~~Applying pySTEPS optical flow algorithms to improve convection nowcasting over the Maritime Continent~~Evaluating satellite-based convection nowcasting over the Maritime Continent using pySTEPS optical flow algorithms

Joseph Smith[1], Cathryn Birch[1], John Marsham[1], Simon Peatman[1], Massimo Bollasina[2], George Pankiewicz[3]

[1]School of Earth and Environment, University of Leeds, Leeds, LS2 9JT
[2]School of GeoSciences, University of Edinburgh, Edinburgh, EH8 9YL
[3]UK Met Office, Exeter, EX1 3PB

Correspondence to: Joseph Smith (eejasm@leeds.ac.uk)

Abstract. The Maritime Continent (MC) regularly experiences powerful convective storms that produce intense rainfall, flooding and landslides, which numerical weather prediction models struggle to forecast. Nowcasting uses observations to make more accurate predictions of convective activity over short timescales (~0-6 hours). Optical flow algorithms are effective nowcasting methods as they are able to accurately track clouds across observed image series and predict forward trajectories. Optical flow is generally applied to weather radar observations, however, the radar coverage network over the MC is not complete and the signal cannot penetrate the high mountainous regions. In this research, we apply optical flow algorithms from the pySTEPS nowcasting library to satellite imagery to generate both deterministic and probabilistic nowcasts over the MC. The deterministic algorithm shows skill up to 4 hours on spatial scales of 10 km and coarser, and outperforms a persistence nowcast~~forecast~~ for all lead times. Lowest skill is observed over the mountainous regions during the early afternoon and highest skill is seen during the night over the sea. A key feature of the probabilistic algorithm is its attempt to reduce uncertainty in the lifetime of small scale convection. Composite analysis of 3-hour lead time nowcasts, initialised in the morning and afternoon, produces reliable ensembles but with an under-dispersive distribution, and produced area under the curve scores (i.e. ratio of hit rate to false alarm rate across all probability thresholds) of 0.80 and 0.71 over the sea and land, respectively. When directly comparing the two approaches, the probabilistic nowcast shows greater skill at ≤ 60 km spatial scales, whereas the deterministic nowcast shows greater skill at larger spatial scales ~200 km. Overall, the results show promise for the use of pySTEPS and satellite retrievals as an operational nowcasting tool over the MC.

Keywords: Nowcasting, Maritime Continent, convection, pySTEPS, satellite data

1 Introduction

The "Early Warnings For All Initiative" was launched by the United Nations in November 2022 and calls for the whole world to be covered by early warning systems by the end of 2027 ("World Meteorological Organization," 2023). It focuses on poorer countries in Asia, Africa, South and Central America and the Pacific, and motivates the development of early warning weather systems for these regions.

The Maritime Continent (MC) is a region of Southeast Asia that includes the countries of Indonesia, Malaysia, Philippines, Papua New Guinea, Brunei and East Timor. It is a complex mix of land and ocean with major islands such as Sumatra, Java, Borneo and New Guinea making it the largest archipelago on Earth (Figure 1a). It is also one of the wettest places on Earth with its complex topography and location across the equator making it a hotspot for extreme weather. The region often experiences natural disasters such as flooding and landslides that have disastrous effects on already very poor areas. Easterly trade winds blow warm water across the Pacific into the

MC creating a 'warm pool' around the region (Dayem et al., 2007), which, when combined with its proximity to the equator and the inter-tropical convergence zone, provides favourable conditions for deep convection. The large amounts of latent heat released from this convection means that the region is often referred to as the 'boiler box' of the tropics as it plays a crucial role in contributing to the global atmospheric circulations (the Hadley and Walker Cells), in turn affecting both local and global weather systems (Ramage, 1968).

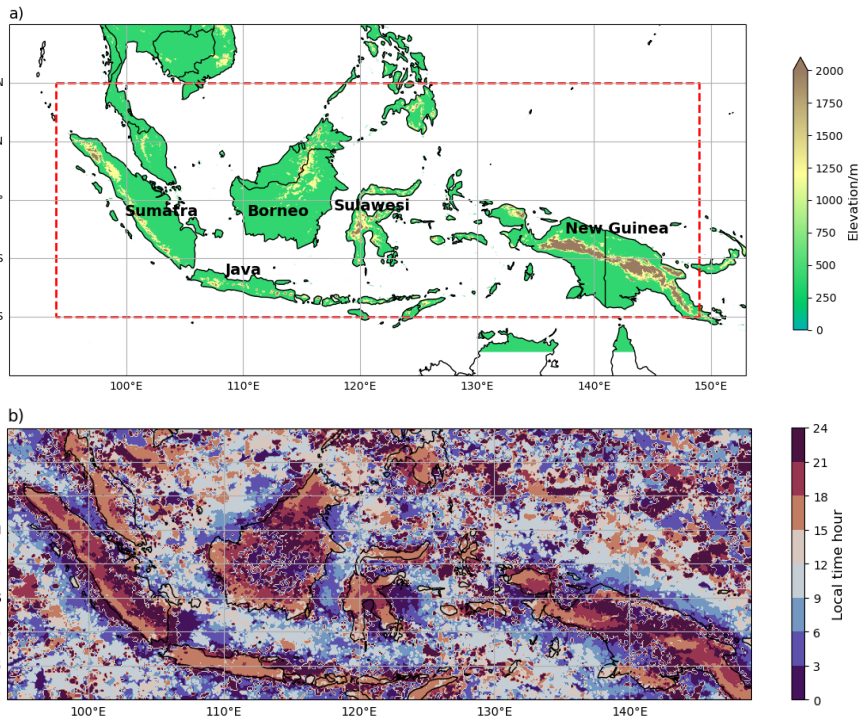



Figure 1. a) Orographic map of the Maritime Continent showing the domain over which the nowcasts were generated and the domain over which they were evaluated (red dotted line). ~~Maps of the selected islands within the MC showing a) the orography and~~ b) ~~Tt~~he diurnal cycle of peak rainfall within the evaluation region (interpolated to local solar time) using the Global Precipitation Measurement dataset (Hou et al., 2014) from December, January and February 2001 – 2020.

The MC's strong diurnal cycle (shown in Figure 1b by the spatial variation in timings of peak rainfall) is one of its dominant drivers of convective activity (Yamanaka, 2016). It has typical characteristics of most diurnal cycles across the tropics (Yang and Slingo, 2001), starting with peak solar insolation around midday. This starts to form a land-sea temperature contrast due to the lower heat capacity of the land. A sea breeze then develops blowing on land, often triggering convection which builds into the late afternoon and evening. Many islands in the MC also contain mountains close to the coast with altitudes of over 2000 m (e.g. Sumatra). Orographic lifting driven by the mountains can further enhance the convection (Mori et al., 2004). Into the late evening and overnight this convection propagates offshore until the early morning the following day, leaving clear skies over land in the morning for strong solar insolation to restart the process.

Numerical weather prediction (NWP) models struggle to represent the moist convection that dominates weather in the MC, with coarser resolution models often initiating convection too early in the day (Porson et al., 2019) or under-estimating the amount of rainfall (Qian, 2008). Ferrett et al., (2021) shows that the skill of ensemble forecasts from a higher-resolution, convective-scale configuration of the Met Office Unified Model, with 4.5 km horizontal grid spacing over Indonesia, only start to show skill (during the first day after initialisation) when coarsened up to spatial scales of ~150 km. Mesoscale convective systems are defined as having spatial scales of

at least ~100 km and so convective-scale models cannot be relied upon to skilfully resolve impactful storms over the MC.

90 Nowcasting is the process of obtaining current observations of the atmosphere and using them to generate rapid, short-term (typically ~0-6 hours ahead) predictions of the future atmospheric state (Roberts et al., 2022). It requires real time observations (e.g. from weather radar) as an input and the application of predictive techniques to forward propagate these observations. Unlike NWP models, nowcasting tools do not use large sets of complex numerical equations in order to model the atmosphere. Instead, they use cutting edge computational techniques

95 such as optical flow and artificial intelligence algorithms, enabling them to generate useful output at near-instantaneous timescales (Ayzel et al., 2020; Han et al., 2019; Gijben and de Coning, 2017).

Currently, there are a number of state-of-the-art nowcasting systems in operation around the world, typically based in developed countries, which take advantage of large weather radar networks to provide near-instant, wide spread data coverage of precipitation. There are, of course, many operational systems in use globally but Table 1 covers

100 some of the most advanced ones.

Nowcasting system	Input sources	Region of application	Reference
Short-Term Ensemble Prediction System (STEPS)	Weather radar, NWP	UK	(Bowler et al., 2006)
Integrated Nowcasting System through Comprehensive Analysis (INCA)	Weather radar, NWP, satellite, surface station observations	Europe (Alpine regions)	(Haiden et al., 2011)
Short-range Warning of Intense Rainstorms in Localised Systems (SWIRLS)	Weather radar, NWP	China	(Srivastava et al., 2021)
Auto-Nowcast System (ANC)	Weather radar, NWP, satellite, surface station observations, wind profiler, atmospheric sounding, lightning detector	US	(Mueller et al., 2003)
McGill Algorithm for Precipitation nowcasting using Lagrangian Extrapolation (MAPLE)	Weather radar	Canada, US	(Germann and Zawadzki, 2002)
Spectral-Prognosis (S-PROG)	Weather radar	Australia	(Seed, 2003)
Global Synthetic Weather Radar (GSWR)	Satellite, lightning, NWP	US	(Reen et al., 2020)
Support to Nowcasting and Very Short Range Forecasting (NWC SAF)	Satellite	Europe	(Marcos, 2015)

Table 1. Information on some of the state-of-the-art nowcasting systems that are currently in operational use around the world.

Weather radar networks are expensive to implement, maintain and often not suitable for regions with mountainous

105 terrain (e.g. the MC), meaning the types of nowcasting systems listed in Table 1 cannot always be implemented. There is, therefore, a widespread interest within nowcasting research in the use of satellite data as the main source of input, especially in the tropics, which can provide constant, widespread coverage of the Earth's atmosphere from space. The advancement of satellite technology in recent years has given us access to data on increasingly higher spatial and temporal resolutions (e.g. Line et al., (2016) use 1-minute retrievals of 1 km resolution imagery

110 for forecasting), allowing finer detail of cloud structures to be observed and more accurate tracking of weather

systems (Sieglaff et al., 2013). This provides the basis for extrapolation nowcasting methods that use the tracked history of weather features (e.g. storms) to calculate motion vectors, which the features are then propagated along to create future predictions (Burton et al., 2022; Vila et al., 2008; Line et al., 2016). The vast volume of satellite data also makes nowcasting a suitable candidate for the application of machine learning methods to make future predictions of the atmosphere. Most commonly, studies have trained machine learning models to take in consecutive satellite images as input and then output (nowcast) the future consecutive images (Lebedev et al., 2019; Lagerquist et al., 2021).

The MC itself has received little attention in the field of nowcasting, despite the region experiencing regular intense convective activity affecting the lives of millions. The Indonesian network consists of 42 weather radars (Permana et al., 2019) but is sparse relative to the size of the country, the country is highly mountainous, and experiences communication issues between sites, meaning real-time full radar coverage of the region is not possible (Permana et al., 2019). One of the radars within the network was used by Ali et al., (2021) to nowcast two rainfall events over southern Borneo. On the other hand, satellite data was used by Harjupa et al., (2022) to apply The Rapidly Developing Cumulus Area algorithm (Sobajima, 2012) to a region of western Java to predict heavy rainfall for 77 events. The limited sample size and domain of these studies makes it difficult to understand how effective the methods are for other regions of the MC. There is, therefore, a need to test nowcasting tools that can be applied and evaluated across the entire MC domain.

pySTEPs (Pulkkinen et al., 2019) is a free, open-source Python library that provides modules for a variety of optical flow-based nowcasting methods (see section 2.1a for optical flow description). The library is designed for use on radar data and has been used to show skilful prediction of stratiform precipitation in the mid-latitudes (Han et al., 2022; Imhoff et al., 2020). To the best of the authors' knowledge, the only study that has applied pySTEPs to satellite data over the tropics is Burton et al., (2022), who produced nowcast skill up to a 4 hour lead time over West Africa. It is this result that motivates the application of pySTEPS to the MC.

This paper presents the evaluation of both deterministic and probabilistic nowcasts produced by applying pySTEPS to satellite data over the MC. The aim is to highlight their strengths and weaknesses and demonstrate their potential use as an operational nowcasting system.

2 Data and Methods

2.1 Data

This study uses brightness temperature (BT; the temperature a black body would need in order to emit the radiance detected by a satellite) data from the Himawari-8/9 satellites as input to the nowcasting algorithms. Himawari-8 and -9 are passive geostationary satellites with 16 band channels ranging from 0.47 µm to 13.3 µm, covering parts of the visible, near-infrared and infrared (IR) spectrum (Bessho et al., 2016). Hourly BT data from channel 13 on-board Himawari-8/9 has been used, which detects IR radiation with a wavelength of 10.4 µm. The full-disc images are transformed to a Cartesian grid with a grid spacing of 2 km, using the gdalwarp command from the Geospatial Data Abstraction Library (Rouault et al., 2023). Convective clouds can be clearly identified in BT maps as they have cold tops relative to the surface of the Earth. The data is selected from the December, January and February (DJF) season, which is the peak season for convection over the MC (Birch et al., 2016), for five seasons from 2015/16 – 2019/20 (data retrievals of Himawari-8 began in 2015 (Bessho et al., 2016)).

Initialisation time (LT)	Number of nowcasts
0000	441
0300	422
0600	422
0900	441
1200	440
1500	429
1800	441

Initialisation time (LT)	Number of nowcasts
0000	441
0300	422
0600	422
0900	441
1200	440
1500	439
2100	440

Table 2. The number of nowcasts that were produced at each initialisation time throughout the day.

In this study, nowcasts were produced using three BT maps as input, spaced evenly apart by 1 hour, starting with the latest observation. 3,476 nowcasts were produced for initialisation times every 3 hours from 0000 LT to 2100 LT to incorporate the diurnal variability of weather over the MC, with the number of nowcasts at each initialisation time shown in Table 2. In order to avoid issues of new convection entering the edge of the domain, which cannot be reproduced (optical flow can only propagate convection that exists in the domain at the nowcast initialisation time), the nowcasts were first produced using BT data on a 15°S – 15°N, 90°E – 153°E domain, and then evaluated on a 10°S – 10°N, 94°E – 149°E domain, which still includes the major islands of the MC (Figure 1a).

2.2 Methods

2.2.1 pySTEPS/optical flow

pySTEPS provides a well-documented framework that allows users to employ optical flow algorithms for both deterministic and probabilistic nowcasting approaches, as well as a range of verification techniques. Optical flow is a computer vision technique that generates velocity fields to describe the apparent motion of objects across consecutive images (Horn and Schunck, 1981). The key assumption of optical flow is that each pixel intensity remains constant across all images as it is advected. Given $I(x, y, t)$ is the intensity of a pixel at time $t = 0$, this results in:

$$I(x, y, t) = I(x, y, t + \Delta t) \,, \tag{1}$$

where Δt is the time between image frames. Applying a Taylor series expansion to equation (1) leads to:

$$\frac{\partial I}{\partial x} U + \frac{\partial I}{\partial y} V + \frac{\partial I}{\partial t} = 0 \,, \tag{2}$$

where $U = \frac{dx}{dt}$ and $V = \frac{dy}{dt}$ are the velocity components of the motion field. (2) is known as the optical flow equation. $\frac{\partial I}{\partial x}$, $\frac{\partial I}{\partial y}$ and $\frac{\partial I}{\partial t}$ can be calculated as they represent the image gradients over space and time, whereas U and V are unknown meaning (2) represents an underdetermined system that cannot be solved directly. Optical flow methods attempt to get round this by applying various spatial constraints to U and V. Section b) and c) will describe the two optical flow methods within pySTEPS that were used in this study.

2.2.2 Lucas-Kanade deterministic algorithm

The Lucas-Kanade algorithm (LK; Lucas and Kanade, 1981) is an optical flow method that assumes, for a given pixel, the eight immediately surrounding pixels move along with that given pixel. This assumption results in nine separate versions of (2) (eight from the surrounding pixels and one from the given pixel itself), representing an overdetermined system. A least squares fit method is then applied to the nine equations to obtain the optimum solution for the given pixel. To create a motion field the algorithm first identifies the key features within an image

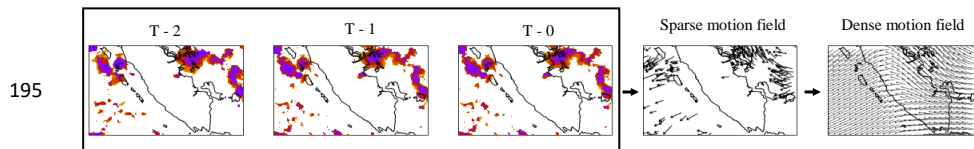



Figure 2. An example (using BT over Sumatra) of how a motion field is generated using the LK method. Features are identified and tracked across the three input images (T-0, T-1 and T-2) to generate the sparse motion field. The sparse motion field is then interpolated onto the rest of the domain to produce the dense motion field.

by using the Shi-Tomasi corner detection algorithm (Jianbo Shi and Tomasi, 1994). The velocity field components are then calculated for each feature within the image (using the LK assumption) to create a sparse motion field, which is then interpolated onto the rest of the image (where there are no velocity vectors) to generate a dense motion field. Once the motion field has been determined, the latest observation needs to be advected along the motion field. pySTEPS implements the backward-in-time semi-Lagrangian advection scheme (Germann and Zawadzki, 2002). The nowcast is the resulting advected field. In this study, three consecutive images (the current BT observation and the two images prior) are inputted into the LK algorithm, as shown in Figure 2. Sparse motion vectors are generated from each successive image pair (T-2/T-1 and T-1/T-0) and then combined together onto one field (sparse motion field). If a pixel in the sparse motion field has two motion vectors associated with it, they are averaged together to produce one motion vector.

The LK algorithm is a deterministic optical flow nowcasting approach that describes the evolution of a field (in this study BT fields) by moving current observations along motion fields. However, this simplistic approach also means it is unable to predict the initiation/growth/decay (IGD) of convection within BT fields.

2.2.3 Short Term Ensemble Prediction System algorithm

The pySTEPS library also contains modules for more advanced, probabilistic nowcasting approaches that attempt to address the IGD problem. The Short Term Ensemble Prediction System (STEPS; Bowler et al., 2006) was jointly developed by the Met Office, UK and the Bureau of Meteorology Research Centre, Australia, and aims to address the issue of unpredictability in the lifetime of convection by injecting fields of varying stochastic noise. It does this by applying a fast Fourier Transform to the current BT field (T-0) to decompose it into cascades of different length scales. Varying intensities of Gaussian noise fields are then injected into each cascade field depending on the length scale. Cascades containing the small length scale features will receive greater intensity of noise injection, as these features represent the greatest uncertainty in growth and decay of convection. In contrast, the large length scale features receive a lower intensity of noise, as these features represent the least uncertainty in growth and decay. The cascades are then recomposed to produce the new BT field, which is ready for extrapolation. In this work the motion field for extrapolation is generated using the LK algorithm (as in Figure 2, by using T-0, T-1 and T-2 BT fields). Stochastic noise perturbations are also applied to the motion fields to try to capture the uncertainty in the extrapolation of the BT fields. The magnitude of the perturbation increases with respect to lead time as the motion field increases in uncertainty. Finally, the new BT field is extrapolated along the motion field to create one ensemble member of the nowcast. Ensemble members are generated by using new realizations of the noise perturbations to create multiple versions of the nowcast.

2.2.4 Verification methods

The stochastic nature of convection in the MC makes it extremely challenging to nowcast the precise location (pixel-to-pixel) of convective activity. When evaluating a nowcasts' ability to predict convection on a pixel-to-pixel basis, the nowcast may be broadly correct but slightly misaligned in location. If we simply take the difference between the nowcast and the verification, this leads to the Double Penalty Problem: firstly, the model is penalised for a miss and secondly it is penalised for a false alarm in the slightly misaligned location. To overcome this

problem Roberts and Lean (2008) developed a method known as the Fractional Skill Score (FSS), which enables a forecast to be verified on a range of spatial scales as opposed to a pixel-by-pixel basis, allowing leeway for minor misalignments. The FSS firstly creates two binary fields from the nowcast field and the observation field by using a threshold value of 235 K ~~(this value was used to try and include the entirety of the convective system(Roca et al., 2017) (Feng et al., 2021)(Machado and Laurent, 2004)))~~ - any pixel with a value below this is set to 1 and any pixel with a value above this is set to 0. This threshold was chosen based on previous convection tracking studies in the tropics (most of which use a threshold in the range 233 K to 241 K) and aims to include the entirety of the convective system in the BT image (Goyens et al., 2012;) (Fiolleau and Roca, 2013;) (Roca et al., 2017;) Feng et al., 2021). A $n \times n$ kernel is then convolved with both binary fields, where n is the desired spatial scale set by the user, and the fraction of pixels within the kernel that have a value of 1 is calculated. The mean squared error (MSE) between the fraction of 1's in the observation kernel, $O_{(n)}$, and the fraction of 1's in the nowcast kernel, $M_{(n)}$, is then calculated:

$$MSE_{(n)} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [O_{(n)i,j} - M_{(n)i,j}]^2 , \qquad (3)$$

where N_x and N_y are the number of pixels in the longitude and latitude direction. Because $MSE_{(n)}$ is highly dependent upon the frequency of the event it must be compared to the MSE of a relatively low-skill reference nowcast in order to provide any usefulness, which is defined in Murphy and Epstein, 1989 by:

$$MSE_{(n)ref} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [O_{(n)i,j}^2 + M_{(n)i,j}^2] . \qquad (4)$$

The final FSS is then calculated as:

$$FSS = 1 - \frac{MSE_{(n)}}{MSE_{(n)\,ref}} . \qquad (5)$$

The nowcast can be evaluated at different spatial scales by changing the value of n. In this study 10 km, 20 km, 60 km, 100 km and 200 km were chosen as the spatial scales. This range of scales allows a nowcast to be evaluated in its ability to predict convection on a range of scales. An FSS of 1 can be interpreted as a perfect score whereas an FSS of 0 can be interpreted as a nowcast with no skill. A threshold value for FSS above which a nowcast is useful is given by:

$$FSS_{(useful)} \geq 0.5 + \frac{f}{2} , \qquad (6)$$

where f is the fractional coverage of pixels with a value of 1 over the entire domain. As f becomes small then $FSS_{(useful)}$ can be approximated by:

$$FSS_{(useful)} \geq 0.5 . \qquad (7)$$

This is the basic approach to the FSS, which calculates a single score for each nowcast to describe the skill over the whole domain. However, often the skill of a nowcast will vary across the domain due to differences in environments (e.g. land, sea and mountains) and the different interactions that result from these changing environments. Woodhams et al., (2018) developed an adapted version of FSS known as the localised fractional skill score (LFSS), which enables the skill of a nowcast to be evaluated at each pixel across the domain, resulting in a spatial map of FSSs. The LFSS is calculated by adapting (3) and (4) so that instead of dividing the sum of the squared error between $O_{(n)}$ and $M_{(n)}$ over the spatial domain, it is divided over the time domain (i.e. the number of time steps). This results in replacing (3) and (4) with,

$$MSE_{(n)} = \frac{1}{N_t} \sum_{j=1}^{N_t} [O_{(n)t} - M_{(n)t}]^2 , \qquad (8)$$

and,

$$MSE_{(n)ref} = \frac{1}{N_t} \sum_{j=1}^{N_t} [O_{(n)t}^2 + M_{(n)t}^2] , \qquad (9)$$

where N_t is the number of time steps.

The skill of a given ensemble nowcast produced by the STEPS algorithm is evaluated by generating the probabilistic nowcast from the ensemble members and then comparing it to the observations at different probability thresholds. At each probability threshold, any pixel in the probabilistic nowcast with a value equal to or greater than the threshold is assigned a value of 1 and any pixel with a value less than the threshold is assigned a value of 0 (creating a binary field). The observation field is then converted into a binary field in the same way except using a threshold value of 235 K. The two binary fields are compared at each corresponding pixel to obtain the number of hits (both pixel values equal 1), misses (nowcast pixel value is equal to 0 but observation pixel value is equal to 1), false alarms (nowcast pixel value is equal to 1 but observation pixel value is equal to 0) and correct negatives (both pixel values equal 0). These metrics are then used to calculate the probability of detection (POD) hit rate [hits/(hits + misses)] and false alarm ratethe probability of false detection (POFD) [false alarms/(correct negatives + false alarms)]. This is repeated at multiple probability thresholds and the PODshit rates are plotted against the POFDs false alarm rates to produce a receiver operating characteristic (ROC) curve. The nowcasts with highest skill will minimise the false alarm ratePOFD and maximise the PODhit rate, resulting in a ROC curve in the top left corner of the diagram with a large area under the curve (AUC) score.

Reliability diagrams are used to evaluate how well STEPS's probabilistic predictions compare to the actual observed frequency of events (in this work an event counts as a pixel with a value less than or equal to 235 K). For a given ensemble forecast, the predicted event probabilities are first evenly binned, creating a sub-group of nowcasts for each bin. The frequency for which events are observed is then calculated for each sub-group. The mean event probability within each bin is plotted against the observed event frequency to create a reliability diagram. A perfectly reliable nowcast will predict event likelihoods consistent with the observed frequency i.e. a diagonal X = Y line.

In order to gain a useful understanding of the optical flow nowcast skill, a persistence nowcast is used for comparison. A persistence nowcast is considered as having the baseline of minimum skill and is produced by using the latest observation as the next prediction i.e. it assumes that the current weather will persist and be identical at the next nowcast lead time.

3 Results

3.1 Deterministic nowcasting - Lucas-Kanade algorithm

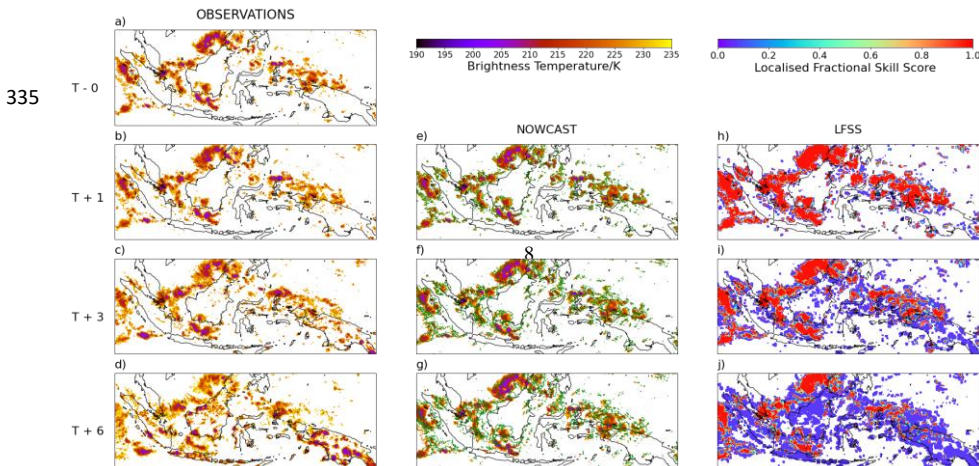
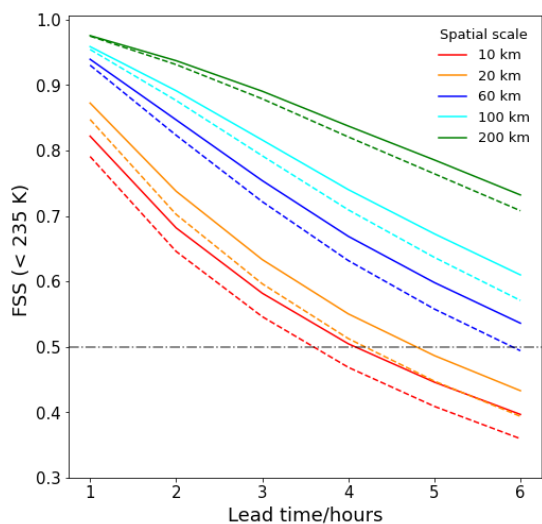
Figure 3e–g shows an example of a nowcast (each lead time is produced using the same T-0, T-1 and T-2 observations) produced by the LK algorithm for a qualitative assessment of the skill against observations (Figure 3a–d), whilst Figure 3h–j provides the LFSS (evaluated on a 20 km scale to show clearly defined differences in skill) at each timestep of the nowcast (where $N_t = 1$) for a quantitative assessment. This particular set of observations contains convection on a range of scales, with regions of propagation and regions of initiation, providing a good example to evaluate the LK algorithm on a range of capabilities (for this reason the same example is used throughout the paper). At T-0, the observed organised, large scale convection e.g. north of Borneo, approximately maintains its shape through T+1 and T+3 and then starts to change in structure at T+6 e.g. east of Sumatra. There is also the development of relatively smaller scale convection observed during the T+3 and T+6 hour lead times e.g. over New Guinea.

Visually, the deterministic LK approach appears to predict the propagation of large scale, organised convection
well. The T+1 nowcast best resembles the corresponding observation due to the least amount of new convection
developing during this time, as well as little propagation of the organised convection (the nowcast at T+1 closely
matches the persistence nowcast). This is seen in Figure 3h, which shows high skill in the organised convective
regions over the majority of the domain. As the lead time increases, westward propagation of the large scale
convection is observed, which appears to be effectively tracked by the LK nowcast at T+3 e.g. east of Sumatra.
This is confirmed in Figure 3i, which shows the skill of the nowcast at T+3 remaining high over the regions of
organised convection. There is, however, a clear increase in areas of low skill at T+3 and T+6 due to the LK
nowcast being unable to reproduce the IGD of convection. At T+6, the change in structure of convection also
contributes to the majority of the domain experiencing low skill. The nowcast at this lead time shows the least
resemblance to the observations with only the largest regions of predicted convection providing any skill (Figure
3j).

Figure 3 highlights some of the key advantages and disadvantages of the LK algorithm. Overall, it does well at
predicting the propagation of large scale, organised convection. However, because of the principle that underlies
optical flow, that each pixel maintains its intensity between timesteps, it is unable to capture the IGD of
convection. Smaller scale convection exhibits higher rates of change in its evolution (Venugopal et al., 1999) and
so has the greatest uncertainty associated with it. Initially, this justifies why the majority of low skill is seen at
smaller scales (Figure 3h-i). However, at T+6 the difference in small scale features between the observations and
the nowcast becomes more widespread and so the low skill spreads further across the domain (Figure 3j).

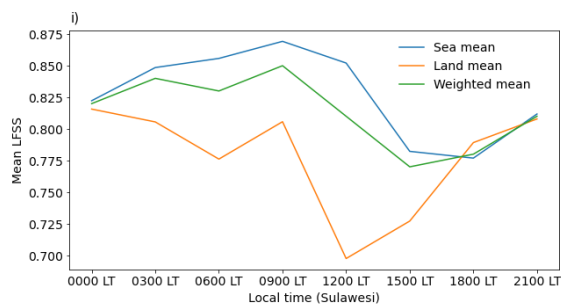
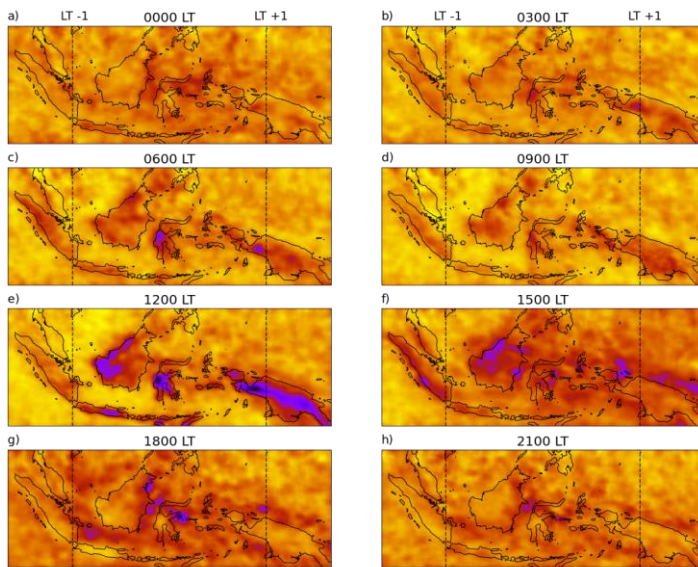
Figure 4 shows the mean FSSs for all 3,467 LK nowcasts (Table 2) and their corresponding persistence
~~forecasts~~nowcasts, evaluated at spatial scales of 10 km, 20 km, 60 km, 100 km and 200 km. The 10 km spatial
scale is the smallest scale of evaluation, hence it consistently produces the lowest scores. However, the model still
shows good skill on this scale ($FSS \geq 0.5$) at a lead time of 4 hours. Doubling the spatial scale to 20 km increases
the skilful lead time by ~1 hour. At the 60 km, 100 km and 200 km spatial scales, the LK nowcasts show skill
across all lead times with FSS scores of ~0.54, ~0.61 and ~0.75, respectively at the 6 hour lead time.

Skill reduces with lead time for all spatial scales, and increases with spatial scale at all lead times. On average,

385

390

the LK nowcasts outperform the persistence nowcasts~~forecast~~ at all lead times and for all spatial scales. Both the persistence and the LK nowcasts maintain the same structure and intensity at each lead time (hence relatively smaller skill difference at 1–2 hour lead time), however, the LK nowcasts propagate the convection across the domain whereas the persistence nowcasts remains stationary. This explains the increasing added value of the LK nowcast with lead time, as the observations move further from the persistence nowcast and the skill difference increases. Greater added value of the LK nowcast over persistence is also seen at smaller spatial scales e.g. the skill gap between the persistence nowcasts and the LK nowcasts at the 6 hour lead time is greater for 10 km spatial scale compared to 200 km spatial scale. The trend of FSSs across the shown lead times also varies for different spatial scales. At smaller spatial scales the nowcast is being evaluated on its ability to predict smaller scale

395

400

420

425

convection, which changes most rapidly/unpredictably, resulting in a higher rate of decrease in skill. The FSS evaluation on higher spatial scales smooths out these smaller scale convection changes and so a much lower rate
430 of decrease in skill is seen. The rate of decrease of the FSS evaluated on the 10 km spatial scale appears to show its largest rate of skill decrease at 1 – 2 hours and then decelerates over the following time intervals. On the other hand the rate of skill decrease for the 200 km spatial scale increases with lead time.

Figure 5 shows the mean LFSS for 3 hour lead time nowcasts over the MC, evaluated at a 100 km spatial scale. Evaluation on this scale has been used as it is able to clearly highlight the variations of skill across the domain.
435 For all nowcast initialisation times there is consistent noise in the skill over the sea. This may be representative of the stochastic nature of convection initiation over the sea. and Iit would therefore be expected that, if the period of evaluation was extended beyond 2015–2020 (i.e. increasing the number of nowcasted events used in the averaging in equations (8) and (9)), the LFSS noise field would become smoother over the sea would smooth out. During the overnight and morning initialisation times (Figures 5a–d) there is, on average, high skill over the
440 majority of the domain (Figure 5i). This is to be expected as at these times the majority of convection has formed large-scale, organised cloud systems with very few small-scale convection initiations occurring. These larger-scale systems will most often be propagating offshore of the islands (as part of the diurnal cycle), which the LK algorithm is most effective at capturing. Furthermore, overnight and during the morning there is, on average, relatively less convective activity than during the day meaning that reduced skill from inaccurate propagation
445 predictions is minimised.

Between 0900 LT and 1200 LT there is a significant drop in the mean skill over land (Figure 5i). Figures 5e–f show these distinct regions of low skill over land, which are tightly constrained to the coastal and mountainous regions of the islands and are closely tied to the diurnal cycle of convection over the MC. Convection begins to initiate and develop over the coastal and mountainous regions of the islands in the early afternoon, which is not
450 present at 1200 LT. The LK nowcast is unable to capture this new convection, resulting in low overland skill for the 1200 LT initialisation. Over these locations at this time of day, the LK algorithm would not be a skilful nowcasting tool. The low skill is still seen at 1500 LT but to a lesser extent. At this initialisation time the T–0 observations that are inputted into the LK algorithm will, on average, contain the majority of the convection that has initiated over the early afternoon. This convection will likely remain stationary over this time but be will be
455 growing in size. Therefore the persisting low skill at 1500 LT is representative of the LK nowcasts inability to predict the growth of convection.

The development of convection starts to slow as storms reach their mature stage in the evening. Less growth results in higher skill over the land. At 2100 LT the LFSS map looks similar to the overnight LFSS maps with high skill across the entire domain. Over this 3 hour forecast period the LK algorithm has shown good skill at
460 being able to nowcast the propagation of mesoscale convective systems offshore, which developed overland during the afternoon.

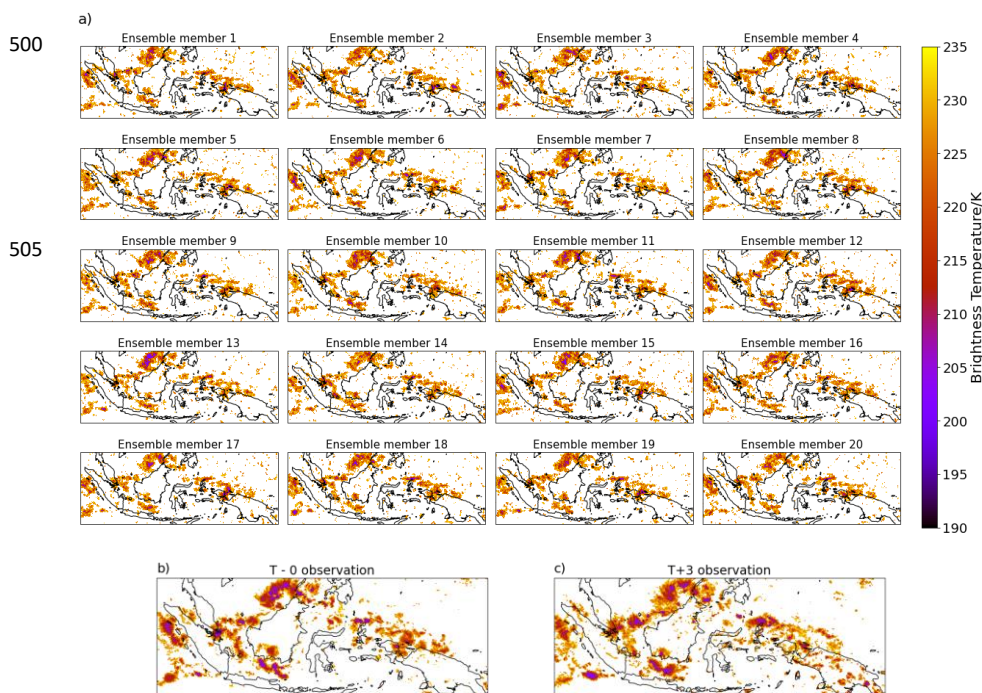
Understanding that the LK algorithm is unable to predict the IGD of convection means that, by identifying anomalous regions of low skill, LFSS maps can be a useful tool for identifying local effects due to land-sea interactions. An example of this is seen in a region of low skill over the Northeast coast of Borneo at 1800 LT.

465 On average, at this time of the day a land breeze begins to develop along the entire concave shaped coastline, potentially causing convergence near the middle. This convergence may then lead to the initiation of convection that the LK algorithm is unable to predict.

3.2 Ensemble nowcasting - STEPS algorithm

Figure 6a provides an example of a 20-member ensemble nowcast (generated using the same T-2, T-1 and T-0 set 470 as for the LK algorithm example in Figure 3), with a lead time of 3 hours, produced by pySTEPS's implementation of STEPS. Over the 3 hour period the main differences between the T-0 and T+3 observations are around New Guinea where new convection develops over the land and the convection north of the island becomes more scattered (Figure 6b–c). Visually, each ensemble member provides a good prediction of the large scale convection e.g. north of Borneo, with little difference between the members in the predicted shape and structure. The main 475 differences between the ensemble members come from differences in the stochastic fields injected for each member. This is seen in the varying levels of BT intensities in the large scale convection in each member. For example, the BT intensity over northern Borneo in ensemble member 13 is greater than in ensemble member 2. Differences are also seen between each ensemble member in the distribution of the predicted small scale convection e.g. over the Philippine Sea. Over Borneo, some members have predicted small scale convection which 480 approximately aligns with the new convection observed at T+3 e.g. ensemble member 18, whereas some members have predicted no convection here at all e.g. ensemble member 17.

One of the key aims of STEPS is to address the uncertainty in the evolution of small scale convection. In all members the convection northwest of New Guinea appears much noisier than in the T-0 observation. STEPS has recognized this as a region of uncertainty and addressed it by injecting noise at this scale. When compared to the 485 T+3 observation, it can be seen that the convection does in fact become more scattered and dissipated and so, although STEPS has not been able to precisely predict the new shape of the scattered convection, it has been able to capture the unpredictable nature of the evolution of this small scale convection.

The 20 member ensemble in Figure 6 has been used to produce the probabilistic nowcast in Figure 7 (extended to 1, 3 and 6 hour lead times). This probabilistic nowcast uses a threshold of 235 K, therefore including all the pixels 490 that were used to produce each ensemble nowcast. At T+1 the probabilistic nowcast shows a high degree of certainty in its prediction of the shape and location of convection, meaning that there is little variance between ensemble members. The T+1 lead time is the first timestep prediction that the algorithm makes and so it contains the least amount of stochastic noise in the extrapolation motion field, hence the least amount of member variance. As the lead time increases more stochastic noise is injected into the extrapolation motion fields and so the 495 uncertainty of the probabilistic nowcast increases. This can be seen in the reduction in high probabilities over the regions enclosed within the 20% contour. A reduction of high uncertainties within the 20% contour is also matched with a greater spread of low ensemble probability across the entire domain (outside the 20% contour) at T+3 and T+6.

510

515

520   Figure 7 also shows the number of small scale features in the probabilistic nowcast reducing at longer lead times.
        For example, aAt T+1 the convection northwest of New Guinea appears scattered in small blobs, whereas at T+6
        STEPS has smoothed out this small scale convection into a larger region. Again, this is evidence of the algorithm's
        attempt to address the uncertainty in the evolution of small scale convection by replacing it with stochastic noise.

        Figure 8a and b show the mean ROC and reliability curves for STEPS nowcasts initialised in the morning (0900
525   LT) and afternoon (1500 LT), evaluated over the sea and the land. For both surface types and initialisation times,
        the POD increases at each threshold meaning that, even as more uncertain regions enter the evaluation, the
        proportionnumber of hits to continues to exceed the number of misses increases. Furthermore, the POD exceeds
        the POFD at each threshold indicating that STEPS has skill in predicting regions < 235 K BT over the MC. The
        greatest POD – POFD difference, which is considered the optimum threshold for a probabilistic nowcast, is shown
530   at the ≥10% likelihood threshold for all of the curves.

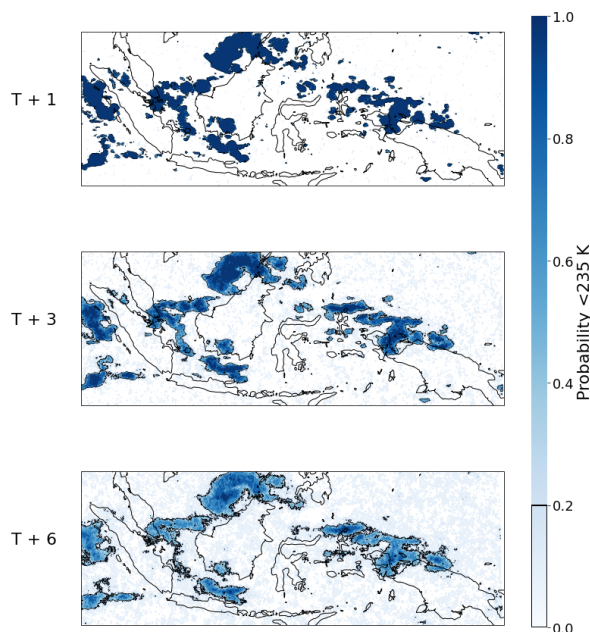



Figure 7. An example of a probabilistic BT nowcast for lead times of 1, 3 and 6 hours produced by the STEPS implementation
of pySTEPS on 11 December 2019. A BT threshold of < 235 K is used in order to include all pixels in the nowcast.

For both initialisation times in Figure 8a, STEPS produces higher Area Under the Curve (AUC) scores over the sea (0.80 and 0.78 for 0900 LT and 1500 LT, respectively) than over the land (0.71 and 0.68 for 0900 LT and 1500 LT, respectively), meaning that STEPS has more skill over the sea at these times. This can be explained by lower POD scores over the land, which is due to the STEPS algorithm being unable to capture the new convection

that most often develops there (increasing the number of misses). Furthermore, a comparison of each region within initialisation times suggests that STEPS is slightly more skilful in the morning (average AUC of 0.76) than in the afternoon (average AUC of 0.73). The morning – afternoon difference in POFD is also much greater over the land than the sea. This is likely due to a greater decrease in correct negative scores over the land, caused by new convection initiating where there was previously none, which STEPS is unable to predict.

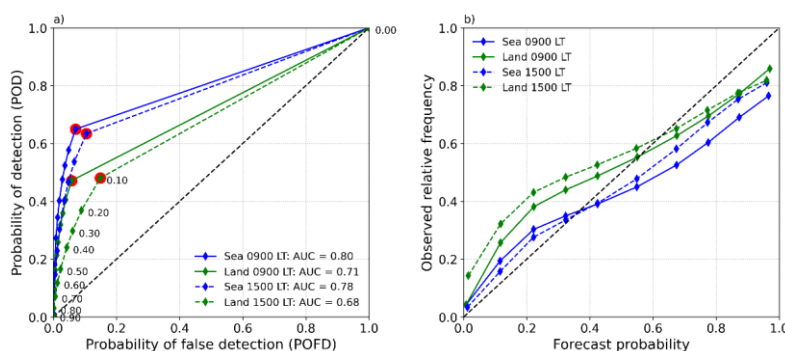



Figure 8. Composite a) ROC and b) reliability curves over the sea and land for STEPS 3-hour lead time nowcasts, initialised at 9am (441 nowcasts) and 3pm (405 nowcasts), with a threshold of < 235 K. The numbers next to the green points in a) represent the thresholds used to evaluate the nowcast at different likelihoods.

The positive slope between all the points for each reliability curve in Figure 8b means that, over both surface

distribution of ensemble predictions is therefore under-dispersive, meaning that the spread of predictions falls within the spread of observations i.e. it does not provide an optimum estimate of uncertainty.

When comparing the two surface types it can be seen that the under-prediction at lower nowcast probabilities is greater over the land, meaning the ensemble distribution has less variance and STEPS is better at capturing uncertainty over the sea for these initialisation times. However, at higher nowcast probabilities, more over-prediction is seen over the sea compared to over the land, meaning that STEPS becomes too confident at predicting higher likelihood events over the sea.

The under-dispersive feature of STEPS over the MC is due to low ensemble member variance, which (as previously mentioned) can be exemplified by visually assessing the lack of diversity between the ensemble members in Figure 6. The main source of ensemble member variance comes from the differences in the stochastic noise fields that STEPS injects into the nowcasts, and so, increasing the range of noise field intensities, or simply adding more members, would likely help to reduce this under-dispersive feature.

3.3 Comparison of STEPS, LK and persistence

By applying a threshold to a probabilistic STEPS nowcast, it is possible to produce a deterministic STEPS nowcast (in the form of a binary field) that can be evaluated using the FSS and directly compared to the corresponding LK and persistence nowcast. Figure 9 shows the mean FSSs for 3,467 LK nowcasts (solid), persistence nowcasts (dashed) and STEPS deterministic nowcasts produced using a threshold of ≥10% (dotted; STEPS10). The choice of threshold was based on the results of Figure 8, which show that, for both the morning and evening initialisation times, the optimum likelihood threshold was ≥10%.

At 10 km spatial scale, STEPS10 shows skill up to ~5 hours and has the highest skill across all lead times, outperforming LK and persistence. At 60 km scales STEPS10 still outperforms persistence across all lead times but approximately equals the skill of LK from 1 – 3 hours lead time. Onwards of 3 hours lead time, STEPS10 shows higher skill than LK.

The added value of STEPS10 over both LK and persistence decreases between the 10 km and 60 km spatial scales and, by the 200 km scale, STEPS10 shows the least skill out of all the nowcasts. This is likely due to less of the propagation being detected within the 200 km scale evaluation compared to the 10 km and 60 km scales – hence LK tends towards persistence at greater scales. Unlike LK, STEPS changes the internal structure of the convection

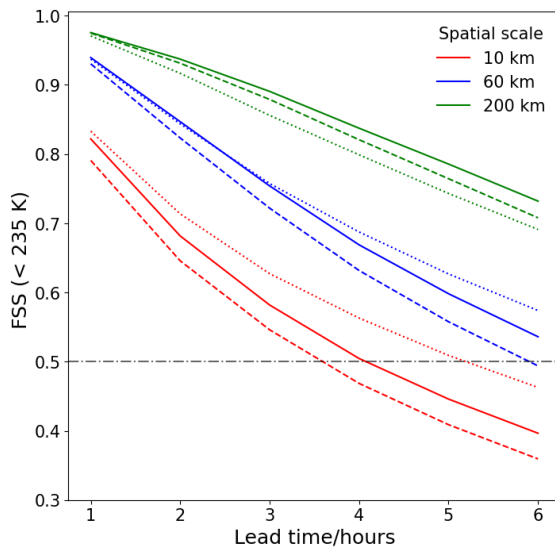



Figure 9. Composite FSSs against lead time for 3,457 LK (solid line), STEPS10 (dotted line) and persistence (dashed line) nowcasts, evaluated at a threshold of 235 K for 10 km, 60 km and 200 km spatial scales. The grey horizontal line marks the 0.5 FSS line, which is considered the cut-off for nowcast skill.

through the injection of stochastic noise. This change in the internal structure of convection (as opposed to change due to propagation) will be detected on the 200 km spatial scale and may contribute to a drop in performance relative to LK and persistence.

4 Conclusion

A deterministic (LK algorithm) and probabilistic (STEPS algorithm) implementation of the pySTEPS optical flow nowcasting library have been applied to satellite data over the MC to produce nowcasts with lead times of up to 6 hours.

630 Overall, the LK algorithm predicts the propagation of convection across the domain with skill (FSS≥0.5) up to 4 hours on the 10 km scale (the smallest scale of evaluation) and up to at least 6 hours on the 60 km scale. Similarly, Burton et al., (2022) used the LK algorithm to nowcast convective rain rates over West Africa using retrievals of BTs. Although it is difficult to precisely compare the two sets of results (rainfall retrievals likely have more fine-scale variability and errors in rainfall retrievals will affect nowcast skill (Hill et al., 2020)) they are nevertheless

635 comparable. Burton et al., (2022) show skill up to about 3 hours on the 64 km scale whereas we would expect somewhat higher skill for BT nowcasts over the MC.

Similar to the findings in Burton et al., (2022), the LK algorithm was unable to predict the initiation/growth/decay of convection, which is a manifestation of the optical flow assumption – each pixel maintains its intensity across all timesteps. This inability to predict IGD of convection is clearly seen when analysing maps of LFSS over the

640 MC. Over the sea the LK algorithm shows, on average, good skill at all nowcast initialisation times due to convection being mostly propagating in nature. Over land, however, the model shows high skill in the morning and evening but much lower skill in the afternoon. In the early afternoon this is due to the initiation of convection, which is closely constrained to the mountains. Later in the afternoon the low skill is due to the growth of the convection that initiated earlier on and persists over the mountains. Over the mountainous regions of the MC

645 during the afternoon, the LK algorithm would not be a useful nowcasting tool.

The STEPS algorithm aims to address the issue of unpredictability in the initiation/growth/decay of convection by injecting varying intensities of stochastic noise at different length scales to produce an ensemble nowcast. When analysing a probabilistic nowcast produced by a STEPS ensemble, it can be seen that the injection of noise has a smoothing effect, removing small scale convection and maintaining the shape of the larger, more predictable

650 convective regions.

A composite analysis of STEPS nowcasts was performed using ROC and reliability curves for 3 lead time predictions. The ROC curve is a measure of a nowcast's ability to discriminate between convective events that happened and convective events that did not (the higher the area under a ROC curve the more efficient it is at this), whereas the reliability curve measures how well the probabilistic predictions compare to observations. A

655 composite analysis of STEPS nowcasts using ROC and reliability curves showed that the algorithm is able to produceOverall, the analysis showed that STEPS can produce both skilful and reliable ensemble predictions for 3 hour lead times. When comparing land and sea at different times of day, it was shown that STEPS has highest skill over the sea during the morning (0900 LT initialisation time) with an AUC score of 0.8 (compared to an AUC score of 0.71 over the land). Imhoff et al., (2020) also applied the pySTEPS implementation of STEPS to

660 radar data over the Netherlands to produce nowcasts for a range of lead times. In their work they produced composite ROC curves for nowcasts with a lead time of 95–120 minutes, which had an AUC score of 0.81. The Netherlands experiences far less convective activity than the MC with the majority of its weather coming from propagating frontal clouds. This, therefore, further highlights the effectiveness of STEPS for the MC as it tries to predict convective clouds, which have a more unpredictable nature. However, the disadvantages of STEPS were

665 revealed when analysing the reliability curves. A common feature across both times and regions was the under-dispersive ensemble distributions, which were more extreme over land. This highlights STEPS's inability to predict the low likelihood events e.g. new initiations, and capture the whole uncertainty of the observed system.

To compare STEPS with LK and persistence, a deterministic version of STEPS was produced by thresholding the probabilistic nowcasts at ≥10% (STEPS10). When evaluated, the STEPS10 nowcasts had higher skill at spatial

670 scales of 10 km (across all lead times) and 60 km (from 3 – 6 hours). Therefore, not only does STEPS provide insight into the uncertainty of a convective system, but it can also derive a better single deterministic nowcast than LK and persistence at these scales. However, at a higher spatial scale of 200 km (where relatively less convection propagation is detected), LK nowcasts had the highest overall skill and the injection of stochastic noise produced by STEPS likely caused the STEPS10 nowcasts to have the lowest overall skill.

675 Continuous nowcasting over the entire MC is a requirement for early-warning systems, which does not currently exist. By providing both nowcast examples and composite nowcast analysis, this paper has shown the effective application of a deterministic (LK) and probabilistic (STEPS) algorithm to satellite data, showing their potential to be used operationally over the whole of the MC. The work highlights the key strengths and weaknesses of both algorithms, providing important information to a potential forecaster using these tools.

680

Code availability. Code is available from the first author upon request.

Data availability. Himawari data are available from the ICARE Data and Services Center at the University of Lille, at https://www.icare.univ-lille1.fr

685

Author contribution. JS performed the analysis and wrote the paper. CB, JM and SP participated in technical discussions and provided guidance for the analysis. MB and GP provided guidance for writing the paper.

Competing interests. The contact author has declared that none of the authors has any competing interests.

690

Acknowledgements. The author would like to acknowledge the SENSE CDT for providing funding and technical training for this project.

References

695   Ali, A., Supriatna, S., Sa'adah, U., 2021. RADAR-BASED STOCHASTIC PRECIPITATION NOWCASTING USING THE SHORT-TERM ENSEMBLE PREDICTION SYSTEM (STEPS) (CASE STUDY: PANGKALAN BUN WEATHER RADAR). IJReSES 18, 91. https://doi.org/10.30536/j.ijreses.2021.v18.a3527

Ayzel, G., Scheffer, T., Heistermann, M., 2020. RainNet v1.0: a convolutional neural network for radar-based precipitation nowcasting. Geosci. Model Dev. 13, 2631–2644. https://doi.org/10.5194/gmd-13-2631-2020

700   Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., Kumagai, Y., Miyakawa, T., Murata, H., Ohno, T., Okuyama, A., Oyama, R., Sasaki, Y., Shimazu, Y., Shimoji, K., Sumida, Y., Suzuki, M., Taniguchi, H., Tsuchiyama, H., Uesawa, D., Yokota, H., Yoshida, R., 2016. An Introduction to Himawari-8/9— Japan’s New-Generation Geostationary Meteorological Satellites. Journal of the Meteorological Society of Japan 94, 151–183. https://doi.org/10.2151/jmsj.2016-009

705   Birch, C.E., Webster, S., Peatman, S.C., Parker, D.J., Matthews, A.J., Li, Y., Hassim, M.E.E., 2016. Scale Interactions between the MJO and the Western Maritime Continent. Journal of Climate 29, 2471–2492. https://doi.org/10.1175/JCLI-D-15-0557.1

Bowler, N.E., Pierce, C.E., Seed, A.W., 2006. STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. Q. J. R. Meteorol. Soc. 132, 2127–2155.
710   https://doi.org/10.1256/qj.04.100

Burton, R.R., Blyth, A.M., Cui, Z., Groves, J., Lamptey, B.L., Fletcher, J.K., Marsham, J.H., Parker, D.J., Roberts, A., 2022. Satellite-based nowcasting of West African mesoscale storms has skill at up to four hours lead time. Weather and Forecasting. https://doi.org/10.1175/WAF-D-21-0051.1

Dayem, K.E., Noone, D.C., Molnar, P., 2007. Tropical western Pacific warm pool and maritime continent
715   precipitation rates and their contrasting relationships with the Walker Circulation. J. Geophys. Res. 112, D06101. https://doi.org/10.1029/2006JD007870

Feng, Z., Leung, L.R., Liu, N., Wang, J., Houze, R.A., Li, J., Hardin, J.C., Chen, D., Guo, J., 2021. A Global High-Resolution Mesoscale Convective System Database Using Satellite-Derived Cloud Tops, Surface Precipitation, and Tracking. JGR Atmospheres 126, e2020JD034202. https://doi.org/10.1029/2020JD034202

720   Ferrett, S., Frame, T.H.A., Methven, J., Holloway, C.E., Webster, S., Stein, T.H.M., Cafaro, C., 2021. Evaluating convection-permitting ensemble forecasts of precipitation over Southeast Asia. Weather and Forecasting. https://doi.org/10.1175/WAF-D-20-0216.1

Fiolleau, T., Roca, R., 2013. Composite life cycle of tropical mesoscale convective systems from geostationary and low Earth orbit satellite observations: method and sampling considerations. Quart J Royal Meteoro Soc 139, 941–953. https://doi.org/10.1002/qj.2174

Germann, U., Zawadzki, I., 2002. Scale-Dependence of the Predictability of Precipitation from Continental Radar Images. Part I: Description of the Methodology. Mon. Wea. Rev. 130, 2859–2873. https://doi.org/10.1175/1520-0493(2002)130<2859:SDOTPO>2.0.CO;2

Gijben, M., de Coning, E., 2017. Using Satellite and Lightning Data to Track Rapidly Developing Thunderstorms in Data Sparse Regions. Atmosphere 8, 67. https://doi.org/10.3390/atmos8040067

Goyens, C., Lauwaet, D., Schröder, M., Demuzere, M., Van Lipzig, N.P.M., 2012. Tracking mesoscale convective systems in the Sahel: relation between cloud parameters and precipitation. Intl Journal of Climatology 32, 1921–1934. https://doi.org/10.1002/joc.2407

Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., Gruber, C., 2011. The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region. Weather and Forecasting 26, 166–183. https://doi.org/10.1175/2010WAF2222451.1

Han, D., Lee, J., Im, J., Sim, S., Lee, S., Han, H., 2019. A Novel Framework of Detecting Convective Initiation Combining Automated Sampling, Machine Learning, and Repeated Model Tuning from Geostationary Satellite Data. Remote Sensing 11, 1454. https://doi.org/10.3390/rs11121454

Han, L., Zhang, J., Chen, H., Zhang, W., Yao, S., 2022. Toward the Predictability of a Radar-Based Nowcasting System for Different Precipitation Systems. IEEE Geosci. Remote Sensing Lett. 19, 1–5. https://doi.org/10.1109/LGRS.2022.3185031

Harjupa, W., Abdillah, M.R., Azura, A., Putranto, M.F., Marzuki, M., Nauval, F., Risyanto, Saufina, E., Jumianti, N., Fathrio, I., 2022. On the utilization of RDCA method for detecting and predicting the occurrence of heavy rainfall in Indonesia. Remote Sensing Applications: Society and Environment 25, 100681. https://doi.org/10.1016/j.rsase.2021.100681

Hill, P.G., Stein, T.H.M., Roberts, A.J., Fletcher, J.K., Marsham, J.H., Groves, J., 2020. How skilful are Nowcasting Satellite Applications Facility products for tropical Africa? Meteorol Appl 27. https://doi.org/10.1002/met.1966

Horn, B.K.P., Schunck, B.G., 1981. Determining optical flow. Artificial Intelligence 17, 185–203. https://doi.org/10.1016/0004-3702(81)90024-2

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2014. The Global Precipitation Measurement Mission. Bull. Amer. Meteor. Soc. 95, 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1

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 Resour. Res. 56. https://doi.org/10.1029/2019WR026723

Jianbo Shi, Tomasi, 1994. Good features to track, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94. Presented at the Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, IEEE Comput. Soc. Press, Seattle, WA, USA, pp. 593–600. https://doi.org/10.1109/CVPR.1994.323794

Lagerquist, R., Stewart, J., Ebert-Uphoff, I., Christina, K., 2021. Using Deep Learning to Nowcast the Spatial Coverage of Convection from Himawari-8 Satellite Data. Monthly Weather Review 149, 3897–3921. https://doi.org/10.1175/MWR-D-21-0096.1

Lebedev, V., Ivashkin, V., Rudenko, I., Ganshin, A., Molchanov, A., Ovcharenko, S., Grokhovetskiy, R., Bushmarinov, I., Solomentsev, D., 2019. Precipitation Nowcasting with Satellite Imagery, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Presented at the KDD '19: The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, ACM, Anchorage AK USA, pp. 2680–2688. https://doi.org/10.1145/3292500.3330762

770   Line, W.E., Schmit, T.J., Lindsey, D.T., Goodman, S.J., 2016. Use of Geostationary Super Rapid Scan Satellite Imagery by the Storm Prediction Center*. Weather and Forecasting 31, 483–494. https://doi.org/10.1175/WAF-D-15-0135.1

Lucas, B.D., Kanade, T., 1981. An Iterative Image Registration Technique with an Application to Stereo Vision 10.

775   Marcos, 2015. NWC SAF convective precipitation product from MSG: A new day-time method based on cloud top physical properties. Tethys. https://doi.org/10.3369/tethys.2015.12.01

Mori, S., Jun-Ichi, H., Tauhid, Y.I., Yamanaka, M.D., Okamoto, N., Murata, F., Sakurai, N., Hashiguchi, H., Sribimawati, T., 2004. Diurnal Land–Sea Rainfall Peak Migration over Sumatera Island, Indonesian Maritime Continent, Observed by TRMM Satellite and Intensive Rawinsonde Soundings. Mon. Wea. Rev. 132, 2021–
780   2039. https://doi.org/10.1175/1520-0493(2004)132<2021:DLRPMO>2.0.CO;2

Mueller, C., Saxen, T., Roberts, R., Wilson, J., Betancourt, T., Dettling, S., Oien, N., Yee, J., 2003. NCAR Auto-Nowcast System. Wea. Forecasting 18, 545–561. https://doi.org/10.1175/1520-0434(2003)018<0545:NAS>2.0.CO;2

Murphy, A.H., Epstein, E.S., 1989. Skill Scores and Correlation Coefficients in Model Verification. Mon. Wea.
785   Rev. 117, 572–582. https://doi.org/10.1175/1520-0493(1989)117<0572:SSACCI>2.0.CO;2

Permana, D.S., Hutapea, T.D., Praja, A.S., Paski, J.A.I., Makmur, E.E.S., Haryoko, U., Umam, I.H., Saepudin, M., Adriyanto, R., 2019. The Indonesia In-House Radar Integration System (InaRAISE) of Indonesian Agency for Meteorology Climatology and Geophysics (BMKG): Development, Constraint, and Progress. IOP Conf. Ser.: Earth Environ. Sci. 303, 012051. https://doi.org/10.1088/1755-1315/303/1/012051

790   Porson, A.N., Hagelin, S., Boyd, D.F.A., Roberts, N.M., North, R., Webster, S., Lo, J.C., 2019. Extreme rainfall sensitivity in convective-scale ensemble modelling over Singapore. Q.J.R. Meteorol. Soc. 145, 3004–3022. https://doi.org/10.1002/qj.3601

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). Geosci. Model Dev.
795   12, 4185–4219. https://doi.org/10.5194/gmd-12-4185-2019

Qian, J.-H., 2008. Why Precipitation Is Mostly Concentrated over Islands in the Maritime Continent. Journal of the Atmospheric Sciences 65, 1428–1441. https://doi.org/10.1175/2007JAS2422.1

Ramage, C.S., 1968. ROLE OF A TROPICAL "MARITIME CONTINENT" IN THE ATMOSPHERIC CIRCULATION [1]. Mon. Wea. Rev. 96, 365–370. https://doi.org/10.1175/1520-
800   0493(1968)096<0365:ROATMC>2.0.CO;2

Reen, B.P., Cai, H., Raby, J.W., 2020. Preliminary Investigation of Assimilating Global Synthetic Weather Radar.

Roberts, A.J., Fletcher, J.K., Groves, J., Marsham, J.H., Parker, D.J., Blyth, A.M., Adefisan, E.A., Ajayi, V.O., Barrette, R., de Coning, E., Dione, C., Diop, A., Foamouhoue, A.K., Gijben, M., Hill, P.G., Lawal, K.A.,
805   Mutemi, J., Padi, M., Popoola, T.I., Rípodas, P., Stein, T.H.M., Woodhams, B.J., 2022. Nowcasting for AFRICA : advances, potential and value. Weather 77, 250–256. https://doi.org/10.1002/wea.3936

Roberts, N.M., Lean, H.W., 2008. Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. Monthly Weather Review 136, 78–97. https://doi.org/10.1175/2007MWR2123.1

810   Roca, R., Fiolleau, T., Bouniol, D., 2017. A Simple Model of the Life Cycle of Mesoscale Convective Systems Cloud Shield in the Tropics. Journal of Climate 30, 4283–4298. https://doi.org/10.1175/JCLI-D-16-0556.1

Rouault, E., Warmerdam, F., Schwehr, K., Kiselev, A., Butler, H., Łoskot, M., Szekeres, T., Tourigny, E., Landa, M., Miara, I., Elliston, B., Chaitanya, K., Plesea, L., Morissette, D., Jolma, A., Dawson, N., Baston, D., de Stigter, C., Miura, H., 2023. GDAL. https://doi.org/10.5281/ZENODO.5884351

815   Seed, A.W., 2003. A Dynamic and Spatial Scaling Approach to Advection Forecasting. J. Appl. Meteor. 42, 381–388. https://doi.org/10.1175/1520-0450(2003)042<0381:ADASSA>2.0.CO;2

Sieglaff, J.M., Hartung, D.C., Feltz, W.F., Cronce, L.M., Lakshmanan, V., 2013. A Satellite-Based Convective Cloud Object Tracking and Multipurpose Data Fusion Tool with Application to Developing Convection. Journal of Atmospheric and Oceanic Technology 30, 510–525. https://doi.org/10.1175/JTECH-D-12-00114.1

820   Sobajima, A., 2012. Rapidly Developing Cumulus Areas Derivation Algorithm Theoretical Basis Document.

Srivastava, K., Lau, Sharons.Y., Yeung, H.Y., Cheng, T.L., Bhardwaj, R., Kannan, A.M., Bhowmik, S.K.R., Singh, H., 2021. Use of SWIRLS nowcasting system for quantitative precipitation forecast using Indian DWR data. MAUSAM 63, 1–16. https://doi.org/10.54302/mausam.v63i1.1442

Venugopal, V., Foufoula-Georgiou, E., Sapozhnikov, V., 1999. Evidence of dynamic scaling in space-time
825   rainfall. J. Geophys. Res. 104, 31599–31610. https://doi.org/10.1029/1999JD900437

Vila, D.A., Machado, L.A.T., Laurent, H., Velasco, I., 2008. Forecast and Tracking the Evolution of Cloud Clusters (ForTraCC) Using Satellite Infrared Imagery: Methodology and Validation. Weather and Forecasting 23, 233–245. https://doi.org/10.1175/2007WAF2006121.1

Woodhams, B.J., Birch, C.E., Marsham, J.H., Bain, C.L., Roberts, N.M., Boyd, D.F.A., 2018. What Is the
830   Added Value of a Convection-Permitting Model for Forecasting Extreme Rainfall over Tropical East Africa? Monthly Weather Review 146, 2757–2780. https://doi.org/10.1175/MWR-D-17-0396.1

World Meteorological Organization [WWW Document], 2023. . Early Warnings For All Initiative scaled up into action on the ground. URL https://public.wmo.int/en/media/press-release/early-warnings-all-initiative-scaled-action-ground

835   Yamanaka, M.D., 2016. Physical climatology of Indonesian maritime continent: An outline to comprehend observational studies. Atmospheric Research 178–179, 231–259. https://doi.org/10.1016/j.atmosres.2016.03.017

Yang, G.-Y., Slingo, J., 2001. The Diurnal Cycle in the Tropics. Mon. Wea. Rev. 129, 784–801. https://doi.org/10.1175/1520-0493(2001)129<0784:TDCITT>2.0.CO;2

840

845