

Prévision multi-échelles des précipitations

Aniss ZEBIRI^{1,4}, Dominique BÉRÉZIAT¹, Etienne HUOT², Isabelle HERLIN³, Emmanuel BUISSON⁴

¹Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, LIP6, F-75005 Paris, France

²Université de Versailles Saint-Quentin-en-Yvelines, LATMOS, 11 boulevard d'Alembert, 78280 Guyancourt, France

³Inria

⁴Weather Measures, 10 allée Evariste Galois – Pascalis, 63100 Clermont-Ferrand, France

Aniss.Zebiri@lip6.fr, Dominique.Bereziat@lip6.fr, Etienne.Huot@latmos.ipsl.fr,
Isabelle.Herlin@inria.fr, Emmanuel.Buisson@weather-measures.com

Résumé – Cet article propose une méthode de prévision à court terme des précipitations, basée sur deux types d'images radar avec des résolutions spatiales complémentaires. Cette méthode est implémentée grâce à l'élaboration d'une version multi-échelles d'un algorithme opérationnel. Différentes variantes de la méthode permettent également d'augmenter l'horizon de prévision. Une évaluation quantitative permet de valider l'approche est de démontrer sont potentiel.

Abstract – The paper describes a rain nowcasting method, based on multi-resolution radar images. A multi-scale version of an operational algorithm is given. Alternatives to this method are also proposed to increase the temporal horizon of the forecasts. A quantitative evaluation validates the approach.

1 Introduction

La prévision des précipitations est un enjeu sociétal majeur. On peut, à titre d'illustration, citer l'optimisation des pratiques agricoles ou la prévention des inondations. Cette prévision s'appuie à la fois sur les modèles météorologiques et sur les observations issues de divers capteurs, dont des imageurs. Les météorologues s'accordent sur le fait que les images fournies par les radars sont les plus appropriées pour quantifier les précipitations sur de larges domaines, à des échelles spatiales et temporelles fines.

Il existe différentes approches pour intégrer les images dans le processus de prévision de la pluie, depuis l'utilisation de modèles météorologiques jusqu'à la simple extrapolation calculée à partir de quelques images. C'est le choix de l'horizon temporel et de l'échelle spatiale de la prévision à calculer qui déterminent l'approche la mieux adaptée.

Les modèles météorologiques sont utilisés quotidiennement par les centres de prévision opérationnelle. Ces modèles reposent sur les équations mathématiques de l'évolution de l'atmosphère et leurs prévisions sont calculées par intégration temporelle de ces équations à partir d'une condition initiale connue. Les modèles de prévision opérationnelle s'appuient de plus sur l'assimilation de données, afin de prendre en compte de manière optimale les observations des différentes variables d'état. Le calcul des prévisions nécessite alors des temps de calcul de plusieurs heures, interdisant en conséquence des horizons temporels à court terme.

Une approche basée sur l'extrapolation d'images se révèle

souvent assez efficace si l'horizon de prévision est de l'ordre de quelques minutes. On distingue ainsi deux types de méthodes d'extrapolation selon la façon d'estimer le déplacement inter-image : à partir du suivi de structures ou à partir d'une estimation dense du mouvement. En 2019, les méthodes les plus utilisées de manière opérationnelle restent celles basées sur le suivi de cellules précipitantes [2, 4]. Les méthodes denses reposent sur le calcul du flot optique [1] et les prévisions sont extrapolées à partir de la dernière image disponible en utilisant l'information de mouvement.

Depuis une douzaine d'années, une alternative a fait son apparition : l'approche par *assimilation d'images* [3]. Dans le contexte de l'estimation du mouvement, on peut la considérer comme une extension du flot optique et plus généralement comme une application des techniques d'assimilation de données aux images. Cette "nouvelle" approche permet de déterminer une solution optimale en s'appuyant sur différentes types d'information. Elle permet la prise en compte d'un modèle représentant la dynamique du vecteur d'état et offre la possibilité de combiner dans un même formalisme le suivi de structures et le flot optique [6, 7, 11].

Notre travail s'appuie sur cette approche d'assimilation d'images. Nous nous intéressons à l'amélioration de la prévision de la pluie à court terme, en utilisant les images acquises par deux réseaux de capteurs radars présentant des résolutions différentes. Un des réseaux est opéré par la startup *Weather Measures*¹. Il fournit des images *locales*, à haute résolution spa-

1. <https://www.weather-measures.fr/>

tiale (200 m), sur un domaine limité à la région de Clermont-Ferrand. L'autre réseau est national et opéré par Météo France. Il couvre le territoire métropolitain et fournit des images *régionales*, avec une résolution plus grossière (1 km). Les deux réseaux délivrent simultanément, et toutes les 5 minutes, des acquisitions.

La section 2 présente la méthode utilisée pour la prévision de la pluie en s'appuyant sur les deux résolutions d'images, locale et régionale. La section 3 propose deux variantes à l'algorithme afin d'augmenter l'horizon temporel de prévision, tandis que la section 4 analyse leurs résultats. La conclusion est donnée section 5, accompagnée de quelques perspectives scientifiques.

2 Prévision des précipitations

Notre méthode d'estimation à court terme des précipitations repose sur deux étapes successives : 1- une estimation du mouvement par assimilation d'images radar avec deux résolutions spatiales ; 2- une extrapolation de la dernière observation à l'horizon temporel désiré à partir du mouvement précédemment calculé. À chaque nouvelle acquisition, le mouvement est réestimé et une nouvelle prévision est calculée dans un processus de fenêtre glissante. La méthode est résumée ci-dessous et le lecteur est invité à lire [13] pour plus de détails.

Le vecteur d'état $\mathbf{X} = (\mathbf{w}, \mathbf{I})$ est défini comme une fonction de carré intégrable sur le domaine spatio-temporel $\Omega \times [t_0, t_N]$, avec \mathbf{w} le vecteur vitesse et \mathbf{I} une image synthétique ayant les mêmes caractéristiques que les observations radar. Nous cherchons les solutions du système d'équations suivant :

$$\mathbf{X}(\mathbf{x}, t_0) = \mathbf{X}_b(\mathbf{x}) + \epsilon_B(\mathbf{x}), \forall \mathbf{x} \in \Omega (1)$$

$$\frac{\partial \mathbf{X}}{\partial t}(\mathbf{x}, t) + \mathbb{M}(\mathbf{X})(\mathbf{x}, t) = 0, \forall (\mathbf{x}, t) \in \Omega \times [t_0, t_N] (2)$$

$$\mathbb{H}(\mathbf{X}, \mathbf{O})(\mathbf{x}, t) = \epsilon_R(\mathbf{x}, t), \forall \mathbf{x} \in \Omega, t = t_i (3)$$

Dans l'équation (1), le vecteur $\mathbf{X}_b(\mathbf{x})$ est appelé ébauche et représente la connaissance initiale sur le vecteur d'état. $\epsilon_B(\mathbf{x})$ est un terme d'erreur, supposé stochastique et suivant une loi gaussienne de matrice de covariance $B(\mathbf{x})$. L'équation (2) décrit l'évolution temporelle du vecteur d'état, par le modèle \mathbb{M} . L'état complet du système ne dépend que des N observations $\mathbf{O}(\mathbf{x}, t_i)$ et de l'état initial $\mathbf{X}(\mathbf{x}, t_0)$. L'équation (3) décrit le lien entre le vecteur d'état et les observations discrètes en temps $\mathbf{O}(\mathbf{x}, t_i)$, $i = 1 \dots N$, au moyen d'un opérateur \mathbb{H} appelé opérateur d'observation. Le terme d'erreur $\epsilon_R(\mathbf{x}, t_i)$ est supposé gaussien de matrice de covariance $R(\mathbf{x}, t_i)$. Dans la suite, nous omettons la variable \mathbf{x} pour plus de lisibilité dans les équations. Le problème peut être formulé par approche variationnelle : une fonction de coût est définie et son gradient est calculé par une méthode adjointe [5].

Les deux échelles d'observation, régionale ou locale, sont incluses dans le vecteur $\mathbf{O}(t_i) = (O^r(t_i) \ O^l(t_i))^T$. La composante \mathbf{I} du vecteur d'état incorpore également ces deux échelles : $\mathbf{I}(t) = (I^r(t) \ I^l(t))^T$, mais ramenées à la résolution la plus fine. L'opérateur d'observation compare la composante \mathbf{I} du

vecteur d'état aux observations. Il s'écrit donc : $\mathbb{H}(\mathbf{X}, \mathbf{O})(t) = (I^r(t) - \uparrow O^r(t) \ I^l(t) - O^l(t))^T$ où \uparrow est un opérateur de sur-échantillonnage. Pour estimer le mouvement à deux échelles, la dynamique décrite par \mathbf{w} est supposée stationnaire. C'est une hypothèse physiquement cohérente avec le déplacement des structures observées sur quatre acquisitions radar espacées chacune de 5 minutes. Les structures précipitantes sont de plus considérées comme des traceurs passifs, ce qui se traduit par l'hypothèse que \mathbf{I} est advecté par \mathbf{w} . L'opérateur \mathbb{M} s'écrit donc :

$$\mathbb{M}(\mathbf{X})(t) = (\vec{0} \ \nabla I^r(t) \cdot \mathbf{w} \ \nabla I^l(t) \cdot \mathbf{w})^T$$

L'ébauche du vecteur d'état est défini comme suit : $\mathbf{X}_b = (\mathbf{w}^{\text{pre}} \ \uparrow O^r(t_0) \ O^l(t_0))^T$ où \mathbf{w}^{pre} est l'estimation du mouvement obtenu sur la fenêtre précédente dans l'algorithme de fenêtre glissante. Les matrices B et R sont diagonales. Leurs valeurs sont déterminées en fonction des caractéristiques des radars et de la quantité de pluie mesurée. La méthode calcule la valeur initiale du vecteur d'état $\mathbf{X}(t_0)$ qui transporte le terme d'ébauche \mathbf{I}_b "au mieux", c'est-à-dire que l'advection de \mathbf{I} aux temps t_1, t_2 , et t_3 permet de minimiser l'erreur ϵ_R .

Une fois \mathbf{w} calculé vient l'étape de prévision. Celle-ci est obtenue en intégrant en temps l'équation (2), avec comme conditions initiales \mathbf{w} et les dernières observations disponibles : $\uparrow O^r(t_3)$ et $O^l(t_3)$. L'horizon d'intégration choisi est de 45 minutes. Sur un tel intervalle de temps, l'hypothèse de stationnarité du mouvement n'est plus pertinente et doit être remplacée par la constance lagrangienne du mouvement, qui se traduit par l'opérateur $\mathbb{M}(\mathbf{X}) = (\nabla \mathbf{w} \cdot \mathbf{w} \ \nabla I^r \cdot \mathbf{w} \ \nabla I^l \cdot \mathbf{w})^T$. L'hypothèse d'advection pour les composantes I^r et I^l est conservée.

3 Alternative à la méthode de prévision

Comme discuté dans [13], la méthode proposée est compétitive par rapport à l'état de l'art et a permis d'améliorer significativement les performances de prévision, par rapport à une version à échelle unique. Cette approche possède toutefois plusieurs pistes d'amélioration.

La première concerne l'étape d'estimation du mouvement, qui est sensible aux gradients spatiaux dans les données. Le vecteur vitesse, s'il est correctement estimé en direction, est ainsi sous-estimé en intensité pour les zones de précipitations faibles ou nulles. De ce défaut peut résulter un mauvais positionnement spatial des structures lors de la prévision. Pour éviter cela, nous faisons l'hypothèse d'un vent synoptique, c'est-à-dire avec une intensité locale constante, et nous seuillons la norme du champ de vitesse à 70% de la valeur maximale : les normes inférieures à ce seuil sont relevées à cette valeur et les autres vecteurs restent inchangés.

La deuxième piste d'amélioration concerne l'étape de prévision. Pour des raisons de coût algorithmique, l'équation (2) est discrétisée avec un schéma semi-Lagrangien [10]. Un tel schéma a l'avantage d'être inconditionnellement stable, mais présente l'inconvénient de lisser les structures. Les quantités de

pluie sont en conséquence sous-estimées là où elle doivent être les plus élevées et sur-estimées là où elles doivent être faibles voire nulles. Pour éviter cela, nous proposons de renormaliser les histogrammes des images prédites à l'aide de l'histogramme de la dernière observation disponible. Cette méthode est couramment utilisée pour corriger les distributions des mesures radar par des mesures de pluviomètres [8].

4 Résultats

Dans cette section nous comparons la méthode décrite dans la section 2 avec les variantes de la section 3. L'évaluation porte sur la qualité de prévision, qui est comparée aux observations locales. Nous utilisons pour cela quatre métriques. Les trois premières s'intéressent aux positions des structures à partir d'un seuil de détection de la pluie. Ce seuil est fixé à 0,2 mm (soit 2,4 mm/h), qui correspond à une quantité significative de pluie tombée. L'horizon de prévision choisi est de 45 minutes après la dernière observation disponible. Ce temps est volontairement très élevé au regard des observations disponibles, et cela explique l'écart de qualité des prévisions par rapport aux observations. Malgré tout, cette valeur élevée permet de démontrer la pertinence de notre méthode dans des conditions extrêmes.

La métrique POD est définie comme $\frac{VP}{VP+FN}$, où VP est le nombre de pixels vrais positifs et FN le nombre de faux négatifs. Elle quantifie le recouvrement entre les structures nuageuses observées et les structures prédites réellement observées. La métrique FMS correspond à $\frac{VP}{FP+FN+VP}$, où FP est le nombre de pixels faux positifs. Elle quantifie le recouvrement entre les structures nuageuses observées et les structures prédites. La métrique FAR s'exprime par $\frac{FP}{FP+VN}$. Elle quantifie le recouvrement entre les fausses structures nuageuses prédites et la totalité des structures prédites. La quatrième métrique est MAE, qui mesure le taux moyen d'erreur entre quantité de pluie prédite et observée. Les métriques POD et FMS doivent avoir des valeurs le plus élevées possibles, au contraire des métriques FAR et MAE.

Nous avons utilisé soixante images radar divisées en deux épisodes de trente images chacun. Chaque épisode concerne le suivi radar d'un évènement pluvio-orageux. La figure 1 montre les valeurs obtenues sur les quatre métriques pour la méthode de référence et sa variante-1 avec seuillage des vitesses, pour les deux épisodes pluvieux. La table 1 donne les moyennes des quatre métriques sur l'ensemble des données, pour la référence et la variante-1. Ces statistiques démontrent une nette amélioration des performances, pour ce qui est de la localisation spatiale des structures pluvieuses. La figure 2 illustre un exemple typique de résultat, où l'on peut observer que les structures sont effectivement mieux localisées, améliorant globalement les métriques spatiales (POD, FMS et FAR). En revanche, la métrique MAE n'est pas globalement améliorée.

Nous avons également étudié les résultats de la variante-2 obtenue lorsqu'on ajoute la renormalisation de la prévision à l'aide de l'histogramme de la dernière observation (voir table 1, lignes Var. 2). On constate que les métriques POD et FMS

continuent de s'améliorer. Pour mémoire ces métriques s'attachent surtout à quantifier la qualité des vrais positifs. En revanche la métrique FAR est dégradée pour l'épisode 1, en raison de l'apparition de faux négatifs lors du réhaussement de l'histogramme pour les plages de valeurs les plus faibles. La MAE est également dégradée pour la même raison et pour les deux épisodes. C'est en fait une conséquence attendue du schéma semi-Lagrangien qui est, nous l'avons expliqué, trop lissant, et qui n'est pas corrigé par la normalisation d'histogramme.²

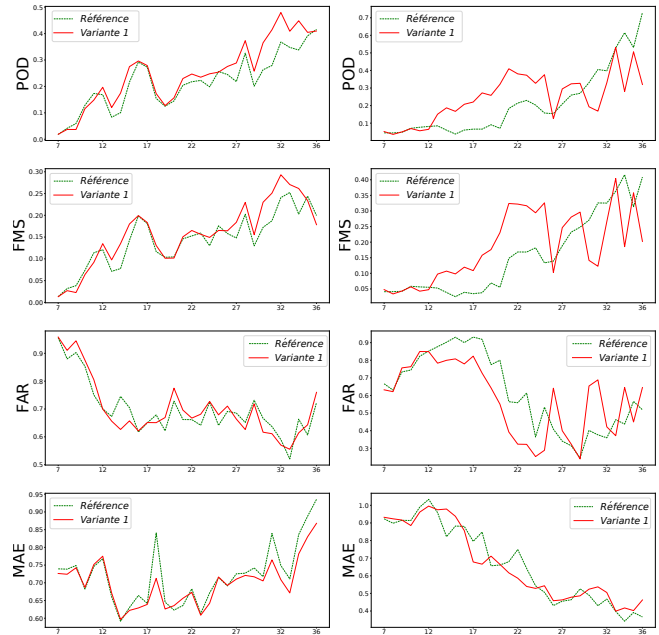


FIGURE 1 – Comparaison entre la méthode de référence et sa variante-1 pour les 4 métriques et pour l'épisode 1 (colonne de gauche) et 2 (colonne de droite). L'axe des abscisses représente le numéro de plan dans la séquence d'images radar.

TABLE 1 – Moyenne des métriques pour les deux épisodes avec les variantes 1 et 2.

		POD	FMS	FAR	MAE
Épi. 1	Réf.	0.2157	0.1429	0.7014	0.7200
	Var. 1	0.2472	0.1573	0.6987	0.7001
	Var. 2	0.2498	0.1791	0.7250	1.0515
Épi. 2	Réf.	0.2106	0.1559	0.5815	0.6677
	Var. 1	0.2459	0.1852	0.6184	0.6687
	Var. 2	0.3790	0.2475	0.5772	1.1845

5 Conclusion

Dans cet article, nous avons présenté des variantes d'une méthode multi-échelles d'estimation des précipitations à horizon temporel proche. L'objectif est d'obtenir des résultats de

². Le lecteur est invité à consulter <https://sites.google.com/view/grets2019> qui contient des animations et quelques informations complémentaires sur les données et les résultats.

qualité pour un horizon de prévision supérieur à 45 minutes. Les deux modifications que nous avons proposées, renormalisation des vitesses et des histogrammes de prévision, ont permis d'améliorer significativement les résultats pour ce qui concerne la localisation spatiale des structures nuageuses. Entre 45 à 60 minutes les performances se dégradent, mais restent acceptables. Toutefois, la nature du schéma numérique utilisé pour l'intégration du modèle lors de la phase de prévision est trop lissant et a pour conséquence de dégrader quantitativement les prévisions de précipitation au-delà de 60 minutes. Les perspectives de ce travail sont d'utiliser des schémas numériques conservatifs et explicites, réputés moins lissants, pour l'intégration du modèle. Nous travaillons actuellement sur l'utilisation de schémas qui réduisent la diffusion implicite comme [9] et des schémas de type CIP [12]. L'inconvénient de ces schémas concerne le coût algorithmique et le fait qu'ils soient soumis aux conditions CFL, ce qui implique de réduire la taille du pas de temps et d'en augmenter le nombre. Les contraintes industrielles imposent un en moins de 5 minutes ce qui nécessite d'utiliser des GPU adaptés à la fine granularité du problème. Dans la présente étude, on donne la même importance aux images régionale et locale. Nous reconsidérons la pertinence de cette hypothèse, notamment en fonction des classes d'événements.

Références

- [1] N. Bowler, C. Pierce, and A. Seed. Development of a rainfall nowcasting algorithm based on optical flow techniques. *Journal of Hydrology*, 288 :74–91, 03 2004.
- [2] M. Dixon and G. Wiener. Titan : Thunderstorm identification, tracking, analysis, and nowcasting—a radar-based methodology. *Journal of Atmospheric and Oceanic Technology*, 10 :785, December 1993.
- [3] E. Huot, Y. Xu, G. Korotaev, L. Wu, I. Herlin, and F.-X. Le Dimet. Estimation de vitesses par assimilation de données variationnelle. In *Colloque GRETSI*, pages 1285–1288, 2007.
- [4] J. Johnson, P. MacKeen, A. Witt, E. Mitchell, G. Stumpf, M. Eilts, and K. Thomas. The storm cell identification and tracking algorithm : An enhanced wsr-88d algorithm. *Weather and Forecasting*, 13(2) :263–276, 1998.
- [5] F.-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations : Theoretical aspects. *Tellus*, 38A :97–110, 1986.
- [6] Y. Lepoittevin, D. Béréziat, I. Herlin, and N. Mercier. Continuous tracking of structures from an image sequence. In *VISAPP*, pages 386–389, 2013.
- [7] N. Papadakis and E. Mémin. A variational technique for time consistent tracking of curves and motion. *Journal of Mathematical Imaging and Vision*, 31(1) :81–103, May 2008.
- [8] E. Rabiei and U. Haberlandt. Applying bias correction for merging rain gauge and radar data. *Journal of Hydrology*, 522, 03 2015.
- [9] P. Smolarkiewicz. A simple positive definite advection scheme with small implicit diffusion. *Monthly Weather Review*, 111 :479–487, March 1983.
- [10] A. Staniforth and J. Côté. Semi-lagrangian integration schemes for atmospheric models - a review. *Monthly Weather Review*, 119 :2206–2223, September 1991.
- [11] C. Thomas, T. Corpetti, and E. Mémin. Data assimilation for convective cells tracking on meteorological image sequences. *IEEE Trans. on Geoscience and Remote sensing*, 48(8) :3162–3177, 2010.
- [12] T. Yabe and T. Aoki. A universal solver for hyperbolic equations by cubic-polynomial interpolation i. one-dimensional solver. *Computer Physics Communications*, 66 :219–232, 09 1991.
- [13] A. Zebiri, D. Béréziat, E. Huot, and I. Herlin. Rain Nowcasting from Multiscale Radar Images. In *VISAPP*, 2019.

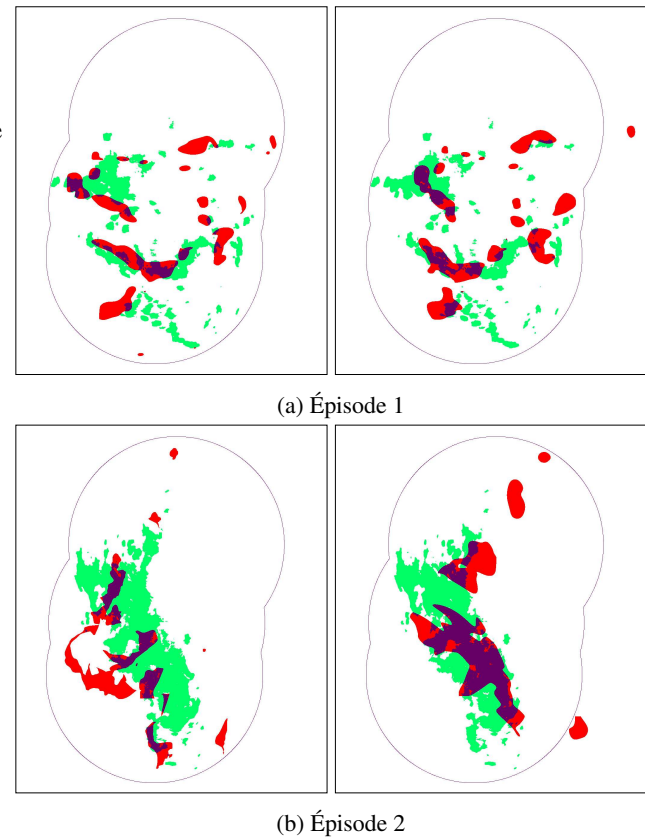


FIGURE 2 – Prévisions à 45 minutes (en rouge) obtenues par la méthode de référence (colonne de gauche) et sa variante-1 (colonne de droite) comparées aux observations locales (en vert). L'intersection entre prévision et observation, en couleur grenat, devrait être idéalement la plus grande possible.