

Étude statistique durant la campagne ParisFog

Ce chapitre présente une étude en mode prévision sur l'ensemble des situations mesurées durant ParisFog. Il vise à déterminer l'apport de la technique nudging pour décrire le forçage à méso-échelle pour prévoir un épisode de brouillard à 36 heures d'échéance. Les simulations sont effectuées avec le modèle colonne durant les cinq mois qu'ont duré la campagne. Nous présentons d'abord la méthodologie retenue puis les résultats obtenus.

Sommaire

4.1	Objectif de l'étude	89
4.2	Méthodologie	90
4.2.1	Condition initiale	91
4.2.2	Condition de forçage	92
4.2.3	Méthode statistique de vérification	92
4.3	Prévision du brouillard avec nudging	94
4.3.1	Variables de surface	94
4.3.2	Prévision du brouillard	101
4.4	Prévision du brouillard sans nudging	105
4.4.1	Variables de surface	105
4.4.2	Prévision du brouillard	106
4.5	Étude statistique avec le modèle de <i>Louis</i>	109
4.6	Discussion et conclusion	109

4.1 Objectif de l'étude

Lorenz (Lorenz, 1965) a montré que la prévision parfaite du temps ne peut être atteinte, car *i*) les équations du comportement atmosphérique ne sont pas encore parfaitement connues et sont seulement des approximations, et *ii*) l'état initial n'est pas parfaitement déterminé. En fait, même un modèle parfait ne pourrait produire une prévision

parfaite, car les erreurs dans les conditions initiales iront en s'amplifiant lors de la prévision et celle-ci divergera de la réalité, le système atmosphérique étant chaotique.

Cette difficulté provient pour une part de la complexité des mécanismes physiques mis en oeuvre lors du passage à la saturation. Ce phénomène à seuil est très sensible aux conditions initiales, mais aussi à l'équilibre délicat entre les échanges radiatif, la turbulence et les processus microphysiques dépendant entre autres de la nature et de la taille des aérosols servant de noyaux de condensation. Dans ce domaine les modèles 3D de méso-échelle n'ont pas encore la résolution verticale suffisante pour décrire cet équilibre notamment en cas de stratification thermique stable même si l'on a vu que cela restait une difficulté pour la modélisation 1D.

Aussi, aujourd'hui encore l'approche 1D est utilisée à des fins opérationnelles (Bergot et al., 2005). Dans ce domaine des progrès sensibles ont été réalisés en améliorant sensiblement la connaissance de l'état initial par assimilation de données locales à l'aide de techniques variationnelles. Il reste que, dans ces approches, le couplage avec un modèle de méso-échelle par détermination directe des termes d'advection horizontale n'a pas donné les améliorations escomptées en raison notamment de la nécessité de lisser les gradients horizontaux obtenus directement à partir des sorties des modèles opérationnels. Dans ce chapitre, nous nous proposons de revisiter ces conclusions en utilisant des techniques d'assimilation de données déjà anciennes mais simples de mise en oeuvre, basée sur des techniques de Cressman pour l'interpolation sur la verticale et de "nudging" pour le forçage synoptique à méso-échelle. La démarche que nous avons suivie est la suivante.

Dans une première étape le modèle 1D a été validé sur les situations les mieux documentées de l'expérience ParisFog. Ce travail présenté dans Zhang et al. (in revision) a permis de définir une version de référence du modèle. Cette version est à l'état de l'art en ce qui concerne la turbulence, la représentation des nuages dans les modèles de transfert radiatif, la modélisation de la nucléation et de la sédimentation à l'aide d'un modèle semi-spectral donnant accès à la distribution en taille de gouttes d'eau. Dans ce travail la technique d'assimilation de données par Cressman et nudging a été utilisée pour simuler l'évolution de la couche de brouillard dans des conditions les plus proches possibles des observations afin de tester les différentes paramétrisations déjà évoquées.

Nous proposons d'utiliser le modèle 1D en mode prévision durant l'ensemble de la campagne ParisFog afin d'étudier l'apport d'un couplage avec le modèle de méso-échelle MM5. Nous présentons d'abord la méthodologie utilisée (conditions initiales, forçage atmosphérique, . . .), les scores statistiques utilisés, puis les résultats obtenus pour différentes options concernant le nudging sont présentés.

4.2 Méthodologie

Afin de tester cette méthode nous avons travaillé sur l'ensemble de la période ParisFog, de novembre 2006 à mars 2007, pour laquelle nous disposons de toutes les données nécessaires à cette étude. Ainsi nous réalisons chaque jour une prévision à 36 h en partant

de 00 TU ou 12 TU, heures où les données du radiosondage de Trappes sont disponibles pour initialiser le modèle.

4.2.1 Condition initiale

La reconstruction d'un champ initial est une étape importante, la restriction du nombre et de la qualité des données oriente la manière d'initialiser le modèle. On va donc lister d'abord les différentes données disponibles pendant la période simulée.

1. Les données issues du modèle MM5 (The Fifth-Generation NCAR/Penn State Mesoscale Model) sur 4 jours, de J à $J + 3$. Il est installé et utilisé en prévision expérimentale quotidienne au LMD à l'École Polytechnique pour une prévision à 96 h à partir des réseaux de 00 TU. Il prévoit et simule l'ensemble des paramètres météorologiques avec notamment une description explicite des nuages avec une résolution horizontale d'environ 5 km. On peut observer des écarts importants entre les valeurs données par ce modèle et les mesures in-situ au voisinage du sol.
2. Les radiosondages de Trappes. Ces radiosondages fournissent des données sur la température, l'humidité, le vent et la pression. Ces données sont en général assez fiables malgré quelques données manquantes. Au voisinage du sol, dans les premières centaines de mètres, elles peuvent aussi s'écarter des mesures du SIRTA à Palaiseau en raison de la variation des conditions locales entre Trappes et Palaiseau distants d'environ 23 km.
3. Les mesures des mâts de 30 mètres installés en zone 1 et zone 3 sur le site du SIRTA.

Pour nos simulations, le modèle est initialisé à 00 TU (ou 12 TU) en utilisant les radiosondages du Trappes en couplant avec les données du mât pour les niveaux bas. Pour compléter l'initialisation, on utilise les données provenant du dispositif dans le sol du SIRTA pour calculer la température profonde dans le sol, intervenant dans le bilan énergétique de surface comme température de rappel. Celle-ci est considérée comme égale à la moyenne sur les 5 derniers jours de la température mesurée à -50 cm.

Un contrôle des données a été effectué afin de retirer les situations où des données sont manquantes ou manifestement erronées, comme par exemple dans le cas de température en K négative. Au total 55 échantillons sont retenus pour cette étude, dont 4 échantillons seront retirés lors du calcul de l'occurrence des situations de brouillard (en raison des données erronées de visibilité).

Normalement, les modèles de prévision doivent être construits de sorte que les données utilisées pour caler les paramètres diffèrent des données servant à évaluer les résultats. Un jeu d'apprentissage et un jeu de validation sont donc envisagés. En raison de faible nombre d'échantillons que l'on a, nous avons choisi de prendre directement le jeu de validation. De plus, un test de sensibilité sur les paramètres du modèle de sol-atmosphère concernant la constante C_{sol} est effectué pour estimer l'erreur de prévision entre différentes valeurs de cette constante. Les écarts de l'erreur quadratique moyenne de la température et de l'humidité relative à 2 mètres pour deux valeurs de C_{sol} : $C_{sol} = 0.4 \cdot 10^{-5} \text{ m}^2 \cdot K \cdot J^{-1}$

TAB. 4.1 : Les simulations retenues pour l'étude statistique.

<i>Date</i>	<i>Date</i>	<i>Date</i>	<i>Date</i>	<i>Date</i>
20061108	20070103	20070122	20070206	20070313
20061109	20070104	20070123	20070209	20070314
20061218	20070105	20070124	20070210	20070315
20061219	20070110	20070127	20070228	20070316
20061220	20070111	20070130	20070301	20070319
20061223	20070112	20070131	20070302	20070320
20061224	20070113	20070201	20070306	20070321
20061227	20070114	20070202	20070307	20070322
20061228	20070115	20070203	20070310	20070325
20061229	20070116	20070204	20070311	20070326
20061230	20070121	20070205	20070312	20070327

et $C_{sol} = 0.6 \cdot 10^{-5} \text{ m}^2.K.J^{-1}$, sont de l'ordre de $0.3 \text{ }^\circ\text{C}$ et 4% , respectivement. L'écart est relativement faible, on peut donc négliger cette différence. La valeur $C_{sol} = 0.4 \cdot 10^{-5} \text{ m}^2.K.J^{-1}$ est retenue (faible erreur quadratique moyenne).

4.2.2 Condition de forçage

Le forçage météorologique de grande échelle est introduit par un terme de nudging en utilisant comme profil de rappel les prévisions MM5 à 36 h par pas de 1 h . Notons qu'au dessus du domaine de calcul (2500 m), les profils de vent, de température et d'humidité sont utilisés afin de définir les couches nuageuses supérieures dont la description est nécessaire aux calculs radiatifs pour le réchauffement/refroidissement des couches atmosphériques et pour estimer les flux radiatifs atteignant le sol. Dans cette étude, un profil linéaire est utilisé qui correspond à une variation linéaire du coefficient de nudging de 10^{-4} s^{-1} à la surface et à 10^{-3} s^{-1} à 15000 m .

4.2.3 Méthode statistique de vérification

Selon les paramètres que l'on va évaluer (mesure d'un paramètre comme la température ou occurrence d'un phénomène comme la présence de brouillard) deux types de méthodes de vérification sont utilisés pour cette étude afin d'évaluer la capacité de prévision de notre modèle. Pour les paramètres directement mesurés en continu on utilisera le biais et l'erreur quadratique moyenne. Le deuxième type de score pour l'occurrence du brouillard est présenté sous forme d'une table de contingence à deux dimensions. Les catégories prévues et observées sont simplement classifiées dans une table de 2 lignes et de 2 colonnes (voir Tab. 4.2). Dans cette table, il y a une ligne pour chaque catégorie observée et une colonne pour chaque catégorie prévue. Une prévision "*hit*" signifie que le modèle prévoit une situation de brouillard et qu'il est présent ; une prévision "*false alarm*" signifie que

TAB. 4.2 : Table de contingence à 2×2 dimensions. a représente "hit" événements, b "missed", c "false alarms" et d "correct rejections".

	Prévision <i>oui</i>	Prévision <i>non</i>	Total
Observation <i>oui</i>	a (hit)	b (missed)	$a + b$
Observation <i>non</i>	c (false)	d (correct rejection)	$c + d$
Total	$a + c$	$b + d$	$a + b + c + d$

le modèle a prévu une situation de brouillard et qu'il n'est pas présent ; une prévision "missed" signifie que le modèle a prévu une journée sans brouillard mais qu'il était présent. A partir de cette table on peut former des scores de réussite de la prévision en termes d'indicateurs. Les formulations des indicateurs utilisés dans cette étude sont présentées ci dessous.

Méthode continue

1. Biais (B), il est une moyenne arithmétique, c'est-à-dire la somme des valeurs numériques divisée par le nombre de ces valeurs numériques. Il indique une déviation par rapport à l'observation, cependant il ne reflète pas l'amplitude de l'erreur. Autrement dit, il caractérise seulement l'erreur systématique du modèle (premier moment de la distribution des erreurs de prévision).

$$B = \frac{1}{N} \sum_{i=1}^N (F_i - O_i), \quad (4.1)$$

où F_i représente la valeur prévue, O_i représente la valeur observée, et N est la taille de l'échantillon.

2. L'erreur quadratique moyenne (Root-Mean-Square Error, RMSE). L'RMSE mesure l'écart moyen des scores (la dispersion des valeurs) par rapport à la moyenne arithmétique, et reflète ainsi la variabilité à l'intérieur de l'échantillon de données. Plus il est faible plus le modèle est précis (deuxième moment de la distribution des erreurs de prévision).

$$RMSE = \left[\frac{1}{N} \sum_{i=1}^N (F_i - O_i)^2 \right]^{\frac{1}{2}} \quad (4.2)$$

Scores associé avec une table de contingence à 2 dimensions

Une table de contingence est une façon de représenter des données d'énumération (de comptage) d'individus classés en catégories. Une table de contingence est déjà le résultat d'un traitement des données, puisque les individus (observations) ont dû être préalablement classés, puis comptés. A partir de cette table de contingence on peut définir plusieurs scores :

1. Le pourcentage de bonnes prévisions (*percent correct*, PC). La somme des éléments diagonaux est divisée par le nombre total d'événement. Sa valeur en cas de prévision parfaite est égale à 1.

$$PC = \frac{a + d}{a + b + c + d}$$

2. La probabilité de détection (*probability of detection*, POD.). Elle mesure la capacité d'une prévision correcte (hit rate). La POD est souvent utilisé conjointement avec la proportion de fausse alarme (*fause alarm ratio*, FAR) pour un phénomène rare. Ces deux indicateurs varient entre 0 et 1. Leur valeur pour la prévision parfaite est égale à 1 pour la *POD* et 0 pour la *FAR*.

$$POD = \frac{a}{a + b}, \quad FAR = \frac{c}{a + c}$$

3. L'indice de succès critique (*critical success index*, CSI, ou *threat score*, TS). L'avantage de cet indicateur est d'être sensible à la fois à l'événement de fausse alarme et à l'événement de mauvaise détection. Il donne une représentation précise de ces deux situations. La fourchette des valeurs est comprise entre 0 et 1. La valeur de la prévision parfaite est égale à 1.

$$CSI = \frac{a}{a + b + c}$$

4. Indice de qualité d'Heidke (*Heidke skill score*, HSS). L'indice d'Heidke varie entre $-\infty$ et 1. Il vaut 1 dans le cas d'une prévision parfaite, il est négatif quand la prévision est moins bonne que celle obtenue par le hasard.

$$HSS = \frac{2(ad - bc)}{(a + b)(b + d) + (a + c)(c + d)}$$

5. Hanssen-Kuiper Score (appelé souvent *true skill statistic*, TSS). Il mesure la capacité d'un modèle de prévision en séparant les situations de brouillard des situations de non-brouillard. La fourchette de cet indicateur varie entre -1 et 1. Une valeur de -1 signifie que le modèle n'a aucune capacité prédictive (quand toutes les prévisions sont ratées). La valeur de la prévision parfaite est égale à 1.

$$TSS = \frac{ad - bc}{(a + b)(c + d)}$$

Il existe encore d'autres méthodes possibles (Jolliffe and Stephenson, 2003), mais les scores présentés ci-dessus ont été retenus parce qu'ils sont relativement simples à utiliser et d'emploi courant.

4.3 Prédiction du brouillard avec nudging

4.3.1 Variables de surface

A noter que les simulations de périodes validation sont en réalité une succession de simulations de 36 h pour chacune des quelles les conditions initiales sont définies en deux cas différents : un démarrant à 00 TU (IC-00) et l'autre démarrant à 12 TU (IC-12).

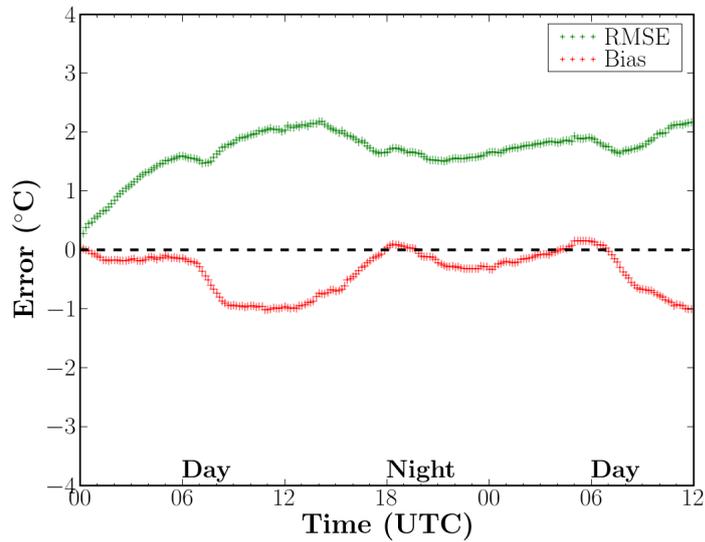
Cas IC-00

Dans cette section, les simulations sont démarrées à 00 *TU*. Avant de s'intéresser aux résultats concernant la prévision du brouillard, il est tout d'abord nécessaire de s'assurer de la bonne qualité prédictive générale de notre modèle 1D. La comparaison des paramètres météorologiques au voisinage du sol sont présentées en fonction de l'heure sur les figures suivantes.

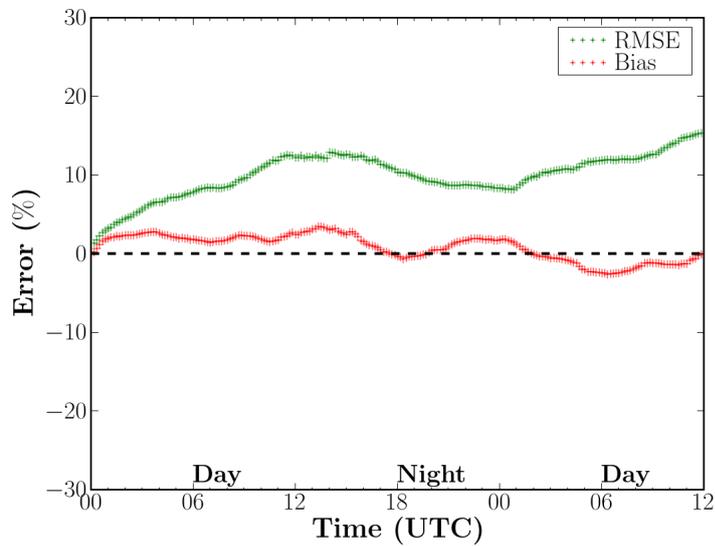
Le biais et l'erreur quadratique moyenne de la température et de l'humidité relative à 2 mètres en fonction d'échéance sont présentés dans la figure 4.1. Dans l'ensemble, les valeurs obtenues par notre modèle en démarrant à 00 *TU* se comparent assez favorablement aux mesures. On constate que la température et l'humidité relative à 2 *m* prévues sont très peu biaisées. Le biais de température évolue en effet entre -1 et 0 °C. L'erreur quadratique moyenne évolue entre 1 et 2 °C. Cela signifie que la température est un peu sous estimée pendant la période de prévision. Ce qui reste dans le même ordre de grandeur que l'erreur quadratique moyenne habituellement rencontrée dans des modèles de prévision numérique comme ALADIN (référence : *Contrôle des produits numérique utilisés pour la prévision météorologique* par Météo-France). Pour l'humidité relative à 2 *m*, le biais reste assez faible et inférieur à 5% et l'erreur quadratique moyenne est dans des ordres de grandeur classique entre 8% et 15%, comparativement à une erreur de 7.5% et 15% pour d'autres modèles.

Le biais et l'erreur quadratique moyenne de la vitesse du vent et de l'énergie cinétique turbulente à 10 mètres en fonction d'échéance sont présentés dans la figure 4.2. Les mesures considérées ici sont celles des anémomètres ultrasoniques de la zone 3. À 10 *m*, la vitesse du vent est peu biaisée pour une plage horaire allant de 00 *TU* à 12 *TU* correspondant à une transition entre régime stable et instable. Au delà de 12 *TU*, le modèle surestime la vitesse de 0.5 $m.s^{-1}$, tandis que le biais de l'énergie cinétique turbulente devient positif (autour de 0.1 $m^2.s^{-2}$). L'erreur quadratique moyenne est de l'ordre de 1.5 $m.s^{-1}$. Ces résultats sont satisfaisants puisque notre modèle permet de prendre en compte les spécificités locales du site à travers le réglage des constantes de sol, même si le modèle de sol utilisé est assez simplifié. Il est difficile de connaître les raisons du biais positif de l'énergie cinétique turbulente après 18 *TU*. Il peut être lié à la surestimation de la vitesse du vent dû au forçage de grande échelle.

La comparaison des flux radiatifs au niveau 2 mètres fournis par le modèle et les mesures. Le flux d'infrarouge descendant et le flux solaire descendant à 2 mètres en fonction du temps sont présentés à la figure 4.3. Pour le rayonnement infrarouge, on observe un assez bon accord entre le calcul et les mesures (Fig. 4.3(a)) avec un biais pratiquement nul et une erreur quadratique moyenne de 11 $W.m^{-2}$, sauf durant la période nocturne, le modèle n'ayant pas prévu correctement la présence de la couverture nuageuse pendant toute cette période. On observe une erreur importante pour le flux solaire (Fig. 4.3(b)) avec une valeur maximale d'erreur quadratique moyenne d'environ 80 $W.m^{-2}$. Cette erreur est due aux erreurs sur le contenu en aérosols et en gouttelettes d'eau. En fait, par temps clair, le rayonnement solaire est bien plus élevé. Par contre, quand l'atmosphère est chargée de

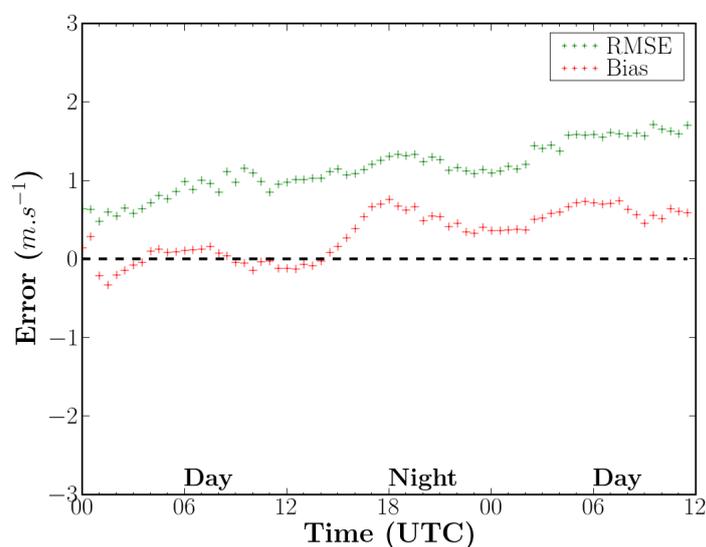


(a) Température à 2 m

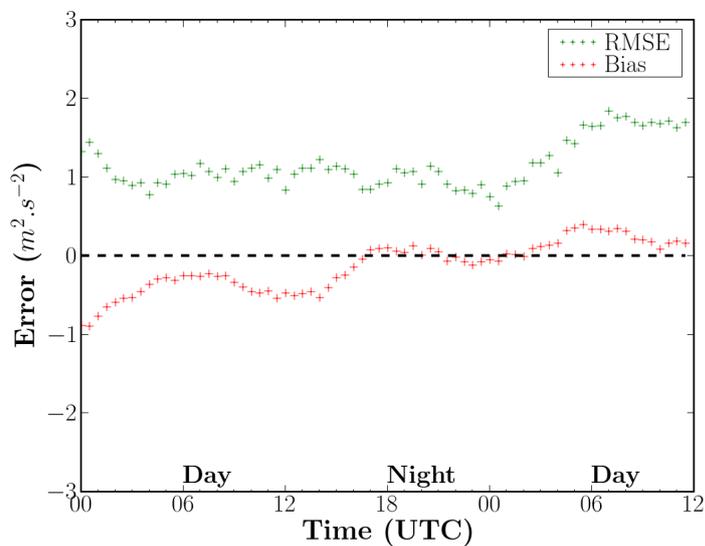


(b) Humidité relative à 2 m

FIG. 4.1 : Prédiction avec nudging, cas démarré à 00 TU (IC-00). Biais (rouge) et erreur quadratique moyenne (verte) de prévision à 36 heures par rapport aux mesures pour la température et l'humidité relative à 2 m.



(a) Vitesse du vent à 10 m



(b) Énergie cinétique turbulente à 10 m

FIG. 4.2 : Prédiction avec nudging, cas démarré à 00 TU (IC-00). Biais (rouge) et erreur quadratique moyenne (verte) de prédiction à 36 heures par rapport aux mesures pour la vitesse du vent et l'énergie cinétique turbulente à 10 m.

gouttelettes d'eau ou de poussières (aérosols), le rayonnement solaire au sol est atténué. Au total, le rayonnement reçu en atmosphère turbide est moins élevé qu'en atmosphère limpide, à cause de la forte atténuation subie par le rayonnement solaire direct.

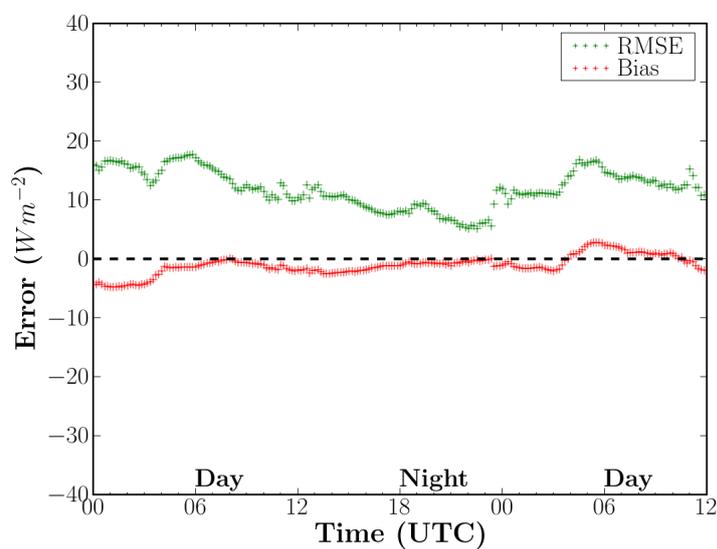
Test sur les journées de ciel clair (Fig. 4.4)

Afin de tester notre paramétrisation du rayonnement, un test "ciel clair" est effectué. C'est-à-dire en l'absence de nuage et de brouillard dans le domaine de calcul, la teneur en eau liquide q_l est forcée à 0. Au total, 5 journées ont été simulées : les 20061220, 20070203, 20070311, 20070312 et 20070326. On constate que les erreurs quadratiques moyennes pour les deux flux radiatifs sont bien réduites, notamment pour le flux solaire ($RMSE_{solar} < 15 W.m^{-2}$). On constate toutefois un fort biais entre 00 TU et 04 TU sur le flux infrarouge (Fig. 4.4(a)). La mesure par télémètre indique que pendant cette période du 3 février 2007 on observe une valeur plus élevée de ce flux en raison de la présence d'une couche nuageuse située entre 400 m et 1000 m et non représentée par le modèle. Cela conduit à une erreur importante dans la statistique.

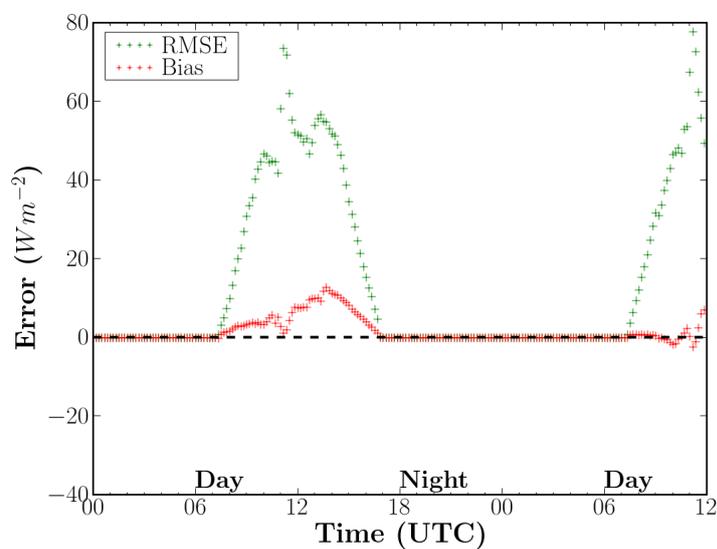
Cas IC-12

Pour le cas démarrant à 12 TU (IC-12), le biais et l'erreur quadratique moyenne des paramètres météorologiques au voisinage du sol en fonction de l'échéance sont présentés dans les figures 4.5, 4.6 et 4.7. On observe des erreurs importants au départ de la simulation (12 TU). C'est en fait dû à la présence d'une couche nuageuse à haute altitude non représentée par le modèle. Les valeurs d'erreur quadratique moyenne sont un peu plus élevées que celles du cas IC-00. Le biais de la température varie avec le cycle diurne (Fig. 4.5(a)). Au début de la simulation, en raison de la surestimation du flux de rayonnement solaire (comme mentionné dans le cas IC-00), la température est un peu surestimée dans l'après-midi, de l'ordre de 1 °C. Le modèle tend à refroidir durant la période du lever du soleil (de l'ordre de 1 °C). Cela peut s'expliquer par le choix des paramètres de l'interface sol-atmosphère, mais aussi par la tendance du modèle à créer un peu trop facilement de l'eau liquide. Après, la surestimation du flux solaire tend à réchauffer le sol puis la température au contact jusqu'au coucher du soleil. En ce qui concerne l'humidité relative, elle varie en fonction de la température comme attendu (Fig. 4.5(b)).

Pour le vent et l'énergie cinétique turbulente (Fig. 4.6), les biais sont similaires à ceux du cas IC-00 mais sont renforcés. Le vent est très biaisé (Fig. 4.6(a)) avec un écart moyen supérieur à $1 m.s^{-1}$. On constate que le plus mauvais biais négatif de la vitesse du vent se trouve vers 13 TU (de l'ordre de $1 m.s^{-1}$) et le plus mauvais biais positif se trouve vers 18 TU (de l'ordre de $2 m.s^{-1}$). De plus, la valeur d'erreur quadratique moyenne de vitesse est de l'ordre de $2 m.s^{-1}$. Cette forte valeur du $RMSE$ est cependant fortement influencée par quelques situations pour lesquelles une grosse erreur a été constatée. Ceci peut être lié à la paramétrisation dynamique dans le modèle. Or il est probablement dû à la façon du calcul des scores qui est une incertitude non physique. Puisque la valeur d'erreur quadratique moyenne est sensible à l'influence de grosse erreur que le petit nombre d'échantillons. En

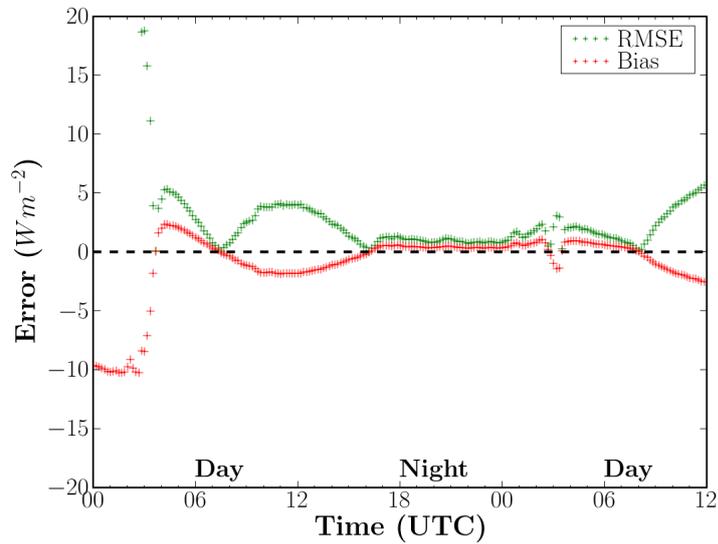


(a) Flux IR descendant à 2 m

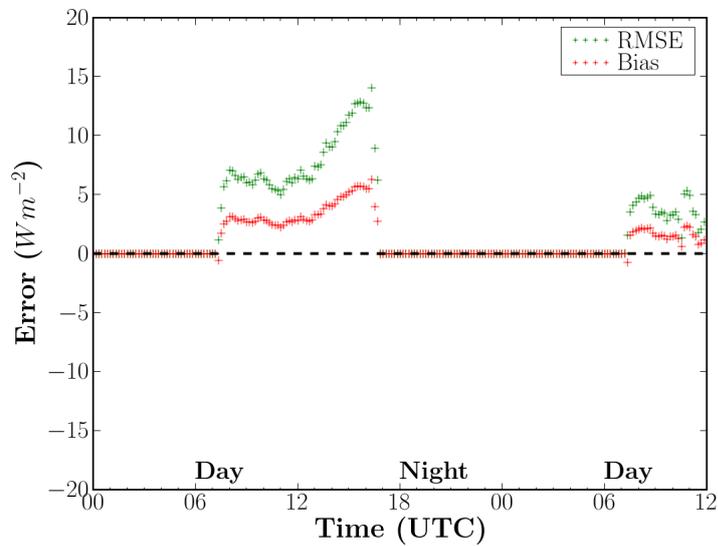


(b) Flux solaire descendant à 2 m

FIG. 4.3 : Prédiction avec nudging, cas démarré à 00 TU (IC-00). Biais (rouge) et erreur quadratique moyenne (verte) de prédiction à 36 heures par rapport aux mesures pour les flux radiatifs descendants à 2 m.



(a) Flux IR descendant à 2 m



(b) Flux solaire descendant à 2 m

FIG. 4.4 : Test sur les journées de ciel clair (20061220, 20070203, 20070311, 20070312 et 20070326). Les biais (rouge) et les erreurs quadratiques moyennes (rouge) pour les flux radiatifs descendants à 2 m.

TAB. 4.3 : Table de contingence à 2 dimensions. Cas démarré à 00 TU avec nudging pour 36 h d'échéance.

	Prévision <i>oui</i>	Prévision <i>non</i>	Total
Observation <i>oui</i>	10 (hit)	1 (missed)	11
Observation <i>non</i>	15 (false)	25 (correct rejection)	40
Total	25	26	51
Scores	$PC = 0.68, POD = 0.91, FAR = 0.60, CSI = 0.38, TSS = 0.53, HSS = 0.36$		

raison de la surestimation de la vitesse du vent et l'effet du nudging, la turbulence est finalement assez peu biaisée.

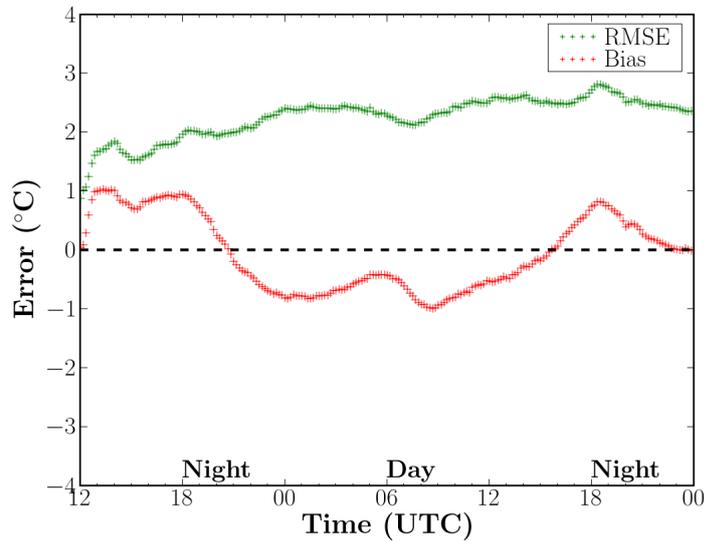
Comme nous l'avons vu lors du cas IC-00 pour la distribution d'erreur du flux radiatif, on surestime le flux solaire descendant à 2 m (de l'ordre de $80 W.m^{-2}$ au maximum) (non présenté). Cette erreur est due à la mauvaise prévision de la couverture nuageuse. Celle-ci peut aussi expliquer les valeurs du *RMSE* pour le flux infrarouge.

4.3.2 Prévision du brouillard

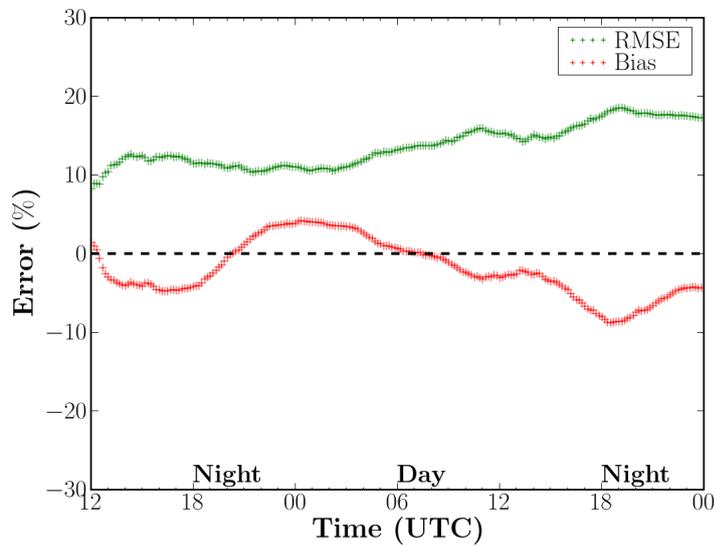
Dans cette section, les résultats de l'occurrence des situations de brouillard sont présentés. On s'intéresse à la présence du brouillard en termes de tableaux de contingence et à la durée du brouillard en termes d'heures de formation et de dissipation. La classification des situations de brouillard est présentée au *chapitre 2*. Ici, on rappelle qu'une situation de brouillard est définie par une visibilité inférieure à 1000 m sur une période d'au moins de 30 minutes.

Les tableaux de contingences sur l'occurrence du brouillard montrent que, pendant la période simulée, dans le cas IC-00 (Tab. 4.3) le brouillard est plus fréquent que dans le cas IC-12 (Tab. 4.4). Environ 68% de prévision sont correctes dans le cas IC-00 et 41% dans le cas IC-12. En gros, plus de 90% des situations du brouillard observé sont correctement prévues dans le cas IC-00, et seulement 45% dans le cas IC-12. De plus, il y a environ 60% de *FAR* dans le cas IC-00, et plus de 80% de *FAR* dans le cas IC-12. Une valeur de *CSI* supérieure à 0.38 signifie qu'il y a plus d'environ 1/3 d'épisodes de brouillard (observé et/ou prévu) qui est correctement prévu. On remarque que le modèle surestime le champ d'humidité, ceci conduit à augmenter la fréquence de formation du brouillard. Qui peut expliquer pourquoi nous avons une valeur élevée de *POD* pour le cas IC-00. La valeur de *TSS* (supérieure à 0.50 pour le cas IC-00) montre que notre modèle a toutefois une assez bonne capacité à prévoir le brouillard.

La figure 4.8 illustre la distribution des erreurs sur les heures de formation et dissipation du brouillard pour le cas IC-00 à 36 h d'échéance. Une occurrence de brouillard est définie sur une période de 36 h. Sur l'heure de formation on observe que, dans environ 65% des cas, l'intervalle de l'heure de formation du brouillard entre prévision et observation est inférieur à 3 h. Dans 10% des cas, l'heure de formation est retardé d'environ 20 h.

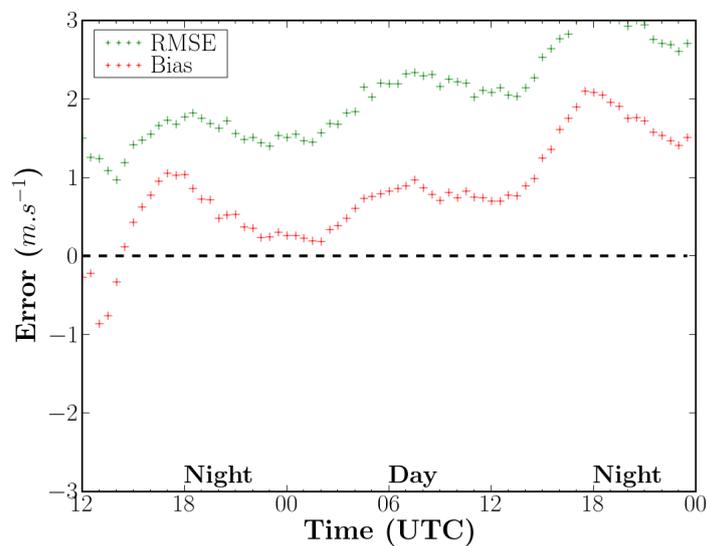


(a) Température à 2 m

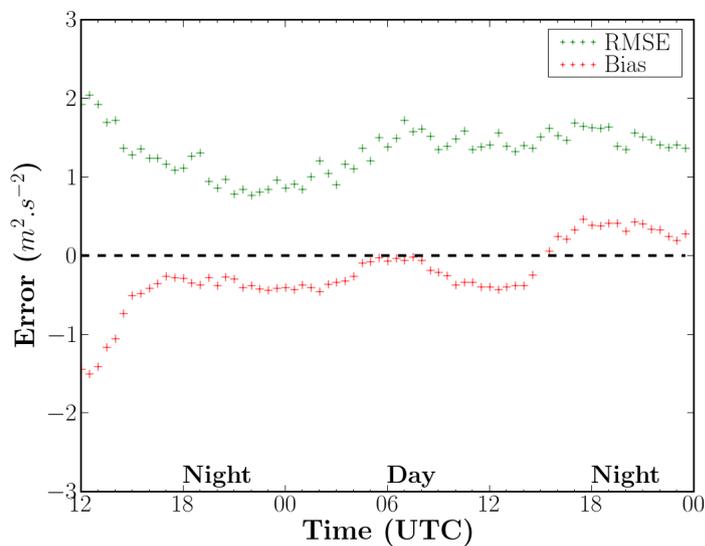


(b) Humidité relative à 2 m

FIG. 4.5 : Prédiction avec nudging, cas démarré à 12 TU (IC-12). Biais (rouge) et erreur quadratique moyenne (verte) de prévision à 36 heures par rapport aux mesures pour la température et l'humidité relative à 2 m.



(a) Vitesse du vent à 10 m



(b) Énergie cinétique turbulente à 10 m

FIG. 4.6 : Prédiction avec nudging, cas démarré à 12 TU (IC-12). Biais (rouge) et erreur quadratique moyenne (verte) de prédiction à 36 heures par rapport aux mesures pour la vitesse du vent et l'énergie cinétique turbulente à 10 m.

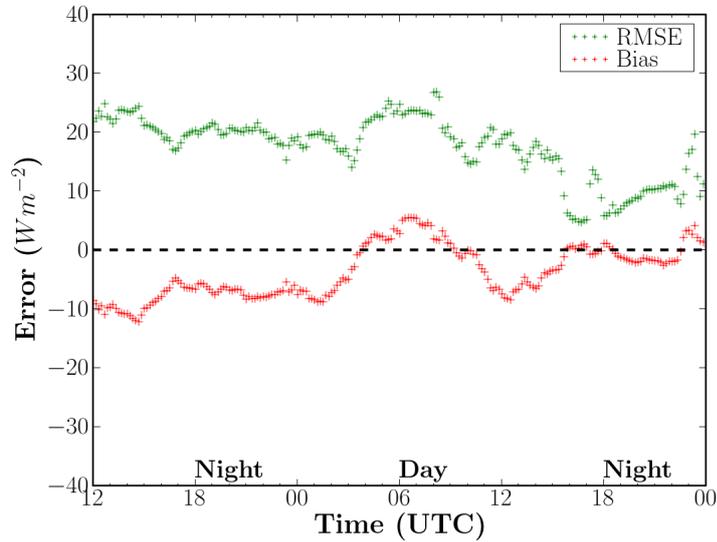


FIG. 4.7 : Prédiction avec nudging, cas démarré à 12 TU (IC-12). Biais (rouge) et erreur quadratique moyenne (verte) de prédiction à 36 heures par rapport aux mesures pour le flux infrarouge descendant à 2 m.

TAB. 4.4 : Table de contingence à 2 dimensions. Cas démarré à 12 TU avec nudging pour 36 h d'échéance.

	Prédiction <i>oui</i>	Prédiction <i>non</i>	Total
Observation <i>oui</i>	5 (hit)	6 (missed)	11
Observation <i>non</i>	24 (false)	16 (correct rejection)	40
Total	29	22	51
Scores	$PC = 0.41, POD = 0.45, FAR = 0.82, CSI = 0.14, TSS = -0.15, HSS = -0.09$		

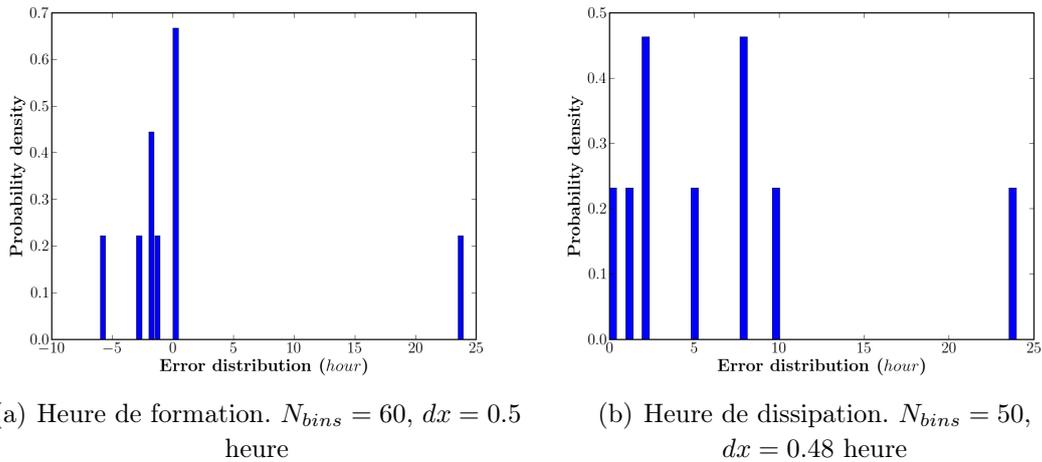


FIG. 4.8 : *Distribution des erreurs de prévisions pour l'heure de formation (a) et l'heure de dissipation (b) du brouillard. Cas démarré à 00 TU avec nudging. La densité de probabilité de l'intervalle $[a, b]$ est donnée par $\int_a^b f(x)dx = 1$ (Idem pour les figures suivantes)*

Ceci peut être lié à la prévision des conditions météorologiques de plus grande échelle. Sur l'heure de dissipation, le modèle est en retard sur toutes les situations de brouillard prévues. Ceci reste à étudier de manière détaillée afin d'améliorer la qualité de prévision du modèle. On peut toutefois suspecter la tendance de notre modèle à donner trop d'eau liquide. Pour le cas IC-12, la distribution est plus discrète que celle dans le cas IC-00, notamment pour l'heure de dissipation (Fig. 4.9). Dans l'ensemble, les scores du réseau de 12 TU sont moins bons que ceux du réseau de 00 TU. Ceci semble finalement assez compréhensible puisque le brouillard apparaît plutôt en fin de soirée où la nuit et donc à des périodes proches de 00 TU.

4.4 Prédiction du brouillard sans nudging

Dans cette section, on présente quelques résultats sans la prise en compte des prévisions à méso-échelle par l'utilisation de la technique de nudging.

4.4.1 Variables de surface

Pour le cas démarré à 00 TU. On constate que sans la prise en compte du nudging les résultats sont sensiblement dégradés, avec une augmentation du biais sur la température pouvant atteindre 1 °C (Fig. 4.10(a)). Cependant la forme de l'évolution temporelle de la température reste identique à celle du cas avec nudging. Le mélange vertical sur une épaisseur de l'ordre de quelques centaines de mètres conduit à influencer les paramètres météorologiques près du sol par la prise en compte du nudging (notamment par une valeur plus élevée du coefficient de nudging). Les différences "sans nudging" et "avec nudging" sont également marquées pour l'humidité relative à 2 mètres (Fig. 4.10(b)),

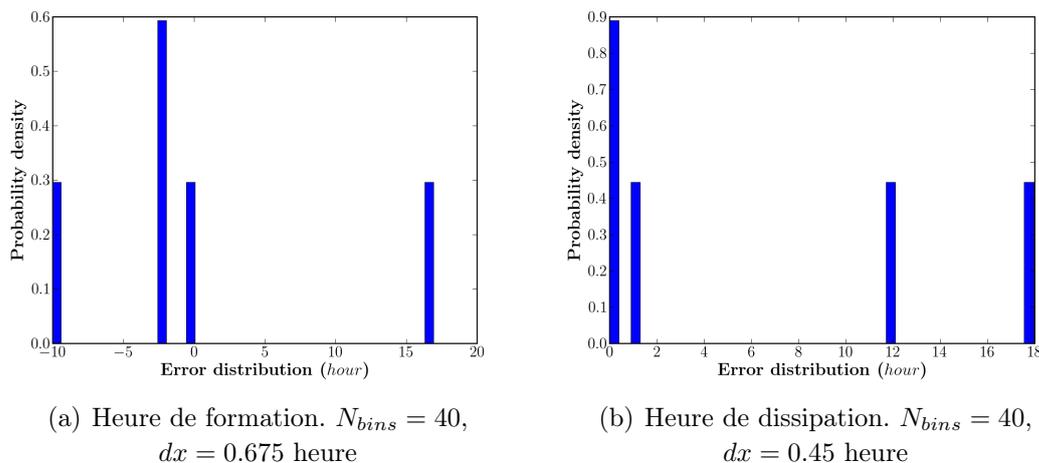


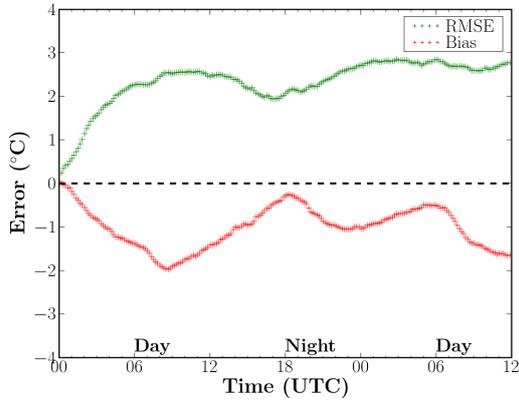
FIG. 4.9 : Distribution des erreurs de prévisions pour l'heure de formation (a) et l'heure de dissipation (b) du brouillard. Cas démarré à 12 TU avec nudging.

l'erreur quadratique moyenne est de l'ordre de 10 à 20 % au lieu de 8 à 15 %. Sur la dynamique (Fig. 4.10(c) et Fig. 4.10(d)), l'apport du nudging est clair. Au voisinage du sol, la prise en compte du nudging modifie la turbulence créée par le rappel au forçage de grande échelle.

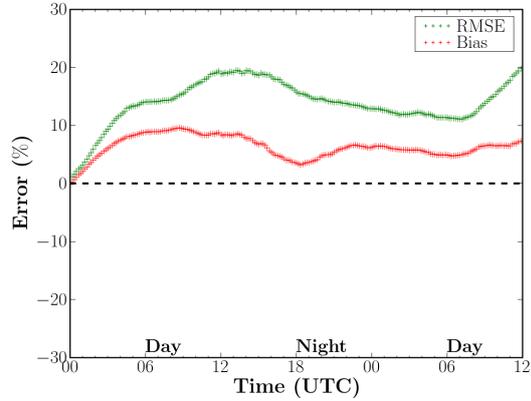
Le biais et l'erreur quadratique moyenne entre le modèle et les mesures pour le cas démarré à 12 TU, pour la température à 2 m, l'humidité relative à 2 m, la vitesse du vent à 10 m et l'énergie cinétique turbulente à 10 m sont présentés dans la figure 4.11. Ici encore, l'absence du nudging dégrade les résultats. L'erreur quadratique moyenne de la température est de l'ordre de 2 à 3 °C, et l'humidité relative de 10 à 20 %. Pour la dynamique, l'erreur quadratique moyenne de la vitesse du vent est de l'ordre de $2.3 \text{ m}\cdot\text{s}^{-1}$, et sur l'énergie cinétique turbulente de $2.1 \text{ m}^2\cdot\text{s}^{-2}$. Ceci permet de mettre en évidence le rôle positif de nudging, qui illustre aussi la qualité de la prévision MM5.

4.4.2 Prédiction du brouillard

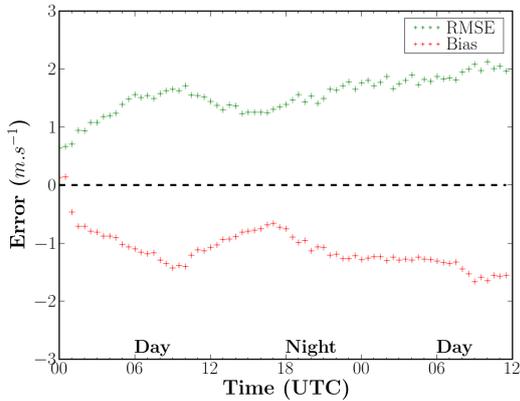
Les tableaux de contingences sur l'occurrence du brouillard montrent que la qualité de prévision en l'absence du nudging est bien moins bonne que dans le cas avec le nudging. Il s'en suit une dégradation du taux de bonnes prévisions par rapport au cas avec nudging de l'ordre de 30%. La probabilité de détection passe de 0.91 à 0.55 par rapport au cas avec nudging (cas IC-00). Les valeurs de *TSS* pour les deux cas, IC-00 et IC-12, sont négatives. Ce mauvais résultat de prévision est à relier à une valeur relativement importante de la proportion de fausse alarme. Sans la prise en compte du nudging on peut sous-estimer le mélange vertical et ne pas reproduire les assèchements par advection. Ceci conduit à augmenter le nombre de situations de brouillard (*FAR* augmente), qui peut être lié au biais négatif sur la température à 2 m (et donc au biais positif sur l'humidité relative à 2 m).



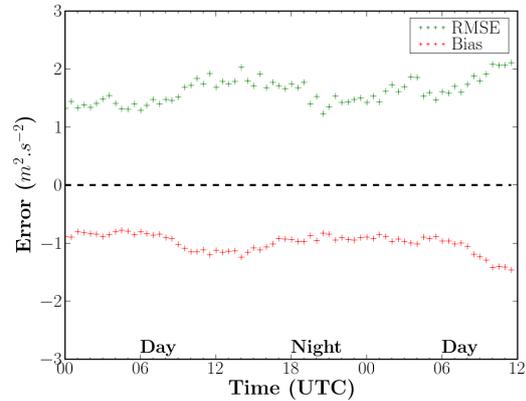
(a) Température à 2 m



(b) Humidité relative à 2 m



(c) Vitesse du vent à 10 m



(d) Énergie cinétique turbulente à 10 m

FIG. 4.10 : Prédiction sans nudging, cas démarré à 00 TU (IC-00). Biais (rouge) et erreur quadratique moyenne (verte) de prédiction à 36 heures par rapport aux mesures pour les paramètres météorologiques au voisinage du sol.

TAB. 4.5 : Table de contingence à 2 dimensions. Cas démarré à 00 TU sans nudging pour 36 h d'échéance.

	Prédiction <i>oui</i>	Prédiction <i>non</i>	Total
Observation <i>oui</i>	6 (hit)	5 (missed)	11
Observation <i>non</i>	25 (false)	15 (correct rejection)	40
Total	31	20	51
Scores	$PC = 0.41, POD = 0.55, FAR = 0.81, CSI = 0.17, TSS = -0.08, HSS = -0.05$		

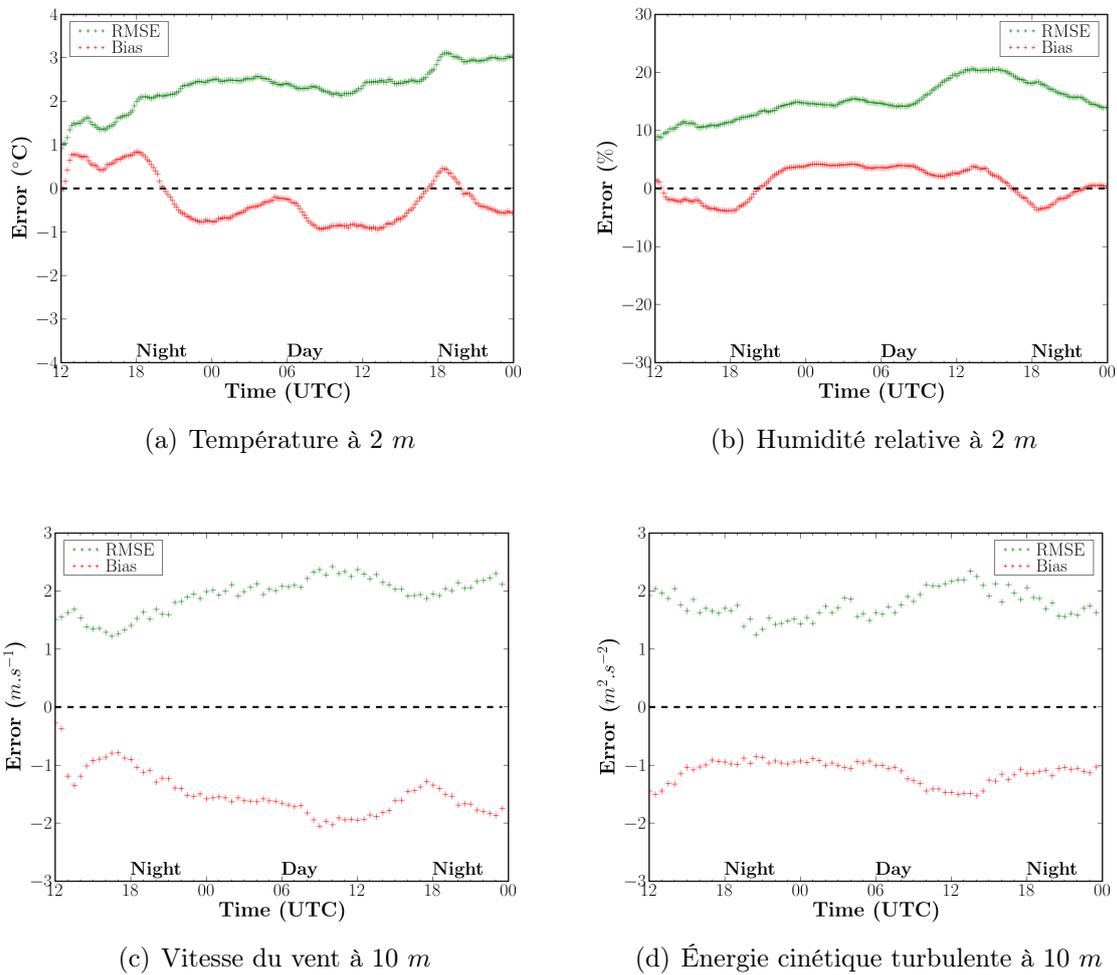


FIG. 4.11 : Prédiction sans nudging, cas démarré à 12 TU (IC-12). Biais (rouge) et erreur quadratique moyenne (verte) de prédiction à 36 heures par rapport aux mesures pour les paramètres météorologiques au voisinage du sol.

TAB. 4.6 : Table de contingence à 2 dimensions. Cas démarré à 12 TU sans nudging pour 36 h d'échéance.

	Prédiction <i>oui</i>	Prédiction <i>non</i>	Total
Observation <i>oui</i>	5 (hit)	6 (missed)	11
Observation <i>non</i>	23 (false)	17 (correct rejection)	40
Total	28	23	51
Scores	$PC = 0.43, POD = 0.45, FAR = 0.82, CSI = 0.15, TSS = -0.12, HSS = -0.08$		

TAB. 4.7 : Table de contingence à 2 dimensions. Cas démarré à 00 TU avec le modèle de *Louis* avec nudging pour 36 h d'échéance.

	Prévision <i>oui</i>	Prévision <i>non</i>	Total
Observation <i>oui</i>	6 (hit)	5 (missed)	11
Observation <i>non</i>	25 (false)	15 (correct rejection)	40
Total	31	20	51
Scores	$PC = 0.41, POD = 0.55, FAR = 0.81, CSI = 0.17, TSS = -0.08, HSS = -0.05$		

4.5 Étude statistique avec le modèle de *Louis*

Comme nous l'avons déjà mentionné lors de l'analyse de nos résultats dans l'étude 1D sur la POI-13, l'épaississement de la couche de mélange est assez sensible au choix de la fermeture turbulente. Cette étude statistique a été reconduite pour le modèle diagnostique de *Louis* en démarrant à 00 TU.

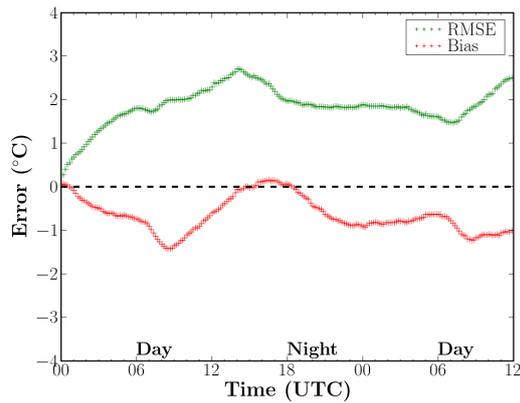
On constate que le modèle de *Louis* donne des écarts comparables au modèle $k - \varepsilon$ (Fig. 4.12). Les erreurs quadratiques moyennes de la température et de l'humidité relative sont de l'ordre de 2 °C et 11 % respectivement. Elles sont du même ordre de grandeur que les erreurs dans le modèle de prévision habituel. En ce qui concerne la vitesse du vent (Fig. 4.12(c)), le modèle de *Louis* est plus précis que le modèle $k - \varepsilon$, puisqu'il conduit à une erreur quadratique moyenne inférieure à 1 $m.s^{-1}$ au lieu de 1 à 2 $m.s^{-1}$ pour le modèle $k - \varepsilon$. En revanche, le modèle de *Louis* sous-estime la vitesse alors que le modèle $k - \varepsilon$ la surestime (excepté au moment du coucher du soleil). De plus, le modèle de *Louis* sous-estime l'énergie cinétique turbulente pendant toute la simulation, qui peut conduire à diminuer le mélange sur la verticale avec un impact sur la formation du brouillard.

Les tableaux de contingence (Tab. 4.7) permettent de le mettre en évidence. Le pourcentage de bonne prévision est de moins de 50% avec une valeur négative de TSS . La surestimation de l'humidité relative peut conduire à une surestimation de l'eau liquide. Cela conduit à une valeur élevée du FAR .

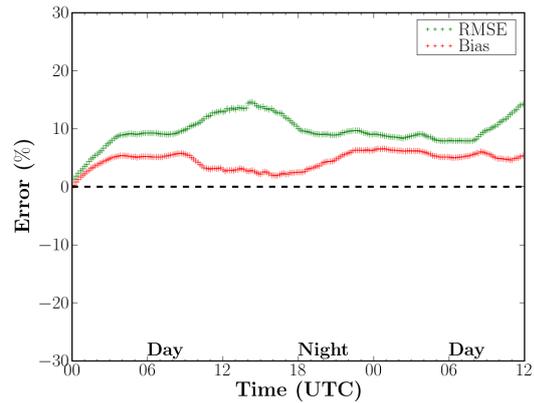
4.6 Discussion et conclusion

Une étude statistique sur l'ensemble de la période de ParisFog (novembre 2006-mars 2007) en mode prévision a été présentée. Dans ces simulations, le modèle 1D est couplé avec les données de prévisions issues du modèle méso-échelle MM5 sous une condition de nudging.

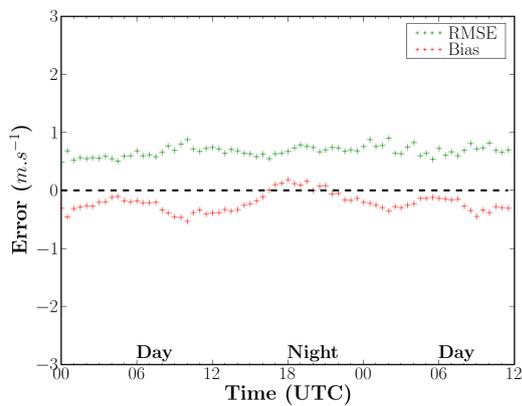
Sur les paramètres classiques, les scores obtenus en mode couple se comparent globalement assez favorablement aux scores des modèles opérationnels de prévision notamment en démarrant sur le réseau de 00 TU. En ce qui concerne la prévision du brouillard, les scores issus des tableaux de contingences ont permis de montrer que notre modèle a une



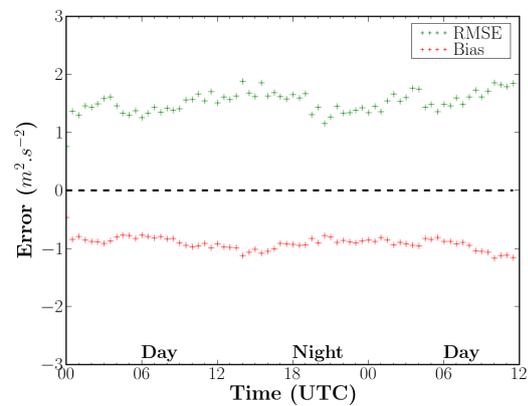
(a) Température à 2 m



(b) Humidité relative à 2 m



(c) Vitesse du vent à 10 m



(d) Énergie cinétique turbulente à 10 m

FIG. 4.12 : Prédiction avec le modèle de diagnostique de Louis avec nudging, cas démarré à 00 TU (IC-00). Biais (rouge) et erreur quadratique moyenne (verte) de prédiction à 36 heures par rapport aux mesures pour les paramètres météorologiques au voisinage du sol.

assez bonne capacité à prévoir le brouillard à courte terme (au moins 36 h à l'avance) avec une valeur de *true skill statistic TSS* supérieur à 0.5. Le pourcentage de bonne prévision est du même ordre de grandeur que celui trouvé dans le modèle COBEL-ISBA (Bergot et al., 2005). En revanche, l'heure de dissipation prévue est souvent retardée par rapport à celle observée. Une étude plus détaillée sur ce point est à envisager.

Sans la prise en compte du nudging, la dégradation des résultats est significative. L'erreur quadratique moyenne augmente environ d'un facteur 1.5 pour les paramètres météorologiques. Cela dégrade la précision du modèle. De plus, la valeur de *TSS* devient négative, ce qui montre une mauvaise prévision. Ceci permet de mettre en évidence le rôle positif de nudging.

Le modèle de *Louis* en mode prévision est moins performant que le modèle $k - \varepsilon$. La bonne valeur de l'erreur quadratique moyenne sur la vitesse du vent ne compense pas les erreurs observées pour la température, l'humidité et l'énergie cinétique turbulente. Les erreurs conduisent à une mauvaise prévision du brouillard ($TSS < 0$). C'est pourquoi pour la suite nous choisirons le modèle de turbulence $k - \varepsilon$.

La prévision du brouillard de rayonnement sur un site avec des fortes hétérogénéités spatiales est une tâche difficile. L'utilisation de technique de nudging et d'un forçage méso-échelle a permis de s'affranchir de certaines difficultés et d'obtenir finalement une prévision relativement bonne. Il reste que ce travail demanderait à être approfondie en réalisant une étude plus fine des cas d'échec de la prévision sur cette période. La qualité des données mesurées sur le site du SIRTÀ durant ParisFog devrait le permettre. Par ailleurs un couplage avec les modèles opérationnels de Météo-France pourrait améliorer les scores. Ceci reste à déterminer.

