Review of paper:

Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0.1)

by Georgy Ayzel, Maik Heistermann, and Tanja Winterrath

Summary

The authors present a new open source python library for benchmarking radar-based nowcasting systems based upon optical flow techniques. They propose 2 local (sparse) and 2 global (dense) approaches for motion field estimation and extrapolation of radar echoes. The nowcasts are verified using 11 precipitation events over Germany and compared to the operational nowcasting product of the DWD. Despite some variability of forecast accuracy, the dense group of techniques provides the best performance.

General comments

I am very glad to see this open source initiative in the field of radar-based precipitation nowcasting and I fully support it. I enjoyed reading the paper and invested significant time to provide a useful review. Hopefully, the authors will appreciate my efforts and suggestions on how to improve the manuscript.

Below you will find a table with the detailed comments line by line.
My main comments on the paper are summarized here:

- The forecast verification is well done, but in my opinion it should include a verification of the statistical properties of the advected rainfall fields to understand the degree of numerical diffusion, which can be a major problem in precipitation nowcasting if not properly handled. Such effect usually leads to an undesired smoothing of the precipitation fields, which reduces the more interesting high rainfall intensities and complicates the inter-comparison of models.
- As the paper presents new optical flow and advection techniques, it must include some additional figures showing examples of motion fields and precipitation nowcasts, e.g.:
    - A multi-panel figure with vector plots of the motion fields retrieved by the different methods overlaid on top of radar images (for example one “rotational” precipitation event).
    - A multi-panel figure showing examples of observed and nowcasted precipitation fields at different lead times, e.g. 30 or 60 minutes. This would be very useful to understand the quality and realism of the advected rainfall fields, and check whether there are any artefacts due to numerical diffusion and interpolation processes.
- Some statements in the literature review are a bit imprecise and could be improved.

Specific comments

Page 1, line 3, Page 2, line 14	"extrapolate the motion" -> "extrapolate the radar echoes". The motion field is usually kept fixed and only the radar echoes are extrapolated, although in some cases it may be beneficial to extrapolate the motion field together with the precipitation echoes.
Page 2, line 4	The cited approaches (analogue, local Lagrangian and stochastic) were mentioned in the context of probabilistic precipitation nowcasting. They all provide empirical estimates of the probability density function in different ways. Please update accordingly.

Page 2, lines 5-6	Foresti et al. (2015) did not use the correlation coefficient as a measure of similarity to retrieve the analogues (as done e.g. by Atencia et al., 2015), but rather the Euclidian distance in the space of principal components. Please adjust the statement.
Page 2, lines 8-10	I think there is some confusion about the definition of "local Lagrangian method". The cited paper (Foresti et al., 2015) follows the definition of Germann and Zawadzki (2004), which defines the “local Lagrangian” as one possible method to derive a probabilistic nowcast. This is achieved by collecting the precipitation values upstream in a local neighbourhood, whose size is increased as a function of lead time.
Page 3, line 1	I fully agree that optical flow libraries have been around for long, but they cannot be directly applied for the retrieval of radar echo motion without important adaptations and tests. For example, they must be tuned to represent the typical range of advection speeds of real precipitation fields, they must be spatially dense and extrapolate well also in regions without precipitation, etc. This is why papers like yours are important contributions to make the necessary adaptations and tests.
Page 3, line 8 Page 4, line 24	It would be interesting to know why you decided not to include in the list of benchmark extrapolation techniques the backward-in-time semi-Lagrangian scheme, which is generally accepted to be the most appropriate method (Germann and Zawadzki, 2002). The forward scheme is known to produce holes in the precipitation field in presence of divergent vectors, which need to be interpolated. This inevitably leads to additional numerical diffusion.
Page 3, line 26	I cannot understand properly why you mention the concept of scale-dependence in the context of local LK methods. Please explain how local optical flow techniques account for scale-dependence.
Page 4, line 5	I am a bit worried that the use of warping and interpolation of discontinuities in the advected radar field can lead to serious numerical diffusion effects. The most appropriate method to test this issue is to compute the Fourier spectrum of the original and advected fields to check whether there is loss of power at the high spatial frequencies (see Fig. 10 in Germann and Zawadzki, 2002). A simpler approach would be to compare the histogram of nowcasted rainfall fields at different lead times with the one of the last observed radar image. The variance and histogram should be conserved during the extrapolation.
Page 4, line 26	I agree that the constant-vector approach does not explicitly allow to account for rotation. However, if the advection is applied recursively in short time steps the rotation can be approximated by a set of short straight lines (at the cost of stronger diffusion). Despite this fact, I believe that a good implementation of the semi-Lagrangian scheme should consistently give better (or comparable) results than the constant-vector approach.

Page 4, line 29	Also here I would study the effect of numerical diffusion caused by the interpolation. Numerical diffusion can also have undesired consequences when comparing (benchmarking) different nowcasting models. In fact, a precipitation nowcast that loses power at the high spatial frequencies will be generally smoother. This behavior will be rewarded in terms of some verification scores (in particular the MAE/RMSE), which affects the comparison with other models. A fair comparison of different nowcast systems should be done at similar spatial scales, for example using Fourier or wavelet decompositions.
Page 6, line 10	"programmatic realization" is a strange expression.
Page 6, line 31	“rainfall depth product”. Is it the instantaneous intensity in mm/hr or an accumulation?
Page 6, line 23	It would be very interesting to move the CSI verification at a threshold of 5 mm/hr from the supplementary material to the actual paper. These rainrates are the ones that are relevant to trigger warnings for severe weather.
Page 8, lines 5-10	You are correct. Detailed motion fields provide better skill at short lead times, while smoother motion fields are more adapted for longer lead times. Similarly to precipitation fields, the motion fields also have an intrinsic predictability (persistence). This can be exploited by gradually smoothing the motion field in a way that is consistent with its predictability.
Page 8, line 14	All the proposed solutions to the problem of low predictability at convective scales are based on the optical flow and are all valid options. However, precipitation, and in particular the one of convective nature, has a large unpredictable component that we will likely never be able to predict. Therefore, the nowcasting community needs to admit the incapability of providing accurate deterministic precipitation forecasts and find ways to estimate and communicate the inherent uncertainty. I am glad that you presented this issue in the conclusion at page 9, lines 22-25, but it would be a good idea to make this point stronger.
Page 9, line 20-21	I also believe that we should not discard the Sparse models. One possibility is to make them “dense” by interpolating the motion vectors before applying the advection scheme (hopefully semi-Lagrangian).
Figures 1 and 2	These are extremely clean and nice presentations of the methods.
Figure 3	You may add in the caption that the figure shows the forward-in-time semi-Lagrangian method.
Figure 5	You may consider writing a more descriptive figure caption, e.g. “Verification of the different optical flow based nowcasts in terms of MAE for 11 precipitation events over Germany”.
Page 7 line 15, Figures 6-7	Is there an explanation on why the RADVOR nowcasting method performs poorly in the first 5-10 minutes? The effect seems quite systematic and I have a

	hard time explaining it with the faster movement of precipitation fields.
Page 9, lines 27-30	With respect to the use of open source libraries to promote the developments in the field of nowcasting, you could also mention how you would imagine the contribution from rainymotion to the developments of other projects, as for example the probabilistic nowcasting library pysteps (https://pysteps.github.io/). In my opinion, any improvement in optical flow methods, e.g. using the rainymotion library, will also have a positive impact on the quality of probabilistic nowcasts. This could represent an interesting synergy between the two libraries, in line with the open source philosophy.

References

- Atencia, A., and I. Zawadzki, 2015: A comparison of two techniques for generating nowcasting ensembles. Part II: Analogs selection and comparison of techniques. Mon. Wea. Rev., 143 (7), 2890–2908.
- Germann, U. and I. Zawadzki, 2002: Scale-Dependence of the Predictability of Precipitation from Continental Radar Images. Part I: Description of the Methodology. Mon. Wea. Rev., 130, 2859–2873.
- Germann, U. and I. Zawadzki, 2004: Scale Dependence of the Predictability of Precipitation from Continental Radar Images. Part II: Probability Forecasts. J. Appl. Meteor., 43, 74–89.