Response letter

Dear Referees, dear Editor,

We would like to thank you again for your positive comments and constructive suggestions for the improvement of our manuscript. In this document, we would like to provide our responses to the comments of each of the three referees including the decision on specific changes in the manuscript. For that purpose, we use the following color code:

black: original referee comment
blue: our original response in the Interactive discussion
green: our final response and the specific changes made in the revised manuscript

Before addressing each referee comment, we also provide a summary of the most important changes to both the manuscript and the rainymotion library in the course of this revision.

Addressing the referee comments has, in our opinion, substantially improved the paper, and we hope that the quality of the paper now allows for publication in GMD.

Sincerely,
Georgy (on behalf of the authors)

Summary of major changes to the manuscript and the rainymotion library

Benchmarking of the new global optical flow techniques

Based on the reviewers' comments (comments #9 and #12 by Dr. Foresti, #4-7 by Dr. Pulkkinen), we complemented the previously used Farnebäck local optical flow algorithm with procedures to replace zero velocities and a further variational refinement of the obtained velocity field, and also included a set of global optical flow algorithms, such as DIS, DeepFlow, and PCAFlow. We updated the rainymotion library accordingly, and conducted an extensive benchmark experiment to evaluate efficiency of these different optical flow algorithms for precipitation nowcasting. Results showed that the DIS optical flow algorithm provides better results both in terms of verification metrics and computational performance. Thus, we selected the DIS optical flow algorithm as the default option for the tracking step in the Dense group of models in the rainymotion library.

Benchmarking of different advection approaches

Based on the reviewers' comments (comment #9 by Dr. Foresti, and #5 by Dr. Pulkkinen), we performed a benchmark experiment to verify the performance two different implementations (namely forward and backward) of the constant-vector and the semi-Lagrangian advection schemes. Results showed that the backward scheme performs *slightly* better for low rainfall intensity rates (under 0.5 mm/h) and longer lead times (from 30 minutes). For rainfall intensities above 0.5 mm/h and shorter lead times (up to 30 minutes) there are no significant differences between both schemes. Thus, we decided to implement the backward scheme as the default option for precipitation advection in the revised version of the rainymotion library.

Increasing computational performance by a new interpolation approach

The linear interpolation of the advected rainfall pixels to the original radar grid was a serious bottleneck in the computational performance of the Dense group of the rainymotion models for large grids (900x900 pixels in the case of the RY data). In the revised version of the rainymotion library, we replaced the linear interpolation (implemented via scipy.interpolate module) by inverse distance weighting interpolation (implemented via the wradlib.ipol module). This led to a substantial improvement in computational efficiency (by a factor of about 15), without a drop in verification efficiency.

Exemplary investigation of effects of numerical diffusion

Based on the comments #1, #8, #11, #12, and #13 by Dr. Foresti, we introduced an exemplary analysis of the effects of numerical diffusion in Figure 5 of the revised manuscript, based on the loss of power spectral density as compared to the observations. For that case study, we did not find any substantial loss of power at small spatial scales - at least for the lead times of up to one hour investigated in our study. We hypothesize that this encouraging result is due to the fact that we interpolate only *once per lead time* (in the Dense group of models, for both the forward and the backward scheme), and that the warping procedure intrinsically conserves power at small scales while interpolation effects are negligible (Sparse group of models).

Referee comment #1 (by Loris Foresti)

Main comments

1. 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.

RESPONSE: We entirely agree that it would be interesting to verify the statistical properties of the advected rainfall fields. It will be done as suggested in comment #11: using a periodogram of rainfall intensities of advected precipitation fields for different lead times.

ACTION: A new figure (Figure 5 of the revised manuscript) has been added to address the issue of numerical diffusion (see also ACTION reg. comment #11). We also investigated the issue of numerical diffusion in Section 4 (Results) and Section 5.2 (Advection schemes properties and effectiveness) in the revised manuscript.

2. 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.:
    2.1. 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).
    2.2. 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.

RESPONSE: We will add requested figures in the revised version of the manuscript.

ACTION: A new figure (Figure 5 of the revised manuscript) has been added to show examples of motion fields as well as nowcasts for different models and lead times.

3. Some statements in the literature review are a bit imprecise and could be improved.

RESPONSE: We will revise the literature review in the introductory section based on several referee comments (comments #5, #6, #7, and #24 by Dr. Foresti, #1.2 and #2 by Dr. Pulkkinen, #1, #2, and #3 by Dr. Uijlenhoet).

ACTION: We have updated the introductory section of the revised manuscript in accordance with the reviewer's comments (see also the actions related to other reviewers' comments: #5, #6, #7, and #24 by Dr. Foresti, #1.2 and #2 by Dr. Pulkkinen, #1, #2, and #3 by Prof. Uijlenhoet).

Specific comments

4. 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.

RESPONSE: We suggest to rephrase this to "[...] and then to displace the precipitation field to the imminent future (minutes to hours) based on that motion, [...]".
ACTION:
1. We rephrased the corresponding sentence in the abstract of the revised version (Page 1, lines 2-5) as follows:

"[...] A common heuristic prediction approach is to track the motion of precipitation features from a sequence of weather radar images, and then to displace the precipitation field to the imminent future (minutes to hours) based on that motion, assuming that the intensity of the features remains constant ("Lagrangian persistence"). [...]"

2. We rephrased the corresponding sentence in the introductory section of the revised version (Page 2, lines 17-18) as follows:

"[...] In the second step, we use that velocity field to advect the most recent rain field, i.e. to displace it to the imminent future based on its observed motion. [...]"

5. 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.

RESPONSE: Please see response to comment #7.
ACTION: Please see action to comment #7.

6. 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.

RESPONSE: Please see response to comment #7.
ACTION: Please see action to comment #7.

7. 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.

RESPONSE: Comments #5, #6, and #7 are related to one paragraph (Page 2, lines 4-9). We will rewrite the whole paragraph in accordance with the referee’s suggestions and try to make the main message of this paragraph (classification of methods used for radar-based precipitation nowcasting) clearer.
We suggest to rephrase the corresponding paragraph to:
“A variety of radar-based precipitation nowcasting techniques can be classified on three major groups based on assumptions we make regarding precipitation field characteristics (Germann and Zawadski, 2002). The first group -- climatological persistence -- provides nowcasts by using climatological values (mean or median). The second group -- Eulerian persistence -- is based on using the latest available observation as a prediction, and is thus independent from the forecast lead time. The third group -- Lagrangian persistence -- allows the extrapolation of the most recent observed precipitation field under the assumption that the motion field is persistent (Germann and Zawadzki, 2002; Woo and Wong, 2017). In addition, we can classify nowcasting methods based on introduced prediction uncertainty: In contrast to deterministic approaches, ensemble nowcast attempt to account for predictive uncertainty by including different realizations of the motion field and the evolution of rainfall intensity itself. In this study, we focus our model development around the group of Lagrangian persistence models which provide deterministic precipitation nowcasts. Yet, the unified availability of different tracking and extrapolation techniques in the rainymotion library could directly be used to construct ensembles that account for the uncertainty of rainfield displacement.”
ACTION: We have updated the corresponding paragraph in the introductory section (Page 2, lines 4-13) in accordance with the proposed solution in our response:
“[...] A variety of radar-based precipitation nowcasting techniques can be classified into three major groups based on assumptions we make regarding precipitation field characteristics (Germann and Zawadski, 2002). The first group -- climatological persistence -- provides nowcasts by using climatological values (mean or median). The second group -- Eulerian persistence -- is based on using the latest available observation as a prediction, and is thus independent from the forecast lead time. The third group -- Lagrangian persistence -- allows the extrapolation of the most recent observed precipitation field under the assumption that intensity of precipitation features and the motion field are persistent (Germann and Zawadzki, 2002; Woo and Wong, 2017). In addition, we can classify nowcasting methods based on how predictive uncertainty is accounted for: In contrast to deterministic approaches, ensemble nowcasts attempt to account for predictive uncertainty by including different realizations of the motion field and the evolution of rainfall intensity itself (Berenguer et al., 2011). In this study, we focus our

model development around the group of Lagrangian persistence models which provide deterministic precipitation nowcasts. [...]"

8. 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.

RESPONSE: We agree that the original manuscript does not sufficiently address how, on the one hand, parameters of different optical flow techniques affect the specific problem of precipitation field tracking, and, on the other hand, how the results of different optical flow techniques might need further post-processing in order to enhance their usefulness for the extrapolation step (e.g. filling or interpolating zero velocities that might occur in regions of zero rainfall). Given that we also introduce further optical flow/tracking as well as extrapolation techniques (comments #9 and #12 by Dr. Foresti, #4-7 by Dr. Pulkkinen) in the revised version of the manuscript and the rainymotion library, the revised manuscript will address these requirements more precisely and comprehensively.
ACTION: We have added statements which clarify the possibility to optimize the rainymotion models parameters in order to provide better (in terms of verification efficiency) nowcasts or to represent the typical range of advection speeds of real precipitation fields as follows:

- Page 3, lines 7-12: *"[...] That is all the more surprising since open source implementations of fundamental optical flow algorithms (Brox et al., 2004; Bruhn et al., 2005b) have been around for up to 20 years -- with the OpenCV library (https://opencv.org) just being the most widely known. Such libraries provide efficient implementations of various optical flow algorithms for a vast number of research and application contexts. Yet, none can be applied in the QPN context out of the box -- without the need to address additional and specific challenges such as underlying assumptions and constraints of velocity fields, pre- and postprocessing steps, or model parameterization and verification. [...]"*
- Page 3, lines 32-33; Page 4, line 1: *"[...] However, the rainymotion library provides an opportunity to investigate how different optical flow model parameters can affect nowcasting results or how they can be tuned to represent, e.g. the typical range of advection speeds of real precipitation fields. [...]"*

9. 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.

RESPONSE: We originally implemented the forward scheme because it is more intuitive to advect the precipitation field "forward in time" and "downstream in space". Based on the referee's comment, though, we decided to complement the revised version of the rainymotion library with a backward method for the optical flow calculation.

On that basis, we repeated our benchmark experiments by using the backward scheme both for the Dense (constant-vector) and DenseRotation (semi-Lagrangian) models. Results show that the backward scheme performs slightly better for low rainfall intensity rates (under 0.5 mm/h) and longer lead times (from 30 minutes). For rainfall intensity rates over 0.5 mm/h and shorter lead times (up to 30 minutes) there are no significant differences between both schemes. Based on the new results we decided to implement the backward scheme as a default option for precipitation motion field calculation in the revised version of the rainymotion library. We will update the revised version of the manuscript and the supplementary material in accordance with the new results.
However, we also want to note that the intercomparison of different advection schemes provided in Germann and Zawadzki (2002) cannot, in our opinion, be interpreted in a way that "backward-in-time semi-Lagrangian scheme [...] is generally accepted to be the most appropriate method". In the corresponding paper, the forward-in-time scheme is concerted with a gaussian redistribution of advected rainfall in contrast to the interpolation used for backward-in-time scheme. In our library, we adapt the same "interpolate only once" idea of Germann and Zawadzki (2002) -- regardless of the direction used for velocity field (optical flow) calculation that allows intercomparison of forward and backward schemes in a similar setting. Although the new results of our intercomparison are consistent with the referee's statement on the backward scheme being superior (for low rainfall intensities and longer lead times), in our opinion further research is needed to compare the efficiency of different implementations in detail.

ACTION: We have updated Section 2 of the revised manuscript and the corresponding sections in the Supplementary information (Sections S4-S6), based on the results of a new benchmarking experiment with both forward and backward schemes as well as with both the constant-vector and the semi-Lagrangian advection schemes, and also modified the rainymotion library accordingly.

10. 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.

RESPONSE: Our intention was to highlight that for the Sparse group of rainymotion's tracking models we use distinct "corners" instead of storm cells -- this eliminates the need to specify arbitrary and scale dependent characteristics of "precipitation features" while the identification of "corners" depends only on the gradient sharpness in a cell's neighborhood. Of course this will not solve the issue of scale-dependence of the average motion itself. We will clarify these aspects in the revised manuscript.

ACTION: We have updated the corresponding paragraph of Section 2.1 (The Sparse group) as follows (Page 4, Lines 5-8):
"[...] That approach is less arbitrary and scale dependent and thus more universal than classical approaches that track storm cells as contiguous objects (e.g. Wilson et al., 1998) because it eliminates the need to specify arbitrary and scale dependent characteristics of "precipitation features" while the identification of "corners" depends only on the gradient sharpness in a cell's neighborhood. [...]"

11. 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.

RESPONSE: We will update the revised version of the manuscript with figures that describe the level of numerical diffusion for the different models by using the nowcasts' power spectral density for different lead times.

ACTION: We have added a new figure (Figure 5 of the revised manuscript) to describe the level of numerical diffusion for different models and lead times (0, +30, and +60 minutes) based on the power spectral density of the nowcasts. We also discussed those results regarding the relevance of numerical diffusion effects in Section 4 (Results) and Section 5.2 (Advection schemes properties and effectiveness).

12. 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.

RESPONSE: We agree that an accurate implementation of the semi-Lagrangian scheme should yield a skill that is at least equivalent to the constant-vector approach. We have found two possible reasons why our original implementation did not achieve that: 1. Errors in the estimation of motion fields (e.g. with anomalies, artefacts etc.) could affect the forecast in the semi-Lagrangian advection scheme more than in the constant-vector scheme, since displacement vectors from regions of higher uncertainty might be “activated” more frequently; 2. Higher complexity of semi-Lagrangian scheme implementation which involves interpolation on two levels: when we advect each pixel and try to find the new velocity vector for any new pixel location, and during the final interpolation of intensities.
We attempted to address the estimation of field motion using the (local) Farnebäck optical flow method by implementing a variational refinement procedure to smooth the velocity field, and to get rid of spurious velocities, and by implementing different global optical flow methods that usually provide more smooth and robust motion fields (see also comment #4.1 and #7 by Dr. Pulkkinen). As a result, we included both the variational refinement and different global methods for the tracking step in the rainymotion library, and included these approaches in our benchmarking experiments.
According to these new results, the implementation of the Dense Inverse Search (DIS) global optical flow method (Kroeger et al., 2016) provides better results than the Farnebäck method with variational refinement and other global methods such as DeepFlow (Weinzaepfel et al., 2013) and PCAFlow (Wulff and Black, 2015). Based on these new findings, we decided to use the DIS method as a default method for the precipitation motion field calculation in the revised version of the rainymotion library. We also found that using the DIS method, our results show no significant difference between Dense and DenseRotation models. That confirms the strong influence of motion field estimation on the performance of the DenseRotation model.

We will update manuscript accordingly and show the intercomparison of of different optical flow methods in the supplementary material.

Kroeger, T., Timofte, R., Dai, D., & Van Gool, L. (2016, October). Fast optical flow using dense inverse search. In *European Conference on Computer Vision* (pp. 471-488). Springer, Cham.
Weinzaepfel, P., Revaud, J., Harchaoui, Z., & Schmid, C. (2013). DeepFlow: Large displacement optical flow with deep matching. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1385-1392).
Wulff, J., & Black, M. J. (2015). Efficient sparse-to-dense optical flow estimation using a learned basis and layers. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 120-130).

ACTION: We have updated Section 2 of the revised manuscript and the corresponding section in the Supplementary Information with the new results of benchmarking different global optical flow algorithms and the modified local Farnebäck algorithm. After introducing various changes (in terms of the tracking step, but also with regard to the implementation of the Dense models' extrapolation step, the Dense and the DenseRotation models provide, in effect, the same skill for the selected events and over the lead time of one hour. We also discuss the corresponding differences between implemented advection schemes in Section 5.2 (Advection schemes properties and effectiveness).

13. 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.

RESPONSE: Please see response to comment #11.

ACTION: Please see action to comment #11.

14. Page 6, line 10. "programmatic realization" is a strange expression.

RESPONSE: We will rephrase to "In this study we used the RV product data as an operational baseline and did not re-implement the underlying algorithm itself."

ACTION: We have revised the corresponding statement in accordance with the statement proposed in the response above (Page 7, lines 20-21).

15. Page 6, line 31. "rainfall depth product". Is it the instantaneous intensity in mm/hr or an accumulation?

RESPONSE: The RY product represents rainfall depth in mm for a five minute interval which is however derived from an instantaneous intensity considered representative for that interval.

ACTION: No specific action is needed.

16. 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.

RESPONSE: We will update Figure 6 to represent the CSI for the threshold of 1 mm/h and Figure 7 to represent the CSI for the threshold of 5 mm/h.

ACTION: Figure 7 in the revised version of manuscript represents the CSI for a threshold of 1 mm/h, and Figure 8 represents the CSI for a threshold of 5 mm/h.

17. 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.

RESPONSE: We agree. In fact, users can use the library to implement such ideas.

ACTION: No specific action is needed.

18. 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.

RESPONSE: We will try to emphasize this point in the revised version of the manuscript.

ACTION: We have updated the corresponding paragraph of the Section 6 (Summary and conclusions) to clarify the point of communicating inherent nowcasting uncertainties (Page 11, lines 26-29):
"[...] Admittedly, deterministic nowcasts in a Lagrangian framework do neither account for precipitation intensity dynamics nor for the uncertainties in representing precipitation field motion. At least for the latter, the rainymotion library provides ample opportunities to experiment with forecast ensembles, based on various tracking and extrapolation techniques [...]"

19. 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).

RESPONSE: We thank referee for this comment. More generally, future research should analyse in more detail which steps in our Sparse model chain contribute the most uncertainty. We still think the combination of Sparse optical flow and warping is very efficient and promising, but should be understood better.

ACTION: No specific action is needed.

20. Figures 1 and 2. These are extremely clean and nice presentations of the methods.

RESPONSE: We thank referee for this comment.

ACTION: No specific action is needed.

21. Figure 3. You may add in the caption that the figure shows the forward-in-time semi-Lagrangian method.

RESPONSE: Figure 3 caption will be updated accordingly.

ACTION: We have updated Figure 3 of the revised manuscript to illustrate the four implemented advection methods (forward/backward constant vector, forward/backward semi-Lagrangian).

22. 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".

RESPONSE: Figure 5 caption will be updated accordingly.

ACTION: We have updated the captions of Figures 6-8 in the revised manuscript accordingly.

23. 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.

RESPONSE: We briefly described the possible reasons of this RADVOR behavior on Page 7 lines 14-15 and Page 8 lines 3-8. In our opinion, the use of smoothed displacement fields that focus on a large scale motion patterns particularly cause a loss of skill in the RV product for the first 5-10 minutes. We will update the corresponding paragraph of Section 5.1 (Model comparison) to make that point clearer.

ACTION: We have updated Section 5.1 according to the new results.

24. 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.

RESPONSE: At the moment of paper submission (6th July), no reference to pySTEPS was known to us (first commit on GitHub from 9th July). We will update the introductory section to add a reference to pySTEPS, but we will also include the perspectives mentioned by the referee in the "Summary and conclusions".

ACTION: We have updated Section 6 (Summary and conclusions) to highlight PySTEPS and the importance and perspectives of open source software for advancing the field of radar science as follows (Page 12, lines 5-10):
“[...] Recent studies show that open source community-driven software advances the field of weather radar science (Heistermann et al., 2015a, b). Just a few months ago, the pySTEPS (https://pysteps.github.io) initiative was introduced "to develop and maintain an easy to use, modular, free and open source python framework for short-term ensemble prediction systems." As another evidence of the dynamic evolution of QPN research over the recent years, these developments could pave the way for future synergies between the pySTEPS and rainymotion projects -- towards the availability of open, reproducible, and skillful methods in quantitative precipitation nowcasting. [...]”

Referee comment #2 (by Seppo Pulkkinen)

Relation to previous work and literature review

1. There are two important classes of optical flow methods that are only briefly mentioned or not mentioned at all:
    1.1. In the variational methods, a smoothness constraint is added to the optical flow equations and they are solved "globally" over the whole domain. The key practical difference to the "local" methods, such as Farnebäck and Lucas-Kanade is that the motion field is automatically filled to areas of no precipitation.
    1.2. In the spectral methods, the Fourier transform is applied to the inputs and the optical flow equations are solved in the spectral domain. The authors could add a citation to [3].

[3] E. Ruzanski, V. Chandrasekar and Y. Wang, The CASA Nowcasting System, Journal of Atmospheric and Oceanic Technology, 28(5), 640-655, 2011.

RESPONSE: We thank referee for the clarification. We will add the corresponding methods and references to the introduction of the revised version of the manuscript, particularly since we added a variational approach and several global methods to the rainymotion library and our benchmark experiment - please see our response to comment #12 of Dr. Foresti.

ACTION: We have updated the corresponding paragraph of the introductory section (Page 2, lines 29-31) to account for the aforementioned group of optical flow methods:
"[...] There is also a distinct group of spectral methods where the Fourier transform is applied to the inputs, and an OFC resolves in the spectral (Fourier) domain (Ruzanski et al., 2011). [...]"

2. There are several widely used optical flow algorithms developed in the machine vision literature. The authors could cite the Brox and CLG algorithms ([1] and [2]). These have also publicly available C implementations (see the IPOL journal).

[1] T. Brox, A. Bruhn, N. Papenberg and J. Weickert, High Accuracy Optical Flow Estimation Based on a Theory for Warping, ECCV 2004: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part IV, 25-35, 2004.
[2] A. Bruhn, J. Weickert and C. Schnörr, Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Methods, International Journal of Computer Vision, 61(3), 211-231, 2005.

RESPONSE: We will add those references to the paragraph where we mention only openCV as an open software library with optical flow algorithms implementation (Page 3 line 1 of the discussion paper).

ACTION: The references have been added to the revised version of the manuscript as follows (Page 3, lines 7-10):
"[...] That is all the more surprising since open source implementations of fundamental optical flow algorithms (Brox et al., 2004; Bruhn et al., 2005b) have been around for up to 20 years – with the OpenCV library (https://opencv.org) just being the most widely known. Such libraries provide efficient implementations of various optical flow algorithms for a vast number of research and application contexts [...]"

3. The paper cites to a large number of references where more advanced probabilistic nowcasting methods are described. Therefore, in the third paragraph of Section 6 the authors should be more concrete about future plans to include such features into rainymotion, and not just present ideas of potential improvements. Or will rainymotion be restricted only to deterministic extrapolation nowcasting based on Lagrangian persistence?

RESPONSE: We would like to thank the referee for this suggestion, yet we are hesitant whether more detailed perspectives on future developments should be elaborated in the paper. Based on its current design, rainymotion's focus is to track the motion of rainfields and to extrapolate future rainfall on that basis. At the same time, the rather low-level implementation easily allows for the flexibility to manipulate the displaced precipitation fields in order to represent -- stochastically or deterministically -- the dynamics of precipitation intensity. Yet, there are no specific plans to implement such features to the rainymotion, but we will, in the revised version of the manuscript, highlight more explicitly the possibility to include such developments.

ACTION: We have updated Section 6 (Summary and conclusions) of the revised manuscript with the elaboration of using the rainymotion for providing ensemble nowcasts as follows (Page 11, lines 28-29):
"[...] At least for the latter, the rainymotion library provides ample opportunities to experiment with forecast ensembles, based on various tracking and extrapolation techniques [...]"

Methodology

4. Precisely speaking the Farnebäck optical flow algorithm is not global or dense and should not be called such. This misuse of terminology originates from the OpenCV library.
    4.1. The Farnebäck method is dense only in the sense that it produces gridded output instead of motion vectors for sparse feature points as Lucas-Kanade does. If you look at the paper of Farnebäck, the method is formulated as local feature matching, where the solution of the optical flow equations is done by using a polynomial approximation. As a result, the method produces zero motion velocities to areas of no precipitation. You can verify this by plotting motion fields produced by the Farnebäck method.

RESPONSE: We thank the referee for pointing that misuse in terminology which we have not been aware of so far. In our paper we use the term "local" and "sparse" in the sense that these methods provide motion vectors at specific locations only. In contrast, we use "global" and "dense" for pointing out that the motion vectors are calculated for an every radar image pixel. We will revise the paper in a way that terminology is both accurate and easy to understand. As a possible solution we propose to change "local" and "global" to "sparse" and "dense" in the revised version of the manuscript and provide a more detailed description of what we consider as "sparse" and "dense" models in the Section 2 (Model). Furthermore, we have actually added global optical flow techniques to the set of tracking models (please see our response to comment #12 of Loris Foresti), so the revised manuscript version will explicitly address the issue of global vs. local optical flow.

ACTION: We have revised the manuscript regarding used terminology based on provided suggestions. Since various truly global optical flow techniques have been incorporated into the rainymotion library and then extensively benchmarked (please see our response to comment #12 of Loris Foresti), we explicitly address the issue of global vs. local optical flow in the revised version of the manuscript.

4.2. It follows from the above that when a pixel is advected into area of no precipitation and a new motion vector is taken at that location (as in the DenseRotation method), it's motion to stops at the boundary. This could explain why Dense has in many cases better performance than DenseRotation.

RESPONSE: At the time of submission of this manuscript and thus when the results for this paper had been produced, the implementation of the extrapolation algorithm of the dense optical flow models did in fact not account for the case that pixels are advected into regions of zero velocities. In the meantime, however, we have revised the algorithm so that zero velocities are discarded and replaced by interpolation, and the results will be updated accordingly. Yet, the hypothesis that the insufficient treatment of zero velocities was responsible for the Dense model outperforming the DenseRotation model could not yet be corroborated (please also refer to our response to comment #12 of Dr. Foresti).

ACTION: The manuscript and the rainymotion library were substantially revised, particularly regarding the advection schemes for the Dense group of models, based on the corresponding reviewer suggestions. Please also see action to comment #12 of Dr. Foresti.

5. In Germann and Zawadzki (2002), the authors conclude that the backward semi-Lagrangian has better performance than the forward method. In fact, a majority of existing nowcasting methods use the former that is widely regarded as the best approach. However, here the authors use only the latter. If possible, the authors could also implement the backward method and include it in the performance comparison.

RESPONSE: Dr. Foresti in his comment #9 also raised this issue. Please refer to our answer there.

ACTION: Please see action to comment #9 of Dr. Foresti.

6. Using the backward method would require filling the gaps in the motion field on areas of no precipitation. Otherwise, no precipitation would be advected into areas where it does not exist at the nowcast start time. A simple distance-weighted interpolation should be sufficient for this purpose. For the above reason, using gap-filling would also improve the performance of the forward semi-Lagrangian method.

RESPONSE: In the revised version of the rainymotion library we implemented the referee's suggestion of "[...] filling the gaps in the motion field on areas of no precipitation" by utilizing inverse distance weighted interpolation to fill zero-gaps in the motion field. However, the benefit of this implementation on the performance of the forward semi-Lagrangian method (the DenseRotation model) is not so distinct probably because of the reasons we highlighted in the response on the comment #12 from Dr. Foresti (motion field estimation errors by the Farnebäck algorithm and additional interpolation).

ACTION: The manuscript and the rainymotion library have been revised in accordance with the above response.

7. Note that the gap-filling is automatically done in the variational methods without the need for separate post-processing of the motion field. Therefore, such methods are truly dense and global. The authors could consider implementing a variational method and include it in the performance comparison.

RESPONSE: We thank the referee for his recommendations regarding the implementation of variational optical flow models in the rainymotion library. We incorporated global optical flow methods which are available in opencv library as additional options for motion field calculation in the rainymotion library (see also our response to the comment #12 from Dr. Foresti), verified their skill for nowcasting and have to conclude that using more advanced global optical flow methods advances an efficiency of a semi-Lagrangian advection scheme. Based on the new obtained results we decided to replace the Farnebäck method by the global Dense Inverse Search (DIS, Kroeger et al., 2016) as a default tracking option. We will also update the supplementary material with intercomparison results of different optical flow methods.
Kroeger, T., Timofte, R., Dai, D., & Van Gool, L. (2016, October). Fast optical flow using dense inverse search. In *European Conference on Computer Vision* (pp. 471-488). Springer, Cham.

ACTION: The manuscript has been revised according to the above response.

Software library

8. Sections 2.4 and 3: Is the library restricted only to using the DWD data? Please add discussion about how to use the library with other file formats? For instance, by using wradlib this should be easily done because it supports a large number of different formats.

RESPONSE: There is no restriction in using different data formats because of rainymotion works directly with numpy arrays, and the data preprocessing routine is fully on the user-side. There is a set of available open software libraries for radar data reading and preprocessing (the list available on https://openradarscience.org/). We will add the corresponding information to the Section 2.4 (The rainymotion Python library).

ACTION: We have updated Section 2.4 (The rainymotion Python library) in accordance with provided response as follows (Page 6, lines 28-31):
"[...] Since the rainymotion uses standard format of numpy arrays for data manipulation, there is no restriction in using different data formats which can be read, transformed, and converted to numpy arrays using any tool from the set of available open software libraries for radar data manipulation (the list is available on https://openradarscience.org). [...]"

Verification

9. Section 2.6: MAE could be computed conditionally over those pixels where both the nowcast and the verifying observation exceed the detection threshold. Otherwise, there would be overlap with the CSI statistic as both penalize incorrect forecasts of precipitation/no precipitation.

RESPONSE: In our study we decided to use MAE as a score from a continuous category and implement it directly without making specific thresholds (like we do for categorical category of verification scores). In our opinion, this admittedly arbitrary decision of using different verification score categories helps to represent a diversity of obtained results.

ACTION: No specific action is needed.

10. A large number of CSI and MAE statistics are shown for different lead times. There could be more analysis of the results.
    10.1. There is no indication about what can be considered as a good CSI or MAE value for the nowcast to be usable. Can you give some thresholds?

RESPONSE: At the best of our knowledge, there is no convention regarding what to consider as "good" or "bad" for any verification metric commonly used in radar-based QPN. For our benchmarking experiment, the focus is on the differences of scores between the different models, not on their absolute values.

ACTION: No specific action is needed.

    10.2. The differences between the methods (excluding Persistence) are relatively small in terms of CSI and MAE statistics. Based on such differences, the authors should be more careful when claiming that some method is better than another. For instance the maximum mean difference between Dense and DenseRotation is only 0.01 according to Table 3.

RESPONSE: Table 3 represents statistics which are averaged over all the analyzed events and two lead time periods (5--30, and 35--60) and primarily highlight the difference between the Dense group of rainymotion models (Dense and DenseRotation) and and the RV product (as mentioned on the Page 7, lines 27--31) -- which is more distinct than the difference between Dense and DenseRotation models themselves. For the verification procedure we also carried out the Student's independent two-sample *t*-test to find whether differences between mean CSI and MAE values for the specific lead times are significant or not (not shown in the manuscript). We found that the results of the visual inspection of the verification plots are well consistent with the formal statistical evaluation: if there is a clear difference in the plots, it is typically significant in a statistical sense.
We will update Table 3 with the new results and adjust our statements about considering one model better/worse than another correspondingly.

ACTION: Table 3 of the revised manuscript has been updated with the new verification results.

11. Figures 5-7 and p. 9, lines 19-21. The authors should indeed take a closer look on why the performance of the sparse methods is poor. Some comments about this:
    11.1. The relevant parameter here is the number of features used in the tracking and nowcasting. If this number is too small, the motion vectors of the features are not representative of the large-scale motion field. Can you check this by adjusting the thresholds in the feature detector?
    11.2. In addition, can you specify somewhere how many feature points are used with the sparse methods because this is a key parameter?
    11.3. Another point missed in the paper is that the corner detector tends to pick features that have high intensities and gradients. Therefore, a very careful quality control is needed to ensure that the features are precipitation and not some

random artefacts in the radar data. Can you be sure that the quality control is sufficient?

11.4. Even if the features are precipitation, they represent small-scale phenomena that can have very different motion from the large-scale advection field. Thus, the representativity of such features can be very poor.

RESPONSE: We agree with the referee that the sensitivity of the Sparse group of models to specific key parameters needs to be investigated more closely. Yet, we consider such an analysis beyond the scope of this study. Another study is underway that specifically and systematically focuses on the error of the forecast location of detected features based on a vast set of tracking and extrapolation techniques, and including different parameterisations as mentioned by the referee (such as the maximum number of features detected, or different approaches to filter spurious or non-representative velocities at small spatiotemporal scales). In the present manuscript under discussion, however, our aim is to present two basic and open architectures of nowcasting models based on optical flow which can serve as a baseline for future developments - as part of the rainymotion library itself or in combination with the library, and to demonstrate that these are skillful. Still, the parameters of the Shi-Tomasi corner detector provide us a possibility to control the maximum number of features, their quality (which is based on the minimal eigenvalue) and a minimum euclidean distance between the nearest identified points. A calibration of these parameters had been performed on different events and the most robust values had been set up as default parameters as follows: maximum number of features -- 200; quality level -- 0.2 (the corners with the quality measure less than the product of quality level and minimal eigenvalue will be rejected); minimum euclidean distance -- 7 pixels (the corners which have stronger neighbors in a neighborhood less than 7 pixels will be rejected). As for quality control of the actual radar data, we rely on the DWD's processing workflow that produces the RY product and which eliminates vast parts of spurious echoes. Yet, even in the presence of residual static or dynamic clutter, the tracking algorithm has proven to be robust against producing zero velocities.

ACTION: We have updated Section 5.1 (Model comparison) with the comparison of the Dense and Sparse group of the rainymotion in accordance with the reviewer suggestions as follows (Page 10, lines 5-11):
[...] Despite their skill over Eulerian persistence, the Sparse group models are significantly outperformed by the Dense group models for all the analyzed events and lead times. The reason for this behaviour remains yet unclear. It could, in general, be a combination of errors introduced in corner-tracking and extrapolation as well as image warping as a surrogate for formal advection. While the systematic identification of error sources will be subject to future studies, we suspect that the the local features ("corners") identified by the Shi-Tomasi corner detector might not be representative for the overall motion of the precipitation field: the detection focuses on features with high intensities and gradients, the motion of which might not represent the dominant meso-γ scale motion patterns. [...]

12. Forecasting the occurrence of precipitation/no precipitation for high intensities is highly relevant for practical applications. Therefore, I would suggest moving the results with the 5 mm/h threshold from the supplementary material to the main paper.

RESPONSE: We support referee's recommendation (see also comment #16 by Loris Foresti) and will transfer the corresponding figure from the supplementary to the main paper.

ACTION: The new Figure 8 represents the CSI for the threshold of 5 mm/h in the revised version of the manuscript.

Figures

13. Since the motion field determination plays a key role in the paper, the authors should show at least one figure with an observed precipitation field and the computed motion field plotted on the same figure. Even better would be a figure showing motion vectors of features and motion fields computed by using different methods.

RESPONSE: We agree with the referees' recommendation (see also comment #2 by Loris Foresti) and will add the requested figures to the revised version of the manuscript.

ACTION: The new figure (Figure 5) has been added to the revised version of the manuscript to show examples of nowcasts, velocity vectors of features and velocity fields for different models and lead times.

14. Figure 4: Are names of individual functions relevant here? Consider removing them.

RESPONSE: In our opinion, it is informative to show the key functions that we used from various libraries in order to put together the main functionality of rainymotion. It illustrates that the combination is, from a technical perspective, not too complex.

ACTION: We have updated Figure 4 in the revised manuscript based on the changes in the rainymotion library.

Minor details

15. Page 4, lines 3-6 and Figure 1. How exactly is the affine transformation matrix calculated. In particular, is a single matrix estimated for all features or is this done separately for each feature?

RESPONSE: The transformation matrix is calculated on the basis of all identified features. We will add this clarification to Section 2.1 (Local optical flow models).

ACTION: We have updated the corresponding statement in the revised manuscript (Page 4, lines 22-23).

16. Page 5, line 24. Why the HDF5 file format was chosen? Please add some justification for this.

RESPONSE: For all internal projects we use HDF5 database and corresponding file format as an efficient data storage with powerful set of archiving options (i.e. compression rate, chunk size) instead of using default binary files provided by the DWD. However, we propose to remove the reference to HDF5 file format and h5py library because of it is neither integral part of our analysis, nor the rainymotion library, but just a subjective choice we made regarding our research workflow. We will update the Section 3 correspondingly.

ACTION: The references to HDF5 file format and h5py library have been removed from the revised version of the manuscript.

17. Page 9, lines 7-9. I don't understand what this means. Can you clarify?

RESPONSE: The statement “It might also be considered to combine the warping procedure for the extrapolation step with the Dense optical flow procedure for the tracking step in order to dramatically enhance computational performance” describes the idea to detect corners, then predict the future locations of these corners using the motion field from dense optical flow, and then construct the Affine Transformation Matrix for the warping based on the corner locations at forecast time and lead time t_n. That way, we would combine the robustness of the dense optical flow technique with the computational efficiency of the warping technique. We will clarify that idea in the revised manuscript.

ACTION: The statement under consideration has been updated as follows (Page 11, lines 20-22):
“[...] It might also be considered to combine the warping procedure for the extrapolation step with the Dense optical flow procedure for the tracking step (i.e. to advect "corners" based on a "Dense" velocity field obtained by implementing one of the dense optical flow techniques). [...]”

18. Page 9, line 24. Stochastic accounting <- stochastic modeling?

RESPONSE: We thank referee for pointing out that mistake which will be corrected in the revised version of the manuscript.

ACTION: The issue has been fixed in the revised version of the manuscript.

Referee comment #3 (by Remko Uijlenhoet)

1. References to important papers from Marc Berenguer, Daniel Sempere-Torres and Geoff Pegram are missing (SBMcast, etc.). These are very relevant papers in the context of this manuscript, which discuss the issue of spectral decomposition of precipitation fields and scale-dependent radar nowcasting.
2. Reference to Berne et al. (2004; JoH) is missing. This is a (by now) classical paper on space-time scales of rainfall fields required for (urban) hydrological applications.
3. Reference to pySTEPS appears to be missing (https://github.com/pySTEPS). This is the open source Python version of STEPS. Highly relevant given the topic and focus of this manuscript.

RESPONSE: We will include the suggested references in the introductory section. As for the missing reference to pySTEPS, we refer to our response to comment #24 of Loris Foresti.

ACTION:

1. Reference to the relevant paper in the context of an ensemble and scale-dependent radar-based precipitation nowcasting -- Berenguer et al., 2011 -- has been added in the Introductory Section.
    - Page 2, lines 9-12: *"[...] In addition, we can classify nowcasting methods based on how predictive uncertainty is accounted for: In contrast to deterministic approaches, ensemble nowcasts attempt to account for predictive uncertainty by including different realizations of the motion field and the evolution of rainfall intensity itself (Berenguer et al., 2011). [...]"*
    - Page 2, lines 19-21: *"[...] Different algorithms can be used for each step, tracking and forecasting, in order to compute an ensemble forecast (Berenguer et al., 2011; Grecu and Krajewski, 2000; Foresti et al., 2016). [...]"*
2. As, in the present study, we do not have the aim to describe the importance of radar-based precipitation nowcasting in closely related fields (i.e. hydrological forecasting), we decided not to refer to the paper by Berne et al., 2004.
3. A reference to the pySTEPS project has been added in Section 6 (Summary and conclusions).

4. Please provide some more detailed background information concerning: Shi–Tomasi corner detector (Shi and Tomasi, 1994); Lucas–Kanade optical flow algorithm (Lucas and Kanade, 1981); affine transformation matrix (Schneider and Eberly, 2003); warping and interpolation (Wolberg, 1990).

RESPONSE: We will try to illustrate in more detail the main features of these techniques in the revised version of the manuscript.

ACTION: We have provided more detailed description of methods used in Sparse group models as follows (Section 2.1; Page 4, lines 10-30):

"[...] The first model (SparseSD, for Sparse Single Delta) uses only the two most recent radar images for identifying, tracking, and extrapolating features. Assuming that t denotes both the nowcast issue time and the time of the most recent radar image, the implementation can be summarized as follows:

1. Identify features in a radar image at time t-1 using the Shi–Tomasi corner detector (Shi and Tomasi, 1994). This detector determines the most prominent corners in the image based on the calculation of the corner quality measure (min(λ1, λ2), where λ1 and λ2 are corresponding eigenvalues) at every image pixel (see Section S1 of the Supplementary Information for detailed description of algorithm parameters);
2. Track these features at time t using the local Lucas–Kanade optical flow algorithm (Lucas and Kanade, 1981). This algorithm tries to identify the location of feature we previously identified on the radar image at time t-1 on the radar image at time t based on the solving a set of optical flow equations in the local feature neighborhood using the least-squares approach (see Section S1 of the Supplementary Information for detailed description of algorithm parameters);
3. Linearly extrapolate the features' motion in order to predict the features' locations at each lead time n;
4. Calculate the affine transformation matrix for each lead time n based on the locations of all identified features at time t and t+n using the least-squares approach (Schneider and Eberly, 2003). This matrix uniquely identifies the required transformation of the last observed radar image at time t so that the nowcast images at times t+1...t+n provide the smallest possible difference between the locations of detected features at time t and the extrapolated features at times t+1...t+n;
5. Extrapolate the radar image at time t by warping: for each lead time, the warping procedure uniquely transforms each pixel location of the radar image at time t to its future location in the nowcast radar images at times t+1...t+n, using the affine transformation matrix. Remaining discontinuities in the predicted image are linearly interpolated in order to obtain nowcast intensities on a grid that corresponds to the radar image at time t (Wolberg, 1990) [...]"

5. Corrections, grammar and typos
    5.1. "Supplementary" –> "Supplementary Information" (several times in the manuscript).
    5.2. P.4, l.8: "24 recent radar images" –> "24 most recent radar images".
    5.3. P.5, l.20: "models' description" –> "model description".
    5.4. P.6, l.25–26: "rainfall rates prediction" –> "rainfall rate prediction".
    5.5. P.8, l.6: Insert comma before "which".

RESPONSE: Will be fixed.

ACTION: The corresponding issues have been fixed.

6. Is "RV" the same ad "RadVor"?

RESPONSE: RADVOR is the entire nowcasting workflow used by the DWD. RV is a main product along that processing chain which is the forecast precipitation depth in five minute intervals over a lead time of two hours. The official main product of RADVOR, though, is the RQ product which is the precipitation depth accumulated over an interval of one hour for a lead time of two hours. It is basically obtained from the RV product, but includes an additional adjustment of the distribution function. In summary, the RV product is the part of DWD's nowcasting chain that is best comparable to our nowcasting products and the best "end product" that is available at an interval of five minutes.
ACTION: No specific action is needed.

7. General: (much) more detailed captions; figures + captions should be as self-contained as possible.

RESPONSE: We will update the figure captions to make them more self-contained.

ACTION: We have updated captions for figures 6-8 according to comment #22 of Dr. Loresti.

8. Journal (Nature), issue, page numbers missing from reference to Bauer et al. (2015).

RESPONSE: We will update the corresponding reference to the Bauer et al. (2015) paper as following:
Bauer, P., Thorpe, A., Brunet G.: The quiet revolution of numerical weather prediction, Nature, 525, 47–55, https://doi.org/10.1038/nature14956, https://www.nature.com/articles/nature14956, 2015.

ACTION: The reference to Bauer et al. (2015) has been updated.

Other changes made in the manuscript

1. All the figures and tables (except Table 1) in the manuscript and the Supplementary Information have been updated. Additionally, new sections of the Supplementary information (Sections S5, S6) have been added based on the new results which have been obtained using the revised version of the rainymotion library.
2. Subsections of Section 2 (Models) have been updated to fit the proposed substitution of terminology from local/global optical flow to sparse/dense.
3. The rainymotion library source code and documentation have been substantially updated (https://github.com/hydrogo/rainymotion/commits/v0.1).

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

Georgy Ayzel[1], Maik Heistermann[1], and Tanja Winterrath[2]

[1]Institute for Environmental Sciences and Geography, University of Potsdam, Potsdam, Germany
[2]Deutscher Wetterdienst, Department of Hydrometeorology, Offenbach, Germany

Correspondence: Georgy Ayzel (ayzel@uni-potsdam.de)

Abstract. Quantitative precipitation nowcasting (QPN) has become an essential technique in various application contexts, such as early warning or urban sewage control. A common heuristic prediction approach is to track the motion of precipitation features from a sequence of weather radar images, and then to ~~extrapolate that motion~~ displace the precipitation field to the imminent future (minutes to hours) based on that motion, assuming that the intensity of the features remains constant ("Lagrangian persistence"). In that context, "optical flow" has become one of the most popular tracking techniques. Yet, the present landscape of computational QPN models still struggles with producing open software implementations. Focusing on this gap, we have developed and extensively benchmarked a stack of models based on different optical flow algorithms for the tracking step, and a set of parsimonious extrapolation procedures based on image warping and advection. We demonstrate that these models provide skillful predictions comparable with or even superior to state-of-the-art operational software. Our software library ("rainymotion") for precipitation nowcasting is written in the Python programming language, and openly available at GitHub (https://github.com/hydrogo/rainymotion). That way, the library may serve as a tool for providing fast, free and transparent solutions that could serve as a benchmark for further model development and hypothesis testing – a benchmark that is far more advanced than the conventional benchmark of Eulerian persistence commonly used in QPN verification experiments.

1 Introduction

How much will it rain within the next hour? The term "quantitative precipitation nowcasting" refers to forecasts at high spatio-temporal resolution (60-600 seconds, 100-1000 meters) and short lead times of only a few hours. Nowcasts have become important for broad levels of the population for planning various kinds of activities. Yet, they are particularly relevant in the context of early warning of heavy convective rainfall events, and their corresponding impacts such as flash floods, landslides, or sewage overflow in urban areas.

While recent advances in numerical weather prediction (NWP) allow us to forecast atmospheric dynamics at very high resolution (Bauer et al., 2015), computational costs are typically prohibitive for the requirements of operational nowcasting applications with frequent update cycles. Furthermore, the heuristic extrapolation of rain field motion and development, as observed by weather radar, still appears to outperform NWP forecasts at very short lead times. Today, many precipitation nowcasting systems are operational at regional or national scales, utilizing various radar products, algorithms, and blending techniques in order to provide forecasts up to 1-3 hours: ANC (Mueller et al., 2003), MAPLE (Germann and Zawadzki, 2002), RADVOR (Winterrath et al., 2012), STEPS (Bowler et al., 2006), STEPS-BE (Foresti et al., 2016), and SWIRLS (Cheung and Yeung, 2012; Woo and Wong, 2017). For an extensive review of existing operational systems, please refer to Reyniers (2008).

~~There are three main groups of techniques for~~ A variety of radar-based precipitation nowcasting ~~: analog, local Lagrangian, and stochastic (Foresti et al., 2016). Analog-based methods utilize a measure of similarity (e.g., correlation coefficient)to find the most similar sequences in archived radar observations in order to construct a precipitation forecast (Foresti et al., 2015). Stochastic nowcasts add random perturbations to deterministic ones to account for the uncertainty (Bowler et al., 2006; Foresti et al., 2016). In this study, we focus on the second~~ techniques can be classified into three major groups based on assumptions we make regarding precipitation field characteristics (Germann and Zawadzki, 2002). The first group – climatological persistence – provides nowcasts by using climatological values (mean or median). The second group – Eulerian persistence – is based on using the latest available observation as a prediction, and is thus independent from the forecast lead time. The third group – ~~local Lagrangian~~ Lagrangian persistence – ~~which~~ allows the extrapolation of the most recent ~~radar images~~ observed precipitation field under the assumption that ~~the velocity field is~~ intensity of precipitation features and the motion field are persistent (Germann and Zawadzki, 2002; Woo and Wong, 2017). In addition, we can classify nowcasting methods based on how predictive uncertainty is accounted for: In contrast to deterministic approaches, ensemble nowcasts attempt to account for predictive uncertainty by including different realizations of the motion field and the evolution of rainfall intensity itself (Berenguer et al., 2011). In this study, we focus our model development around the group of Lagrangian persistence models which provide deterministic precipitation nowcasts.

~~Local~~ Lagrangian methods consist of two computational steps: tracking and forecasting (extrapolation) (Austin and Bellon, 1974). In the tracking step, we compute a velocity field from a series of consecutive radar images, either on a per pixel basis (Germann and Zawadzki, 2002; Grecu and Krajewski, 2000; Liu et al., 2015; Zahraei et al., 2012), or for contiguous objects (Zahraei et al., 2013). In the second step, we use that velocity field to advect the most recent rain field, i.e. to ~~extrapolate its motion into~~ displace it to the imminent future based on its observed motion. That step has been implemented based on semi-Lagrangian schemes (Germann and Zawadzki, 2002), interpolation procedures (Liu et al., 2015), or mesh-based models (Bellerby, 2006; Zahraei et al., 2012). Different algorithms can be used for each step, tracking and forecasting, in order to compute an ensemble forecast ~~(Grecu and Krajewski, 2000; Foresti et al., 2016)~~(Berenguer et al., 2011; Foresti et al., 2016; Grecu and Krajew
.

One of the most prominent techniques for the tracking step is referred to as "optical flow". The original term was inspired by the idea of an *apparent* motion of brightness patterns observed when a camera or the eyeball is moving relative to the objects (Horn and Schunck, 1981). Today, optical flow is often understood as a group of techniques to infer motion patterns or velocity fields from consecutive image frames, e.g. in the field of precipitation nowcasting (Bowler et al., 2004; Liu et al., 2015; Woo and Wong, 2017). For the velocity field estimation, we need to accept both the brightness constancy assumption and one of a set of additional optical flow constraints (OFC). The spatial attribution of OFC marks the two main categories of optical flow models: local (differential) and global (variational) (Cheung and Yeung, 2012; Liu et al., 2015). Local models try to set an OFC only in some neighborhood, while global models apply an OFC for a whole image. There is also a distinct group of spectral methods where the Fourier transform is applied to the inputs, and an OFC resolves in the spectral (Fourier) domain (Ruzanski et al., 2011). Bowler et al. (2004) introduced the first local optical flow algorithm for precipitation nowcasting, and gave rise to a new direction of models. Bowler's algorithm is the basis of the STEPS (Bowler et al., 2006) and STEPS-BE (Foresti et al., 2016) operational nowcasting systems. Liu et al. (2015) proposed using a local Lucas–Kanade optical flow method (Lucas and Kanade, 1981) independently for ~~every~~ each pixel of satellite imagery and compared its performance with a global Horn–Schunck (Horn and Schunck, 1981) optical flow algorithm. Yeung et al. (2009), Cheung and Yeung (2012), and Woo and Wong (2017) used different global optical flow algorithms (Bruhn et al., 2005a; Wong et al., 2009) for establishing the SWIRLS product for operational nowcasting in Hong-Kong.

Hence, for around two decades, optical flow algorithms have been doing their best for state-of-the-art operational nowcasting systems around the globe. Should research still care about them? It should… and the reason is that – despite the abundance of publications about different flavours of optical flow techniques for nowcasting applications – an open and transparent benchmark model is yet not available, except for the most trivial one: Eulerian persistence.

That is all the more surprising since open source ~~libraries such as OpenCV ()~~ implementations of fundamental optical flow algorithms (Brox et al., 2004; Bruhn et al., 2005b) have been around for ~~almost~~ up to 20 years ~~, providing~~ – with the OpenCV library (https://opencv.org) just being the most widely known. Such libraries provide efficient implementations of various optical flow algorithms for ~~an endless~~ a vast number of research and application contexts. Yet, none can be applied in the QPN context out of the box – without the need to address additional and specific challenges such as underlying assumptions and constraints of velocity fields, pre- and postprocessing steps, or model parameterization and verification.

The aim of this paper is thus to establish a set of benchmark procedures for quantitative precipitation nowcasting as an alternative to the trivial case of Eulerian persistence. This study does not aim to improve the standard of precipitation nowcasting beyond the state-of-the-art, but to provide an open, transparent, reproducible and easy-to-use approach that can compete with the state-of-the-art, and against which future advances can be measured. To that end, we developed a group of models that are based on two optical flow formulations of algorithms for the tracking step – ~~local (Lucas and Kanade, 1981) and global (Farnebäck, 2003)~~ sparse (Lucas and Kanade, 1981) and dense (Kroeger et al., 2016) – together with two parsimonious extrapolation techniques based on image warping and spatial interpolation. These models are verified against Eulerian persistence, as a trivial benchmark, and against the operational nowcasting system of the Deutscher Wetterdienst (the German Weather Service, DWD), as a representative of state-of-the-art models. The different optical flow implementations are published as an open source Python library (*rainymotion*, https://github.com/hydrogo/rainymotion) that entirely relies on free and open source dependencies, including detailed documentation and example workflows (https://rainymotion.readthedocs.io).

The paper is organized as follows. In Section 2, we describe the algorithmic and technical aspects of the suggested optical flow models. Section 3 describes the data we used, and provides a short synopsis of events we used for the benchmark experiment. We report the results in Section 4, and discuss them in various contexts in Section 5. Section 6 provides summary and conclusions.

2 Models

The benchmark models developed in this study consist of different combinations of algorithms for the two major steps of Lagrangian nowcasting frameworks, namely tracking and extrapolation (Austin and Bellon, 1974). Table 1 provides an overview of the models. The values of model parameters adopted in the benchmark experiment have been heuristically determined and not yet been subject to systematic optimization. However, the *rainymotion* library provides an opportunity to investigate how different optical flow model parameters can affect nowcasting results, or how they can be tuned to represent, e.g., the typical range of advection speeds of real precipitation fields. For a description of parameters, please refer to Section S1 in the Supplementary Information or the *rainymotion* library documentation (https://rainymotion.readthedocs.io/).

2.1 ~~Local optical flow models (the~~ The Sparse group~~)~~

The central idea around this group of methods is to identify distinct features in a radar image that are suitable for tracking. In this context, a "feature" is defined as a distinct point ("corner") with a sharp gradient of rainfall intensity. That approach is less arbitrary and scale dependent and thus more universal than classical approaches that track storm cells as contiguous objects (e.g., Wilson et al., 1998) because it eliminates the need to specify arbitrary and scale dependent characteristics of "precipitation features" while the identification of "corners" depends only on the gradient sharpness in a cell's neighborhood. Inside this group, we developed two models that slightly differ with regard to both tracking and extrapolation.

The first model (SparseSD, for Sparse Single Delta) uses only the two most recent radar images for identifying, tracking, and extrapolating features. Assuming that *t* denotes both the nowcast issue time and the time of the most recent radar image, the implementation can be summarized as follows:

1. Identify features in a radar image at time *t-1* using the Shi–Tomasi corner detector (Shi and Tomasi, 1994). This detector determines the most prominent corners in the image based on the calculation of the corner quality measure ($min(\lambda_1, \lambda_2)$, where λ_1 and λ_2 are corresponding eigenvalues) at each image pixel (see Section S1 of the Supplementary Information for a detailed description of algorithm parameters);

2. Track these features at time *t* using the local Lucas–Kanade optical flow algorithm (Lucas and Kanade, 1981). This algorithm tries to identify the location of a feature we previously identified at time *t-1* in the radar image at time *t*, based on solving a set of optical flow equations in the local feature neighborhood using the least-squares approach (see Section S1 of the Supplementary Information for a detailed description of algorithm parameters);

3. Linearly extrapolate the features' motion in order to predict the features' locations at each lead time *n*;

4. Calculate the affine transformation matrix ~~(Schneider and Eberly, 2003)~~ for each lead time *n* based on the ~~features' locations~~ locations of all identified features at time *t* and *t+n* ~~;~~ using the least-squares approach (Schneider and Eberly, 2003) . This matrix uniquely identifies the required transformation of the last observed radar image at time *t* so that the nowcast images at times *t+1...t+n* provide the smallest possible difference between the locations of detected features at time *t* and the extrapolated features at times *t+1...t+n*;

5. ~~Warp~~ Extrapolate the radar image at time *t* by warping: for each lead time, the warping procedure uniquely transforms each pixel location of the radar image at time ~~*nusing the corresponding affine matrix, and linearly interpolate remaining discontinuities (Wolberg, 1990).*~~ *t* to its future location in the nowcast radar images at times *t+1*...*t+n*, using the affine transformation matrix. Remaining discontinuities in the predicted image are linearly interpolated in order to obtain nowcast intensities on a grid that corresponds to the radar image at time *t* (Wolberg, 1990).

To our knowledge, this study is the first to apply image warping directly as a simple and fast algorithm to represent advective motion of a precipitation field. In Section S2 of the Supplementary Information, you can find a simple synthetic example which shows the potential of the warping technique to replace an explicit advection formulation for temporal extrapolation.

For a visual representation of the SparseSD model~~routine~~, please refer to Fig. 1.

The second model (Sparse) uses the 24 most recent radar images, and we consider ~~here~~ only features that are persistent over the whole period (of 24 ~~timesteps)for capturing the most steady movement. Its~~ time steps). The implementation can be summarized as follows:

1. Identify features on a radar image at time *t-23* using the Shi–Tomasi corner detector (Shi and Tomasi, 1994);

2. Track these features ~~on radar images at the time~~ in the radar images from *t-22* to *t* using the local Lucas–Kanade optical flow algorithm (Lucas and Kanade, 1981);

3. Build linear regression models which independently ~~parametrize~~ parameterize changes in coordinates through time (from *t-23* to *t*) for every successfully tracked feature;

4. Continue with steps 3-5 of SparseSD.

For a visual representation of the Sparse model~~routine~~, please refer to Fig. 2.

~~To our knowledge, this study is the first to apply image warping directly as a simple and fast algorithm to represent advective motion of a precipitation field. In Section S2 of the Supplementary, you can find a simple synthetic example which shows the potential of the warping technique to replace an explicit advection formulation for temporal extrapolation.~~

2.2 ~~Global optical flow models (the~~ The Dense group~~)~~

The Dense group of models uses~~the~~, by default, the Dense Inverse Search algorithm (DIS) – a global optical flow algorithm proposed by ~~Farnebäck (2003)~~ Kroeger et al. (2016) – which allows us to explicitly estimate the velocity of each image pixel based on an analysis of two consecutive radar images. The DIS algorithm was selected as the default optical flow method for motion field retrieval because it showed, in our benchmark experiments, a higher accuracy and also a higher computational efficiency in comparison with other global optical flow algorithms such as DeepFlow (Weinzaepfel et al., 2013), and PCAFlow (Wulff and Black, 2015). We also tested the local Farnebäck algorithm (Farnebäck, 2003), which we modified by replacing zero velocities by interpolation, and by smoothing the obtained velocity field based on a variational refinement procedure (Brox et al., 2004) (please refer to Section S5 in the Supplementary Information for verification results of the corresponding benchmark experiment with various dense optical flow models). However, the *rainymotion* library provides the option to choose any of the specified above optical flow methods for precipitation nowcasting.

The two models in this group differ only with regard to the extrapolation (or advection) step. The first model (Dense) uses a constant-vector advection scheme (Bowler et al., 2004), while the second model (DenseRotation) uses a ~~forward~~ semi-Lagrangian advection scheme (Germann and Zawadzki, 2002). The main difference between ~~the proposed~~ both approaches is that a constant-vector scheme does not allow for the representation of rotational motion (Bowler et al., 2004); a semi-Lagrangian scheme allows for the representation of large-scale rotational movement while assuming the motion field itself to be persistent (Fig. 3). ~~Both~~

There are two possible options of how both advection schemes may be implemented: forward in time (and downstream in space) or backward in time (and upstream in space) (Fig. 3). It is yet unclear which scheme can be considered as the most appropriate and universal solution for radar-based precipitation nowcasting, regarding the conservation of mass on the one hand and the attributed loss of power at small scales on the other hand (e.g., see discussion in Bowler et al., 2004; Germann and Zawadzki, 2002) . Thus, we conducted a benchmark experiment with any possible combination of forward vs. backward and constant-vector vs. semi-Lagrangian advection. Based on the results (see Section S6 in the Supplementary Information), we use the backward scheme as the default option for both the Dense and DenseRotation models. However, the *rainymotion* library still provides the option to use the forward scheme, too.

Both the Dense and DenseRotation models utilize a linear interpolation procedure in order to interpolate advected rainfall intensities at their predicted locations to the ~~original~~ native radar grid. ~~This interpolation procedure follows the same idea of distributing~~ The interpolation procedure "distributes" the value of a rain pixel to its neighborhood, as proposed in different modifications by Bowler et al. (2004), Liu et al. (2015), and Zahraei et al. (2012). The Dense group models' implementation can be summarized as follows:

1. Calculate a ~~continuous displacement field using a global Farnebäck~~ velocity field using the global DIS optical flow algorithm ~~(Farnebäck, 2003)~~ (Kroeger et al., 2016), based on the radar images at time *t-1* and *t*;

2. Use a backward constant-vector (Bowler et al., 2004) or a backward semi-Lagrangian scheme (Germann and Zawadzki, 2002) to extrapolate (advect) each pixel according to the ~~obtained~~ displacement (velocity) field, in one single step for each lead time *t+n*. For the semi-Lagrangian scheme, we update the velocity of ~~each displaced pixel~~ the displaced pixels at each prediction time step ~~by retrieving the velocity closest to the predicted pixellocation~~*n* by linear interpolation of the velocity field to a pixel's location at that time step;

3. As a result of the advection step, we basically obtain an irregular point cloud that consists of the original radar pixels displaced from their original location. We use the intensity of each displaced pixel at its predicted location at time *t+n* in order to interpolate the intensity at each grid point of the original (native) radar grid (Liu et al., 2015; Zahraei et al., 2012), using the inverse distance weighting interpolation technique. It is important to note that we minimize numerical diffusion by first advecting each pixel over the target lead time before applying the interpolation procedure (as in the

"interpolate once" approach proposed by Germann and Zawadzki (2002)). That way, we avoid rainfall features to be smoothed in space by the effects of interpolation.

2.3 Persistence

The (trivial) benchmark model of Eulerian persistence assumes that for any lead time n, the precipitation field is the same as for time t. ~~Considering~~ Despite its simplicity, it is quite a powerful predictor for very short lead times, and, at the same time, its verification performance is a good measure of temporal decorrelation for different events.

2.4 The *rainymotion* Python library

We have developed the *rainymotion* Python library ~~that implements~~ to implement the above models. Since the *rainymotion* uses standard format of *numpy* arrays for data manipulation, there is no restriction in using different data formats which can be read, transformed, and converted to *numpy* arrays using any tool from the set of available open software libraries for radar data manipulation (the list is available on https://openradarscience.org). The source code is available in a Github repository (https://github.com/hydrogo/rainymotion), and has a documentation page (https://rainymotion.readthedocs.io) which includes installation instructions, ~~models'~~ model description, and usage examples. The library code and accompanying documentation are freely distributed under the MIT software license which allows unrestricted use. The library is written in the Python 3 programming language (https://python.org) and its core is entirely based on open source software libraries (Fig. 4): *wradlib* (Heistermann et al., 2013), *OpenCV* (Bradski and Kaehler, 2008), *SciPy* (Jones et al., 2018), *NumPy* (Oliphant, 2006), *Scikit-learn* (Pedregosa et al., 2011), and *Scikit-image* (Van der Walt et al., 2014). For ~~manipulation of the data stored in HDF databases we also use the *h5py* library (), to generate~~ generating figures we use the *Matplotlib* library (Hunter, 2007), and we use the *Jupyter notebook* (https://jupyter.org) interactive development environment for code and documentation development and distribution. For managing the dependencies without any conflicts, we recommend to use the *Anaconda* Python distribution (https://anaconda.com) and follow *rainymotion* installation instructions (https://rainymotion.readthedocs.io).

2.5 Operational baseline (RADVOR)

The DWD operationally runs a stack of models for radar-based nowcasting ~~providing~~ and provides precipitation forecasts for a lead time up to 2 hours. The operational ~~quantitative precipitation nowcasts are~~ QPN is based on the RADVOR module (Bartels et al., 2005; Rudolf et al., 2012). The tracking algorithm estimates the motion field from the latest sequential clutter-filtered radar images using a pattern recognition technique on different spatial resolutions (Winterrath and Rosenow, 2007; Winterrath et al., 2012). The focus of the tracking algorithm is on the meso-β scale (spatial extent: 25–250 km) to cover mainly large-scale precipitation patterns, but the meso-γ scale (spatial extension: 2.5–25 km) is also incorporated to allow the detection of smaller-scale convective structures. The resulting displacement field is interpolated to a regular grid, and a weighted averaging with previously derived displacement fields is implemented to guarantee a smooth displacement over time. The extrapolation of the most recent radar image according to the obtained velocity field is performed using a semi-Lagrangian approach. The described operational model is updated every 5 minutes and produces precipitation nowcasts at a temporal resolution of 5 minutes and a lead time of 2 hours (RV product). In ~~the presented~~ this study we used the RV product data as an operational baseline and did not ~~implement a programmatic realization of the DWD~~ re-implement the underlying algorithm itself.

2.6 Verification

For the verification we use two general categories of scores: continuous (based on the differences between nowcast and observed rainfall intensities) and categorical (based on standard contingency tables for calculating matches between boolean values which reflect the exceedance of specific rainfall intensity thresholds). We use the mean absolute error (MAE) as a continuous score:

$$MAE = \frac{\sum_{i=1}^{n} |now_i - obs_i|}{n} \tag{1}$$

where now_i and obs_i are nowcast and observed rainfall rate in the *i*-th pixel of the corresponding radar image, and *n* the number of pixels.

And we use the critical success index (CSI) as a categorical score:

$$CSI = \frac{hits}{hits + false\ alarms + misses} \tag{2}$$

where *hits*, *false alarms*, and *misses* are defined by the contingency table and the corresponding threshold value (for details see Section S4 of the Supplementary Information).

Following ~~the study of Bowler et al. (2006)~~ studies of Bowler et al. (2006) and Foresti et al. (2016) we have applied threshold rain rates of 0.125, 0.25, 0.5~~and~~ , 1 and 5 mm h^{-1} for calculating the CSI.

These two metrics inform us about the models' performance from the two perspectives: MAE captures errors in rainfall ~~rates~~ rate prediction (the less the better), and CSI captures model accuracy (the fraction of the forecast event that was correctly predicted; does not distinguish the source of errors; the higher the better). You can find results represented in terms of additional categorical scores (false alarm rate, probability of detection, equitable threat score) in Section S4 of the Supplementary Information.

3 Radar data and verification events

We use the so-called RY product of the DWD as input to our nowcasting models. The RY product represents a quality-controlled rainfall depth product that is a composite of the 17 operational Doppler radars maintained by the DWD. It has a spatial extent of 900×900 km and covers the whole area of Germany. Spatial and temporal resolution of the RY product is 1×1 km and 5 minutes, respectively. This composite product includes various procedures for correction and quality control (e.g. clutter removal). We used the $\omega radlib$ (Heistermann et al., 2013) software library for reading the DWD radar data.

For the analysis, we have selected 11 events during the summer periods of 2016 and 2017. These events are selected for covering a range of event characteristics with different rainfall intensity, spatial coverage, and duration. Table 2 shows the studied events. You can also find links to animations of event intensity dynamics in Section S3 of the Supplementary Information.

4 Results

For each event, all models (Sparse, SparseSD, Dense, DenseRotation, Persistence) were used to compute nowcasts with lead times from 5 to 60 minutes (in 5 minute steps). Operational nowcasts generated by the RADVOR system were provided by the DWD with the same temporal settings. An example of nowcasts for lead times 0, 5, 30, and 60 minutes is shown in Fig. 5.

~~Figure 5~~ To investigate the effects of numerical diffusion, we calculated, for the same example, the power spectral density (PSD) of the nowcasts and the corresponding observations (bottom panel in Figure 5) using Welch's method (Welch, 1967). As had been shown in (Germann and Zawadzki, 2002), the most significant loss of power spectra (lower PSD values) refers to small-scale precipitation patterns in the range of 8 to 64 km, so we constrained the PSD plots to highlight that range. The power spectra show that, compared to the observations, the loss of power is small for all lead times, scales, and models (Sparse and Dense). At least for this example, it appears that both the warping and the "interpolate only once" approaches are successful in limiting the effects of numerical diffusion and thus the loss of power at small scales – at least for lead times up to one hour.

Figure 6 shows the model performance (in terms of MAE) as a function of lead time. For each event, the Dense group of models is superior to the other ones. The RV product achieves an ~~average rank between models of the Sparse and Dense groups~~efficiency that is comparable to the Dense group. The SparseSD model outperforms the Sparse model for short lead times (up to 10-15 minutes), and vice versa for longer lead times. For some events (1-4, 6, 10, 11), the performance of the RV product appears to be particularly low in the first 10 minutes, compared to the other models. These events are characterized by particularly fast rainfall field movement.

Figure ~~6~~ 7 has the same structure as Fig. ~~5~~6, but shows the CSI with a threshold value of ~~0.125~~ 1 mm h^{-1}. For two events (7 and 10) the RV product ~~performs better than optical flow based models for lead times beyond 30 minutes, and the Sparse group outperforms~~ achieves a comparable efficiency with the Dense group for lead times beyond ~~45~~ 30 minutes. For the remaining events, the Dense group ~~outperforms~~ tends to outperform all other methods and the RV product achieves an average rank between models of the Sparse and Dense groups. For the Dense group of models, it ~~is clear~~ appears that accounting for ~~the rotation in the field only improves the forecast when strong rotation exists (e.g., event #7, which is consistent with the performance of the RV product that also follows a semi-Lagrangian approach). For the majority of events, however, the constant-vector advection scheme (the Dense model) appears to perform slightly better then the semi-Lagrangian scheme (the DenseRotation model)~~field rotation does not affect the results of the benchmark experiment much – the Dense and DenseRotation models perform very similarly, at least for the selected events and the analyzed lead times. The behavior of the Sparse group models is mostly consistent with the MAE.

Figure ~~7~~ 8 shows the model performance using the CSI with a threshold value of ~~1~~ 5 mm h^{-1}. For the majority of events, the resulting ranking of models is the same as for the CSI with a threshold of ~~0.125~~ 1 mm h^{-1}. For ~~three events (~~events #2 ~~, 7, 10)~~and #3, the performance of the RV product relative to the Dense models is a little bit better, while for other events (e.g. #7), the Dense models outperform the RV product ~~performs better for lead times beyond 20-30 minutes~~more clearly than for the CSI of 1 mm h^{-1}.

Table 3 summarizes the results of the Dense group models in comparison to the RV ~~model~~ product for different verification metrics averaged over all the selected events and two lead time periods: 5–30, and 35–60 minutes. Results show that the Dense group always slightly outperforms the RV model ~~for lead times up to 30 minutes. For lead times from 30 minutes to 1 hour~~in terms of CSI metric for both lead time periods and all analyzed rainfall intensity threshold used for CSI calculation. In terms of MAE, differences between model performances are less pronounced. For the CSI metric, the absolute differences between all models tend to ~~decrease~~ be consistent with increasing rainfall thresholds.

You can find more figures illustrating the models' efficiency for different thresholds and lead times in Section S4 of the Supplementary Information.

5 Discussion

5.1 Model comparison

~~For the~~ All tested models show significant skill over the trivial Eulerian persistence over a lead time of at least one hour. Yet, a substantial loss of skill over lead time is present for all analyzed events, as expected. We have not disentangled the causes of that loss, but predictive uncertainty will always result from errors in both the representation of field motion and the total lack of representing precipitation formation, dynamics, and dissipation in a framework of Lagrangian persistence. Many studies specify a lead time of 30 minutes as a predictability limit for convective structures with fast dynamics of rainfall evolution (Foresti et al., 2016; Grecu and Krajewski, 2000; Thorndahl et al., 2017; Wilson et al., 1998; Zahraei et al., 2012). Our study confirms these findings.

For the majority of analyzed events, there is a clear pattern that the Dense group of optical flow models outperforms the operational RV ~~model for shorter lead times(up to 30 minutes) and sometimes (events 2, 7, 10) underperforms for longer lead times (from 30 minutes to 1 hour). That behavior of~~ nowcast product. For the analyzed events and lead times, the differences between the Dense and the DenseRotation models (or, in other words, between constant-vector and semi-Lagrangian schemes), are negligible. The absolute difference in performance between the Dense group models and the RV product ~~could be a result of accounting for meso-β scale features in the velocity field computation (Winterrath et al., 2012) which is designed to capture movement patterns at a larger scale, concerted with a weighted averaging of the derived displacement vectors over the three recent time steps (in order to guarantee a steady displacement over time).~~ appears to be independent from rainfall intensity threshold and lead time (Table 3), which implies that the relative advance of the Dense group models over the RV product increases both with lead time and rainfall intensity threshold. A gain in performance for longer lead times by taking into account more time steps from the past can ~~also~~ be observed when comparing the SparseSD model (looks back five minutes in time) against the Sparse model (looks back two hours in time).

~~The natural properties of the precipitation formation process limit predictability. Many studies specify a lead time of 30 minutes as a predictability limit for convective structures with fast dynamics of rainfall evolution (Foresti et al., 2016; Grecu and Krajewski, . Our study confirms these findings~~Despite their skill over Eulerian persistence, the Sparse group models are significantly outperformed by the Dense group models for all the analyzed events and lead times. The reason for this behaviour remains yet unclear. It could, in general, be a combination of errors introduced in corner-tracking and extrapolation as well as image warping as a surrogate for formal advection. While the systematic identification of error sources will be subject to future studies, we suspect that the the local features ("corners") identified by the Shi–Tomasi corner detector might not be representative for the overall motion of the precipitation field: the detection focuses on features with high intensities and gradients, the motion of which might not represent the dominant meso-γ scale motion patterns.

There are a couple of possible directions for enhancing the performance for longer lead times using the Dense group of models. ~~The first way is to adopt the RV scheme for a weighted averaging of calculated velocity fields separately derived for the last~~ A first is to use a weighted average of velocity fields derived from radar images three (or more) ~~time steps . Another~~ steps back in time (as done in RADVOR to compute the RV product). A second option is to calculate separate velocity fields for low and high intensity subregions of the rain field, and advect these subregions separately (like proposed in Golding (1998)), or find an optimal weighting procedure. ~~The third possible way is to implement different smoothness constraints when implementing optical flow algorithms like~~ A third approach could be to optimize the use of various optical flow constraints in order to improve the performance for longer lead times, as proposed in Germann and Zawadzki (2002), Bowler et al. (2004), or Mecklenburg et al. (2000). The flexibility of the *rainymotion* software library allows users to incorporate such algorithms for benchmarking any hypothesis, and e.g. implement different models or ~~parameterisations~~ parameterizations for different lead times. Bowler et al. (2004) also showed a significant performance increase for longer lead times by using NWP model winds for the advection step. However, Winterrath and Rosenow (2007) did not obtain any improvement compared to RADVOR for longer lead times by incorporating NWP model winds ~~in~~ into the nowcasting procedure.

5.2 Advection schemes properties and effectiveness

~~In our study, we have shown the advantages~~ Within the Dense group of models, we could not find any significant difference between the performance and PSD of the constant-vector ~~advection scheme (implemented in the Dense model, provided in Bowler et al. (2004)) over the semi-Lagrangian scheme (implemented in the DenseRotationmodel). For the majority of events (except the event #7) and lead times it appears that a linear extrapolation together with a constant-vector advection are preferable, particularly in a context where we have a complex motion pattern with an absence of distinct large-scale rotation. However, when a precipitation field has a clear rotational component (e. g. counter-clock wise cyclonic rotation during the event #7) , accounting for rotation increases the forecast efficiency (DenseRotation outperforms Dense). It is also possible that the positive effect of using a~~ (Dense model) and the Semi-Lagrangian scheme (DenseRotation). That confirms findings presented by Germann and Zawadzki (2002) who found that the constant vector and the modified semi-Lagrangian ~~advection scheme may be more evident~~ schemes have very similar power spectra, presumably since they share the same interpolation procedure. The theoretical superiority of the Semi-Lagrangian scheme might, however, materialize for other events with substantial, though persistent rotational motion. A more comprehensive analysis should thus be subject to future studies.

Interpolation is included in both the post-processing of image warping (Sparse models) and in the computation of gridded nowcasts as part of the Dense models. In general, such interpolation steps can lead to numerical diffusion and thus to the degradation or loss of small-scale features (Germann and Zawadzki, 2002). Yet, we were mostly able to contain such adverse effects for both the Sparse and the Dense group of models by carrying out only one interpolation step for any forecast at a specific lead time. We showed that numerical diffusion was negligible for lead times ~~longer than one hour , and for smoothed velocity fields (as~~ up to one hour for any model, however, as had been shown in Germann and Zawadzki (2002)~~).~~ , for longer lead times these effects can be significant, depending on the implemented extrapolation technique.

5.3 Computational performance

Computational performance might be an important criterion for end users aiming at ~~update cycles with high frequency~~frequent update cycles. We ran our nowcasting models on a standard office PC with an Intel® Core™ i7-2600 CPU (8 cores, 3.4 GHz), and on a standard laptop with an Intel® Core™ i5-7300HQ CPU (4 cores, 2.5 GHz). The average time for generating one nowcast for one hour lead time (at 5 minute resolution) for the Sparse group is 2-3 s, and for the Dense group is ~~150-180~~ 7-12 s. The Dense group ~~consumes more computational resources mostly because of expensive interpolation operation implemented for images of high resolution~~ is computationally more expensive due to interpolation operations implemented for large grids (900×900 pixels). There is also ~~ample~~ potential for increasing the computational performance of the interpolation. ~~It might also be considered to combine the warping procedure for the extrapolation step with the Dense optical flow procedure for the tracking step in order to dramatically enhance computational performance. For that purpose, however, the errors introduced by the warping procedure need to be understood better.~~

6 Summary and conclusions

Optical flow is a technique for deriving a velocity field from consecutive ~~series of imageswhich~~ images. It is widely used in image analysis, and became increasingly popular in meteorological applications over the past 20 years. In our study, we examined the performance of optical flow based models for radar-based precipitation nowcasting, as implemented in the open-source *rainymotion* library, for a wide range of rainfall events using radar data provided by the DWD.

~~The comparison of the models' verification performance with the performance of the operational baseline~~ Our benchmark experiments, including an operational baseline model (the RV product provided by the DWD)~~shows~~ , show a firm basis for using optical flow in radar-based precipitation nowcasting studies. For the majority of the analyzed events, models ~~which use the global optical flow algorithm for deriving a displacement vector field~~ from the Dense group outperform the operational ~~model~~baseline. The Sparse group of models showed significant skill, yet they performed generally poorer than both the Dense group and the RV product. We should, however, not prematurely discard the group of Sparse models before we have not gained a better understanding of error sources with regard to the tracking, extrapolation and warping steps. It might also be considered to combine the warping procedure for the extrapolation step with the Dense optical flow procedure for the tracking step (i.e. to advect "corners" based on a "Dense" velocity field obtained by implementing one of the dense optical flow techniques). This opens the way for merging two different model development branches in the future releases of the *rainymotion* library.

There is a clear and rapid model performance loss over lead time for events with high rainfall intensities. This issue continues to be unresolved by standard nowcasting approaches, but some improvement in this field may be achieved with using strategies such as ~~nowcasts~~ merging with NWP results and stochastic ~~accounting~~ modelling of rainfall field evolution. ~~We~~ Admittedly, deterministic nowcasts in a Lagrangian framework do neither account for precipitation intensity dynamics nor for the uncertainties in representing precipitation field motion. At least for the latter, the *rainymotion* library provides ample opportunities to experiment with forecast ensembles, based on various tracking and extrapolation techniques. Furthermore, we suppose that using new data-driven models based on machine and deep learning ~~approaches may provide additional gain in~~ may increase the performance by utilizing and structuring common patterns in the massive archives of radar data.

We do not claim that the developed models will compete with well-established and excessively tuned operational ~~services~~ models for radar-based precipitation nowcasting~~, but~~ . Yet, we hope our models may serve as an essential tool for providing a fast, free and open source solution that can serve as a benchmark for further model development and hypothesis testing – a benchmark that is far more advanced than the conventional benchmark of Eulerian persistence.

Recent studies show that open source community-driven software advances the field of weather radar science (Heistermann et al., 2015a, . Just a few months ago, the *pySTEPS* (https://pysteps.github.io) initiative was introduced "to develop and maintain an easy to use, modular, free and open source python framework for short-term ensemble prediction systems." As another evidence of the dynamic evolution of QPN research over the recent years, these developments could pave the way for future synergies between the *pySTEPS* and *rainymotion* projects – towards the availability of open, reproducible, and skillful methods in quantitative precipitation nowcasting.

Code and data availability. The *rainymotion* library is free and open source. It is distributed under the MIT software license which allows unrestricted use. The source code is provided through a GitHub repository https://github.com/hydrogo/rainymotion, the snapshot of the *rainymotion* v0.1 is also available on Zenodo: https://doi.org/10.5281/zenodo.2561583, and the documentation is available on a website https://rainymotion.readthedocs.io. The DWD provided the sample data of the RY product, and it is distributed with the *rainymotion* repository to provide a real case and reproducible example of precipitation nowcasting.

Author contributions. GA performed the benchmark experiments, analyzed the data and wrote the manuscript. MH coordinated and supervised the work, analyzed the data and wrote the manuscript. TW assisted in the data retrieval and analysis, and shared her expertise in DWD radar products.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. ~~This publication~~ Georgy Ayzel was financially supported by Geo.X, the Research Network for Geosciences in Berlin and Potsdam. The authors thank Loris Foresti, Seppo Pulkkinen, and Remko Uijlenhoet for their constructive comments and suggestions that helped to improve the paper.

References

Austin, G. L. and Bellon, A.: The use of digital weather radar records for short-term precipitation forecasting, Quarterly Journal of the Royal Meteorological Society, 100, 658–664, https://doi.org/10.1002/qj.49710042612, http://doi.wiley.com/10.1002/qj.49710042612, 1974.

Bauer, P., Thorpe, A., and Brunet, G.: The quiet revolution of numerical weather prediction, Nature, 525, 47–55, https://doi.org/10.1038/nature14956, http://www.nature.com/doifinder/10.1038/nature14956, 2015.

Bellerby, T. J.: High-resolution 2-D cloud-top advection from geostationary satellite imagery, IEEE Transactions on Geoscience and Remote Sensing, 44, 3639–3648, https://doi.org/10.1109/TGRS.2006.881117, http://ieeexplore.ieee.org/document/4014303/, 2006.

Berenguer, M., Sempere-Torres, D., and Pegram, G. G. S.: SBMcast - An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation, Journal of Hydrology, 404, 226–240, https://doi.org/10.1016/j.jhydrol.2011.04.033, https://www.sciencedirect.com/science/article/pii/S0022169411002940, 2011.

Bowler, N. E., Pierce, C. E., and Seed, A.: Development of a precipitation nowcasting algorithm based upon optical flow techniques, Journal of Hydrology, 288, 74–91, https://doi.org/10.1016/j.jhydrol.2003.11.011, https://www.sciencedirect.com/science/article/pii/S0022169403004591, 2004.

Bowler, N. E., Pierce, C. E., and Seed, A. W.: STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP, Quarterly Journal of the Royal Meteorological Society, 132, 2127–2155, https://doi.org/10.1256/qj.04.100, http://dx.doi.org/10.1256/qj.04.100, 2006.

Bradski, G. and Kaehler, A.: Learning OpenCV: Computer vision with the OpenCV library, " O'Reilly Media, Inc.", 2008.

Brox, T., Bruhn, A., Papenberg, N., and Weickert, J.: High accuracy optical flow estimation based on a theory for warping, in: European conference on computer vision, pp. 25–36, Springer, 2004.

Bruhn, A., Weickert, J., Feddern, C., Kohlberger, T., and Schnörr, C.: Variational optical flow computation in real time, IEEE transactions on image processing : a publication of the IEEE Signal Processing Society, 14, 608–15, https://doi.org/10.1109/TIP.2005.846018, https://ieeexplore.ieee.org/document/1420392/, 2005a.

Bruhn, A., Weickert, J., and Schnörr, C.: Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods, International journal of computer vision, 61, 211–231, 2005b.

Cheung, P. and Yeung, H. Y.: Application of optical-flow technique to significant convection nowcast for terminal areas in Hong Kong, in: The 3rd WMO International Symposium on Nowcasting and Very Short-Range Forecasting (WSN12), pp. 1–10, http://www.hko.gov.hk/publica/reprint/r1025.pdf, 2012.

Farnebäck, G.: Two-frame motion estimation based on polynomial expansion, Image Analysis, 2003, 363–370, https://doi.org/10.1007/3-540-45103-x_50, http://link.springer.com/10.1007/3-540-45103-X, 2003.

Foresti, L., Panziera, L., Mandapaka, P. V., Germann, U., and Seed, A.: Retrieval of analogue radar images for ensemble nowcasting of orographic rainfall, Meteorological Applications, 22, 141–155, https://doi.org/10.1002/met.1416, http://doi.wiley.com/10.1002/met.1416, 2015.

Foresti, L., Reyniers, M., Seed, A., and Delobbe, L.: Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in Belgium, Hydrology and Earth System Sciences, 20, 505–527, https://doi.org/10.5194/hess-20-505-2016, https://www.hydrol-earth-syst-sci.net/20/505/2016/, 2016.

Germann, U. and Zawadzki, I.: Scale-Dependence of the Predictability of Precipitation from Continental Radar Images. Part I: Description of the Methodology, Monthly Weather Review, 130, 2859–2873, https://doi.org/10.1175/1520-0493(2002)130<2859:SDOTPO>2.0.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1520-0493{%}282002{%}29130{%}3C2859{%}3ASDOTPO{%}3E2.0.CO{%}3B2, 2002.

Golding, B. W.: Nimrod: A system for generating automated very short range forecasts, Meteorological Applications, 5, 1–16, https://doi.org/10.1017/S1350482798000577, http://doi.wiley.com/10.1017/S1350482798000577, 1998.

Grecu, M. and Krajewski, W. F.: A large-sample investigation of statistical procedures for radar-based short-term quantitative precipitation forecasting, Journal of Hydrology, 239, 69–84, https://doi.org/10.1016/S0022-1694(00)00360-7, https://www.sciencedirect.com/science/article/pii/S0022169400003607, 2000.

Heistermann, M., Jacobi, S., and Pfaff, T.: Technical Note: An open source library for processing weather radar data (wradlib), Hydrology and Earth System Sciences, 17, 863–871, https://doi.org/10.5194/hess-17-863-2013, https://www.hydrol-earth-syst-sci.net/17/863/2013/, 2013.

Heistermann, M., Collis, S., Dixon, M. J., Giangrande, S., Helmus, J. J., Kelley, B., Koistinen, J., Michelson, D. B., Peura, M., Pfaff, T., and Wolff, D. B.: The emergence of open-source software for the weather radar community, Bulletin of the American Meteorological Society, 96, 117–128, https://doi.org/10.1175/BAMS-D-13-00240.1, http://journals.ametsoc.org/doi/10.1175/BAMS-D-13-00240.1, 2015a.

Heistermann, M., Collis, S., Dixon, M. J., Helmus, J. J., Henja, A., Michelson, D. B., and Pfaff, T.: An open virtual machine for cross-platform weather radar science, Bulletin of the American Meteorological Society, 96, 1641–1645, https://doi.org/10.1175/BAMS-D-14-00220.1, http://journals.ametsoc.org/doi/10.1175/BAMS-D-14-00220.1, 2015b.

Horn, B. K. and Schunck, B. G.: Determining optical flow, Artificial intelligence, 17, 185–203, 1981.

Hunter, J. D.: Matplotlib: A 2D graphics environment, Computing in science & engineering, 9, 90–95, 2007.

Jones, E., Oliphant, T., and Peterson, P.: SciPy: open source scientific tools for Python, https://scipy.org/, last access: 29 June 2018, 2018.

Kroeger, T., Timofte, R., Dai, D., and Van Gool, L.: Fast optical flow using dense inverse search, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9908 LNCS, pp. 471–488, https://doi.org/10.1007/978-3-319-46493-0_29, http://arxiv.org/abs/1603.03590, 2016.

Liu, Y., Xi, D. G., Li, Z. L., and Hong, Y.: A new methodology for pixel-quantitative precipitation nowcasting using a pyramid Lucas Kanade optical flow approach, Journal of Hydrology, 529, 354–364, https://doi.org/10.1016/j.jhydrol.2015.07.042, https://www.sciencedirect.com/science/article/pii/S002216941500548X, 2015.

Lucas, B. D. and Kanade, T.: An iterative image Registration Technique with an Application to Stereo Vision, in: Proceedings DARPA Image Understanding Workrhop, pp. 674–679, Morgan Kaufmann Publishers Inc., https://doi.org/10.1145/358669.358692, https://ri.cmu.edu/pub_files/pub3/lucas_bruce_d_1981_2/lucas_bruce_d_1981_2.pdf, 1981.

Mecklenburg, S., Joss, J., and Schmid, W.: Improving the nowcasting of precipitation in an Alpine region with an enhanced radar echo tracking algorithm, Journal of Hydrology, 239, 46–68, https://doi.org/10.1016/S0022-1694(00)00352-8, https://www.sciencedirect.com/science/article/pii/S0022169400003528, 2000.

Mueller, C. M., Axen, T. S., Oberts, R. R., Ilson, J. W., Etancourt, T. B., Ettling, S. D., and Ien, N. O.: NCAR Auto-Nowcast System, Weather and Forecasting, 18, 545–561, https://doi.org/10.1175/1520-0434(2003)018<0545:NAS>2.0.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1520-0434{%}282003{%}29018{%}3C0545{%}3ANAS{%}3E2.0.CO{%}3B2, 2003.

Oliphant, T. E.: A guide to NumPy, vol. 1, Trelgol Publishing USA, 2006.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al.: Scikit-learn: Machine learning in Python, Journal of machine learning research, 12, 2825–2830, 2011.

Reyniers, M.: Quantitative precipitation forecasts based on radar observations: Principles, algorithms and operational systems, Institut Royal Météorologique de Belgique, https://www.meteo.be/meteo/download/fr/3040165/pdf/rmi_scpub-1261.pdf, 2008.

Ruzanski, E., Chandrasekar, V., and Wang, Y.: The CASA nowcasting system, Journal of Atmospheric and Oceanic Technology, 28, 640–655, https://doi.org/10.1175/2011JTECHA1496.1, http://journals.ametsoc.org/doi/abs/10.1175/2011JTECHA1496.1, 2011.

Schneider, P. J. and Eberly, D. H.: Geometric tools for computer graphics, Boston, https://www.sciencedirect.com/science/book/9781558605947, 2003.

Shi, J. and Tomasi, C.: Good features to track, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94, pp. 593–600, IEEE Comput. Soc. Press, https://doi.org/10.1109/CVPR.1994.323794, http://ieeexplore.ieee.org/document/323794/, 1994.

Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J. E., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., Rasmussen, M. R., and Molnar, P.: Weather radar rainfall data in urban hydrology, Hydrology and Earth System Sciences, 21, 1359–1380, https://doi.org/10.5194/hess-21-1359-2017, http://www.hydrol-earth-syst-sci.net/21/1359/2017/, 2017.

Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T.: scikit-image: image processing in Python, PeerJ, 2, e453, https://doi.org/10.7717/peerj.453, https://peerj.com/articles/453/, 2014.

Weinzaepfel, P., Revaud, J., Harchaoui, Z., and Schmid, C.: DeepFlow: Large Displacement Optical Flow with Deep Matching, in: 2013 IEEE International Conference on Computer Vision, pp. 1385–1392, IEEE, https://doi.org/10.1109/ICCV.2013.175, http://ieeexplore.ieee.org/document/6751282/, 2013.

Welch, P.: The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Transactions on audio and electroacoustics, 15, 70–73, 1967.

Wilson, J. W., Crook, N. A., Mueller, C. K., Sun, J., and Dixon, M.: Nowcasting Thunderstorms: A Status Report, Bulletin of the American Meteorological Society, 79, 2079–2099, https://doi.org/10.1175/1520-0477(1998)079<2079:NTASR>2.0.CO;2, https://journals.ametsoc.org/doi/abs/10.1175/1520-0477%281998%29079%3C2079%3ANTASR%3E2.0.CO%3B2, 1998.

Winterrath, T. and Rosenow, W.: A new module for the tracking of radar-derived precipitation with model-derived winds, Advances in Geosciences, 10, 77–83, https://doi.org/10.5194/adgeo-10-77-2007, http://www.adv-geosci.net/10/77/2007/, 2007.

Winterrath, T., Rosenow, W., and Weigl, E.: On the DWD quantitative precipitation analysis and nowcasting system for real-time application in German flood risk management, in: Weather Radar and Hydrology (Proceedings of a symposium held in Exeter, UK, April 2011) IAHS Publ., vol. 351, pp. 323–329, https://www.dwd.de/DE/leistungen/radolan/radolan_info/Winterrath_German_flood_risk_management_pdf.pdf?__blob=publicationFile&v=4, 2012.

Wolberg, G.: Digital Image Warping, IEEE Computer Society Press, 1990.

Wong, W. K., Yeung, L. H. Y., Wang, Y. C., and Chen, M.: Towards the blending of NWP with nowcast - Operation experience in B08FDP, in: WMO symposium on nowcasting, http://my.hko.gov.hk/publica/reprint/r844.pdf, 2009.

Woo, W.-C. and Wong, W.-K.: Operational Application of Optical Flow Techniques to Radar-Based Rainfall Nowcasting, Atmosphere, 8, 48, https://doi.org/10.3390/atmos8030048, http://www.mdpi.com/2073-4433/8/3/48, 2017.

Wulff, J. and Black, M. J.: Efficient Sparse-to-Dense Optical Flow Estimation Using a Learned Basis and Layers, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

Yeung, L. H. Y., Wong, W. K., Chan, P. K. Y., Lai, E. S. T., Yeung, L. H. Y., Wong, W. K., Chan, P. K. Y., and Lai, E. S. T.: Applications of the Hong Kong Observatory nowcasting system SWIRLS-2 in support of the 2008 Beijing Olympic Games, in: WMO Symposium on Nowcasting, vol. 30, http://my.hko.gov.hk/publica/reprint/r843.pdf, 2009.

Zahraei, A., lin Hsu, K., Sorooshian, S., Gourley, J. J., Lakshmanan, V., Hong, Y., and Bellerby, T.: Quantitative Precipitation Nowcasting: A Lagrangian Pixel-Based Approach, Atmospheric Research, 118, 418–434, https://doi.org/10.1016/j.atmosres.2012.07.001, https://www.sciencedirect.com/science/article/pii/S0169809512002219, 2012.

Zahraei, A., lin Hsu, K., Sorooshian, S., Gourley, J. J., Hong, Y., and Behrangi, A.: Short-term quantitative precipitation forecasting using an object-based approach, Journal of Hydrology, 483, 1–15, https://doi.org/10.1016/j.jhydrol.2012.09.052, https://www.sciencedirect.com/science/article/pii/S0022169412008694, 2013.

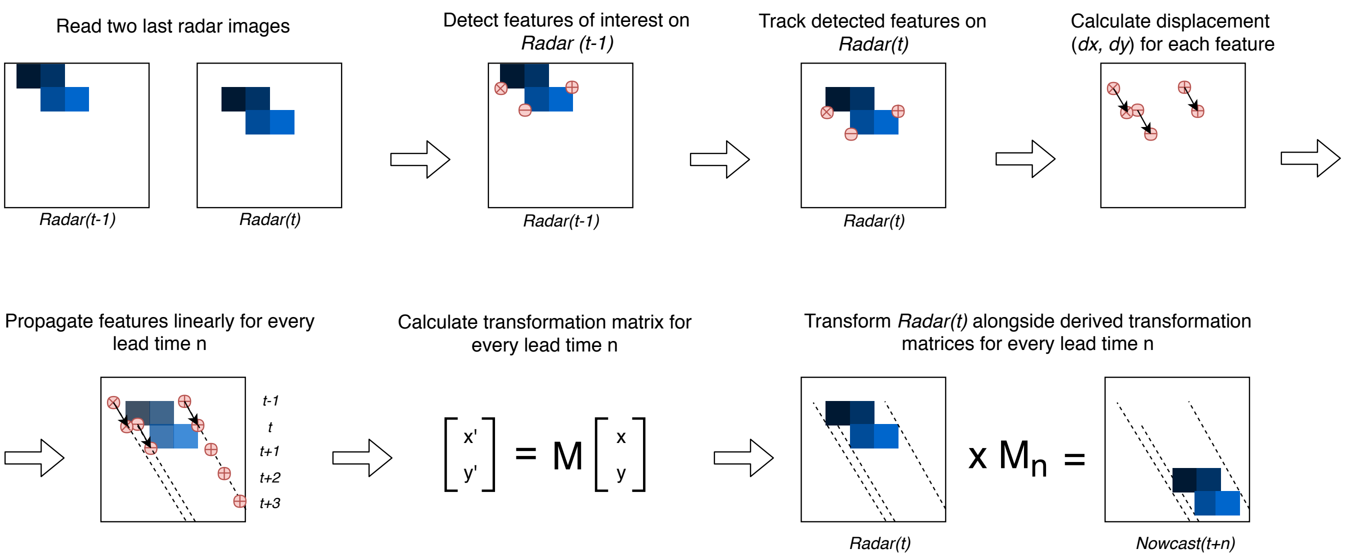

Figure 1. Scheme of the SparseSD model

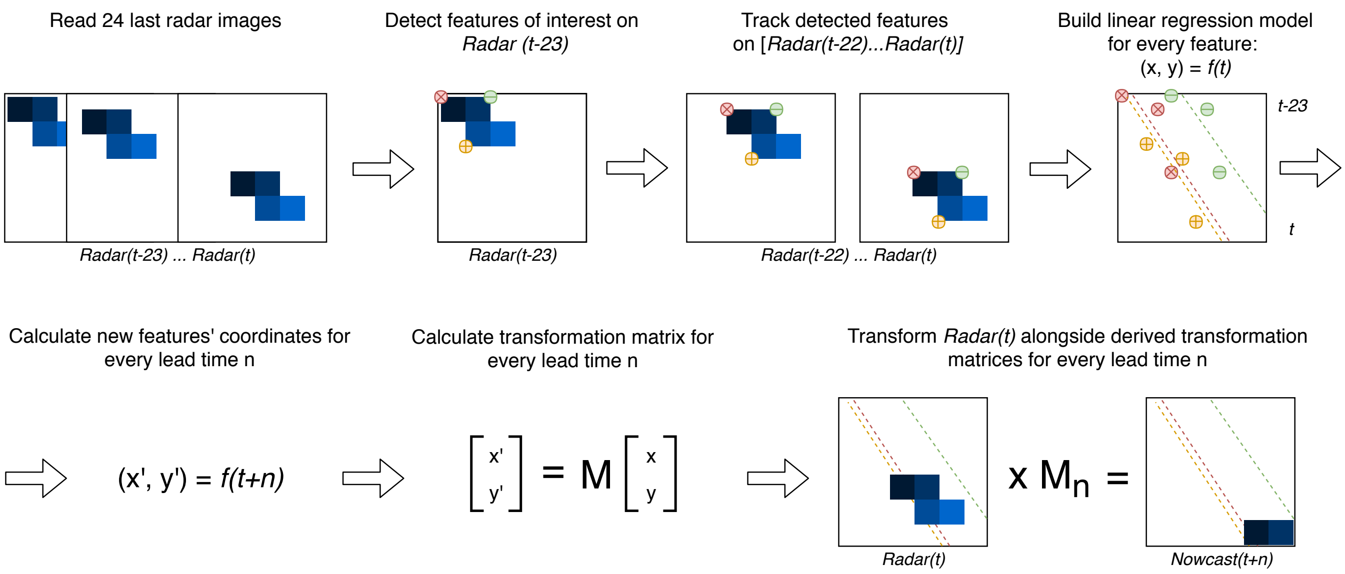

Figure 2. Scheme of the Sparse model

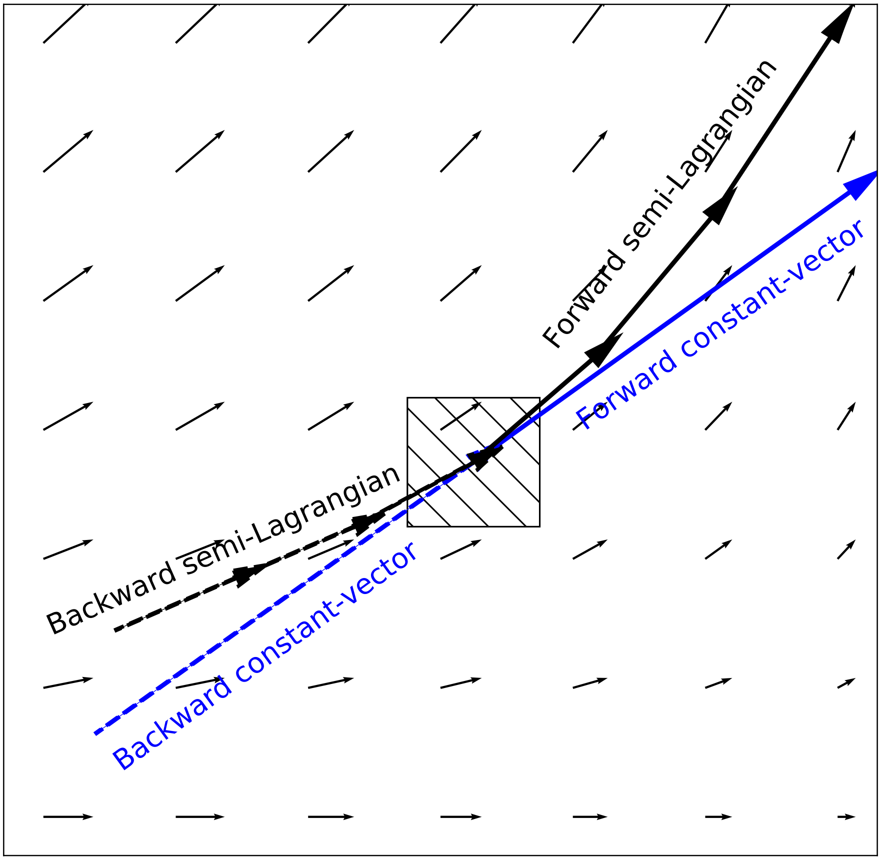

Figure 3. Displacement vectors of ~~two~~ four proposed advection schemes: forward/backward constant vector and forward/backward semi-Lagrangian

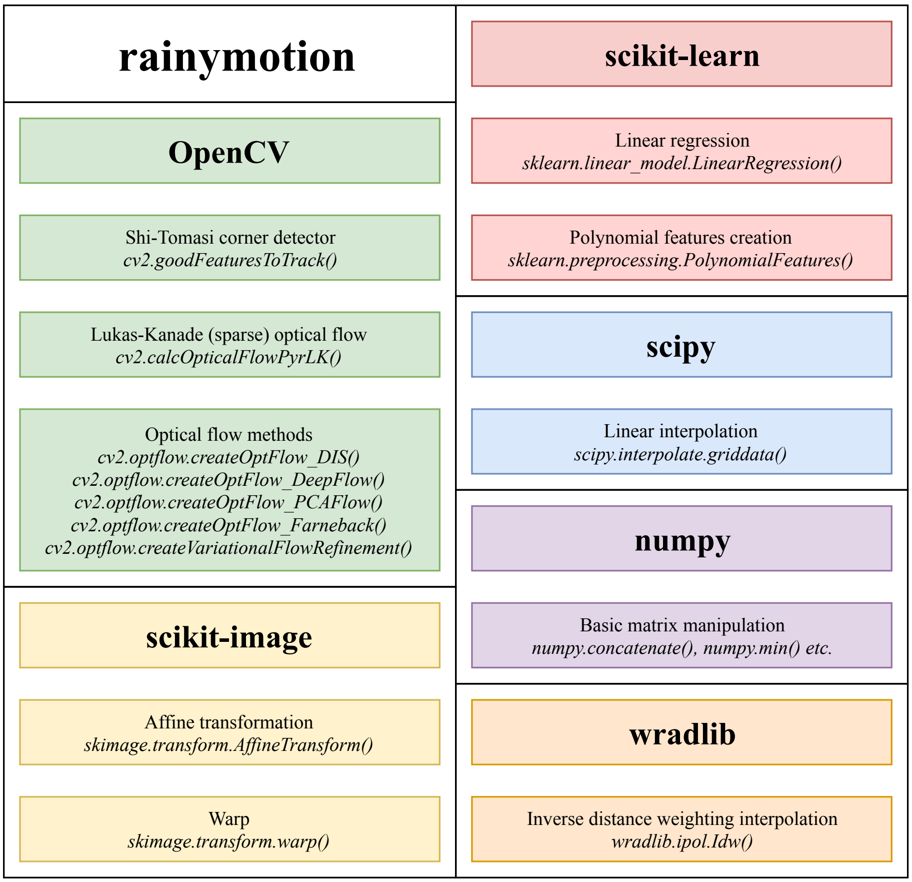

Figure 4. Key Python libraries for *rainymotion* library development

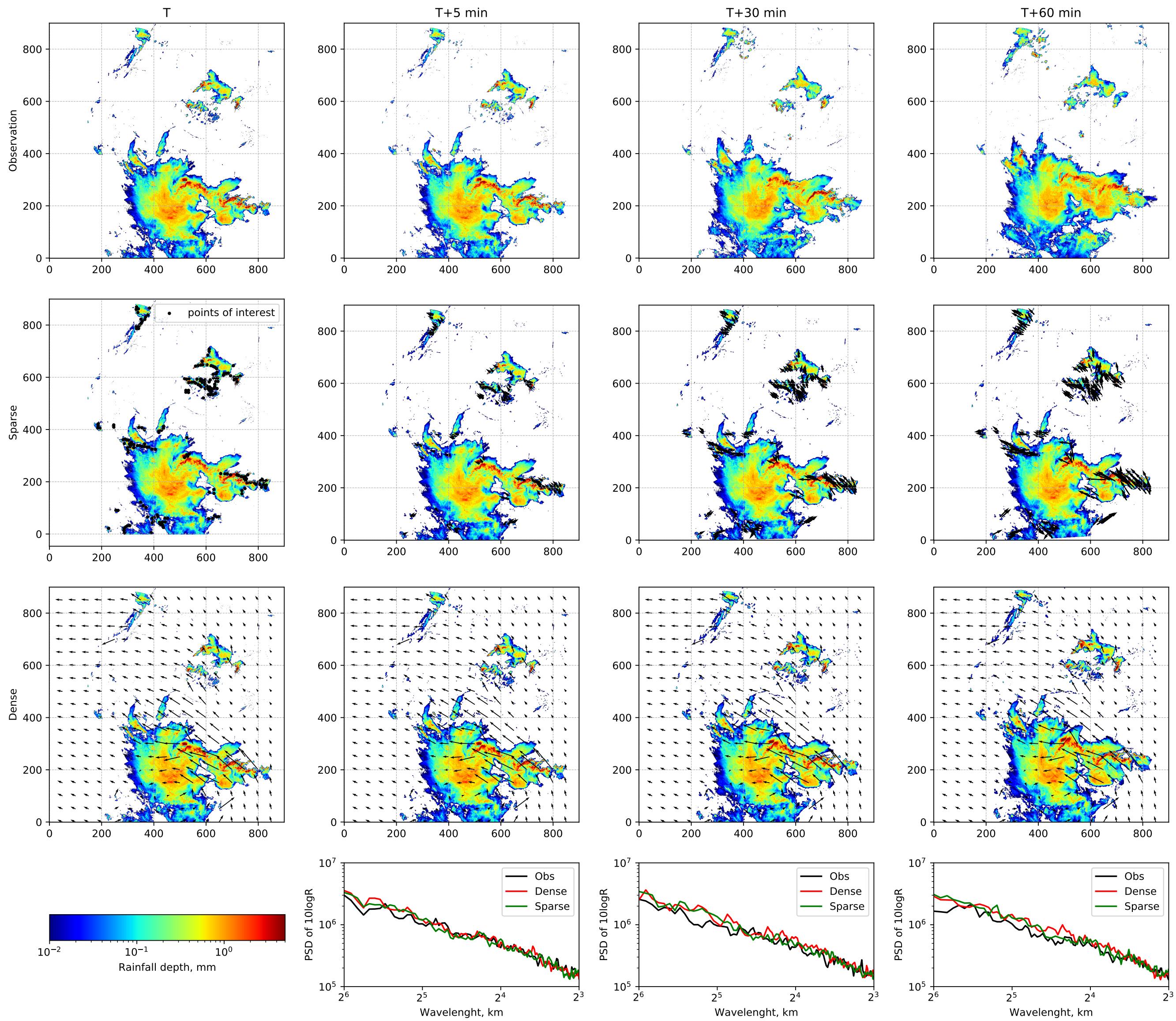

Figure 5. ~~Lead time wise MAE~~Example of the nowcasting models output (SparseSD and Dense models) for the timestep "2016-05-29 19:15" and corresponding level of numerical diffusion (the last row)

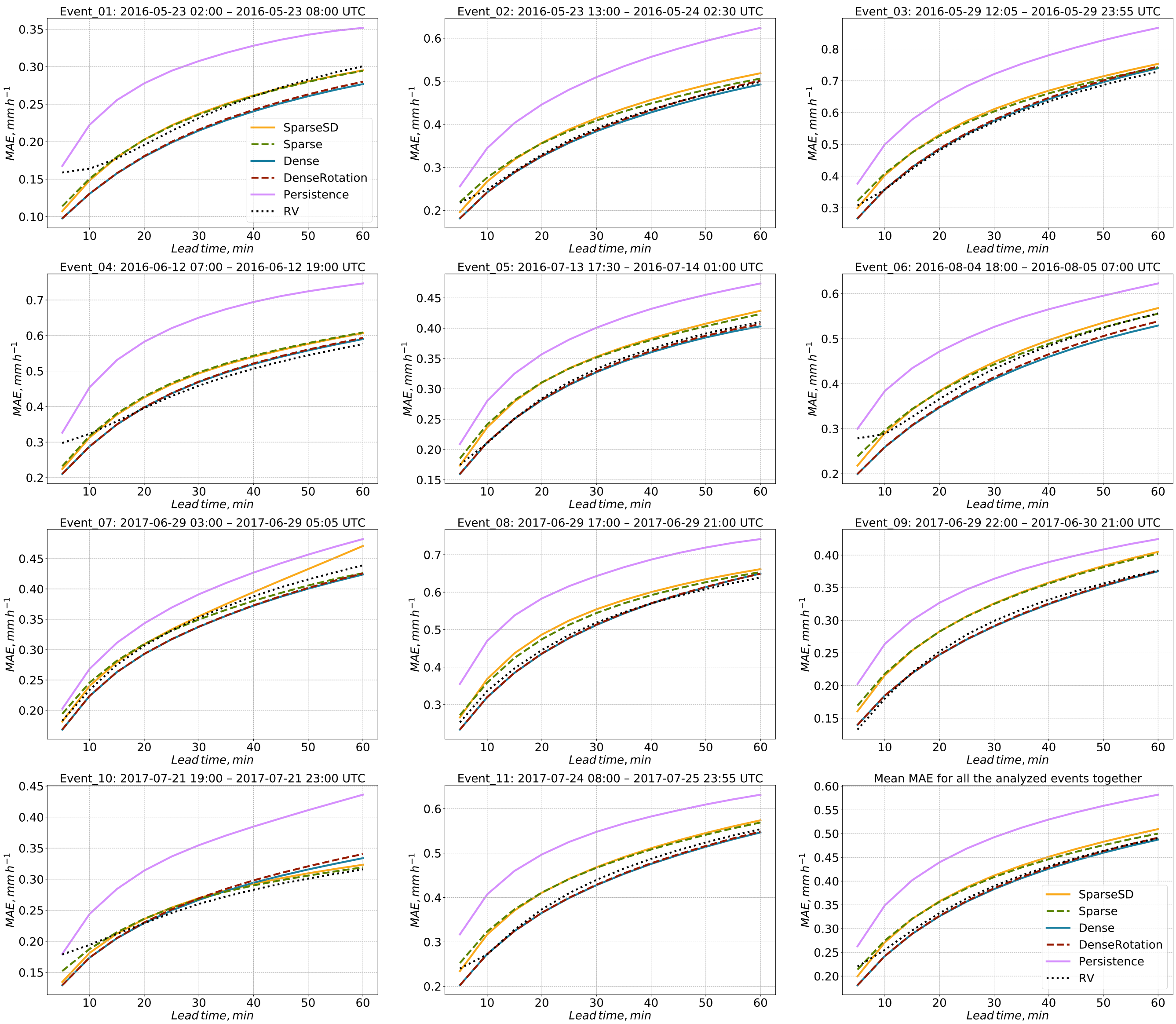

Figure 6. ~~Lead time wise CSI for~~ Verification of the ~~threshold~~ different optical flow based nowcasts in terms of ~~0.125 mm h^{-1}~~ MAE for 11 precipitation events over Germany

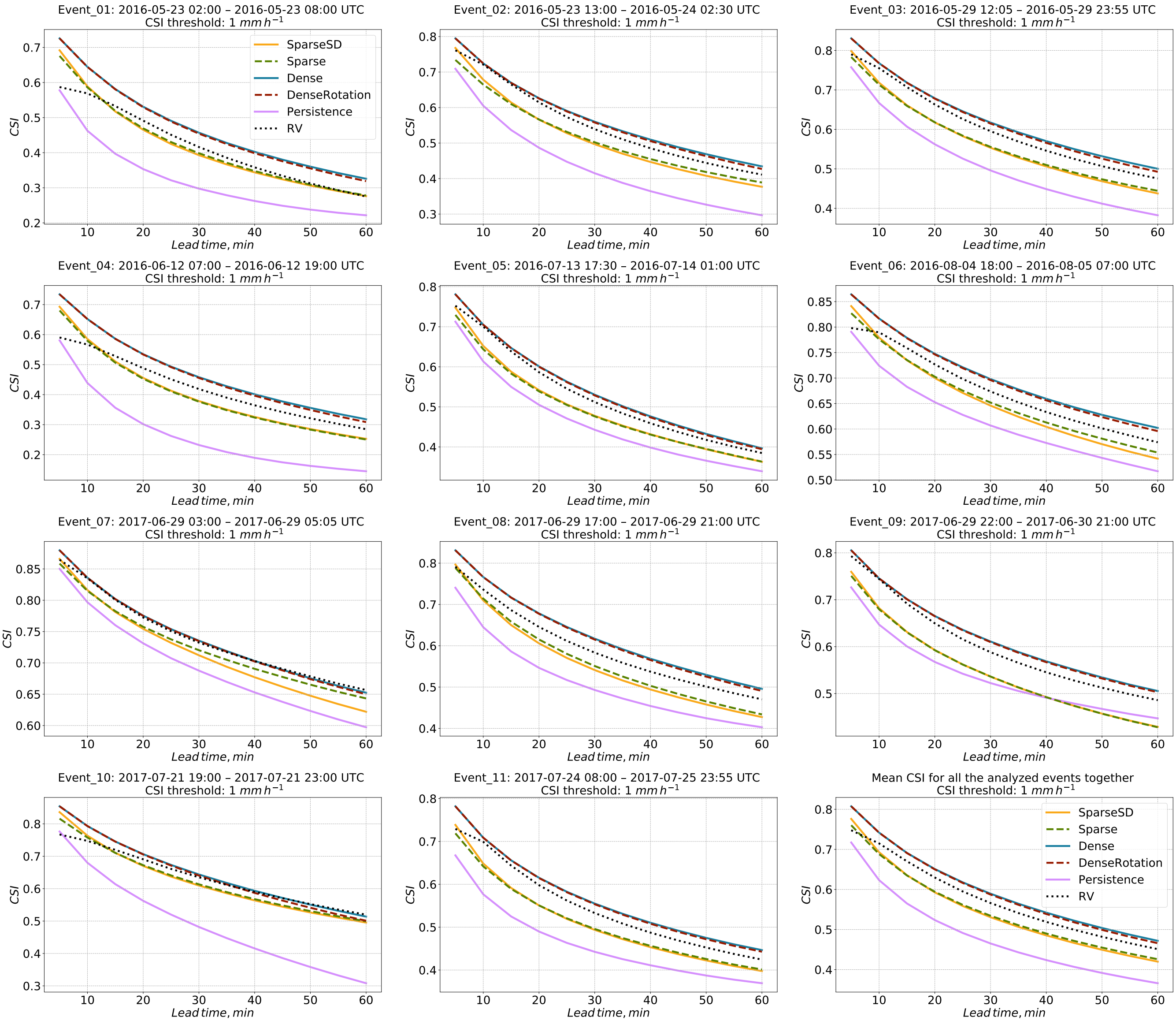

Figure 7. ~~Lead time wise~~ Verification of the different optical flow based nowcasts in terms of CSI for the threshold of ~~1.0~~ 1 mm h^{-1} for 11 precipitation events over Germany

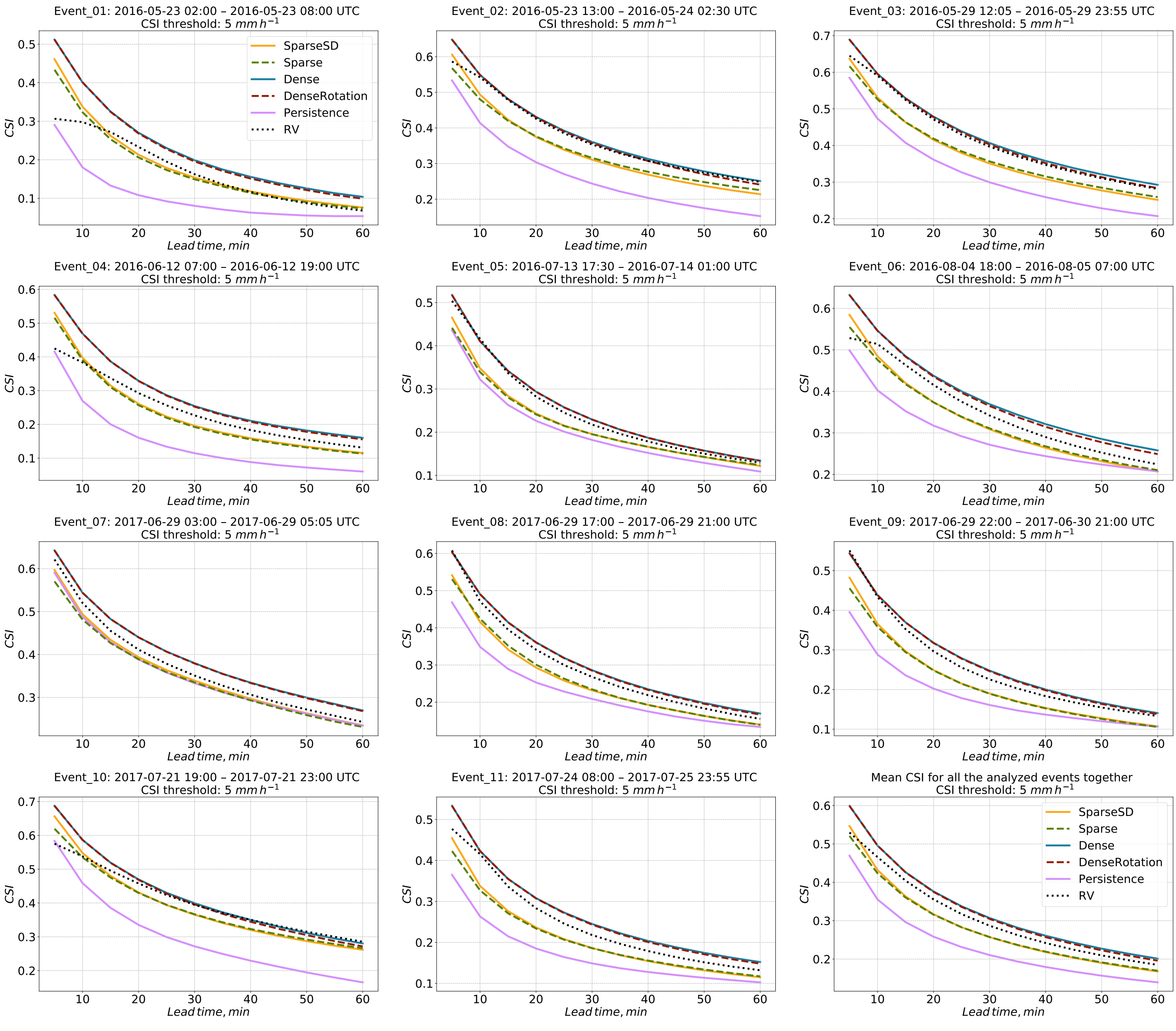

Figure 8. Verification of the different optical flow based nowcasts in terms of CSI for the threshold of 5 mm h^{-1} for 11 precipitation events over Germany

Table 1. Overview of the developed models

Model name	Input radar images	~~Tracking~~ Default tracking algorithm	Extrapolation	~~Number of parameters~~ Computational time, s
SparseSD	2	Shi–Tomasi corner detector, Lucas–Kanade optical flow ~~(local)~~	Constant delta-change, affine warping	~~7~~ ∼2-3
Sparse	3-24	Shi–Tomasi corner detector, Lucas–Kanade optical flow ~~(local)~~	Linear regression, affine warping	~~10~~ ∼2-3
Dense	2	~~Farnebäck optical flow (global)~~ DIS optical flow	~~Constant-vector~~ Backward constant-vector advection scheme	~~7 ∼150-180~~ 7-9
DenseRotation	2	~~Farnebäck optical flow (global)~~ DIS optical flow	~~Forward~~ Backward semi-Lagrangian advection scheme	~~7 ∼150-180~~ 10-12

Table 2. Characteristics of the selected events

Event ~~shortcut~~ #	Start	End	Duration, hours	Maximum extent, km^2	Extent >1 mm h^{-1}, %
Event 1	2016-05-23 2:00	2016-05-23 8:00	6	159318	42
Event 2	2016-05-23 13:00	2016-05-24 2:30	13.5	135272	56
Event 3	2016-05-29 12:05	2016-05-29 23:55	12	160095	72
Event 4	2016-06-12 7:00	2016-06-12 19:00	12	150416	53
Event 5	2016-07-13 17:30	2016-07-14 1:00	7.5	145501	62
Event 6	2016-08-04 18:00	2016-08-05 7:00	13	168407	74
Event 7	2017-06-29 3:00	2017-06-29 5:05	2	140021	70
Event 8	2017-06-29 17:00	2017-06-29 21:00	4	182561	60
Event 9	2017-06-29 22:00	2017-06-30 21:00	23	160822	75
Event 10	2017-07-21 19:00	2017-07-21 23:00	4	63698	77
Event 11	2017-07-24 8:00	2017-07-25 23:55	16	253666	63

Table 3. Mean model metrics for different lead time periods

	Lead time (from–to), min	
Model	5–30	35–60
	MAE, mm h^{-1}	
Dense	~~**0.29**~~ 0.30	~~**0.44**~~ 0.45
DenseRotation	~~**0.29**~~ 0.30	~~**0.44**~~ 0.45
RV	0.31	0.45
	CSI, threshold=0.125 mm h^{-1}	
Dense	~~**0.78**~~ 0.78	~~**0.63**~~ 0.64
DenseRotation	~~**0.78**~~ 0.78	~~0.62~~ 0.64
RV	0.76	0.61
	CSI, threshold=0.25 mm h^{-1}	
Dense	~~**0.76**~~ 0.76	~~**0.60**~~ 0.61
DenseRotation	~~**0.76**~~ 0.76	~~0.59~~ 0.61
RV	0.74	0.59
	CSI, threshold=0.5 mm h^{-1}	
Dense	~~**0.73**~~ 0.73	~~0.55~~ 0.57
DenseRotation	~~0.72~~ 0.73	~~0.55~~ 0.57
RV	0.70	0.55
	CSI, threshold=1 mm h^{-1}	
Dense	~~**0.68**~~ 0.68	~~0.49~~ 0.52
DenseRotation	~~0.67~~ 0.68	~~0.49~~ 0.51
RV	0.65	0.49
	CSI, threshold=5 mm h^{-1}	
Dense	0.42	0.24
DenseRotation	0.42	0.23
RV	0.39	0.22