Evaluation of the model representation of the evolution of convective systems using satellite observations of outgoing longwave radiation

K. J. Pearson,[1] R. J. Hogan,[2] R. P. Allan,[2] G. M. S. Lister,[3] and C. E. Holloway[4]

[1]   We introduce a technique for assessing the diurnal development of convective storm systems based on outgoing longwave radiation fields. Using the size distribution of the storms measured from a series of images, we generate an array in the length scale-time domain based on the standard score statistic. It demonstrates succinctly the size evolution of storms as well as the dissipation kinematics. It also provides evidence related to the temperature evolution of the cloud tops. We apply this approach to a test case comparing observations made by the Geostationary Earth Radiation Budget instrument to output from the Met Office Unified Model run at two resolutions. The 12 km resolution model produces peak convective activity on all length scales significantly earlier in the day than shown by the observations and no evidence for storms growing in size. The 4 km resolution model shows realistic timing and growth evolution, although the dissipation mechanism still differs from the observed data.

1. Introduction

[2]   Deep convection constitutes the fundamental building block of tropical weather, producing cloud systems that organize and interact on many spatial scales. These span the range from individual convection cells, through clusters of intense thunderstorms embedded in mesoscale systems and disturbances such as African Easterly Waves, to superclusters related to global circulation patterns such as the Madden Julian Oscillation [*Leary and Houze*, 1979; *Machado et al.*, 1993]. Similarly, a large range of temporal scales are involved from the diurnal cycle through to interannual modes of variability.

[3]   The upscaling of energy from the smaller and high-frequency scales to the larger and low-frequency regimes is still poorly understood. The Cascade project (Woolnough, manuscript in preparation, 2010) is designed to address this topic by running high-resolution cloud system resolving models (1.5 km) over large domains (~4000 km) using the Met Office Unified Model (UM). The intention is to capture both the small-scale convection physics and large-scale cir-culation patterns in the same simulation. As one element, we are examining the diurnal development of convective systems over Africa and we shall address this in detail in a forthcoming paper. We present here a method developed to test the diurnal cycle as represented in numerical models against observation.

[4]   The fundamental starting point for representation of the diurnal behavior of tropical convective systems is simply to plot cloud fraction or some other suitable proxy variable as a function of time of day. Such an approach, as in the early study of satellite infrared measurements by *Gruber* [1976] for example, rapidly reveals the tendency for peak activity to occur in the early evening. Repeating this for multiple locations which can be overplotted is then trivial and additionally begins to reveal spatial information. A more sophisticated form of this idea, in the era of high-resolution satellite imagery, is to break a scene down to the pixel level and separately Fourier analyze each in time as in *Yang and Slingo* [2001]. These authors compared the results of a simulation run using the UM with the Cloud Archive User Service (CLAUS) data set [*Hodges et al.*, 2000]. Their model was run for a simulated year for all longitudes in the 30°S–30°N latitude band with a resolution of 3.75° in longitude and 2.5° in latitude. They conducted their comparative analysis through two variables acting as proxies for convective activity: window channel brightness temperature and precipitation rate. At this resolution, these are likely detecting mesoscale convective systems or larger aggregations. For both variables, the phase of the diurnal Fourier components indicated that the models produced peak convective activity much earlier in the day than the observations.

[1]Environmental Systems Science Centre, University of Reading, Reading, UK.
[2]Department of Meteorology, University of Reading, Reading, UK.
[3]Computational Modeling Services, National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK.
[4]Climate Directorate, National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK.
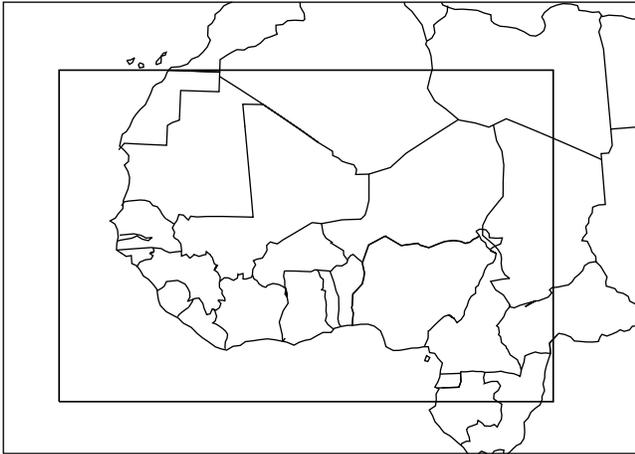
Figure 1. The West Africa test area covered by the 12 km model and the subset of the GERB data with the nested domain of the 4 km model.

[5] A generalization of the Fourier approach that includes spatial variability is principal component analysis with empirically orthogonal functions as used by *Murakami* [1980], *Lau and Chan* [1985], and *Comer et al.* [2007] inter alia. This latter paper used data from the Geostationary Earth Radiation Budget (GERB) instrument [*Harries et al.*, 2005] to study the diurnal cycle of outgoing long-wave radiation (OLR). It showed that 95% of the variation could be explained by the first two components and that the development of cloudiness was strongly influenced by orographic features.

[6] While such studies yield spatiotemporal information, they do not relate directly to the evolutionary processes in a system's lifecycle. It is possible to tackle this using a tracking method that allows individual storms to be followed from image to image over time [e.g., *Williams and Houze*, 1987; *Futyan and Del Genio*, 2007]. This enables all the measurable parameters to be labeled with position and time coordinates and plotted accordingly. While this approach is now quite common, the implementation can be time-consuming and sensitive to the many parameters that are required. In addition, to make sense of such information, the data still need to be reaggregated in some form else we are left with individual case studies of individual systems. The method we set out here is fast and has essentially a single parameter that rapidly produces a representation of the evolution of storm systems.

2. Method

[7] The model data being tested was generated using the nonhydrostatic UM version 7.1 in a configuration previously used by *Lean et al.* [2008, 2009]. In summary, the model solves the atmospheric dynamical equations using a semi-implicit, semi-Lagrangian scheme on a rotated latitude-longitude grid [*Davies et al.*, 2005; *Cullen et al.*, 1997]. It utilizes parameterizations to represent, inter alia, subsurface and surface fluxes [*Essery et al.*, 2001], boundary layer turbulence [*Lock et al.*, 2000], mixed-phase microphysics [*Wilson and Ballard*, 1999] with multiple hydrometeor components and, in general, convection [*Gregory and*

Rowntree, 1990]. The model was run in a limited area mode forced with constant sea surface temperatures and initiated with analysis fields from the European Centre for Medium-Range Weather Forecasts. These fields were subsequently updated only at the lateral boundaries.

[8] The observed data for comparison was provided by GERB. This is a broadband radiometer on the Meteosat-8 satellite producing data at a nadir resolution of 50 km. The satellite is positioned on the equator at 3.5°W providing a field of view covering the whole of Africa and stretching to Northern Europe, the Middle East and parts of South America. GERB makes radiation measurements every 6 min enabling it to follow the evolution of medium- to large-scale weather systems. We made use of a hybrid product (NRT V003 ARCH) that additionally includes information from the Spinning Enhanced Visible and InfraRed Imager (SEVIRI), also on Meteosat-8, to produce high-resolution OLR measurements with approximately 10 km resolution. The product is described in *Dewitte et al.* [2008] where it is termed "Standard High-Resolution Image (SHI)" and it has been used previously to study the effect of Saharan dust on the atmospheric radiation balance [*Slingo et al.*, 2006].

[9] The test area covered West Africa and used 10 days of model data running from 26 July to 4 August and 17 days of GERB observations from 22 July to 7 August during 2006. The model was run at two resolutions: with 460 by 340 pixels of 0.11° and with 1110 by 776 pixels of 0.036°, corresponding to resolutions at the center of the domain of 12 km and 4 km, respectively. The convection scheme differed between the two models in that the 12 km model scheme used a convectively available potential energy closure method [*Gregory and Rowntree*, 1990], whereas the 4 km model used a version in which the mass flux at the cloud base was limited. In practice, the 4 km resolution model was tuned such that almost all of the convection was represented explicitly rather than through parameterization. This configuration leads to an improved representation of larger storms at the expense of missing some weak showers [*Lean et al.*, 2008]. The 12 km model included two components in the microphysics scheme (cloud liquid and frozen water) and 38 vertical levels rather than three hydrometeor components (additionally prognostic rain) and 70 vertical levels in the 4 km model. The 12 km model was run with a 300 s time step and the 4 km with a 60 s time step for the first 5 days and a 30 s time step thereafter. In addition to an improvement in the direct representation of atmospheric processes, *Hohenegger et al.* [2009] showed how moving to an explicit convection, cloud-resolving model resulted in a radical change to the behavior of land-atmosphere feedbacks.

[10] The GERB data were subsetted and regridded to match the area and resolution of the 12 km model data (approximately 5°S–35°N, 25°W–25°E). The 4 km model was nested inside the 12 km model region and covered a smaller area (approximately 0°N–30°N, 20°W–20°E). These regions are shown in Figure 1. Both the model output and the observed data products were generated at 15 min intervals. As the models were driven only at the boundary by analysis fields, we are not able to match individual observed clouds to the model. Thus, comparisons with observation rely on indirect statistical methods.
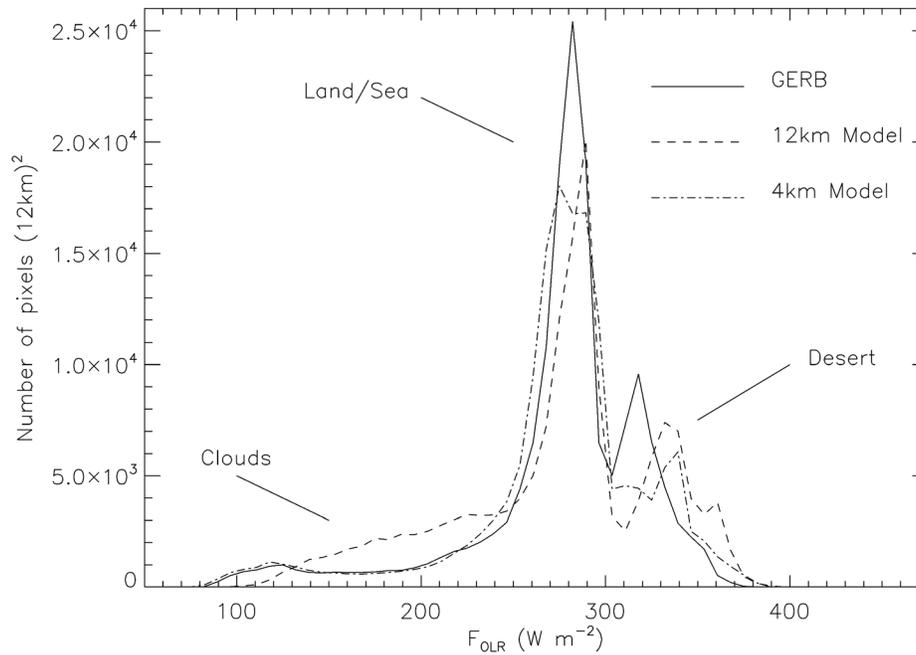
Figure 2. The frequency distribution of OLR fluxes for the GERB data (solid), 12 km model (dashed), and 4 km model (dot-dashed) at 27 July 2006::15:00. The 4 km model values have been normalized to ensure the same total number of pixels as the other two data sets. The mismatch between model and observations in the "Desert" region most likely results from the lack of mineral dust aerosols in the model [*Haywood et al.*, 2005].

[11] Pixels were identified as cloud on the basis of being below a threshold OLR flux (F_{th}). An alternative approach would be to use SEVIRI window channel data for the observations and apply a more sophisticated algorithm to identify clouds which could then be compared to internally flagged cloudy regions from the model. However, as we intend to study the model representation of convection, we prefer the approach that the burden should rest with the numerical simulations to produce directly observable quantities. These can then be processed in an identical way for both observations and models. With the observations then acting as "truth," any differences must rest with some aspect of the modeling. The alternative approach leaves open the possibility that the cloud detection algorithm applied to the observations is the source of the discrepancies.

[12] It is a common approach to identify cold cloud regions with sites of deep convection. However, studies such as those by *Rickenbach* [1999] and *Rickenbach et al.* [2008] indicate that as propagating systems mature, the locations of cold cloud can become separated from those of precipitation and that these cloudy regions can continue to grow despite becoming decoupled from the original convection cells. This can complicate detailed interpretation of the convective behavior. However, understanding the diurnal cycle of the cloud cover itself is of direct relevance to studies of the radiative processes and feedbacks that are important to both weather and climate prediction. In particular, both the sign and magnitude of cloud radiative forcing varies according to the diurnal phase. During the day, the reflection of incoming radiation by cloud tends to dominate over the tendency to trap the OLR emitted by the surface that is the dominant effect at night. Cloud feedbacks are the largest uncertainties in climate sensitivity estimates [*Randall et al.*, 2007].

[13] *Mapes and Houze* [1993] show the wide range of blackbody temperatures (188–267 K) that could be used to determine a cloudy region. Among these, *Fu et al.* [1990] use temperature thresholds of 215 K for deep convective clouds and 267 K for convective anvil cloud. Our fluxes are not directly comparable, however, as they are broadband measurements that include absorption features, in contrast to an effective temperature based on narrow window channel measurements. *Fu et al.* [1990] also quote a value of 240 W m^{-2} as an OLR threshold often used to diagnose deep convection. We use F_{th} = 150 W m^{-2} here initially which corresponds to an effective *broadband* emission temperature ($T_{b,e}$) of 227 K. The OLR flux distributions plotted in Figure 2 show that this conservatively restricts the sample to cold cloud tops, minimizing any possible noncloud contamination. While increasing this threshold increases the range of length scales being sampled, it reduces the ability to identify the coldest regions at small scales arising from deep convection. Contiguous regions of pixels below F_{th} sharing an edge (i.e., not including diagonal adjacency) were then identified as clusters using the intrinsic IDL routine label_region and the area (A), based on the nominal equatorial resolution, and length scale ($L = \sqrt{A}$) calculated.

[14] To examine the character of the diurnal signal, the cluster length scale distribution at each time (t) was normalized by calculating the standard score statistic (Z) for the number of systems (N) in each length scale bin

$$Z(L,t) = \frac{N(L,t) - \overline{N}(L)}{\sigma(L)}, \qquad (1)$$

using the measured mean ($\overline{N}$) and measured standard deviation (σ) over time for each length scale bin separately.
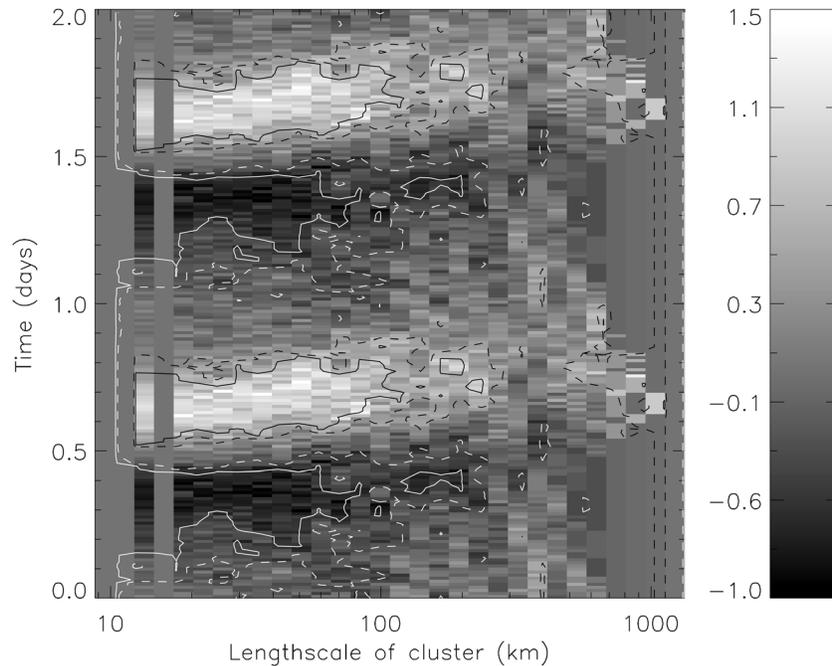
Figure 3. Normalized time series of cloud cluster size distributions $Z(L,t)$ observed by GERB with $F_{th} = 150$ W m^{-2}. The data have been folded on a period of 1 day and repeated for clarity. Overplotted are the probability contours $p = 0.01, 0.1, 0.9, 0.99$ for the observed number of systems or fewer to occur randomly from a Poisson distribution with the observed mean number for each length scale. The probability field has been smoothed in the time direction by a 1 h running mean.

Taking the Fourier transform of this array in the time direction confirms the presence of a dominant signal on a period of 1 day. Accordingly, while calculations were carried out on the full field, the plotted results are folded on the daily cycle (i.e., diurnally composited).

3. Discussion

[15] Equation (1) measures the deviation from the mean number of storms at each size in multiples of the measured standard deviation for that size bin. As such, it is essentially "self-normalized" by the data set and makes no assumption about any underlying spatiotemporal distribution of the data. The values generated from equation (1) for the GERB data are shown in the gray scale image in Figure 3. The length scale bins in this and subsequent figures are logarithmic with edge values given by $L_i = 2^{2+i/4}$ km.

[16] In the following analysis, we proceed on the basis that the data is composed of two dominant components. Fundamentally, we assume that there is an underlying random distribution for the number of clouds in each scene which, being essentially counting statistics, is Poisson. We also assume that there is an additional diurnal signal that results in no net increase in the number of storms when averaged over the daily cycle but has a potentially large variance.

[17] In order to find regimes where the diurnal forcing is having an impact on the number of systems, the significance of each measured value of $N(L,t)$ can be calculated from a Poisson distribution with the appropriate mean $\overline{N}(L)$ for each length scale bin. The contours for several probability levels are overplotted on Figure 3. These contours indicate

the size regimes and time of day where there is significant evidence for the effect of a nonrandom signal in addition to the underlying Poisson distribution. We should note in passing that, in a situation with large $\overline{N}$ and a weak diurnal signal, the Poisson distribution could be approximated as Gaussian with $\sigma = \sqrt{\overline{N}}$ and the probability calculation would reduce to a simpler case based directly on Z.

[18] The images formed from equation (1) for the three data sets are plotted alongside each other in Figure 4. This figure provides an elegant tool to compare and contrast the diurnal behavior of convection in the different data sets. Similar to the way that a Hovmöller plot demonstrates the trajectory of some anomaly against a position coordinate, these plots represent the evolution in length scale space. The bright areas show, for each data set, the times when there are more storm systems than average at each length scale. The GERB data show a broad upward stripe beginning shortly after midday, fitting with the paradigm of small convective systems forming, growing, and merging during the day. In contrast, the 12 km model data exhibit a flat response across all length scales: showing that storm systems of all sizes switch on and off together. There is a tendency for parameterized schemes to generate deep convection rapidly, shortly after sunrise [*Guichard et al.*, 2004; *Grabowski et al.*, 2006] rather than generating steadily deepening convection cells. The parameterized convection scheme also triggers convection independently in grid cells. As a result, cold cloud tops are generated very early in the day with an initial cluster size distribution reflecting the similarity or otherwise of the conditions in adjacent grid cells.

[19] The 4 km model bears much closer comparison to the observations with a progression in the excess number of
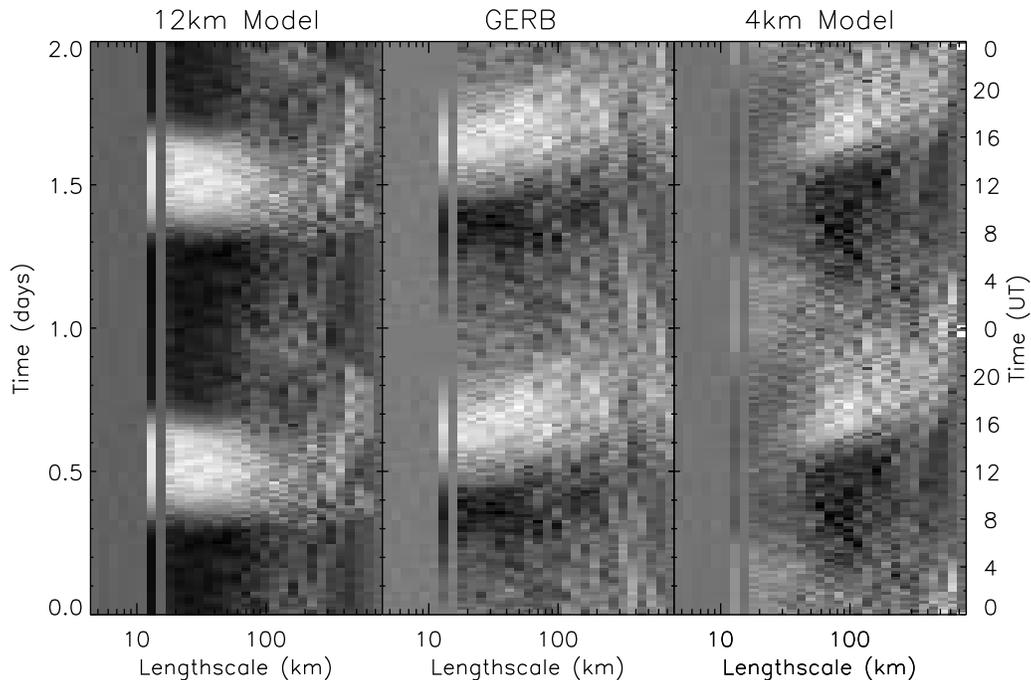
Figure 4. Normalized time series of cloud cluster size distributions $Z(L,t)$ observed by GERB and produced by the two model runs with a threshold flux for cloud detection of 150 W m^{-2}. The data have been folded on a period of 1 day and repeated. The gray scale for each frame has been set independently to ensure each utilizes the entire contrast range.

systems from medium to large length scales at a similar rate to the observations. An intriguing additional feature, however, is that the peak for smaller systems occurs early in the morning. From animations of the OLR this appears to correspond to the way that the large systems in the model tend to "shatter" overnight into many small systems rather than the more organic "dissolving" behavior exhibited by the GERB data. Two series of snapshots illustrating this for the 4 km model and GERB data are shown in Figures 5 and 6, respectively.

[20] The improvement in the representation of convection with resolution is consistent with the results of *Guichard et al.* [2004] who compared the behavior of seven single column models (SCM) with parameterized convection to that of three cloud resolving models (CRM) with horizontal resolutions of 2 km or better. The SCMs generally showed precipitation beginning early in the day and continuing until evening. In contrast, the CRMs produced precipitation strongly peaked in the late afternoon and early evening. While our 4 km model has too coarse a resolution to be truly cloud resolving and still includes convective parameterization, it generates less than 1% of the total rainfall through this scheme. In contrast, 94% of the rainfall is generated by the convection scheme in the 12 km model.

[21] As a test of our understanding, we can apply a Fourier analysis as described earlier to the data. In this case, we calculate the Fast Fourier Transform of Z(L,t) in the time direction and examine the phase ($\phi(L)$) of the coefficient corresponding to the diurnal signal. This phase is plotted against length scale in Figure 7 for the three data sets. The diagram confirms the principal interpretation of Figure 4 given above, with the 12 km model response flat across

all length scales and earlier in the day than the observations, and the 4 km model results diverging from the observations at short length scales. The apparent precision of this type of plot makes it appealing to use as the primary analysis tool rather than the underlying image of $Z(L,t)$. However, in situations with only slightly greater complexity, $\phi(L)$ can begin to vary significantly and loses its utility. In contrast, plots of $Z(L,t)$ retain the subtlety of the evolutionary behavior and remain useful.

[22] An equivalent diagram to Figure 4 is plotted in Figure 8 but now using F_{th} = 210 W m^{-2} ($T_{b,e}$ = 247 K). At this level, the clusters are additionally sampling significantly lower, warmer clouds. Both models show the same behavior as at the lower threshold flux level. In contrast, the GERB panel is now remarkably different, exhibiting a clear semidiurnal signal. The previous late afternoon evolution appears to have been retained, if less distinctly, at all length scales. However, there is now an additional peak number of systems occurring at the same early morning time as in the 4 km model for small- and medium-scale systems. This covers a larger range of length scales than the 4 km model but there is no slope apparent that might suggest a progression from larger to smaller systems mirroring the earlier growth for the cold systems. Instead, the large cold systems that have been generated during the day appear to break up as warmer systems with a spectrum of sizes in the early morning. The consistency of the behavior of the 4 km model at both flux thresholds suggests that there we are seeing growth of cold systems during the day toward large scales that appear to suddenly break up into small fragments of equally cold clouds.
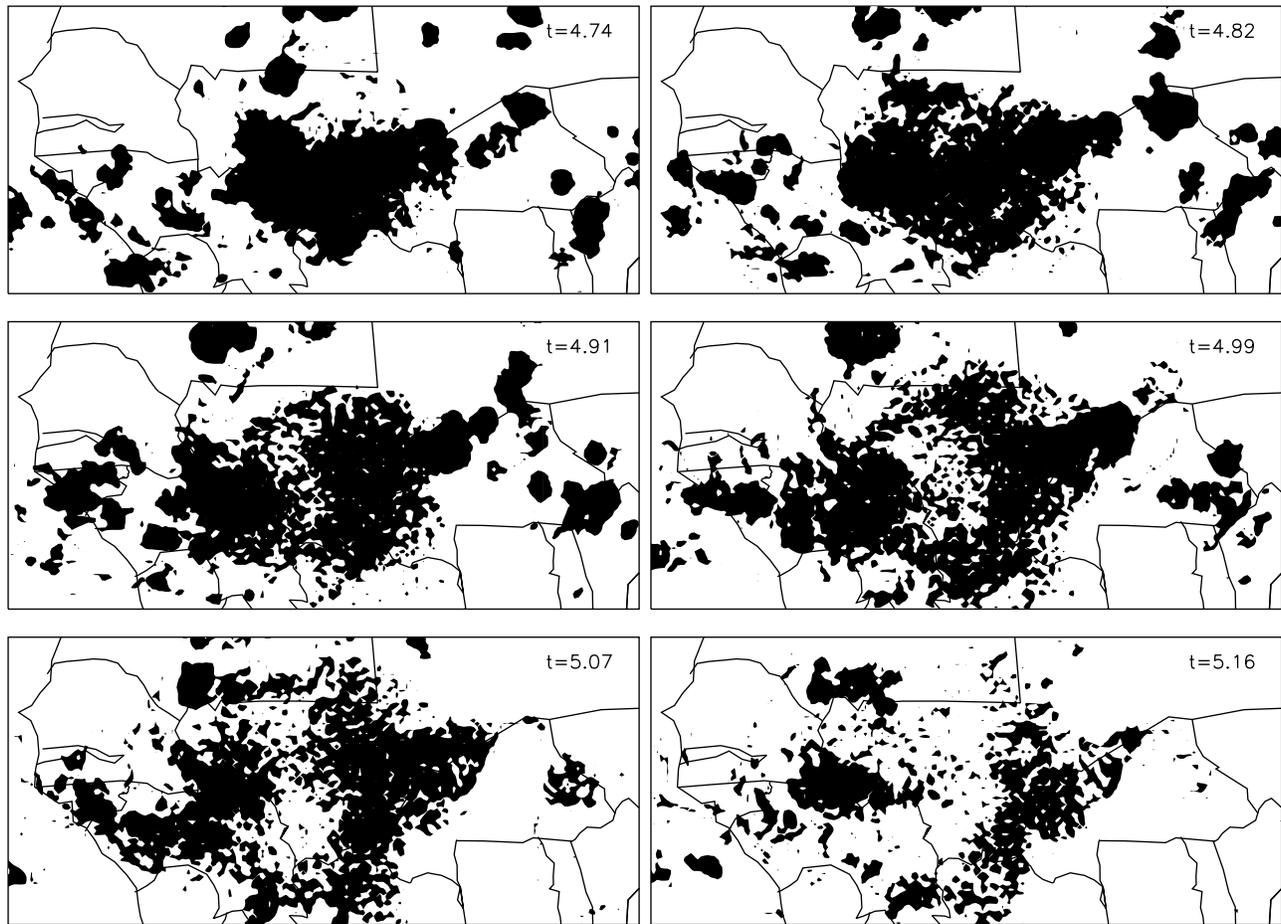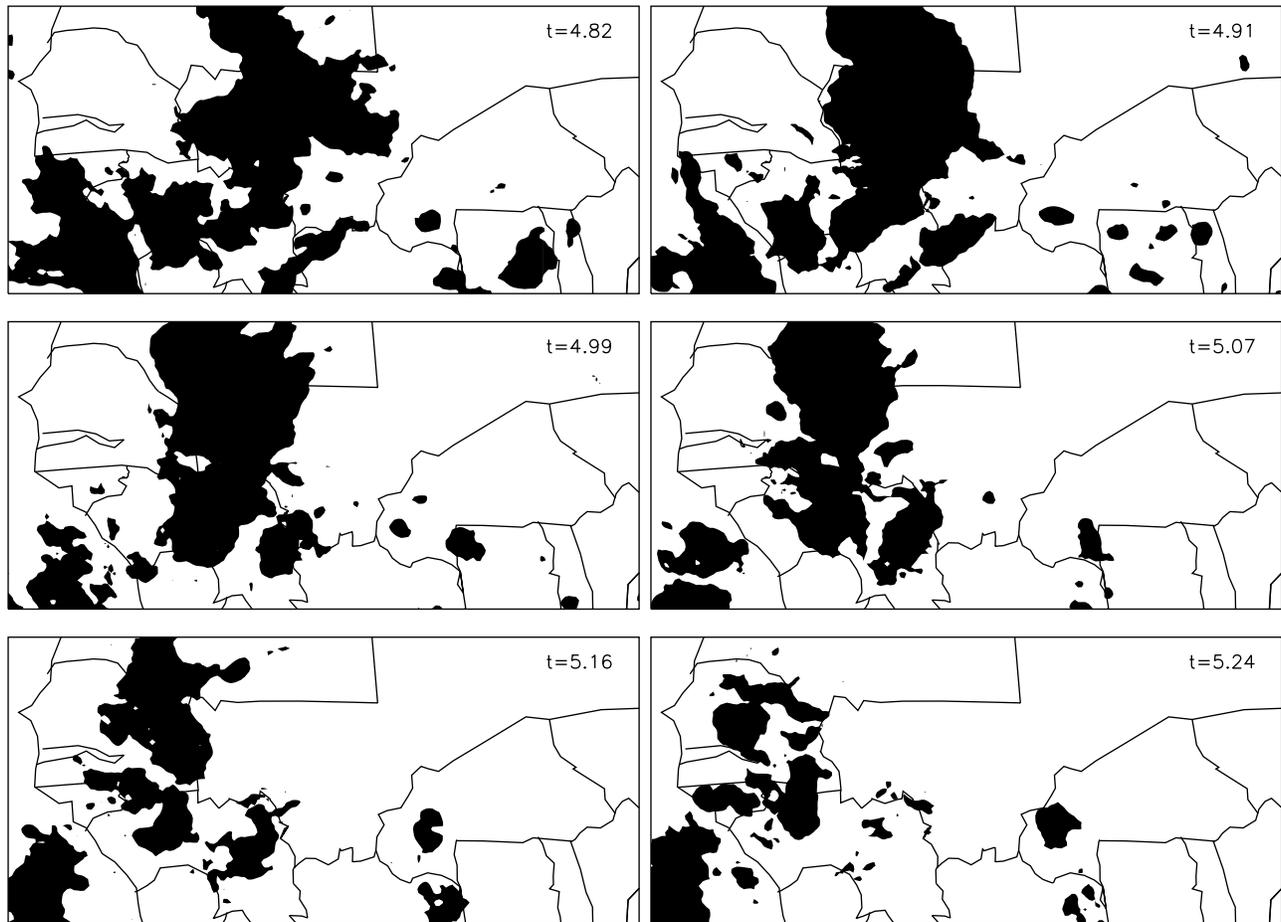
Figure 5. Snapshots of pixels in the 4 km model data falling below $F_{\text{th}} = 150$ W m^{-2} two hourly illustrating the overnight breakup of a large system into many smaller fragments. The data were rebinned to 12 km resolution.

[23] Several authors have reported semidiurnal cycles in total cloudiness and precipitation in the tropics [e.g., *Gruber*, 1976; *Augustine*, 1984; *Liu and Zipser*, 2008]. Almost invariably, these occur over the ocean. *Minnis and Harrison* [1984] do report two maxima in the diurnal behavior of cloudiness over parts of South America but they attribute this to the peculiar local interaction of the coastline and mountainous terrain. In direct contrast, versions of Figure 8 generated separately for each quarter of the GERB domain show a clear double peaked behavior arising from the NE quadrant (that is almost exclusively land) but no apparent signal from the SW quadrant (containing a substantial fraction of ocean). This result is consistent, however, with *Allan et al.* [2007] who examined the cloud masks derived from SEVIRI data for equatorial Africa (7°E–45°E,10°S–10°N) during July 2006. They showed that, while the convective cloud fraction exhibited a single-peaked diurnal signal, the total cloud fraction that also included warmer clouds had two peaks at 3 and 12 UTC.

[24] The underlying distributions of the two statistics $(\overline{N}, \sigma)$ used to generate Figures 4 and 8 themselves yield useful information about the relative behavior of modeled and observed systems. The mean size distributions plotted in Figure 9 appear to follow a power law of the form $\overline{N} = N_0 L^{-\Gamma}$ over a large range of length scales found by several other

studies [e.g., *Machado et al.*, 1992]. *Zhao and Di Girolamo* [2007] highlight the difficulties involved with comparing parameters derived from fitting to such a form from different data sets due to variations such as scene selection and cloud detection criteria, regional peculiarities, the overall meteorological state, and the details of the fitting procedures themselves. Nonetheless, having produced three data sets in a consistent way from the same physical quantity, we can be confident, at least, of being able use this as a fair test of our models against the GERB data.

[25] The derivation of values for the parameters N_0 and Γ requires a weighted fit to the data which in turn requires an estimate for the standard error $(\sigma_{\overline{N}})$ on each data point. The plotted error bars represent the measured standard deviation σ of each bin. However, the significant variability evinced by the size of these error bars is due largely to the systematic diurnal cycle signal superimposed on to the random variability. To make progress we must appeal to our assumption that the diurnal signal averages to zero and that the mean size distribution is thus subject solely to an underlying intrinsic Poisson noise with $\sigma = \sqrt{\overline{N}}$. We would then normally proceed by calculating the standard error on the mean in each bin from the population variance $\sigma_{\overline{N}}^2 = \sigma^2/n_{\text{f}} = \overline{N}/n_{\text{f}}$ where n_{f} is the total number of image frames in the data set. However, given that systems can persist for many hours, it is

Figure 6. Snapshots of pixels in the GERB data falling below $F_{th} = 150$ W m^{-2} two hourly illustrating the overnight weakening of large-scale systems. The data cover the same area and similar time period to Figure 5.

not valid to assume that each image is an independent trial. Using a mean lifetime for a convective storm system (t_c) and interval between images (Δt), we have instead effectively $n_f \Delta t/t_c$ independent trials. The standard error on the mean number of systems in each length scale bin we adopt is thus $\sigma_{\overline{N}} = \sqrt{\overline{N} t_c/n_f \Delta t}$. In order to place conservative error estimates on the fitted parameters, we use a long $t_c = 6$ h. This results in 40 "trials" over the course of the 10 days of model data. It is therefore reasonable to appeal to the central limit theorem and assume that the distribution of mean values has a Gaussian distribution. Values derived from a nonlinear least squares fit to the data sets are listed in Table 1. The fitted region was restricted to 30 km $< L <$ 200 km to avoid apparent breaks in the power law. The reduced χ^2 values are all agreeably <1, consistent with our conservative error assumptions.

[26] At the lower threshold, the fitted power law index (Γ) is much smaller for the observed data than for the model. This is apparent in Figure 9 with the slope of the GERB data clearly differing from the models. The two model distributions appear to have similar slopes. However, the derived error estimates for Γ from the models indicate that they are formally inconsistent with each other. Although the 4 km model does represent an improvement in Γ, it consistently overpredicts the amount of cloud at all length scales. The model values change little between the two thresholds. However, the observed value of Γ is much larger at the higher flux threshold, now falling between the two model values and in agreement with them both at around the 2σ level.

[27] The above remarks concerning the difficulties of intercomparison notwithstanding, *Zhao and Di Girolamo* [2007] present values for Γ ranging from 1.19 to 2.18 for cumulus cloud in the size range of interest [*Cahalan and Joseph*, 1989; *Sengupta et al.*, 1990; *Benner and Curry*, 1998]. These arise from high-resolution observational studies (~50 m) with similar numbers of pixels to ours but consequently much smaller domains. *Machado et al.* [1992] use Meteosat data with a similar resolution and covering a similar region to ours, finding $\Gamma = 1.3$ for their warm detection threshold ($T_{IR} = 253$ K) and $\Gamma = 0.8$ for $T_{IR} = 218$ K. In all these cases, we have accounted for the authors' use of the number *density* and our logarithmic binning by subtracting 1 from the quoted values.

[28] As we move to longer length scales, the diurnal signal in Figure 4 becomes less distinct. Figure 10 shows the measured $\sigma^2/\overline{N}$ plotted against L for the three data sets using $F_{th} = 150$ W m^{-2}. Since the main source of variability is the
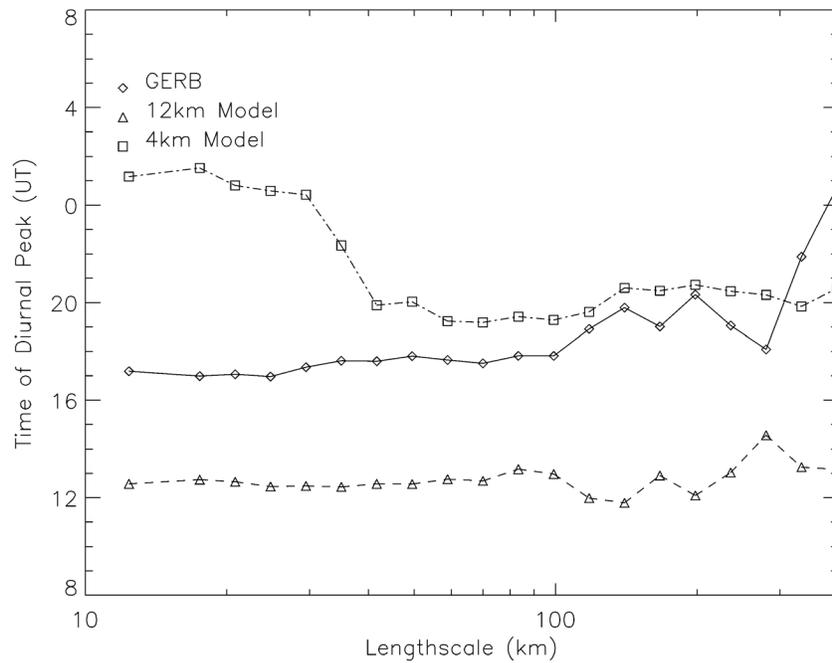
Figure 7. The phase of the maximum of the diurnal component derived from the Fourier transform of the cloud cluster size distribution time series. Each data set is plotted with a different style: GERB (diamonds, solid line), 12 km model (triangle, dashed line), and 4 km model (square, dot-dashed line).

diurnal signal, this plot essentially measures the strength of that signal against length scale. All three data sets asymptote to $\sigma^2 \approx N$ that one would expect for pure Poisson noise, supporting our earlier assumption regarding the underlying

distribution. However, both models inject a greater signal than apparent in the observations at length scales $\lesssim 100$ km, although the 4 km model does represent an improvement. The 12 km model generates a large number of single pixel
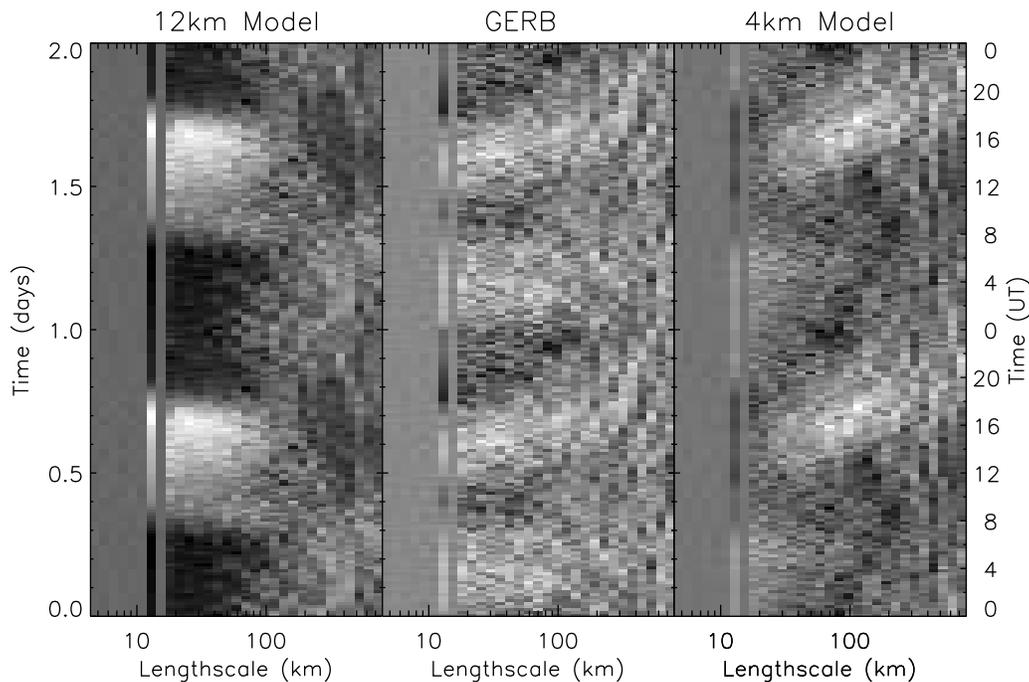
Figure 8. Normalized time series of cloud cluster size distributions $Z(L,t)$ observed by GERB and produced by the two model runs with a threshold flux for cloud detection of 210 W m^{-2}. The data have been folded on a period of 1 day and repeated. The gray scale for each frame has been set independently to ensure each utilizes the entire contrast range.
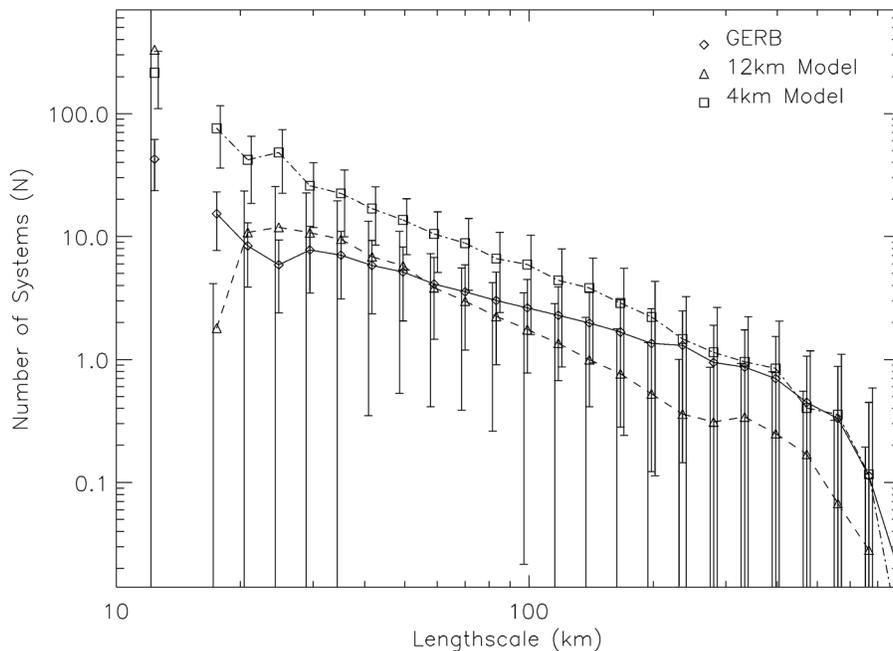
Figure 9. Mean number of clouds identified in each 15 min image for GERB (diamond, solid line), 12 km model (triangle, dashed line), and 4 km model (square, dot-dashed line) data using $F_{th} = 150$ W m^{-2}. The length scale bins are arranged logarithmically with edge values given by $L_i = 2^{2+i/4}$ km. The error bars indicate the standard deviation measured for each data set and are displaced slightly for the model data sets for clarity.

storms with a strong diurnal variability that runs off of the scale in Figure 10.

4. Summary

[29] A technique for assessing the diurnal development of convective storm systems has been presented based on comparisons of OLR simulated by high-resolution versions of the Met Office UM and observed by GERB. Applying an OLR thresholding technique to define the distribution of convective systems, a direct representation of the aggregated evolutionary behavior of convective storms can be generated by considering anomalies in the size distribution through the standard score statistic. The values in the length scale-time domain reveal the time of peak activity for storms of different sizes from which the growth and dissipation behavior can be inferred. Greater precision and insight can be gained by using the underlying statistics of the distributions to define how data sets differ in the way they exhibit convection. The 12 km model test case produced convection significantly earlier in the day than observed and

showed no evidence for size growth. In contrast, the 4 km model produced realistic evolution at medium to large scales but the small-scale behavior was dominated by the large number of systems suddenly produced when the storms broke up overnight rather than the gentler dissipation in the observed data. Both models generated a greater variability in the number of storms of all sizes.

[30] Now that we can diagnose the realism or otherwise of a model's evolutionary behavior, we can begin to probe the mechanisms that cause this. Is the crucial process mesoscale circulation such as downdrafts and subsequent cold pools propagating as density currents and organizing convection? How do the clouds "communicate" with each other? Is it through direct tendencies or the large-scale environment? What is the balance between the growth of individual systems and mergers? We also intend to investigate the relationship between the African Easterly Wave present in the domain and the growth behavior, the representation of initiation in different regions in the domain and the structure of the mature systems.

Table 1. The Derived Parameters for the Best Fit Power Laws of the Form $\overline{N} = N_0 L^{-\Gamma}$ to the Mean Size Distribution of Cloud Clusters for Two Values of F_{th}[a]

Data Set	150 W m^{-2}			210 W m^{-2}		
	N_0	Γ	χ^2/ν	N_0	Γ	χ^2/ν
GERB	189 ± 30	0.931 ± 0.038	0.17	2590.9 ± 1.1	1.40482 ± 0.00011	0.44
4 km model (rebinned)	2240 ± 310	1.304 ± 0.034	0.50	2880 ± 390	1.352 ± 0.033	0.36
4 km model (full resolution)	2960 ± 400	1.360 ± 0.033	0.35	3450 ± 470	1.392 ± 0.033	0.34
12 km model	3140 ± 690	1.633 ± 0.055	0.27	2290 ± 490	1.517 ± 0.053	0.29

[a]The fitted region was restricted to 30 km < L < 200 km to avoid apparent breaks in the power law. The value of N_0 for the 4 km model was corrected to account for the smaller area it covered.
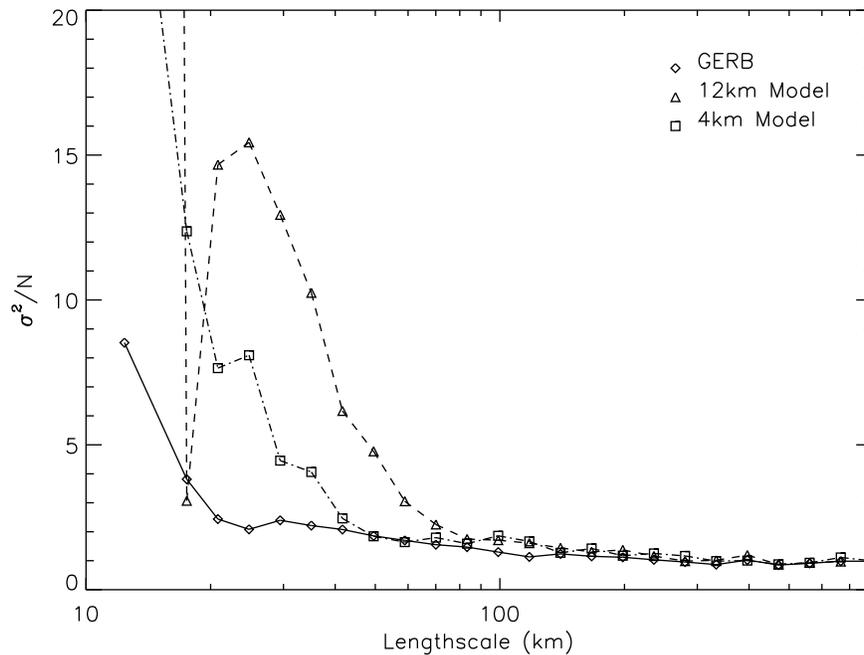
Figure 10. The measured variance of the number of storms of a particular length scale in the different data sets using $F_{th} = 150$ W m^{-2}. Each data set is plotted with a different style: GERB (diamonds, solid line), 12 km model (triangle, dashed line), and 4 km model (square, dot-dashed line).

References

Allan, R. P., A. Slingo, S. F. Milton, and M. E. Brooks (2007), Evaluation of the Met Office global forecast model using Geostationary Earth Radiation Budget (GERB) data, *Q. J. R. Meteorol. Soc.*, *133*, 1993–2010.

Augustine, J. A. (1984), The diurnal variation of large-scale inferred rainfall over the tropical Pacific Ocean during August 1979, *Mon. Weather Rev.*, *112*, 1745–1751.

Benner, T. C., and J. A. Curry (1998), Characteristics of small tropical cumulus clouds and their impact on the environment, *J. Geophys. Res.*, *103*(D22), 28,752–28,767, doi:10.1029/98JD02579.

Cahalan, R. F., and J. H. Joseph (1989), Fractal statistics of cloud fields, *Mon. Weather Rev.*, *117*, 261–272.

Comer, R. E., A. Slingo, and R. P. Allan (2007), Observations of the diurnal cycle of outgoing longwave radiation from the Geostationary Earth Radiation Budget instrument, *Geophys. Res. Lett.*, *34*, L02823, doi:10.1029/2006GL028229.

Cullen, M. J. P., T. Davies, M. H. Mawson, J. A. James, S. C. Coulter, and A. Malcolm (1997), An overview of numerical methods for the next generation UK NWP and climate model, in *Numerical Methods in Atmospheric and Ocean Modelling: The André J. Robert Memorial Volume*, edited by C. A. Lin et al., pp. 425–444, Can. Meteorol. and Oceanogr. Soc., Ottawa, Ont.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood (2005), A new dynamical core for the Met Office's global and regional modelling of the atmosphere, *Q. J. R. Meteorol. Soc.*, *131*, 1759–1782.

Dewitte, S., L. Gonzalez, N. Clerbaux, A. Ipe, C. Bertrand, and B. De Paepe (2008), The Geostationary Earth Radiation Budget Edition 1 data processing algorithms, *Adv. Space Res.*, *41*, 1906–1913.

Essery, R., M. Best, and P. Cox (2001), MOSES 2.2 technical documentation, *Hadley Cent. Tech. Note 30*, Met Office, Exeter, UK.

Fu, R., A. D. Del Genio, and W. B. Rossow (1990), Behavior of deep convective clouds in the tropical Pacific deduced from ISCCP radiances, *J. Clim.*, *3*, 1129–1152.

Futyan, J., and A. D. Del Genio (2007), Deep convective system evolution over Africa and the tropical Atlantic, *J. Clim.*, *20*, 5041–5060.

Grabowski, W. W., et al. (2006), Daytime convective development over land: A model intercomparison based on LBA observations, *Q. J. R. Meteorol. Soc.*, *132*, 317–344.

Gregory, D., and P. R. Rowntree (1990), A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure, *Mon. Weather Rev.*, *118*, 1483–1506.

Gruber, A. (1976), An estimate of the daily variation of cloudiness over the GATE A/B area, *Mon. Weather Rev.*, *104*, 1036–1039.

Guichard, F., et al. (2004), Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models, *Q. J. R. Meteorol. Soc.*, *130*, 3139–3172.

Harries, J. E., et al. (2005), The Geostationary Earth Radiation Budget project, *Bull. Am. Meteorol. Soc.*, *86*, 945–960.

Haywood, J. M., R. P. Allan, I. Culverwell, A. Slingo, S. Milton, J. Edwards, and M. Clerbaux (2005), Can desert dust explain the outgoing longwave radiation anomaly over the Sahara during July 2003?, *J. Geophys. Res.*, *110*, D05105, doi:10.1029/2004JD005232.

Hodges, K. I., D. W. Chappel, G. J. Robinson, and G. Yang (2000), An improved algorithm for generating global window brightness temperatures from multiple satellite infrared imagery, *J. Atmos. Oceanic Technol.*, *17*, 1296–1312.

Hohenegger, C., P. Brockhaus, C. S. Bretherton, and C. Schär (2009), The soil moisture-precipitation feedback in simulations with explicit and parameterized convection, *J. Clim.*, *22*, 5003–5020.

Lau, K. M, and P. H. Chan (1985), Aspects of the 40–50 day oscillation during the northern winter as inferred from outgoing longwave radiation, *Mon. Weather Rev.*, *113*, 1889–1909.

Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell (2008), Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom, *Mon. Weather Rev.*, *136*, 3408–3434.

Lean, H. W., N. M. Roberts, P. A. Clark, and C. Morcrette (2009), The surprising role of orography in the initiation of an isolated thunderstorm in southern England, *Mon. Weather Rev.*, *137*, 3026–3046.

Leary, C. A., and R. A. Houze Jr. (1979), The structure and evolution of convection in a tropical cloud cluster, *J. Atmos. Sci.*, *36*, 437–457.

Liu, C., and E. J. Zipser (2008), Diurnal cycles of precipitation, clouds, and lightning in the tropics from 9 years of TRMM observations, *Geophys. Res. Lett.*, *35*, L04819, doi:10.1029/2007GL032437.

Lock, A. P., and A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith (2000), A new boundary layer mixing scheme: Part I. Scheme description and single-column model tests, *Mon. Weather Rev.*, *128*, 3187–3199.

Machado, L. A. T., M. Desbois, and J.-Ph. Duvel (1992), Structural characteristics of deep convective systems over tropical Africa and the Atlantic Ocean, *Mon. Weather Rev.*, *120*, 392–406.

Machado, L. A. T., J.-P. Duvel, and M. Desbois (1993), Diurnal variations and modulation by easterly waves of the size distribution of convective cloud clusters over west Africa and the Atlantic Ocean, *Mon. Weather Rev.*, *121*, 37–49.

Mapes, B. E., and R. A. Houze Jr. (1993), Cloud clusters and superclusters over the oceanic warm pool, *Mon. Weather Rev.*, *121*, 1398–1415.

Minnis, P., and E. F. Harrison (1984), Diurnal variability of regional clouds and clear-sky radiative parameters derived from GOES data: Part II. November 1978 cloud distribution, *J. Clim. Appl. Meteorol.*, *23*, 1012–1031.

Murakami, T. (1980), Empirical orthogonal function analysis of satellite-observed outgoing longwave radiation during summer, *Mon. Weather Rev.*, *108*, 205–222.

Randall, D. A., et al. (2007), Climate models and their evaluation, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 589–662, Cambridge Univ. Press, Cambridge, UK.

Rickenbach, T. M. (1999), Cloud-top evolution of tropical oceanic squall lines from radar reflectivity and infrared satellite data, *Mon. Weather Rev.*, *127*, 2951–2976.

Rickenbach, T. M., P. Kuchera, M. Gentry, L. Carey, A. Lare, R.-F. Lin, B. Demoz, and D. O'C. Starr (2008), The relationship between anvil clouds and convective cells: A case study in south Florida during CRYSTAL-FACE, *Mon. Weather Rev.*, *136*, 3917–3932.

Sengupta, S. K., R. M. Welch, M. S. Navar, T. A. Berendes, and D.W. Chen (1990), Cumulus cloud field morphology and spatial patterns derived from high spatial resolution Landsat imagery, *J. Appl. Meteorol.*, *29*, 1245–1267.

Slingo, A., et al. (2006), Observations of the impact of a major Saharan dust storm on the atmospheric radiation balance, *Geophys. Res. Lett.*, *33*, L24817, doi:10.1029/2006GL027869.

Williams, M., and R. A. Houze Jr. (1987), Satellite-observed characteristics of winter monsoon cloud clusters, *Mon. Weather Rev.*, *115*, 505–519.

Wilson, D. R., and S. P. Ballard (1999), A microphysically based precipitation scheme for the UK Meteorological Office Unified Model, *Q. J. R. Meteorol. Soc.*, *125*, 1607–1636.

Yang, G., and J. Slingo (2001), The diurnal cycle in the tropics, *Mon. Weather Rev.*, *129*, 784–801.

Zhao, G., and L. Di Girolamo (2007), Statistics on the macrophysical properties of trade wind cumuli over the tropical western Atlantic, *J. Geophys. Res.*, *112*, D10204, doi:10.1029/2006JD007371.

R. P. Allan and R. J. Hogan, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK.

C. E. Holloway, Climate Directorate, National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK.

G. M. S. Lister, Computational Modeling Services, National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK.

K. J. Pearson, Environmental Systems Science Centre, University of Reading, 3 Earley Gate, Reading RG6 6AL, UK. (kjp@mail.nerc-essc.ac.uk)