(REVIEW ARTICLE)

# Mini-review of Pysteps: An open-source python library for precipitation nowcasting

Nasooh Ismail *

*Nanjing University of Information Science and Technology, Nanjing, China.*

## Abstract

Weather forecasting, particularly in short timeframes has been a longstanding challenge in meteorology, addressed in part by nowcasting methods. Leveraging radar data and innovative methodologies, nowcasting tools have evolved significantly, with open-source python platforms like Pysteps making is accessible to researchers to try advanced techniques. This review focus on Pysteps, a modular and user-friendly framework, offering optical flow based deterministic nowcasts and Short-Term Ensemble Prediction System (STEPS) ensemble nowcasts. Recent studies highlights its efficacy, including blending with Numerical Weather Prediction (NWP) models for improved performance beyond the nowcasting timeframe. Pysteps emerges as a versatile solution, facilitating both research innovation and operational forecasting needs, with wide range of input data and modularity.

**Keywords:** Pysteps; Blending; Optical Flow; Nowcasting; Open-source; Python package

## 1. Introduction

Forecasting the weather with precision and rapidly has always been a challenging endeavor in the ever-changing domain of meteorology. However, meteorologists have significantly advanced in minimizing the difference between present measurements and future estimates by implementing state-of-the-art equipment and innovative methodologies. Leading this work is the practice of nowcasting, a crucial element of meteorological research that is revolutionizing our comprehension of and ability to forecast imminent weather phenomena. Nowcasting in meteorology refers to the process of determining the present weather conditions and making short-term forecasts for a period of 0 to 6 hours. This is typically done by extrapolating data from surface weather stations, wind profiler data, and any other accessible weather data. It is possible to reasonably predict smaller local details, such as individual storms, with a certain degree of accuracy within this timeframe. Weather radar echoes and satellite data play a vital role in nowcasting due to their higher resolution and ability to continuously differentiate the size, shape, intensity, speed, and direction of movement of specific weather features.

Radar-based nowcasting emerged during a time when synoptic and mesoscale weather forecasting were dominant. Forecasting rainfall during that period was challenging for NWP models due to computational limitations that restricted the resolution of the models. While the models were able to accurately simulate mesoscale weather patterns such as fronts, they struggled to replicate smaller-scale convective patterns within these systems [1]. Radar extrapolation methods for nowcasting assume that the movement of a precipitation area is mostly affected by Lagrangian advection [1]. The efficacy of radar-based nowcasting, similar to other forecasting techniques, depends on several aspects such as meteorological conditions, geographical location, and spatial and temporal scales. The community of nowcasting experts promptly acknowledged the need of computing prediction uncertainty [2]. Probabilistic techniques in weather forecasting account for departures from pure advection by considering forecasts as predictions for a specific area rather than a single point. This may be achieved by either blurring or replacing the unpredictable scales with random noise, as suggested by Prudden [1].

* Corresponding author: Nasooh Ismail.

Nowcasting, the capacity to forecast imminent meteorological occurrences within a brief timeframe, is essential for several industries, such as aviation, agriculture, disaster management, and infrastructure development. In recent years, the area of nowcasting has been transformed by technological breakthroughs and data analytics. This has led to the emergence of many methodologies, such as extrapolation, optical flow methods, and machine learning algorithms. There is a diverse selection of open-source precipitation nowcasting platforms currently available. TrajGRU is a model developed by X.Shi in 2017 [3]. It is a Trajectory GRU (TrajGRU) model that has the ability to actively train the location-variant structure for recurrent connections. This is achieved by using a vast quantity of radar echo data. Rainymotion by Ayzel developed in 2019 [4] is a tool that combines optical flow algorithms with radar data. It was specifically created to serve as a benchmark for the advancement of precipitation nowcasting models. RainNet is a deep convolutional neural network developed also by Ayzel later in 2020 [5]. It is designed specifically for predicting future precipitation using large sets of radar data. MetNet developed by Sønderby in 2020 [6] utilizes radar and satellite data, to generate a probabilistic precipitation map using neural networks. LDCast is a latent diffusion model (LDM) developed by Leinonen in 2023 [7]. Not all of the currently available nowcasting software is referred and is limited to software with an active GitHub repository.

In this context, the presence of open-source software has made nowcasting tools more accessible to a wider audience, allowing both scholars and practitioners to use advanced approaches. Pysteps developed by Pulkkinen in 2019 [2] is a standout platform known for its versatility and user-friendly interface. It is particularly intended for probabilistic precipitation nowcasting. The complete pysteps suite may be effortlessly deployed in a Linux environment by using Conda. From an operational perspective, pysteps includes essential modules for importing data, preprocessing, conducting nowcasts, and verifying/plotting the results. An important benefit of this platform is its flexible modular structure, for example the 'io' module, which is the importer module can easily be modified to the user's needs. This module currently has importers for various radar datasets, but users may adapt it to their own requirements or develop new importers if necessary. Pysteps distinguishes itself from competing solutions by using a modular structure that places emphasis on simplicity, flexibility, and operational readiness, rather than relying significantly on machine learning methods. The documentation page of pysteps provides comprehensive information on the different modules, including detailed descriptions and example galleries. This review article discusses the commonly applied modules in pysteps and presents recent research based pysteps in the area of nowcasting.

## 2. Overview of Pysteps

Pysteps has two main aims. The primary objective is to offer a comprehensive and flexible framework for academics who wish to create innovative techniques for nowcasting and stochastic space-time modelling of precipitation, with a focus on providing thorough documentation and modularity[2]. Furthermore, it aims to provide a user-friendly and customizable platform for professionals in many domains such as meteorology and hydrology. Pysteps has the capacity to fulfil a vital function in comprehensive early warning systems for severe weather occurrences [2]. Pysteps is notable for its effective Python version of the Short-Term Ensemble Prediction System (STEPS), which is a probabilistic field-based nowcasting technique. It also includes the deterministic forerunner S-PROG [8, 9]. This approach takes into account the dynamic scaling of rainfall prediction by breaking down rainfall fields into a multiplicative cascade that represents different geographical scales. STEPS provides an ensemble of rainfall predictions by combining geographically and temporally associated stochastic disturbances into a deterministic extrapolation nowcast at each scale [10]. This methodology guarantees that larger-scale characteristics undergo a more gradual transformation compared to smaller-scale ones, thereby including the uncertainty associated with the increase and decrease of rainfall [10].

The workflow of Pysteps encompasses several key stages. As Imhoff [10] describes, the process for an ensemble nowcast begins by using the 'io' module to read input files and then the optical flow algorithms from the 'motion' module is used to calculate motion fields. Next, the motion field is advected using the 'extrapolation' module. The decomposition of these fields into spatial scales is then achieved using the 'cascade' module. The autoregressive (AR) parameters are then calculated and then applied in the time domain, together with the spatial cascades in the space domain. In order to incorporate the unpredictability in rainfall forecasts, stochastic noise is introduced into both the AR model and the advection fields through the use of the 'noise' module. The cascades are subsequently recomposed and iterated using the AR model, and then are again extrapolated. Ultimately, the generated ensembles undergo postprocessing to enhance the accuracy of the nowcasts through the application of the 'postprocessing' module. Pysteps additionally features the 'nowcasts' module, which integrates all the processes mentioned above for improved efficiency and easier use case.
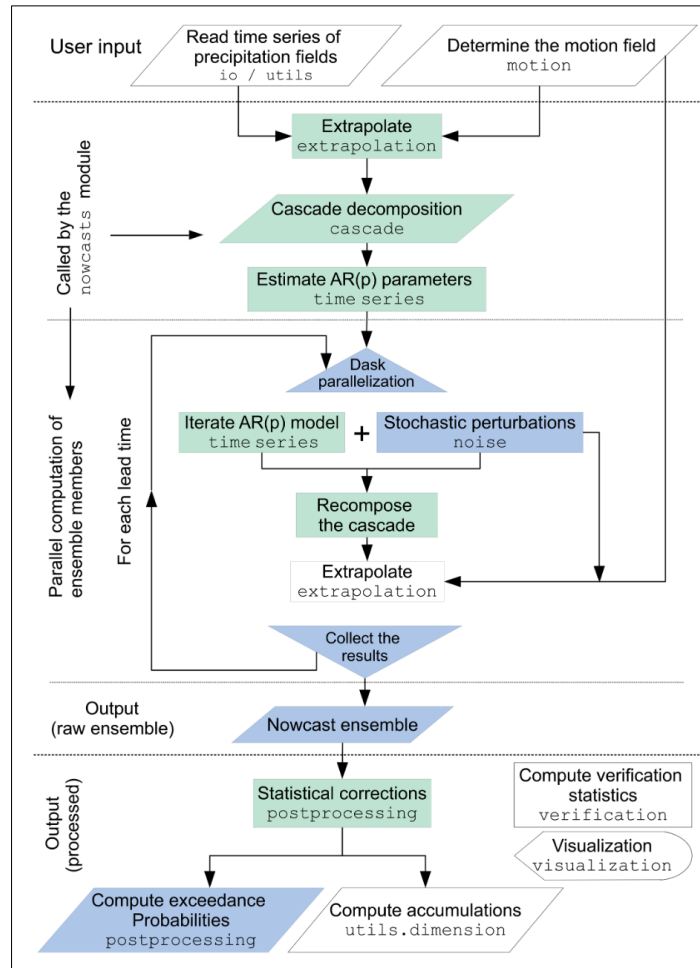
**Figure 1** Workflow of pysteps [2]

## 3. Recent research based on Pysteps

**Table 1** Non exhaustive list of modules available in pysteps library

| Module name | Function |
|---|---|
| io | Reading input data, can be used to read different radar and nc datasets |
| motion | Motion field computation using optical flow methods |
| extrapolation | Advection based extrapolation on radar echoes |
| timeseries | Methods and models for timeseries analysis |
| noise | Stochastic noise generators for perturbing motion fields and precipitation fields |
| cascade | The decomposition of precipitation fields based on spatial scales |
| nowcasts | Contains different nowcasting techniques to use |
| postprocessing | Statistical post processing |
| verification | Deterministic and ensemble verification |
| visualization | Plotting module |
| utils | Extra utility functions such as geomapping |

Pysteps primarily focuses on the probabilistic prediction of precipitation fields, however deterministic nowcasting is also offered as an alternative [2]. Due to its modular nature, Pysteps allows for the customization of different procedures. Pysteps is mainly centered on the S-PROG [11] and STEPS algorithm [8, 9]. Multiple modules provide a range of options for the user to choose from when generating a deterministic or probabilistic nowcast. The motion module comprises three optical flow techniques: a local Lucas-Kanade approach [12]. The Variational Echo Tracking (VET) algorithm, developed by Laroche & Zawadzki in 1994 [13], and the Dynamic and Adaptive Radar Tracking of Storms (DARTS) technique, proposed by Ruzanski & Chandrasekar, 2012 [14]. The extrapolation module incorporates the use of the semi-Lagrangian advection method, as outlined by Germann & Zawadzki [15], as well as Eulerian Persistence. The time-series module contains functions related to Auto-regressive (AR) models and correlation approaches. The 'nowcasts' module consists of many methodologies, such as extrapolation models including the advection-based LINDA model with autoregression, Lagrangian Probability, S-PROG, STEPS, and a localized version of STEPS known as short space ensemble prediction system (SSEPS) [2]. The subsequent paragraphs present recent research conducted utilizing certain modules of Pysteps.

## 3.1. Optical Flow based nowcasting

The optical flow approach in meteorology is a technique employed to analyze the motion of objects in satellite images or radar echo data. This method monitors the movement of patterns in sequential data frames. By examining the changes in these patterns over time, meteorologists may get a deeper understanding of and make more accurate predictions about weather propagation. This includes predicting the formation and movement of storms, the behavior of air fronts, and other significant weather. Pysteps library has several modules for optical flow methods to estimate motion fields. The library provides many techniques for optical flow analysis, such as the Lukas-Kanade [12] method, which use a pyramid-based approach to handle motion [19]. In addition, the Pysteps framework benefits from the use of global variational echo-tracking (VET) and spectral techniques, such as DARTS, which improve the precision and reliability of motion field estimates.

An essential aspect of any extrapolation-based nowcasting system is the determination of the advection field through optical flow. Pysteps provides a convenient way to assess the influence of the optical flow technique and scale filtering on the forecast skill, as demonstrated by original work by Pulkkinen [2]. The Lukas-Kanade (LK) method is a widely used optical flow technique in the pysteps library. It may be seen as a deterministic approach to rainfall nowcasting. The LK algorithm is an optical flow technique that makes the assumption that the eight neighboring pixels of a given pixel move in the same direction as that given pixel [12]. The approach begins by using the Shi-Tomasi corner identification algorithm [16] to identify the prominent elements in a picture, which are crucial for creating a motion field. The velocity field components are computed for each feature in order to build a sparse motion field. This sparse motion field is then interpolated onto the remaining areas of the picture, where there are no velocity vectors, to form a dense motion field. The Lukas Kanade approach is highly effective for situations when there is a large spatial extent and rainfall persistence is more significant than its beginning and expansion. In the original study, Pulkkinen [2] showed that when employing three optical flow methods (LK, VET, and DARTS), the LK and VET approaches generate smooth fields that are quite comparable, especially near the areas of precipitation. The author states that although the choice of the optical flow technique has a minor impact on the nowcast, the LK and DARTS approach has considerably reduced processing demands.

The study conducted by Imhoff in 2023 [17], examined 1533 instances of rainfall in the Dutch region. They used radar datasets from the Royal Netherlands meteorological institute (KNMI), specifically selecting data from 4 different event durations and 12 lowland catchments. The aim of the study was to assess the accuracy of nowcasting. The author of this article utilized Pysteps deterministic, and Pysteps probabilistic algorithms with 20 ensemble members. Pysteps deterministic case being Lukas-Kanade optical flow approach with a backward semi-Lagrangian advection method [15]. This method is similar to the S-PROG algorithm included in the pysteps package. In the pysteps deterministic experiment, the author discovered that there is excessive dissipation, resulting in the disappearance of intense rainfall centers. However, a general pattern of rainfall at a larger scale remains. Furthermore, it was observed that this method is most effective in predicting lead-times for events of longer duration.

Burton [18] used the LK technique from the pysteps library to perform convection nowcasting across west Africa using a satellite-derived product. This is the first instance of convection nowcasting being conducted in this manner. The author suggests that the optical flow method can be effective in predicting weather patterns with a lead time of 2 hours on a 10-kilometer scale and a lead time of 4 hours on larger scales of 200 km. This was especially beneficial for that particular region because of the limited radar coverage and the inability of NWP outputs to accurately predict convection at a local scale. The author's conclusion is that when LK method was used, highly accurate nowcasts can be

obtained when there is minimal initiation or growth of convection. Additionally, the author highlights the high feasibility of using the pysteps library as a tool for operational nowcasting of convection over west Africa.

Smith [19] subsequently employed a comparable approach, using brightness temperature (BT) data obtained from the Himawari-8/9 satellites over the maritime continent. The author employed optical flow algorithms from the Pysteps nowcasting library to produce deterministic and probabilistic nowcasts over the maritime continent. The deterministic predictions are made using the LK optical flow approach. The author's conclusion is that this technique demonstrates proficiency at geographical scales of 10 km and larger, up to 4 hours, and is superior to a persistence prediction for all lead times. While sharing similarities with the study conducted by [18], the author also acknowledges that the skill of the LK nowcasts diminishes when new convection develops, particularly in the afternoon.

### 3.2. Short Term Ensemble Prediction System (STEPS) based nowcasting

The relationship between the duration of precipitation and its size in space was explored by Seed [11] through the introduction of the Spectral Prognosis (S-PROG) model. This model served as the basis for the creation of the Short-Term Ensemble Prediction System (STEPS)[8, 9]. The STEPS nowcasting technique is a very renowned module that is integrated into the pysteps package. In the original literature, Pulkkinen [2] discuss the primary concept of decomposing the precipitation field using a fast Fourier transform (FFT). This decomposition results in a multiplicative cascade, where each cascade level represents a distinct spatial scale. These levels are then treated separately in the nowcasting model. This nowcasting approach aims to address the inherent uncertainty in the development and dissipation of new convection, which is a limitation of LK optical flow-based nowcasting.

In the study by Imhoff in 2020 [20], the author conducted probabilistic nowcasting in the Netherlands by using commercial microwave connections (CML) instead of radar datasets. They collected data from CMLs for a period of 12 summer days. The STEPS nowcast technique, which is available in the pysteps library, was used for the nowcasting process. This enhances the reputation of the pysteps open-source framework as a versatile tool capable of processing various data sets in addition to radar sets. In a separate study conducted by also by Imhoff in 2023 [10], the author employed the Pysteps probabilistic method, which involves using STEPS nowcasts, to assess its performance compared to the deterministic LK method. The study concludes that the Pysteps method is dependable for predicting rainfall amounts of ≥1.0 mm hr−1 up to a minimum of 120 minutes in advance. Nevertheless, the ensembles are considered to be underdispersed, and it is suggested that the dispersion of the ensemble should be broader. Despite the author still acknowledges that STEPS is a significant advancement in addressing these uncertainties in rainfall

The study conducted by Smith in 2023 [19] also compared the LK optical flow technique with the STEPS ensemble nowcast over the maritime continent. The study concluded that the probabilistic STEPS algorithm is capable of generating ensemble forecasts that are both accurate and reliable for 3-hour lead periods. The author also highlights that ensembles generated by STEPS exhibits an under-dispersive characteristic, which is attributed to the low variation among ensemble members and that the primary factor contributing to this variance is the divergence in the stochastic noise fields introduced by STEPS during the generation of nowcasts.

### 3.3. Blending (blending ensemble nowcasts with NWP models)

Nowcasting offers real-time data on present meteorological conditions, including radar data for precipitation, satellite imaging for cloud cover, and surface measurements for temperature and wind. NWP models employ intricate mathematical equations to replicate forthcoming atmospheric conditions by considering beginning states and physical principles. In order to achieve seamless forecasts beyond the nowcasting timeframe, it is essential to blend these two products [21].

Blending these two sources entails merging the advantages of both techniques. The nowcast data, characterized by its exceptional resolution and immediate updates, has the capability to rectify short-term inaccuracies in NWP models and offer specific information that would be overlooked by more generalized models. NWP models, on the other hand, provide a structured approach for anticipating weather patterns on a broader scale and may generate predictions that extend beyond the scope of nowcasting. Pysteps enables the blending of nowcasts and Numerical Weather Prediction (NWP) models, using their 'blending' module to improve prediction accuracy. This results in a high-quality blended product that extends the lead time beyond the nowcast timeframe for important weather events and improves decision-making.

After the release of pysteps in 2019, Imhoff [10] introduced a significant update to the platform. This update involved a new method of blending implementation, which combines the extrapolation nowcast with NWP and stochastic noise components. The blending is done at different spatial scales, with varying blending weights per cascade level. The

fundamental idea entails breaking down the distribution of rainfall into a combination of many spatial cascades with each cascades exhibiting characteristics at a particular scale. Each step of evolution is examined and treated separately. This decomposition uses both NWP and radar extrapolation fields, enhanced by correlated noise with distinct characteristics. STEPS blends these three decomposed domains with different weights depending on the size and lead time. After the blending process is completed, the levels are mixed again to provide a final product [10]. The author has provided a comprehensive explanation of how this new approach has been included into the pysteps framework, which serves as more evidence of the expanding and enhancing pysteps community. This novel approach primarily demonstrates comparable or superior performance than only relying on nowcasting and provides additional benefits compared to NWP during the initial hours of the forecast. Imhoff [10] also states that the open-source blending strategy in pysteps can serve as a foundation for future implementations of various blending approaches and collaborations.

## 4. Conclusion

In summary, Pysteps emerges as a versatile and powerful tool for probabilistic precipitation nowcasting, offering a unique blend of flexibility, accessibility, and performance. The easy-to-use interface entices individual researchers or even forecasting communities to use this library where not only radar fields, but other data such as satellite imagery and CLM data can be used for operational nowcasting. Its modular design and open-source nature not only foster innovation within the research community but also empower operational users with reliable forecasting capabilities.

## Compliance with ethical standards

### Disclosure of conflict of interest

The author affirm that there is no conflict of interest in this mini-review article. The article is aimed at providing an insight to Pysteps platform and its versatility.

## References

[1] Prudden, R., Adams, S., Kangin, D., Robinson, N., Ravuri, S., Mohamed, S., & Arribas, A. (2020, May 11). A review of radar-based nowcasting of precipitation and applicable machine learning techniques. arXiv. Retrieved from http://arxiv.org/abs/2005.04988

[2] Pulkkinen, S., Nerini, D., Pérez Hortal, A. A., Velasco-Forero, C., Seed, A., Germann, U., & Foresti, L. (2019). Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geoscientific Model Development*, *12*(10), 4185–4219. https://doi.org/10.5194/gmd-12-4185-2019

[3] Shi, X., Gao, Z., Lausen, L., Wang, H., Yeung, D.-Y., Wong, W., & Woo, W. (2017). Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model. Presented at the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

[4] Ayzel, G., Heistermann, M., & Winterrath, T. (2019). Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0.1). *Geoscientific Model Development*, *12*(4), 1387–1402. https://doi.org/10.5194/gmd-12-1387-2019

[5] Ayzel, G., Scheffer, T., & Heistermann, M. (2020). RainNet v1.0: a convolutional neural network for radar-based precipitation nowcasting. *Geoscientific Model Development*, *13*(6), 2631–2644. https://doi.org/10.5194/gmd-13-2631-2020

[6] Sønderby, C. K., Espeholt, L., Heek, J., Dehghani, M., Oliver, A., Salimans, T., ... Kalchbrenner, N. (2020, March 30). MetNet: A Neural Weather Model for Precipitation Forecasting. arXiv. Retrieved from http://arxiv.org/abs/2003.12140

[7] Leinonen, J., Hamann, U., Nerini, D., Germann, U., & Franch, G. (2023, April 25). Latent diffusion models for generative precipitation nowcasting with accurate uncertainty quantification. arXiv. Retrieved from http://arxiv.org/abs/2304.12891

[8] Bowler, N. E., Pierce, C. E., & Seed, A. W. (2006). STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quarterly Journal of the Royal Meteorological Society*, *132*(620), 2127–2155. https://doi.org/10.1256/qj.04.100

[9] Seed, A. W., Pierce, C. E., & Norman, K. (2013). Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme: Formulation of a Scale Decomposition Nowcast Scheme. *Water Resources Research*, *49*(10), 6624–6641. https://doi.org/10.1002/wrcr.20536

[10] Imhoff, R. O., De Cruz, L., Dewettinck, W., Brauer, C. C., Uijlenhoet, R., Van Heeringen, K., … Weerts, A. H. (2023). Scale-dependent blending of ensemble rainfall nowcasts and numerical weather prediction in the open-source pysteps library. *Quarterly Journal of the Royal Meteorological Society*, *149*(753), 1335–1364. https://doi.org/10.1002/qj.4461

[11] Seed, A. W. (2003). A Dynamic and Spatial Scaling Approach to Advection Forecasting. *Journal of Applied Meteorology*, *42*(3), 381–388. https://doi.org/10.1175/1520-0450(2003)042<0381:ADASSA>2.0.CO;2

[12] Lucas, B. D., & Kanade, T. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision.

[13] Laroche, S., & Zawadzki, I. (1994). A Variational Analysis Method for Retrieval of Three-Dimensional Wind Field from Single-Doppler Radar Data. *Journal of the Atmospheric Sciences*, *51*(18), 2664–2682. https://doi.org/10.1175/1520-0469(1994)051<2664:AVAMFR>2.0.CO;2

[14] Ruzanski, E., & Chandrasekar, V. (2012). An Investigation of the Short-Term Predictability of Precipitation Using High-Resolution Composite Radar Observations. *Journal of Applied Meteorology and Climatology*, *51*(5), 912–925. https://doi.org/10.1175/JAMC-D-11-069.1

[15] Germann, U., & Zawadzki, I. (2002). Scale-Dependence of the Predictability of Precipitation from Continental Radar Images. Part I: Description of the Methodology. *Monthly Weather Review*, *130*(12), 2859–2873. https://doi.org/10.1175/1520-0493(2002)130<2859:SDOTPO>2.0.CO;2

[16] Shi, J., & Tomasi, C. (1994). Good features to track. Presented at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR94), Seattle.

[17] Imhoff, R. O., Brauer, C. C., Overeem, A., Weerts, A. H., & Uijlenhoet, R. (2020). Spatial and Temporal Evaluation of Radar Rainfall Nowcasting Techniques on 1,533 Events. *Water Resources Research*, *56*(8), e2019WR026723. https://doi.org/10.1029/2019WR026723

[18] Burton, R. R., Blyth, A. M., Cui, Z., Groves, J., Lamptey, B. L., Fletcher, J. K., … Roberts, A. (2022). Satellite-Based Nowcasting of West African Mesoscale Storms Has Skill at up to 4-h Lead Time. *Weather and Forecasting*, *37*(4), 445–455. https://doi.org/10.1175/WAF-D-21-0051.1

[19] Smith, J. A., Birch, C., Marsham, J., Peatman, S., Bollasina, M., & Pankiewicz, G. (2023). *Applying pySTEPS optical flow algorithms to improve convection nowcasting over the Maritime Continent* (preprint). Atmospheric, Meteorological and Climatological Hazards. https://doi.org/10.5194/egusphere-2023-1404

[20] Imhoff, R. O., Overeem, A., Brauer, C. C., Leijnse, H., Weerts, A. H., & Uijlenhoet, R. (2020). Rainfall Nowcasting Using Commercial Microwave Links. *Geophysical Research Letters*, *47*(19), e2020GL089365. https://doi.org/10.1029/2020GL089365

[21] Sun, J., Xue, M., Wilson, J. W., Zawadzki, I., Ballard, S. P., Onvlee-Hooimeyer, J., … Pinto, J. (2014). Use of NWP for Nowcasting Convective Precipitation: Recent Progress and Challenges. *Bulletin of the American Meteorological Society*, *95*(3), 409–426. https://doi.org/10.1175/BAMS-D-11-00263.1