé de façon régulière et ainsi permettre une optimisation pour le suivi du développement des cultures.

Cependant, les résultats obtenus dans ce travail restent encore insuffisants pour une application à grande échelle. Il est donc nécessaire de continuer les tests de validation sur les autres dates de l'expérimentation d'Alpilles.

Le bon fonctionnement d'un système d'assimilation est lié à la quantité et à la qualité des informations qu'on lui fournit, i.e. à celles du modèle, des observations et des informations a priori, ainsi qu'à la description adéquate des statistiques d'erreur de ces informations.

Parmi les pistes à considérer pour la suite de ce travail, l'élaboration de stratégies pour définir l'information a priori et les statistiques d'erreur semble importante. Nous avons vu que l'approche actuelle adoptée se révèle parfois limitée car elle fournit dans certains cas soit des valeurs d'informations a priori pas très adaptées soit des statistiques d'erreur trop faibles.

Concernant la qualité du modèle, il faut souligner ici que le modèle PBLs contient une représentation assez simplifiée de l'atmosphère. Il ne peut donc pas représenter toutes les conditions atmosphériques. Des modèles 1D plus élaborés type MOET pourraient être proposés mais l'implémentation des méthodes variationnelles est alors plus complexe, donc il y a un compromis entre représentation plus réaliste et simplicité pour une application opérationnelle. Par ailleurs, une autre piste de recherche serait d'analyser plus en détail l'erreur liée à l'indice

foliaire, qui est assez importante sur l'estimation des flux de chaleur. Estimer le *LAI* avec le système d'assimilation ou coupler le système à un modèle de transfert radiatif pour améliorer l'estimation des flux de surface pourrait être une piste à explorer.

Enfin, on pourrait envisager également l'introduction de plus d'observations pour ajuster mieux les paramètres d'intérêt (assimiler des températures de surface acquises à différentes heures de la journée, telles que celle fournies par Météosat). Cependant, ceci suppose de travailler à des échelles spatiales plus larges que le paysage agricole, ou bien d'utiliser des méthodes de désagrégation.

## **BIBLIOGRAPHIE**

Aboudrare, A. (2000). Stratégies de stockage et d'utilisation de l'eau pour le tournesol pluvial dans la région de Meknes : Thèse de Doctorat. In (p. 175): Institut Agronomique et Vétérinaire Hassan II

Albertson, J.D., Scanlon, T.M., Cahill, A.T., & Kustas, W.P. (2001). Estimating surface energy fluxes with remotely sensed data. In, *International Symposium on Remote Sensing and Hydrology 2000* (pp. 145-150). SANTAFE, NM

Allen, R.G., Pereira, L.S., Raes, D., & Smith, M. (1998). Crop evapotranspiration: guidelines for computing crop water requirements. *FAO Irrigation and Drainage Paper* (p. xxvi + 300)

Allen, R.G., Tasumi, M., & Morse, A. (2005a). Satellite-based evapotranspiration by METRIC and Landsat for western states water management. In, *the US Bureau of Reclamation Evapotranspiration Workshop*. Ft. Collins, CO, USA

Allen, R.G., Tasumi, M., Morse, A., & Trezza, R. (2005b). A Landsat-based energy balance and evapotranspiration model in Western US water rights regulation and planning. *Irrigation and Drainage Systems*, 19, 251-268

Allen, R.G., Tasumi, M., & Trezza, R. (2007). Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-model. *J. Irrig. and Drain. Eng.*, 133, 380-394

Ament, F., & Simmer, C. (2006). Improved representation of land-surface heterogeneity in a non-hydrostatic numerical weather prediction model. *Bound. Layer Meteor.*, 121, 153-174

Anderson, M.C., Norman, J.M., Diak, G.R., Kustas, W.P., & Mecikalski, J.R. (1997). A two-source time integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sensing Environment*, 60, 195-216

André, J.C. (1980). Simulation numérique détaillée de la couche limite atmosphérique, comparaison avec la situation des 2 et 4 juillet 1977 à Voves. *La Météorologie*, 22, 6-49

Asrar, G., Fuchs, M., Kanemasu, E.T., & Hatfield, J.L. (1984). Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. *Agronomy Journal*, 76, 300-306

Avissar, R., Avissar, P., Mahrer, Y., & Bravdo, B. (1985). A Model to simulate response of plant stomata to environmental conditions. *Agricultural and Forest Meteorology*, 34, 21-29

Avissar, R., & Pielke, R.A. (1989). A Parameterization of Heterogeneous Land Surfaces for Atmospheric Numerical Models and Its Impact on Regional Meteorology. *Monthly Weather Review*, 117, 2113-2136

Baccour, C., Baret, F., Béal, D., Weiss, M., & Pavageau, K. (2006). Neural network estimation of LAI, fAPAR, fCover and LAIxCab, from top of canopy MERIS reflectance data: Principles and validation. *Remote Sensing Environment*, 105, 313-325

Baret, F. (2002). Multi-sensor and multi-temporal remote sensing observations to characterise canopy functioning. The ReSeDA projet. *Agronomie*, 22, 165

Baret, F., Clevers, J.G.P.W., & Steven, M.D. (1995). The robustness of canopy gap fraction estimates from red and near-infrared reflectances: a comparison of approaches. *Remote Sensing Environment*, 54, 141-151

Baret, F., & Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and PAR assessment. *Remote Sensing of Environment*. *Remote Sensing Environment*, 35, 161-173

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J.L., & LEROY, M. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm. *Remote Sensing Environment*, 110, 275-286

Bastiaanssen, W.G.M., & Bos, M.G. (1999). Irrigation performance indicators based on remotely sensed data: a review of literature. *Irrigation and Drainage Systems*, 13, 291-311

Bastiaanssen W.G.M., Miltenburg I., Evans R., Molly R., Bastiaanssen F. & Van Der Pol E. (2009). An operational satellite-based irrigation monitoring and scheduling tool for saving water in irrigation Presented at the Irrigation and Drainage

Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., & Holtslag, A.A. (2001). A remote sensing surface energy balance algorithm for land (SEBAL). *Journal of Hydrology*, 212-213, 198-212

Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., & Holtslag, A.A.M. (1998a). A remote sensing surface energy balance algorithm for land (SEBAL): 1. Formulation. *Journal of Hydrology*, 212-213, 198-212

Bastiaanssen, W.G.M., Noordman, E.J.M., Pelgrum, H., Davids, G., & Allen, R.G. (2005). SEBAL for spatially distributed ET under actual management and growing conditions. *ASCE Journal of Irrigation and Drainage Engineering*, 131, 85-93

Bastiaanssen, W.G.M., Pelgrum, H., Wang, H., Ma, Y., Moreno, J.F., & Roerink, G.J. (1998b). A remote sensing surface energy balance algorithm for land (SEBAL): 2. Validation. *Journal of Hydrology*, 212-213, 213-229

Bischof, C., Carle, A., Hovland, P., Peyvand, K., & Mauer, A. (1995). ADIFOR 2.0 User's Guide. In. <http://www.mcs.anl.gov/research/projects/adifor/>

Blümel, K. (1999). A simple formula for estimation of the roughness length for heat transfer over partly vegetated surfaces. *Journal Applied Meteor.*, 38, 814-829

Blyth, E.M. (1995). Using a simple SVAT scheme to describe the effect of scale on aggregation. *Bound. Layer Meteor.*, 72, 267-285

Blyth, E.M., Dolman, A.J., & Wood, N. (1993). Effective Resistance to Sensible-and Latent-Heat Flux in Heterogeneous Terrain. *Q. J. R. Meteorol. Soc.*, 119, 423-442

Boni, G., Entekhabi, D., & Castelli, F. (2001). Land data assimilation with satellite measurements for the estimation of surface energy balance components and surface control on evaporation. *water resources research*, 37, 1713-1722

Bouttier, F., & Courtier, P. (1999). Data assimilation concepts and methods. In, *ECMWF Meteorological Training Course Lecture Series* (pp. 1-58). Reading, England

Bouysse F., Cassé V., Pailleux J. (1999). Variational surface analysis from screen level atmospheric parameters. *Tellus A vol. 51, (4) 453-468*

Braud, I., & Chanzy, A. (2000). Soil properties, initial and boundary conditions for use within SVAT models in the framework of the Alpilles-ReSeDA intercomparison. In, *Note Alpilles 13 July 2000* (p. 43). <http://avignon.inra.fr/reseda>: LTHE, Grenoble, INRA-CSE, Avignon

Braud, I., Dantas Antonio, A.C., Vauclin, M., Thony, J.L., & Ruelle, P. (1995). A Simple Soil Plant Atmosphere Transfer model (SisPAT): Development and field verification. *Journal of Hydrology*, 166, 213-250

Braud, I., Noilhan, J., Bessemoulin, P., Mascart, P., Haverkamp, R., & Vauclin, M. (1993). Bare-Ground surface heat and water exchanges under dry conditions - Observations and Parameterization. *Bound. Layer Meteor.*, 66, 173-200

Braud, I., Varado, N., & Olioso, A. (2005). Comparison of root water uptake modules using either the surface energy balance or potential transpiration. *Journal of Hydrology*, 301, 267-286

Brisson, N., Launay, M., Mary, B., & Beaudoin, N. (2008). Conceptual basis, formalisations and parameterization of the STICS crop model. (p. 297)

Brown, K.W., & Rosenberg, N.J. (1973). A resistance model to predict evapotranspiration and its application to a sugar beet field. *Agronomy Journal*, 65, 341-347

Bruckler, L., & Witono, H. (1989). Use of remotely sensed soil moisture content as boundary conditions in soil atmosphere water transport modelling. 2. Estimating soil water balance. *water resources research*, 25, 2437-2447

Brunet, Y., Nunez, M., & Lagouarde, J.P. (1991). A simple method for estimating regional evapotranspiration from infrared surface-temperature data. *ISPR. J. Photogrammetry Remote Sens.*, 46, 311-327

Brunet, Y., Pradel, E., & Camenen, L. (1995). Introduction aux transferts turbulents. *Actes de l'école chercheurs en Bioclimatologie INRA, Tome 1* (pp. 405-423)

Brutsaert, W. (1982). *Evaporation into the atmosphere: theory, history, and applications*: Reidel publishing company. London, England (pp. 299)

Bsaibes, A., Courault, D., Baret, F., Weiss, M., Olioso, A., Jacob, F., Hagolle, O., Marloie, O., Bertrand, N., Desfond, V., & Kzemipour, F. (2009). Albedo and LAI estimates from FORMOSAT-2 data for crop monitoring. *Remote Sensing of Environment. Remote Sensing Environment*

Buis, S., Piacentini, A., & Déclat, D. (2006). PALM: a computational framework for assembling high-performance computing applications. *Concurrency and Computation: Practice and Experience*, 18, 231-245

Businger, J.A., Wyngaard, J.C., Izumi, Y., & Bradley, E.F. (1971). Flux-Profile relationships in the atmospheric surface layer. *J. Atmos. Sci.*, 28, 181-189

Calvet, J.C., Noilhan, J., Roujean, J.L., Bessemoulin, P., Cabelguenne, M., Olioso, A., & Wigneron, J.P. (1998). An interactive vegetation SVAT model tested against data from six contrasting sites. *Agricultural and Forest Meteorology*, 92, 73-95

Calvet, J.C., Noilhan, J., & Wigneron, J.P. (2001). Root-zone soil moisture analysis using microwave radiometry. In ISPRS-CNES (Ed.), *Proceedings du 8ème Symposium International Mesures Physiques et Signatures en Télédétection* (pp. 629-637). Aussois, France

Caparrini, F., Castelli, F., & Entekhabi, D. (2004). Estimation of Surface Turbulent Fluxes through Assimilation of Radiometric Surface Temperature Sequences. *Journal of Hydrometeorology*, 5, 145-159

Carlson, T.N., & Buffum, M.J. (1989). On estimating total daily evapotranspiration from remote surface temperature measurements. *Remote Sensing Environment*, 29, 197-207

Carlson, T.N., & Lynn, B. (1991). The effects of plant water storage on transpiration and radiometric surface temperature. *Agricultural and Forest Meteorology*, 57, 171-186

Carmichael, G.R., Sandu, A., Chai, T., Daescu, D.N., Constantinescu, E.M., & Tang, Y. (2008). Predicting air quality: Improvements through advanced methods to integrate models and measurements. *Journal of computational physics*, 227, 3540 – 3571

Cayrol, P., Kergoat, L., Moulin, S., Dedieu, G., & Chehbouni, A. (2000). Calibrating a coupled SVAT / vegetation growth model with remotely sensed reflectance and surface temperature. A case study for the HAPEX-Sahel grassland sites. *Journal of applied meteorology*, 39, 2452-2472

Cellier, P., Richard, G., & Robin, P. (1996). Partition of sensible heat fluxes into bare soil and the atmosphere. *Agricultural and Forest Meteorology*, 82, 245-265

Chamberlain, A.C. (1968). Transport of gases to and from surfaces with bluff and wave-like roughness elements. *Q. J. R. Meteorol. Soc.*, 94, 318-332

Chanzy, A. (1991). Modélisation simplifiée de l'évaporation d'un sol nu utilisant l'humidité et la température de surface accessibles par télédétection. Thèse de doctorat INRA. In (p. 221): Inst. Nat. Agronomique de Paris-Grignon

Chanzy, A., & Bruckler, L. (1993). Significance of soil surface moisture with respect to daily bare soil evaporation. *water resources research*, 29, 1113-1125

Chanzy, A., Mumen, M., & Richard, G. (2008). Accuracy of top soil moisture simulation using a mechanistic model with limited soil characterization. *water resources research*, 44, w03432

Chebouni, A., Escadafal, R., Duchemin, B., Boulet, G., Simonneaux, V., Dedieu, G., Mougnot, B., Khabba, S., Kharrou, H., Maisongrande, P., Merlin, O., Chaponnière, A., Ezzahar, J., Er-raki, S., Hoedjes, J., Hadria, R., Abourida, A., Cheggour, A., Raibi, F., Boudhar, A., Benhadj, I., Hanich, L., Benkaddour, A., Guemouria, N., Chehbouni, A., Lahrouni, A., Oliosio, A., Jacob, F., Williams, D.G., & Sobrino, J.A. (2008). An integrated modeling and remote sensing approach for hydrological study in arid and semi-arid regions: the SUDMED Program. *International Journal of Remote Sensing*, 29, 5161-5151-5181

Choisnel, E. (1977). Le bilan d'énergie et le bilan hydrique du sol, n° spécial: "Evapotranspiration". *La Météorologie*, 11, 103-133

Choudhury, B.J. (1983). Simulating the effects of weather variables and soil water potential on a corn canopy temperature. *Agric. Meteorol.*, 29, 169-182

Choudhury, B.J., Idso, S.B., & Reginato, R.J. (1987). Analysis of an empirical model for soil heat flux under a growing wheat crop for estimating evaporation by an infrared-temperature based energy balance equation. *Agricultural and Forest Meteorology*, 39, 283-297

Choudhury, B.J., & Monteith, J.L. (1988). A four layer model for the heat budget of homogeneous land surfaces. *Q. J. R. Meteorol. Soc.*, 114, 373-398

Clapp, R., & Hornberger, G. (1978). Empirical equations for some hydraulic properties. *water resources research*, 14, 601-604

Clevers, J.G.P.W., Vonder, O.W., Jongschaap, R.E.E., Desprats, J.-F., King, C., Prevoit, L., & BRUGUIER, N. (2002). Using SPOT data for calibrating a wheat growth model under Mediterranean conditions *Agronomie*, 22, 687-694

Clothier, B.E., Clawson, K.L., Pinter, P.J.J., Moran, M.S., Reginato, R.J., & Jackson, R.D. (1986). Estimation of soil heat flux from net radiation during the growth of alfalfa. *Agricultural and Forest Meteorology*, 37, 319-329

Collin, J. (2006). Apport de la télédétection optique à la définition d'indicateurs de performance pour l'utilisation de l'eau en agriculture. Thèse de doctorat. In (p. 137): Université de Strasbourg

Coudert, B., & Ottlé, C. (2007). An improved SVAT model calibration strategy based on the optimisation of surface temperature temporal dynamics. *Geophysical Research Letters*, 34, L04402

Coudert, B., Ottlé, C., Boudevillain, B., Demarty, J., & Guillevic, P. (2006). Contribution of thermal infrared remote sensing data in multiobjective calibration of a dual-source SVAT model. *Journal of Hydrometeorology*, 7, 404-420

Courault, D., Bsaibes, A., Kpemlie, E., Hadria, R., Hagolle, O., Marloie, O., Hanocq, J.F., Olioso, A., Bertrand, N., & Desfonds, V. (2008). Assessing the Potentialities of FORMOSAT-2 Data for Water and Crop Monitoring at Small Regional Scale in South-Eastern France. *Sensors*, 8, 3460-3481

Courault, D., Clastre, P., Guinot, J.P., & Seguin, B. (1994). Analyse des sécheresses de 1988 à 1990 en France à partir de l'analyse combinée de données satellitaires NOAA-AVHRR et d'un modèle agrométéorologique. *Agronomie*, 14, 41-56

Courault, D., Drobinski, P., Brunet, Y., Lacarrère, P., & Talbot, C. (2007). Impact of surface heterogeneity on a buoyancy-driven convective boundary layer in light winds. *Bound. Layer Meteor.*, 124, 383-403

Courault, D., Jacob, F., Benoit, V., Weiss, M., Marloie, O., Hanocq, J.F., Fillol, E., Olioso, A., Dedieu, G., Gouaux, P., Gay, M., & French, A.N. (2009a). Influence of agricultural practices on micrometeorological spatial variations at local and regional scales. *International Journal of Remote Sensing*, 30, 1183-1205

Courault, D., Kpemlie, E., Hadria, R., Bsaibes, A., Buis, S., Olioso, A., & Hagolle, O. (2009b). Data assimilation: effect of agricultural land-use modifications on surface fluxes and microclimate. In, *iLEAPS Newsletter Issue N° 7*: <http://www.ileaps.org/index.php> (pp. 24-25)

Courault, D., Lacarrère, P., Clastre, P., Lecharpentier, P., Jacob, F., Marloie, O., Prévot, L., & Olioso, A. (2003). Estimation of surface fluxes in a small agricultural area using the threedimensional atmospheric model Meso-NH and remote sensing data. *Canadian journal of remote Sensing*, 29, 741-754

Courault, D., Lagouarde, J.P., & Aloui, B. (1996). Evaporation for maritime catchment combining a meteorological model with vegetation information and airborne surface temperatures. *Agricultural and Forest Meteorology*, 82, 93-117

Courault, D., Olioso, A., Lagouarde, J.P., Monestiez, P., & Allard, D. (2004). Influence des cultures sur les variables climatiques. ECOSPACE, AIP Organisation spatiale des activités agricoles et processus environnementaux. In P. Monestiez, S. Lardon & B. Seguin (Eds.), *Collection Science Update, INRA Editions* (pp. 303-320). PARIS

Courault, D., Seguin, B., & Olioso, A. (2005). Review on estimation of evapotranspiration from remote sensing data: from empirical to numerical modelling approaches. *Irrigation and Drainage Systems*, 19, 223-249

Courtier, P., & Talagrand, O. (1990). Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus*, 42A, 531-549

Crow, W.T., & Kustas, W.P. (2005). Utility of assimilating surface radiometric temperature observations for evaporative fraction and heat transfer coefficient retrieval. *Bound. Layer Meteor.*, 115, 105-130

De Moor, G. (1978). Les théories de la turbulence dans la couche limite atmosphérique. In, *Support de cours. Établissement d'études et de recherches* (pp. 65-100)

Deardoff, J.W. (1978). Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. *Journal of Geophysical Research*, 83, 1889-1903

Del Moral, J.C., Noyer, G.R., & Salut, G. (1995). Résolution particulière et traitement non-linéaire du signal : Applications radar/sonar. *Traitement du signal*, 12, 287-301

Demarty, J., Otle, C., Braud, I., Olioso, A., Frangi, J.P., Bastidas, L.A., & Gupta, H.V. (2004). Using a multiobjective approach to retrieve information on surface properties used in a SVAT model *Journal of Hydrology*, 287, 214-236

Demarty, J., Otlé, C., Braud, I., Olioso, A., Frangi, J.P., Gupta, H.V., & Bastidas, L.A. (2005). Constraining a physically based soil-vegetation-atmosphere transfer model with surface

water content and thermal infrared brightness temperature measurements using a multiobjective approach. *water resources research*, 41, W01011

Demarty, J., Ottlé, C., François, C., Braud, I., & Frangi, J.P. (2002). Effect of aerodynamic resistance modeling on SiSPAT-RS simulated surface fluxes. *Agronomie*, 22, 641-650

Deschamps, P.Y., & Phulpin, T. (1980). Atmospheric correction of infrared measurements of sea surface temperature using channels at 3.7, 11 and 12 m. *Bound. Layer Meteor.*, 18, 131-143

Dickinson, R.E. (1984). Modeling Evapotranspiration for Three-Dimensional Global Climate Models. In J.E. Hanson & T. Takahashi (Eds.), *Climate Processes and Climate Sensitivity* (pp. 58-72): Amer. Geophys. Union, Geophys. Monogr.

Dickinson, R.E., Henderson-Sellers, A., Kennedy, P.J., & Wilson, M.F. (1986). Biosphere-atmosphere transfer scheme (BATS) for the NCAR Community Climate Model. In, *Technical Note TN 275 TSTR, NCAR, Boulder, CO*.

Doucet, A., Andrieu, C., & Godsill, S. (2000). On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. *Statistics and Computing*, 10, 197-208

Droogers, P., & Bastiaanssen, W.G.M. (2002). Irrigation performance using hydrological and remote sensing modeling. *Journal of Irrigation and Drainage Engineering ASCE*, 128, 11-18

Duchemin, B., Hadria, R., Er-raki, S., Ezzahar, J., Boulet, G., Maisongrande, P., Chehbouni, A., Escadafal, R., Hoedjes, J., Karroui, H., Khabba, S., Mougenot, B., Olioso, A., Rodriguez, J.C., Simonneaux, V., & Timouk, F. (2006). Monitoring wheat phenology and irrigation in Center of Morocco: on the use of relationship between evapotranspiration, crops coefficients, leaf area index and remotely-sensed vegetation indices *Agricultural Water Management*, 79, 1-27

El Akkraoui, A., Gauthier, P., Pellerinc, S., & Buis, S. (2008). Intercomparison of the primal and dual formulations of variational data assimilation. *Q. J. R. Meteorol. Soc.*, 134, 1015-1025

Elizondo, D., Faure, C., & Cappelaere, B. (2000). Automatic- versus Manual-differentiation for non-linear inverse modelling. In, *Rapport de recherche n° 3981, INRIA, Sophia Antipolis* (p. 41). <ftp://ftp.inria.fr/INRIA/publication/publi-pdf/RR/RR-3981.pdf>

Evensen, G. (2003). The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53, 343-367

Fan, L., Liu, S., Bernhofer, C., Liu, H., & Berger, F.H. (2007). Regional land surface energy fluxes by satellite remote sensing in the Upper Xilin River Watershed (Inner Mongolia, China). *Theoretical and Applied Climatology*, 88, 231-245

Federer, C.A. (1979). A Soil Plant Atmosphere Model for Transpiration and Availability of Soil Water. *water resources research*, 15, 555-562

Findeling, A. (2001). Etude et modélisation de certains effets du semis direct avec paillis de résidus sur les bilans hydrique, thermique et azoté d'une culture de maïs pluvial au Mexique. Thèse de doctorat. In (p. 375): Ecole National du Génie Rural des Eaux et Forêts, Montpellier

Flerchinger, G.N., Kustas, W.P., & Wertz, M.A. (1998). Simulating surface energy fluxes and radiometric surface temperatures for two arid vegetation communities using the SHAW model. *Journal Applied Meteor.*, 37, 449-460

Fletcher, R. (2000). Practical methods of optimization. In S. Wiley (Ed.). New York

François, C., Cayrol, P., Kergoat, L., & Moulin, S. (2001). Assimilation techniques of remote sensing measurements into vegetation models: Overview, limits and promises. In, *Proceedings of the 8th International Symposium on Physical Measurements and Signatures in Remote Sensing* (pp. 649-658). Aussois, France

François, C., Ottlé, C., & Prévot, L. (1997). Analytical parameterization of canopy directional emissivity and directional radiance in the thermal infrared. Application on the retrieval of soil and foliage temperatures using two directional measurements. *International Journal of Remote Sensing*, 18, 2587-2621

French, A.N., Jacob, F., Anderson, M.C., Kustas, W.P., Timmermans, W., Gieske, A.S.M., Su, Z., Su, H., McCabe, M., Li, F., Prueger, J.H., & Brunsell, N. (2005). Surface energy fluxes with the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER) at the Iowa 2002 SMACEX site (USA). *Remote Sensing Environment*, 99, 55-65

Friedl, M.A. (2002). Forward and inverse modeling of land surface energy balance using surface temperature measurements. *Remote Sensing Environment*, 79, 344-354

Giering, R., & Kaminski, T. (1998). Recipes for adjoint code construction. *ACM Trans. Mathematical Software*, 4, 437-474

Gilbert, J.-C., & Lemaréchal, C. (1989). Some numerical experiments with variable storage quasi-Newton algorithms. *Mathematical Programming*, 45, 407-435

Gomez, M., Olioso, A., Sobrino, J.A., & Jacob, F. (2005). Retrieval of evapotranspiration over the Alpillles/ReSeDA experimental site using airborne POLDER sensor and a thermal camera. *Remote Sensing Environment*, 96, 399-408

Gowda, P.H., Chavez, J.L., Colaizzi, P.D., Evett, S.R., Howell, T.A., & Tolk, J.A. (2008). ET mapping for agricultural water management: present status and challenges. *Irrigation Science*, 26, 223-237

Hadria, R., Duchemin, B., Lahrouni, A., Khabba, S., Er-raki, S., Dedieu, G., Chehbouni, A., & Olioso, A. (2006). Monitoring of irrigated wheat in a semi-arid climate using crop modelling and remote sensing data: Impact of satellite revisit time frequency. *Int J. Remote Sens*, 27, 1093-1117

Hasager, C.B. (1997). Surface fluxes in heterogeneous landscape. In, *Ph.D. dissertation* (p. 177): Riso National Laboratory -R-922(EN)

Hasager, C.B., Jensen, N.O., & Olioso, A. (2002a). Land cover, surface temperature and leaf area index maps from satellites used for the aggregation of momentum and temperature roughnesses. In J.A. Sobrino (Ed.), *First International symposium on Recent Advances in Quantitative Remote Sensing* (pp. 466-473). University of Valencia, Spain

Hasager, C.B., Olioso, A., & Jacob, F. (2002b). Parametrisation of aggregated roughness and sensible heat flux from field scale to hydrological scale by microscale modelling in the Alpillles experiment in France. In, *EGS 2002*. Nice, France 21-27 April 2002: Geophys. Res. Abstr. (CD-ROM)

Hascoët, L., & Pascual, V. (2004). Rt-0300 - tapenade 2.1 user's guide. In, *Rapport technique, INRIA, Sophia Antipolis* (pp. 1-78). <http://tapenade.inria.fr:8080/tapenade/index.jsp>

Hoedts, T. (2005). Simulation des variations microclimatiques d'une région agricole à l'aide d'un modèle 1D patché. In INRA-EMMAH (Ed.), *Rapport de stage de fin d'études: Ing. Calcul Scientifique* (pp. 1-55): ISITV, Université Toulon-Var

Idso, S.B., Jackson, R.D., Pinter, P.J.J., Reginato, R.J., & Hatfield, J.L. (1981). Normalizing the stress-degree-day parameter or environmental variability. *Agric. Meteorol.*, 24, 24-45

Inoue, Y., & Olioso, A. (2006). Estimating the dynamics of ecosystem CO<sub>2</sub> flux and biomass production in agricultural fields on the basis of synergy between process models and remotely sensed signatures. *Journal of Geophysical Research*, 111, D24, D24S91

Intsiful, J., & Kunstmann, H. (2008). Upscaling of Land-Surface Parameters Through Inverse Stochastic SVAT-Modelling *Bound. Layer Meteor.*, 129, 137-158

Itier, B., & Riou, C. (1982). Une nouvelle méthode de détermination de l'évapotranspiration réelle par thermographie infrarouge. *J. Rech. Atmos.*, 16, 113-125

Jackson, R.D., Kustas, W.P., & Choudhury, B.J. (1988). A reexamination of the crop water stress index. *Irrigation Science*, 9, 309-317

Jackson, R.D., Moran, M.S., Gay, L.W., & Raymond, L.H. (1987). Evaluating evaporation from field crops using airborne radiometry and ground based meteorological data. *Irrigation Science*, 8, 81-90

Jackson, R.D., Reginato, R.J., & Idso, S.B. (1977). Wheat canopy temperature: a practical tool for evaluating water requirements. *water resources research*, 13, 651-656

Jacob, F. (1999). Utilisation de la télédétection courtes longueur d'onde et infrarouge thermique à haute résolution spatiale pour l'estimation des flux d'énergie à l'échelle de la parcelle agricole. Thèse de doctorat. In (p. 250): Université Paul Sabatier Toulouse 3

Jacob, F., Gu, X.F., Hanocq, J.F., Tallet, N., & Baret, F. (2003). Atmospheric corrections of single channel and multidirectional airborne thermal infrared data : application to the ReSeDA Experiment. *International Journal of Remote Sensing*, 24, 3269-3290

Jacob, F., & Olioso, A. (2005). Derivation of diurnal courses of albedo and reflected solar irradiance from airborne POLDER data acquired near solar noon. *Journal of Geophysical Research*, 110, D10104

Jacob, F., Olioso, A., Gu, X.F., Su, Z., & Seguin, B. (2002a). Mapping surface fluxes using airborne visible, near infrared, thermal infrared remote sensing data and a spatialized surface energy balance model. *Agronomie*, 22, 669-680

Jacob, F., Olioso, A., Weiss, M., Baret, F., & Hautecoeur, O. (2002b). Mapping short-wave albedo of agricultural surfaces using airborne POLDER data. *Remote Sensing Environment*, 80, 36-46

Jacob, F., Schmugge, T., Olioso, A., French, A.N., Courault, D., Ogawa, K., Petitcolin, F., Chehbouni, A., Pinheiro, A., & Privette, F. (2008). Modeling and inversion in thermal infrared remote sensing over vegetated land surfaces. In S. Liand (Ed.), *In Advances in Land Remote Sensing: System, Modeling, Inversion and Application* (pp. 245-292): Springer

Jacob, F., Weiss, M., Olioso, A., & French, A.N. (2002c). Assessing the narrowband to broadband conversion to estimate visible, near infrared and shortwave apparent albedo from airborne POLDER data. *Agronomie*, 22, 537-546

Jacobs, C., & De Bruin, H. (1997). Predicting regional transpiration at elevated atmospheric CO<sub>2</sub>: Influence of the PBL-vegetation interaction. *Journal of applied meteorology*, 36, 1663-1675

Jacquemin, B., & Noilhan, J. (1990). Sensitivity study and validation of land surface parametrization using the Hapex-Mobilhy data set. *Bound. Layer Meteor.*, 52, 93-134

Jarvis, P.G. (1976). The interpretation of the variations in Leaf Water Potential and stomatal conductance found in canopies in the field. *Phil. Trans. R. Soc.*, B273, 593-610

Jensen, N.O., & Hummelshøj, P. (1995). Derivation of canopy resistance for water vapour fluxes over a spruce forest, using a new technique for the viscous sublayer resistance. *Agricultural and Forest Meteorology*, 73, 339-352

Jensen, N.O., & Hummelshøj, P. (1997). Erratum to “Derivation of canopy resistance for water vapour fluxes over a spruce forest, using a new technique for the viscous sublayer resistance”. *Agricultural and Forest Meteorology*, 85, 289

Ji, X.B., Kang, E.S., Zhao, W.Z., Zhang, Z.H., & Jin, B.W. (2009). Simulation of heat and water transfer in a surface irrigated, cropped sandy soil. *Agricultural Water Management*, 96, 1010-1020

Jia, L. (2004). Modeling heat exchanges at the land-atmosphere interface using multi-angular thermal infrared measurements: Thesis. In (p. 199): Wageningen University, Netherlands

Jia, L., Su, Z., Van den Hurk, B.J.J.M., Menenti, M., Moene, A., De Bruin, H.A.R., Yrisarry, J.J.B., Ibanez, M., & Cuesta, A. (2003). Estimation of sensible heat flux using the Surface Energy Balance System (SEBS) and ATSR measurements. *Physics and Chemistry of the Earth*, 28, 75-88

Jones J. W. & Zur B. (1984). Simulation of Possible Adaptive Mechanisms in Crops Subjected to Water Stress. *Irrig Sciences* 5, 251-264

Judge, J., England, A.W., Metcalfe, J.R., McNichol, D., & Goodison, B.E. (2008). Calibration of an integrated land surface process and radiobrightness (LSP/R) model during summertime. *Advances in Water Resources*, 31, 189-202

Kalma, J.D., & Jupp, D.L.B. (1990). Estimating evaporation from pasture using infrared thermometry: evaluation of one layer resistance model. *Agricultural and Forest Meteorology*, 51, 223-246

Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *J. Bas. Eng.*, 83, 35-45

Kerr Y. H., Lagouarde J.P., Imbernon J. (1992). Accurate land surface temperature retrieval from AVHRR data with use of an improved split window algorithm. *Rem. Sen. Envir.*, 41, 197-209

Kim, H., & Patel, V. (2000). Test of turbulence models for wind flow over terrain with separation and recirculation. *Bound. Layer Meteor.*, 94, 5-21

Kimura, F., & Shimizu, Y. (1994). Estimation of Sensible and Latent Heat Fluxes from Soil Surface Temperature using a Linear Air-Land Heat Transfer Model. *Journal Applied Meteor.*, 33, 477-489

Kpemplie, E. (2006). Application d'un algorithme d'assimilation variationnelle pour inverser un modèle 1D de transfert atmosphérique à partir de données thermiques. In INRA-EMMAH (Ed.), *Rapport de stage de fin d'études: Ing. Calcul Scientifique* (pp. 1-113): ISITV, Université Toulon-Var

Kumar, S.V., Reichle, R.H., & Peters-Lidard, C.D. (2008). A land surface data assimilation framework using the land information system: Description and applications *Adv. Water Resour.*, 31, 1419-1432

Kustas, W.P., & Albertson, J.D. (2003). Effects of surface temperature contrast on land-atmosphere exchange: A case study from Monsoon 90. *water resources research*, 39, 1159

Kustas, W.P., Anderson, M.C., Norman, J., & Li, F. (2007). Utility of radiometric-aerodynamic temperature relations for heat flux estimation. *Bound. Layer Meteor.*, 122, 167-187

Kustas, W.P., Choudhury, B.J., Moran, M.S., Reginato, R.J., Jackson, R.D., Gay, L.W., & Weaver, H.L. (1989). Determination of sensible heat flux over sparse canopy using thermal infrared data. *Agricultural and Forest Meteorology*, 44, 197-216

Kustas, W.P., & Daughtry, C.S.T. (1990). Estimation of the soil heat flux/ net radiation ratio from spectral data. *Agricultural and Forest Meteorology*, 49, 205-223

Kustas, W.P., Daughtry, C.S.T., & Oevelen, P.J.V. (1993). Analytical Treatment of the Relationships between Soil Heat Flux/Net Radiation Ratio and Vegetation Indices. *Remote Sensing Environment*, 46, 319-330

Kustas, W.P., & Googrich, D.C. (1994). Preface to Monsoon 90 experiment. *water resources research*, 30, 205-233

Kustas, W.P., Jackson, T.J., Prueger, J.H., Hatfield, J.L., & Anderson, M.C. (2003). Remote sensing field experiments for evaluating soil moisture retrieval algorithms and modeling land-atmosphere dynamics in central Iowa. *EOS - Transactions of the American Geophysical Union*, 84 45, 485-493

Kustas, W.P., Moran, M.S., Humes, K., Stannard, I., Pinter, P.J.J., Hipps, L.E., Swiatec, E., & Googrich, D.C. (1994). Calculation of basin-scale surface fluxes by combining remotely sensed data and atmospheric properties in a semiarid landscape. *water resources research*, 30, 1241-1259

Kustas, W.P., Moran, M.S., Jackson, R.D., Gay, L.W., Duell, L.F.W., Kunkel, K.E., & Matthias, A.D. (1990). Instantaneous and daily values of the surface energy balance over agricultural fields using remote sensing and a reference field in an arid environment. *Remote Sensing Environment*, 32, 125-141

Lagarde, T., Piacentini, A., & Thual, O. (2001). A new representation of data assimilation methods: The PALM flow charting approach. *Quart. J. Roy. Meteor. Soc.*, 127, 189-207

Lagouarde, J.P. (1991). Use of NOAA-AVHRR data combined with an agrometeorological model for evaporation mapping. *Int J. Remote Sens*, 12, 1853-1864

Lagouarde, J.P., & Brunet, Y. (1989). Spatial integration of surface latent heat flux and evaporation mapping. *advanced space Research*, 9, 259-264

Lagouarde, J.P., & Brunet, Y. (1994). Un modele simple d'evapotranspiration d'une foret de pin maritime. In, *Atelier de Modélisation de l'Atmosphère* (pp. 21-28). Toulouse (France)

Lagouarde, J.P., & MacAnaney, K.J. (1992). Daily sensible heat flux estimation from a single measurement of surface temperature and maximum air temperature. *Bound. Layer Meteor.*, 59, 341-362

Lauvernet, C. (2005). Assimilation variationnelle d'observations de télédétection dans les modèles de fonctionnement de la végétation : utilisation du modèle adjoint et prise en compte de contraintes spatiales: Thèse de Doctorat. In. <http://tel.ccsd.cnrs.fr/index.php>: Université Joseph Fourier Grenoble 1

Lauvernet, C., Baret, F., Hascoët, L., Buis, S., & Le Dimet, F.X. (2008). Multitemporal-patch ensemble inversion of coupled surface-atmosphere radiative transfer models for land surface characterization. *Remote Sensing Environment*, 112, 851-861

Lauvernet, C., Le Dimet, F.X., Baret, F., Deboissezon, H., Favard, J.C., Vintila, R., Lazar, C., & Badea, A. (2002). Assimilation of high temporal frequency SPOT data to describe canopy functioning - The case of wheat crops in the ADAM experiment in Romania. In J.A. Sobrino (Ed.), *First International Symposium on Recent Advances in Quantitative Remote Sensing* (pp. 921-926). Publicacions de la Universitat de Valencia, España

Le Dimet, F.X., & Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, 38A, 97-110

Le Moigne, P. (2008). Surfex: Meso-NH training course. In, [http://mesonh.aero.obs-mip.fr/mesonh/dir\\_tutorial/lemoigne\\_surfex\\_oct2008.pdf](http://mesonh.aero.obs-mip.fr/mesonh/dir_tutorial/lemoigne_surfex_oct2008.pdf)

Lellouche, J.M., Tranchant, B., Nouel, L., Benkiran, M., & Grener, E. (2002). Operational data assimilation in the Mercator North Atlantic and Mediterranean High Resolution Prototype PAM. In, *AGU Fall Meeting* (p. C269+)

Lettau, H. (1951). Lettau H., 1951, Theory of surface temperature and heat-transfer oscillations near level ground surface. *EOS - Transactions of the American Geophysical Union*, 32, 189-200

Li, F., Kustas, W.P., Prueger, J.H., Neale, C.M.U., & Jackson, T.J. (2005). Utility of remote sensing based two-source energy balance model under low- and high-vegetation cover conditions. *Journal of Hydrometeorology*, 6, 878-891

Li, Z.L., Tang, R., Wan, Z., Bi, Y., Zhou, C., Tang, B., Yan, G., & Zhang, X. (2009). A Review of Current Methodologies for Regional Evapotranspiration Estimation from Remotely Sensed Data. *Sensors*, 9, 3801-3853

Liu, D.C., & Nocedal, J. (1989). On the limited memory BFGS method for large-scale optimization. *Mathematical Programming*, 45, 503-528

Liu, S., Lu, L., Mao, D., & Jia, L. (2007). Evaluating parameterizations of aerodynamic resistance to heat transfer using field measurements. *Hydrology and Earth System Sciences*, 11, 769-783

Lucht, W., Schaaf, C., & Strahler, A. (2000). An algorithm for the retrieval of albedo from space using semi-empirical BRDF models. *IEEE Trans. Geosci. Remote Sens.*, 38, 977-997

Mallet, I., & Payart, J. (2008). DOCUMENTATION STAGE MESO-NH. In: [http://mesonh.aero.obs-mip.fr/mesonh/dir\\_tutorial/doc\\_octo2008.pdf](http://mesonh.aero.obs-mip.fr/mesonh/dir_tutorial/doc_octo2008.pdf)

Margulis, S.A., & Entekhabi, D. (2003). Variational assimilation of radiometric surface temperature and reference-level micrometeorology into a model of the atmospheric boundary layer and land surface. *Monthly Weather Review*, 131, 1272-1288

Margulis, S.A., Kim, J., & Hogue, T. (2005). A comparison of the triangle retrieval and variational data assimilation methods for surface turbulent flux estimation. *Journal of Hydrometeorology*, 6, 1063-1072

Margulis, S.A., McLaughlin, D., Entekhabi, D., & Dunne, S. (2002). Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. *water resources research*, 38, 1-18

Massart, S., Buis, S., Erhard, P., & Gacon, G. (2007). Use of 3D-VAR and Kalman filter approaches for neutronic state and parameter estimation in nuclear reactors. *Nuclear Science and Engineering*, 155, 409-424

Massman, W.J. (1997). An analytical one-dimensional model of momentum transfer by vegetation of arbitrary structure. *Bound. Layer Meteor.*, 83, 407-421

Massman, W.J. (1999a). model study of kB-1 for vegetated surfaces using 'localized near-field' Lagrangian theory. *Journal Hydrology*, 223, 27-43

Massman, W.J. (1999b). Molecular diffusivities of Hg vapor in air, O<sub>2</sub> and N<sub>2</sub> near STP and the kinematic viscosity and the thermal diffusivity of air near STP. *Atmospheric Environment*, 33, 453-457

Mazauric, C. (2003). Assimilation de données pour les modèles d'hydraulique fluviale. Estimation de paramètres, analyse de sensibilité et décomposition de domaine. Thèse de Doctorat. In. <http://tel.ccsd.cnrs.fr/index.php>: Université Joseph Fourier - Grenoble 1

McLaughlin, D. (1995). Recent developments in hydrologic data assimilation. *Reviews of Geophysics*, 33, 977-984

McLaughlin, D., Zhou, Y., Entekhabi, D., & Chatdarong, V. (2006). Computational Issues for Large-Scale Land Surface Data Assimilation Problems. *Journal of Hydrometeorology*, 7, 494-510

Menenti, M., & Choudhury, B.J. (1993). Parametrization of land surface evapotranspiration using a location dependent potential evapotranspiration and surface temperature range. In, *Proceedings of IAHS conference on Land Surface Processes* (pp. 561-568). Yokohama

Menenti, M., & Ritchie, J.C. (1994). Estimation of effective aerodynamic roughness of Walnut Gulch watershed with laser altimeter measurements. *water resources research*, 30, 1329-1337

Milly, P.C.D. (1982). Moisture and heat transport in hysteretic inhomogeneous porous media: a matric head-based formulation and a numerical model. *water resources research*, 18, 489-498

Monin, A.S., & Obukhov, A.M. (1954). Basic laws of turbulent mixing in the ground layer of the atmosphere. *Akad. Nauk SSSR, Trudy Geofiz. Inst.*, 24, 163-187

Monteith, J.L. (1965). Evaporation and the environment. In F. GE (Ed.), *The state and movement of water in living organisms, 19th symposium of the society for experimental biology* (pp. 205-234). University Press, Cambridge

Monteith, J.L. (1973). *Principles of Environmental Physics*. (p. 241): Elsevier, New York

Monteith, J.L., & Unsworth, M.H. (1990). *Principles of environmental physics*. (p. 291): Edward Arnold Press, London

Moran, M. S., Jackson, R. D., Raymond, L. H., Gay, L. W., & Slater, P. N. (1989). Mapping surface energy balance components by combining Landsat thematic mapper and ground-based meteorological data. *Remote Sensing of Environment*, 30, 77-87

Moran, M.S., Kustas, W.P., Vidal, A., Stannard, I., Blanford, J.H., & Nichols, W.D. (1994). Use of ground-based remotely sensed data for surface energy balance evaluation of semiarid rangeland. *water resources research*, 30, 1339-1349

Morel, T., & group, T.P. (2006). Manuel de formation au logiciel PALM\_MP In, *CERFACS/ Global Change and Climate Modelling Team* (p. 72). Toulouse, France

Moulin, S., Kergoat, L., Cayrol, P., Dedieu, G., & Prévot, L. (2002). Calibration of a coupled canopy functioning and SVAT model in the ReSeDA experiment. Towards the assimilation of SPOT/HRV observations into the model. *Agronomie*, 22, 681-686

Mukabana, J.R., & Pielke, R.A. (1996). Investigating the Influence of Synoptic-Scale Monsoonal Winds and Mesoscale Circulations on Diurnal Weather Patterns over Kenya Using a Mesoscale Numerical Model. *Monthly Weather Review*, 124, 224-244

Mumen, M. (2006). Caractérisation du Fonctionnement Hydrique des Sols à l'aide d'un Modèle Mécaniste de Transferts d'Eau et de Chaleur mis en oeuvre en Fonction des Informations Disponibles dur le Sol. Thèse doctorat INRA. In (p. 169): Université d'Avignon et des pays de Vaucluse

Munoz-Sabater, J.M., Jarlan, L., Calvet, J.C., Bouyssel, F., & De Rosnay, P. (2007). From near-surface to root-zone soil moisture using different assimilation techniques. *Journal of Hydrometeorology*, 8, 194-206

Murray, T., & Verhoef, A. (2007). Moving towards a more mechanistic approach in the determination of soil heat flux from remote measurements: II. Diurnal shape of soil heat flux. *Agricultural and Forest Meteorology*, 147, 88-97

Naud, C., Makowski, D., & Jeuffroy, M.H. (2007). Application of an interacting particle filter to improve nitrogen nutrition index predictions for winter wheat. *Ecological Modelling*, 207, 251-263

Nieuwenhuis, G.J.A., Smidt, E.H., & Thunissen, H.A.M. (1985). Estimating of regional evapotranspiration of arable crops from thermal infrared images. *Int J. Remote Sens*, 6, 1319-1334

Noilhan, J., & Mahfouf, J.F. (1996). The ISBA land surface parameterisation scheme. *Glob. Planet. Chang.*, 13, 145-159

Noilhan, J., & Planton, S. (1989). A simple parametrization of land surface processes for meteorological models. *Monthly Weather Review*, 117, 536-549

Norman, J., Anderson, M.C., & Kustas, W.P. (2006). Are single-source, remote-sensing surface-flux models too simple? In, *Earth Observation for Vegetation Monitoring and Water Management* (pp. 170-177). Melville, New York, USA: American Institute of Physics

Norman, J., Kustas, W.P., & Humes, K. (1995). A two-source approach for estimating soil and vegetation energy fluxes from observattions of directional radiometric surface temperature. *Agricultural and Forest Meteorology*, 77, 263-293

Norman, J.M. (1979). Modeling the complete crop canopy. In, *Modification of the Aerial Environment of Plants (B. J. Barfield and J. F. Gerber, Eds.)* (pp. 249-277): American Society of Agricultural Engineering, St. Joseph, MI

Norman, J.M., Kustas, W.P., Prueger, J.H., & Diak, G.R. (2000). Surface flux estimation using radiometric temperature: A dual-temperature-difference method to minimize measurement errors. *water resources research*, 36, 2263-2274

Olioso, A. (1992). Simulation des échanges d'énergie et de masse d'un couvert végétal, dans le but de relier la transpiration et la photosynthèse aux mesures de réflectance et de température de surface. Thèse de doctorat. In (p. 264): Université Montpellier II, France

Olioso, A. (1994). Modélisation des échanges d'énergies et de masse d'un couvert végétal dans le but de relier la transpiration et la photosynthèse aux mesures de réflectance et de températures de surface. In, *Sixième Symposium International sur les Mesures Physiques et Signatures en Télédétection*. val d'Isère, France. CNES. p1135-1142

Olioso, A., Braud, I., Chanzy, A., Courault, D., Demarty, J., Kergoat, L., Lewan, E., Ottlé, C., Prévot, L., Zhao, W.G., Calvet, J.C., Cayrol, P., Jongscaap, R., Moulin, S., Noilhan, J., & Wigneron, J.P. (2002a). SVAT modeling over the Alpilles-ReSeDA experiment: comparing SVAT models over wheat fields. *Agronomie*, 22, 651-668

Olioso, A., Braud, I., Chanzy, A., Demarty, J., Ducros, Y., Gaudu, J.-C., Gonzalez-Sosa, E., Lewan, E., Marloie, O., Ottlé, C., Prévot, L., Thony, J.L., Autret, H., Bethenod, O., Bonnefond, J.M., Bruguier, N., Buis, J.P., Calvet, J.C., Caselles, V., Chauki, H., Coll, C., François, C., Goujet, R., Jongscaap, R., Kerr, Y., King, C., Lagouarde, J.P., Laurent, J.P., Lecharpentier, P., Mc Aneney, J., Moulin, S., Rubio, E., Weiss, M., & Wigneron, J.P. (2002b). Monitoring energy and mass transfers during the Alpilles-ReSeDA experiment. *Agronomie*, 22, 597-610

Olioso, A., Carlson, T.N., & Brisson, N. (1996a). Simulation of diurnal transpiration and photosynthesis of a water stressed soybean crop. *Agricultural and Forest Meteorology*, 81, 41-59

Olioso, A., Chauki, H., Courault, D., & Wigneron, J.P. (1999a). Estimation of evaptranspiration and photosynthesis by assimilation of remote sensing data into SVAT models. *Remote Sensing Environment*, 68, 341-356

Olioso, A., Chauki, H., Wigneron, J.P., Bergaoui, K., Bertuzzi, P., Chanzy, A., Bessemoulin, P., & Calvet, J.C. (1999b). Estimation of energy fluxes from thermal infrared, spectral reflectances, microwave data and SVAT modeling. *Physics and Chemistry of the Earth*, 24, 829-836

Olioso, A., Hasager, C.B., Jacob, F., Wassenaar, T., Chehbouni, A., Marloie, O., Lecharpentier, P., & Courault, D. (2002c). Mapping surface sensible heat flux from thermal infrared and reflectances data using various models over the Alpillés test site. In J.A. Sobrino (Ed.), *Proceedings of the First International Symposium on Recent Advances in Quantitative Remote Sensing, 16-20 September 2002* (pp. 450-457). Valencia, Spain: Universitat de València, España

Olioso, A., Inoue, Y., Demarty, J., Wigneron, J.P., Braud, I., Ortega-Farias, S., Lecharpentier, P., Otlé, C., Calvet, J.C., & Brisson, N. (2002d). Assimilation of remote sensing data into crop simulation models and SVAT models. In J.A. Sobrino (Ed.), *Proceedings of the First International Symposium on Recent Advances in Quantitative Remote Sensing, 16-20 September 2002* (pp. 329-338). Valencia, Spain: Universitat de València, España

Olioso, A., Inoue, Y., Ortega-Farias, S., Demarty, J., Wigneron, J.P., Braud, I., Jacob, F., Lecharpentier, P., Otlé, C., Calvet, J.C., & Brisson, N. (2005). Future directions for advanced evapotranspiration modeling: Assimilation of remote sensing data into crop simulation models and SVAT models. *Irrigation and Drainage Systems*, 19, 377-412

Olioso, A., & Jacob, F. (2002). Estimation de l'évapotranspiration à partir de mesures de télédétection. *La houille Blanche*, 1, 62-67

Olioso, A., Rivalland, V., Faivre, R., Weiss, M., Demarty, J., Wassenaar, T., Baret, F., Cardot, H., Rosselo, P., Jacob, F., Hasager, C.B., & Inoue, Y. (2006). Monitoring Evapotranspiration over the Alpillés Test Site by Introducing Remote Sensing Data at Various Spatial Resolutions into a Dynamic SVAT Model. In J.F. Moreno, M.A. Jochum Osann & G. D'Urso (Eds.), *AIP Conference proceedings 'Earth observation for vegetation monitoring and water management'* (pp. 234-241). Napoli, 24-25 Octobre 2005: American Institute of Physics

Olioso, A., Taconet, O., & BenMehrez, M. (1996b). Estimation of heat and mass fluxes from IR brightness temperature In, *5th IEEE International Geoscience and Remote Sensing Symposium (IGARSS 95) - Quantitative Remote Sensing for Science and Applications* (pp. 1184-1190). FLORENCE, ITALY, JUL 10-14, 1995: IEEE, Geoscience and Remote Sensing Soc

Opoku-Duah, S., Donoghue, D.N.M., & Burt, T.P. (2008). Intercomparison of evapotranspiration over the Savannah Volta Basin in West Africa using remote sensing data. *Sensors*, 8, 2736-2761

Ottlé, C., Richard, D., Le Hegarat, S., Guerin, C., Noilhan, J., Le Moigne, P., & Habets, F. (2002). Modélisation hydrométéorologique du bassin du Rhône: apport de la télédétection spatiale. *La houille Blanche*, 1, 57-61

Ottlé, C., & Vidal-Madjar, D. (1994). Assimilation of soil moisture inferred from infrared remote sensing in a hydrological model over the HAPEX-MOBILHY region. *Journal of Hydrology*, 158, 241-264

Owen, P.R., & Thomson, W.R. (1963). Heat transfer across rough surfaces. *Journal Fluid Mech.*, 15, 321 – 334

Paulson, C.A. (1970). The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *Journal of applied meteorology*, 9, 857-861

Perego, S. (1999). Metphomod - A numerical mesoscale model for simulation of regional photosmog in complex terrain : model description and application during pollumet 1993 (Switzerland) *Meteorol. Atmos. Phys.*, 70, 43-69

Petropoulos, G., Carlson, T.N., Wooster, M.J., & Islam, S. (2009). A review of Ts/VI remote sensing based methods for the retrieval of land surface energy fluxes and soil surface moisture. *progress in Physical Geography*, 33, 224-250

Pham, D.T. (2001). Stochastic methods for sequential data assimilation in strongly non-linear systems. *Monthly Weather Review*, 129, 1194-1207

Pham, D.T., Verron, J., & Roubaud, M.C. (1998). A singular evolutive extended kalman filter for data assimilation in oceanography. *Journal of Marine Systems*, 16, 323-340

Philp, J.R., & De Vries, D.L. (1957). Moisture movement in porous materials under temperature gradients. *Trans. Am. Geophys. Union*, 38, 222-232

Pielke, R.A., Cotton, W., Walko, R., Tremback, C.J., Nicholls, M.E., Moran, M.D., Wesley, D.A., Lee, T.J., & Copeland, J.H. (1992). A comprehensive meteorological modeling system - RAMS. *Meteorol. Atmos. Phys.*, 49, 69-91

Pipunic, R.C., Walker, J.P., & Western, A. (2008). Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: A comparative synthetic study. *Remote Sensing Environment*, 112, 1295-1305

Prata, A. J., J. Caselles, J. Coll, C. Otle & J. Sobrino (1995). Thermal remote sensing of the land surface from satellite: Current status and future prospects. *Remote Sens. Rev.*, 12, 175–224

Prévoit, L., Baret, F., Chanzy, A., & Olioso, A. (1998). Assimilation of multisensor and multi-temporal remote sensing data to monitor vegetation and soil: The Alpilles-ReSeDA project. In IEEE (Ed.), *IGARSS 98* (pp. E04-03). Seattle, WA, USA: vol. CDRom

Prévo, L., Chauki, H., Troufleau, D., Weiss, M., Baret, F., & Brisson, N. (2003). Assimilating optical and radar data into the STICS crop model for wheat. *Agronomie*, 23, 297-3003

Raupach, M.R. (1991). Vegetation - Atmosphere interaction in homogeneous and heterogeneous terrain. Some implications of mixed-layer dynamics. *Vegetatio*, 91, 105-120

Raupach, M.R. (1992). Drag and drag partition on rough surfaces. *Bound. Layer Meteor.*, 60, 375-395

Raupach, M.R. (1994). Simplified Expressions for Vegetation Roughness Length and Zero-Plane Displacement. *Bound. Layer Meteor.*, 71, 211-216

Raupach, M.R. (1995). Vegetation atmosphere interaction and surface conductance at leaf, canopy and regional scales. *Agricultural and Forest Meteorology*, 73, 151-179

Raupach, M.R., A.S., T., & Edwards, I. (1980). A wind-tunnel study of turbulent-flow close to regularly arrayed rough surfaces. *Bound. Layer Meteor.*, 18, 373-397

Récan, M. (1982). Simulation numérique du comportement thermique et hydrique d'un sol nu. Application à l'étude de l'évaporation par télédétection. Thèse de docteur ingénieur. In (p. 112): Inst. Nat. polytechnique de Toulouse

Reginato, R.J., Jackson, R.D., & Pinter, P.J.J. (1985). Evapotranspiration calculated from remote multispectral and ground station meteorological data. *Remote Sensing Environment*, 18, 75-89

Reichle, R. H., and D. Entekhabi (2001). Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach. *water resources research* , 37, 2353–2364

Reichle, R.H. (2002). Extended versus ensemble Kalman filtering for land data assimilation. *Journal of Hydrometeorology*, 3, 728-740

Riou, C., Itier, B., & Seguin, B. (1988). The influence of surface roughness on the simplified relationship between daily evaporation and surface temperature. *Int J. Remote Sens*, 9, 1529-1533

Rivalland, V., Olioso, A., Claverie, M., Weiss, M., & Baret, F. (2006). Neural Net Techniques Used to Estimate Temporal and High Resolution Canopy Biophysical Variables from 3 Remote Sensing Data Sources. In J.A. Sobrino (Ed.), *Proceedings of the Second Recent Advances in Quantitative Remote Sensing* (pp. 567-572). Valencia: Servicio de Publicaciones: Universitat de Valencia

Robinson, A.R., & Lermusiaux, P.F.J. (2000). Overview of data assimilation, Harvard Reports in Physical/Interdisciplinary (Ocean Sciences). In

[http://people.seas.harvard.edu/~robinson/PAPERS/red\\_report\\_62.html](http://people.seas.harvard.edu/~robinson/PAPERS/red_report_62.html) (Ed.), *The Division of Engineering and Applied Sciences*: Harvard university Cambridge, Massachusetts, USA

Roerink, G.J., Su, Z., & Menenti, M. (2000). S-SEBI: a simple remote sensing algorithm to estimate the surface energy balance. *Physics and Chemistry of the Earth*, 25, 147-157

Ross, J. (1981). The radiation regime and plant architecture in plant stands. In H. Lieth (Ed.), *Taks for Vegetation Sciencies 3*. The Hague, Netherlands: Dr. W. Junk

Rubio, E., Caselles, V., Coll, C., Valour, E., & Sospedra, F. (2003). Thermal-infrared emissivities of natural surfaces: improvements on the experimental set-up and new measurements. *Int J. Remote Sens*, 24, 5279-5390

Samaali, M. (2002). Evaluation d'un modèle de couche limite atmosphérique 3D dans un cas homogène: Application à une parcelle agricole de soja. Thèse de doctorat INRA. In. <http://www.inra.fr/theses/these-integrale/Theses/samaali/html/these.html>: Université de Provence Polytechnique de Marseille

Sanchez, J.M., Kustas, W.P., Caselles, V., & Anderson, M.C. (2008). Modelling surface energy fluxes over maize using a two-source patch model and radiometric soil and canopy temperature observations. *Remote Sensing of Environment*, 112, 1130-1143

Santanello, J.J., & Friedl, M. (2003). Diurnal Covariation in Soil Heat Flux and Net Radiation. *Journal Applied Meteor.*, 42, 851-862

Seguin, B., Baelz, S., Monget, J.M., & Petit, V. (1982). Utilisation de la thermographie IR pour l'estimation de l'évaporation régionale. II: Résultats obtenus à partir de données satellites. *Agronomie*, 2, 113-118

Seguin, B., & Itier, B. (1983). Using midday surface temperature to estimate daily evaporation from satellite thermal IR data. *International Journal of Remote Sensing*, 4, 371-383

Sellers, P.J., Randall, D.R., Collatz, G.J., Berry, J.A., Field, C.B., Dazlich, D.A., Zhang, C., Collelo, G.D., & Bounoua, L. (1996). A revised land-surface parameterization (SiB2) for atmospheric GCMs. Part 1: Model formulation. *Journal of Climate*, 9, 676-705

Seuffert, G., Wilker, H., Viterbo, P., Mahfouf, J.-F., Drusch, M., & Calvet, J.-C. (2003). Soil moisture analysis combining screen-level parameters and microwave brightness temperature. *Geophysical Research Letters*, 30, 1498

Sinclair, T.R., & Seligman, N.G. (1996). Crop modeling: From infancy to maturity. *Agronomy Journal*, 88, 698-704

Sobrino, J.A., Gomez, M., Jimenez-Munoz, C., & Oliso, A. (2007). Application of a simple algorithm to estimate daily evapotranspiration from NOAA-AVHRR images for the Iberian Peninsula. *Remote Sensing Environment*, 110, 139-148

Sobrino, J.A., Gomez, M., Jimenez-Munoz, J.C., Oliso, A., & Chehbouni, G. (2005). A simple algorithm to estimate evapotranspiration from DAIS data: Application to the DAISEX campaigns. *Journal of Hydrology*, 315, 117-125

Sobrino, J.A., Raissouni, N., Oliso, A., Hasager, C.B., Belaid, M.A., Rahman, S.A., & Chehbouni, A. (2001). WATERMED – Water use efficiency in natural vegetation and agricultural areas by remote sensing in the mediterranean basin. In, *IEEE International Symposium on Geoscience and Remote Sensing (IGARSS)* (pp. 3158-3160). Sydney, Australia

Soer, G. (1980). Estimation of regional evapotranspiration and soil moisture conditions using remotely sensed crop surface temperatures. *Remote Sensing Environment*, 9, 27-45

Soudani, K., François, C., Le Maire, G., Le Dantec, V., & Dufrêne, E. (2006). Comparative analysis of IKONOS, SPOT, and ETM+ data for leaf area index estimation in temperate coniferous and deciduous forest stands. *Remote Sensing Environment*, 102, 161-175

Steinmetz, S., Lagouarde, J.P., Delécolle, R., Guérif, M., & Seguin, B. (1989). Evaporation and water stress using thermal infrared measurements. A general review and a case study on winter durum wheat in southern France. In, *Symposium on Physiology Breeding of Winter Cereals for Stressed Mediterranean Environments*. Montpellier, France: ICARDA – INRA

Stull, R. (1998). An introduction to Boundary Layer Meteorology (p. 1.250): Kluwer Academic Publishers, Netherlands

Su, H., McCabe, M., Wood, E.F., Su, Z., & Prueger, J.H. (2005). Modeling evapotranspiration during SMACEX: Comparing two approaches for local- and regional-scale prediction. *Journal of Hydrometeorology*, 6, 910-922

Su, Z. (2002). The surface energy balance system (SEBS) for estimation of turbulent heat fluxes. *Hydrology and Earth System Sciences*, 6, 85-99

Su, Z. (2005). Hydrological applications of remote sensing. Surface fluxes and other derived variables surface energy balance. In M.C. Anderson (Ed.), *Encyclopedia of hydrological sciences*. Hoboken, NJ, USA: John Wiley and Sons

Su, Z., Li, X., Zhou, Y., Wan, L., Wen, J., & Sintonen, K. (2003). Estimating areal evaporation from remote sensing. In, *Proc. IEEE Int.* (pp. 1166-1168)

Su, Z., Schmugge, T., Kustas, W.P., & Massman, W.J. (2001). An evaluation of two models for estimation of the roughness height for heat transfer between the land surface and the atmosphere. *Journal Applied Meteor.*, 40, 1933–1951

Su, Z., Timmermans, W., Gieske, A.S.M., Jia, L., Elbers, J.A., Oliso, A., Timmermans, J., Van der Velde, R., Jin, X., Van der Kwast, H., Nerry, F., Sabol, D., Sobrino, J.A., Moreno, J.F., & Bianchi, R. (2008). Quantification of land-atmosphere exchanges of water, energy and carbon dioxide in space and time over the heterogeneous Barrax site. *International Journal of Remote Sensing*, 29, 5215-5235

Sun, N. (1994). Inverse problems in groundwater modeling. *kluwer Academic Publishers*, 39-91

Taconet, O., Bernard, R., & Vidal-Madjar, D. (1986). Evapotranspiration over an agricultural region using a surface flux/temperature model based on NOAA-AVHRR data. *J. Clim. Appl. Meteor.*, 25, 284-307

Taconet, O., Oliso, A., BenMehrez, M., & Brisson, N. (1995). Seasonal estimation of evaporation and stomatal conductance over a soybean field using surface IR temperatures. *Agricultural and Forest Meteorology*, 73, 321-337

Talagrand O., Courtier P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Q. J. R. Meteorol. Soc.*, vol. 113, 1311-1328

Talagrand, O. (1997). Assimilation of observations, an introduction. *J. Meteor. Soc.*, 75, 191-209

Tanner, B.C.W., & Pelton, W.L. (1960). Potential evaporation estimates by the approximate energy balance method of Penman. *Journal of Geophysical Research*, 65, 3391-3413

Tarantola, A. (1987). Inverse problem theory, methods for data fitting and model parameter estimation. In Elsevier (Ed.). Amsterdam

Tarantola, A. (2005). Inverse Problem Theory and Methods for Model Parameter Estimation. In SIAM (Ed.) (p. 358)

Tasumi, M., Allen, R.G., & Trezza, R. (2008). At surface reflectance and albedo from satellite for operational calculation of land surface energy balance. *Journal of hydrologic engineering*, 13, 51-63

Teixeira, A.H.D.C., Bastiaanssen, W.G.M., Ahmadd, M.D., & Bos, M.G. (2009a). Reviewing SEBAL input parameters for assessing evapotranspiration and water productivity for the Low-Middle São Francisco River basin, Brazil Part A: Application to the regional scale. *Agricultural and Forest Meteorology*, 149, 462-476

Teixeira, A.H.D.C., Bastiaanssen, W.G.M., Ahmadd, M.D., & Bos, M.G. (2009b). Reviewing SEBAL input parameters for assessing evapotranspiration and water productivity

for the Low-Middle São Francisco river basin, Brazil Part B: Application to the regional scale. *Agricultural and Forest Meteorology*, 149, 477-490

Tennekes, H., & Driedonks, A.G.M. (1981). Basic entrainment equations for the atmospheric boundary layer. *Bound. Layer Meteor.*, 20, 515-531

Thépaut, J.N., & Courtier, P. (1991). Four dimensional variational data assimilation using the adjoint of a multilevel primitive equation model. *Quart. J. Roy. Meteor. Soc.*, 117, 1225-1254

Thoreson B., Clark B., Soppe R., Keller A., Bastiaanssen W.G.M. & Eckhardt, J. (2009). Comparison of evapotranspiration estimates from remote sensing (SEBAL), water balance and crop coefficient approach, Proceedings ASCE World Environmental and Water Resources Congress 2009: Great Rivers

Tikhonov, A.N. (1963). Solution of incorrectly formulated problems and the regularization method. *Sov Math Doklady*, 4, 1035-1038

Timmermans, W., Bertoldi, G., Albertson, J.D., Olioso, A., Su, Z., & Gieske, A.S.M. (2008). Accounting for Atmospheric Boundary Layer variability on flux estimation from RS observations. *International Journal of Remote Sensing Environment*, 29, 5275-5290

Timmermans, W., Kustas, W.P., Anderson, M.C., & French, A.N. (2007). An intercomparison of the Surface Energy Balance Algorithm for Land (SEBAL) and the Two-Source Energy Balance (TSEB) modeling schemes. *Remote Sensing Environment*, 108, 369-384

Timmermans, W.J., Olioso, A., & Albertson, J.D. (2009). Surface heterogeneity influence on land-atmosphere energy exchanges. In, *The International Conference on Land Surface Radiation and Energy Budgets. Mars 18-20*. Beijing

Troufleau, D., Lhomme, J., Monteny, B., & Vidal, A. (1997). Sensible heat flux and radiometric surface temperature over sparse Sahelian vegetation. I. An Experimental analysis of the kB parameter. *Journal of Hydrology*, 188-189, 815-838

Tuzet, A., Perrier, A., & Leuning, R. (2003). A coupled model of stomatal conductance, photosynthesis and transpiration. *Plant, Cell and Environment*, 26, 1097-1116

Van de Wiel, B.J.H., Moene, A., Ronda, R.J., De Bruin, H.A.R., & Holtslag, A.A. (2002a). Intermittent Turbulence and Oscillations in the Stable Boundary Layer Over Land Part II: A System Dynamics Approach. *J. Atmos. Sci.*, 59, 2567-2581

Van de Wiel, B.J.H., Ronda, R.J., Moene, A., De Bruin, H.A.R., & Holtslag, A.A. (2002b). Intermittent turbulence and oscillations in the stable boundary layer over land. Part I: A bulk model Source. *J. Atmos. Sci.*, 59, 942-958

Van den Hurk, B.J.J.M. (2001). Energy balance based surface flux estimation from satellite data, and its application for surface moisture assimilation. *Meteorol. Atmos. Phys.*, 76, 43-52

Van Wijk, W.R., & De Vries, D.A. (1963). Periodic temperature variations in a homogeneous soil. In W.R. Van Wijk (Ed.), *Physics of Plant Environment* (pp. 103-143). Amsterdam, Holland

Verhoef, A., De Bruin, H.A.R., & Van den Hurk, B.J.J.M. (1997). Some practical notes on the parameter for sparse vegetation *Journal Applied Meteor.*, 36, 560-572

Verhoef, W., & Bach, H. (2003). Remote sensing data assimilation using coupled radiative transfer models. *Physics and Chemistry of the Earth*, 28, 3-13

Vidal, A. (1990). Estimation de l'évapotranspiration par télédétection, application au contrôle de l'irrigation. In, *Etudes du CEMAGREF, Série Hydraulique agricole n°8* (pp. 1-180). Montpellier, France: CEMAGREF-DICOVA

Vidal, A., Kerr, Y., Lagouarde, J.P., & Seguin, B. (1987). Remote-sensing and water-balance-combined use of an agrometeorological model and of NOAA-AVHRR satellite thermal IR data. *Agricultural and Forest Meteorology*, 39, 155-175

Vidal, A., & Perrier, A. (1989). Technical note: Analysis of a simplified relation for estimating daily evapotranspiration from satellite thermal IR data. *Int J. Remote Sens*, 10, 1327-1337

Walker, J.P., & Houser, P.R. (2001). A methodology for initializing soil moisture in a global climate model: Assimilation of near surface soil moisture observations. *Journal of Geophysical Research*, 106, 761-774

Wallach, D. (2006). Evaluating crop models. In D. Wallach, D. Makowski & J.W. Jones (Eds.), *Working with Dynamic Crop Models: Evaluation, Analysis, Parameterization, and Application* (pp. 11-55)

Wassenaar, T., Olioso, A., Hasager, C.B., Jacob, F., & Chehbouni, A. (2002). Estimation of evapotranspiration on heterogeneous pixels. In J.A. Sobrino (Ed.), *Proceedings of the First International Symposium on Recent Advances in Quantitative Remote Sensing, 16-20 September 2002* (pp. 458-465). Valencia, Spain: Publications de la Universitat de València, España

Webb, E.K. (1970). Profile relationships: The log-linear range, and extension to strong stability. *Quart. J. Roy. Meteor. Soc.*, 96, 67-90

Weiss, M., & Baret, F. (1999). Evaluation of Canopy Biophysical Variable Retrieval Performances from the Accumulation of Large Swath Satellite Data. *Remote Sensing Environment*, 70, 293-306

Weiss, M., Baret, F., Garrigues, S., & Lacaze, R. (2007). LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: validation and comparison with MODIS collection 4 products. *Remote Sensing Environment*, 110, 317-331

Weiss, M., Baret, F., Leroy, M., Hauteceur, O., Bacour, C., Prevot, L., & Bruguier, N. (2002). Validation of neural net techniques to estimate canopy biophysical variables from remote sensing data. *Agronomie*, 22, 547-553

Weiss, M., Troufleau, D., Baret, F., Chauki, H., Prévot, L., Olioso, A., Bruguier, N., & Brisson, N. (2001). Coupling canopy functioning and radiative transfer models for remote sensing data assimilation. *Agricultural and Forest Meteorology*, 108, 113-128

Wiegand, C.L., & Bartholic, J.F. (1970). Remote sensing in evapotranspiration research on the Great Plains. In, *Proceedings of Evapotranspiration in the Great Plains Seminar, Research Committee, Great Plains Agricultural Publication N°50*

Wigneron, J.P., Olioso, A., Calvet, J.C., & Bertuzzi, P. (1999). Estimating root-zone soil moisture from surface soil moisture data and soil-vegetation-atmosphere transfer modeling. *water resources research*, 35, 3735-3745

Witono, H., & Bruckler, L. (1989). Use of remotely sensed soil moisture content as boundary conditions in soil atmosphere water transport modelling. 1. Field validation of a water flow model. *water resources research*, 25, 2423-2435

Wood, E.F., Su, H., McCabe, M., & Su, Z. (2003). Estimating Evaporation from Satellite Remote Sensing. In, *IGARSS* (pp. 21-25)

Xu, Q. (1996). Generalized adjoint for physical processes with parametrized discontinuities. part I: Basic issues and heuristic examples. *J. Atmos. Sci.*, 53, 1123-1142

Zhou Y., McLaughlin D. & Entekhabi D (2006). Assessing the Performance of the Ensemble Kalman Filter for Land Surface Data Assimilation. *Monthly Weather Review*, 134, 2128-2142

Zhao, W.G., Olioso, A., Lagouarde, J.P., Kerr, Y., Mc Aneney, J., Bonnefond, J.M., & Marloie, O. (2002). Estimation of aerodynamic parameters under non-neutral stability conditions from Alpillés measurement data. *Agronomie*, 22, 619-625

## ***ANNEXES***

## **ANNEXE I: Article Courault et al. 2008**

Courault, D., Bsaibes, A., Kpemlie, E., Hadria, R., Hagolle, O., Marloie, O., Hanocq, J.F., Oliosio, A., Bertrand, N., & Desfonds, V. (2008). **Assessing the Potentialities of FORMOSAT-2 Data for Water and Crop Monitoring at Small Regional Scale in South-Eastern France.** *Sensors*, 8, 3460-3481

**Abstract:** Water monitoring at the scale of a small agricultural region is a key point to insure a good crop development particularly in South-Eastern France, where extreme climatic conditions result in long dry periods in spring and summer with very sparse precipitation events, corresponding to a crucial period of crop development. Remote sensing with the increasing imagery resolution is a useful tool to provide information on plant water status over various temporal and spatial scales. The current study focussed on assessing the potentialities of FORMOSAT-2 data, characterized by high spatial (8m pixel) and temporal resolutions (1-3 day/time revisit), to improve crop modeling and spatial estimation of the main land properties. Thirty cloud free images were acquired from March to October 2006 over a small region called Crau-Camargue in SE France, while numerous ground measurements were performed simultaneously over various crop types. We have compared two models simulating energy transfers between soil, vegetation and atmosphere: SEBAL and PBLs. Maps of evapotranspiration were analyzed according to the agricultural practices at field scale. These practices were well identified from FORMOSAT-2 images, which provided accurate input surface parameters to the SVAT models.

**Keywords:** FORMOSAT-2, crop monitoring, evapotranspiration, LAI, albedo, surface temperature.

# Poster iLEAPS: 17 – 21 Novembre 2008, Hyères, France:

Kpemlie, E., Olioso, A., Courault, D., Bsaibes, A. (2008): **Assimilation of remote sensing data into a coupled land surface – boundary layer model for mapping evapotranspiration and microclimate.** <http://www.ileaps.org/index.php>

## ASSIMILATION OF REMOTE SENSING DATA INTO A COUPLED LAND SURFACE - BOUNDARY LAYER MODEL FOR MAPPING EVAPOTRANSPIRATION AND MICROCLIMATE

E. Kpemlie, A. Olioso, D. Courault, S. Buis, A. Bsaibes  
 INRA, UMR EMMAH (1114), Domaine St Paul, Site  
 Agroparc, 84914 AVIGNON, Cedex 9 France  
 Emmanuel.Kpemlie@avignon.inra.fr



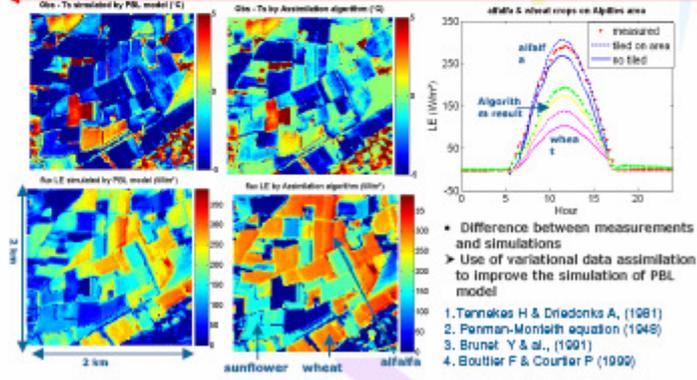
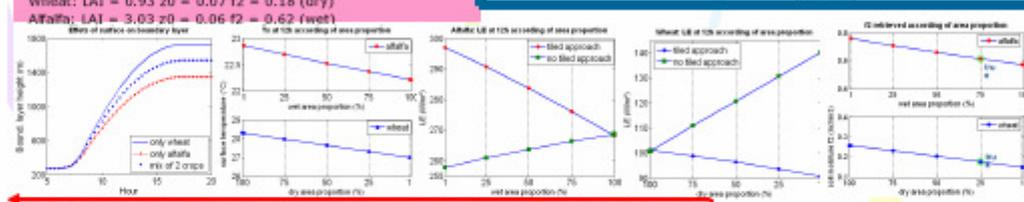
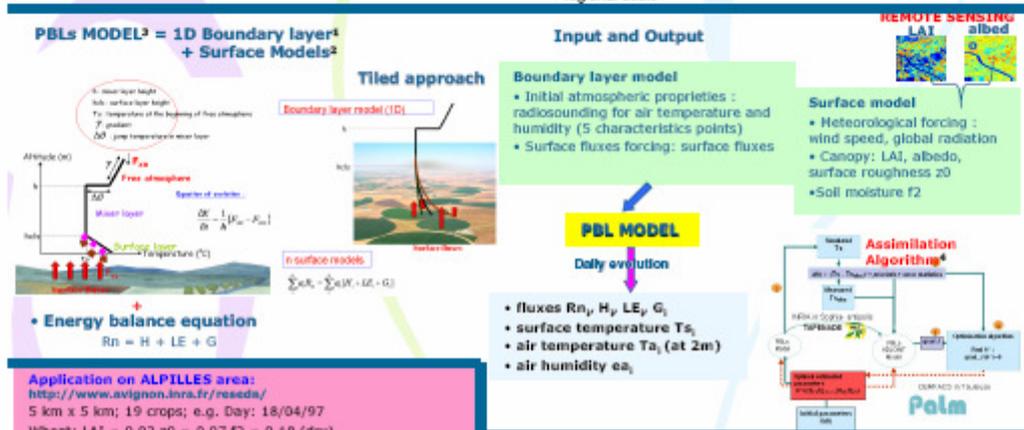
**Abstract:** Evapotranspiration and microclimate depend on :  
 - feedback effects between the land surface and the atmosphere  
 - the spatial and the temporal variability of surface characteristics.

We have firstly developed a simplified coupled atmospheric boundary layer – land surface model which accounts for the landscape heterogeneity according to a tiled approach. This requires the introduction into the model of the various proportions of the main vegetation types encountered in the region, (each surface being characterised by several parameters such LAI, albedo, roughness (z0), normalized soil moisture (f2)).

Secondly, the knowledge of the spatial variability of some surface parameters requires the use of remote sensing data and assimilation procedures.

On one hand, LAI and albedo can be easily estimated from remote sensing images in the solar domain. On the other hand, it's still difficult to estimate spatial variations of soil moisture (f2) and aerodynamic roughness (z0) from remote data.

So, we have implemented appropriate procedures for assimilating thermal data into the model in order to retrieve these 2 parameters to improve the evapotranspiration at regional scale



**ANNEXE II : Résultats de l’algorithme A : (Complément chapitre V section 1)**

✚ **Résultats de l’algorithme d’assimilation avec la **version 1 de PBLs** : sur les données d’expérience jumelle (625 pixels)**

**Tab. I.1a. RMSE des résultats de l’algorithme A (estimer que  $z_0$ ) avec erreur d’observation de 1.5 K**

<b>A</b>	no error		$\alpha$		LAI		$u_a = 0.8$		$\alpha, LAI \& u_a$	
	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo
<i>Rn</i>	4.2	3.6	11.1	11	12	11	6	5	14	13
<i>H</i>	8	7	11	10	43	44	12	9	46	44
<i>LE</i>	5	4	5	4	46	45	7	5	47	46
<i>G</i>	0.6	0.5	1.7	1.6	1.8	1.7	1.0	0.7	2	1.9
<i>Ts</i>	0.7	0.6	0.7	0.5	1.1	0.8	1.0	0.7	1.2	0.9
<i>Ta</i>	0.4	0.36	0.4	0.37	0.5	0.45	0.4	0.4	0.6	0.5
<i>z0</i>	0.100	0.090	0.100	0.093	0.100	0.094	0.100	0.093	0.100	0.097

**Tab. I.1b. RMSE des résultats de l’algorithme A (estimer que  $z_0$ ) avec erreur d’observation de 0.5 K**

<b>A</b>	no error		$\alpha$		LAI		$u_a = 0.8$		$\alpha, LAI \& u_a$	
	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo
<i>Rn</i>	4.2	2.3	11.1	10	12	11	6	2	14	11
<i>H</i>	8	4	11	9	43	43	12	5	46	46
<i>LE</i>	5	3	5	3	46	46	7	3	47	47
<i>G</i>	0.6	0.3	1.7	1.6	1.8	1.6	1.0	0.4	2	1.7
<i>Ts</i>	0.7	0.4	0.7	0.3	1.1	0.5	1.0	0.4	1.2	0.5
<i>Ta</i>	0.4	0.23	0.4	0.25	0.5	0.3	0.4	0.3	0.6	0.4
<i>z0</i>	0.100	0.058	0.100	0.063	0.100	0.085	0.100	0.066	0.100	0.088

✚ Résultats de l'algorithme d'assimilation avec la **version 2 de PBLs** : sur les données d'expérience jumelle (625 pixels)

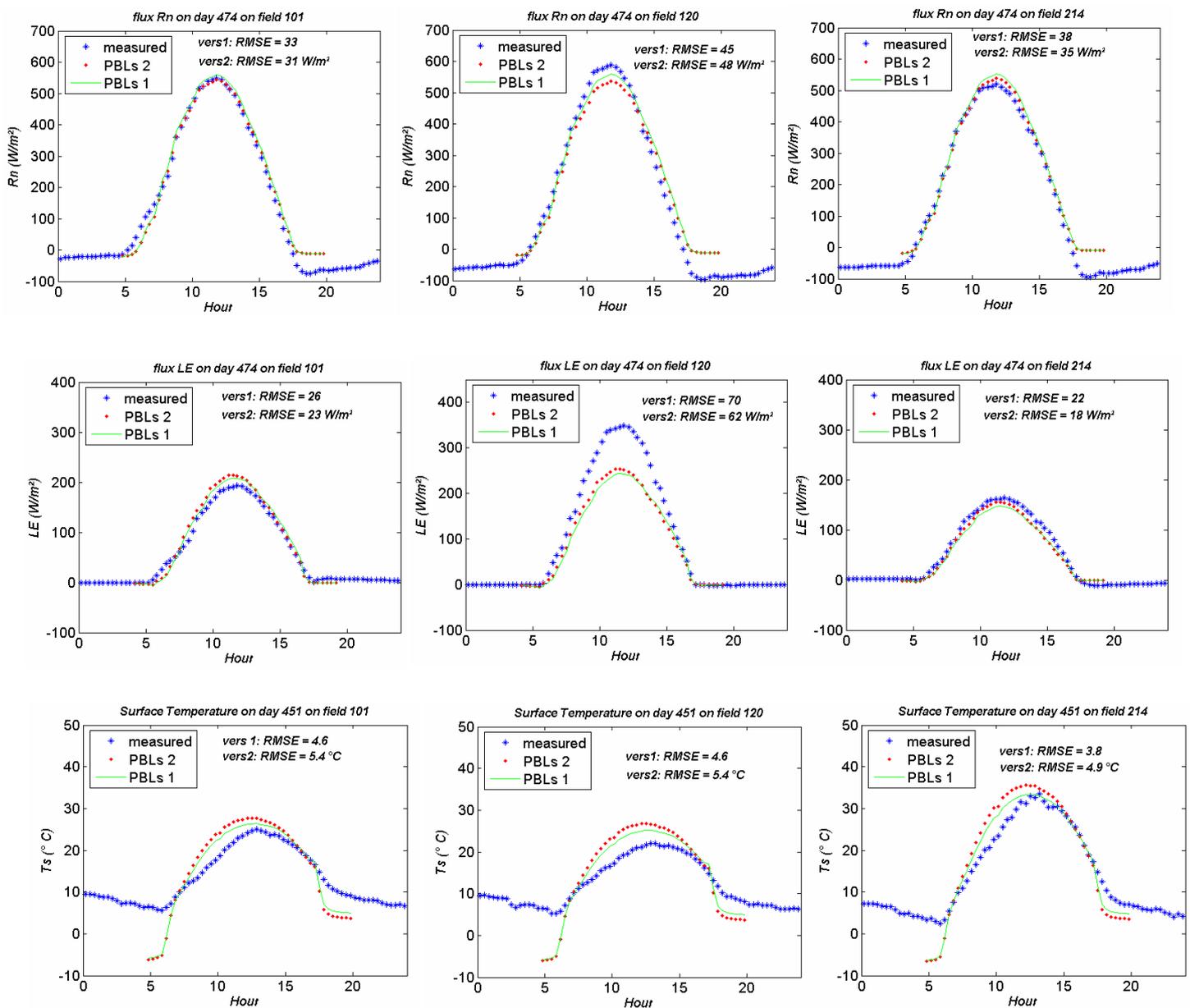
**Tab. II.2a.** *RMSE* des résultats de l'algorithme A (estimer que  $z_0$ ) avec erreur d'observation de 1.5 K

A	no error		$\alpha$		LAI		$u_a = 0.8$		$\alpha, LAI \& u_a$	
	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo
<i>Rn</i>	4	3	x	x	x	x	9	8	16	15
<i>H</i>	6	5	x	x	x	x	13	10	57	56
<i>LE</i>	3.5	3.1	x	x	x	x	6	5	55	54
<i>G</i>	0.8	0.7	x	x	x	x	1.9	1.5	18	17
<i>Ts</i>	0.7	0.5	x	x	x	x	1.5	1.2	1.7	1.3
<i>Ta</i>	0.4	0.4	x	x	x	x	0.6	0.5	0.6	0.5
<i>z0</i>	0.100	0.097	x	x	x	x	0.100	0.100	0.100	0.100

**Tab. II.2b.** *RMSE* des résultats de l'algorithme A (estimer que  $z_0$ ) avec erreur d'observation de 0.5 K

A	no error		$\alpha$		LAI		$u_a = 0.8$		$\alpha, LAI \& u_a$	
	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo	PBLs	Algo
<i>Rn</i>	4	2.3	x	x	x	x	9	4	16	13
<i>H</i>	6	4	x	x	x	x	13	6	57	56
<i>LE</i>	3.5	2	x	x	x	x	6	3	55	54
<i>G</i>	0.8	0.5	x	x	x	x	1.9	0.9	18	17
<i>Ts</i>	0.7	0.3	x	x	x	x	1.5	0.7	1.7	0.8
<i>Ta</i>	0.4	0.3	x	x	x	x	0.6	0.3	0.6	0.4
<i>z0</i>	0.100	0.072	x	x	x	x	0.100	0.100	0.100	0.100

### ANNEXE III : Comparaison PBLs1 et PBLs2 sur les parcelles Alpilles : (Complément chapitre III section 3)

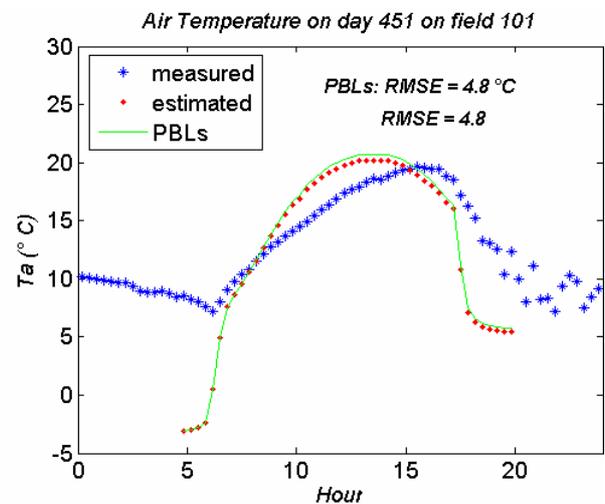
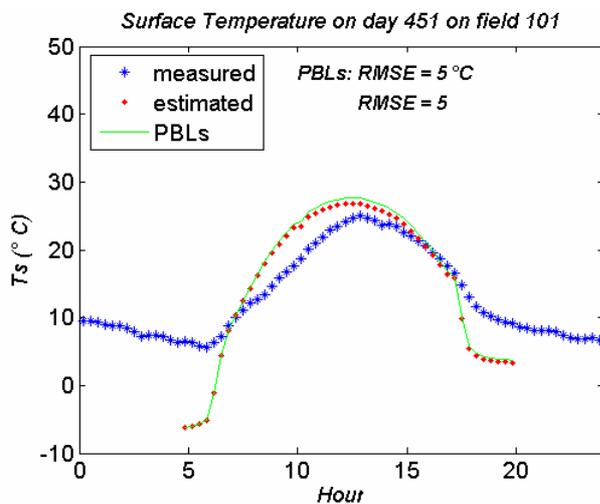
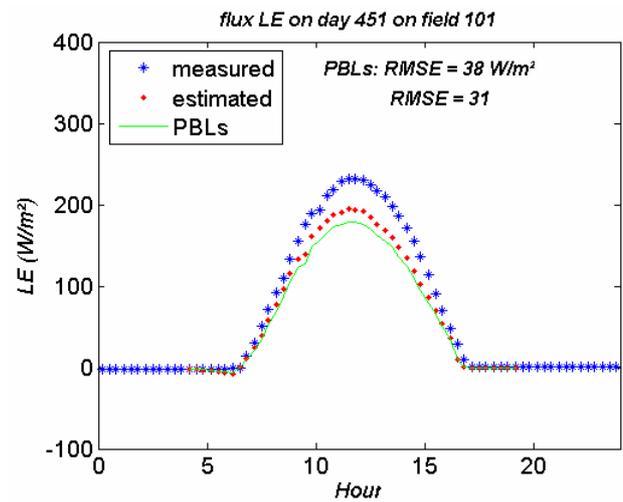
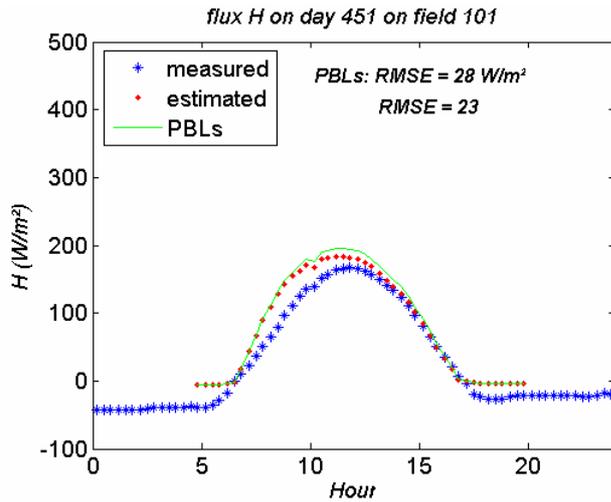
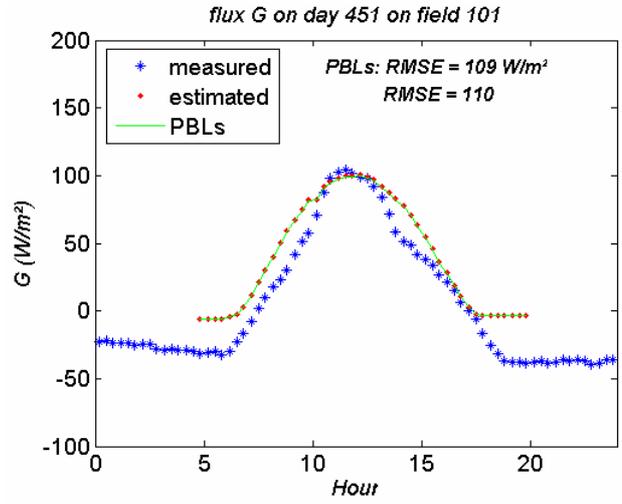
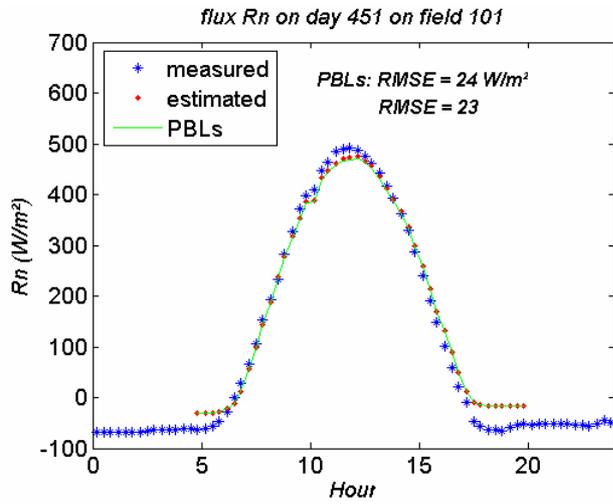


**Fig. III.26. Comparaison des simulations des flux  $R_n$ ,  $LE$  et  $T_s$  par les deux versions de PBLs (vers 1 et vers 2) aux flux mesurés sur les parcelles 101, 120, et 214 les dates du 18/04/97 (DOE 474) et du 26/03/97 (DOE 451).**

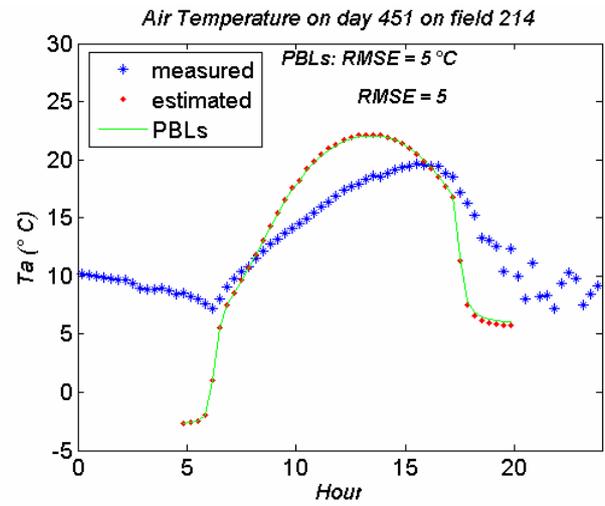
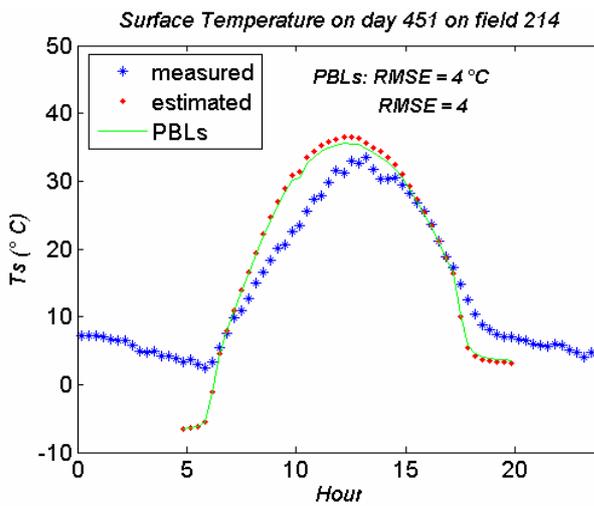
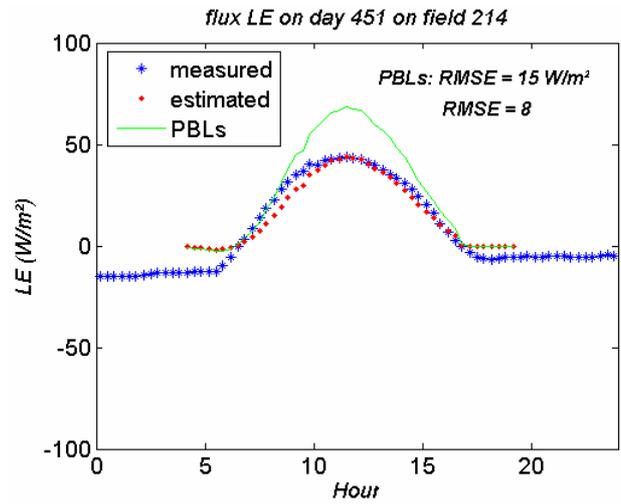
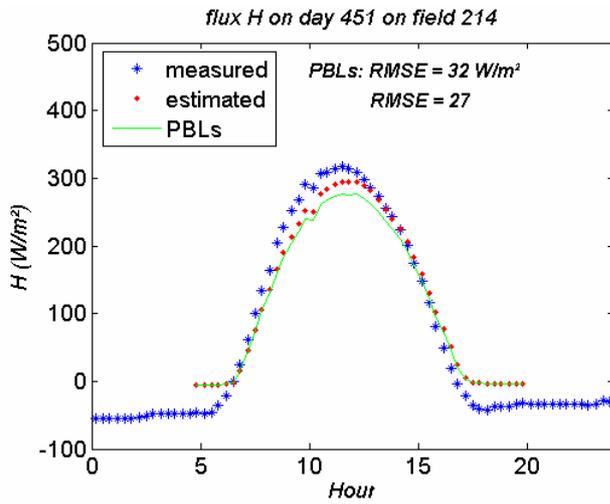
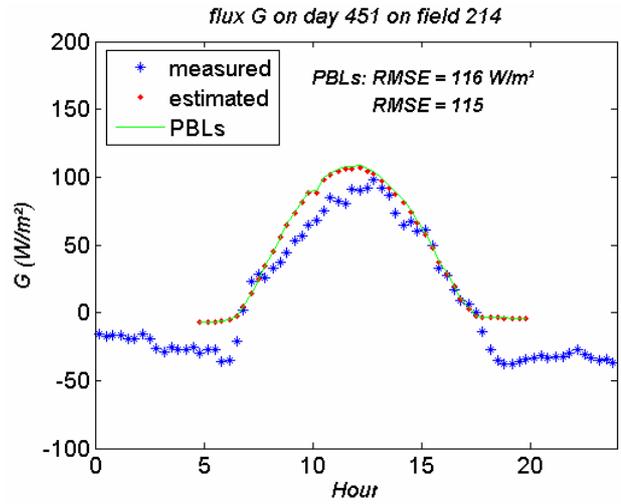
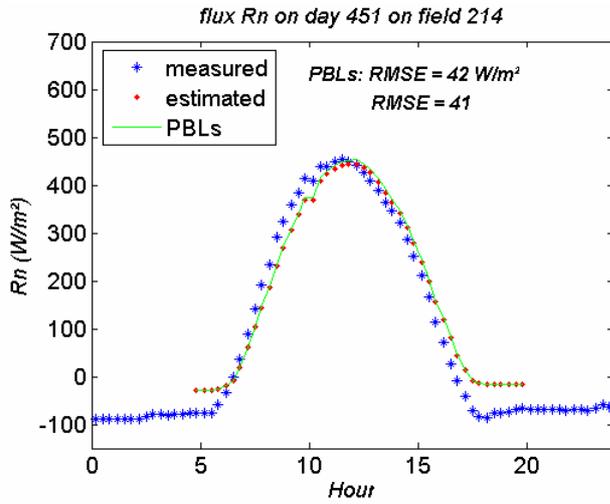
***ANNEXE IV*** : Performances du système d'assimilation sur les données  
**Alpilles** : (Complément chapitre V section 2)

# ANNEXE IV.1 : Validation de l'algorithme d'assimilation sur les parcelles d'Alpilles

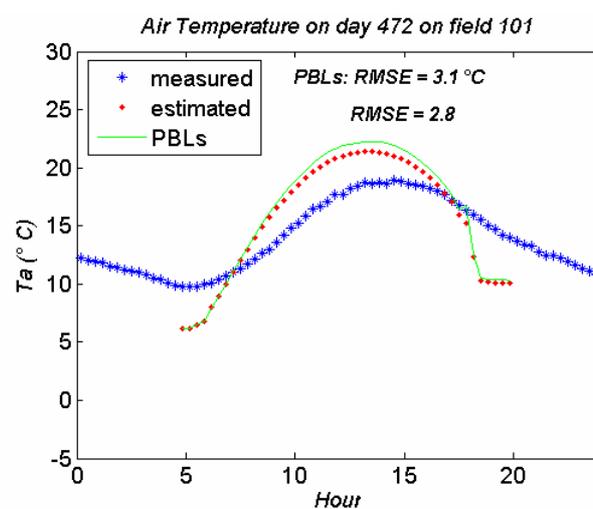
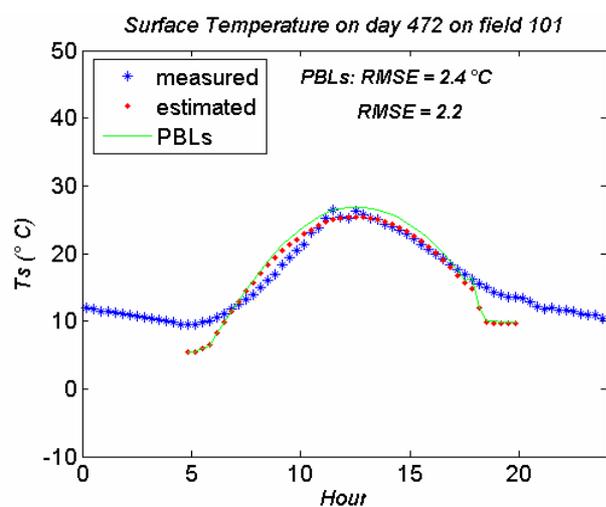
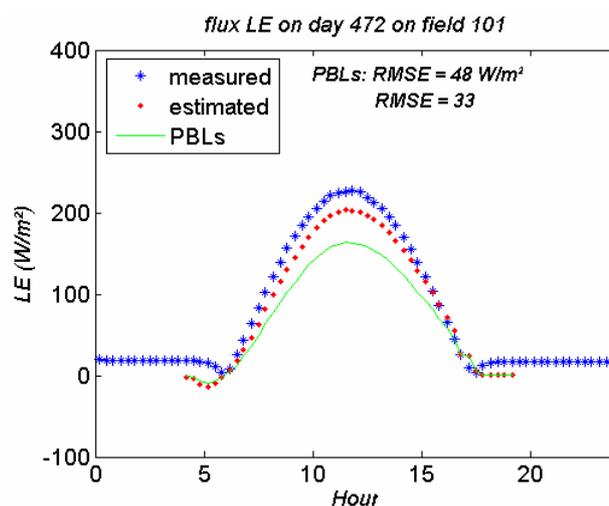
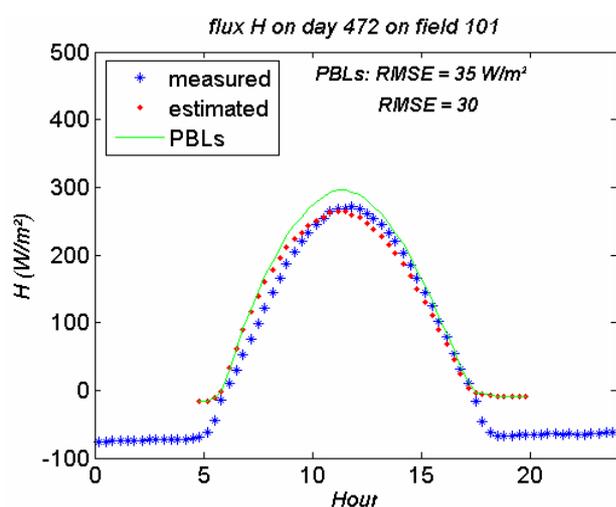
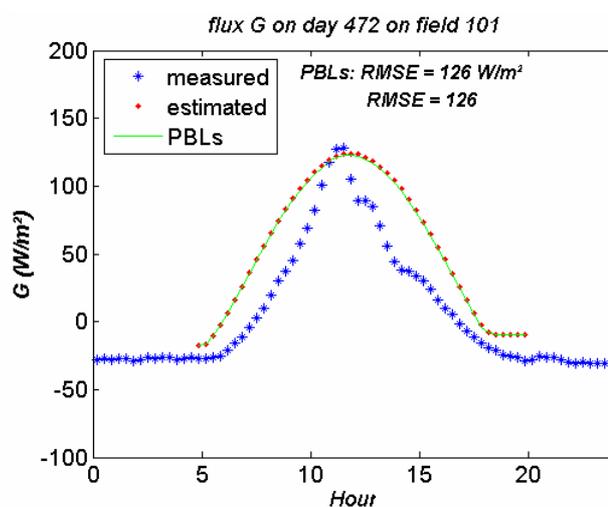
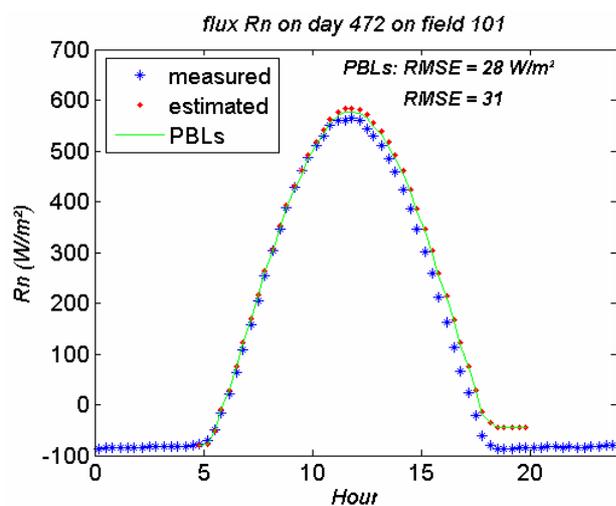
## 1. Date du 26 Mars 1997 (DOE 451) : parcelles 101 (blé)



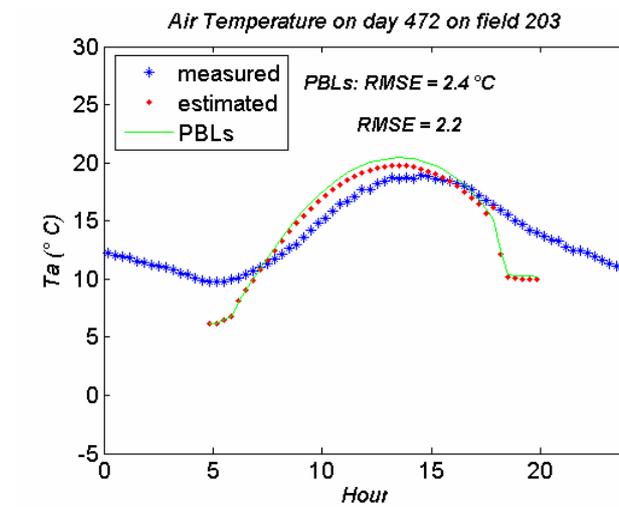
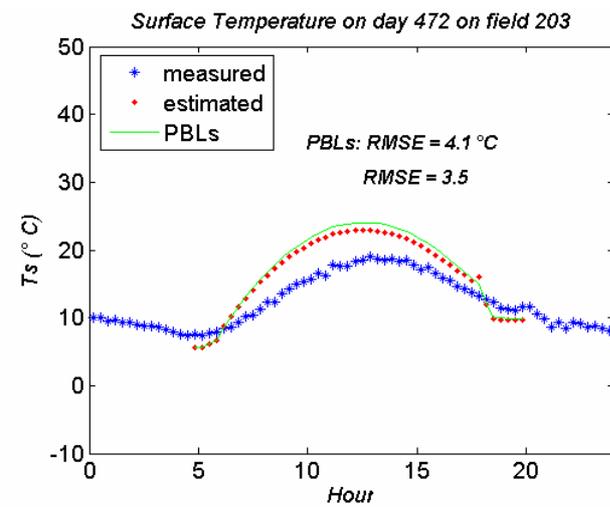
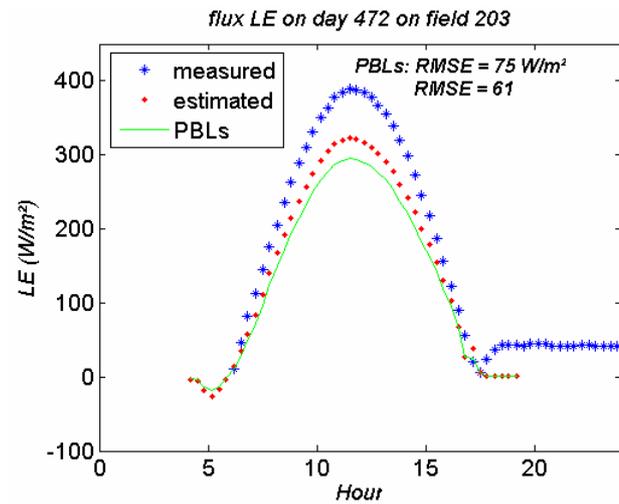
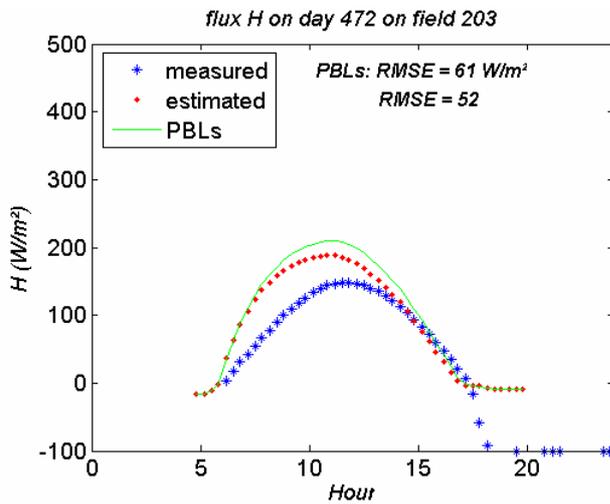
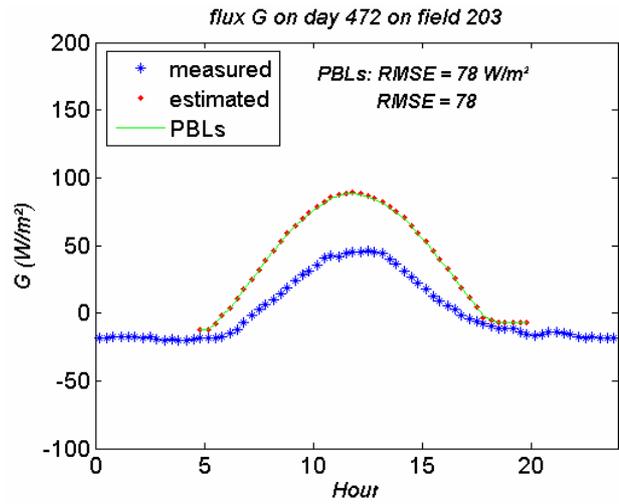
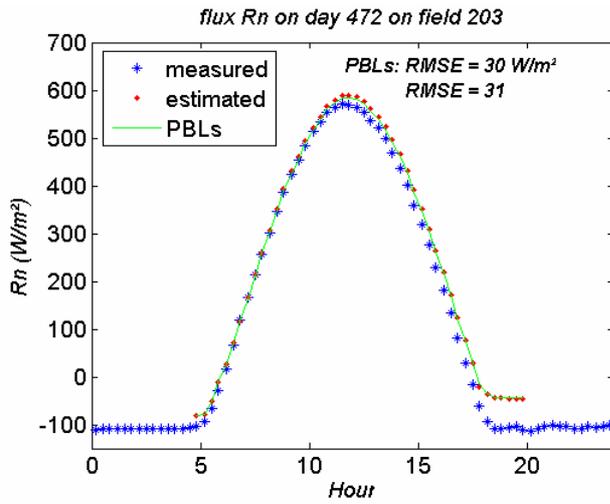
- Date du 26 Mars 1997 (DOE 451) : parcelle 214 (blé)



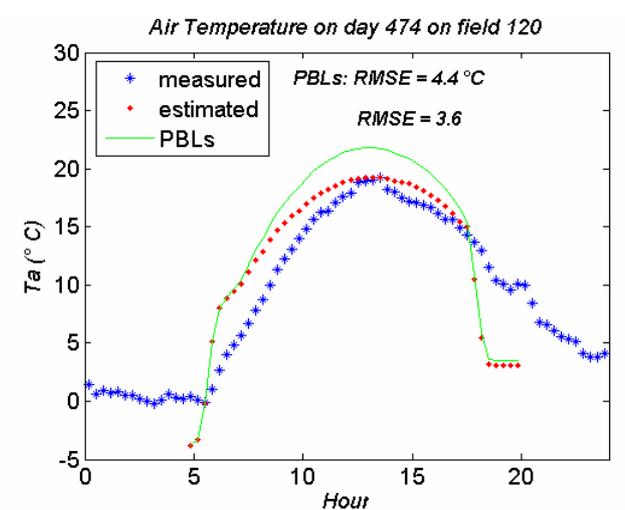
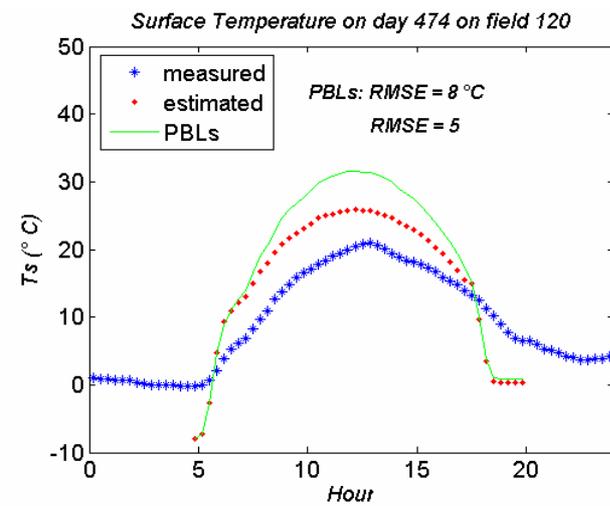
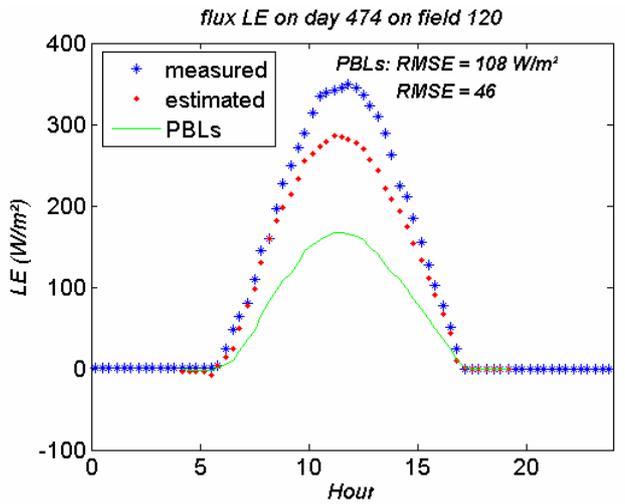
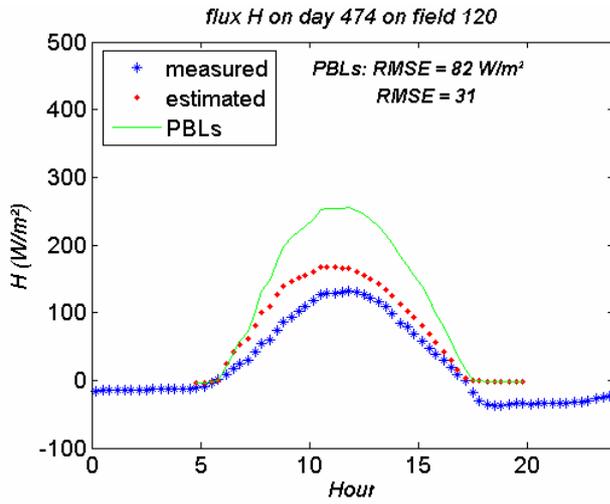
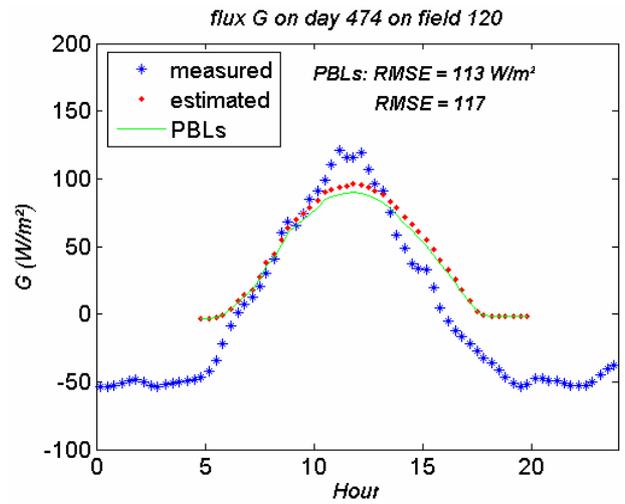
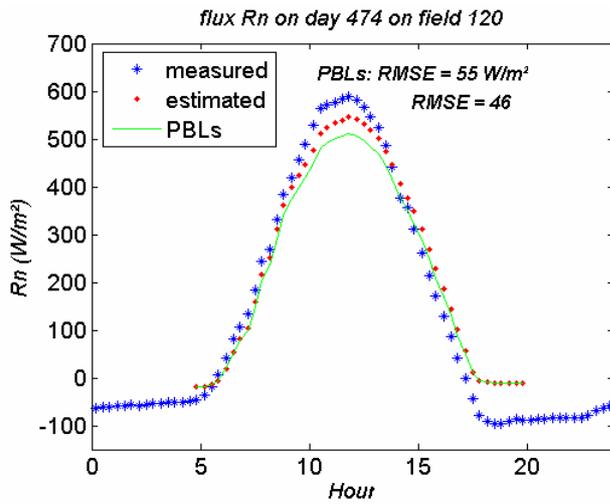
2. Date du 16 Avril 1997 (DOE 472) : parcelles 101 (blé)



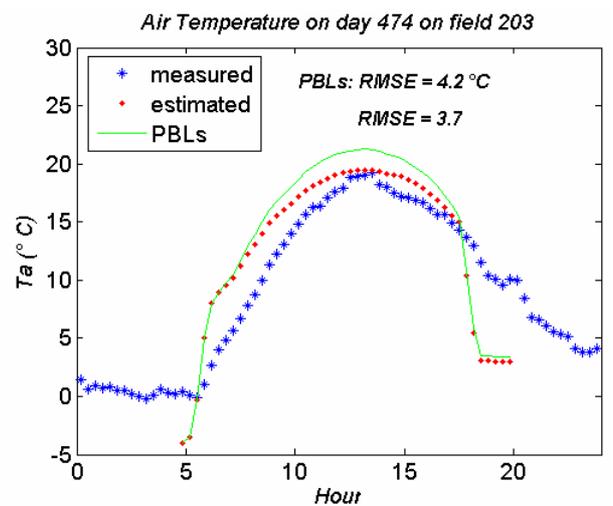
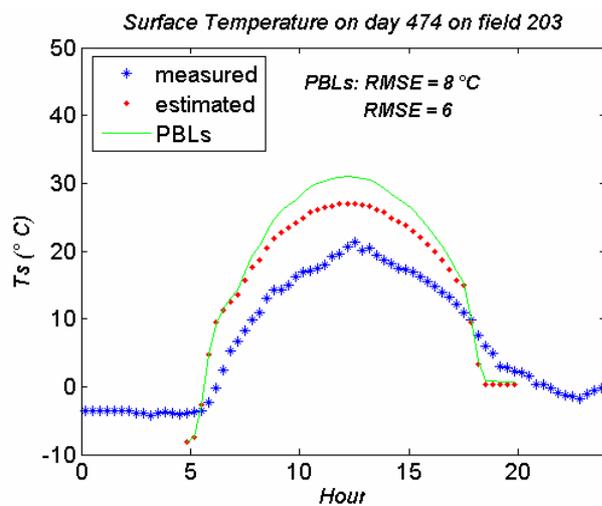
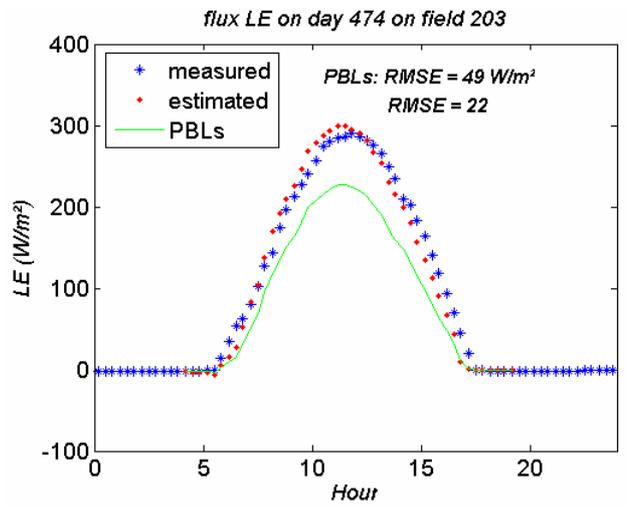
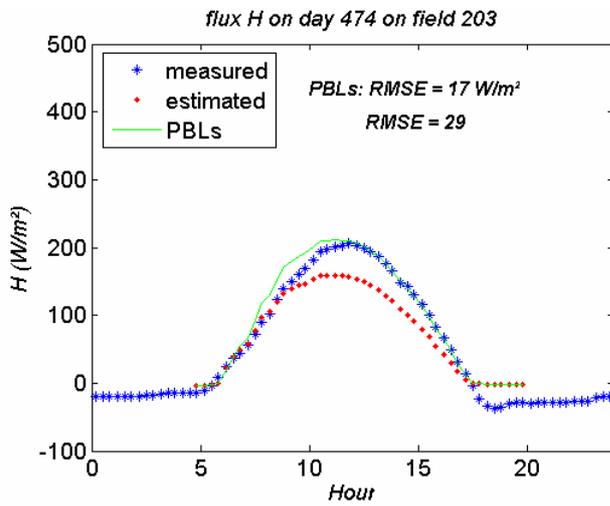
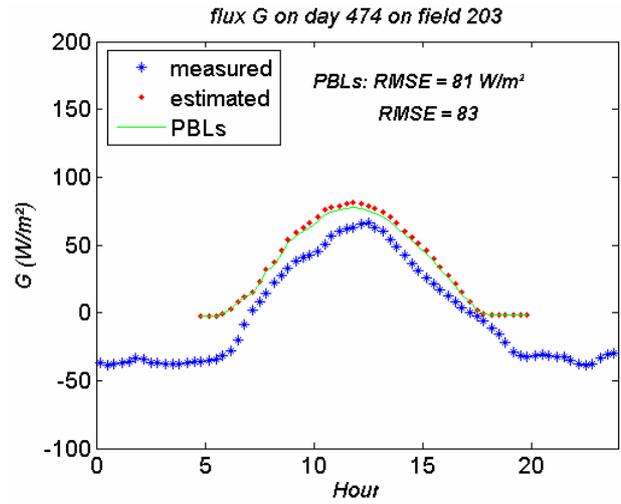
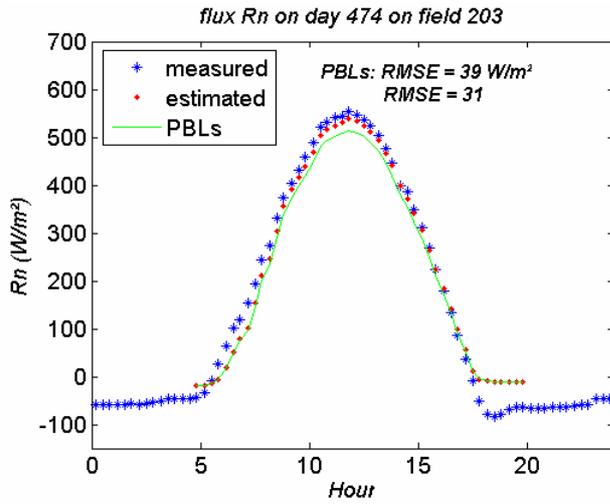
- Date du 16 Avril 1997 (DOE 472) : parcelle 203 (luzerne)



3. **Date du 18 Avril 1997 (DOE 474)** avec  $f_2$  de l'eq. (V.2) : parcelle 120 (blé)



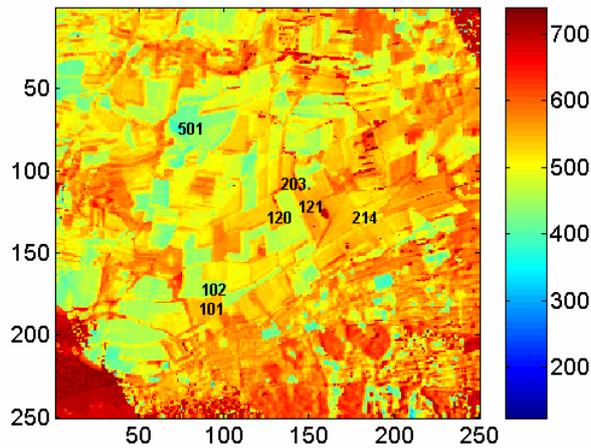
- Date du 18 Avril 1997 (DOE 474) : parcelle 203 (luzerne)



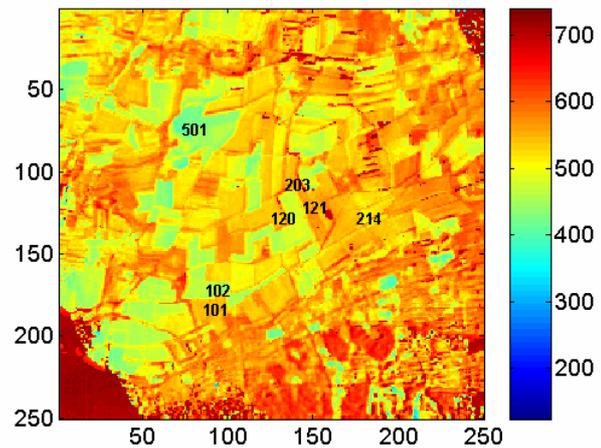
## ANNEXE IV.2 : Cartographie des flux de surface et du microclimat sur la zone Alpilles

### 1. Estimation pour la date du 18 avril 1997 (DOE 474) avec $f_2$ issu de ISBA : Algorithme B

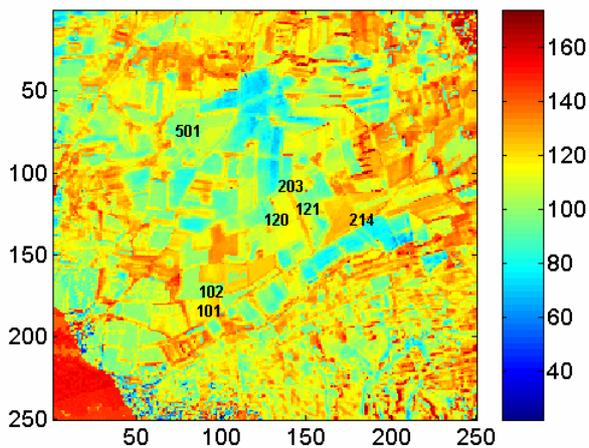
*flux Rn estimated by PBLs on day 474*



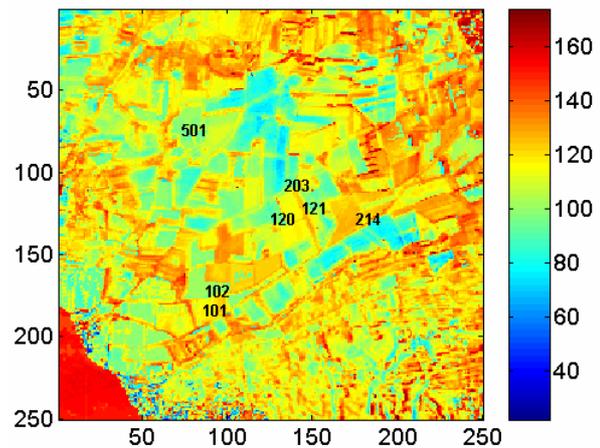
*flux Rn estimated by algorithm on day 474*



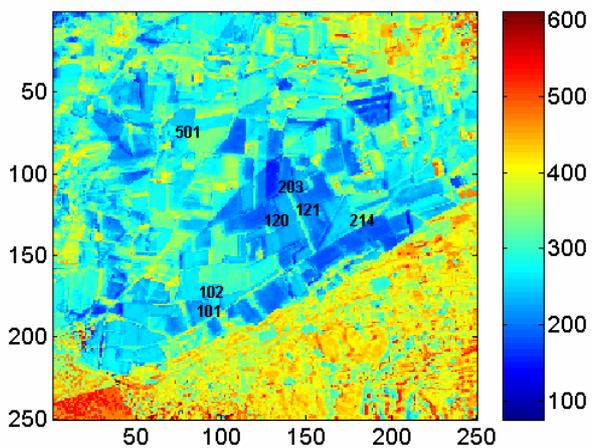
*flux G estimated by PBLs on day 474*



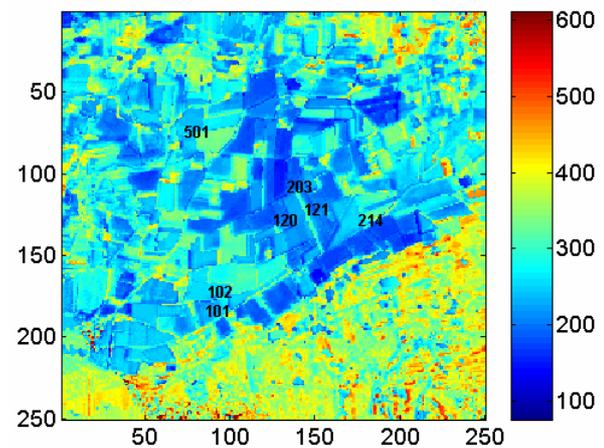
*flux G estimated by algorithm on day 474*



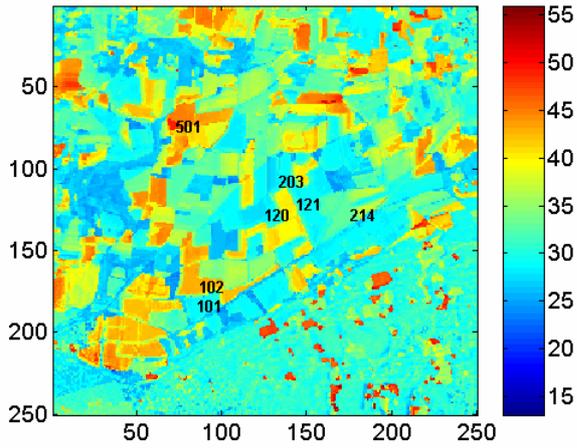
*flux H estimated by PBLs on day 474*



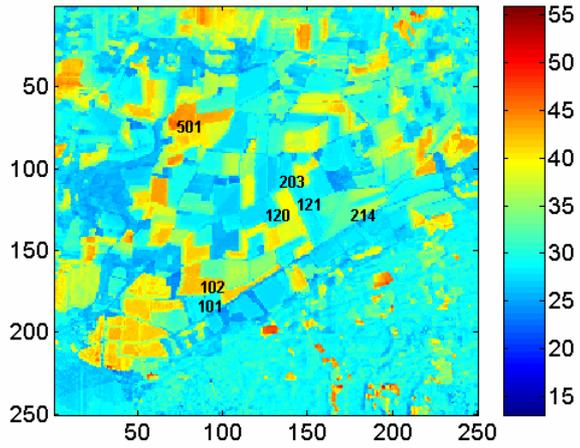
*flux H estimated by algorithm on day 474*



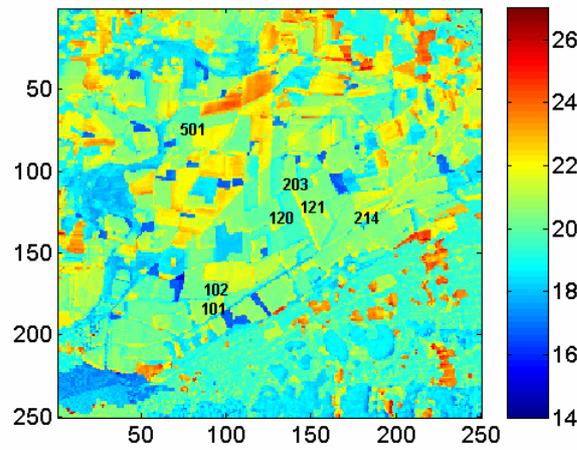
*Ts estimated by PBLs on day 474*



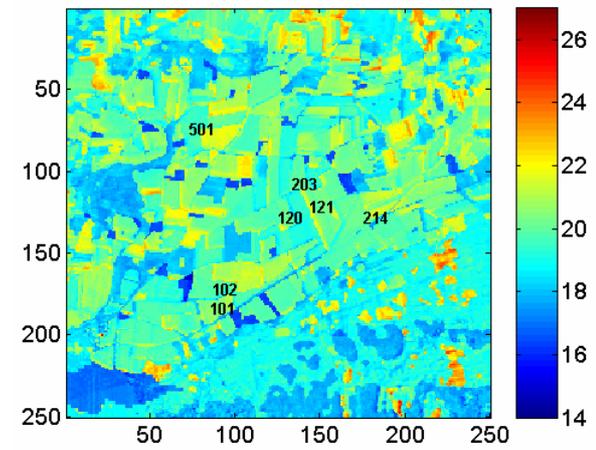
*Ts estimated by algorithm on day 474*



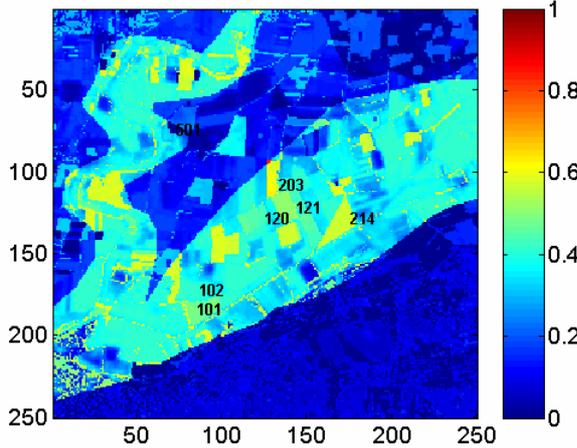
*Ta at 2 m estimated by PBLs on day 474*



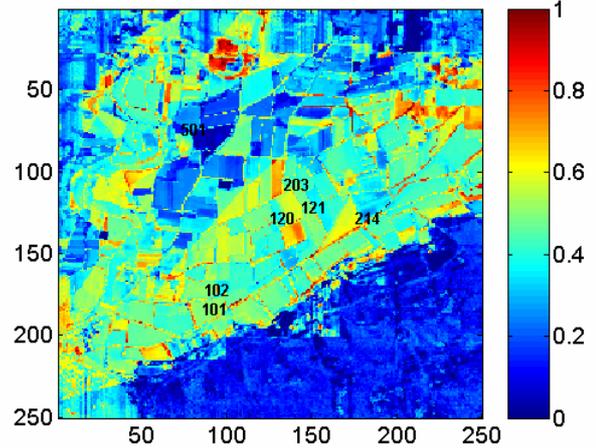
*Ta at 2 m estimated by algorithm on day 474*



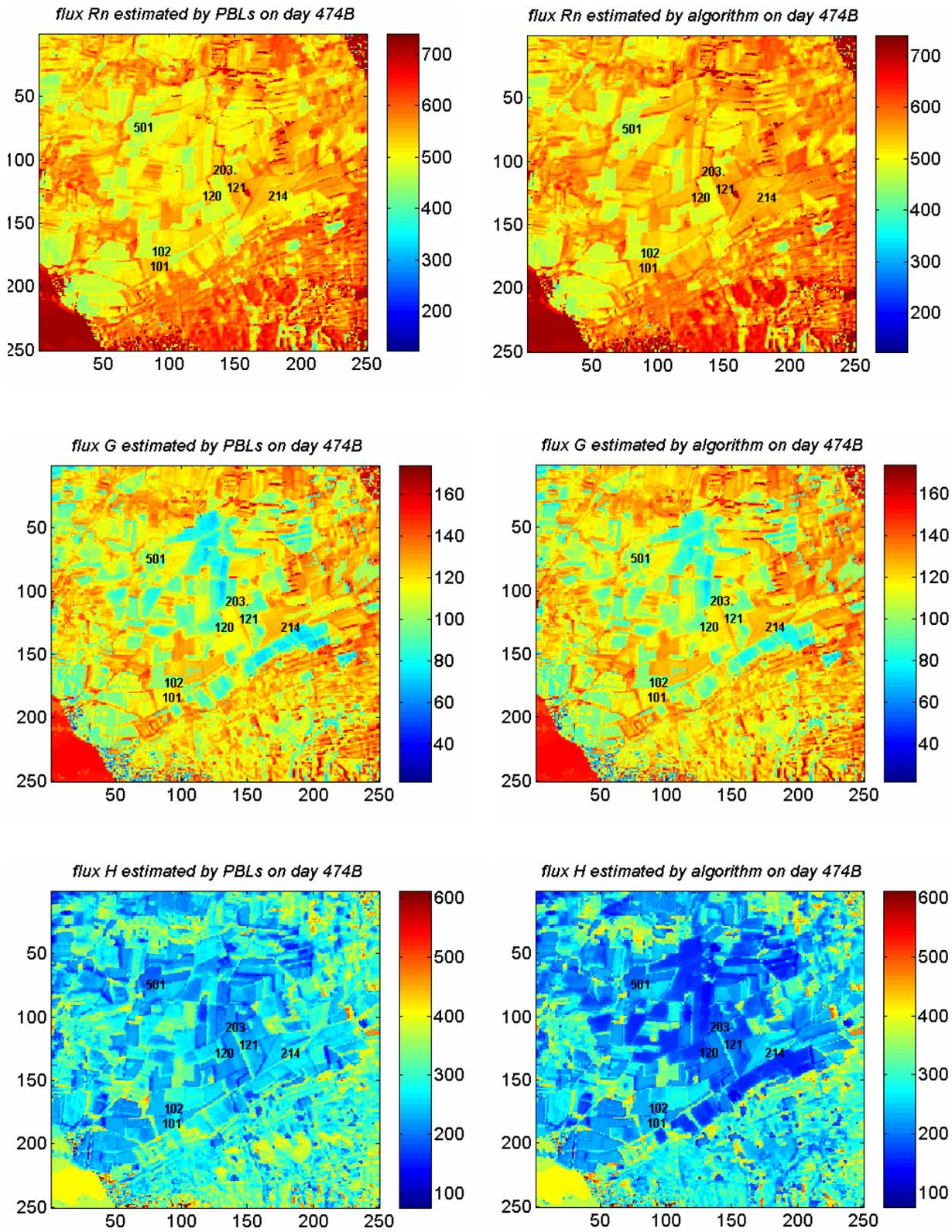
*Background soil moisture from ISBA on day 474*



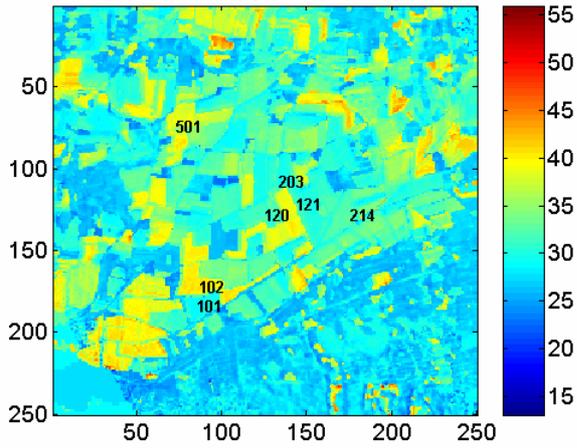
*soil moisture estimated by algorithm on day 474*



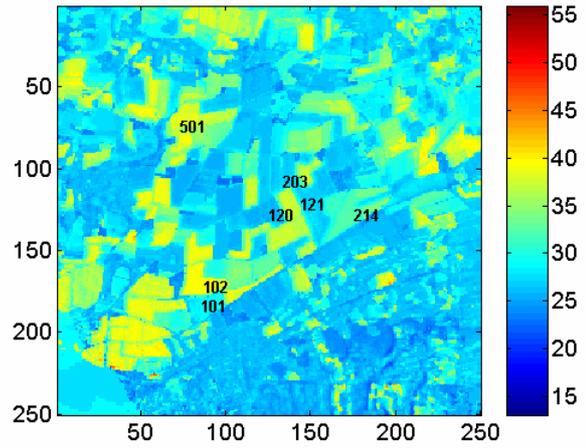
2. Estimation pour la date du 18 avril 1997 (DOE 474) avec  $f_2$  de l'éq. (V.2) : **Algorithme B**



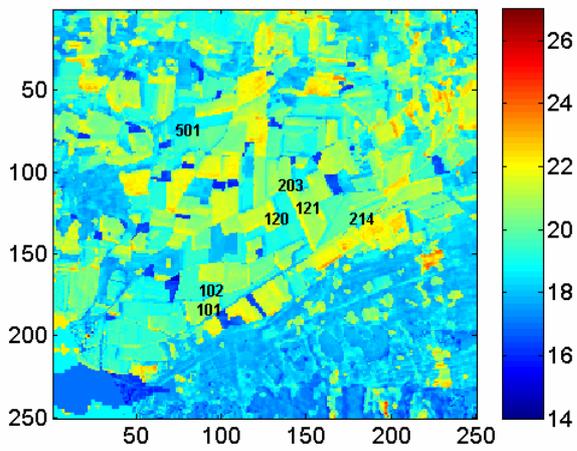
*Ts estimated by PBLs on day 474B*



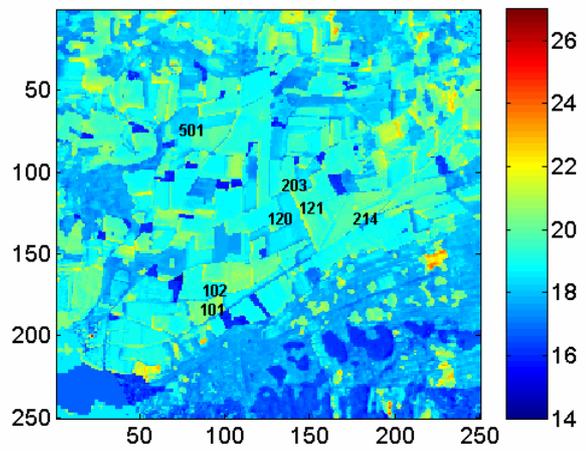
*Ts estimated by algorithm on day 474B*



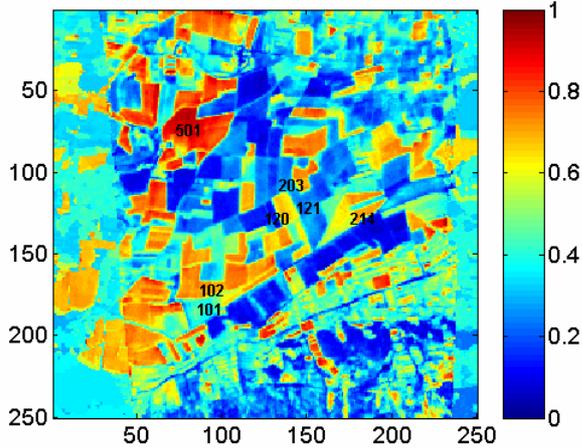
*Ta at 2 m estimated by PBLs on day 474B*



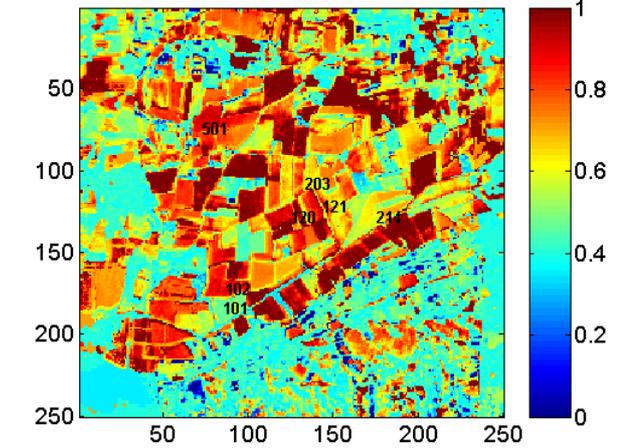
*Ta at 2 m estimated by algorithm on day 474B*



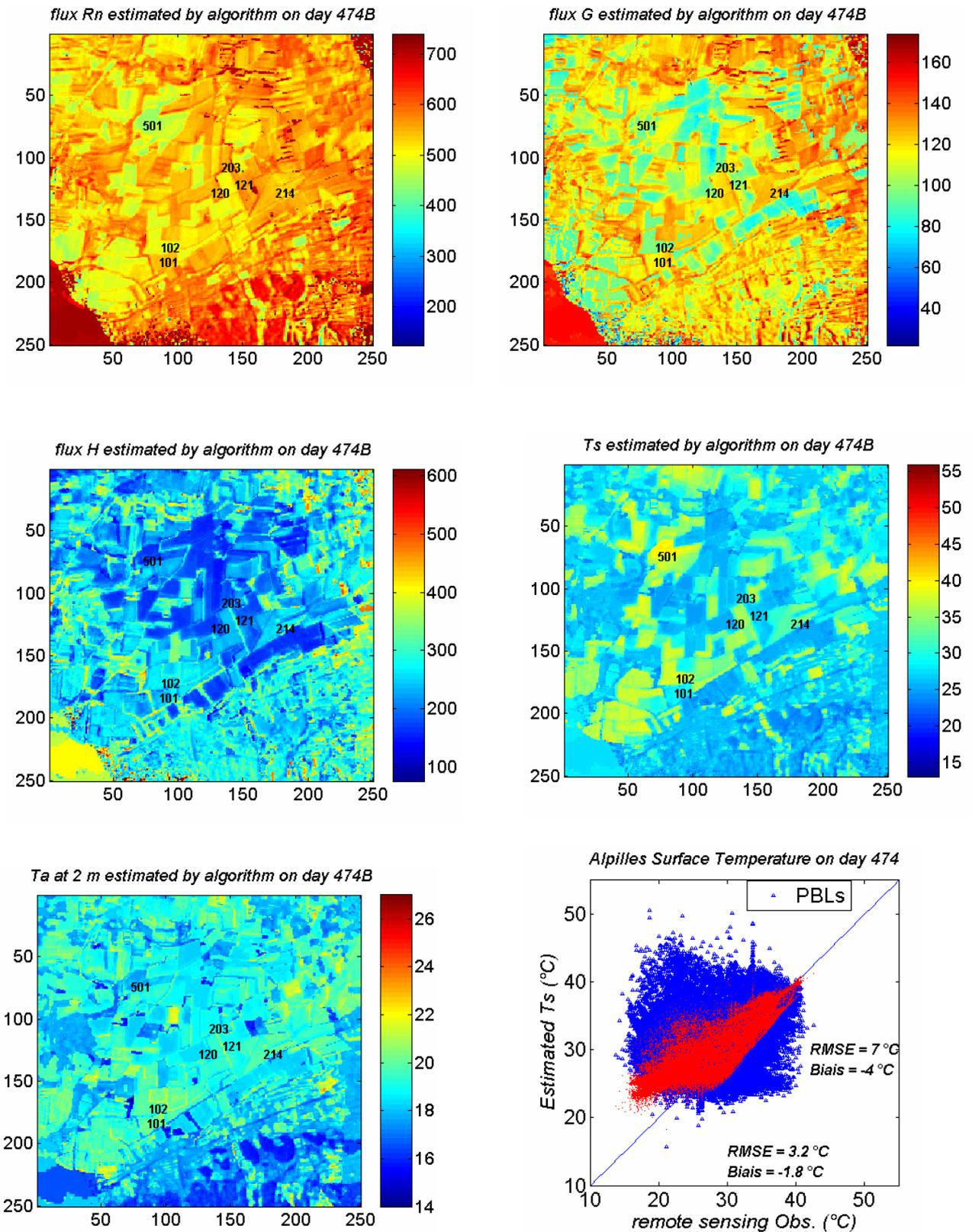
*Background soil moisture from Eq. on day 474*



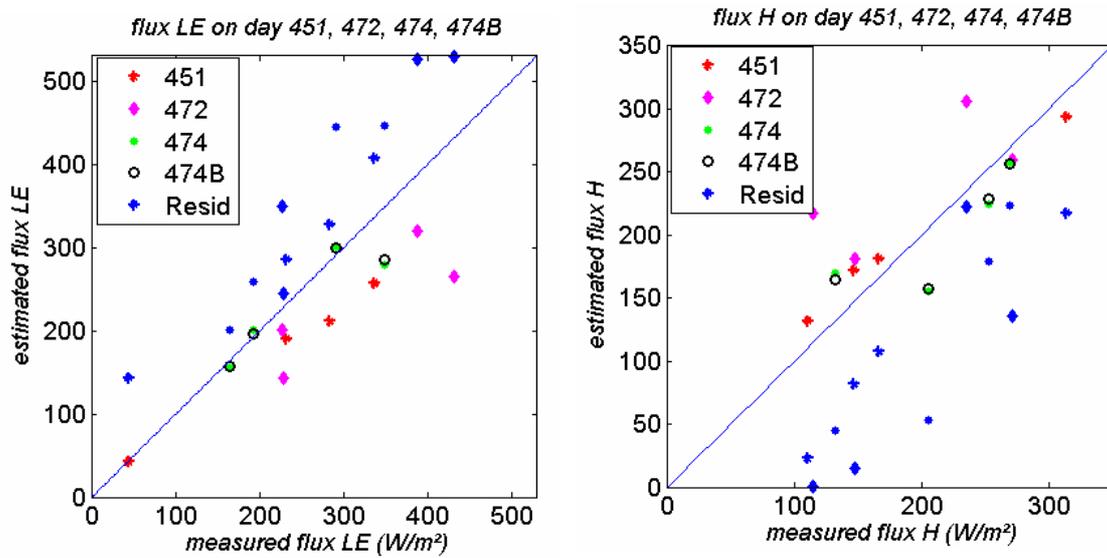
*soil moisture estimated by algorithm on day 474B*



3. Estimation pour la date du 18 avril 1997 (**DOE 474**) avec  $f_2$  de l'éq. (V.2) : **Algorithme C**  
 (résultats de PBLs ci-dessus comparés aux résultats de l'algorithme B)

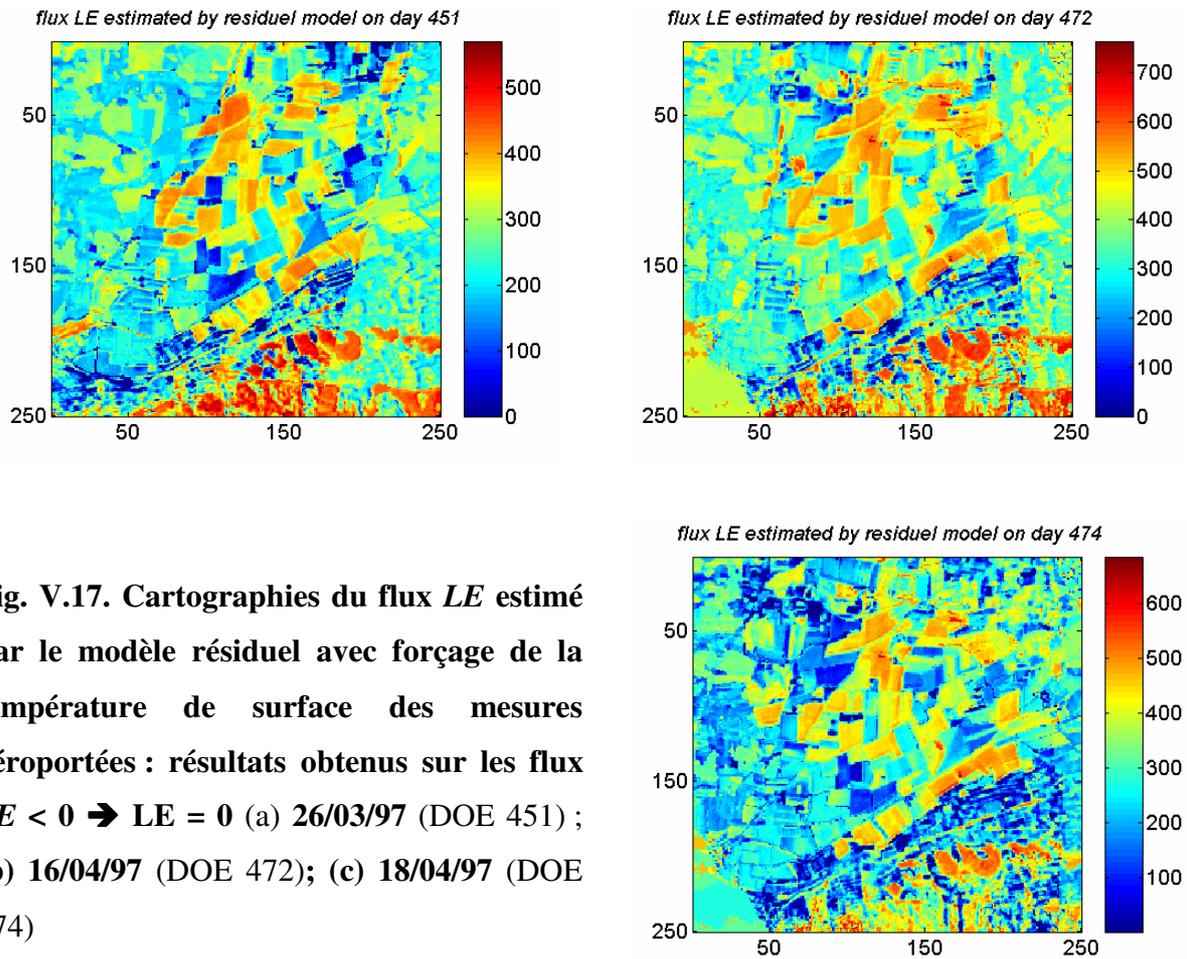


### ANNEXE IV.3 : Résultats de la méthode résiduelle sur la zone Alpilles



**Fig. V.15.** Flux *LE* (à gauche) et *H* estimés (à droite) par l’algorithme à midi solaire avec  $\sigma_i = 1,5$  K sur l’ensemble des parcelles (101, 120, 203, 214) de la zone Alpilles pour les dates 451(en étoile rouge), 472(en losange magenta), 474(en point vert), et 474B (en cercle noir) ( $f_2$  obtenue par Eq. V.2). les résultats du modèle résiduel sont (en bleu) : chaque symbole correspond à une date.

## 2. Cartographie du flux LE sur les dates 451, 472 et 474 par le modèle résiduel



**Fig. V.17.** Cartographies du flux *LE* estimé par le modèle résiduel avec forçage de la température de surface des mesures aéroportées : résultats obtenus sur les flux  $LE < 0 \rightarrow LE = 0$  (a) 26/03/97 (DOE 451) ; (b) 16/04/97 (DOE 472); (c) 18/04/97 (DOE 474)

## Curriculum Vitae

### **FORMATION**

2006-2009 **Doctorat Sciences Agronomiques, INRA & Université d'Avignon**

2003-2006 **Formation ingénieur en Calcul Scientifique, ISITV**

Institut des sciences de l'ingénieur de Toulon et du Var

2001-2003 **DEUG MIA5** mathématique et informatique appliquées aux sciences

Université du Sud Toulon et du Var

2000-2001 **Baccalauréat scientifique**

Lycée moderne de Grand-Bassam (Côte d'Ivoire)

### **Application du modèle et présentation de la méthode dans les contributions scientifiques**

**E. Kpemlie, A. Olioso, D. Courault, S. Buis, A. Bsaibes (à soumettre) Variational Data Assimilation of remote sensing data into a coupled land surface – boundary layer model for mapping evapotranspiration and microclimate**

**E. Kpemlie, A. Olioso, D. Courault, S. Buis, A. Bsaibes (2008) Assimilation de données de télédétection dans un modèle d'interaction surface – Atmosphère pour cartographier l'évapotranspiration et le microclimat** Milieux Poreux et Transferts Hydriques – Novembre 2008 – Bulletin du G.F.H.N. – N°54. p63-68

Courault, D., Bsaibes, A., **Kpemlie, E.**, Hadria, R., Hagolle, O., Marloie, O., Hanocq, J.F., Olioso, A., Bertrand, N., & Desfonds, V. (2008). **Assessing the Potentialities of FORMOSAT-2 Data for Water and Crop Monitoring at Small Regional Scale in South-Eastern France.** *Sensors*, 8, 3460-3481

**Poster iLEAPS: 17 – 21 Novembre 2008, Hyères, France: effet rétroaction surface – atmosphère**

**Kpemlie, E.**, Olioso, A., Courault, D., Bsaibes, A. (2008): **Assimilation of remote sensing data into a coupled land surface – boundary layer model for mapping evapotranspiration and microclimate.** <http://www.ileaps.org/index.php>

Courault, D., **Kpemlie, E.**, Hadria, R., Bsaibes, A., Buis, S., Olioso, A., & Hagolle, O. (2009). **Data assimilation: effect of agricultural land-use modifications on surface fluxes and microclimate.** In, *iLEAPS Newsletter Issue N° 7*: <http://www.ileaps.org/index.php> (pp. 24-25)

EGU General Assembly 2008, Vienna Austria: Présentation orale apport de l'assimilation  
Geophysical Research Abstracts, Vol. 10, EGU2008 –A-00000, 2008

**E. Kpemlie, D. Courault, S. Buis, A. Olioso, A. Bsaibes, (2008) Assimilation of remote sensing data into a coupled land surface – boundary layer model for mapping evapotranspiration and microclimate**

**Kpemlie, E.**, Courault, D., Buis, S., Olioso, A., & Bsaibes, A. (2007). **Data assimilation using remote sensing data into SVAT model for mapping evapotranspiration and microclimate.** In, *Continental Biosphere Vegetation and Water Cycle:Analyses and Prospects.*, pp 4. Paris



## **RESUME**

*La connaissance du microclimat et de l'évapotranspiration ou flux de chaleur latente qui représente la consommation réelle en eau de la culture à l'échelle des parcelles agricoles est une donnée importante pour comprendre le développement des cultures. La plupart des modèles permettant d'estimer l'évapotranspiration sont utilisées sur des surfaces homogènes sans tenir compte des interactions surface - atmosphère et de la variabilité spatiale du domaine agricole. Nous avons utilisé un modèle de couche limite atmosphérique afin de prendre en compte ces interactions. Une approche dite « patchée » permet d'introduire la variabilité spatiale des surfaces dans le modèle à partir des diverses proportions et des caractéristiques des principaux couverts végétaux qui composent le paysage. Une méthode d'assimilation variationnelle a été implémentée afin d'estimer certains paramètres du modèle difficile à connaître précisément. La méthode est basée sur le calcul de l'adjoint du modèle et utilise une température de surface observée par télédétection. L'approche développée est comparée à des approches plus simples considérant chaque type de surface indépendamment, mettant en évidence le rôle de la prise en compte de la variabilité spatiale de la surface sur la simulation du microclimat et des flux de surface.*

**Mots clés :** évapotranspiration, humidité du sol, rugosités de surface, modèle PBLs, modèle adjoint, Alpilles – ReSeDA, assimilation variationnelle, température de surface

## **ABSTRACT**

*Knowledge of climate at regional scale and evapotranspiration (or latent heat flux which represents the actual water consumption of culture) is a key to understand the development of crops. Most of the methods aiming at estimating evapotranspiration assume homogeneous or decoupled atmospheric variables over the modelling domain without accounting for the feedback between surface and atmosphere. In order to analyse such dependencies and to predict microclimate and land surface fluxes we have developed a coupled atmospheric boundary layer - land surface model which accounts for the landscape heterogeneity using a tiled approach. We have implemented appropriate procedures (variational data assimilation) for assimilating remote sensing data into the model allowing to retrieve some input parameters difficult to estimate spatially (soil moisture and aerodynamic roughness). The developed method is compared to classical approaches considering each type of surface independently. Results are discussed in this paper.*

**Key words:** evapotranspiration, soil moisture, surface roughness, PBLs model, adjoint model, Alpilles - ReSeDA, variational assimilation, surface temperature