Modeling high rainfall regions for flash flood nowcasting

Yann Lepoittevin, Isabelle Herlin

Modeling high rainfall regions for flash flood nowcasting

Yann Lepoittevin
Inria, Project team Clime
Le Chesnay, France
Email: Yann.Lepoittevin@inria.fr

Isabelle Herlin
Inria, Project team Clime
Le Chesnay, France
Email: Isabelle.Herlin@inria.fr

Abstract—The paper concerns the description of a nowcasting system used for flash floods mitigation, which is based on ground radar images acquired at local scale. The forecast of rain quantity is usually obtained either with numerical weather prediction models or with image-based methods. The first solution is suitable for temporal horizons above three hours due to the computational time required for applying data assimilation techniques to these high dimension models. The second solution is therefore applied for nowcasting up to two hours. Such nowcasting system involves two main parts. The first one estimates the motion field and the second one forecasts the rainfall location and quantity using the radar images and the estimated motion field. The paper focuses on the estimation process, which is obtained by assimilating radar images with an image model describing the temporal evolution of the studied system. The state vector of this model includes the motion field, the image brightness and a description of the objects displayed on the image data. These objects correspond to the rainy regions, whose displacement must be correctly assessed if the estimated motion field is accurate. Results demonstrate that this extended description of the image data allows to produce an improved forecast.

I. Introduction

The paper concerns the nowcasting of flash floods from ground-based radar images acquired at the city scale. Illustration is obtained with data acquired for the city of Clermont-Ferrand in France, which is strongly affected by flood events in spring and summer. Rainfalls forecast is usually obtained in two ways: either from numerical weather prediction models or from image-based methods. The first solution is used for temporal horizons above three hours according to the computational time required for applying data assimilation techniques to these high dimension models. The second solution is usually applied for nowcasting up to two hours, for instance in Mecklenburg et al. [1] and Ebert et al. [2], as modeling the long term appearance or disappearance of clouds requires involving physical laws. A nowcasting system involves two main parts, as described in Lepoittevin et al. [3]. The first one estimates the current motion field and the second one forecasts the rainfall location and quantity, using as input the motion field and the radar acquisitions. The core of the paper concerns the first part, which is obtained by assimilating radar images with an image model. The innovation concerns the definition of the state vector associated to the image model and of its evolution law. As the clouds displayed on the image data are transported by motion accordingly with the image brightness

values and provide additional information on the observed dynamics, the state vector, which describes the studied system, includes both the brightness function and a representation of the objects. Such extended description is required in order to better assess the rainfall displacement and their future location.

In this paper, motion estimation relies on three consecutive radar acquisitions that are named $\mathbf{A}_{e_1}$ to $\mathbf{A}_{e_3}$. I_{f_i} denotes the ith forecasted image and $\mathbf{A}_{f_i}$ the radar image acquired at the same date. As no ground-truth on motion is available, the only validation of the method comes from the comparison of I_{f_i} and $\mathbf{A}_{f_i}$. However, their discrepancy corresponds not only to the inaccuracy of the estimated motion field but also of physical processes that create appearance or disappearance of clouds. The estimation and forecast parts are illustrated by Fig 1.

The input data are X band images from ground radars, acquired every five minutes, with an angular resolution of one degree and a radial resolution of 150 meters. The reflectivity measures are first converted into precipitations rates with the Z-R formula from Marshal-Palmer [4] and second transformed from the polar grid to the Cartesian one, as displayed on the left column of Fig. 2. Motion is then estimated (see Section II) from acquisitions $\mathbf{A}_{e_1}$ to $\mathbf{A}_{e_3}$ and used for the nowcast (see Section III) with $\mathbf{A}_{e_3}$ as initial value (see Fig. 1).

II. Estimation component

Most estimation methods of the literature on nowcasting are either based on optical flow computation, which was first described by Horn and Schunk [5], or on clouds tracking. They usually rely on only two consecutive image acquisitions. Our approach is alternatively based on an image evolution model and an image assimilation approach, as the one described by Béréziat et al. [6]. The temporal evolution of motion is ruled either by a stationarity equation or by advection laws, depending on the image acquisition period and on the duration of the time interval used for the estimation. This paper illustrates the approach with the stationarity option, which is valid in the applicative context because the acquisition period is equal to 5 minutes and only three acquisitions are used for the estimation on a 10 minutes period.

The studied image domain is denoted Ω. A state vector $\mathbf{X}$ is defined on the space-time domain D, depending on the location $\mathbf{x}$ and date t. The vector $\mathbf{X}$ has three components: it

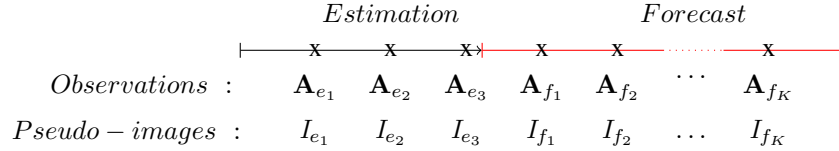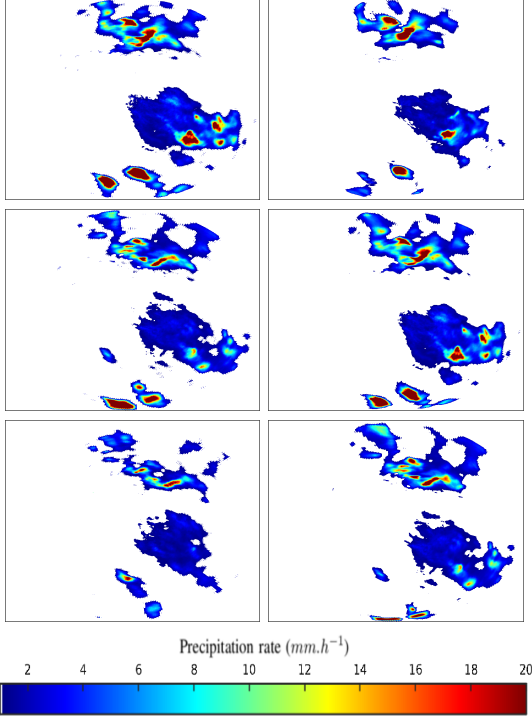| Observations :	$\mathbf{A}_{e_1}$	$\mathbf{A}_{e_2}$	$\mathbf{A}_{e_3}$	$\mathbf{A}_{f_1}$	$\mathbf{A}_{f_2}$	$\cdots$	$\mathbf{A}_{f_K}$
Pseudo − images :	I_{e_1}	I_{e_2}	I_{e_3}	I_{f_1}	I_{f_2}	$\ldots$	I_{f_K}

Fig. 1: Estimation and forecast.

Fig. 2: Left: radar acquisitions $\mathbf{A}_{e_1}$, $\mathbf{A}_{e_2}$ and $\mathbf{A}_{e_3}$. Right: estimations I_{e_1}, I_{e_2} and I_{e_3}.

includes the motion function $\mathbf{w}$, an image function I and an object representation ϕ. I is defined with the same characteristics than the radar data. ϕ is the distance to the boundary B of the rainy regions identified on the image I. Fig. 3 displays, on the left, the rainy regions obtained by thresholding the radar image and, on the right, the distance map ϕ to the boundary B of these regions. The two components I and ϕ are both transported by motion $\mathbf{w}$. The evolution equations verified by the state vector $\mathbf{X}$ are the partial differential equations describing the stationarity of $\mathbf{w}$ and the transport of I and ϕ by $\mathbf{w}$:

$$\begin{cases} \dfrac{\partial \mathbf{w}}{\partial t}(\mathbf{x}, t) = 0 \\ \dfrac{\partial I}{\partial t}(\mathbf{x}, t) = -\mathbf{w}(\mathbf{x}, t) \cdot \nabla I(\mathbf{x}, t) \\ \dfrac{\partial \phi}{\partial t}(\mathbf{x}, t) = -\mathbf{w}(\mathbf{x}, t) \cdot \nabla \phi(\mathbf{x}, t) \end{cases} \quad (1)$$

Motion is then estimated by an iterative process so that its final value allows: 1- the image function I to be close to the effective acquisition $\mathbf{A}$, 2- the object component ϕ to be close
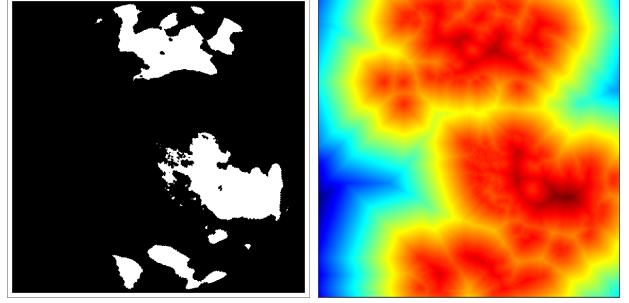

Fig. 3: Left: threshold of the precipitation rate. Right: distance map to the boundary B of the rainy regions.

to a value $D_{\mathbf{A}}$, which is the distance map to the boundary $B_{\mathbf{A}}$ of the rainy regions on the radar image $\mathbf{A}$. Equivalently, motion is estimated so that the boundary B associated with ϕ is close to $B_{\mathbf{A}}$.

The estimation relies on a 4D-Var data assimilation approach, which was initially described by Le Dimet and Talagrand [7]. This method computes the minimum of a cost function J, depending on the initial value of the state vector $\mathbf{X}(0)$, with a gradient descent technique. The cost function value is obtained from a forward integration of the evolution model described by System (1) and its gradient value from the backward integration of the adjoint model. J is defined by:

$$J(\mathbf{X}(0)) = \int_D \frac{\varepsilon_I(\mathbf{x}, t)^2}{R_I(\mathbf{x}, t)} d\mathbf{x}dt + \int_D \frac{\varepsilon_\phi(\mathbf{x}, t)^2}{R_\phi(\mathbf{x}, t)} d\mathbf{x}dt$$
$$+ \int_\Omega \varepsilon_b(\mathbf{x})^T B(\mathbf{x})^{-1} \varepsilon_b(\mathbf{x}) d\mathbf{x} \quad (2)$$

The first term is the integral, on the space-time domain D, of a function of the error term ε_I, which is defined as the difference between the image function and the radar acquisition:

$$\varepsilon_I(\mathbf{x}, t) = I(\mathbf{x}, t) - \mathbf{A}(\mathbf{x}, t) \quad (3)$$

R_I is the variance of the Gaussian variable ε_I. It is given the same value than the variance of the acquisition error.

The second term is the integral, on the space-time domain D, of a function of the error term ε_ϕ, which is defined as the difference between the distance map ϕ, computed on the image function I, and the distance map $D_{\mathbf{A}}$, computed on the radar acquisitions:

$$\varepsilon_\phi(\mathbf{x}, t) = \phi(\mathbf{x}, t) - D_{\mathbf{A}}(\mathbf{x}, t) \quad (4)$$

R_ϕ is the variance of the Gaussian variable ε_ϕ.

Let $\mathbb{P}$ be the projection of the state vector $\mathbf{X}$ on the vector with two components I and ϕ. The third term is the integral, on the image domain Ω, of a function of the difference ε_b between the estimated value of $\mathbb{P}(\mathbf{X}(0))$ and an initial value $\mathbb{P}(\mathbf{X}_b)$:

$$\varepsilon_b = \mathbb{P}(\mathbf{X}(0)) - \mathbb{P}(\mathbf{X}_b) \tag{5}$$

$\mathbf{X}_b$ is the background of the data assimilation process. It defined in the following way. I_b is the first image $\mathbf{A}_{e_1}$ of the estimation process and ϕ_b is the distance map to the boundary of the rainy regions computed on $\mathbf{A}_{e_1}$. The operational system is implemented with a sliding-window approach that estimates the state vector $\mathbf{X}$ on temporal windows of three acquisitions with one frame shift between two consecutive windows. $\mathbf{w}_b$ is then defined as the result of motion estimation on the previous window. The covariance $B(\mathbf{x})$, which is involved in Eq. (2), is diagonal with values equal to $R_I(\mathbf{x}, t)$ and $R_\phi(\mathbf{x}, t)$.

As no ground-truth is available, validation of the estimated motion field is postponed to the forecast process. However, a first analysis is at hand by comparing the estimated image function I and the radar images $\mathbf{A}$. The left column of Fig. 2 displays three consecutive acquisitions, $\mathbf{A}_{e_1}$, $\mathbf{A}_{e_2}$ and $\mathbf{A}_{e_3}$, used as input for the motion estimation process, while the right column gives the estimations I_{e_1}, I_{e_2} and I_{e_3}. An additional analysis comes from the comparison between the estimated distance map ϕ and the distance map $D_\mathbf{A}$ that is computed on the radar image acquired at the same date.
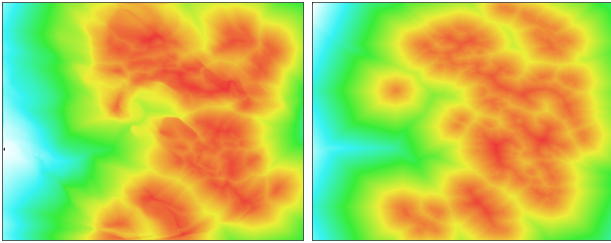
Fig. 4: Left: estimated distance map ϕ_{e_3}. Right: distance map to the contours on the radar image $\mathbf{A}_{e_3}$.

The estimated motion field is displayed on the left column of Fig. 5 and demonstrates the impact on ϕ. A same display
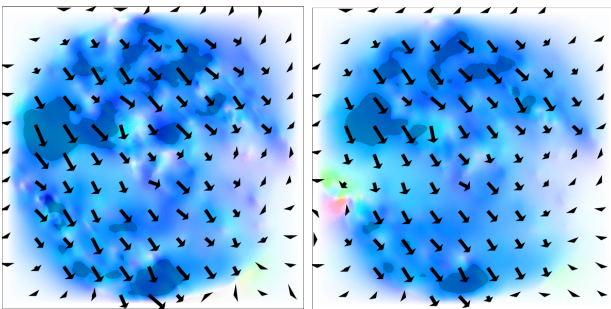
Fig. 5: Left: estimation obtained with ϕ. Right: without.

is given on Fig. 6 for another time interval. Understanding the visible impact of ϕ on the results requires a further description
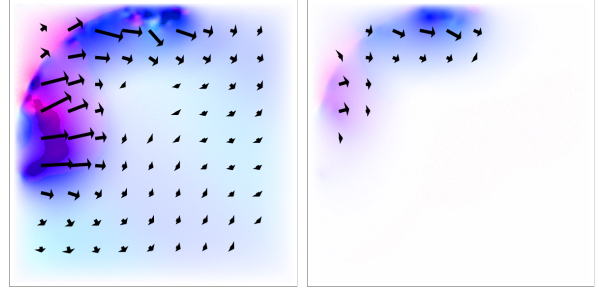
Fig. 6: Left: estimation obtained with ϕ on another time window. Right: without.

of its role during the estimation. At each location $\mathbf{x}$ of the image domain, the value of $\phi(\mathbf{x})$ is the distance to the closest point $\mathbf{c}$, belonging to the boundary B associated with ϕ. Therefore, ϕ relates each pixel $\mathbf{x}$ of Ω to one point $\mathbf{c}$ of B. During the data assimilation process, ϕ (and therefore B) is iteratively modified from its background value, in order to correctly describe the boundary $B_\mathbf{A}$ of the rainy regions on the radar acquisitions. Modifying the position of a given point $\mathbf{c}$ of the boundary B has a strong influence on all pixels $\mathbf{x}$, which have $\mathbf{c}$ as closest point. The estimation of the state vector $\mathbf{X}(\mathbf{x})$ and of motion $\mathbf{w}(\mathbf{x})$ at these pixels is then linked to the estimation at point $\mathbf{c}$. If $\mathbf{c}$ is moved for being closer to one point $\mathbf{c}_\mathbf{A}$ of the boundary $B_\mathbf{A}$, this impacts the estimation of the state vector (and consequently of motion) on all pixels $\mathbf{x}$ that are linked to $\mathbf{c}$. This explains why motion is estimated even far away of the rainy regions as proven on Fig. 6. An influence value is defined for each point $\mathbf{c}$ as:

$$\xi(\mathbf{x}) = \frac{1}{|d(\mathbf{x}, \mathbf{c}) - \phi(\mathbf{x})|} \tag{6}$$

with $d(\mathbf{x}, \mathbf{c})$ being the Euclidean distance between pixel $\mathbf{x}$ and point $\mathbf{c}$. Fig. 7 gives the influence map related to the point $\mathbf{c}$ displayed by the gray circle on the black boundary.

Low influence   High influence

Fig. 7: Influence map of the circled pixel.

III. Forecast component

In a sliding-window framework, as soon as the third observation $\mathbf{A}_{e_3}$ is available, a motion field $\mathbf{w}$ is estimated and used for nowcasting from $\mathbf{A}_{e_3}$, before the next acquisition $\mathbf{A}_{f_1}$, as displayed on Fig. 1. As the time horizon ranges from 30 minutes to 2 hours, in order to have enough time for applying preventive measures, the temporal duration of nowcasting is higher than the one of the estimation. The advection option is then applied for describing the temporal evolution of the motion field:

$$\frac{\partial \mathbf{w}}{\partial t}(\mathbf{x}, t) = -\mathbf{w}(\mathbf{x}, t) \cdot \nabla \mathbf{w}(\mathbf{x}, t) \qquad (7)$$

Temporal integration of Eq. 7 follows a three step process:

- Motion is integrated from one acquisition date i to the next one $i + 1$ thanks to a Godunov scheme with a one second time step.
- A semi-Lagrangian scheme is then applied at date $i + 1$ to compute I and ϕ from their values at date i.
- The two previous points are iterated from $i + 1$ to $i + 2$.

The use of these two integration schemes allows to get benefit of both the stability of the semi-Lagrangian scheme and the accuracy of the Godunov scheme when applied to the non linear advection.

From an operational point of view, comparison of the forecasted images with the effective acquisitions has no interest at the pixel level. The required output is the location and approximate date of high cumulative rain quantities that can result in floods. The cumulative rain quantity is obtained from the time integration of the instantaneous rainfall rates over a given period that is set to 1 hour for illustrating the next discussion. The performance of the method is measured by the Root Mean Square Error (RMSE) between these forecasted rain quantities and the observed ones. For the data set corresponding to the the studied meteorological events (more than 150 frames), the value of the RMSE is around 2 millimeters, which is accurate enough for operational use.

Another interesting output is the detection of threshold exceedence, when the cumulative rain quantity is above a given value that requires emergency measures. The percentage of a correct forecast at 1 hour of these threshold exceedence events (the forecast corresponds to a real event) is around 90%.

The last performance criteria concerns the detection of such events. The involvement of an object representation ϕ in the state vector has a positive or neutral impact on that criteria. In average, the percentage of detection at 1 hour is close to 60% with the object component and decreases to 50% without, as displayed in Fig. 8.

IV. Conclusion and Perspectives

This paper describes an operational nowcasting system for flash floods associated to high rainfalls. The short horizon of nowcasting is obtained thanks to an image-based method applied on ground radar acquisitions. In a similar way than usual image-based systems, our approach relies on two main
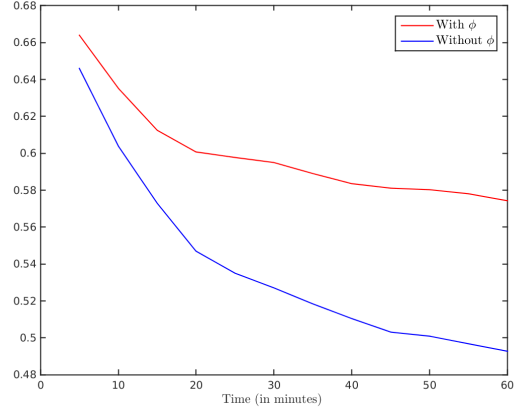

Fig. 8: Percentage of detection with and without ϕ.

parts: one related to wind fields estimation and one to the forecast. The innovation comes from the estimation part that relies on a data assimilation technique, with an image evolution model including a description of the high rainfall regions. This object description allows a better monitoring of the rainfall displacement and results in an improved nowcast.

Two main perspectives will be studied. First, the whole system will be rewritten in polar coordinates, according to the acquisition sensor. This will suppress unnecessary preprocessing. Second, the nowcast system will be coupled with a Numerical Weather Prediction model (NWP), in order to slowly switch from the image-based nowcast to the NWP output and provide a forecast on the whole temporal horizon as done in Ridal et al. [8].

References

[1] S. Mecklenburg, A. Jurczyk, J. Szturc, and K. Osrodka, "Quantitative precipitation forecasts (QPF) based on radar data for hydrological models," in *COST action*, 2002.
[2] E. Ebert, L. Wilson, B. Brown, P. Nurmi, H. Brooks, J. Bally, and M. Jaeneke, "Verification of nowcasts from the WWRP Sydney 2000 forecast demonstration project," *Weather and Forecasting*, 2004.
[3] Y. Lepoittevin and I. Herlin, "Assimilation of radar reflectivity for rainfall nowcasting," in *International Geoscience and Remote Sensing Symposium*, 2015.
[4] J. S. Marshall and W. M. K. Palmer, "The distribution of raindrops with size," *Journal of meteorology*, 1948.
[5] B. Horn and B. Schunk, "Determining optical flow," *Artificial Intelligence*, 1981.
[6] D. Béréziat and I. Herlin, "Solving ill-posed image processing problems using data assimilation," *Numerical Algorithms*, 2011.
[7] F. Le Dimet and O. Talagrand, *Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects.* Tellus, 1986.
[8] M. Ridal, M. Lindskog, N. Gustafsson, and G. Haase, "Optimized advection of radar reflectivities," *Atmospheric Research*, 2011.