Article:

Peatman, S.C., Birch, C.E. orcid.org/0000-0001-9384-2810, Schwendike, J. et al. (5 more authors) (Accepted: 2025) Physical controls on the variability of offshore propagation of convection from Sumatra. Journal of Geophysical Research: Atmospheres. ISSN 2169-897X (In Press)

Physical controls on the variability of offshore propagation of convection from Sumatra

Simon C. Peatman[1,2], Cathryn E. Birch[1], Juliane Schwendike[1], John H. Marsham[1], Emma Howard[3,4,5], Steven J. Woolnough[4,5], Jack M. Mustafa[1,6], Adrian J. Matthews[6,7]

[1]Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds, United Kingdom
[2]Centre for Climate Research Singapore, Meteorological Service Singapore
[3]Bureau of Meteorology, Brisbane, Australia
[4]Department of Meteorology, University of Reading, Reading, UK
[5]National Centre for Atmospheric Science, University of Reading, Reading, UK
[6]Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom
[7]School of Mathematics, University of East Anglia, Norwich, United Kingdom

Key Points:

- The offshore propagation of convection south-west of Sumatra is a key feature of the mean diurnal cycle but occurs on only 28% of DJF days
- The diurnal cycle over land occurs when large-scale onshore wind causes convergence over the mountains and low-level humidity causes moist instability
- Offshore propagation arises due to the mid-level wind, convergence due to land breezes or cold pools, and inflow of low-level unstable air

Corresponding author: Simon Peatman, `simon_peatman@nea.gov.sg`

Abstract

Previous work has explained the physical mechanisms behind nocturnal offshore propagation of convection south-west of Sumatra. Low-level moisture flux convergence due to the land breeze front controls the progression of convection, typically a squall line, away from the coast overnight. However, the diurnal convection over the mountains occurs on only 57% of days in December–February (DJF) and propagates offshore on only 49% of those days. We investigate day-to-day variability in dynamical and thermodynamical conditions to explain the variability in diurnal convection and offshore propagation, using a convection-permitting simulation run for 900 DJF days. A convolutional neural network is used to identify regimes of diurnal cycle and offshore propagation behaviour. The diurnal cycle and offshore propagation are most likely to occur ahead of an active Madden-Julian Oscillation, or during El Niño or positive Indian Ocean Dipole; however, any regime can occur in any phase of these large-scale drivers, since the major control arises from the local scale. When the diurnal cycle of convection occurs over land, low-level wind is generally onshore, providing convergence over the mountains; and low-level humidity over the mountains is high enough to make the air column unstable for moist convection. When this convection propagates offshore, mid-level offshore winds provide a steering flow, combined with stronger convergence offshore due to more onshore environmental winds. Low-level moisture around the coast also means that, as the convection propagates, the storm-relative inflow of air into the system adds greater instability than would be the case on other days.

Plain Language Summary

In Sumatra, a large island of west Indonesia, rainfall tends to form by convection over the mountains during the afternoon and evening. This is known as the diurnal cycle. Furthermore, the rainfall often propagates overnight, both offshore (towards the south-west) and onshore (towards the north-east). A previous paper investigated the physics behind this offshore propagation; overnight the land breeze converges offshore with what remains of the sea breeze from earlier in the day, and the convergence of air masses causes uplift at the front between them, sustaining the line of convection which duly propagates offshore. However, neither the diurnal cycle nor its offshore propagation occur every day. This study investigates the physical conditions that control whether these phenomena occur on any given day. The Madden-Julian Oscillation, El Niño–Southern Oscillation

and Indian Ocean Dipole all have an impact, but they alone cannot predict if the diurnal convection will occur or if it will propagate. Instead, these phenomena are controlled by the wind direction and low-level humidity, which cause convergence of air over the mountains and an unstable vertical profile when there is a diurnal cycle; and offshore mid-level winds, convergence of air over the sea and inflow of unstable air when there is offshore propagation.

1 Introduction

Located in the Indo-Pacific warm pool, the Maritime Continent (the south-east Asia archipelago; Ramage, 1968) experiences intense deep convection, with the diurnal cycle being the greatest form of variability (e.g., Yang & Slingo, 2001; Qian, 2008; Biasutti et al., 2012). Typically, the diurnal cycle of precipitation peaks over the islands during the afternoon and evening, whereas over the sea the peak is during the early hours of the morning. However, this diurnal cycle does not occur on all days and when it does occur the amplitude may vary considerably between days. Many studies have described scale interactions in which the local-scale diurnal cycle is forced by large-scale weather phenomena. For example, the strongest diurnal cycle tends to occur ahead of the arrival of an active Madden-Julian Oscillation (MJO) envelope (Oh et al., 2012; Peatman et al., 2014; Sakaeda et al., 2017; Vincent & Lane, 2017; Sakaeda et al., 2020; Peatman et al., 2021); the El Niño–Southern Oscillation (ENSO) can enhance the local diurnal cycle of rainfall over the islands in the El Niño phase, even though on the large scale the Maritime Continent rainfall is suppressed (Rauniyar & Walsh, 2013); and previous studies stated that the negative phase of the Indian Ocean Dipole (IOD) causes wetter extremes (Kurniadi et al., 2021) and a stronger diurnal cycle over the southern half of Sumatra (Fujita et al., 2013).

There is growing evidence that these scale interactions result, at least in part, from the way in which large-scale drivers control coastal winds. Peatman et al. (2021) showed that over south-west Sumatra a stronger diurnal cycle and stronger offshore propagation tend to occur when coastal winds are offshore, while strong onshore winds result in a very weak diurnal cycle and no offshore propagation. Similar results were found for offshore propagation from Borneo and Java. A more comprehensive study by Aoki and Shige (2024) investigated precipitation rates and offshore propagation under onshore and offshore wind conditions of varying strengths, across the global tropics. The Maritime

Continent experiences the most intense diurnal cycle of rainfall over land when daily mean 850 hPa wind is moderate or weak. Strong onshore winds (fifth panel of their figure 7e) are associated with large-scale rain rather than a localized diurnal cycle, in agreement with Peatman et al. (2021).

Over certain regions of the Maritime Continent, organized convection that is initiated over land is observed to propagate offshore overnight (e.g., Mori et al., 2004; Love et al., 2011; Sakaeda et al., 2020; Peatman et al., 2023), including to the south-west of Sumatra. Although the Hovmöller diagrams in Aoki and Shige (2024)'s figure 7 do not show hours of the following day, so it is not possible to see nocturnal offshore propagation, the authors make further arguments relating to the Doppler shifting and advection of gravity waves to explain an asymmetry between onshore and offshore propagation under strong coastal wind conditions.

The physical mechanisms of the offshore propagation, in the Maritime Continent and other tropical locations, have been investigated using both observations and models (Houze et al., 1981; Mapes, Warner, & Xu, 2003; Mapes, Warner, Xu, & Negri, 2003; Love et al., 2011; Peatman et al., 2023). The proposed mechanism of Houze et al. (1981, see their figure 16) for offshore propagation from Borneo involved low-level convergence between a land breeze and the monsoon wind, triggering convection successively offshore. On the other hand, Mapes, Warner, Xu, and Negri (2003) explained offshore propagation over the Panama Bight in terms of gravity waves emitted by the boundary layer triggering offshore convection (see their figure 11). For south-west Sumatra, Love et al. (2011) noted a transition from a convective profile with mid-tropospheric heating during early afternoon to a stratiform profile with upper-tropospheric heating and mid-tropospheric cooling at later times. They attributed the offshore propagation to gravity waves forced by this change in the heating profile.

Peatman et al. (2023) considered the land-sea breeze circulation induced by the land-sea temperature contrast and its effect on the propagation (see their figure 14). They found that, in examples of clear and coherent propagation, a land breeze provides strong low-level moisture flux convergence as it converges with the remnant of the onshore sea breeze from earlier in the day. An organized squall line that forms over the Barisan mountains, aligned along Sumatra's south-west coast, propagates offshore collocated with the convergence line due to the land breeze. There may also have been a contribution to the moisture flux convergence from cold pools caused by the convection, but it was not possible to disentangle this from the effect of the land breeze.

Furthermore, gravity waves propagating offshore can trigger isolated convection, which in a composite is averaged out to appear as a faster mode of propagation. However, Peatman et al. (2023) demonstrated that gravity waves are not responsible for the coherent, organized squall line which forms over the Sumatran mountains and propagates offshore as an organized system on any given day.

Although we know that the diurnal cycle and its offshore propagation vary according to the large scale, and we understand the physical mechanism of the offshore propagation when it occurs, there remains a lack of understanding of the physical mechanisms that cause the day-to-day variability in offshore propagation at the local scale. The present study addresses this by identifying a range of diurnal cycle and offshore propagation behaviours, and uses a convection-permitting simulation to understand the dynamical and thermodynamical conditions associated with each. We opt to use a convection-permitting model instead of a reanalysis for this research since reanalyses typically rely on convection parametrization schemes which may not represent the location and timing of convective storms correctly, and because running a model allows us to output a comprehensive set of diagnostics that we require.

The model and methodologies used are explained in section 2, results are presented in section 3 and a discussion is found in section 4.

2 Data and methods

2.1 Observational data

Gridded precipitation observations are taken from the Global Precipitation Measurement (GPM)'s Integrated Multi-satellitE Retrievals for GPM (IMERG) dataset, version 6 (Huffman et al., 2019), which is provided on a 0.1° grid every 30 minutes. The phase of the MJO for any given day is taken from the Real-time Multivariate MJO index (RMM; Webster & Hoyos, 2004), discarding days on which the RMM amplitude is less than 1 as the MJO is defined as being weak. Orography is shown using the Global Land One-km Base Elevation (GLOBE) project (Hastings et al., 1999).

The December to February (DJF) seasons used in this study (see section 2.2) were judged in table 1 of Howard et al. (2024) to belong, overall, to a particular phase of ENSO (El Niño, La Niña or neutral) using the Niño3.4 index; and the IOD (positive or negative) using the Dipole Mode Index. These ENSO and IOD phases are used in the present study also.

2.2 Convection-permitting MetUM simulation

We use a convection-permitting configuration of the Met Office Unified Model (MetUM), which was described and evaluated by Howard et al. (2024), and is summarized here. The model setup is illustrated in figure 1. The outer domain is 85–160°E and 20°S–20°N, with a 0.09375° (latitude) by 0.140625° (longitude) grid, equating to approximately 12 km grid spacing at the equator. At the lateral boundaries, forcing is provided by the European Centre for Medium-Range Weather Forecasting (ECMWF) Reanalysis 5 (ERA5; Hersbach et al., 2020), every 6 hours. This 12 km model has parametrized convection.

The nested inner domain is 90–155°E and 15°S–15°N, with a 0.02° grid, equating to approximately 2.2 km grid spacing at the equator; and is driven at the lateral boundaries by the 12 km model. This 2.2 km model has explicit convection and uses the tropical version of the Regional Atmosphere and Land 2 (RAL2T) science configuration (Bush et al., 2023).

Both atmospheric models are coupled to a K-Profile Parameterisation (KPP) ocean model (Large et al., 1994) on the same horizontal grid as the 12 km configuration. This is a mixed-layer model – that is, all columns are independent one-dimensional models simulating vertical mixing but there is no horizontal transport, allowing the representation of air-sea interactions with little computational expense.

Only the 2.2 km convection-permitting model is used in the present study, with output available up to every 5 minutes.

The model was run for 10 DJFs, chosen to cover a range of large-scale conditions, including different phases of ENSO and the IOD, and different levels of MJO activity. For details, see table 1 of Howard et al. (2024). For consistency, these same 10 DJFs were used for all observational parts of the present study.

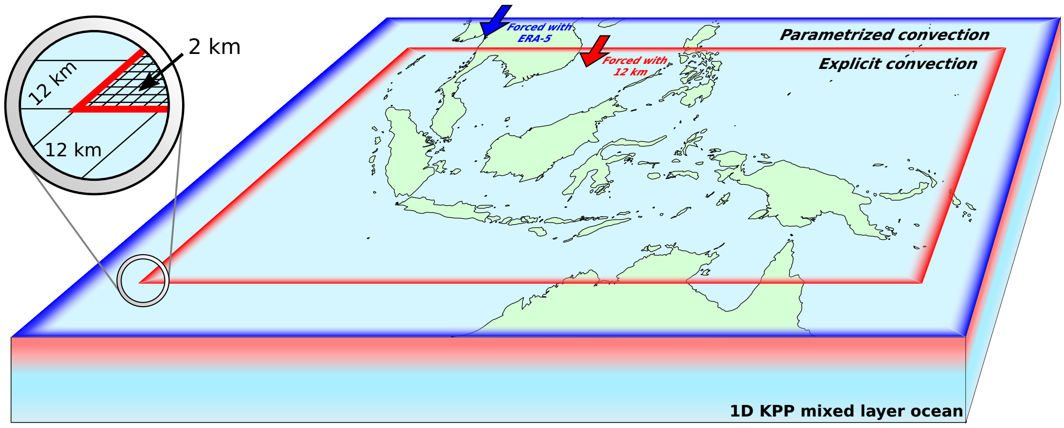

Figure 1. Schematic diagram of the MetUM simulation used in this study. For more details, see Howard et al. (2024).

Note that the MetUM run is used as a research tool to investigate the physical mechanisms of the diurnal cycle of convection and its propagation, not as a forecast tool. Therefore, the days on which the diurnal convection occurs over Sumatra and the days on which it propagates offshore may differ between the model and observations, but this is not a problem provided the model has a reasonably realistic distribution of diurnal cycle behaviour. Due to forcing at the boundaries by ERA5, large-scale phenomena should be represented with a high degree of accuracy (see Howard et al., 2024, for details), but on the local scale we do not expect such a strong match with the observations.

2.3 Subjective classification of propagation regimes

In order to investigate the day-to-day variability in offshore propagation, Hovmöller diagrams of observed precipitation from IMERG were created, using the red box in figure 2, averaging over the long side. This was done for the same 900 days as were covered by the MetUM simulation. Eight examples are shown in figure 3.

A visual inspection of the 900 Hovmöllers suggested four broad regimes, listed in table 1. The examples in figure 3 include two instances of each regime, indicated by the coloured rectangles above the panels. When there is a diurnal cycle of convection, with precipitation peaking over the mountains in the late afternoon or evening, the precipitation may (figures 3c,f,h) or may not (figure 3g) propagate onshore, north-eastward. However, we focus solely on the propagation to the south-west in this study.

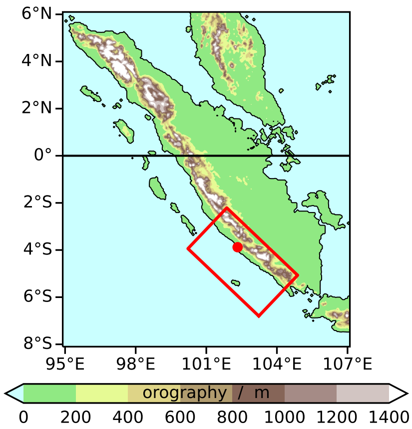

Figure 2. Orography of Sumatra from the GLOBE dataset, with the Hovmöller box used in figures 3 and 4 drawn in red. The red dot is Bengkulu.

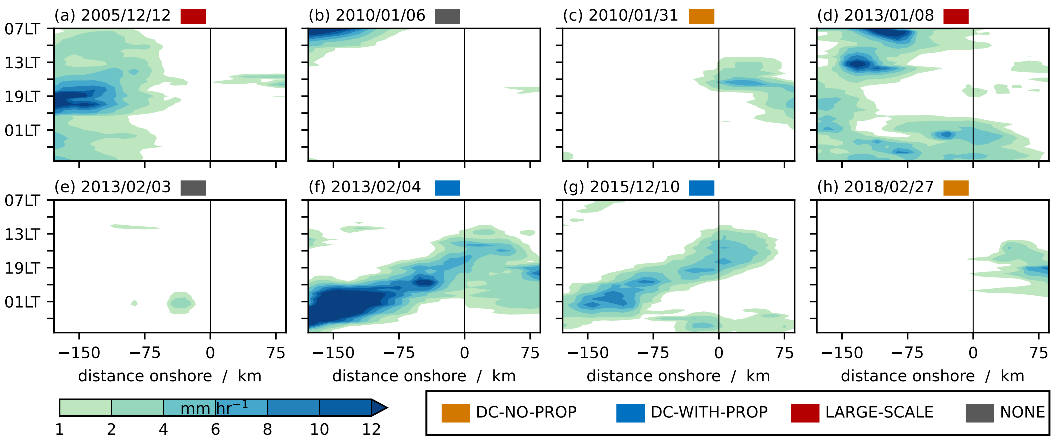

Figure 3. Example offshore-propagating Hovmöller diagrams of precipitation from IMERG observations for the red box shown in figure 2, averaged over the alongshore direction, with local time (LT) running down the page from 07 LT one day to 07 LT the next. The black vertical line at $x = 0$ is the south-west coast of Sumatra, with the land, mainly covered by mountains, to the right ($x > 0$) and the sea to the left ($x < 0$). Eight selected days are shown, from within the model run period. Coloured rectangles indicate the regime for the day shown (see section 2.3 and table 1).

The 900 Hovmöller diagrams were each classified subjectively as belonging to one of the four regimes. In order to achieve a degree of robustness, this exercise was performed independently by two of the authors. Where they agreed on the classification (which was true for 701 of the days), this classification was taken as definitive. For the remaining 199 days, the process was repeated until a majority verdict was reached.

Table 1. Names of regimes found in Hovmöllers of precipitation and a description of each.

Regime	Description
DC-NO-PROP	Diurnal cycle of precipitation occurs over mountains but precipitation does not propagate offshore*
DC-WITH-PROP	Diurnal cycle of precipitation occurs over mountains and precipitation does propagate offshore
LARGE-SCALE	Hovmöller is dominated by large-scale rainfall for much or all of the day
NONE	None of the above – very little or no rainfall

*This does not necessarily preclude precipitation propagating onshore, to the north-east.

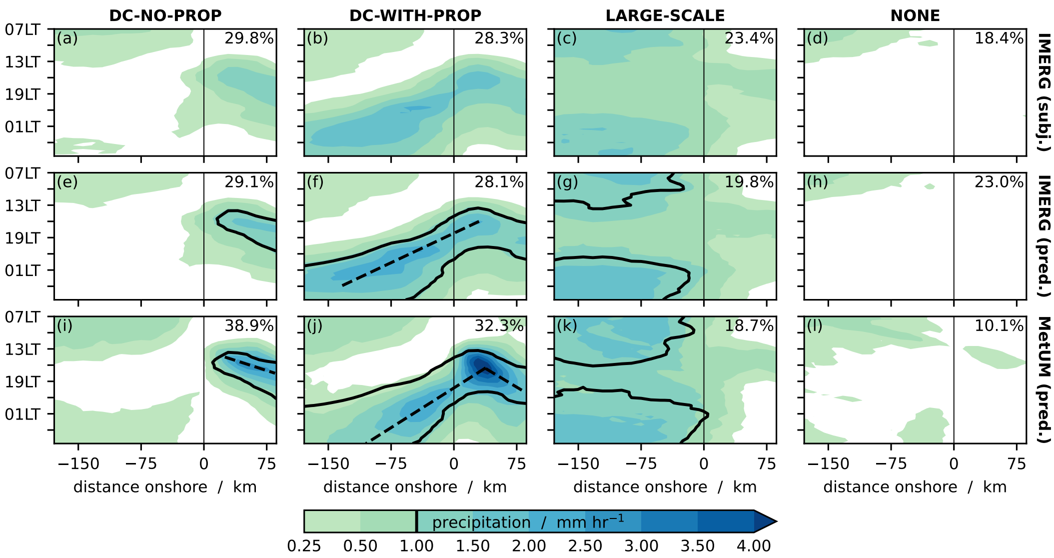

Figure 4. Composite Hovmöller diagrams of precipitation for (a–d) IMERG (subjective classification), (e–h) IMERG (CNN-predicted) and (i–l) the MetUM simulation (CNN-predicted). Numbers in the top-right of each panel indicate the percentage of days in that regime. The black 1 mm hr^{-1} contour in (e–h) is taken from the shading in (a–d), and in (i–l) is taken from the shading in (e–h), for comparison. The black dashed lines indicate propagation at (f) 3.8 m s^{-1} offshore, (i) 5.6 m s^{-1} onshore, and (j) 2.8 m s^{-1} offshore and 3.0 m s^{-1} onshore.

2.4 Supervised machine learning

While the observations were classified into regimes subjectively (section 2.3), for consistency and convenience this process was automated for the MetUM output, using the subjective classifications from the observations to train a supervised machine learning model.

A convolutional neural network (CNN) was used with the architecture shown in figure S1. 75% of the observations (675 days) were randomly selected as the training data set and the remaining 25% (225 days) formed the testing data set. Hovmöller diagrams as in figure 3 were used as input images and the subjectively-classified regimes, to be predicted by the CNN, were represented as one-hot vectors. Prior to training, the training data were augmented by randomly rotating images by up to 20°, or randomly translating them either horizontally or vertically by up to 10%. Data augmentation (e.g., Montserrat et al., 2017; Shorten & Khoshgoftaar, 2019; Poojary et al., 2021) gives greater scope for the CNN to recognize patterns that are not at exactly the same orientation or in exactly the same place as in the 675 input Hovmöllers used.

For the subjective classification in section 2.3 the images were not normalized first, so the images were also not normalized in the CNN. Hence, the CNN makes its predictions based on the magnitude of the precipitation as well as its spatial pattern.

Repeating the creation and training of a CNN will not produce an identical result, owing to the randomness in the selection of the training data set, the data augmentation, the initialisation of the CNN's hyperparameters and the batching of input data during training. 50 separate CNNs were trained and saved. The CNN with the highest accuracy where the accuracy (0.845) was approximately equal to validation accuracy (0.844) was chosen and the remaining CNNs were discarded. By way of comparison, the highest accuracy achieved by any of the 50 CNNs was 0.869, but the validation accuracy was only 0.769 so it was likely overfitting to the training data set.

The CNN produces outputs $x_0, \ldots, x_3$ (figure S1k), each in the interval $[0, 1]$ and with $\sum_i x_i = 1$. The largest of these determines the predicted regime (figure S1l). Some input images are classified with low certainty (i.e., $\max(x_i)$ is not close to 1), but it is difficult to quantify the degree of certainty since x_i cannot be interpreted as probabilities (the CNN is not so calibrated). When classifying the observed Hovmöllers subjectively (section 2.3), the two people performing the classifications agreed at the first attempt on 701 occasions out of 900. We take this as an estimate of how certain we can reasonably expect a classification to be. Therefore, of all 900 of the $\max(x_i)$ values found when the CNN classified the observed Hovmöllers, the 701$^{\text{st}}$ largest value was taken as a threshold. For any input image where the CNN's $\max(x_i)$ is below this threshold, the classification is deemed uncertain.

Composites over each of the four regimes in observations are shown in figures 4a–d. The diurnal cycle of convection occurs over the mountains on just over 58% of DJF days, with a little under half of these having offshore propagation. The diurnal precipitation and its offshore propagation show up clearly in long-term composites of the diurnal cycle (e.g., Peatman et al., 2014) so this is typically thought of as the canonical behaviour for Sumatra and the sea to its south-west, but in fact this occurs on only around 28% of DJF days. Composites for the 701 days with certain classifications are shown in figures 4e–h. When classifying the MetUM Hovmöllers, 749 days were above the threshold; composites for these days are in figures 4i–l.

The confusion matrix in figure 5a measures the performance of the CNN by comparing the subjectively-classified regimes and CNN-predicted regimes, both in observations. If the CNN were perfect, all values would lie on the leading diagonal. With very few values off this leading diagonal, and with the composite Hovmöllers in figures 4e–h being very similar to those in figures 4a–d, we conclude that the CNN is successful in performing the classifications.

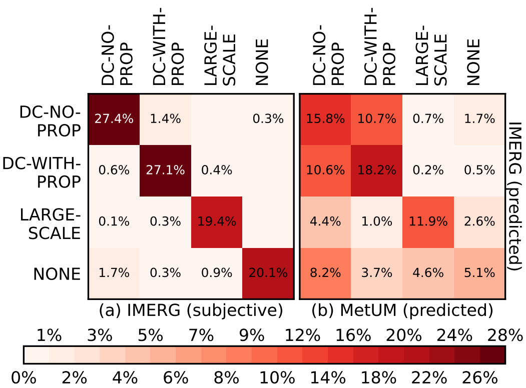

Figure 5. Confusion matrices showing (a) the performance of the CNN in classifying IMERG observations, compared against the subjective classification; (b) the performance of the MetUM simulation at producing days in the same regime as observations.

3 Results

3.1 Evaluation of MetUM simulation

Figures 4i–l show composite Hovmöller diagrams for each regime in the MetUM simulation. By construction there must be a degree of similarity with the observed composites in figures 4e–h, since the CNN seeks similar patterns of pixels in order to perform the classification. However, their remarkable similarity indicates that the model reproduces the observed regimes realistically. In figure 4i there is precipitation offshore overnight (bottom-left of panel) which is not seen in the corresponding figure 4e. This suggests there are MetUM days classified as DC-NO-PROP which could perhaps have been considered DC-WITH-PROP, but the offshore precipitation was too weak to be picked up by the CNN. In the DC-WITH-PROP regime, the modelled offshore propagation (2.8 m s^{-1}) is slightly slower than in observations (3.8 m s^{-1}).

The confusion matrix in figure 5b measures the performance of the MetUM simulation by comparing the regimes against observations, for those days when the CNN was above the certainty threshold for both. If the regime in the MetUM matched that in observations on all days, all values would lie on the leading diagonal.

The percentage of days falling into each regime are broadly similar for the observations and the model (figures 4e–l). The exception is that the NONE regime occurs less frequently in the model (10.1%; figure 4l) than in observations (23.0%; figure 4h), with the model instead having the two diurnal cycle regimes more often. Hence, in the MetUM the diurnal convection over the Sumatra mountains is triggered more often than in observations. When NONE occurs in observations, the MetUM may exhibit any of the other three regimes (bottom row of figure 5b).

The MetUM generally reproduces the LARGE-SCALE regime on the same days as observations. This is likely to be because the LARGE-SCALE regime arises from certain large-scale conditions in the region, and these should match well between the MetUM and observations due to the forcing at the lateral boundaries with ERA5. Looking at the four squares in the top-left of figure 5b, when the observations are in one of DC-NO-PROP or DC-WITH-PROP, the MetUM also tends to be in one of those two regimes. In other words, there is a good match in terms of whether the diurnal cycle of precipitation occurs. However, there is no close match between those two regimes, so the model does not closely replicate the observed occurrence of offshore propagation on any given day. This suggests that the triggering of diurnal cycle rainfall may be more related to large-scale conditions, while the offshore propagation may depend more on localized conditions that are not reproduced realistically at a location far from the lateral boundary forcing.

3.2 Dependence on large-scale drivers

Previous studies have shown that the amplitude of the diurnal cycle varies through an MJO cycle (see section 1; Oh et al., 2012; Peatman et al., 2014; Sakaeda et al., 2017; Vincent & Lane, 2017; Sakaeda et al., 2020; Peatman et al., 2021), so we now examine the frequency of occurrence of each regime by MJO phase (figures 6a–d). LARGE-SCALE has a strong peak in phases 4–5, when the active MJO is over the Maritime Continent, which is consistent with large-scale convection being present. The diurnal cycle regimes tend to occur during the suppressed and pre-active phases, 7–2 for DC-NO-PROP and 8–2 for DC-WITH-PROP. NONE occurs fairly frequently during phases 4–5, so in an active MJO environment the diurnal cycle tends to be suppressed even if the large-scale rainfall is weak or absent over this particular region on a given day; but NONE occurs most frequently in phase 6, just after the active envelope has passed through (the pre-suppression phase). These results are all consistent with Peatman et al. (2014), which found that the diurnal cycle has its greatest amplitude just ahead of the arrival of the large-scale active MJO envelope and is most strongly suppressed just ahead of the large-scale suppressed MJO conditions. However, note that here we find that all regimes can occur in all MJO phases, so the MJO does not uniquely determine the regime.

Convection over the Maritime Continent, on the large scale, tends to be enhanced during La Niña, due to the enhancement of the zonal Walker circulation, and suppressed during the opposite El Niño phase (e.g., Hendon, 2003). Moreover, Peatman et al. (2021) found that ENSO phase affects the diurnal cycle and offshore propagation for Sumatra, both being more enhanced during El Niño. Figures 6i–l are consistent with these findings, with El Niño favouring the diurnal cycle regimes and La Niña the LARGE-SCALE regime. ENSO is correlated with IOD, with El Niño most likely to coincide with IOD+ and La Niña with IOD− (e.g., Stuecker et al., 2017). This is consistent with figures 6q–t, histograms of regimes by IOD phase, which show a similarity between the results for El Niño and IOD+, and La Niña and IOD−.

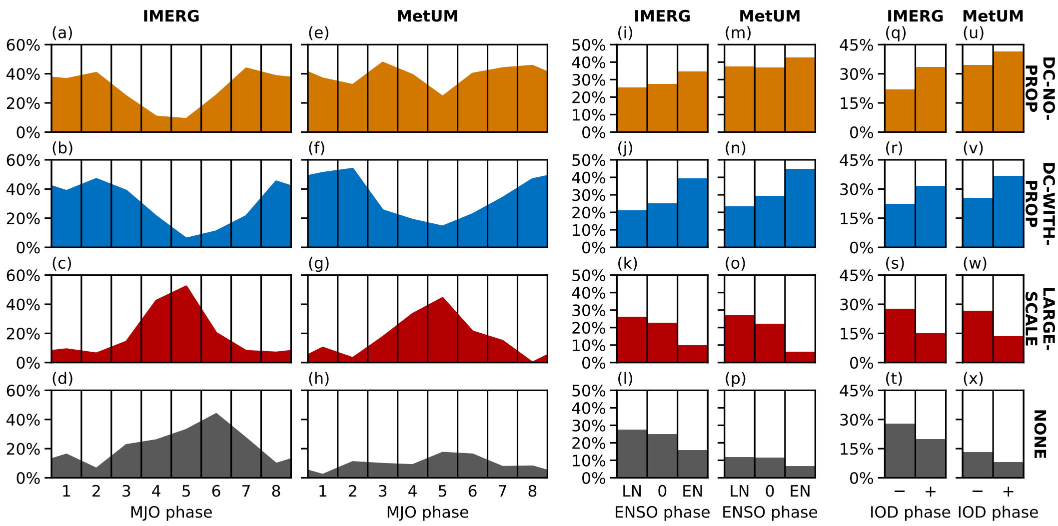

Figure 6. (a–h) For each MJO phase, the percentage of days that fall into each regime. (i–p) Similarly but for ENSO phase (La Niña, LN; neutral phase, 0; and El Niño, EN). (q–x) Similarly but for IOD phase. For each large-scale driver, the first column (a–d,i–l,q–t) is for IMERG observations and the second (e–h,m–p,u–x) for the MetUM simulation.

The remaining panels of figure 6 show equivalent results for the MetUM. For each large-scale driver the results for DC-WITH-PROP and LARGE-SCALE are very similar to the observations. As seen in section 3.1, the MetUM has far too few days in the NONE regime. Here we see that the dearth of NONE days (figure 6h) occurs as a result of active and pre-suppressed MJO days (phases 4–7), when the model is more likely to produce DC-NO-PROP, hence the DC-NO-PROP histogram (figure 6e) failing to reproduce a broad minimum for the active phases, as seen in observations (figure 6a). However, the lack of NONE days and corresponding surfeit of DC-NO-PROP days have no strong dependence on phase when it comes to ENSO or the IOD.

3.3 Physical mechanisms responsible for the diurnal cycle

We now consider the dynamical and thermodynamical conditions that control which of the four regimes a given day falls into. For this we necessarily use only the MetUM simulations. As mentioned in section 1, although there has been much discussion in the literature regarding the role of gravity waves in the offshore propagation, earlier work (Peatman et al., 2023) showed that gravity waves do not play a primary role in the propagation of the organized convection, so here we focus on the mechanism that the latter study found to be most important.

In figures 7a–d are composite maps of the dynamical conditions for each regime at 13 LT. This time was chosen as the earliest time in the day at which the composites clearly differentiate the precipitation patterns between all regimes. In the two diurnal cycle regimes, convection has already been triggered over the mountains and is precipitating; in LARGE-SCALE, precipitation exceeding 1 mm hr^{-1} already covers all of the sea in the domain shown, but there is no triggering of convection over the mountains; and in NONE there is almost no precipitation.

The mean 10 m wind over the low-lying land (on the north-east side of the island – see figure 2) is very similar in all regimes, with predominantly north-westerly flow. However, to the south-west of Sumatra, LARGE-SCALE and NONE have strong alongshore (north-westerly) flow, while the diurnal cycle regimes have much weaker large-scale flow and a sea breeze blowing approximately perpendicular to the coast. This causes low-level convergence over the mountains, seen as a coherent blue region in figures 7a,b and collocated with the diurnal cycle convection. In LARGE-SCALE and NONE the convergence over the mountains is weak and less coherent, contributing to the lack of convection there.

To investigate the thermodynamical conditions we use a moist instability diagnostic, $\theta_e^{300\mathrm{m}} - \theta_{es}^{4500\mathrm{m}}$ (e.g., Birch et al., 2016), where θ_e is equivalent potential temperature, θ_{es} is saturation equivalent potential temperature and heights are measured above the ground. The definition of θ_e is

$$\theta_e = \left(\frac{p_0}{p}\right)^{2/7} T_e \tag{1}$$

$$= \left(\frac{p_0}{p}\right)^{2/7} \left(T + \frac{rL_v}{c_{pd}}\right), \tag{2}$$

where T_e is equivalent temperature (the temperature that an air parcel would have if all water vapour were condensed out and the resulting latent heat used to heat the air parcel), p is pressure, p_0 is a reference pressure (taken to be 1000 hPa), T is temperature, r is humidity mixing ratio, L_v is the latent heat of vaporization of water and c_{pd} is the specific heat capacity of dry air at constant pressure.

$\theta_e^{300\mathrm{m}} - \theta_{es}^{4500\mathrm{m}}$ is an approximate indicator of potential for moist convection (e.g., Garcia-Carreras et al., 2011). Consider an air parcel at low levels (taken here to be 300 m above the ground). If this parcel were lifted pseudo-adiabatically to a mid-tropospheric

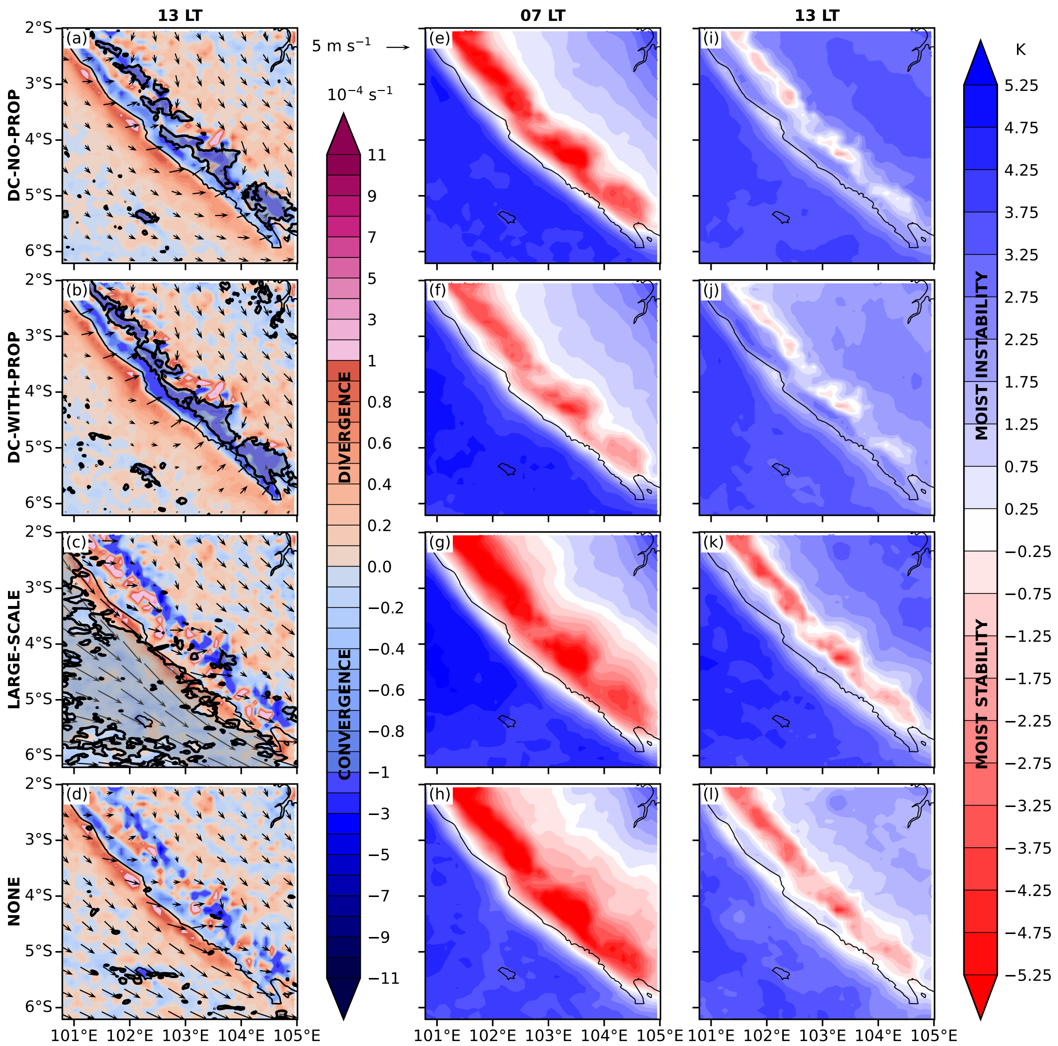

Figure 7. Composite maps at given times for each of the regimes. (a–d) Wind at 10 m (vectors), divergence at 10 m (coloured shading) and precipitation (black contour and transparent grey shading, 1 mm hr^{-1}), at 13 LT. (e–h) $\theta_e^{300\mathrm{m}} - \theta_{es}^{4500\mathrm{m}}$ at 07 LT. (i–l) As (e–h) but for 13 LT.

level where it is bound to be saturated (taken here to be 4500 m) and at that stage it still has positive buoyancy, we can expect it to continue rising and cause deep convection. Hence, the presence of moist instability is suggested by $\theta_e^{300\mathrm{m}} > \theta_{es}^{4500\mathrm{m}}$, or the diagnostic being positive (blue in figures 7e–l).

At both times of day shown, there is large-scale moist instability over the ocean and low-lying land in all four regimes. However, this does not necessarily cause convection to occur in these places, as there is also convective inhibition to be overcome (although note that the instability is strongest in LARGE-SCALE, particularly at 07 LT, in which regime there is indeed rainfall over a wide area). In all regimes there is stability over the mountains at 07 LT, before the onset of diurnally-driven rainfall, although it is weaker in DC-WITH-PROP. By 13 LT the sign of the diagnostic has changed over almost all of the mountains in the diurnal cycle regimes, and convection has duly occurred there. In LARGE-SCALE and NONE, even by this time of day, the atmospheric column is still stable against moist convection, hence the diurnal cycle of convection does not take place.

This raises the question of what causes the differences in $\theta_e^{300\mathrm{m}} - \theta_{es}^{4500\mathrm{m}}$ over the mountains, between the regimes. The mean diurnal cycle of this diagnostic, averaged over the mountains, is plotted in figure 8a. This confirms the distinction between i) the diurnal cycle regimes, where the diagnostic is positive from around 11:30 LT to around 20:00 LT; and ii) LARGE-SCALE and NONE, where the atmospheric column is stable all day.

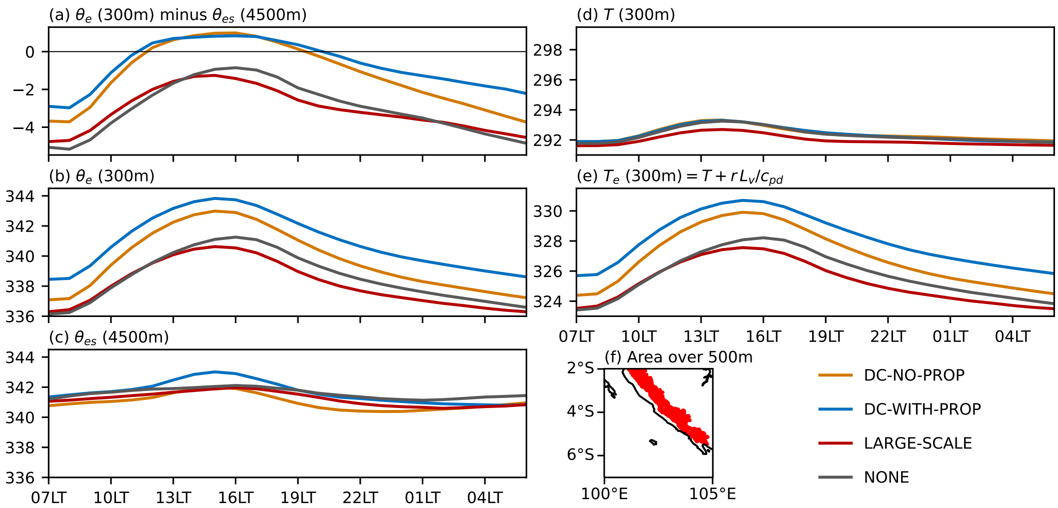

Figure 8. (a–e) Mean diurnal cycles in kelvin, averaged over land greater than 500 m above sea level, for each regime. (a) The moist instability diagnostic $\theta_e^{300\mathrm{m}} - \theta_{es}^{4500\mathrm{m}}$ (where heights are measured above the ground); (b) equivalent potential temperature, $\theta_e^{300\mathrm{m}}$; (c) saturation equivalent potential temperature, $\theta_{es}^{4500\mathrm{m}}$; (d) temperature, T; (e) equivalent temperature, T_e, which is the sum of the temperature in panel (d) and a moisture term rL_v/c_{pd}. Note that panels (b) and (c) have the same vertical scale, and the vertical axes of panels (d) and (e) cover the same size range. (f) Map of the area over 500 m above sea level, over which quantities are averaged.

The mean diurnal cycle of each term is shown in figures 8b ($\theta_e^{300\mathrm{m}}$) and 8c ($\theta_{es}^{4500\mathrm{m}}$). The diurnal variability and the variability between regimes are dominated by $\theta_e^{300\mathrm{m}}$, so it is the low-level conditions that dictate the instability, not the mid-level. We next ask what controls the value of $\theta_e^{300\mathrm{m}}$. This is a function of T_e and p (equation 1). However, the variation in pressure is small enough to have negligible impact on the θ_e (not shown), so the problem reduces to explaining the variability in T_e.

The mean diurnal cycles of T and T_e are shown in figures 8d,e. A visual comparison of figures 8b and 8e shows that they are almost identical in shape, the only difference being the factor of $(p_0/p)^{2/7}$ in the definition of potential temperature (equation 1), confirming that the variation in pressure between the regimes is negligible. Comparing T and T_e (the two panels have vertical scales covering the same size range), we see that for T both the diurnal variability and the variability between regimes are relatively small. Hence, it is the moisture term rL_v/c_{pd} present in figure 8e (see equation 2) that dominates.

Hence, it is variations in low-level humidity that control moist instability over the mountains; and, therefore, it is a combination of the sea breeze (causing convergence) and low-level humidity that cause the diurnal cycle of rainfall to occur.

3.4 Physical mechanisms responsible for the offshore propagation

The dynamical and thermodynamical conditions discussed in the previous section were used to explain the causes of the diurnal cycle regimes *versus* LARGE-SCALE and NONE; but those conditions, at the times of day considered thus far, do not explain the causes of DC-NO-PROP *versus* DC-WITH-PROP. Although there are very slight differences in the location of the precipitation over the mountains at 13 LT (figures 7a,b), with precipitation having formed slightly closer to the coast in DC-WITH-PROP , the substantial differences between the two regimes emerge in the afternoon and evening. Figure 9 shows the dynamical conditions at 18 LT, by which time the difference in propagation behaviour is evident; and at 23 LT, by which time the precipitation in DC-WITH-PROP has extended tens of km offshore.

The 500 hPa winds at 18 LT (figure 10) are mostly westerly in DC-NO-PROP but mostly northerly in DC-WITH-PROP. These mid-level winds conceivably may steer the organized convection through advection. This is consistent with the differences between

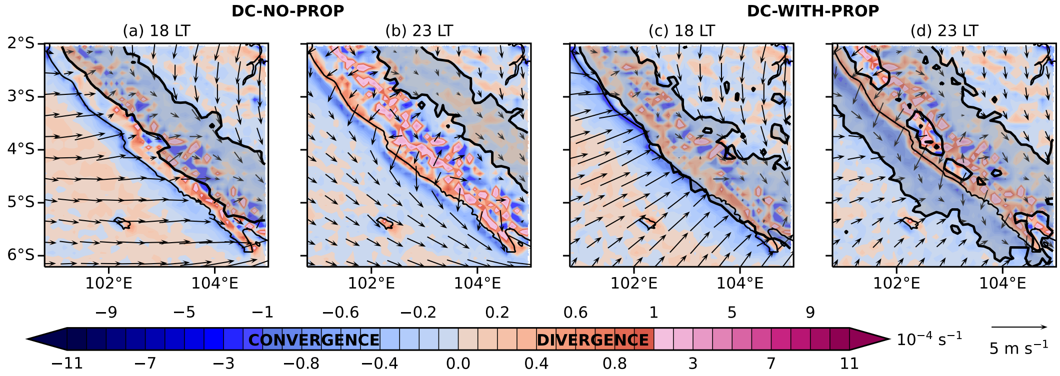

Figure 9. As figures 7a,b but for (a,c) 18 LT and (b,d) 23 LT. The black contour with transparent grey shading is precipitation (1 mm hr^{-1}).

figures 9a and 9c, which show that by 18 LT the precipitation has spread as far as the coast in DC-WITH-PROP, with no equivalent propagation in DC-NO-PROP.

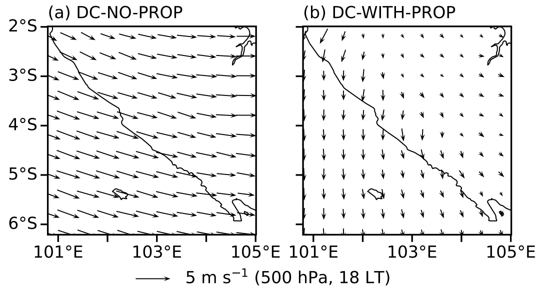

Figure 10. Composite wind at 500 hPa at 18 LT for days with a diurnal cycle (a) and no offshore propagation, and (b) with offshore propagation.

However, for the squall line to continue propagating for several hours, as is observed in DC-WITH-PROP, it is also necessary for the local conditions to sustain the deep convection. In section 1 it was explained that Peatman et al. (2023) attributed this to low-level moisture flux convergence caused by the land breeze. Figure 9 indicates a stronger land breeze for DC-WITH-PROP than DC-NO-PROP. In figure 9a there is very faint convergence along the coast, with a very narrow convergence line that has propagated a few km offshore by 23 LT (figure 9b). Hence, even when the convection does not propagate offshore, there is still a land breeze. However, on days with propagation the convergence at the land breeze front is stronger, with a more intense convergence line seen in figure 9c. Comparing figures 9b and 9d we see that the land breeze has a similar strength in each case, so the difference in convergence is due to the wind direction further offshore. This is directly mainly alongshore for DC-NO-PROP but towards the coast for DC-WITH-PROP. By the time the precipitation has spread around 30–40 km offshore at 23 LT (figure 9d) there is fairly strong convergence in the whole offshore region covered by rainfall. As in Peatman et al. (2023), however, it is not possible to determine to what extent this is caused by cold pools and there may also be a contribution due to convergent inflow into the convection itself.

Further to the above arguments, we consider the properties of air that flows into the storm as it propagates (that is, the storm-relative inflow) and ask what effect this has on the overall convective available potential energy (CAPE), to investigate why the conditions on DC-WITH-PROP days are particularly conducive to sustaining convection. Following Alfaro (2017) we compute the layer-lifting CAPE (CAPE_{ll}), which is the instability of inflowing air, averaged over the troposphere. The storm-relative inflow speed (whether from in front of or behind the storm, relative to its motion) is $|u_{\mathrm{on}}(p) - \mathrm{PS}|$, where u_{on} is the onshore wind speed and PS is the storm's propagation speed, defined as positive onshore. As we consider air parcels flowing into the storm at all heights, the important measure of instability is CAPE(p), the CAPE of an air parcel lifted from pressure level p. Then we define

$$\mathrm{CAPE}_{ll} = \frac{\displaystyle\int_{p_0}^{p_T} |u_{\mathrm{on}}(p) - \mathrm{PS}|\ \mathrm{CAPE}(p)\,\mathrm{d}p}{\displaystyle\int_{p_0}^{p_T} |u_{\mathrm{on}}(p) - \mathrm{PS}|\ \mathrm{d}p}, \tag{3}$$

where p_T is the pressure level of the tropopause, taken to be 120 hPa. Where an air parcel is convectively stable, we take $\mathrm{CAPE}(p) = 0$ by definition. In practice, $\mathrm{CAPE}(p) > 0$ only in the lowest few km of the troposphere. We repeated all calculations using the alongshore flow u_{along} and found the alongshore contribution to be negligible (not shown).

From figure 4j we take PS $= -2.8$ m s^{-1} and compute CAPE_{ll} for the mean profile at each time of day on DC-WITH-PROP days. To test the sensitivity to the PS value, we repeated the calculation (not shown) using PS $= -3.5$ m s^{-1} and -1.8 m s^{-1} (the speeds of the fast and slow edges of the propagation envelope in figure 4j) and there were no substantial differences to the results. For simplicity, the calculation is carried out at a point location over the coastal city of Bengkulu (red dot in figure 2). By way of comparison, we also compute CAPE_{ll} for a hypothetical propagating storm with the same PS on DC-NO-PROP days. The two curves are shown in figure 11a. For DC-WITH-PROP, CAPE$_{ll}$ peaks in the mid-late afternoon and early evening at around 4 times the value for DC-NO-PROP. Hence, the mean conditions for DC-WITH-PROP provide greater instability to sustain propagating convection than would be provided on DC-NO-PROP days.

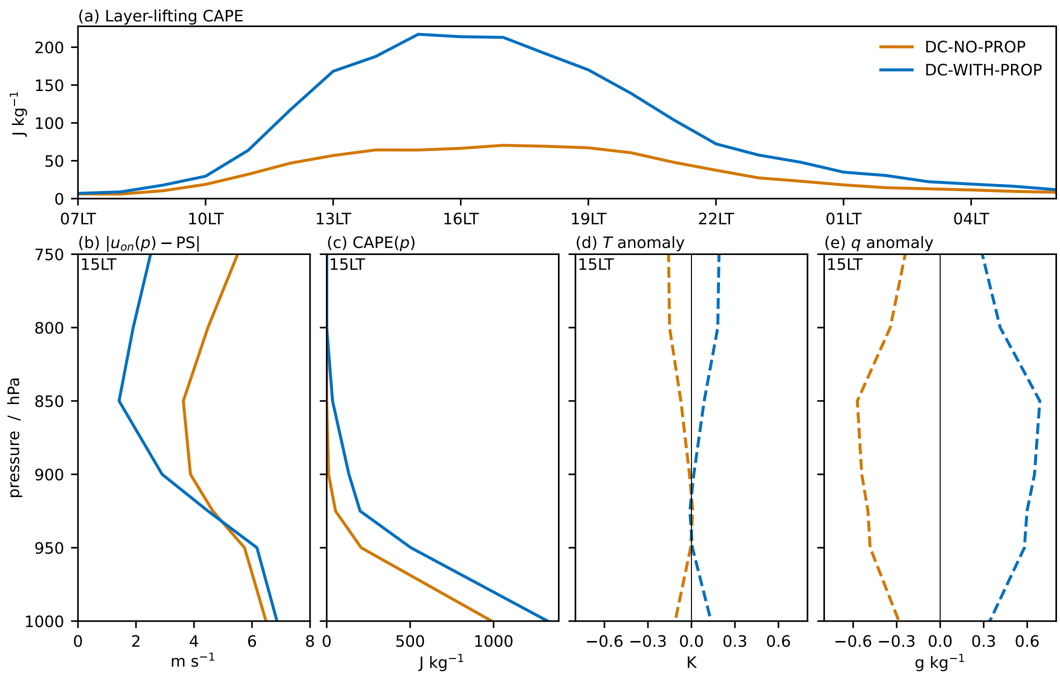

Figure 11. (a) Mean diurnal cycle of layer-lifting CAPE (see equation 3). Bottom row: mean profiles at 15 LT of (b) storm-relative onshore wind, $|u_{\text{on}}(p) - \text{PS}|$, (c) CAPE when an air parcel is lifted from pressure level p, (d) temperature anomaly and (e) specific humidity anomaly. In (d) and (e), the anomalies (indicated by dashed curves) are taken relative to the mean over all days with a diurnal cycle. Calculations are made for mean profiles at Bengkulu (red dot in figure 2).

Figures 11b,c show the two profiles involved in the integral in equation 3, $|u_{\text{on}}(p) - \text{PS}|$ and CAPE(p), at 15 LT, when CAPE$_{ll}$ peaks on DC-WITH-PROP days. Where CAPE(p) is large there is little difference between the two storm-relative wind speeds, whereas the larger difference in storm-relative inflow occurs only where CAPE(p) is very small. Hence, it is the differences in low-level CAPE(p) that have the greatest contribution to CAPE$_{ll}$.

The CAPE is determined by T and specific humidity q, shown as anomalies from all diurnal cycle days in figures 11d,e. By exchanging either the two T profiles or the two q profiles in the CAPE calculation, it is found that only the q makes a substantial difference to the result (not shown). Hence, it is the low-level humidity that ultimately determines that high-instability air flows into propagating convection on DC-WITH-PROP days.

In summary, a number of physical arguments may be made that are consistent with only certain diurnal cycle days having offshore propagation of the convection. The mean mid-level winds have a strong offshore component on DC-WITH-PROP days but are predominantly alongshore, with a slight onshore component, on DC-NO-PROP days, implying there is a steering flow that may begin the process of organized convection propagating from the mountains towards the coast. The land breeze, which previous studies have shown is important for the offshore propagation due to the low-level convergence it provides, appears to be stronger on DC-WITH-PROP days, although there may also be a contribution from cold pools that is not easy to diagnose from the model run used here, as both the land breeze and cold pools are shallow density currents that merge and are not easily distinguished. Finally, the mean low-level moisture around the coast on DC-WITH-PROP days causes larger CAPE with respect to air parcels lifted from near the ground, which contributes more instability to the storm as it propagates, sustaining the convection.

4 Discussion and conclusion

This study has considered the nocturnal offshore propagation of convection south-west of Sumatra, seeking to understand, at a physical process level, why this phenomenon occurs on some days and not others during DJF. Using a convection-permitting model as a research tool, days were identified as having a diurnal cycle of precipitation over land that propagates offshore overnight, a diurnal cycle but with no propagation, large-scale rainfall instead of a diurnal cycle, or no rain at all. In agreement with previous studies (Oh et al., 2012; Peatman et al., 2014; Sakaeda et al., 2017; Vincent & Lane, 2017; Sakaeda et al., 2020; Peatman et al., 2021), the diurnal cycle is most likely to occur ahead of the arrival of an active MJO envelope and the most suppressed local conditions, with little or no rain, ahead of the suppressed MJO envelope. However, the MetUM simulation has too many DC-NO-PROP days, at the expense of NONE. The diurnal cycle is slightly more likely to occur over south-west Sumatra during El Niño or IOD+ than during La Niña or IOD−. Days with large-scale rainfall and days with little or no rainfall are slightly more likely to occur during La Niña or IOD−. However, all regimes may occur in any phase of the MJO, ENSO or IOD, due to local conditions which exert the main control over the rainfall beheviour.

This research has demonstrated the dynamical and thermodynamical conditions responsible for each of the four regimes occurring on any given day. The schematic diagram in figure 12 summarizes the results.

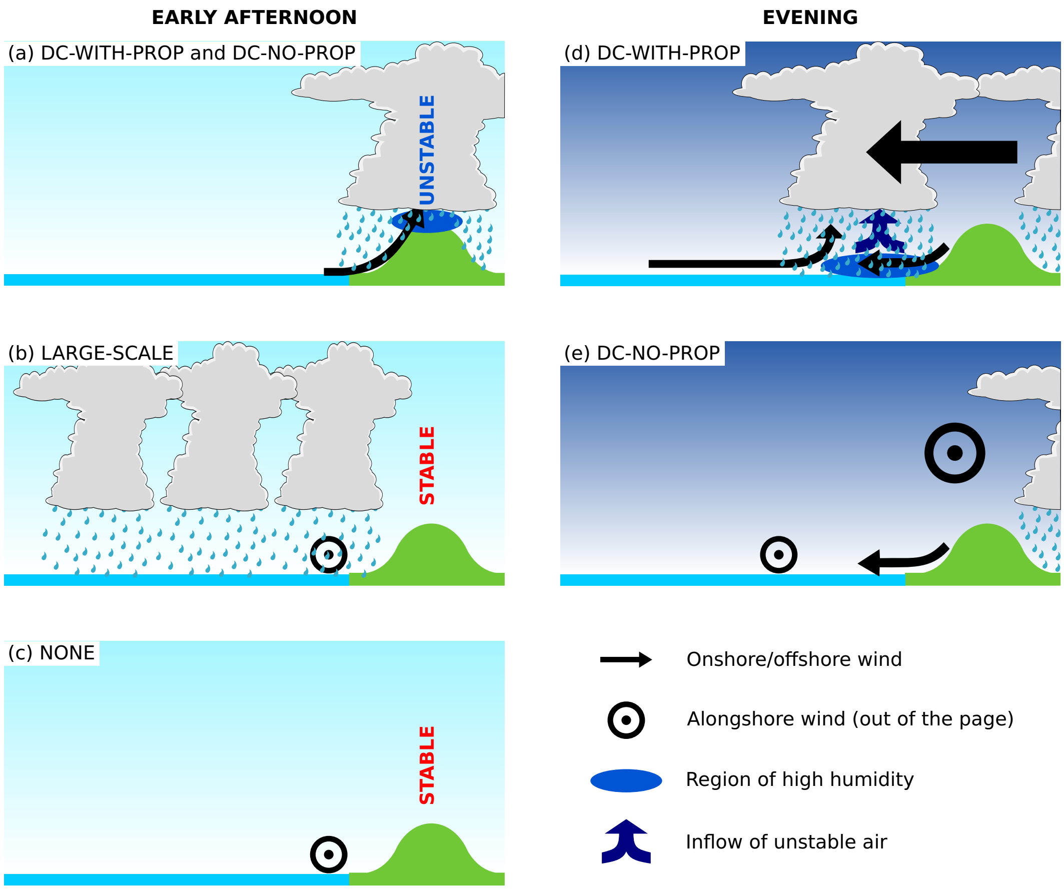

Figure 12. Schematic diagram showing conditions that occur in each of the regimes defined in table 1. Early afternoon: (a) days with a diurnal cycle, (b) LARGE-SCALE and (c) NONE. Evening: (d) DC-WITH-PROP and (e) DC-NO-PROP.

During the morning and up until the early afternoon, the local conditions determine whether the deep convection over the mountains, that constitutes the Sumatra diurnal cycle, will be triggered or not. On days with no diurnal cycle – whether there is large-scale rainfall (figure 12b) or no rainfall nearby (figure 12c) – the vertical profile over the mountains is stable against moist convection. On days with a diurnal cycle (figure 12a), a moist boundary layer over the mountains causes the vertical profile to be unstable, since the higher humidity increases the low-level equivalent potential temperature. Moreover, days with a diurnal cycle have stronger onshore winds, whereas on days with no diurnal cycle the large-scale flow is predominantly alongshore. The onshore, upslope flow provides low-level convergence over the mountains, allowing the triggering of convection in the unstable conditions. The fact that these regimes are more likely to occur ahead (to the east) of the active MJO is in agreement with Birch et al. (2016) who showed that the sea breeze circulation is responsible for local-scale convective triggering ahead of the MJO. However, more research would be required to understand exactly why the onshore sea breeze and low-level humidity are more often associated with El Niño than La Niña, and IOD+ than IOD−.

Understanding the local conditions associated with each regime should in principle assist in forecasting, at least on the day in question and possibly in advance, whether an intense diurnal cycle of convection will occur. However, while local conditions during the morning determine the occurence of the diurnal cycle, they do not determine whether that convection will later propagate offshore.

The important differences in the local conditions do not emerge until the afternoon and evening (figures 12d,e). The wind at mid-levels is predominantly offshore on days with offshore propagation, suggesting it acts as a steering flow as the convection migrates from the mountain peaks to the coast. When there is no offshore propagation, the mid-level wind is predominantly alongshore (from the north-west). However, Peatman et al. (2023) showed the importance of low-level moisture flux convergence due to offshore-propagating density currents (see section 1) for the propagation of the squall line. Here we find that the land-sea breeze circulation is evident on days with and without the offshore propagation of convection, but the wind direction further offshore causes stronger low-level convergence on days with the offshore propagation. Furthermore, due to low-level humidity around the coast, when the convection propagates offshore, the storm-relative inflow of low-level air contributes further instability (CAPE), helping to maintain the convection as it propagates. On days without offshore propagation the coastal low-level humidity is less, so that a storm propagating at the same speed would have less instability added by the storm-relative inflow. Taken together, all these differences between the two regimes present a picture of how the physical mechanisms governing the occurrence of the propagation vary between days.

Peatman et al. (2023) noted a difficulty in the analysis of density currents that comprise the land-sea breeze circulation, namely that cold pools caused by evaporation of rainfall also propagate as density currents, and it is not simple to distinguish the two. In the simulations we use here, the same problem occurs, so we must acknowledge that the relative importance of the land breeze and cold pools, in providing the low-level convergence that forces convection to develop progressively offshore, remains unknown. Future work will use simulations that include a tracer, generated where there is evaporation and subsequently advected around, to quantify the contribution of cold pools to the convergence.

A further limitation of the present study is that, although low-level humidity is found to cause the instability responsible for the diurnal convection over the mountains and the higher-CAPE air flowing into the convection as it propagates, it is not clear what controls this low-level humidity. Analysis not shown here indicates that there is no single cause of this humidity variability, with possible roles for interannual variability and a build-up of humidity over a timescale of 2–3 days, possibly due to evapotranspiration. However, quantifying the various contributions to the humidity is beyond the scope of this study, so further investigation is required.

Notwithstanding these limitations, the present study builds on existing work to aid understanding of how local diurnal cycle and propagation behaviour arises from dynamical and thermodynamical conditions, with potential benefits for forecasters and model developers. State-of-the-art global models are still typically run on grids that cannot resolve coastal processes well, so accurate parametrizations such as land-sea breeze schemes are of great importance to the modelling community. Work such as the present study gives insights into the coastal processes that must be parametrized. This study also demonstrates the power of a convolutional neural network in automating the identification of regimes of coastal precipitation behaviour, provided there is a large enough training data available. Future work may generalize the results of this study by applying a pre-trained machine learning model to coastlines across the global tropics.

We note also that low-level humidity is a major control on the diurnal cycle and its offshore propagation. The next generation of geostationary satellites will provide low-level humidity observations at high spatial and temporal resolution, using infrared sounders (e.g., ESA, 2023). The results of this study provide an example of the value of assimilating these observations into models, given the impact on convection and, therefore, on high-impact precipitation.

Open Research Section

Raw and processed model data are available along with analysis code at `https://doi.org/10.5281/zenodo.13747261` (Peatman et al., 2024).

GPM IMERG precipitation (Huffman et al., 2019) is available from the National Aeronautics and Space Administration (NASA) Goddard Earth Sciences Data and Information Services Center (GES DISC; `https://pmm.nasa.gov/data-access/downloads/gpm`). RMM indices are available from the Bureau of Meteorology (BoM, 2024, `www.bom.gov.au/climate/mjo/graphics/rmm.74toRealtime.txt`). GLOBE orography (Hastings et al., 1999) is available from the National Oceanic and Atmospheric Administration (NOAA; `https://www.ngdc.noaa.gov/mgg/topo/globe.html`).

Acknowledgments

This research was funded by the National Environment Research Council (NERC) through the TerraMaris project, grants NE/R016739/1, NE/R016704/1 and NE/R016712/1. JS was funded by the Met Office Weather and Climate Science for Service Partnership (WCSSP) Southeast Asia. SJW was also supported by the National Centre for Atmospheric Science, a NERC collaborative centre.

Data analysis was carried out using the Python `iris` package, version 3.3.1 (Iris Contributors, 2022). The CNN was created using the Python `keras` package, version 2.11.0 (Chollet et al., 2015).

We thank Thi Lan Dao and Dr Bethan White for their feedback on this manuscript; and Dr Claire Vincent and two anonymous reviewers, for their comments which improved the manuscript prior to publication.

References

Alfaro, D. A. (2017). Low-Tropospheric Shear in the Structure of Squall Lines: Impacts on Latent Heating under Layer-Lifting Ascent. *Journal of the Atmospheric Sciences*, *74*(1), 229–248. doi: 10.1175/JAS-D-16-0168.1

Aoki, S., & Shige, S. (2024). Control of Low-Level Wind on the Diurnal Cycle of

Tropical Coastal Precipitation. *Journal of Climate*, *37*(1), 229–247. doi: 10.1175/JCLI-D-23-0180.1

Biasutti, M., Yuter, S. E., Burleyson, C. D., & Sobel, A. H. (2012). Very high resolution rainfall patterns measured by TRMM precipitation radar: Seasonal and diurnal cycles. *Climate Dynamics*, *39*(1–2), 239–258. doi: 10.1007/s00382-011-1146-6

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*(7), 2471–2492. doi: 10.1175/JCLI-D-15-0557.1

BoM. (2024). *Realtime Multivariate MJO indices [Dataset].* Retrieved from `http://www.bom.gov.au/clim_data/IDCKGEM000/rmm.74toRealtime.txt`

Bush, M., Boutle, I., Edwards, J., Finnenkoetter, A., Franklin, C., Hanley, K., . . . Weeks, M. (2023). The second Met Office Unified Model-JULES Regional Atmosphere and Land configuration, RAL2. *Geoscientific Model Development*, *16*(6), 1713–1734. doi: 10.5194/gmd-16-1713-2023

Chollet, F., et al. (2015). *Keras.* `https://keras.io`.

ESA. (2023). *MTG (Meteosat Third Generation).* Retrieved from `https://www.eoportal.org/satellite-missions/meteosat-third-generation` (Accessed: 5 September 2024)

Fujita, M., Takahashi, H. G., & Hara, M. (2013). Diurnal cycle of precipitation over the eastern Indian Ocean off Sumatra Island during different phases of Indian Ocean Dipole. *Atmospheric Science Letters*, *14*(3), 153–159. doi: 10.1002/asl2.432

Garcia-Carreras, L., Parker, D. J., & Marsham, J. H. (2011). What is the Mechanism for the Modification of Convective Cloud Distributions by Land Surface-Induced Flows? *Journal of the Atmospheric Sciences*, *68*(3), 619–634. doi: 10.1175/2010JAS3604.1

Hastings, D. A., Dunbar, P. K., Elphingstone, G. M., Bootz, M., Murakami, H., Maruyama, H., . . . MacDonald, J. S. (1999). *The Global Land One-kilometer Base Elevation (GLOBE) digital elevation model, version 1.0 [Dataset].* Retrieved from `http://www.ngdc.noaa.gov/mgg/topo/globe.html`

Hendon, H. H. (2003). Indonesian Rainfall Variability: Impacts of ENSO and Local

Air-Sea Interaction. *Journal of Climate*, *16*(11), 1775–1790. doi: 10.1175/1520-0442(2003)016⟨1775:IRVIOE⟩2.0.CO;2

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. doi: 10.1002/qj.3803

Houze, R. A., Geotis, S. G., Marks, F. D., & West, A. K. (1981). Winter monsoon convection in the vicinity of north Borneo. Part I: Structure and time variation of the clouds and precipitation. *Monthly Weather Review*, *109*(8), 1595–1614. doi: 10.1175/1520-0493(1981)109⟨1595:WMCITV⟩2.0.CO;2

Howard, E., Woolnough, S., Klingaman, N., Shipley, D., Sanchez, C., Peatman, S. C., ... Matthews, A. J. (2024). Evaluation of multi-season convection-permitting atmosphere - mixed-layer ocean simulations of the Maritime Continent. *Geoscientific Model Development*, *17*(9), 3815–3837. doi: 10.5194/gmd-17-3815-2024

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., & Tan, J. (2019). *GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06 [Dataset].* doi: 10.5067/GPM/IMERG/3B-HH/06

Iris Contributors. (2022). *Iris.* Retrieved from `https://github.com/SciTools/iris` doi: 10.5281/zenodo.595182

Kurniadi, A., Weller, E., Min, S.-K., & Seong, M.-G. (2021). Independent ENSO and IOD impacts on rainfall extremes over Indonesia. *International Journal of Climatology*, *41*(6), 3640–3656. doi: 10.1002/joc.7040

Large, W. G., McWilliams, J. C., & Doney, S. C. (1994). Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Reviews of Geophysics*, *32*(4), 363–403. doi: 10.1029/94RG01872

Love, B. S., Matthews, A. J., & Lister, G. M. S. (2011). The diurnal cycle of precipitation over the Maritime Continent in a high-resolution atmospheric model. *Quarterly Journal of the Royal Meteorological Society*, *137*, 934–947. doi: 10.1002/qj.809

Mapes, B. E., Warner, T. T., & Xu, M. (2003). Diurnal patterns of rainfall in northwestern South America. Part III: Diurnal gravity waves and nocturnal convection offshore. *Monthly Weather Review*, *131*(5), 830–844. doi: 10.1175/1520-0493(2003)131⟨0830:DPORIN⟩2.0.CO;2

Mapes, B. E., Warner, T. T., Xu, M., & Negri, A. J. (2003). Diurnal patterns of rainfall in northwestern South America. Part I: Observations and context. *Monthly Weather Review*, *131*(5), 799–812. doi: 10.1175/1520-0493(2003)131⟨0799:DPORIN⟩2.0.CO;2

Montserrat, D. M., Lin, Q., Allebach, J., Delp, E. J., Lin, Q., Allebach, J., & Delp, E. J. (2017). Training Object Detection And Recognition CNN Models Using Data Augmentation. *Electronic Imaging*, *29*, 27–36. doi: 10.2352/ISSN.2470-1173.2017.10.IMAWM-163

Mori, S., Jun-Ichi, H., Tauhid, Y. I., & Yamanaka, M. D. (2004). Diurnal land-sea rainfall peak migration over Sumatera Island, Indonesian Maritime Continent, observed by TRMM satellite and intensive rawinsonde soundings. *Monthly Weather Review*, *132*(8), 2021–2039. doi: 10.1175/1520-0493(2004)132⟨2021:DLRPMO⟩2.0.CO;2

Oh, J. H., Kim, K. Y., & Lim, G. H. (2012). Impact of MJO on the diurnal cycle of rainfall over the western Maritime Continent in the austral summer. *Climate Dynamics*, *38*(5–6), 1167–1180. doi: 10.1007/s00382-011-1237-4

Peatman, S. C., Birch, C., Schwendike, J., Marsham, J., Howard, E., Woolnough, S., ... Matthews, A. (2024). *Physical controls on the variability of offshore propagation of convection from Sumatra [Dataset].* doi: 10.5281/zenodo.13747262

Peatman, S. C., Birch, C. E., Schwendike, J., Marsham, J. H., Dearden, C., Webster, S., ... Matthews, A. J. (2023). The Role of Density Currents and Gravity Waves in the Offshore Propagation of Convection over Sumatra. *Monthly Weather Review*, *151*(7), 1757–1777. doi: 10.1175/MWR-D-22-0322.1

Peatman, S. C., Matthews, A. J., & Stevens, D. P. (2014). Propagation of the Madden-Julian Oscillation through the Maritime Continent and scale interaction with the diurnal cycle of precipitation. *Quarterly Journal of the Royal Meteorological Society*, *140*(680), 814–825. doi: 10.1002/qj.2161

Peatman, S. C., Schwendike, J., Birch, C. E., Marsham, J. H., Matthews, A. J., & Yang, G.-Y. (2021). A Local-to-Large Scale View of Maritime Continent Rainfall: Control by ENSO, MJO, and Equatorial Waves. *Journal of Climate*, *34*(22), 8933–8953. doi: 10.1175/JCLI-D-21-0263.1

Poojary, R., Raina, R., & Mondal, A. K. (2021). Effect of data-augmentation on fine-tuned CNN model performance. *IAES International Journal of Artificial*

Intelligence (IJ-AI), *10*(1), 84–92. doi: 10.11591/ijai.v10.i1.pp84-92

Qian, J.-H. (2008). Why precipitation is mostly concentrated over islands in the Maritime Continent. *Journal of the Atmospheric Sciences*, *65*(4), 1428–1441. doi: 10.1175/2007JAS2422.1

Ramage, C. S. (1968). Role of a Tropical "Maritime Continent" in the atmospheric circulation. *Monthly Weather Review*, *96*(6), 365–370. doi: 10.1175/1520-0493(1968)096⟨0365:ROATMC⟩2.0.CO;2

Rauniyar, S. P., & Walsh, K. J. E. (2013). Influence of ENSO on the diurnal cycle of rainfall over the Maritime Continent and Australia. *Journal of Climate*, *26*(4), 1304–1321. doi: 10.1175/JCLI-D-12-00124.1

Sakaeda, N., Kiladis, G., & Dias, J. (2020). The diurnal cycle of rainfall and the convectively coupled equatorial waves over the Maritime Continent. *Journal of Climate*, *33*(8), 3307–3331. doi: 10.1175/JCLI-D-19-0043.1

Sakaeda, N., Kiladis, G. N., & Dias, J. (2017). The diurnal cycle of tropical cloudiness and rainfall associated with the Madden-Julian Oscillation. *Journal of Climate*, *30*(11), 3999–4020. doi: 10.1175/JCLI-D-16-0788.1

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, *6*(1), 60. doi: 10.1186/s40537-019-0197-0

Stuecker, M. F., Timmermann, A., Jin, F.-F., Chikamoto, Y., Zhang, W., Wittenberg, A. T., … Zhao, S. (2017). Revisiting ENSO/Indian Ocean Dipole phase relationships. *Geophysical Research Letters*, *44*(5), 2481–2492. doi: 10.1002/2016GL072308

Vincent, C. L., & Lane, T. P. (2017). A 10-year austral summer climatology of observed and modeled intraseasonal, mesoscale, and diurnal variations over the Maritime Continent. *Journal of Climate*, *30*(10), 3807–3828. doi: 10.1175/JCLI-D-16-0688.1

Webster, P. J., & Hoyos, C. D. (2004). Prediction of Monsoon Rainfall and River Discharge on 15-30-Day Time Scales. *Bulletin of the American Meteorological Society*, *85*(11), 1745–1765. doi: 10.1175/BAMS-85-11-1745

Yang, G.-Y., & Slingo, J. M. (2001). The Diurnal Cycle in the Tropics. *Monthly Weather Review*, *129*(1994), 784–801. doi: 10.1175/1520-0493(2001)129⟨0784:TDCITT⟩2.0.CO;2