

Toulouse le 08 Octobre 2014

Contrôle objet du modèle AROME avec le logiciel MODE développé au NCAR

Travail réalisé par RAZAGUI Abdelhak
sous la direction de Joël STEIN
Période du 13 septembre au 11 Octobre 2014
Compas 2014

Introduction

Le contrôle des modèles de prévision numérique du temps par la méthode des objets est relativement récent. Il s'applique essentiellement pour les modèles à haute résolution comme AROME, WRF etc.. Il permet le suivi spatio-temporel des cellules nuageuses convectives. La méthode a été testée sur l'épisode pluvieux du 16 au 19 septembre 2014 qui a touché le sud de la France où de fortes quantités de pluies ont été enregistrées. Les données d'observation utilisées sont les lames d'eau ANTILOPE (1,2Km) pour contrôler la version opérationnelle de la prévision AROME à 2,5 KM. L'outil de contrôle utilisé dans ce travail est MODE qui fait partie de l'outil d'évaluation MET développé par le NCAR. Les seuils d'intensité de précipitations prévues et observés choisis pour l'identification des différents objets sont de 5 mm/heure. Les objets observés ainsi identifiés dont la distance entre les centres de gravité est inférieure à un certain seuil sont fusionnés en un seul objet. On applique cette étape de fusion pour les objets prévus de la même façon. On peut alors tenter d'associer objets prévus et observés et calculer différents attributs quantitatifs pour cette comparaison. Le réseau de prévision de 18UTC a été choisi de manière à avoir un champ initial au moment même où la convection s'est déjà déclenchée et suivre la qualité de la prévision pendant les 30 heures de prévision.

Principe de la méthode de contrôle

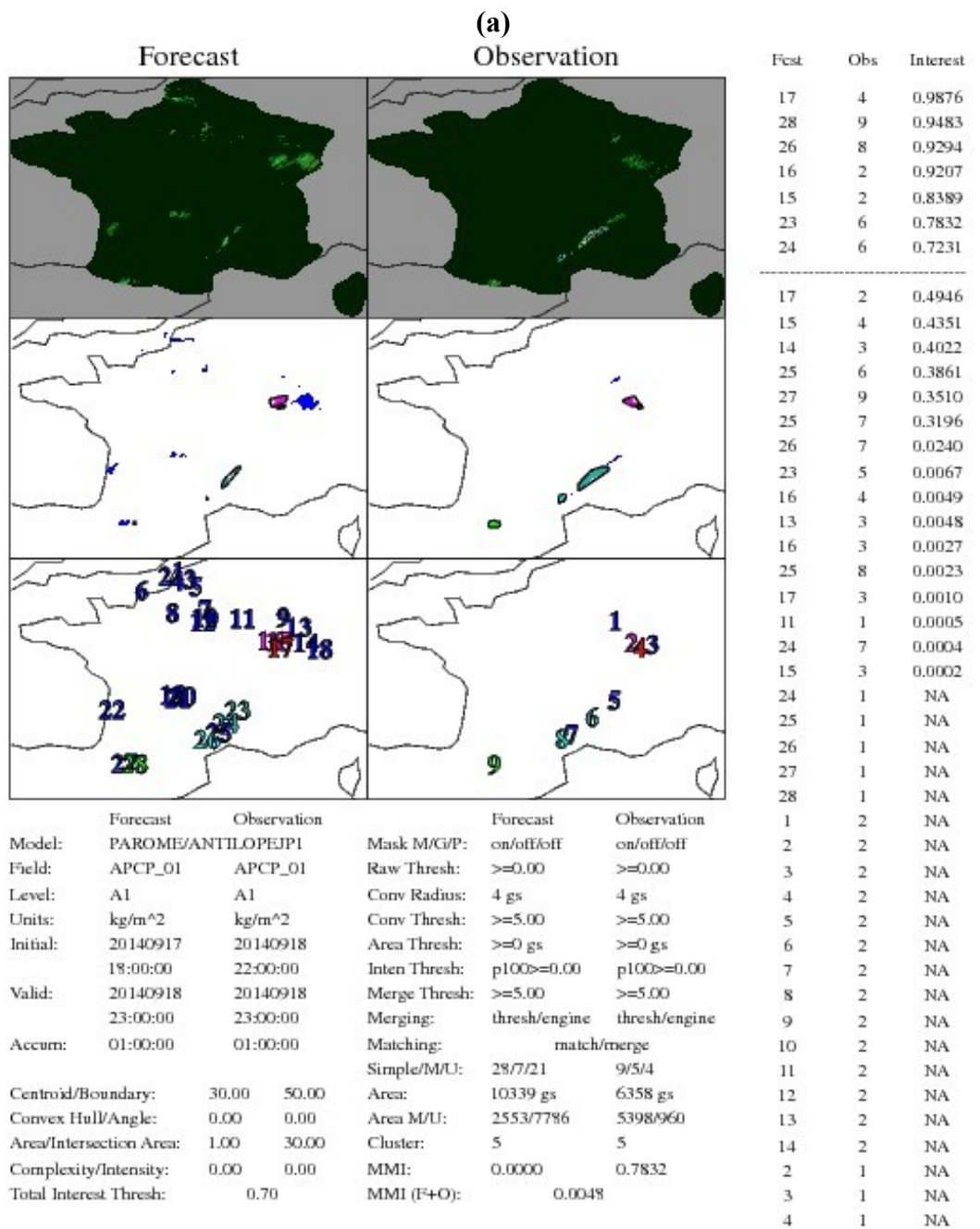
L'algorithme de contrôle est basé sur un ensemble de critères qui sont dans l'ordre

- 1) l'identification des objets dans les données sources prévues et observées par convolution avec une fonction porte sur un disque de rayon R. Cette convolution réalise une opération de moyenne sur le disque car le paramètre de réglage est pris égal à $H = 2 \cdot \pi R^2$. Une fois la convolution obtenue, le domaine où le champ est supérieur au seuil retenu définit un masque. Enfin, le champ initial non filtré est restauré sur chacun des objets définis par le masque
- 2) fusion pour certains objets en un seul objet pour objets observés et prévus séparément
- 3) correspondance des objets prévus et observés.
- 4) calcul d'un paramètre appelé intérêt total qui est la moyenne pondérée de plusieurs termes qualifiant quantitativement la comparaison comme par exemple la distance entre les 2 centres de gravité, rapport des surfaces des deux objets etc... Cette valeur calculée est comparée à une valeur seuil définie auparavant dans un fichier de configuration.

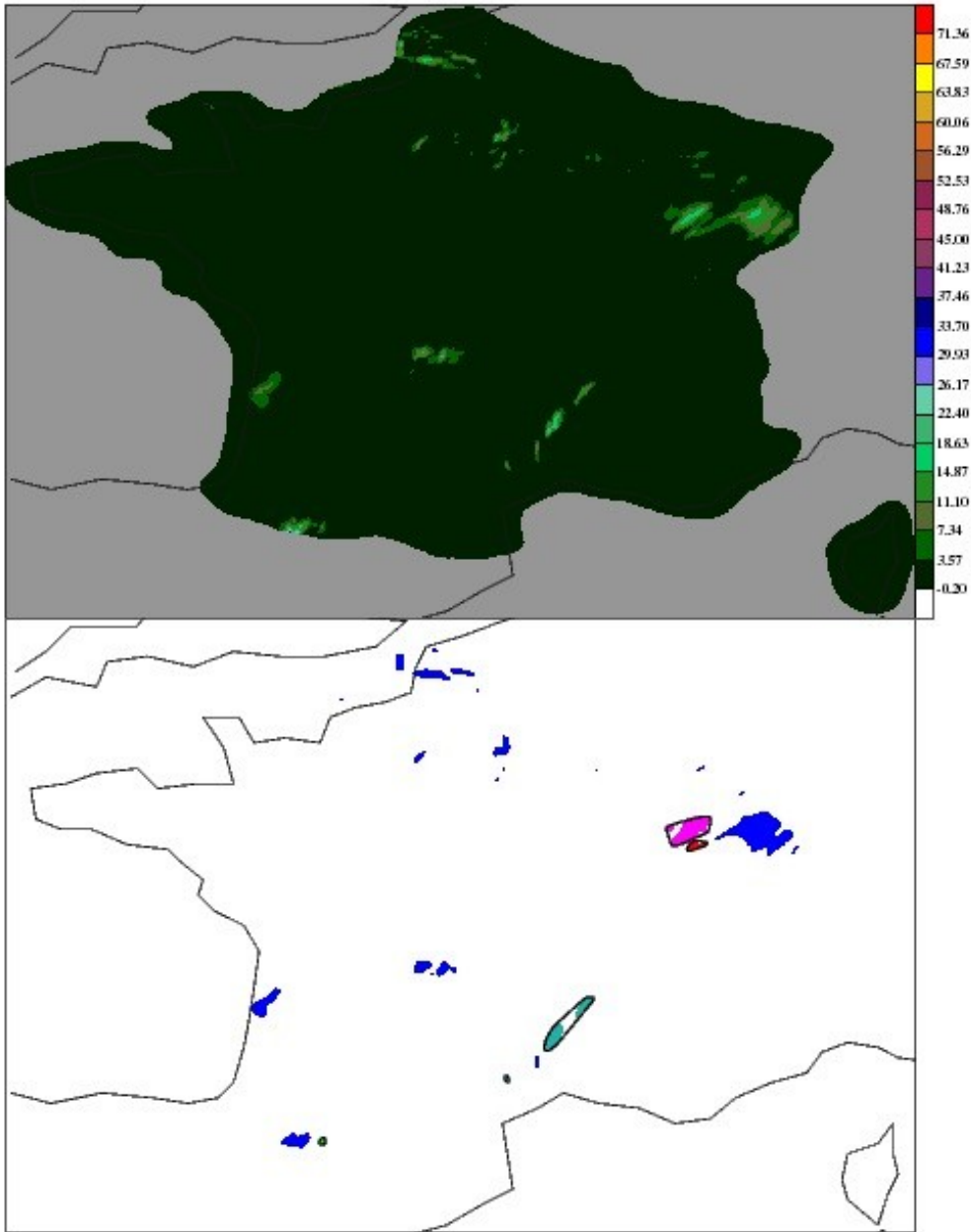
Un fichier de configuration est donné en input à l'outil de contrôle MODE dans lequel tous les attributs (la distance entre les centre de gravité des objets, la distance entre les enveloppe des objets, le recouvrement des objets entre eux et en fin l'orientation des objets l'un par rapport à l'autre) et intérêt total, relative à cette procédure sont fixés selon la nature et l'échelle spatiale du phénomène étudié.

Discussion des résultats et étude de cas

On présente dans les 5 figures suivantes les sorties standards du logiciel MODE qui permettent de comparer les objets prévus et observés avec les paramètres fixés dans le fichier de configuration. Il s'agit de champs de pluies cumulées sur 1 heure entre le 18/9/2014 22 et 23 UTC. Cela correspond à des échéances de prévisions 28 et 29 h respectivement, la simulation AROME ayant démarré le 17/09/2014 à 18 UTC. Le seuil de pluie choisi est de 5 mm/h

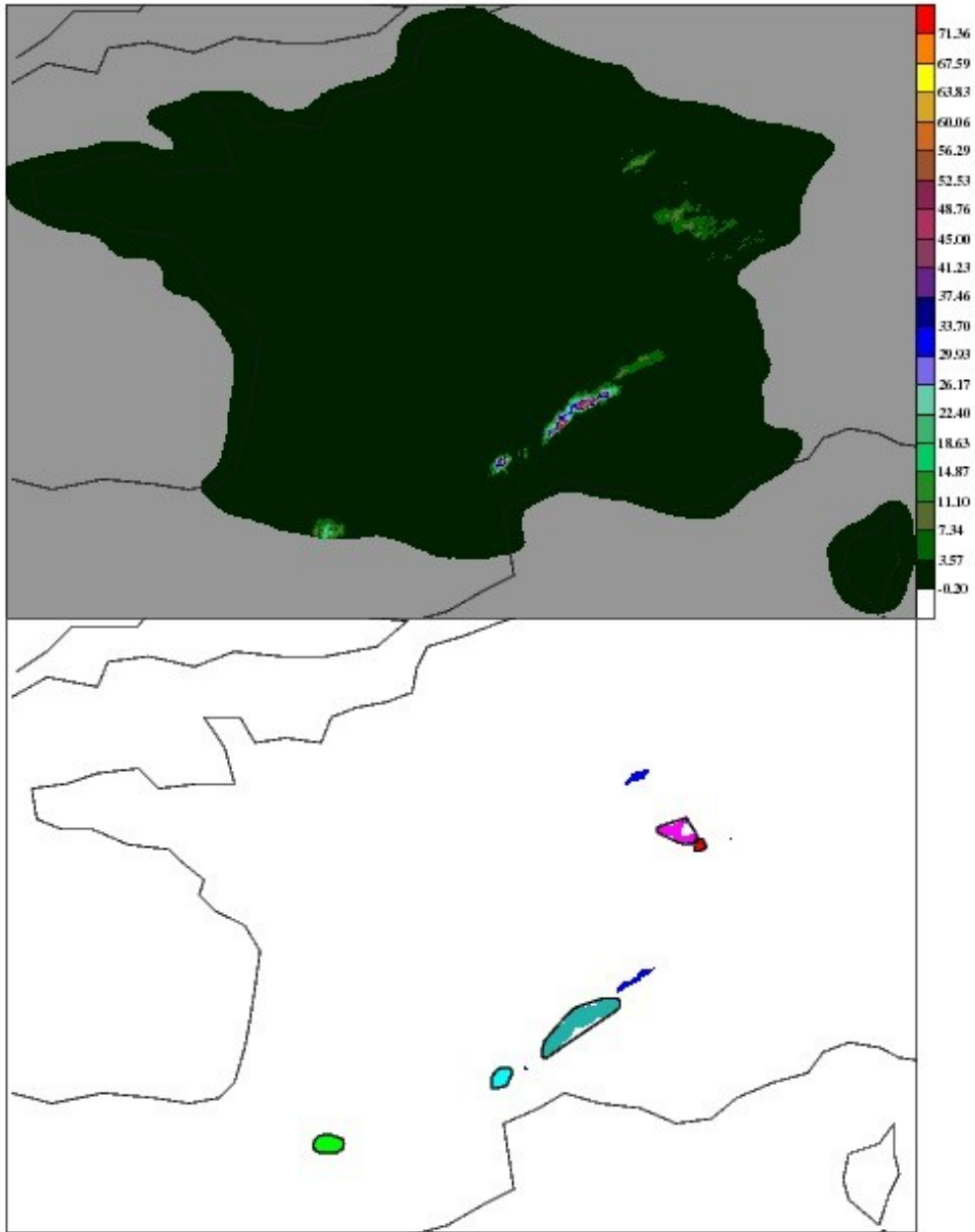


(b)
Forecast



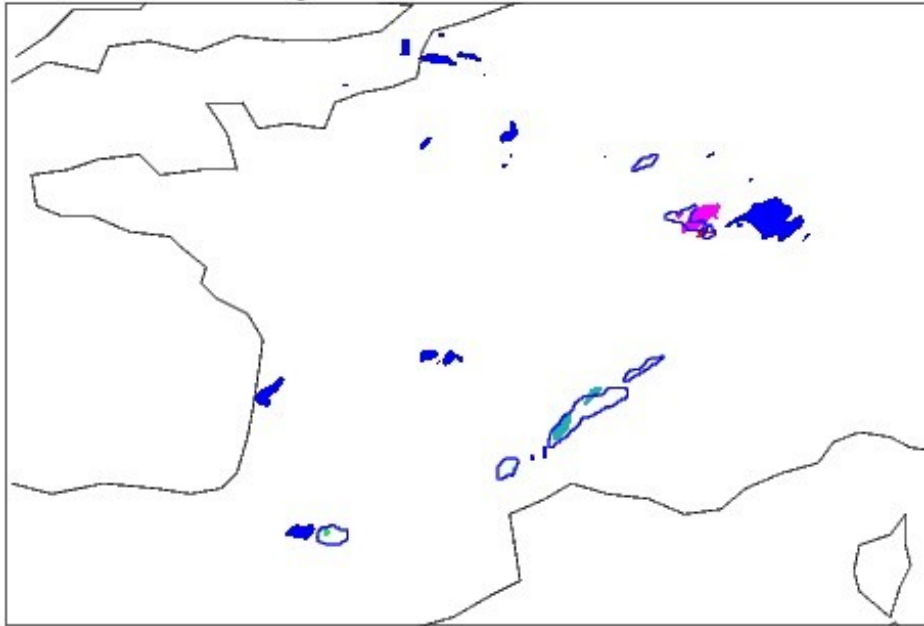
(c)

Observation

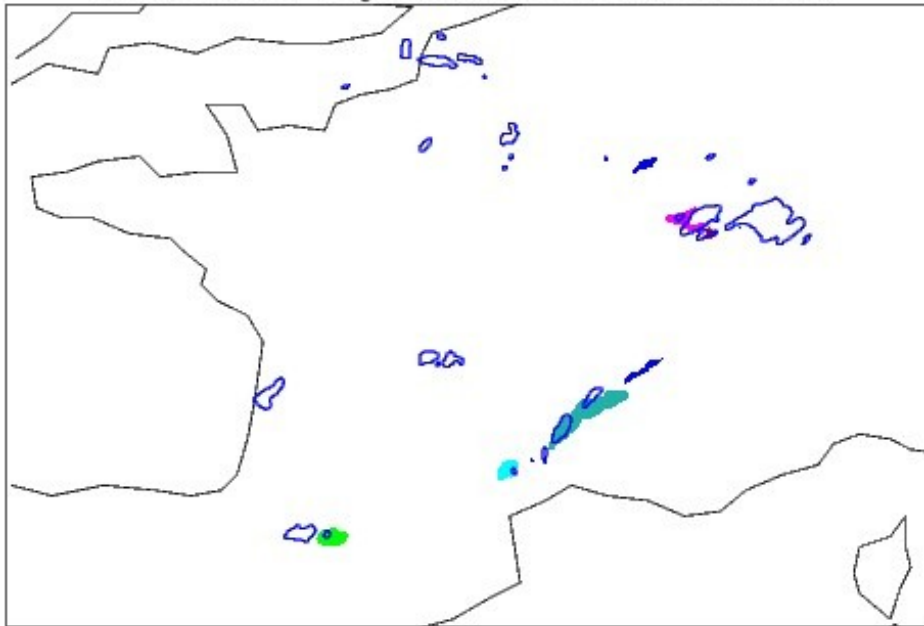


(d)

Forecast Objects with Observation Outlines



Observation Objects with Forecast Outlines



(e)

Cluster Object Information

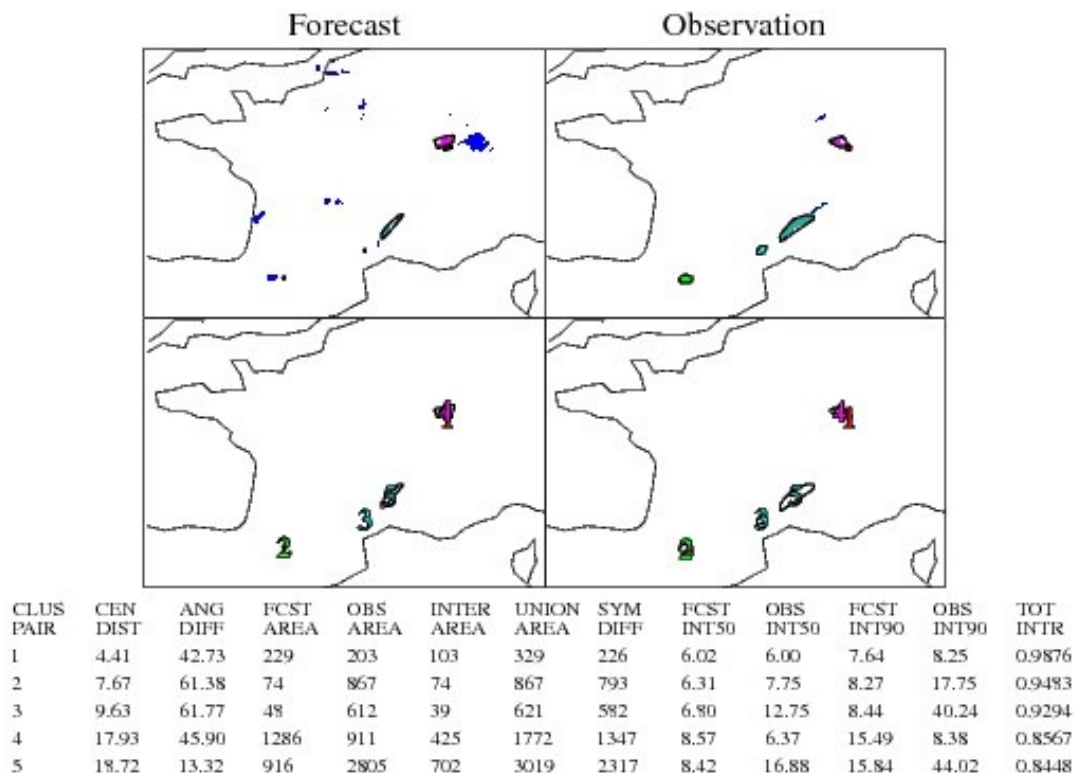
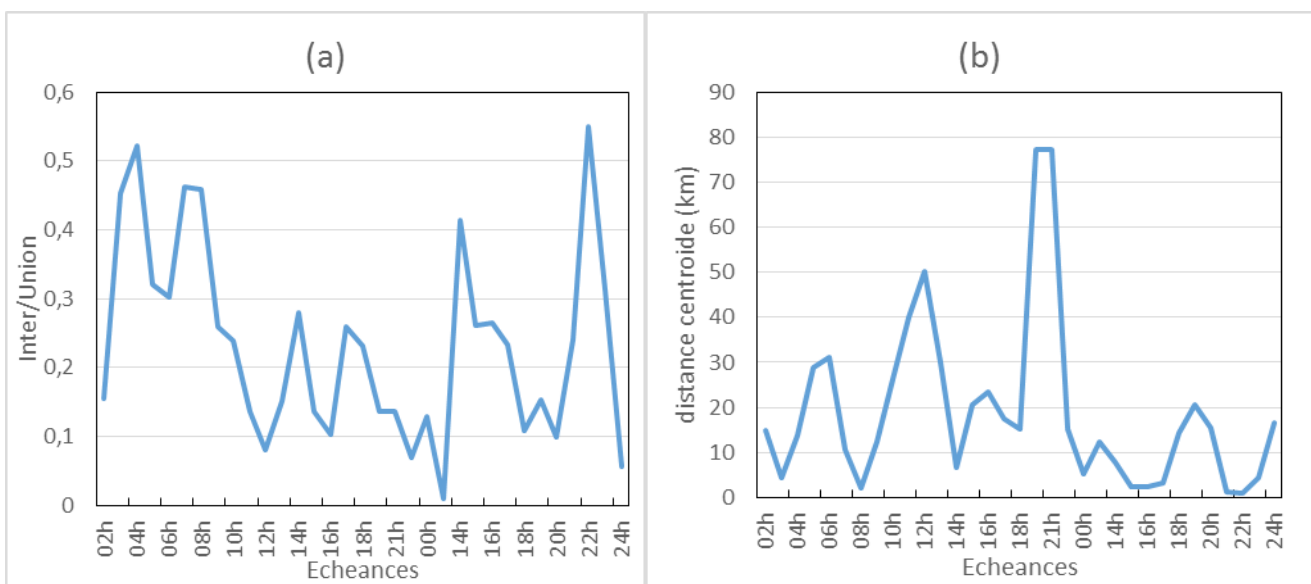


Figure 1. Objets prévus et observés sur les champs de pluies cumulées le 18/09/2014 entre 22 et 23h00TC. Les prévisions sont issues du réseau de 18 UTC d'AROME et les observations de la lame d'eau ANTILOPE.

On remarque bien (figure 1. (a)) que le modèle AROME développe beaucoup plus d'objets qu'il n'en existe en réalité. Cette remarque correspond au fait que le modèle AROME est biaisé positivement pour le champ de précipitation. On note que la localisation des objets observés est globalement respectée avec des maxima sur les Pyrénées, un près des Vosges et les Cévennes. Par contre, il existe des fausses alarmes sur la moitié Ouest de la France. Le tableau à droite de la figure 1a donne un certain nombre de couples d'objets prévus et observés identifiables dont l'intérêt total est significatif (>0.7). La figure 1. (b) et (c) montrent respectivement les champs bruts prévus et observés en haut de la page et les différents objets identifiés en bas de la page avec une couleur par objet. Le haut de la figure 1 (d) montre la projection spatiale des objets observés sur les objets prévus et inversement le bas de la figure 1 (d). Cela montre la surface d'intersection entre les différentes séries d'objets prévus et observés. On remarque que plusieurs objets prévus peuvent avoir une intersection avec le même objet observé et inversement. La figure 1 (e) montre les clusters fusionnés et associés retenus avec un intérêt total significatif. Ainsi, on remarque qu'au lieu des 7 objets significatifs retenus sur le tableau de la figure 1 a, il n'en reste que 5 après le regroupement des objets observés (6 et 8) en un seul objet ainsi que les objets prévus (24 et 26). Il faut bien signaler que le regroupement des objets en seul objet déplace le centre de gravité et augmente la taille de l'objet final, change l'angle des axes de symétries des objets, augmente ou diminue la distance entre centres de gravité des objets prévus et observés (centroïdes), augmente ou diminue la surface d'intersection, et donc affecte la différence symétrique entre les objets observés et prévus et la valeur de l'intérêt total de l'association finale.

On liste dans le paragraphe suivant l'intérêt des différents attributs de la comparaison entre objets prévus et observés. La distance entre les centres de gravité montre le déplacement spatial des deux objets l'un par rapport à l'autre. L'angle entre les axes de symétrie montre comment ces objets sont alignés. Les surfaces calculées des objets prévus et observés et par suite leur rapport donne une idée objective sur la sous ou surestimation de l'étendue spatiale du phénomène. La surface d'intersection mesure les bonnes détections et la différence symétrique entre les deux objets donne les non détections et les fausses alarmes. Les paramètres statistiques percentiles 50 et 90 donnent les intensités de précipitations supérieures au seuil choisi avec 50% des points de grille de l'objet qui ont une intensité supérieures au Q50 et 10 % au Q90. Le rapport des percentiles prévus et observés donne une bonne idée de la qualité de l'accord sur l'intensité du phénomène. La qualité globale de la prévision du modèle AROME est pour ce cas donné reflétée par le paramètre intérêt total variant entre 0.99 à 0.85.

Etude du contrôle en fonction des échéances de prévision



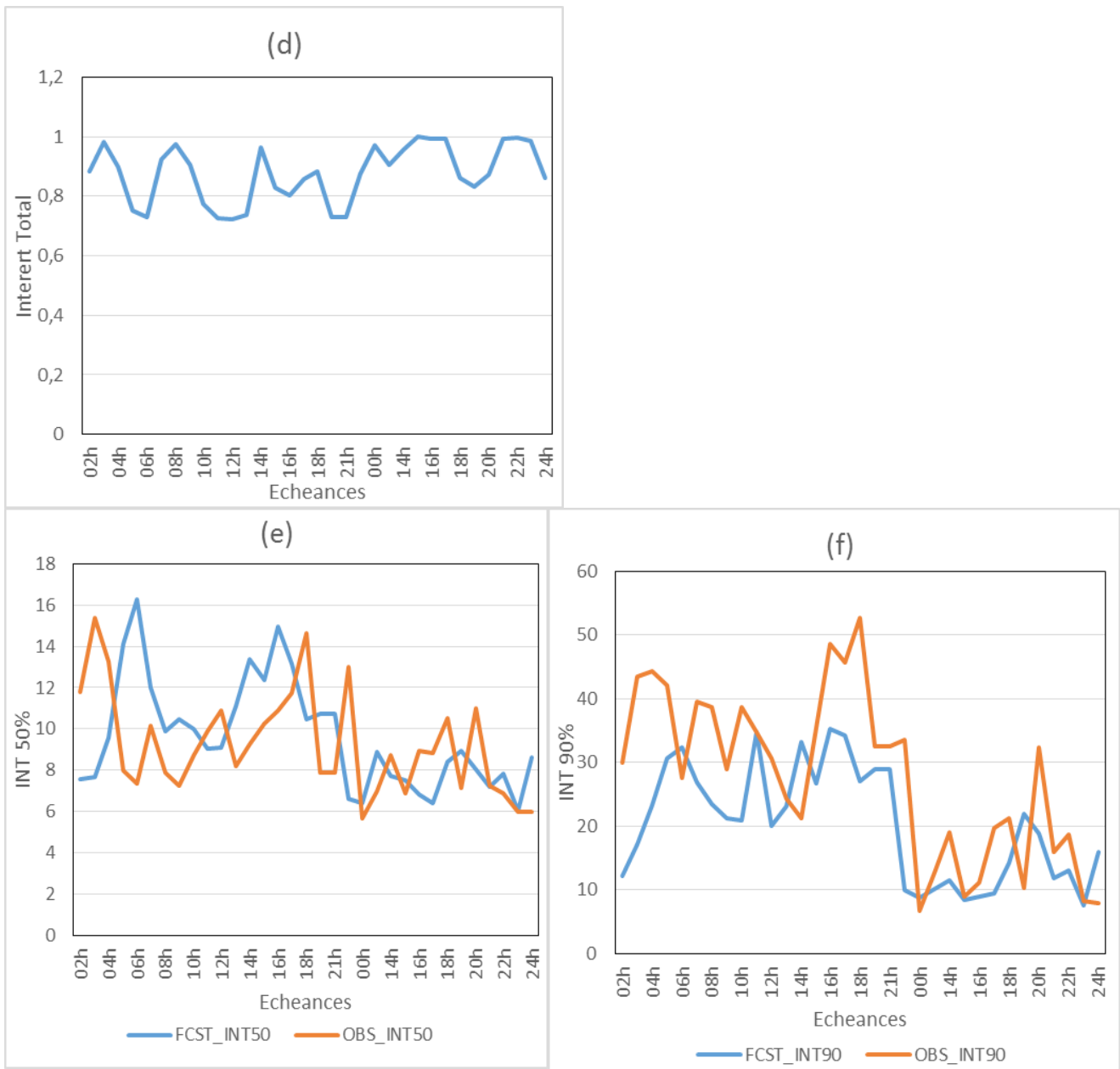


Figure 2. (a) Intersection sur l'union des surfaces des deux objets, (b) Distances entre les centres de gravité, (c) Intérêt total, (d) Intérêt total, percentile 50% et (f) percentile 90% ; L'axe des abscisses correspond à l'échéance en h de la simulation AROME de 1 à 30.

Nous avons suivi l'évolution des précipitations, heure par heure, durant tout l'épisode pluvieux en utilisant la prévision du 17 octobre pour le réseau de 18h sur les trente échéances de prévision du modèle AROME. Le rapport de l'intersection des surfaces des deux objets à leurs unions (Figure 2a) donne la performance du modèle comparable à CSI ($CSI = a / (a + b + c)$) pour chaque objet retenu avec les fortes valeurs de l'ordre de 0.55 sont enregistrées la nuit et la fin de nuit. Ce rapport pour chaque heure correspond à la paire d'objets prévu/observé ayant le plus fort intérêt mais possédant une surface suffisante subjectivement. Ces valeurs sont proches de 0.2 à 0.3 en moyenne avec une forte variabilité temporelle et cela montre que les valeurs de scores classiques sont modestes pour ce cas extrême.

La distance entre centres de gravité est oscillée généralement entre 5 et 20 km mais par 3 fois, on observe de plus forts décalages allant jusqu'à 75 km à 21UTC. Ces valeurs montrent bien le problème de double peine des modèles à haute résolution tel que AROME puisque le phénomène est bien prévu mais pas là où il est observé. Cela justifie l'intérêt des calculs de scores avec tolérance comme c'est fait dans les procédures standards de Météo-France.

Les valeurs moyennes de Q50 observées et prévues sont du même ordre de grandeur soit 9 mm alors que les Q90 sont respectivement de l'ordre de 30 mm au début de l'évènement pour descendre à des intensités maximales plus faibles dans les 10 dernières heures simulées.

La valeur moyenne de l'intérêt total sur tout l'épisode pluvieux (Figure 2f) est de l'ordre de 0.87, soit relativement supérieur au seuil critique qui est de 0.7. Cette valeur résume l'ensemble des statistiques sur la comparaison des objets prévus et observés. Il reste très élevé car il n'est calculé que lorsque l'association objet prévu et observé est possible donc quand le signal donné par la prévision est pertinent. C'est donc un bon indicateur de performance du modèle seulement dans le cas où le modèle dépasse une certaine qualité de prévision en effet on oublie les cas où le modèle est en fausse alarme globale ou en non détection globale car dans ce cas on n'a qu'un seul objet et on ne peut donc pas l'associer.

Conclusion

Cette étude réalisée sur un épisode bien déterminé est un exemple de la problématique du contrôle d'un modèle numérique à haute résolution avec la méthode des objets. Le problème de sa généralisation à une période d'étude demeure néanmoins ouvert dans le cas du modèle AROME. L'analyse faite ici ne tient pas compte des fausses alarmes faite par le modèle AROME comme le montre la figure 1 (a) où l'outil MODE identifie 28 objets sur la prévision alors qu'il n'en existe que 6 dans l'observation. L'analyse objet n'est en fait réalisée uniquement sur les objets associés avec un intérêt total supérieur ou égal à 0.7. Il s'agit donc des cas favorables.

Même si les valeurs de l'intérêt total sont élevées, la distance des centres de gravité entre objets prévu et observé atteint en moyenne de 7 points de grille. Ceci montre que l'évènement est bien prévu mais pas là où il faut. Après analyse des autres réseaux de prévisions d'AROME, nous pouvons conclure que le réseau de 18h reste celui qui donne les résultats les plus meilleurs sur ce cas d'étude. Nous avons testé au cours de cette étude l'impact des seuils sur nos résultats. Nous avons constaté que la diminution des seuils augmente la taille des objets mais dégrade l'association entre les objets prévus et observés comme le montre la variation de l'intérêt total. Ainsi, la fusion d'objets entre eux peut dans certains cas dégrader la valeur de l'intérêt total. Afin de porter un bon jugement objectif sur la performance du modèle AROME, un contrôle sur une longue période s'avère nécessaire pour rendre robuste cette analyse objet. Pour cela, il faudra trouver la bonne représentation ou la bonne moyenne de ces informations fournies par le logiciel MODE.