Review of "Evaluating the diurnal cycle in cloud top temperature from SEVIRI" by Sarah Taylor et al.

Using the CLAAS-2 dataset, the authors characterize the diurnal cycle of cloud top temperature for several regions within the SEVERI's observation disk. Retrievals from SEVERI are compared to CTT inferred from collocated CALIPSO observations in terms of bias and variability. As cloud top temperature is a frequently used quantity in radiative balance and cloud microphysics studies, I feel this is a useful contribution to the literature. I recommend that this study be published after the following minor comments are addressed.

> **NB:** Page and line numbers refer to location in the difference file, rather than the revised manuscript.

Specific Comments:

P3, L17: "... images such as SEVERI observe the radiometric height of the cloud." Perhaps it would be helpful to briefly expand upon what is meant by "radiometric height" in the context of a weighting function.

> P3, L35: A brief discussion of weighting functions and their relation to radiometric height has been added to the text.

P9, L12: CALIOP doesn't observe the CTT, rather it is inferred from CTH measured by the instrument. Please phrase this differently.

> P11, L27: We have rephrased the text to clarify the method by which CTT are obtained from CALIOP.

P12, L32-34: This statement does not make sense to me. Why would a region with a lower surface albedo heat up more slowly? Could this be a consequence of a decreased sensible heat flux in Central Africa due to evapotranspiration?

> P15, L23: We thank the reviewer for their comment on this point. We revised our arguments and agree with the reviewer that the facts do not support this claim. For this reason, we have dropped this line of argument and leave the investigation of this feature to future studies. We have retained a discussion of the differences in the diurnal cycle between these two regions, and highlighted the lack of a robust explanation for the difference.

P14, L12-13: While not incorrect, "temporal distance" seems like an unusual way of expressing this, especially since it is primarily a term used in psychology. Perhaps phrasing it as something like "within ±30 minutes of CALIOP observation" may be clearer.

> P17, L9: We have amended the text to "within ±30 minutes of a CALIOP overpass", in order to clarify our meaning here.

Technical Corrections:

P2, L4: "cloud" should be plural

We have amended the text as suggested.

P9, L20: "night time" should be "nighttime".

We have amended the text as suggested.

Review of "Evaluating the diurnal cycle in cloud top temperature from SEVIRI" by Sarah Taylor et al.

This study attempts to evaluate retrieval biases in the CLAAS-2 Cloud Top Temperature (CTT) dataset derived from SEVIRI via a comparison with co-located CALIOP retrievals. The novelty of the study involves the separation of the evaluation into daytime and nighttime components, in order to establish whether the quality of the retrievals is consistent through the day. This is important given one of the key benefits of SEVIRI based retrievals should be their ability to capture the diurnal cycle in the geophysical quantity of interest. Having qualified the retrieval quality the authors then investigate the diurnal variability seen within the CLAAS dataset and discuss the implications of their evaluation for the robustness of the CTT signals contained within it. Overall I think that the concept of the study is a good one – as noted about, the key benefit of CLAAS above what is possible from instruments in polar orbit should be an accurate representation of the diurnal cycle in cloud parameters. Hence making potential users aware of deficiencies in this representation is useful. The paper is generally well written and the methodology clearly presented. However, I do have some major and minor comments which I list below. Subject to these being satisfactorily addressed, in my opinion, the paper will be suitable for publication.

> **NB:** Page and line numbers refer to location in the difference file, rather than the revised manuscript.

Major Comments:

1. Nowhere is the accuracy of the 'truth' dataset, CTT from CALIOP, actually quantitatively defined. There is also no mention of whether there is any difference in CALIOP retrieval quality from day to night which I suspect there might be. I also wonder whether the CALIOP inferred heights and optical depths are equally accurate for all cloud types since I would expect the sensitivity to be higher for clouds comprised of smaller ice particles and droplets than for those comprised of larger droplets. Quantitative information concerning accuracy must be in the paper rather than phrases like 'very accurate'.

> P8, L22: We have added information on the accuracy of the CALIOP retrieval of CTH and the process by which this is converted to CTT, including a discussion of the potential inaccuracies introduced by the conversion process. We have also contacted NASA's CALIPSO group and included as much information as is available on the accuracy of the CALIOP cloud layer and CTT retrievals. Unfortunately, no studies have been published on the accuracy of the CALIOP CTT measurements themselves.

2. In a similar vein, the transition of height or pressure to CTT using model fields is mentioned as a possible reason for discrepancies between the CALIOP and SEVIRI retrievals, particularly in the SE Atlantic stratus region. However, detail on how the transition is done is rather lacking. It would be useful to know how, for example, timing discrepancies between the background meteorological fields and the satellite overpasses are dealt with. Ideally, in order to isolate the impact of differences in the background model fields one would want to be able to do the analysis using the same set but I appreciate this is beyond the scope of this study. Can any literature be used to give some idea of how well the different meteorological models represent the background state under different regimes as this will affect the CTT comparison?

> P6, L1: We have extended the description of the SEVIRI algorithms used in Sect. 2, including the use of model reanalysis data. In particular, we have added explanations of how spatial and temporal discrepancies are dealt with and references to additional literature relating to the atmospheric models used. We have also added further information about GEOS-5 model parameters used in the CALIOP cloud top properties retrieval (P9, L2). After discussing this issue with the creators of the SEVIRI and CALIOP datasets, we have not been able to discover any studies on how well these two meteorological models represent the background state under different regimes.

3. The introduction focuses almost exclusively on convective cloud yet much of the paper discusses regions of stratiform and potentially mixed cloud. Hence the introduction needs some broadening to reflect this.

> Sect. 1: We have widened the focus of the introduction to include details of the diurnal cycle of all cloud types.

4. It was unclear to me whether pixels were only considered if they were fully cloud covered or if broken cloud scenes were also considered. I thought the former, but there seemed to be a significant amount of discussion on the effects of surface emission. For the majority of deep convective cases, for fully covered pixels this wouldn't actually have an effect. Since the authors limit the cases studied to greater than an optical depth (where spectrally?) of > 1.0 even outside of those regions the surface impact would be limited.

> SEVIRI pixels were considered only if they were fully cloud covered. We agree with the reviewer that once clouds with an optical depth >1.0 are excluded, surface emissions are unlikely to have a large effect. We have removed unnecessary references to surface emissions.
> The impact of surface emissions is however relevant to the discussion of how to choose a COD threshold when comparing SEVIRI and CALIOP data, as well as to the discussion of previous studies, which have not used such stringent COD thresholds. These references to surface emissions have therefore been retained.

5. Given much of the manuscript is spent showing deficiencies in the CLAAS retrieval algorithm outputs I feel a short summary of the algorithm is necessary. If this seems too demanding I think at the very least there should be some description of what changes from day to night. I assume that visible channels are used in the daytime (in addition to IR) which may go some way towards explaining why the biases, relative to CALIOP, are improved during the day.

> P6, L1: We have extended the description of the algorithms used for cloud masking and cloud top pressure retrievals for the CLAAS-2 dataset in section 2.

Minor Comments:

Page 2 line 32: Consider also adding papers by Pearson et al., 2010 (JGR), 2014 (QJRMS)

> P3, L10: We have included the suggested papers in our discussion.

Page 3 line 14: It's a little bit of a stretch to say that SEVIRI 'covers' the Middle East as implied here. It certainly sees it but perhaps not all of it. Similar comments apply to the Atlantic Ocean.

> P3, L31: We have amended the text to clarify that only parts of these regions are covered.

Page 3 line 17: Perhaps a little pedantic but SEVIRI does not observe the radiometric height of the cloud, it observes radiances (or, if we want to be completely technically accurate, digital counts). These can then be inverted to estimate the radiometric height. On a similar theme, the manuscript is written as if CALIOP observes cloud top height, which is not technically true, rather it observes backscatter, and, as with any satellite instrument, the required geophysical variable is inferred in some way.

> P3, L34: We have clarified that SEVIRI does not directly observe radiometric height and that CALIOP does not directly observe cloud top height.

Page 3, line 34: quantify what you mean by optically thin, and give the wavelength that you are defining the opacity at.

> P4, L20: Information on the sensitivity of CALIOP's cloud optical depth measurements, as well the wavelength at which optical depth is defined, has been added. Cloud optical depths from SEVIRI are not used in this study. Instead, SEVIRI data are excluded based on cloud optical depths from collocated CALIOP measurements. A discussion of the cloud optical depth thresholds used in previous studies and a detailed discussion of how the threshold was chosen for this studies are given in Sect. 3.2.1.

Page 4, line 14: this makes it sound as if you are using data from one instrument flying on one satellite through the whole period considered here which is not true. Neither is it true that throughout this record the operational SEVIRI instrument was at 0 degrees longitude.

P5, L3: We have clarified that MSG is a series of satellites, as well as the fact that older versions of the satellites, which continue to fly and are located further to the east, are occasionally used to fill gaps in data coverage.

Page 4, line 31: more detail on how this CTP retrieval works is required. For example, the wording ‘the data available’, is extremely vague.

P6, L1: We have extended the description of the algorithms used for cloud masking and cloud top pressure retrievals in section 2. We have also removed the phrase “and the data available”. In some cases, the NWCSAF/MSGv2012 retrieval process changes depending on whether certain variables are missing from the satellite, or model datasets. However, upon further investigation, the CLAAS-2 dataset used in this study only makes retrievals where all such datasets are complete.

Page 5, line 10: as noted in the major comments, more detail on how this is done is required.

P6, L1: We have extended the description of the algorithms used for cloud masking and cloud top pressure retrievals in section 2.

Figure 1: I find this figure, although useful, somewhat misleading as it implies that clouds are ubiquitous. I would suggest another figure or set of panels indicating frequency of cloud occurrence would be beneficial. Similar information is shown later on in the paper but I think it would be good to have it upfront.

P7, L11: We have moved Fig. 7, which shows the frequency of cloud occurrence for the full CLAAS-2 dataset, so that it follows directly from Fig. 1.

Page 6, line 12: while I agree that the diurnal cycle could well be largest here I don’t think it necessarily follows from the reasoning given. It could be that the cold clouds are there throughout the day. Figure 1 tells you nothing about this on its own.

P7, L30: We agree with the reviewer that this figure says nothing about the diurnal cycle by itself. We have clarified that this figure is only indicative of areas which potentially have a strong diurnal cycle of convection and that a full analysis of the diurnal cycle (presented later in the paper) will be necessary.

Page 6, line 18: Figure 1 implies a spatial and seasonal pattern in cloud type but it doesn’t explicitly show it. Since you state that the cloud retrieval you are using gives cloud type it would be interesting to see if that maps to what you see in Figure 1.

P8, L3: We agree with the reviewer that this would be very interesting. However, since the cloud type information has not been processed by EUMETSAT in the same summarizing way as other cloud properties, this analysis is not currently available and would require substantial additional work, using data not publically available from EUMETSAT at this time. We have added a discussion of the assumed strong seasonal variability in cloud type over Africa due to the cyclic movement of the ITCZ.

Page 6, line 20: I don’t think section 2 really gives a quantitative idea of the implications of cloud type and land surface emissivity for the accuracy of CLAAS CTT. There was more of a general discussion in the introduction to be honest.

P8, L7: We have amended the reference to point to the more in depth discussion in the introduction.

Page 6, line 25: When the authors say ‘data was processed’ are they referring simply to the collocation process? It might be good to state this explicitly.

P8, L11: The reviewer is correct that were referring to the collocation process. This has been clarified.

Page 6, line 30: ‘very accurate’ is not very scientific.

> P8, L16: We have added further details and references regarding the accuracy of CALIOP’s measurement of CTH.

Page 7, line 17: define COD as Cloud Optical Depth (and give wavelength).

> P4, L19: COD has been previously defined on P6, L26. Information on the wavelength at which COD is measured has been added to P4, L19.

Page 7, line 25: I assume the temporal collation window is centred on the CALIOP observation. Is the variation in the timing of SEVIRI scan lines accounted for? What is the spatial match up criterion?

> P9, L16: We have added a paragraph clarifying our collocation methodology to the start of the discussion of the collocation process.

Page 7, line 28: Actually figure 2 suggests that the tightening from 0.3 to 1 does result in a reduced bias while above 1 there is not much change. The authors say this themselves in the next paragraph.

> P10, L6: We mean that when viewing a high COD cloud, CALIOP measures the physical cloud top, while SEVIRI observes the radiometric cloud top, which is below the height of the physical cloud top. The application of COD thresholds can remove cases where CALIOP might report the height of a cloud with an optical depth of, say 0.1, while SEVIRI observes the height of thicker cloud below. These thresholds cannot however account for the difference between the physical properties observed by the two instruments. A discussion of the expected size of this difference is provided in Sect. 1.
>
> We thank the reviewer for pointing out that the distinction between the two issues needs to be clarified and have amended our text accordingly.

Page 7, line 30: How is a cloud layer defined? i.e. are the thresholds simply applied to the topmost layer with a COD as diagnosed by the CALIOP product? Is there any change in vertical resolution in this product? I’m not sure whether it would have an impact but if for example the vertical resolution reduces with height above a certain point, the same COD would actually indicate a more diffuse extinction profile within the layer.

> P10, L8: The reviewer is correct in their assumption that thresholds are applied to the topmost cloud layer with a COD as diagnosed by the CALIOP product. We have clarified this point in our text.
>
> With regard to the vertical resolution of CALIOP, this does change with height. Between the surface and ~8.2 km, the vertical is 30 m. From ~8.2 km to ~20.2 km it is 60 m. The same COD could therefore indicate different extinction profiles above and below ~8.2 km. While several previous studies (Reuter et al., 2009; Stubenrauch et al., 2010; SAFNWC/MSG, 2012; Kniffka et al., 2013; Benas et al., 2016) have used CALIOP layer optical depth as a threshold to exclude low optical depth clouds when comparing cloud top properties from SEVIRI and CALIOP, we do agree that this may introduce a bias. We have therefore added an explanation of the potential biases introduced by the change in vertical resolution of CALIOP (P9, L14).

Page 8, line 13: what is the ‘relatively coarse’ resolution in numbers? How does this compare to the CALIOP vertical resolution? How well do both of the meteorological models actually capture low level temperature inversions? Do they persist throughout the night and day?

> P6, L1: We have extended the description of the algorithms used for cloud masking and cloud top pressure retrievals from SEVIRI in section 2 and included references relating to the model used.
>
> P9, L1: We have also extended the description of the algorithms used by the CALIOP retrieval and included references relating to the model used.

We have contacted the creators of the SEVIRI and CALIOP datasets, but have been unable to obtain further information on the accuracy of these models with regards to low level temperature inversions.

Page 9, line 7: do the deviations show a Gaussian distribution?

P11, L19: The biases between SEVIRI and CALIOP CTTs (SEVIRI minus CALIOP) show a right-skewed Gaussian distribution (below). In general, positive values in the distribution show biases due to differences between the radiometric and physical CTTs retrieved from the two instruments, while negative values indicate a potential retrieval error. In particular, it seems that the distribution may be skewed slightly towards the negative due to the large number of slightly negative values found in the southeast Atlantic Ocean. A brief discussion of the shape of the distribution has been added to the text.

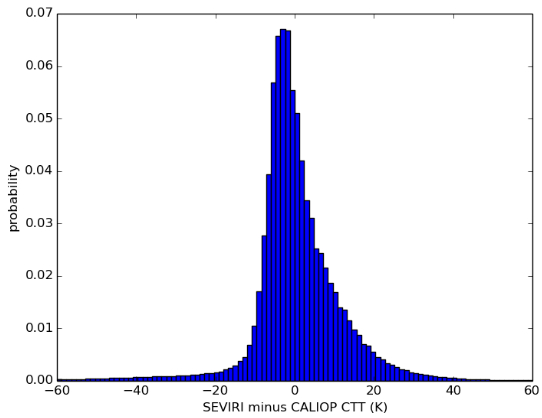

Figure 5: I would weight these differences by occurrence or at the very least discuss them in the context of Figure 4. Otherwise the eye is drawn to the very strong positive daytime bias over the Sahara when really there are few clouds there. Why are there negative differences over the east of Africa during the daytime (Sudan/Ethiopia)?

P12, L5: we agree with the reviewer that it is important to consider Figure 5 in the context of the occurrence of cloud in different regions. We feel that weighting the biases by occurrence would unhelpfully obscure the magnitude of the biases in some regions. We have however amended our discussion to highlight the importance of considering Figure 5 (now Fig. 6) in the context of Figure 4 (now Fig.5).

Re: East Africa. We also found the negative differences over Sudan/Ethiopia during the daytime puzzling. Unfortunately, we have been unable to find an explanation for this feature and further investigation of this feature is left to future studies.

Page 9, line 28: it may be the colour scale but I would say that while the majority of the ocean shows small biases, there are regions where differences look relatively large. One of these regions is discussed in the next paragraph so perhaps merge these two parts together.

P12, L12: Although there are some regions over the ocean with relatively high bias, these still fall within the range of biases expected due to the difference between the radiometric and physical cloud top heights observed by SEVIRI and CALIOP respectively. We have added a brief discussion of this point to the text.

Page 10, line 20: I understand that the authors have used three month means in order to obtain sufficient data to see a coherent cycle at the pixel scale but do they perhaps worry that by doing this you are losing information about how the phase of any diurnal cycle in CTT might vary through the year? The discussion about producing ‘smooth cycles’ is rather vague.

P13, L7: We agree with the reviewer that there is a risk that averaging data over multiple months may obscure information about variation in the diurnal cycle throughout the year. We have attempted to balance the need to have enough data for a statistical analysis of the diurnal cycle in regions with sparse data (e.g. over land) with our interest in analyzing annual variation in the diurnal cycle. We believe that by breaking the dataset down into three-monthly means we are able to show the most important features of the variability throughout the year, while retaining sufficient data for a statistical analysis over the majority of the land.

Our discussion of 'smooth' cycles references our attempt to ensure that in every region there were enough CTT retrievals at each hour in the day to produce mean diurnal cycles in which a single data point (e.g. the data point at 19:00 in Fig. 6) represents a mean over at least 1,000 cloud top temperature retrievals for the majority of the area included in this study. Areas for which this was not possible are shaded in Figures 8 and 9.

We have clarified the process by which the diurnal cycle was calculated and replaced the reference to 'smooth cycles'.

Figure 6: Can the authors provide some idea of the range of values that comprise each hourly mean please, perhaps using quartiles or SDs if the distribution is Gaussian?

Figure 7: The range of values that comprise each hourly mean are now shown in Figure 7, which now includes box plots showing quartiles for each hour. Additional text discussing these values has also been added.

Page 10, line 33 (and in other sections focused on the Sahara): I am a bit bemused about the emphasis on the Sahara in the latter part of the paper. As is shown, there is very little cloud being detected there (as one might anticipate) and I am not sure that I would expect too much of what is there to be deep convective in nature (at least north of the inter-tropical front). Hence why should we expect a marked diurnal cycle? Moreover, when what is happening there is analysed the statistics will be poor.

We agree that few clouds and no marked diurnal cycle are expected in the Sahara. The reason for our focus on this region is the large negative biases seen in the daytime, as well as the related large difference between day and nighttime biases is this area. Although this is of course based on sparse data, the spatial uniformity of the bias in this region is itself interesting. Our purpose in discussing the region is to show that these biases lead to a large, unphysical, diurnal cycle being observed in this region, as well as to show that the large negative daytime biases in CTT in this region may indicate a potential issue with the SEVIRI retrieval algorithm.

Page 12, line 33: I agree that vegetated Central Africa will typically have a lower surface albedo than the Sahara but I don't see why this would cause the Sahara to heat up more quickly after sunrise (I would actually expect the opposite based on albedo alone) or how this would produce lower, warmer clouds. Please explain.

P15, L23: We thank the reviewer for their comment on this point. We revised our arguments and agree with the reviewer that the facts do not support this claim. For this reason, we have removed this discussion and leave the investigation of this feature to future studies. We have retained a discussion of the differences in the diurnal cycle between these two regions, and highlighted the lack of a robust explanation for the difference.

Page 13, line 22: 'Additional biases' – this sentence is very vague. What retrieval errors are being referred to?

We do not claim to have isolated the cause of any retrieval errors in the CLAAS-2 dataset. We simply mean to state that while the majority of the biases described in this paper can potentially be explained by either the expected difference in CTT due to the fact that SEVIRI observes the radiometric and CALIOP measures the physical cloud top, or due to differences in the approach to dealing with atmospheric inversions between the two datasets, there of course may be a few areas where neither of these explanations apply. It may be that in these cases, retrieval errors in the

CLAAS-2 dataset are contributing to the observed biases. It is however, beyond the scope of this paper to identify the cause of further retrieval biases.

P16, L17: We have amended our text to clarify our meaning.

Evaluating the diurnal cycle in cloud top temperature from SEVIRI

Sarah Taylor[1], Philip Stier[1], Bethan White[1], Stephan Finkensieper[2], and Martin Stengel[2]

[1]Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, UK
[2]Deutscher Wetterdienst (DWD), Offenbach, Germany

Correspondence to: S. A. Taylor (sarah.taylor@physics.ox.ac.uk)

Abstract. The variability of convective cloud spans a wide range of temporal and spatial scales and is of fundamental importance for global weather and climate systems. Datasets from geostationary satellite instruments such as SEVIRI provide high time resolution observations across a large area. In this study we use data from SEVIRI to quantify the diurnal cycle of cloud top temperature within the instrument's field of view and discuss these results in relation to retrieval biases.

We evaluate SEVIRI cloud top temperatures from the new CLAAS-2 (CLoud property dAtAset using SEVIRI, Edition 2) dataset against CALIOP data. Results show a mean bias of + 0.44 K with a standard deviation of 11.7 K, which is in agreement with previous validation studies. Analysis of the spatiotemporal distribution of these errors shows that absolute retrieval biases vary from less than 5 K over the southeast Atlantic Ocean up to 30 K over central Africa at night. Night and daytime retrieval biases can also differ by up to 30 K in some areas, potentially contributing to biases in the estimated amplitude of the diurnal cycle. This illustrates the importance of considering spatial and diurnal variations in retrieval errors when using the CLAAS-2 dataset.

Keeping these biases in mind, we quantify the seasonal, diurnal and spatial variation of cloud top temperature across SEVIRI's field of view using the CLAAS-2 dataset. By comparing the mean diurnal cycle of cloud top temperature with the retrieval bias we find that diurnal variations in the retrieval bias can be small, but are often of the same order of magnitude as the amplitude of the observed diurnal cycle, indicating that in some regions the diurnal cycle apparent in the observations may be significantly impacted by diurnal variability in the accuracy of the retrieval.

We show that the CLAAS-2 dataset can measure the diurnal cycle of cloud tops accurately in regions of stratiform cloud such as the southeast Atlantic Ocean and Europe, where cloud top temperature retrieval biases are small and exhibit limited spatial and temporal variability. Quantifying the diurnal cycle over the tropics and regions of desert is more difficult, as retrieval biases are larger and display significant diurnal variability. CLAAS-2 cloud top temperature data are found to be of limited skill in measuring the diurnal cycle accurately over desert regions. In tropical regions such as Central Africa, the diurnal cycle can be described by the CLAAS-2 data to some extent, although retrieval biases appear to reduce the amplitude of the real diurnal cycle of cloud top temperatures.

This is the first study to relate the diurnal variations in SEVIRI retrieval bias to observed diurnal cycles in cloud top temperature. Our results may be of interest to those in the observation and modelling communities when using cloud top properties data from SEVIRI, particularly for studies considering the diurnal cycle of convection.

1 Introduction

Convection is one of the core building blocks of tropical weather and climate, playing important roles in large-scale atmospheric circulations, the hydrological cycle, the global energy budget and the transport of heat, moisture, momentum and aerosols throughout the troposphere (Grabowski and Petch, 2009). The temporal and spatial variability of convective cloud are therefore of fundamental importance for global weather and climate systems.

The diurnal and seasonal cycles of cloud top temperature (CTT), driven by changes in solar insolation, are among the strongest and most fundamental modes of variation in the global weather and climate systems. These cycles play important roles in the hydrological cycle, the global energy budget and the transport of heat, moisture, and momentum throughout the troposphere. While areas of convective cloud are associated with the strongest diurnal and seasonal cycles in CTT, areas of non-convective cloud also have observable diurnal and seasonal cycles (Yang and Slingo, 2001).

Spatial scales of convection range from thousands of kilometers for mesoscale convective systems (MCS) to a few kilometers for individual convective plumes, while time scales of convective variability range from minutes through to seasons. In particular, the diurnal and seasonal cycles of tropical convection, driven by variations in solar forcing, are among the strongest and most fundamental modes of variation in the global weather and climate systems.

A large number of observational studies using data from rain gauges (Wallace, 1975; Gray and Jacobsen, 1977), surface weather reports (Dai, 2001) and both polar-orbiting (Nesbitt and Zipser, 2003; Yang et al., 2008; Stratton and Stirling, 2012) and geostationary (Meisner and Arkin, 1987; Janowiak et al., 1994; Chen and Houze, 1997; Yang and Slingo, 2001; Schröder et al., 2009) satellites have attempted to quantify the diurnal cycle of convection cloud over land. These studies found an early afternoon maximum in convective precipitation, followed by a minimum in cloud top temperature approximately three hours later. The two features are thought to correspond to the beginning and end of the mature stage of convection (Schröder et al., 2009).

However, these large-scale convective features, driven by insolation, display regional and seasonal variations (Yang and Slingo, 2001; Schröder et al., 2009) and can be overridden by other factors such as orography (Yang and Slingo, 2001; Nesbitt and Zipser, 2003; Vondou et al., 2010), land-sea breezes (Chen and Houze, 1997; Yang and Slingo, 2001; Halladay et al., 2012) and the organisation of convection (Nesbitt and Zipser, 2003).

The amplitude of the diurnal cycle of convection CTT is smaller over the ocean than over land (Harrison et al., 1988), due to the ocean's higher heat capacity, and because ocean mixing distributes incoming solar radiation away from the surface. Over large areas of the ocean there is a small diurnal cycle in CTT, caused by stratiform cloud (Yang and Slingo, 2001; Wood, 2002; Duynkerke e During the night, longwave radiative cooling at the cloud top drives turbulent mixing, creating a deeper cloud layer. In the daytime, solar heating of the cloud top causes the cloud layer to become stably stratified, cutting off the transport of heat and moisture from the surface, creating a thinner cloud layer (Duynkerke et al., 2004).

Most studies show a substantial pre-dawn peak in convective cloud over the oceans (Janowiak et al., 1994; Yang and Slingo, 2001; Nesbitt and Zipser, 2003; Bain et al., 2010; Stengel et al., 2014). The mechanisms responsible for this overnight peak in convective cloud remain uncertain (Bain et al., 2010), but are thought to be related to atmospheric instability caused by night

~~time~~ nighttime radiative cooling (Randall et al., 1991) and to the presence of a larger number of MCSs during the night (Chen and Houze, 1997; Nesbitt and Zipser, 2003).

~~However, general~~ General circulation and numerical weather prediction models ~~which parameterise convective processes~~ struggle to realistically simulate spatial and temporal variability of cloud, due to the complexity of the processes involved, from large-scale atmospheric circulations and boundary layer processes, to convection and cloud microphysics. A particular concern is that models fail to capture the observed diurnal cycle of convective cloud (Yang and Slingo, 2001; Guichard et al., 2004; Grabowski et al., 2006; Stratton and Stirling, 2012).

This is generally the result of convection initiating shortly after sunrise, which develops too rapidly, quickly reaching the tropopause and producing precipitation (Guichard et al., 2004; Stratton and Stirling, 2012).

~~Studies by Guichard et al. (2004); Grabowski et al. (2006)~~ Studies by Guichard et al. (2004); Grabowski et al. (2006); Pearson et al. (20 Sato et al. (2009) show that in some cases, cloud resolving models (CRMs), which explicitly resolve convection, are capable of correctly predicting the amplitude and phase of the diurnal cycle in convection. Over land, this accuracy is strongly dependent on horizontal resolution, requiring grid lengths of around 1 km (Guichard et al., 2004), or 500 m (Grabowski et al., 2006). Over the ocean, horizontal resolution appears to be less important, which is likely due to differences in predominant cloud types and lifecycles (Sato et al., 2009).

Observations can both improve our theoretical understanding and provide a useful test of a model's ability to capture the various scales of variation of ~~convective cloud.~~ cloud. Regions of stratiform cloud cover large areas and can persist for days. However, spatial scales of convection range from thousands of kilometers for mesoscale convective systems (MCS) to a few kilometers for individual convective plumes, while time scales of convective variability range from minutes through to seasons.

While low Earth orbit satellites can provide observations at high spatial resolution, their temporal sampling is limited. Polar-orbiting satellites in sun-synchronous orbit, such as those in the A-Train constellation of satellites, observe any given point (except polar regions, which are observed more often) no more than twice per day and always at the same local solar time. Other low Earth orbit satellites, such as the Tropical Rainfall Measuring Mission (TRMM), are able to sample a given point at varying local solar times, thereby providing statistical diurnal cycle observations. Individual low Earth orbit satellites are therefore unable to observe the temporal evolution of ~~convective clouds~~cloud, particularly rapidly evolving convective cloud.

In comparison, while the spatial resolution of geostationary satellites is limited, they provide high temporal resolution observations over a large area. The Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument on board the geostationary Meteosat Second Generation (MSG) satellites has a spatial resolution of 3 km at the satellite nadir and a temporal resolution of 15 minutes, with a record of observations reaching back to 2004. Its field of view, hereafter referred to as the SEVIRI 'disc', covers the entire continents of Africa and Europe, as well as parts of the Middle East, ~~eastern~~ South America, and both the Atlantic ~~Ocean and the western part of the Indian ocean~~and Indian oceans. The continuous nature of SEVIRI's observations makes them ideal for investigating the temporal and spatial variability of cloud across a large area and period of time.

It should be noted that ~~passive imagers~~ the radiances observed by a passive imager such as SEVIRI ~~observe the radiometric height of the cloud. This differs from~~ are weighted averages of the vertical temperature profile of the atmosphere. The contribution of each layer of the atmosphere to the observed radiance can be described for each channel by a weighting function, which varies according to the viewing angle and atmospheric state. The cloud top height inferred from these radiances will therefore be lower and warmer than the physical cloud top height which can be measured ~~very accurately~~ by active lidar instruments such as the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) even for optically thick cloud (Sherwood et al., 2004). SEVIRI is therefore expected to underestimate cloud top pressure (CTP) and cloud top height (CTH), and overestimate ~~cloud top temperature (CTT)~~ CTT relative to CALIOP. This is due to differences in the properties observed by the two instruments, rather than an error in the SEVIRI retrieval.

Although it is often assumed that optically thick clouds have sharp boundaries and can be expected to radiate as black bodies (Sherwood et al., 2004; Stubenrauch et al., 2013), the radiometric height of a cloud may be several kilometers below physical cloud top height depending on its extinction profile at the cloud top and its vertical size (Stubenrauch et al., 2013). In particular, glaciated clouds tend to have poorly defined edges, even when convectively active (Sherwood et al., 2004), and so optical depths increase slowly with distance from the cloud top (Stubenrauch et al., 2013).

Sherwood et al. (2004) found that radiometric cloud tops retrieved from the Geostationary Operational Environmental Satellite 8 (GOES-8) were on average 1 km below, or 5-7 K above, the visible cloud tops observed by NASA's Cloud Physics Lidar (CPL). This bias increased to 2 km for the highest cloud tops, and was not found to vary with cloud albedo. Other studies (Heymsfield et al., 1991; Minnis et al., 2008; Stubenrauch et al., 2010, 2013) show similar biases of between 0.5 and 3 km for high clouds in the tropics.

An additional cause of differences between ~~these~~ the cloud properties observed by CALIOP and SEVIRI is the occurrence of optically thin clouds, or cloud layers. Using 532 nm measurements, CALIOP is capable of detecting cloud layers with optical depths of 0.01 (Vaughan et al., 2009; NASA, 2013), which SEVIRI is not sensitive enough to detect (Heidinger and Pavolonis, 2009; Stube ~~These clouds can usually be detected by CALIOP. Their impact~~ The impact of these clouds on passively measured radiation is ~~however very~~ small and thus not properly detected by passive imager sensors (Heidinger and Pavolonis, 2009; Stubenrauch et al., 2010; SA This is a particular issue in the case of pixels containing semi-transparent cloud types, which have ~~very~~ low optical depths, as it is difficult to fully account for contributions from surface radiation, or low cloud layers underneath (Smith and Platt, 1978). ~~SEVIRI is also not sensitive to clouds with very low optical depths which can be detected by lidar (Heidinger and Pavolonis, 2009; Stubenra~~

In this study we evaluate SEVIRI CTT retrievals dataset against data from CALIOP in order to consider spatial and diurnal variability in retrieval biases and investigate the seasonal and diurnal cycles of CTT across the entire SEVIRI disc. This is the first study to relate the diurnal variations in retrieval bias to observed diurnal cycles in CTT for SEVIRI cloud top properties as contained in the 12-year spanning SEVIRI dataset. The SEVIRI dataset is introduced in Sect. 2. Based on this dataset, spatial and seasonal patterns in CTT are examined in Sect. 3.1. In Sect. 3.2 the SEVIRI cloud top temperature data is compared to CALIOP measurements, extending on existing validation analyses in order to consider the implications of spatial and diurnal variations in retrieval bias for the SEVIRI-based diurnal cycles of CTT. In Sect. 3.3, the diurnal variability of cloud top temperature is quantified across the SEVIRI disc. Conclusions are presented in Sect. 4.

2 Data

The analysis presented here uses data from the ~~SEVIRI instrument on the geostationary MSG satellite. SEVIRI is an imager centred at approximately~~ MSG series of geostationary satellites. Operational MSG satellites are centered near 0 ~~° longitude, with~~ ° longitude (although gaps in coverage are occasionally filled using data from older versions of the satellite, located near 3.5° East (Meirink, 2013)). The SEVIRI imager is the main payload on the MSG series of satellites. It has 12 spectral channels in the visible, near-infrared and infrared~~. The instrument has~~ , a temporal resolution of 15 minutes (5 minutes in rapid scan mode, covering a limited area) and a spatial resolution ranging from 3 km at satellite nadir (1 km in its high-resolution visible channel) to 11 km at the edge of its field of view.

The European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)'s Satellite Application Facility on Climate Monitoring (CMSAF) has produced an updated twelve year dataset of cloud top properties based on SEVIRI measurements, named the CLoud property dAtAset using SEVIRI, Edition 2 (CLAAS-2) (Benas et al., 2016). CLAAS-2 contains the only retrieval of cloud top properties currently available at full SEVIRI spatial and temporal resolution and over a period of several years. The specific dataset used in this study is the instantaneous cloud top parameters product (CTX version 002) (Benas et al., 2016). The dataset is available for the period 2004-2015 and retrieved at full SEVIRI spatial and temporal resolution.

~~The retrieval algorithm applied to produce the CLAAS-2 dataset was developed in the framework of the EUMETSAT Satellite Application Facility on Nowcasting (NWCSAF). The full algorithm (NWCSAF/MSGv2012) is documented in Derrien (2013). To summarize, a multi-spectral threshold method, applying a variety of threshold tests in different channels, is used to obtain a pixel-resolution cloud mask. These tests vary according to conditions such as solar illumination, satellite angle and surface type. Cloud type and cloud top properties are also retrieved for pixels classed as fully cloudy or cloud-contaminated. No further retrievals are made for clear sky pixels, or pixels classified as containing broken clouds.~~

~~The NWCSAF/MSG retrieval calculates the vertical placement of clouds for fully cloudy pixels via a CTP retrieval, which varies according to cloud type, atmospheric conditions, and the data available.~~

~~In the case of optically thick clouds, the vertical position of the cloud top can be calculated from measurements of infrared brightness temperatures in atmospheric window channels by simply correcting for above cloud atmospheric absorption (Smith and Platt, 19~~ ~~However, in the case of partially cloudy pixels, or semi-transparent cloud, surface radiation may be transmitted through the cloud, or gaps in the cloud cover. In such cases a multi-spectral approach is needed (Smith and Platt, 1978; Menzel et al., 1983; Schmetz et a~~

~~For opaque clouds at all heights, CTP is therefore diagnosed by comparing observed 10.8 μm brightness temperatures to values simulated by the Radiative Transfer Model for TOVS (RTTOV) radiative transfer model. For high, semi-transparent cloud the infrared window intercept (Schmetz et al., 1993), or radiance rationing (Menzel et al., 1983) retrieval methods are attempted. Full details of the retrieval algorithm in other cases can be found in Derrien (2013).~~

~~Finally, for all cloud types, cloud top temperature and height are calculated from cloud top pressure using input from ERA-Interim (ECMWF Reanalysis) reanalysis fields.~~

The retrieval algorithm applied to produce the CLAAS-2 dataset was developed in the framework of the EUMETSAT Satellite Application Facility on Nowcasting (NWCSAF). The full algorithms (NWCSAF/MSGv2012) are documented in Derrien (2013). A short summary of the cloud detection and the cloud top pressure retrieval is presented here.

Cloud detection is based on a multi-spectral threshold method, applying a variety of threshold tests in different channels, in order to obtain a pixel-resolution cloud mask. These tests vary according to conditions such as solar illumination (day, night and twilight), satellite angle and surface type (land, ocean and coast). During the daytime SEVIRI visible channel information is available and used. For all pixels identified as cloudy the CTP is derived.

The cloud top pressure retrieval is performed for all cloudy pixels, excluding pixels which contain broken cloud conditions as identified by an intermediate cloud typing procedure. This typing procedure further determines whether the clouds are semi-transparent or optically thick, triggering different approaches for the CTP retrieval. It is assumed that the detection of these intermediate cloud types is more accurate under daytime conditions, due to the availability of visible channel information. This may subsequently result in more accurate CTP retrievals during the daytime, as found later in this paper.

For high, semi-transparent clouds an H2O/IRW (InfraRed Window) intercept method (Schmetz et al., 1993) is attempted. If this is not successful the CTP is determined by a radiance rationing method (Menzel et al., 1983). For all other clouds, the 10.8 μm brightness temperature is simulated using RTTOV and ERA-Interim (Dee et al., 2011) data while modulating the CTP of the cloud. This allows the CTP for which the simulated brightness temperatures best fits that observed to be identified.

Retrieved CTP is used to infer cloud top temperature using ERA-Interim profiles by selecting the temperature at the same level at which the cloud top pressure is located in the pressure profile. It can be assumed that different NWP models would lead to different CTP retrievals and thus to different cloud top temperatures. For both the radiative transfer simulation and the CTP to CTT conversion, 3-hourly ERA-Interim data on a 0.5 degree grid are utilized. A linear temporal interpolation between the previous and subsequent forecast is performed to represent best the meteorological conditions at the observation time. The vertical resolution of the ERA-Interim data used is provided in pressure levels at: 1000, 950, 925, 900, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, and 10 hPa. The discrete vertical resolution will therefore introduce some uncertainty in the retrieval results of cloud top pressure as well as in the conversion from CTP to CTT.

Benas et al. (2016) compare the CLAAS-2 CTT data to measurements from the CALIOP instrument between 2006 and 2015. The comparison is made for the CALIOP cloud layer at which the vertically integrated cloud optical depth (COD) is at least 0.2. For this setting they find a mean bias of 2.1 K and a bias corrected root mean squared error (RMSE) of 16.3 K. Bias and RMSE amount to 11.4 K and 22.1 K when no COD thresholds are applied.

In this study we use CLAAS-2 monthly mean diurnal cycle (MMDC) CTT products, provided at a spatial resolution of 0.25° and a temporal resolution of one hour, to quantify the diurnal cycle of CTT across the SEVIRI disc. We also validate instantaneous CTT retrievals (as produced by the NWCSAF/MSG algorithm and included in the CLAAS-2 dataset) against CALIOP, in order to investigate the implications of both the spatial and diurnal variability in the retrieval bias for the accurate quantification of diurnal cycles in cloud top temperature.

All SEVIRI data used in this study are drawn from the CLAAS-2 cloud top temperature dataset, retrieved from SEVIRI observations using the NMCSAF/MSGv2012 algorithm. For clarity, the term ‘SEVIRI’ will be used to refer to this dataset hereafter.

3 Results

3.1 Mean cloud top temperature

Seasonal mean SEVIRI cloud top temperatures for the period 2005-2015 are shown in ~~(~~Fig. 1~~).~~, while the total number of SEVIRI cloud top temperature retrievals are shown in Fig. 2. As expected, the warmest CTTs are observed over the ocean and the coldest over land, where a strong diurnal cycle in land surface temperatures drives convective initiation. Typical cloud regime patterns, showing deep convection over land in the region of the intertropical convergence zone (ITCZ), shallower convection over the central Atlantic ocean in the trade wind convergence zone and stratocumulus cloud in the southeast Atlantic ocean are evident. Cloud is most frequent over land, in the region of the ITCZ, and over the southeast Atlantic ocean, particularly in SON (Fig. 2). Very few clouds are observed over the Sahara desert and parts of the Middle East, as well as in southern Africa during JJA (Fig. 2).

Seasonal patterns in convection, driven by the movement of the (ITCZ) (Waliser and Gautier, 1993; Yang and Slingo, 2001; Schröder et al., 2009) can be clearly seen. In December, January and February (DJF), the ITCZ is shown as a band of cold cloud running from the southern Indian ocean, through central Africa (where it crosses the equator) and the West African coast before falling back below the equator, towards South America. In June, July and August (JJA), the ITCZ traces a more northerly position, largely located above the equator, stretching from the Gulf States, through the Sahel and trade wind convergence region, towards Venezuela.

In all seasons, regions with the coldest clouds have seasonal mean CTTs of between 200 K and 240 K, indicative of persistent deep convection in these areas. These clouds are concentrated in central Africa, the Amazon and the West African coast. The warmest CTTs are found in the region of persistent stratocumulus cloud in the southeast Atlantic Ocean, where seasonal mean CTTs range between 270 K and 290 K.

The position of the central Atlantic trade wind convergence zone is closely related to the seasonal movement of the ITCZ. Cloud top temperatures in this region fall to between 230 K and 250 K, with particularly cold clouds observed in March, April and May (MAM) and September, October and November (SON) due to the passage of the ITCZ. This indicates the presence of shallower convective cloud, initiated by the convergence of northern and southern hemisphere winds.

The concentration of cold clouds in Central and West Africa suggests that the absolute diurnal cycle of convection (the mean change in CTT throughout the day) ~~is likely to~~ may be strongest in these regions~~, due to the~~. However, this pattern could be the result of either the strong vertical development of deep convective clouds ~~. However, before~~ throughout the day, or the continuous presence of cold cloud. Before attempting to quantify the diurnal cycle of CTT at cloud top using SEVIRI CTT data however, it is necessary to consider the impact of spatial and diurnal variations in retrieval biases, which may have a significant impact on the diurnal cycle derived from this dataset.

3.2 Evaluation of SEVIRI cloud top temperature retrievals with CALIOP data

Biases in the CTT retrieval can be expected to display significant temporal and spatial variation. For example, Fig. 1 shows Figs. 1 and 2 show clear spatial and seasonal patterns in cloud type, while surface occurrence and cloud top temperature, which is likely indicative of variations in cloud type, particularly in the vicinity of the ITCZ. Surface emissions of longwave radiation also display spatial and temporal variations, particularly over land (Harrison et al., 1990; Wild et al., 2014). The implications of cloud type and land surface emissivity for the accuracy of cloud top property retrievals from SEVIRI were discussed in Sect. 21.

This study considers the implications of both the spatial and diurnal variability in the retrieval bias for the accuracy of diurnal cycle measurements across the SEVIRI disc. To this end one year of SEVIRI and CALIOP CTT retrievals were compared across the SEVIRI disc. This analysis was carried out using data from 2007, the first full year for which CALIOP data are available. Data was processed collocated for a single year of the twelve year CLAAS-2 dataset, balancing the need to process sufficient data points to be able to examine the spatial variability of retrieval bias with the large computational expense of collocating two large datasets.

3.2.1 Collocation

Although CALIOP is in sun-synchronous orbit, gathering data only at 13:30 and 01:30 LST (Winker et al., 2006), it provides a very accurate measurement of uses active measurements to observe CTH, and hence CTT, due to its active measurementswith a vertical resolution of between 30 and 60 m. It is therefore an excellent dataset for assessing the accuracy of SEVIRI cloud top retrievals. CALIOP also offers the advantage of a long-running dataset (2006 - present) and global coverage, allowing data to be compared across the entire SEVIRI disc over a long period of time.

CALIOP has a vertical resolution of between The CALIOP dataset used in this study is the same product used in the CLAAS-2 validation report Benas et al. (2016), the Lidar, Level 2, 5 km Cloud Layer, Validated Stage 1 Version 3 product (CAL LID L2 05kmCLay-ValStage1-V3-01) (NASA, 2013). CALIOP CTH is accurate to within 30 and m between the surface and 8.2 km, and to within 60 depending on the cloud's height in the atmosphere and is capable of detecting cloud layers with optical depths of 0.01 (McGill et al., 2007; Vaughan et al., 2009). It m between 8.2 km and 20.2 km above mean sea level (NASA, 2013). The dominant uncertainties in the vertical placement of CALIOP's cloud layers are noise in the backscatter signal and imperfect correction of the attenuation from overlying features (Vaughan et al., 2005), although, as only top-level features are used here, this last point is not relevant to this study. Due to high noise levels caused by solar background signals, the detection of cloud layers is more accurate during the nighttime than in the daytime, although even subvisible cirrus is detectable under both day and nighttime conditions (NASA, 2013; Vaughan et al., 2005). In addition, the amount of signal averaging required before atmospheric features can be retrieved is smaller for strongly scattering features than for weakly scattering features (Vaughan et al., 2005).

CALIOP measures CTH and subsequently uses the GEOS-5 (Goddard Earth Observing System Model, Version 5) atmospheric GCM model (Rienecker et al., 2008) to convert to convert form CTH to CTT (NASA, 2013). This conversion can be seen as a potential source of uncertainty in the CALIOP CTT representation, since the model has a coarser vertical and horizontal resolution than CALIOP. 0.625 degrees longitude by 0.5 degrees latitude grid, with 42 vertical pressure levels, extending to 0.01 hPa (NASA, 2017), compared to CALIOP's horizontal resolution of 5km and vertical resolution of between 30 and 60 m (NASA, 2013). Additionally, rising air parcels such as those found in convective clouds) are usually warmer than the surrounding air, as represented by the grid mean temperature of the model fields. The CALIOP dataset used in this study is the same product used in the CLAAS-2 validation report Benas et al. (2016), the Lidar, Level 2, 5 Cloud Layer, Validated Stage 1 Version 3 product (CAL LID L2 05kmCLay-ValStage1-V3-01) (NASA, 2013)

CALIOP is capable of detecting cloud layers with optical depths of 0.01 (McGill et al., 2007; Vaughan et al., 2009). The optical depth of atmospheric layers is derived using extinction-to-backscatter ratios, which vary depending on the assessed layer type (NASA, 2011). Misclassified layer types can therefore lead to inaccurate estimates of layer optical depth.This is the dominant contributor to optical depth uncertainties.

It should also be noted that, due to the fact that the vertical resolution of CALIOP varies, layer CODs are reported for atmospheric layers with a minimum vertical height of 30 m below 8.2 km and for layers with a minimum vertical height of 60 m above 8.2 km. The same COD above 8.2 km could therefore indicate different extinction profiles above and below 8.2 km. However, CALIOP CODs are shown to compare well with values retrieved from MODIS, although CALIOP slightly underestimates values at lower optical depths (NASA, 2011).

SEVIRI and CALIOP data are collocated by searching for the closest SEVIRI pixel (regardless of whether this pixel contains a cloud in the SEVIRI dataset) to each CALIOP observation. A temporal collocation window, centered around the time of the CALIOP overpass and accounting for the fact that a single SEVIRI timestep includes observations made over the course of a 12 minute time period, is also applied. As discussed below, different temporal collocation windows were tested in order to quantify the sensitivity of the resulting bias to the size of this window.

As SEVIRI's detection efficiency decreases at low optical depths, it is necessary to exclude very thin cloud layers from this comparison. Previous comparisons of SEVIRI and CALIOP data have excluded all cloud with an optical depth of less than 0.3 (Kniffka et al., 2013), 0.2, (Benas et al., 2016) and 0.1 (Stubenrauch et al., 2010; SAFNWC/MSG, 2012), while others have not excluded thin cloud at all (Reuter et al., 2009). In this analysis, mean statistics were calculated for a number of different collocation criteria. Due to the computational expense of collocating the datasets, different collocation criteria were tested using data for every 10th day in 2007.

Table 1 contains information on the seven different sets of collocation criteria tested. It indicates the maximum time window during which retrievals could be collocated, whether multi-layer clouds were included in the comparison and what COD threshold was used. Each set of criteria is identified by an abbreviation, which we use to refer to individual scenarios in the text and by a symbol, which we use to refer to scenarios in the figures.

The mean bias and RMSE for each of the collocation criteria in Table 1 are shown in Fig. 3. Statistics are plotted separately for all collocated data points, land retrievals and ocean retrievals. The number of SEVIRI CTT retrievals collocated with CALIOP for each of the sets of criteria and the number of land and ocean retrievals are shown in Fig. 4.

Adjusting the maximum time window for collocation from 60 to 15 minutes (60-ML-0 and 15-ML-0) does not have a large effect on the mean bias and RMSE (Fig. 3), or on the spatial distribution of the biases (Appendix A). However, Fig. 4 shows that a 15 minute collocation window reduces the number of collocated retrievals by 50%. A temporal collocation window of 60 minutes was therefore chosen for this analysis.

The effects of applying different COD thresholds (60-ML-0, 60-ML-03, 60-ML-1, 60-ML-2) (in order to account for differences in the sensitivity of the SEVIRI and CALIOP instruments to low cloud optical depths) were also considered. However, these thresholds ~~do not~~ are only able to remove biases due to the fact that CALIOP is sensitive to optically thin cloud layers which SEVIRI does not detect. It cannot affect biases between the radiometric cloud top observed by SEVIRI and the physical cloud top observed by CALIOP~~.~~, discussed in Sect. 1. Thresholds were applied to the ~~COD of the~~ top cloud layer from the CALIOP product, using COD as measured by CALIOP ~~and scenes~~ using backscatter from the 532 nm band. Scenes for which this top layer did not meet the threshold value were excluded from the analysis. This differs from the approach implemented by Benas et al. (2016), although it does not lead to a large difference in the spatial distribution of the mean biases (Appendix B).

The mean bias for combined land and ocean data is reduced from 11 K when no COD threshold is used to 2.4 K when a threshold of 0.3 is used (Fig. 3). This is further reduced to 0.4 K for a threshold of 1 and -0.9 for a threshold of 2. Similar decreases in the RMSE were observed (Fig. 3).

The total mean and ocean-only biases are negative for the 60-ML-2 collocation criteria, while the land values display a slight positive bias (Fig. 3). This is due to the fact that the majority of the clouds in this dataset are located in the southeast Atlantic Ocean (Fig. 5), a region of prevalent subtropical subsidence inversions. In the presence of a low level thermal inversion there are two possible solutions when using observed brightness temperatures to infer the vertical placement of cloud tops. As explained in Sect. 2, the SEVIRI retrieval algorithm (NWCSAF/MSGv2012) places the cloud top for low clouds at the pressure level which corresponds to the best fit between observed and simulated brightness temperatures. In cases of low level thermal inversions, the SEVIRI retrieval generally places the cloud above the thermal inversion (Derrien, 2013) and may therefore underestimate CTT in these areas. However, the CALIOP retrieval of CTT is based on direct observations of CTH and is therefore not subject to uncertainty with respect to the vertical placement of the cloud top in the first place (NASA, 2013; Hamann et al., 2014). However, the vertical resolution of the model fields used to convert CALIOP CTH to CTT is relatively coarse, which may introduce some uncertainty in the CALIOP values themselves. If the latter effect is small, the difference in approach to the vertical placement of clouds in the presence of inversions would result in a systematic negative bias in the SEVIRI retrieval in this region.

Finally, the impact of excluding multi-layer cloud from the collocation is considered (60-ML-0, 60-SL-0, 60-ML-1 and 60-SL-1, Table 1). Excluding multi-layer cloud when no COD is used (60-SL-0) results in the mean bias falling from 11.1 K to 3.9 K and the mean RMSE falling from 26.3 K to 20.6 K (Fig. 3), indicating that in the case of multi-layer cloud scenes, observed cloud top brightness temperatures are likely contaminated by longwave emissions from lower-level clouds. However, once a COD threshold of 1.0 is applied, excluding multi-layer cloud (60-SL-1) results in a much smaller change in the bias from 0.4 K to -0.4 K, with the mean RMSE falling from 11.8 K to 10.9 K (Fig. 3). This shows that, when the top cloud layer observed by CALIOP has a COD greater than one, the brightness temperatures observed by SEVIRI are no longer significantly contaminated by longwave emissions from lower-level clouds. CALIOP scenes containing multiple layers of cloud were therefore included.

To summarize, for all further analysis presented in this paper, we collocated SEVIRI and CALIOP data using the 60-ML-1 criteria, consisting of a 60 minute collocation window, inclusion of scenes with multiple layers of cloud, and a COD threshold of 1.0. The resulting cloud top temperature retrievals from SEVIRI and CALIOP were compared for the full year of 2007, as follows. For each 5 by 5 km CALIOP pixel within the SEVIRI disc:

- the highest reported CALIOP cloud layer was selected;
- the SEVIRI pixel with the closest latitude and longitude to that of the CALIOP pixel was selected;
- for this pixel, the nearest SEVIRI retrieval in time (within the allowed 60 minute time window) was identified;
- if multiple CALIOP retrievals fell within a single SEVIRI pixel, the values were averaged. This was only necessary towards the edges of the disc, beyond approximately 50°E and 50°W;
- the retrieval bias (calculated as SEVIRI minus CALIOP CTT) was calculated for each instance of collocated data.

3.2.2 Retrieval biases

The annual mean bias in SEVIRI CTT, calculated across the entire SEVIRI disc from January to December 2007, combining both daytime and nighttime conditions, is 0.44 K with a standard deviation of 11.7 K (Table 2). This is smaller than the 2.1 K mean bias calculated by Benas et al. (2016). The difference is likely due to the fact that a less stringent COD threshold of 0.3 was applied by Benas et al. (2016).

Statistics calculated over a variety of regions and conditions show that the biases between SEVIRI and CALIOP CTTs (SEVIRI minus CALIOP) are normally distributed, although in the case of data over the oceans, this distribution is skewed slightly towards negative values. This skew is likely due to the large number of slightly negative values found in the southeast Atlantic Ocean, as discussed in Sect.3.2.1). These distributions show small mean biases and large standard deviations. Over the ocean the mean bias in SEVIRI CTT is smaller, at -0.12 K, with a standard deviation of 10.5 K, but over land the mean bias rises to 2.38 K, with a standard deviation of 14.9 K (Table 2).

It should also be noted that the sign of the mean bias is negative over the ocean for both daytime and nighttime retrievals, although the bias is always positive over land. As discussed in Sect. 2, even once clouds with a low optical depth are filtered out, observations from SEVIRI are expected to detect a warmer radiometric CTT as compared to the ~~colder~~ higher, and therefore colder, physical cloud top ~~temperature observed~~ heights measured by CALIOP.

The spatial distribution of the number of fully cloudy pixels for which SEVIRI retrieved a CTT value and for which a corresponding CALIOP value was available during 2007 are shown in Fig. 5. Collocated retrievals are concentrated in the southeast Atlantic Ocean (Fig. 5), an area of almost perpetual cloudiness, where atmospheric inversions are prevalent. The negative biases over the ocean are again due to the different retrieval processes for SEVIRI and CALIOP with regard to the vertical placement of clouds in regions of atmospheric inversions (as discussed in Sect. 3.2.1).

Figure 6 shows the spatial distribution of mean SEVIRI minus CALIOP CTT for daytime, nighttime, day and nighttime combined, and the difference between night and daytime biases. It can immediately be seen that for both day and ~~night time~~ nighttime retrievals, there are large areas of very high mean bias, although, due to compensating biases, these are obscured in some areas when day and nighttime biases are plotted together. However, many of the areas of high bias correspond to regions with few cloud retrievals, such as the Sahara (Fig. 5).

SEVIRI is shown to generally overestimate CTT over land, by approximately 10-20 K during the day and 15-25 K (or more) at night (Fig. 6). These values are larger than those reported by Benas et al. (2016) who do not consider day and nighttime biases separately. However, these biases are in agreement with the expected discrepancy of 0.5-3.0 km (approximately 3-20 K assuming a 6.5 K/km lapse rate) between radiometric and physical cloud top height found by Sherwood et al. (2004); Minnis et al. (2008); Stubenrauch et al. (2010, 2013) in the case of high clouds.

~~Over the ocean, biases~~ Biases are relatively small over the majority of the ocean. CTT is underestimated by 5-10 K over large areas of the Atlantic ocean, while in other areas, such as the region of trade wind convergence, SEVIRI overestimates CTT by 5-15 K (Fig. 6)~~.~~ , which is within the expected range of discrepancy due to the differences between the radiometric and physical cloud top height (Sect. 1).

The large areas of slightly negative bias in the Atlantic ocean correspond to the areas of persistent atmospheric inversion. As explained previously, the small systematic bias in the SEVIRI CTT retrieval in this region is due to its treatment of subsidence inversions. As can be seen from Fig. 5, the majority of the collocated retrievals are located in this region. Therefore, the mean biases presented in Table 2 are heavily weighted towards the small negative biases in this region.

The difference in the magnitude of the biases over land and ocean is likely due to differences in the most common cloud regimes observed over land and ocean (Yang and Slingo, 2001; Schröder et al., 2009)~~, as well as the greater difficulty in accounting for variations in surface emissivity over land (Derrien, 2013)~~. As suggested by Sherwood et al. (2004) and Stubenrauch et al. (2013), the differing extinction profiles and vertical heights of convective and stratiform clouds results in larger differences between the radiometric and physical cloud top for tall convective clouds.

Unfortunately, it is not possible to fully characterize the diurnal variation in bias, as CALIOP data is only available for comparison at 01:30 and 13:30 LST. However, the two CALIOP overpasses can give at least an estimate of the potential size of this variation. For the purposes of this study, a bias in CTT which remained constant throughout the day would not be a barrier to quantifying either the amplitude of the diurnal cycle in CTT, or the average time of minimum CTT. However, the fact that biases can change dramatically from the day to nighttime overpasses of CALIOP is more problematic.

Both the mean values and the spatial distribution of the biases change significantly from day to night. Differences between mean nighttime and daytime biases in the SEVIRI CTT retrieval can be as large as 30 K in some areas (Fig. 6). There are strong positive differences over Sub-Saharan Africa and South America, with strong negative differences over the Sahara (although, as discussed above, there are very few CTT retrievals in the Sahara). Differences between night and daytime biases are generally smaller over the ocean and over Europe. In areas where the difference between daytime and nighttime biases (Fig. 6) is greater than, or equal to the observed magnitude of the diurnal cycle in cloud top temperature, the diurnal cycle observed from the retrieval may be a product of diurnal variability in the accuracy of the retrieval. It will therefore be necessary to consider the diurnal variability in retrieval bias when quantifying the diurnal cycle of cloud top temperature.

3.3 Diurnal cycle of CTT

The CLAAS-2 MMDC product was used to calculate three-month mean diurnal cycles of cloud top temperature across the SEVIRI disc. Data were averaged for the period 2005-2015 to produce a diurnal cycle with a temporal resolution of one hour, on a spatial grid of 0.25°. Ten years of data were required in order to produce relatively smooth diurnal cyclesensure that each regional hourly mean value was calculated using a large number (more than one thousand) of CTT retrievals from the ten year period, particularly over land, where cloud retrievals were relatively sparse. There were a few regions where it was not possible to reach this threshold, these are shaded in Figs. 8 and 9.

An example of the resulting diurnal cycle in cloud top temperature is shown in Fig. 7 for a grid box centred on 3.1°S, 16.4°E, (western Democratic Republic of the Congo, see cross in Fig. 10) for the months of SON. The amplitude of the diurnal cycle in CTT was calculated as the maximum minus minimum diurnal mean cloud top temperature, as shown by the arrow indicating an amplitude of 30 K. The local solar time of the minimum cloud top temperature was defined as the time at which the minimum daily mean CTT occurred, as shown by the dashed line at 19:00 local solar time (LST).

While the averaging process produced a coherent diurnal cycle in the majority of cases, the calculated diurnal cycle remained very noisy in a few areas, particularly during seasons when very few clouds were retrieved. The number of CTT retrievals is greatest over areas of the ocean where large, homogeneous stratiform cloud fields result in a large number of cloud-filled pixels and hence a large number of SEVIRI CTT retrievals, and in the region of the ITCZ, where convective cloud is concentrated (Fig. 2). In areas with very few cloud retrievals, such as the Sahara in all seasons and parts of southern Africa in JJA (Fig. 2), it will not be possible to accurately calculate a diurnal cycle of convection.

Spatial distributions of the total number of SEVIRI cloud top temperature retrievals available from the CLAAS-2 dataset during the period 2005-2015. Values are shown for each season.

Maps of the amplitude of the diurnal cycle in SEVIRI cloud top temperature (Fig. 8), calculated as shown in Fig. 7, show the smallest amplitudes located over the southeast Atlantic ocean in all seasons. Over the course of a typical day, stratocumulus cloud tops vary by less than 5 K in this region. Amplitudes increase to between 20 K and 30 K in the trade wind region, where there are more convective clouds. Over Africa and South America amplitudes generally range between 20 K and 50 K, with the seasonal changes tracking the movement of the ITCZ, seen as a migrating band of cold cloud tops in Fig. 1. The diurnal cycle is smaller in Europe where amplitudes range from 15 K in the north during DJF to 50 K in the Mediterranean during JJA.

For illustrative purposes, amplitudes of the diurnal cycle in CTT were plotted for areas with few CTT retrievals, but regions with retrievals at less than 15% of the processed time steps are indicated (Fig. 8). With the exception of southern Africa in JJA, these regions correspond to areas of desert and calculated amplitudes tend to be very large. Amplitudes exceed 60 K in areas of the Sahara and Namibian deserts in all seasons, as well as in Somalia during the December to March dry season (Higgins et al., 1978) and Southern Africa during the May to September dry season (Higgins et al., 1978) (Fig. 8). The large amplitudes observed in these areas are likely to be caused by a mixture of insufficient data (Fig. 2) and, particularly in the Sahara, by a large variation in the size of the retrieval bias between night and daytime conditions (Fig. 6).

In order to consider the effects of systematic differences in day and nighttime CTT retrieval biases in the SEVIRI dataset, the ratio of the amplitude of the diurnal cycle (Fig. 8) to the diurnal variability in the retrieval bias (Fig. 6) was calculated. In regions where this ratio is low, differences between systematic retrieval biases under day and nighttime conditions may contribute strongly to the amplitude of the observed diurnal cycle in cloud top temperatures. A threshold value of 5 was chosen for this ratio, as indicative of regions in which observed diurnal cycles in CTT may simply be artefacts of the diurnal variation in retrieval errors. Areas for which the ratio falls below this threshold are indicated in Fig. 8.

Maps of the phase of the diurnal cycle, defined at each grid box as the local solar time at which the minimum three-monthly mean CTT occurs (Fig. 7) are shown in Fig. 9. Regions with few clouds, or where diurnal variability in retrieval bias may significantly contribute to the observed diurnal cycle are illustrated in Fig. 9, as described for Fig. 8.

Over large areas of the ocean, minimum CTTs are observed at around 16:00. In the southeast Atlantic however the minimum is observed in the morning, at around 09:00.

In South America, minimum CTT generally occurs at around 20:00. Over Sub-Saharan Africa and Europe the minimum is generally observed at around 16:00, with some areas, particularly West Africa, the Sahel and parts of the Congo Basin, showing later peaks at around 18:00. These later peaks broadly track the movement of the ITCZ (Fig. 1) and could be due to more vigorous convection, persisting until later in the day. It could also be due to a mixture of different convective cloud types, including organised MCSs which can persist until the early morning and more isolated local convective cells which peak in the afternoon (Rickenbach et al., 2009; Pfeifroth et al., 2016).

Areas with few cloud retrievals, such as the Sahara desert in all seasons, and southern Africa in JJA, are noisier (Fig. 9). This may be due to the fact that there is simply not enough data in these regions to meaningfully diagnose the phase of the diurnal cycle in CTT. However, the regions with the fewest retrievals (Fig. 5 and Fig. 9) do not match exactly the regions of noise in Fig. 9. This indicates that the noise may also be caused by a mixture of different cloud types with different diurnal cycles.

The relationship between the observed amplitude and phase of the diurnal cycle and the retrieval biases presented in Sect. 3.2 were examined in more detail over the Sahara, central Africa and southeast Atlantic ocean. These areas were chosen because they all exhibit fairly consistent patterns of both retrieval bias and observed diurnal cycle properties and were designed to cover approximately 9×10^6 km^2 each. They also provide examples of desert, rainforest and ocean surface types. The locations of these three areas are illustrated in Fig. 10.

Seasonal mean SEVIRI diurnal cycles and retrieval biases for each of the regions in Fig. 10 were compared (Fig. 11). We have already shown that the SEVIRI dataset has different retrieval biases under daytime and nighttime conditions (Sect. 3.2), due to differences in solar illumination, cloud types, the availability of visible channel observations and, subsequently the exact retrieval algorithms used. Seasonal mean times of sunrise and sunset are therefore indicated for each region and mean retrieval biases, as calculated in Sect. 3.2 for the year 2007, are shown for both day and nighttime CALIOP overpasses.

In the southeast Atlantic Ocean, the bias is shown to be ~~very~~ small with no apparent diurnal cycle in the bias (Fig. 11). Mean CTTs reach a minimum at around 09:00 and persist until 16:00 in the DJF and MAM seasons. In JJA the cold clouds are more short-lived and in SON cloud top temperatures remain constant throughout the day.

In the Sahara, the amplitude of the diurnal cycle is almost 20 K, with a small diurnal cycle in the bias of around 5 K (Fig. 11). The warmest CTTs are observed at 05:00 and at noon, with the coldest cloud at 07:00 and 18:00. Although the diurnal cycle in the bias is less than the amplitude of the diurnal cycle, the bias results in cloud top temperatures retrieved during the day being too warm, indicating that a significant fraction of the amplitude observed over the Sahara may be due to differences between the day and nighttime conditions and hence the differences in the retrieval algorithms applied. Sudden changes in mean CTT around the times of sunrise and sunset in the Sahara are also seen in Fig. 11. In all seasons there is a secondary minimum in cold cloud top temperatures at around 06:00, about an hour before sunrise. This secondary minimum may indicate the presence of MCSs, or simply a change in values due to the change in retrieval algorithm at this point.

In central Africa, the diurnal cycle in the bias is around 7 K and the amplitude of the diurnal cycle is around 15 K throughout the year (Fig. 11). The warmest cloud is observed at 14:00, with the coldest cloud between 20:00 in DJF and 22:00 in JJA. There is a secondary minimum at 09:00 in all seasons except DJF, which may be caused by the presence of MCS, although it is also possible that this secondary peak is produced by the switch from night to daytime retrieval conditions. In contrast to the Sahara however, the diurnal cycles in retrieval bias create a smaller amplitude of the diurnal cycle in cloud top temperatures than would otherwise be observed. This size of the bias during the early morning may indicate a larger difference between the radiometric cloud top measured by SEVIRI and the physical cloud top measured by CALIOP due to differences in cloud types throughout the day. While the ratio of the diurnal cycle in mean CTT to the difference in retrieval biases is small, it appears that in this region the difference in the retrieval bias is acting to reduce, rather than increase, the observed diurnal cycle in CTT.

It is interesting to note that the broad shape of the diurnal cycle curves in the Sahara and Central Africa are similar, although the post-sunrise increase in CTT is delayed in Central Africa relative to the Sahara. ~~This could potentially be caused by the lower surface albedo of Central Africa relative to the Sahara , causing the Sahara to heat up more quickly, producing lower, warmer clouds. Any such effect would be delayed by a few hours in Central Africa due to the buffering effect of the large, dark rainforests in the region~~The reasons for the time lag between diurnal cycles of CTT in Central Africa and the Sahara are not known yet and will be addressed by future studies.

4 Conclusions

In this study we evaluated SEVIRI cloud top temperature data, as retrieved by the NWCSAF/MSGv2012 algorithm and included in the updated CLAAS-2 dataset, against CALIOP and attempted to quantify spatial and diurnal variabilities in retrieval biases. We also quantified the amplitude and phase of the diurnal cycle in cloud top temperatures observed by SEVIRI. Comparing our measurements of the diurnal cycle in mean CTT and retrieval bias we show that diurnal variations in the retrieval bias are often of the same order of magnitude as the amplitude of the observed diurnal cycle. Areas in which there was insufficient data to accurately calculate the diurnal cycle in CTT, or in which the observed cycle was an artefact of retrieval biases, were identified.

SEVIRI and CALIOP data were collocated using a 60 minute collocation window and a COD threshold of 1.0. Scenes with multiple layers of cloud were included. By collocating SEVIRI and CALIOP CTT retrievals for the whole year of 2007, we show that mean errors in the SEVIRI retrieval can vary from less than 5 K to more than 30 K across the SEVIRI disc, and by up to 30 K between the daytime and nighttime overpasses of CALIOP. However, mean errors across the SEVIRI disc are small, at approximately 0.44 K with a standard deviation of 11.7 K. This shows the importance of considering spatial and diurnal variations in retrieval error when using this dataset.

We believe that the difference between the radiometric cloud tops observed by retrieved from SEVIRI and the physical cloud tops observed measured by CALIOP may account for a significant fraction of the biases found in this analysis. As explained in Sect. 1, previous studies indicate that biases of less than 0.5-3.0 km (approximately 3-20 K) could potentially be explained by this difference, even for optically thick clouds. As cloud layers with an optical depth of less than one were not included in the comparison of SEVIRI and CALIOP CTT data, we expect biases to be largest in the case of optically thick clouds with poorly defined edges, such as glaciated clouds. In addition, the small negative bias observed over the southeast Atlantic Ocean is likely related to uncertainties introduced by the use of models to estimate the vertical placement of clouds for both SEVIRI and CALIOP datasets. Additional biases described in this paper may be due to retrieval errors included Any biases which fall outside of both the 3-20 K range and the region of subsidence in the southeast Atlantic Ocean may be caused by other retrieval errors in the SEVIRI dataset, although it is not within the scope of this study to identify the cause of these potential retrieval errors.

Keeping these uncertainties in mind, the seasonal, diurnal and spatial variation of cloud top temperatures were quantified across the SEVIRI disc. By plotting the seasonal mean amplitude and phase of the diurnal cycle in cloud top temperature, we show that SEVIRI is able to capture details of the diurnal cycle of convection, across several continents. We show that the CLAAS-2 dataset measures the diurnal cycle of cloud tops accurately in regions of stratiform cloud such as the southeast Atlantic and Europe, where retrieval biases are small and exhibit limited spatial and temporal variability. Quantifying the diurnal cycle over the tropics and regions of desert is more difficult, as biases are larger and more variable.

Looking at three areas in detail (the southeast Atlantic Ocean, the Sahara desert and Central Africa), we analyse the relationships between the diurnal cycle in cloud top temperature and retrieval biases. We show that retrieval biases in the southeast Atlantic are small enough to detect a small but persistent diurnal cycle of approximately 5 K in the area, with cold clouds peaking between 11:00 and 15:00 local solar time. However, the CLAAS-2 dataset is shown to be of limited skill in measuring the diurnal cycle over the Sahara, which may be due to generally low cloud cover in desert regions and a possible dominance of optically thin clouds such as cirrus outflow from tropical convection when clouds are present. In the Sahara, variability in the bias appears to contribute to an excessively large amplitude of the diurnal cycle, with a large amount of spatial and seasonal variability in the phase. In tropical regions such as Central Africa, a relatively large variability in the retrieval biases appears to dampen the signal from a very strong observed diurnal cycle, with minimum cloud top temperatures occurring consistently between 20:00 and 22:00.

While this study highlights the importance of considering spatial and diurnal variations in retrieval errors when using SEVIRI data, it is also the case that observations from passive imagers in geostationary orbit provide valuable observations of the temporal and spatial variability of cloud on scales which are not available from polar-orbiting satellites such as CALIOP. We therefore see our results as guidance for the observation and modelling communities when using SEVIRI cloud top properties, particularly for studies considering the diurnal cycle of cloud top properties.

Appendix A: Comparing results of different collocation time windows

The insensitivity of the calculated bias in SEVIRI CTT to a change in the collocation window used from 60 minutes (i.e. ~~a maximum temporal distance between SEVIRI and CALIOP retrievals of 30 minutes~~within $\pm$ 15 minutes of a CALIOP overpass) to 15 minutes (i.e. ~~a maximum temporal distance between retrievals of~~ within $\pm$ 7.5 minutes of a CALIOP overpass) is initially surprising. We collocated SEVIRI and CALIOP CTTs, for the full year of 2007, using both 60 minute and a 15 minute collocation windows. This amounts to an extra 22.5 minutes between CALIOP and SEVIRI retrievals in the 60 minute window case, as compared to the 15 minute case.

There is no significant change in either the magnitude, or spatial distribution of the observed biases between the two cases (Fig. 6 and Fig. A.1). However, by reducing the collocation window to 15 minutes, the number of collocated data points is reduced and the spatial patterns become less clear.

The biases shown in Figs. 6 and A.1 consist of biases due to differences in the retrieval processes of the SEVIRI and CALIOP datasets (the retrieval bias) and to spatial and temporal differences in the scenes observed by the two instruments (the collocation bias).

If the size of the retrieval bias increased when moving from a 15 to 60 minute collocation window, we would expect either the mean biases in Fig. 6 to be larger than those shown in Fig. A.1, or, if the mean values are obscuring compensating errors (for example from observations before and after the CALIOP overpass), for the standard deviation of the retrieval biases in the 60 minute case to be larger than those in the 15 minute case. However, there is little difference between the two sets of maps (Fig. A.2). This indicates that the size of the collocation biases does not increase significantly when using a 60 minute time window in place of a 15 minute window.

There are many reasons to think that the collocation bias may be small relative to the large retrieval biases seen in many parts of the SEVIRI disc. For example, over areas of stratiform cloud, cloud top temperature is unlikely to change significantly over the space of the extra 22.5 minutes allowed by a 60 minute collocation window. In more convective areas, some clouds may develop significantly over the course of the larger time window, but the cloud top temperature of mature convective cloud systems and convective anvils will be more stable over time.

Appendix B: Cloud optical depth threshold methodology

When collocating SEVIRI and CALIOP retrievals of CTT, scenes observed by CALIOP were excluded if the top cloud layer had an optical depth of less than 1. This differs from the approach implemented by Benas et al. (2016), who compared CTT values for the first CALIOP layer at which the top-down, vertically integrated cloud optical depth exceeded the threshold value. The number of scenes containing cirrus cloud is therefore reduced in this analysis, compared to that of Benas et al. (2016). This is likely to increase the weighting of statistics presented in Sect. 3.2 towards the southeast Atlantic ocean, where there are few cirrus clouds. However, it does not impact the weighting of statistics elsewhere in the study, where the data is not limited to retrievals which can be collocated to a CALIOP overpass. A comparison of mean SEVIRI day and nighttime retrieval biases (Fig. B.1) with a similar plot in Benas et al. (2016) (Fig. 6-15, row 3, column 3) indicates that this difference in methodology does not lead to a large difference in the spatial distribution of the mean retrieval biases.

Author contributions. S.A.T. and P.S. co-designed the study. S.T. analysed the data and wrote the paper. P.S. and B.A.W. provided suggestions for the methodology, discussed results and commented on the manuscript at all stages. S.F and M.S. developed the CLAAS-2 dataset and provided advice on its use.

Acknowledgements. SEVIRI CLAAS version 2 data were obtained from EUMETSAT's Climate Monitoring Satellite Applications Facility (CMSAF). CALIOP data were obtained from the NASA Langley Research Center Atmospheric Science Data Center. This work was supported funding from the European Research Council under the European Union's Seventh Framework Programme (FP7 2007-2013) ERC grant agreement no. FP7-280025.

References

Bain, C. L., Magnusdottir, G., Smyth, P., and Stern, H.: Diurnal cycle of the Intertropical Convergence Zone in the east Pacific, J .Geophys. Res., 115, 2010.

Benas, N., Finkensieper, S., van Zadelhoff, G., Hanschmann, T., Stengel, M., and Meirink, J. F.: Validation Report SEVIRI cloud products Edition 2 (CLAAS-2), Tech. rep., EUMETSAT Satellite Application Facility on Climate Monitoring, in press, 2016.

Chen, S. S. and Houze, R. A.: Diurnal variation and life-cycle of deep convective systems over the tropical Pacific warm pool, Q. J. Roy. Meteor. Soc., 123, 357–388, 1997.

Dai, A.: Global Precipitation and Thunderstorm Frequencies. Part II : Diurnal Variations, J. Climate, 14, 1,112–1,128, 2001.

Dee, D. P., Uppala, S. M., Simmons, a. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. a., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, a. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, a. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, a. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, http://doi.wiley.com/10.1002/qj.828, 2011.

Derrien, M.: Algorithm Theoretical Basis Document for "Cloud Products" (CMa-PGE01 v3.2, CT-PGE02 v2.2 and CTTH-PGE03 v2.2), SAF/NWC/CDOP2/MFL/SCI/ATBD/01, Issue 3, Rev. 2.1, Tech. rep., EUMETSAT Satellite Application Facility on Nowcasting and Shortrange Forecasting, 2013.

Duynkerke, P. G., de Roode, S. R., van Zanten, M. C., Calvo, J., Cuxart, J., Cheinet, S., Chlond, A., Grenier, H., Jonker, P. J., Köhler, M., Lenderink, G., Lewellen, D., Lappen, C.-L., Lock, A. P., Moeng, C.-H., Müller, F., Olmeda, D., Piriou, J.-M., Sánchez, E., and Sednev, I.: Observations and numerical simulations of the diurnal cycle of the EUROCS stratocumulus case, Q. J. Roy. Meteor. Soc., 130, 3269–3296, 2004.

Grabowski, W. W. and Petch, J. C.: Deep convective clouds, in: Clouds in the Perturbed Climate System, edited by Lupp, J. and Turner, M., pp. 197–216, Massachusets Institute of Technology and Frankfurt Institute for Advanced Studies, Cambridge, Massachusetts, 1 edn., 2009.

Grabowski, W. W., Bechtold, P., Cheng, A., Forbes, R., Halliwell, C., Khairoutdinov, M., Lang, S., Nasuno, T., Petch, J., Tao, W.-K., Wong, R., Wu, X., and Xu, K.-M.: Daytime convective development over land: A model intercomparison based on LBA observations, Q. J. Roy. Meteor. Soc., 132, 317–344, 2006.

Gray, W. M. and Jacobsen, R. W.: Diurnal variation of deep cumulus convection, Mon. Weather Rev., 105, 1,171–1,188, 1977.

Guichard, F., Petch, J., Redelsperger, J.-L., Bechtold, P., Chaboureau, J.-P., Cheinet, S., Grabowski, W., Grenier, H., Jones, C., Köhler, M., Piriou, J.-M., Tailleux, R., and Tomasini, M.: Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models, Q. J. Roy. Meteor. Soc., 130, 3139–3172, 2004.

Halladay, K., Malhi, Y., and New, M.: Cloud frequency climatology at the Andes/Amazon transition: 1. Seasonal and diurnal cycles, J. Geophys. Res., 117, 2012.

Hamann, U., Walther, A., Baum, B., Bennartz, R., Bugliaro, L., Derrien, M., Francis, P., Heidinger, A., Joro, S., Kniffka, A., Le Gléau, H., Lockhoff, M., Lutz, H.-J., Meirink, J. F., Minnis, P., Palikonda, R., Roebeling, R., Thoss, A., Platnick, S., Watts, P., and Wind, G.: Remote sensing of cloud top pressure/height from SEVIRI: analysis of ten current retrieval algorithms, Atmos. Meas. Tech. Discuss., 7, 401–473, 2014.

Harrison, E. F., Brooks, D. R., Minnis, P., Wielicki, B. A., Staylor, W. F., Gibson, G. G., Young, D. F., and Frederick M. Den: First estimates of the diurnal variation of longwave radiation from the multiple-satellite Earth radiation budget experiment (ERBE), B. Am. Meteorol. Soc., 69, 1144–1151, 1988.

Harrison, E. F., Minnis, P., Barkstrom, B. R., Ramanathan, V., Cess, R. D., and Gibson, G. G.: Seasonal Variation of Cloud Radiative Forcing Derived From the Earth Radiation Budget Experiment, J .Geophys. Res., 95703, 687–18, 1990.

Heidinger, A. K. and Pavolonis, M. J.: Gazing at Cirrus Clouds for 25 Years through a Split Window. Part I: Methodology, J. Appl. Meteorol. Clim., 48, 1100–1116, 2009.

Heymsfield, G. M., Spinhirne, J. D., and Fulton, R.: Aircraft overflight measurements of midwest severe storms - Implications on geosynchronous satellite interpretations, Monthly Weather Review, 1991.

Higgins, G., Kassam, A., Kowal, J., Sarraf, S., Arnoldussen, H., Frere, M., Hrabovszki, J., and van Velthuizen, H.: Report on the Agro-ecological zones project. Vol.1 Methodology and results for Africa. World soil resources report No 48., Tech. rep., AGLS, Food and Agriculture Organization of the United Nations, Rome, Italy, 1978.

Janowiak, J. E., Arkin, P. A., and Morrissey, M.: An examination of the diurnal cycle in oceanic tropical rainfall using satellite and in situ data, Mon. Weather Rev., 122, 2,296–2,311, 1994.

Kniffka, A., Lockhoff, M., Stengel, M., and Meirink, J. F.: Validation Report - SEVIRI cloud products, SAF/CM/DWD/VAL/SEV/CLD Issue 1, Rev. 1.2, Tech. rep., EUMETSAT Satellite Application Facility on Climate Monitoring, 2013.

McGill, M. J., Vaughan, M. a., Trepte, C. R., Hart, W. D., Hlavka, D. L., Winker, D. M., and Kuehn, R.: Airborne validation of spatial properties measured by the CALIPSO lidar, J. Geophys. Res., 112, 2007.

Meirink, J. F.: Algorithm Theoretical Basis Document, Cloud Physical Products, SEVIRI. V 1.2, Tech. rep., EUMETSAT Satellite Application Facility on Climate Monitoring. SAF/CM/KNMI/ATBD/SEV/CPP, 2013.

Meisner, B. N. and Arkin, P. A.: Spatial and annual variations in the diurnal cycle of large-sacle tropical convective cloudiness and precipitation, Mon. Weather Rev., 115, 2,009–2,032, 1987.

Menzel, W. P., Stewart, T. R., and Smith, W. L.: Improved cloud motion wind vector and altitude assignment using VAS, J. Clim. Appl. Meteorol., 22, 377–384, 1983.

Minnis, P., Yost, C. R., Sun-Mack, S., and Chen, Y.: Estimating the top altitude of optically thick ice clouds from thermal infrared satellite observations using CALIPSO data, Geophysical Research Letters, 35, 1–6, 2008.

NASA: CALIPSO Quality Statements Lidar Level 2 Cloud and Aerosol Layer Products Version Releases: 3.01, 3.02, Tech. rep., NASA, 2011.

NASA: CALIPSO: Data User's Guide - Data Product Descriptions - Lidar Level 2 Cloud and Aerosol Layer Products, http://www-calipso.larc.nasa.gov/resources/calipso_users_guide/data_summaries/layer/index.php, 2013.

NASA: GEOS-5 FP-IT Product Details, https://gmao.gsfc.nasa.gov/products/GEOS-5_FP-IT_details.php#output_product_highlights, 2017.

Nesbitt, S. W. and Zipser, E. J.: The Diurnal Cycle of Rainfall and Convective Intensity according to Three Years of TRMM Measurements, J. Climate, 16, 1,456–1,475, 2003.

Pearson, K. J., Hogan, R. J., Allan, R. P., Lister, G. M. S., and Holloway, C. E.: Evaluation of the model representation of the evolution of convective systems using satellite observations of outgoing longwave radiation, Journal of Geophysical Research Atmospheres, 115, 1–11, doi:10.1029/2010JD014265, 2010.

Pearson, K. J., Lister, G. M. S., Birch, C. E., Allan, R. P., Hogan, R. J., and Woolnough, S. J.: Modelling the diurnal cycle of tropical convection across the 'grey zone', Quarterly Journal of the Royal Meteorological Society, 140, 491–499, doi:10.1002/qj.2145, 2014.

Pfeifroth, U., Trentmann, J., Fink, A. H., and Ahrens, B.: Evaluating satellite-based diurnal cycles of precipitation in the African tropics, Journal of Applied Meteorology and Climatology, 55, 23–39, 2016.

Randall, D. A., Harshvardhan, and Dazlich, D. A.: Diurnal Variability of the Hydrologic Cycle in a General Circulation Model, 1991.

Reuter, M., Thomas, W., Albert, P., Lockhoff, M., Weber, R., Karlsson, K.-G., and Fischer, J.: The CM-SAF and FUB Cloud Detection Schemes for SEVIRI: Validation with Synoptic Data and Initial Comparison with MODIS and CALIPSO, J. Appl. Meteorol. Clim., 48, 301–316, 2009.

Rickenbach, T., Ferreira, R. N., Guy, N., and Williams, E.: Radar-observed squall line propagation and the diurnal cycle of convection in Niamey, Niger, during the 2006 African monsoon and multidisciplinary analyses intensive observing period, Journal of Geophysical Research Atmospheres, 114, 1–8, 2009.

Rienecker, M. M., Suarez, M. J., Todling, R., Bacmeister, J., Takacs, L., Liu, H.-C., Gu, W., Sienkiewicz, M., Koster, R. D., R., G., Stajner, I., and Nielsen, J. E.: The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1 , 5.1.0, and 5.2.0, Tech. rep., NASA, 2008.

SAFNWC/MSG: Algorithm Theoretical Basis Document for “Cloud Products" (CMa-PGE01 v3.2, CT-PGE02 v2.2 and CTTH-PGE03 v2.2) SAF/NWC/CDOP/MFL/SCI/ ATBD/01, Issue 3, Rev. 2, Tech. rep., NWC-SAF, 2012.

Sato, T., Miura, H., Satoh, M., Takayabu, Y. N., and Wang, Y.: Diurnal Cycle of Precipitation in the Tropics Simulated in a Global Cloud-Resolving Model, J. Climate, 22, 4809–4826, 2009.

Schmetz, J., Holmlund, K., Hoffman, J., Strauss, B., Mason, B., Gaertner, V., Koch, A., and Van De Berg, L.: Operational Cloud-Motion Winds from Meteosat Infrared Images, J. Appl. Meteorol., 32, 1206–1225, 1993.

Schröder, M., König, M., and Schmetz, J.: Deep convection observed by the Spinning Enhanced Visible and Infrared Imager on board Meteosat 8: Spatial distribution and temporal evolution over Africa in summer and winter 2006, J .Geophys. Res., 114, 2009.

Sherwood, S. C., Chae, J. H., Minnis, P., and McGill, M.: Underestimation of deep convective cloud tops by thermal imagery, Geophysical Research Letters, 31, 1–4, 2004.

Smith, W. L. and Platt, C. M. R.: Comparison of Satellite-Deduced Cloud Heights with Indications from Radiosonde and Ground-Based Laser Measurements, 1978.

Stengel, M. S., Kniffka, A. K., Meirink, J. F. M., Lockhoff, M. L., Tan, J. T., and Hollmann, R. H.: CLAAS: The CM SAF cloud property data set using SEVIRI, Atmos. Chem. Phys., 14, 4297–4311, 2014.

Stratton, R. A. and Stirling, A. J.: Improving the diurnal cycle of convection in GCMs, Q. J. Roy. Meteor. Soc., 138, 1121–1134, 2012.

Stubenrauch, C. J., Cros, S., Guignard, A., and Lamquin, N.: A 6-year global cloud climatology from the Atmospheric InfraRed Sounder AIRS and a statistical analysis in synergy with CALIPSO and CloudSat, Atmos. Chem. Phys., 10, 7,197–7,214, 2010.

Stubenrauch, C. J., Rossow, W. B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H., Di Girolamo, L., Getzewich, B., Guignard, A., Heidinger, A., Maddux, B. C., Menzel, W. P., Minnis, P., Pearl, C., Platnick, S., Poulsen, C., Riedi, J., Sun-Mack, S., Walther, A., Winker, D., Zeng, S., and Zhao, G.: Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX radiation panel, Bulletin of the American Meteorological Society, 94, 1031–1049, 2013.

Vaughan, M. A., Winker, D. M., and Powell, K. A.: CALIOP Algorithm Theoretical Basis Document Part 2 : Feature Detection and Layer Properties Algorithms, Tech. Rep. September, NASA, 2005.

Vaughan, M. A., Powell, K. a., Winker, D. M., Hostetler, C. a., Kuehn, R. E., Hunt, W. H., Getzewich, B. J., Young, S. a., Liu, Z., and McGill, M. J.: Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements, Journal of Atmospheric and Oceanic Technology, 26, 2034–2050, 2009.

Vondou, D. A., Nzeukou, A., and Kamga, F. M.: Diurnal cycle of convective activity over the West of Central Africa based on Meteosat images, Int. J. Appl. Earth Obs., 12, 58–62, 2010.

Waliser, D. E. and Gautier, C.: A satellite-derived climatology of the ITCZ, J. Climate, 6, 2,162–2,174, 1993.

Wallace, J. M.: Diurnal variations in precipitation and thunderstorm frequency over the contermonous United States, Mon. Weather Rev., 103, 406–419, 1975.

Wild, M., Folini, D., Hakuba, M. Z., Schar, C., Seneviratne, S. I., Kato, S., Rutan, D., Ammann, C., Wood, E. F., and Konig-Langlo, G.: The energy balance over land and oceans: an assessment based on direct observations and CMIP5 climate models, Climate Dynamics, pp. 3393–3429, 2014.

Winker, D. M., Hostetler, C. A., Vaughan, M. A., and Omar, A. H.: CALIOP algorithm theoretical basis document. Part 1: CALIOP instrument , and algorithms overview, Tech. rep., NASA, 2006.

Wood, R.: Diurnal cycle of liquid water path over the subtropical and tropical oceans, Geophys. Res. Lett., 29, doi:10.1029/2002GL015371, 2002.

Yang, G.-Y. and Slingo, J.: The Diurnal Cycle in the Tropics, Mon. Weather Rev., 129, 784–801, 2001.

Yang, S., Kuo, K.-S., and Smith, E. A.: Persistent Nature of Secondary Diurnal Modes of Precipitation over Oceanic and Continental Regimes, J. Climate, 21, 4115–4131, 2008.

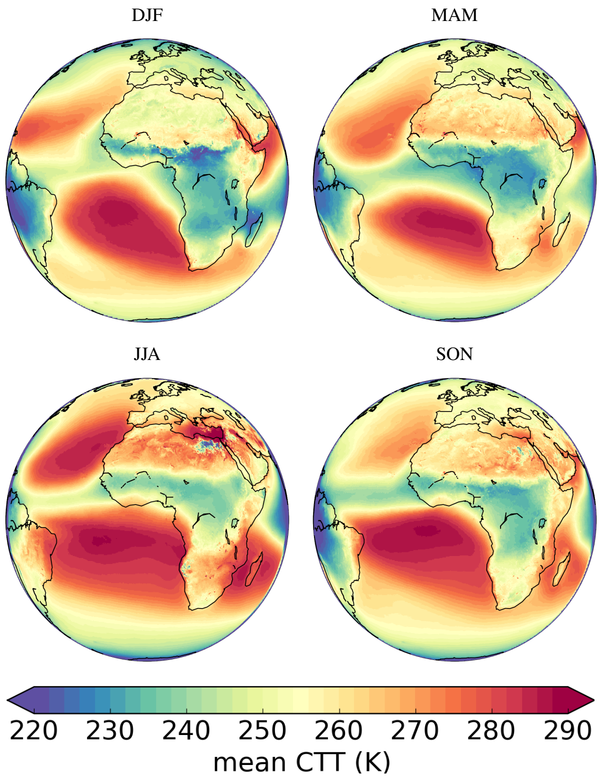

Figure 1. Seasonal mean SEVIRI cloud top temperatures for the period 2005-2015.

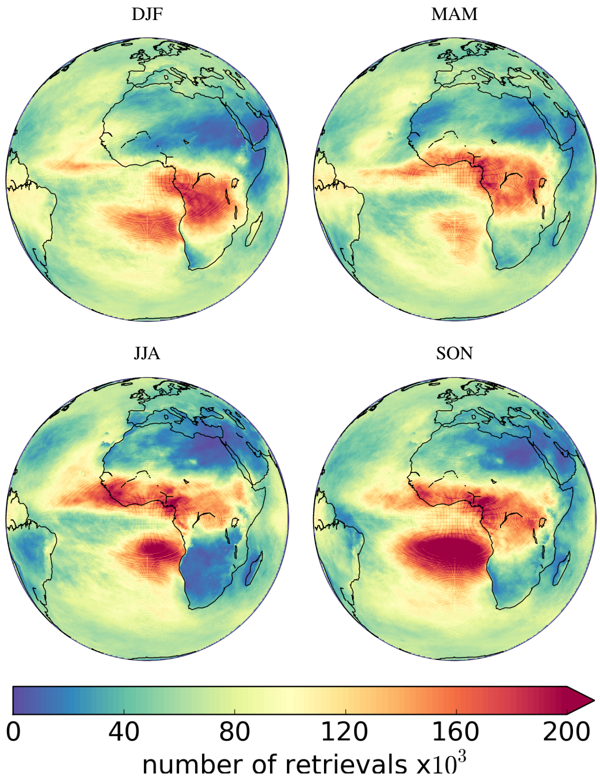

Figure 2. Spatial distributions of the total number of SEVIRI cloud top temperature retrievals available from the CLAAS-2 dataset during the period 2005-2015. Values are shown for each season.

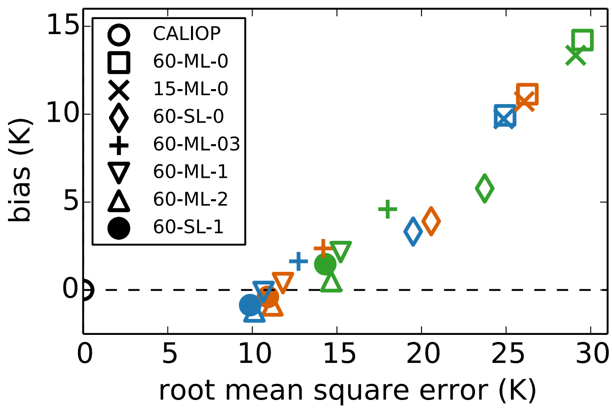

Figure 3. Bias (SEVIRI minus CALIOP CTT) versus root mean square error of SEVIRI cloud top temperature retrievals. Symbols refer to the different sets of collocation criteria, defined in Table 1. Green symbols show retrievals over land, blue over ocean and orange over both. The 'CALIOP' point on the left-hand side indicates where a retrieval which perfectly reproduced the CALIOP observations would be located.

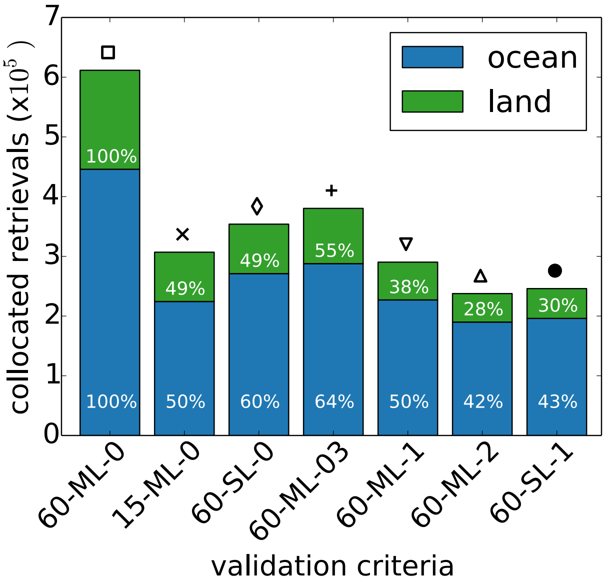

Figure 4. Number of SEVIRI and CALIOP retrievals collocated in 2007 for each of the collocation criteria defined in Table 1. Collocation criteria are identified by both text abbreviation and symbol. Colours show the division between land (green) and ocean (blue) retrievals. Percentages show what fraction of the total number of available retrievals are processed for each set of collocation criteria.

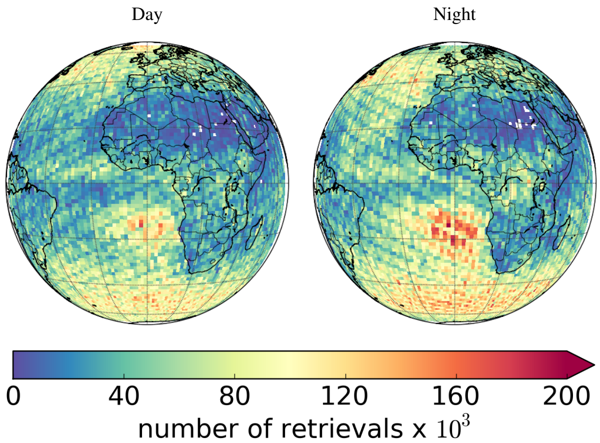

Figure 5. Spatial distribution of the number of collocated SEVIRI and CALIOP retrievals in 2007, shown separately for daytime and nighttime conditions.

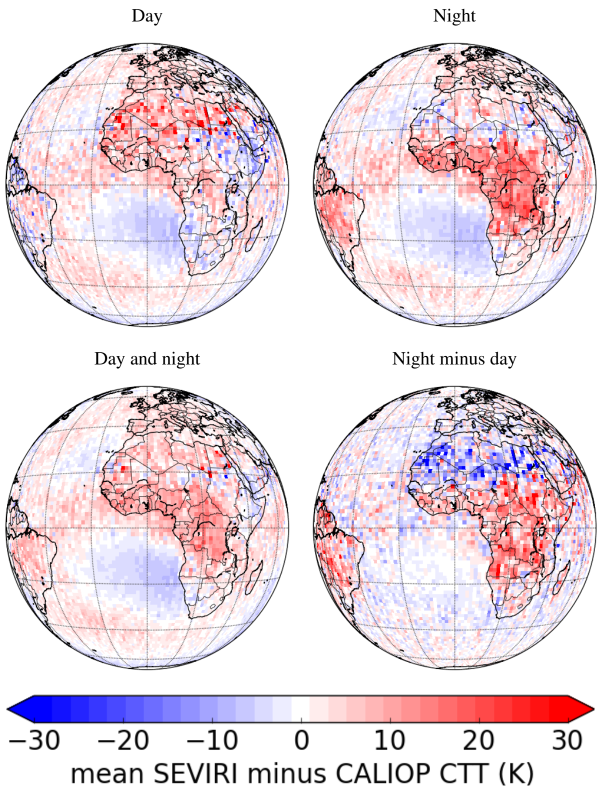

Figure 6. Spatial distribution of the bias in SEVIRI cloud top temperature retrievals during 2007. Biases are shown for the day (13:30 LST CALIOP overpass), night (01:30 LST CALIOP overpass), mean of both day and night biases, and for the difference (night minus day) between night and daytime biases.

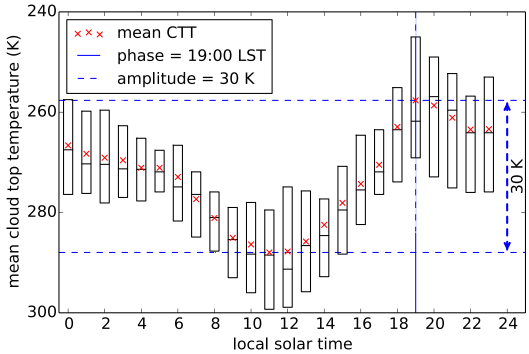

Figure 7. Seasonal mean climatological (September, October, November, 2005 - 20015) diurnal cycle of cloud top temperature at 3.1°S, 16.4°E, (western Democratic Republic of the Congo, see cross in Fig. 10). Box plots are used to show the quartiles the shapes of the distribution for each mean point. The amplitude of the diurnal cycle (defined as minimum minus maximum CTT) and the phase (defined as the local solar time (LST) of minimum CTT) are also illustrated.

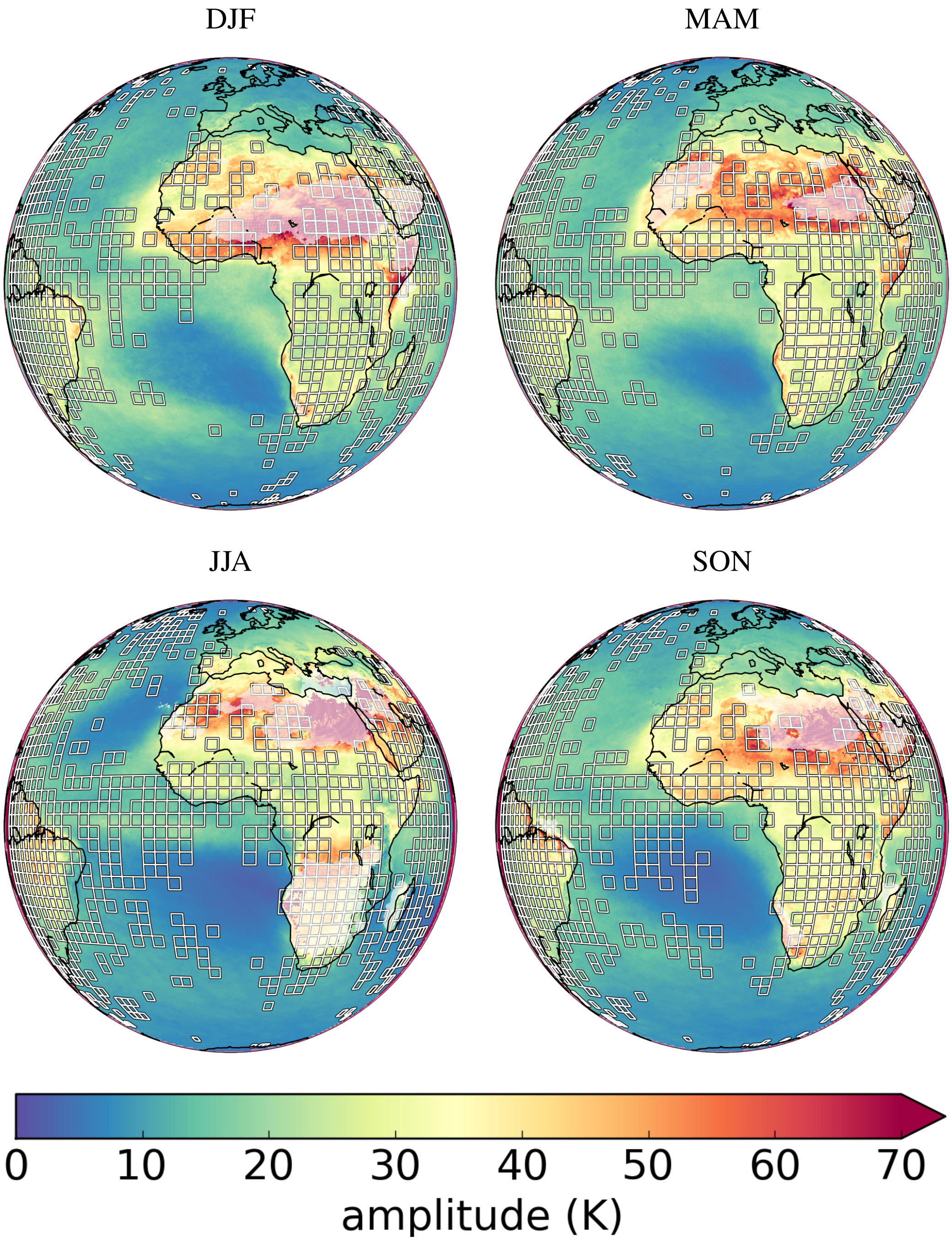

Figure 8. Seasonal mean amplitude of the diurnal cycle in SEVIRI cloud top temperature for the period 2005-2015. The white overlay indicates regions where retrievals are available for fewer than 15% of the processed timesteps. Squares indicate regions where the ratio of the amplitude of the diurnal cycle in CTT to the diurnal variability in CTT retrieval bias is less than 5.

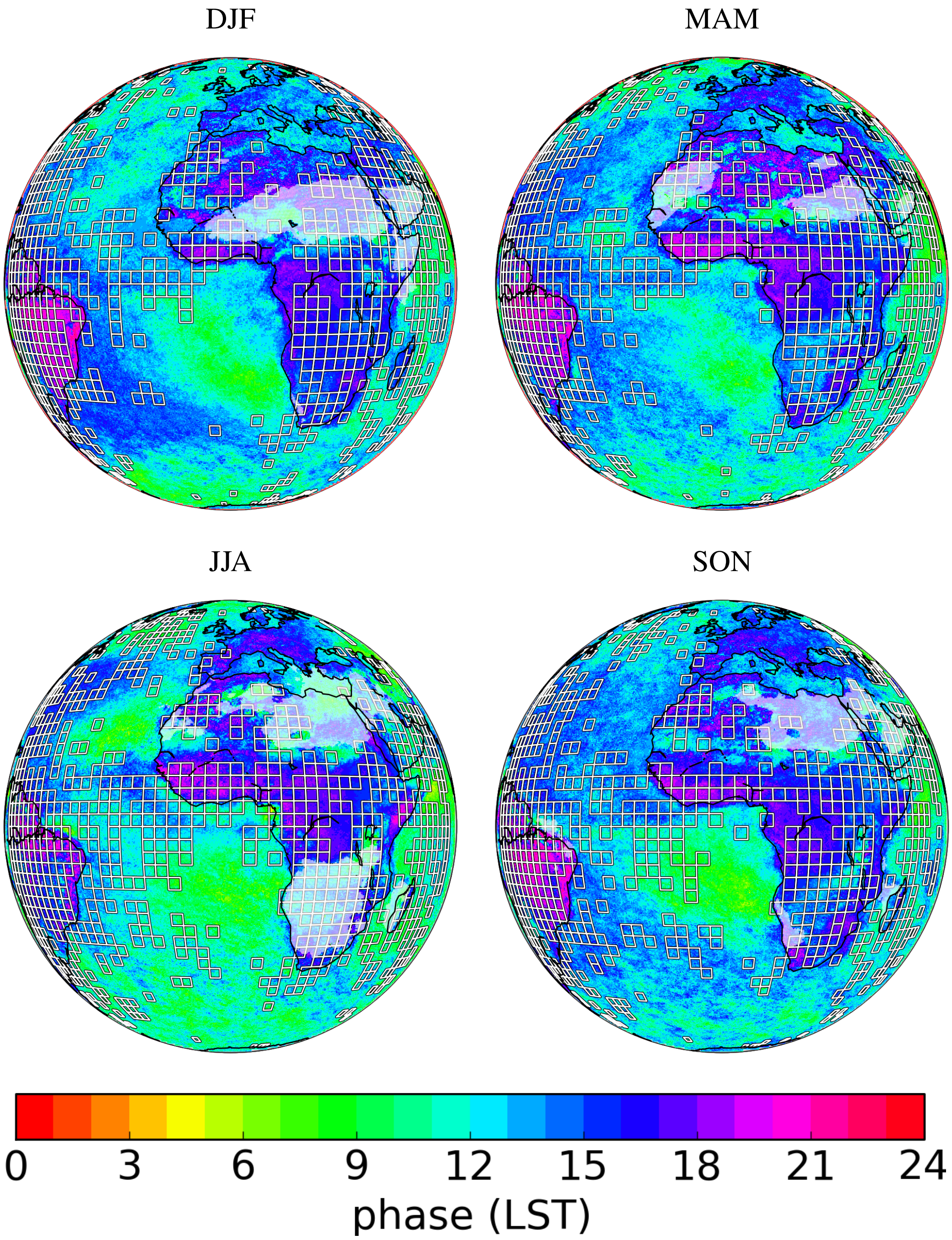

Figure 9. Seasonal mean phase (local solar time of minimum climatological mean CTT) of the diurnal cycle in cloud top temperature for the period 2005-2015. The white overlay indicates regions where retrievals are available for fewer than 15% of the processed timesteps. Squares indicate regions where the ratio of the amplitude of the diurnal cycle in CTT to the diurnal variability in CTT retrieval bias is less than 5.

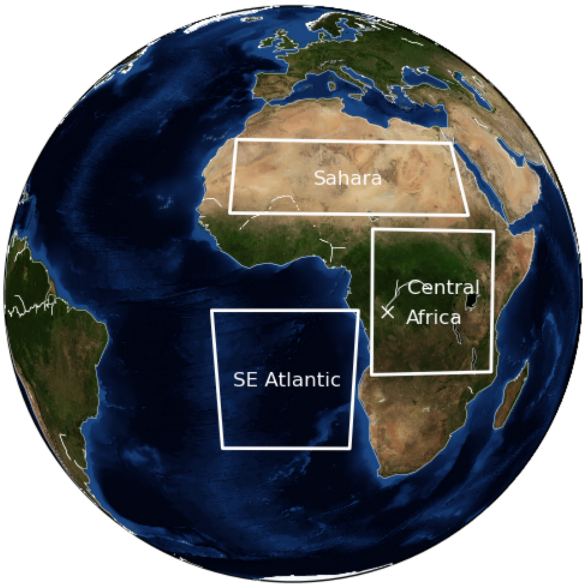

Figure 10. Map of the area observed by SEVIRI showing true-colour surfaces, major rivers and lakes. Labelled boxes show the locations of the regions used in Fig. 11. The white cross shows the location of the data used in Fig. 7.

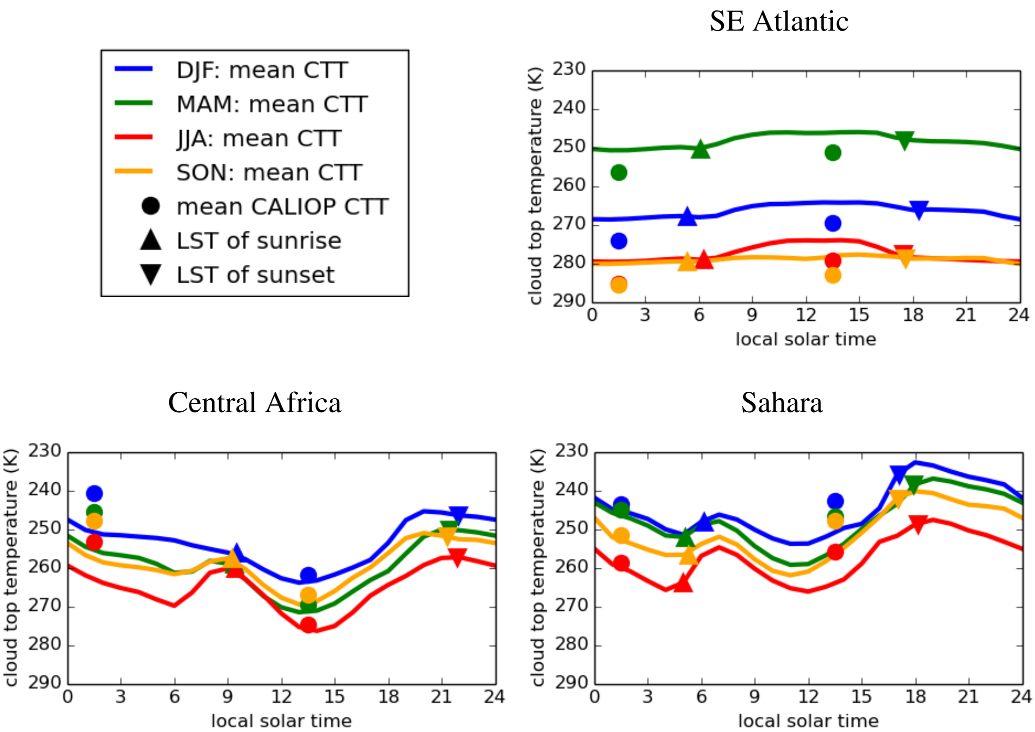

Figure 11. Seasonal mean diurnal cycles of SEVIRI CTT calculated over the period 2005-2015 (solid lines) and mean 2007 SEVIRI minus CALIOP retrieval bias (vertical distance between coloured circles and corresponding coloured lines). Biases are shown at the mean LST for the day and nighttime CALIOP overpasses. Also shown are the local solar times of sunrise (triangles) and sunset (nablas) for each region and season.

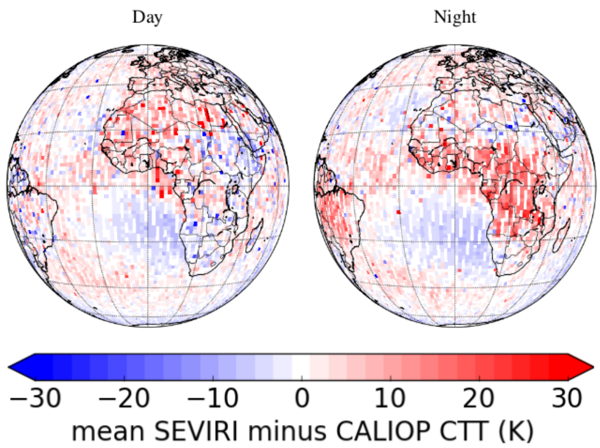

Figure A.1. Spatial distribution of the bias in SEVIRI cloud top temperature retrievals during 2007, using a 15 minute collocation window. Biases are shown separately for the day (13:30 LST CALIOP overpass) and night (01:30 LST CALIOP overpass).

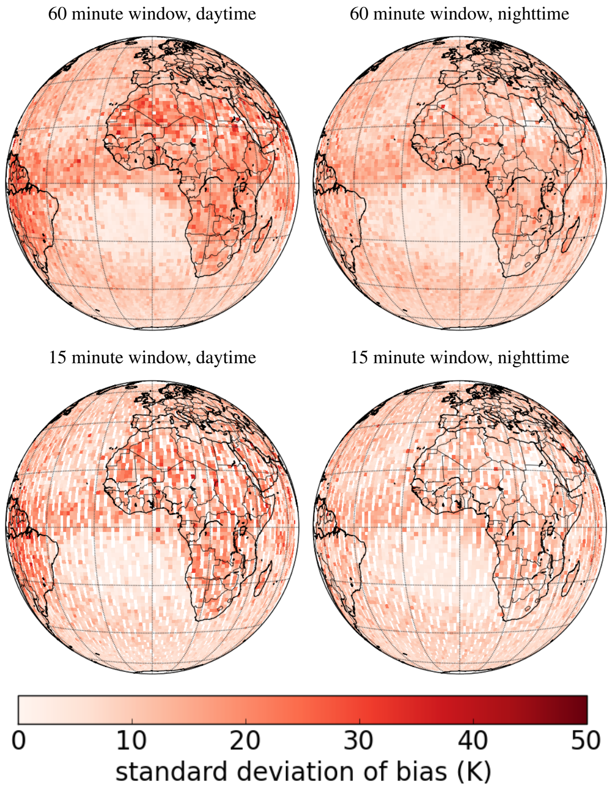

Figure A.2. Spatial distribution of the standard deviation of the biases in SEVIRI cloud top temperature retrievals during 2007. Data are shown for both the day (13:30 LST CALIOP overpass) and nighttime (01:30 LST CALIOP overpass), using a 60 minute collocation window and a 15 minute collocation window.

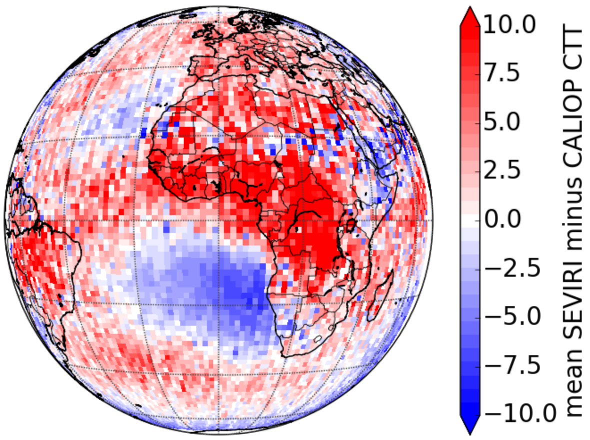

Figure B.1. Spatial distribution of mean daytime (13:30 LST CALIOP overpass) and nighttime (01:30 LST CALIOP overpass) bias in SEVIRI cloud top temperature retrievals during 2007. This plot is the same as the 'Day and night' plot in Fig. 6, but plotted using the same scale as Fig. 6-15 in Benas et al. (2016) for ease of comparison.

Table 1. Descriptions of the seven sets of collocation criteria to be evaluated, the abbreviations by which they are referenced in the text and the symbols by which they are referenced in plots.

Abbreviation	Symbol	Collocation window (mins)	Layers included	COD threshold
60-ML-0	$\square$	60	multi	none
15-ML-0	$\times$	15	multi	none
60-SL-0	$\Diamond$	60	single	none
60-ML-03	$+$	60	multi	> 0.3
60-ML-1	$\triangledown$	60	multi	> 1.0
60-ML-2	$\vartriangle$	60	multi	> 2.0
60-SL-1	$\bullet$	60	single	> 1.0

Table 2. Summary of evaluation results for the 60-ML-1 collocation criteria, showing: number of collocated SEVIRI retrievals, mean bias (SEVIRI minus CALIOP CTT) and standard deviation of the bias. Statistics are reported separately for day, night, land and ocean retrievals.

Surface type	Time of day	Number of retrievals (10^6)	Mean bias (K)	Standard deviation of bias (K)
land and ocean	day and night	2.79	0.44	11.7
	day	1.34	0.05	12.2
	night	1.45	0.80	11.2
land	day and night	0.63	2.38	14.9
	day	0.32	0.72	16.0
	night	0.31	4.11	13.5
ocean	day and night	2.16	-0.12	10.5
	day	1.02	-0.16	10.7
	night	1.14	-0.10	10.4