

# *An improved algorithm for generating global window brightness temperatures from multiple satellite infrared imagery*

Article

Published Version

Hodges, K. I., Chappell, D. W., Robinson, G. J. and Yang, G. (2000) An improved algorithm for generating global window brightness temperatures from multiple satellite infrared imagery. *Journal of Atmospheric and Oceanic Technology*, 17 (10). pp. 1296-1312. ISSN 1520-0426 doi: [https://doi.org/10.1175/1520-0426\(2000\)017<1296:AIAFGG>2.0.CO;2](https://doi.org/10.1175/1520-0426(2000)017<1296:AIAFGG>2.0.CO;2) Available at <https://centaur.reading.ac.uk/169/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: [http://dx.doi.org/DOI:10.1175/1520-0426\(2000\)017<1296:AIAFGG>2.0.CO;2](http://dx.doi.org/DOI:10.1175/1520-0426(2000)017<1296:AIAFGG>2.0.CO;2)

Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

## An Improved Algorithm for Generating Global Window Brightness Temperatures from Multiple Satellite Infrared Imagery

K. I. HODGES, D. W. CHAPPELL, AND G. J. ROBINSON

*Environmental Systems Science Center, University of Reading, Reading, United Kingdom*

G. YANG

*Centre for Global Atmospheric Modeling, University of Reading, Reading, United Kingdom*

(Manuscript received 7 June 1999, in final form 7 January 2000)

### ABSTRACT

An improved algorithm for the generation of gridded window brightness temperatures is presented. The primary data source is the International Satellite Cloud Climatology Project, level B3 data, covering the period from July 1983 to the present. The algorithm takes window brightness, temperatures from multiple satellites, both geostationary and polar orbiting, which have already been navigated and normalized radiometrically to the National Oceanic and Atmospheric Administration's Advanced Very High Resolution Radiometer, and generates 3-hourly global images on a  $0.5^\circ$  by  $0.5^\circ$  latitude–longitude grid. The gridding uses a hierarchical scheme based on spherical kernel estimators. As part of the gridding procedure, the geostationary data are corrected for limb effects using a simple empirical correction to the radiances, from which the corrected temperatures are computed. This is in addition to the application of satellite zenith angle weighting to downweight limb pixels in preference to nearer-nadir pixels. The polar orbiter data are windowed on the target time with temporal weighting to account for the noncontemporaneous nature of the data. Large regions of missing data are interpolated from adjacent processed images using a form of motion compensated interpolation based on the estimation of motion vectors using an hierarchical block matching scheme. Examples are shown of the various stages in the process. Also shown are examples of the usefulness of this type of data in GCM validation.

### 1. Introduction

The aim of this paper is to report a new algorithm for the generation of global window brightness temperatures (WBTs) from multiple infrared (IR) satellite imagery. This form of data is useful both for the validation of general circulation models (GCMs) and performing quantitative climatological studies of convection over wide areas, for example the whole of the Tropics. Thus, these type of data can be very useful for exploring the properties, in terms of the distribution and variability, of organized convective systems from the mesoscale to the synoptic scale over seasonal to decadal periods. This is provided that a sufficiently long time series of this type of data can be generated at suitable spatial and temporal resolution. Of course we can perform regional studies using individual satellites, but generating global data at current forecast GCM resolutions allows comparisons to be made with the vari-

ability predicted by GCMs as well as enabling climatological studies to be performed over wide areas.

Recent developments in operational forecasting have aimed to directly assimilate satellite radiance measurements from atmospheric sounders. This arguably might form the basis of directly comparing radiances from models and satellites via a forward model, taking into account viewing angle dependencies; this would have to be done on a per satellite basis. For example, the current European Centre for Medium-Range Weather Forecasts (ECMWF) code RTATOV (Saunders et al. 1999) is only configured to use the National Oceanic and Atmospheric Administration (NOAA) polar orbiting TIROS Operational Vertical Sounder instrument. However, such comparisons may depend on the sophistication of the underlying radiative transfer model used to perform the forward modeling and the interpolation required in space and time to the satellite observation position. For example, the RTATOV code uses a two-stream model with a plane parallel atmosphere with scattering effects neglected. While this may be a reasonable approximation for clear-sky or stratiform-cloud regions, what the satellite actually measures can in practice be the result of more complex situations in

---

*Corresponding author address:* Dr. K. I. Hodges, ESSC Harry Pitt Building, 3 Earley Gate, Whitenights, University of Reading, P.O. Box 238, Reading RG6 6AL, United Kingdom.  
E-mail: kih@mail.herc-essc.ac.uk

which scattering is important and for which a full 3D radiative transfer code is more appropriate. While this approach is feasible for validating models by comparison of absolute radiance values, a computationally simpler means of validating models can be achieved by comparing the predicted variability with the observed variability on a range of spatial and temporal scales by using the outgoing longwave radiation as a surrogate for WBT. Since most GCMs output this variable routinely there is no need to invoke the additional computational overhead of computing and comparing radiances, which are dependent on satellite and possibly the form of the forward model. Thus, meaningful comparisons can be made between models and observations on this basis and provides the motivation for the generation of global WBT.

Composite satellite imagery is routinely generated by weather forecast groups at a number of operational centers to aid in weather forecasting and, for example, to improve the model forecasts by "bogusing" systems in the forecast initialization process. The process of generating composites generally takes the form of "stitching" multiple IR satellite images together. The composite data are rarely referenced to a grid, which is useful for the types of studies previously mentioned, and even when they are, the methodology is rather crude, often just averaging pixel WBT over grid boxes, whose size may vary with latitude (latitude-dependent grid boxes, equal-area grids, are often used to take account of the areal distortion when projecting spherical data onto a cylindrical projection). Artifacts of the process are often apparent in these composite images, usually appearing as discontinuities in the image intensities. These discontinuities can be caused by prior or non-existent intercalibration between different satellite series and/or limb-darkening effects, as well as by cloud shearing effects caused by cloud motion when the images are not from the same synoptic times.

An alternative to generating the composites, and then gridding the WBT, is to take the calibrated data from the required individual satellites and interpolate/smooth the data directly onto the grid. This has the benefit that data from the individual satellites can be corrected for calibration, navigation, and limb effects before the composite/gridding procedure is performed. Additionally, these corrections produce a better set of gridded data for determining the motion of mesoscale/synoptic systems in tracking studies.

Of course, to generate global WBT at a reasonable temporal sampling rate, that is, 1–3 h, there has to be sufficient satellite coverage over such a time period. This is obviously not a problem currently since there is a profusion of weather satellites operational for which operational centers have access, and which provide nearly global coverage every 30 min. This is an enormous amount of data at full weather satellite IR resolutions (~4 km for geostationary and polar orbiting satellites). However, these data are not generally or systematically

available outside the operational centers. It is also at a significantly higher resolution than current GCM resolutions. In order to generate a historical global WBT dataset for both climatological studies and GCM validation, similar data coverage is required over an extended period of time. Fortunately, the International Satellite Cloud Climatology Project (ISCCP) (Rossow et al. 1997) has been archiving and processing the available weather satellite imagery using data available from July 1983 to the present (excluding the Indian INSAT satellite, apart for a brief period from April 1988 to March 1989, with the result that there is poor synoptic data coverage in the Indian Ocean region). To reduce the amount of data for processing and archiving, the raw imagery is subsampled to a nominal spatial resolution of 30 km and temporal resolution of 3 h to produce the level-B3 imagery. These data are processed into cloud products by ISCCP, typically on an equal-area cylindrical grid of resolution  $\sim 2.5^\circ$  at 3-hourly intervals. While these data are useful for large scale studies, they are of limited use for observing and studying smaller spatial scales, for example, mesoscale phenomena. The trend is for climate models to be run at ever higher spatial resolutions, depending on the simulation period. For example, the ECMWF reanalyses are available at resolutions  $\sim 1.0^\circ$ , and free running models are not far behind. It therefore makes sense to provide data at a resolution that will be useful for validation for the foreseeable future. The fact that the WBT data are not on the identically same grid as the model data is not a problem if we are concerned with studies of variability, as opposed to comparison of absolute values.

A first attempt at generating WBT data from the ISCCP B3 archive was performed by Tanaka et al. (1991), and was used for studies of the diurnal cycle throughout the Tropics by Salby et al. (1991), on a limited period of one year from July 1983 to June 1984. The usefulness of this type of data for studying tropical mesoscale activity was illustrated by Hodges (1998), using Salby's data. In order to conduct such analyses without having to worry about missing data the production of a dataset with few data voids is important. However, some missing data is inevitable, because of failed satellites or corrupted input data, so the best that can be done is to minimize the occurrence of data voids. The grid chosen by Tanaka et al. was  $512 \times 512$  (equivalent to  $0.35^\circ$  latitude  $\times$   $0.7^\circ$  longitude), driven by their desire to perform efficient spatial fast Fourier transform analyses. Their algorithm essentially averaged pixel values for the contributing satellites over grid boxes, with the pixel temperature values weighted by the cosine of the satellite zenith angle; that is, they downweight the limb pixels in comparison with near-nadir pixels. This anticipates the problem of limb darkening and changes in pixel foot print size but does not actually correct the IR values. The data taken for this spatial sampling was constrained to be within 1.5 h of the target synoptic

time. Small voids in the grid were filled by spatially averaging over surrounding nonvoid grid values at two levels; larger voids in the grid were filled by simple linear temporal interpolation using gridded data from adjacent synoptic times (equivalent to 6-h separation). This is a fairly straightforward and simple algorithm for generating the global WBT data and will be used as a reference for the improved algorithm described here.

In this paper an improved algorithm for generating global WBT from the ISCCP B3 archive is described. The algorithm development forms part of the Cloud Archive User Service (CLAUS) project to generate a long-term historical archive of global WBT at resolutions suitable for both climate studies and GCM validation, and to use this data to perform these types of studies. The algorithm essentially follows the approach used by Tanaka et al. (1991), but differs in the techniques that have been used for the spatial gridding, limb-darkening correction, and temporal interpolation. The grid chosen for the study has a spacing of  $0.5^\circ$  latitude  $\times$   $0.5^\circ$  longitude ( $360 \times 720$ ) on a Plate Carrée latitude–longitude projection. The choice of this grid is partially driven by the resolution of the available data and also by the requirements of GCM validation. However, the techniques can be applied just as well to higher-resolution data if required, and we hope to illustrate this in the future with full-resolution imagery. Also, the algorithm can be applied to other data, for example, the water vapor channel, although this channel was not available on all the satellite series until recently. The final product data are supplemented by quality data, which will also be described.

The paper continues with a description of the ISCCP B3 data and then continues with a step-by-step approach to the algorithm methodology, followed by examples of the applicability of the data to GCM validation.

## 2. Description of the ISCCP B3 data

The ISCCP B3 data are generated by ISCCP from the raw full-resolution imagery by a process of subsampling, calibration, and navigation (Rossow et al. 1997). The B3 data are stored as 8-bit values (together with the relevant calibration tables). This can limit some applications due to a lack of accuracy—for example, identifying changes associated with the sea surface or low level stratocumulus. The contributing satellites consist of the geostationary satellites METEOSAT, GOES-E, GOES-W, GMS, and for a limited period, INSAT (this is not used by CLAUS because of the limited availability of the data, although this may be used in subsequent reprocessing), while the polar orbiting satellites are NOAA-AM and NOAA-PM. The full-resolution imagery is sampled in space and time for the geostationary satellites to produce imagery every 3 h at a nominal pixel spacing of 30 km. There is the issue of bias of the raw data, due to how each satellite instrument samples the Earth, as well as bias caused by the sampling

TABLE 1. Summary information of contributing satellites. Note that geostationary satellites were sometimes moved to cover for the failure of other satellites; e.g., *METEOSAT-3* covered for the failure of *GOES-6*.

Satellite	Agency	Series	IR band	Subsatellite longitude
METEOSAT	ESA	3, 4, 5	10.5–12.5 $\mu\text{m}$	$0^\circ$
GMS	JMA	1, 2, 3, 4	10.5–12.5 $\mu\text{m}$	$140^\circ\text{E}$
GOES-E	NOAA	5, 7	10.3–11.4 $\mu\text{m}$ 10.5–12.5 $\mu\text{m}$	$75^\circ\text{W}$
GOES-W	NOAA	6, 7	10.2–11.2 $\mu\text{m}$ 10.5–12.5 $\mu\text{m}$	$135^\circ\text{W}$
NOAA-AM	NOAA	8, 10, 12	10.2–11.2 $\mu\text{m}$	—
NOAA-PM	NOAA	7, 9, 11, 14	10.3–11.3 $\mu\text{m}$	—

scheme used to produce the B3 data, which still needs to be investigated. The effect of the spatial sampling, in contrast to spatial averaging to produce the B3 data also still needs to be explored. For the polar orbiting satellites spatial sampling is performed to produce the same nominal pixel spacing as for the geostationary satellites. Each image is calibrated according to the available spectral channels; the IR channel is used in the CLAUS project, and normalized to the reference NOAA polar orbiter satellites for the channels common to the geostationary satellites (see Table 1 for the relevant channels). Although each image is navigated and corrected for navigation errors, residual problems with satellite intercalibration, image navigation, and bad scan lines are occasionally observed in the B3 data.

Each B3 image has available for each pixel the following: its geolocated (latitude–longitude) position; the satellite zenith angle (the angle between the satellite view direction of a point on the earth and the nadir view direction of the same point); a physical value for each spectral channel (an equivalent blackbody WBT for the IR channels); and flags for scan line quality, on/off planet pixels, and a land/water mask. Instead of using the calibration tables that come with each B3 image based on normalization to the Advanced Very High Resolution Radiometer (AVHRR) on the *NOAA-7* polar orbiting satellite, the newer calibration tables based on normalization to the AVHRR on *NOAA-9* are used (Brest et al. 1997). These new tables, available as version 1 of the ISCCP BT dataset, have been corrected by ISCCP for artifacts found in the older tables. In addition to their use for converting the IR channels of the various satellites to brightness temperatures, they are used to generate polynomial fits between radiance and temperature for use in the limb-darkening correction to be described later.

The ISCCP B3 data consists of all available spectral channels. For the CLAUS project the IR channels listed in Table 1 are extracted from the B3 data and stored in a simple packed format for use by the CLAUS program. The data extraction process is supplemented by quality controls to identify corrupt scan lines and bad pixel



TABLE 2. Statistics for maximum scan line length for *GOES-6* for Jan 1984 as an example of detecting squashed images.

Scan length	392	402	403	453
No. of images	1	226	9	12

values. Bad scan lines are found using two methods: a run length search of scan lines to identify regions that have the identical pixel value and a check of the mean and variance along the scan line, so that any scan line that has a high or low value for the mean and a low variance is also excluded. This approach appears to remove nearly all the bad scan lines that are erroneously flagged as good by ISCCP. Occasionally, images are so badly corrupted that they are wholly unusable: these are identified using the percentage of bad scan lines value obtained from the B3 data. Another form of image corruption appears only for the geostationary satellites, in particular the GOES series. This manifests itself as an apparent bad image navigation. Examination of the images reveals that for some images the on-planet disk is apparently squashed in the vertical direction. Other types of bad navigation appear as vertical translations of a nonsquashed disk. When these data are gridded, this results in large vertical translations in what should be fixed features, such as continents. The squashed images are identified by examining the value for the maximum on-planet scan line length; this is larger than for good images, so that images are excluded if they exceed a scanline length threshold. For example, the *GOES-6* satellite for January 1984 had the maximum scan line statistics shown in Table 2. The single image with length 392, is a partial image so should be accepted; the majority of images have length of 402, but the images with lengths of 453 are bad images and should be excluded. Thus, an acceptance threshold of 20 from the value of 402 accepts the images with length 392, 402, and 403 and rejects those with value 453. The image translations are harder to detect automatically since they require the identification of surface features and the comparison of their ISCCP navigated position with their actual position. A manual quality control is used to identify them, but this is operator dependent. Ideally, these otherwise good images should be renavigated, which may possibly be done in subsequent algorithm improvements. The pixel on/off-planet flag is used to exclude all off-planet pixels. Thus, the resulting reformatted data should contain only data from good images/scan lines and only for on-planet pixels for the IR channel. However, other types of bad data occasionally slip through, in the form of bad images/scan lines that are not flagged and are not easily automatically detected as well as badly navigated images; these then require manual intervention.

### 3. Algorithm overview

The algorithm proceeds by first initializing a grid and a calibration table to convert temperatures to 8-bit in-

tegers for archiving in the chosen image format. The grid is chosen by specifying the number of grid points required in longitude and latitude for the Plate Carrée projection (this is the only projection used at present, although any grid can be used). This was chosen as the most convenient for comparing with global GCM data. The method of performing the gridding is independent of the grid chosen, since a spherical nonparametric regression technique is used (Hodges 1996). The calibration table is used to scale the gridded WBT values so that the data can be stored in image format [the Portable Gray Map (PGM) format is used]. Valid values are scaled to the range (1–255), with the value 0 being used to indicate missing data values. The PGM format is also used for the quality data, with the quality values being packed into bit fields for each quality byte. Each grid point thus has a byte in the WBT image and a byte in each of the quality images; there are two quality images per WBT image.

The algorithm continues by looping over the chosen time range, performing the spatial regression at each synoptic time at three smoothing levels, incorporating corrections for limb darkening for the geostationary satellites and temporal weighting for the polar orbiting satellites to account for the asynoptic nature of the polar orbiter data. Three smoothing levels were chosen as a compromise between processing speed and the requirement of producing as complete an image as possible from the spatial regression. For each synoptic time the required B3-reformatted data within 1.5 h of the target time are identified from the satellite header summaries obtained at the data extraction/reformat stage. This usually results in the identification of a single image for each geostationary satellite, unless more than one series is operational at the same time, for example, *METEOSAT-3* and *METEOSAT-4*, with *METEOSAT-3* being used as cover for the failed *GOES-6*. For the polar orbiting satellites there may be several images that overlap the temporal window.

The spatial regression is performed by looping over the data for each contributing satellite with only the data that fall within the  $\pm 1.5$ -h time window being used. This will include all of the data in a geostationary satellite image that coincides with the target time, but may only be a partial polar orbiter image, so only the data for scan line times that fall within the time window are used. For each satellite the algorithm loops over each of the selected B3 pixels, checking to ensure that it has a valid value and that it falls within a satellite zenith angle mask. The mask is used for the geostationary satellites to exclude the very extreme limb pixels which can be affected by variations in satellite navigation. Each B3 pixel is also corrected for limb-darkening effects if the current satellite is a geostationary satellite. No limb-darkening correction is currently performed for the polar orbiter data as the swath is relatively narrow compared with the geostationary satellite field of view. Also no correction is made for changes in pixel footprint

size with the change in satellite zenith angle. Following the application of these checks and corrections to the B3 pixel the contribution to the grid estimates is computed for each smoothing level and the estimates for the regression and data density are updated. A crude approach to computing the contributions to the grid for each pixel would be to check every grid point for each pixel, a very time-consuming process. However, since the weighting kernels we are using are spatially local in nature a much more efficient scheme has been developed which limits the amount of checking required to just the grid points local to the pixel position on the sphere.

Following the completion of the regression estimation, there are estimates available at three smoothing levels on the chosen grid. The estimate for the smoothing level, which has the highest data density estimate is the one taken as the final gridpoint value. Ideally, a fully adaptive regression estimation should be used which varies the smoothing with the local data density. However, for this amount of data, this approach would be computationally prohibitive (a pilot density estimate is required; see Hodges 1996) as well as invalidating the efficient search method mentioned. Finally the gridded data are scaled to image format and written to file. The quality data are also written to file.

The final stage is to take the images obtained from the spatial regression and fill the remaining data voids using temporal interpolation between adjacent images. This is performed using a motion compensated interpolation method based on a hierarchical determination of the motion vectors, as opposed to the simple linear interpolation method used by Tanaka et al. (1991). However, both methods have been implemented for comparison purposes. The motion compensated interpolation is found to preserve the sharpness of features much better than the simple linear interpolation although on average the differences are not very large between the two methods at the resolution being used. This is in part a consequence of the resolutions in space and time that the ISCCP B3 data provides. The final gridded data are written to file along with the updated quality information.

An important though difficult question to answer is what accuracy we require for validation studies. As the ISCCP B3 data are only 8 bit, this is already a limitation on the available accuracy and appears to preclude the study of the ocean surface or low-level stratocumulus cloud regions in studies so far conducted using data produced by the algorithm described here. However, examination of diurnal and semidiurnal activity and H<sub>y</sub>ashi spectra do not indicate any problems with the data produced by the algorithm, at least in the Tropics for cloud systems and land, so we conclude the algorithm is performing adequately for the required purpose of model validation at least. The usefulness of the data produced by the algorithm in other studies will only

become apparent with time and use in a range of applications.

#### 4. Spatial regression

The spatial regression is required to combine the ISCCP B3 data distributed along scan lines for multiple satellites onto the chosen grid. This is done using the reformatted data and performing a nonparametric regression onto the chosen grid. Since the ISCCP B3 data can be considered as data distributed on a unit sphere, that is, each ISCCP B3 pixel has a latitude–longitude coordinate, the method for gridding the B3 data is to use a spherical nonparametric regression estimator based on local spherical kernels (Hodges 1996). This requires the grid to be converted to three-dimensional Cartesian coordinates representing unit vectors pointing to the grid points on the unit sphere. The kernels essentially control the region of influence of each data point (B3 pixel position) and their contribution to an estimation point (grid point) via a function of the angular separation between a data point and a grid point. This approach is independent of the grid chosen, that is, any distribution of grid points can be selected and a regression estimate performed. Also, there is no bias in the estimate (although there may be in the source data) as the estimation is performed in spherical geometry, which obviates the need to introduce some form of correction, that is, latitudinal-dependent area weighting. It is not apparent that Tanaka et al. (1991) took this into account in their algorithm, since they used fixed angular grid boxes. The ISCCP C1 and C2 data (280-km grid) take this into account by using an equal area grid for their cloud products, although this does not take account of the change in shape of the sampling region, for example, a rectangular latitude–longitude box at the equator is not a rectangular box at higher latitudes on the earth.

The spatial regression is performed by looping over all the identified “good” pixels for each satellite. For each pixel the contributions to the regression and data density estimates are computed for three kernel widths controlled by the spherical kernel parameters. This is performed for those grid points local to the position of the pixel (this will be described shortly). These contributions are then used to update the regression and data density estimates for the three smoothing regimes. Before the updates are computed the pixel value is first compared against a temperature acceptance mask to exclude unlikely temperature values (the mask used by Tanaka et al. 1991 has been used), although it is very rare for a pixel value to be rejected, and then corrected for limb-darkening effects. The limb-darkening effects are caused by a mixture of radiative and geometric effects due to the view angle. For example, for a broken cloud field viewed at a large zenith angle, it is more likely that cold cloud tops are viewed than at nadir where there is a greater possibility of viewing the sur-

## Satellite Zenith Angle Weighting

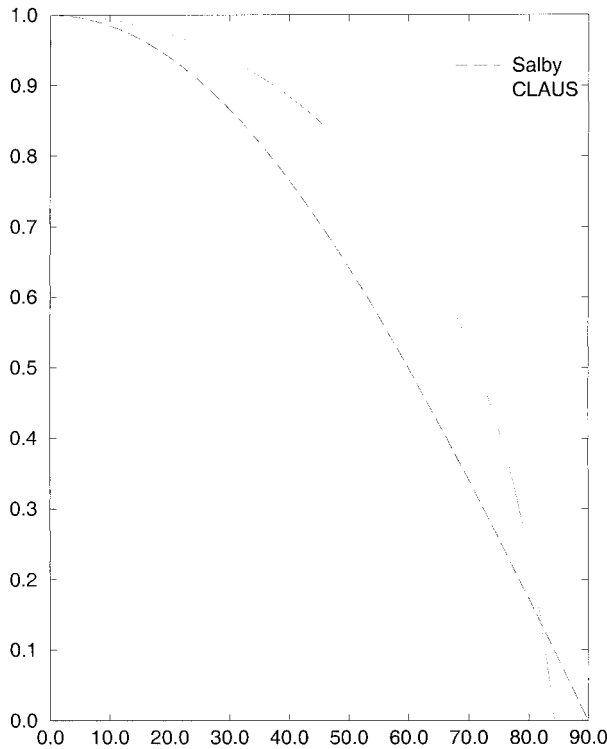


FIG. 1. CLAUS satellite zenith angle weighting function compared with  $\cos(\theta)$  weighting function used by Tanaka et al. (1991).

face. Also, for a large viewing angle the radiance at the top of the atmosphere will be decreased due to the longer pathlength through the atmosphere. This correction is required, in addition to the zenith angle weighting, to correct for artifacts which appear when polar orbiting data are adjacent to geostationary limb data. The zenith angle weighting has little effect unless the pixels from different satellites are mixed. This correction is also required for the temporal interpolation, where a missing geostationary satellite results in a data hole with apparent bright edges due to the limb effects adjacent to the interpolated hole. The correction for these effects will be discussed later.

The regression estimate is computed at grid point  $i$  as

$$\hat{T}_i = \frac{\sum_{j=1}^N T_j w_j^{(o)} w_j^{(z)} k_{i,j}^{(s)}}{\sum_{j=1}^N w_j^{(o)} w_j^{(z)} k_{i,j}^{(s)}}, \quad (1)$$

where  $\hat{T}_i$  is the temperature estimate at the  $i$ th grid point,  $T_j$  is the calibrated temperature at the  $j$ th pixel with  $N$  the number of pixels, and  $w_j^{(o)}$  is the polar orbiter temporal weighting defined as

TABLE 3. Kernel parameters for the three smoothing levels and their equivalent ABRs.

$C$	1.000038078	1.000152328	1.000342792
ABR (deg)	0.5	1.0	1.5

$$w_j^{(o)} = \begin{cases} 1, & \text{if geostationary} \\ \frac{1}{h} \left[ 1 - \left| \frac{t - t_0}{h} \right| \right], & \\ 0, & \text{if polar orbiter } |t - t_0| \leq h, h \leq 1.5 \text{ h.} \end{cases} \quad (2)$$

This progressively downweights polar orbiter pixels that are temporally farther away from the target synoptic time  $t_0$  and thus provides tapering to the polar orbiter data.

The satellite zenith angle weighting  $w_j^{(z)}$  is given by the function

$$w_j^{(z)} = \begin{cases} 1 + 0.43429 \ln[\cos(\theta_j)], & \text{if } \cos(\theta_j) \geq 0.1 \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

where  $\theta_j$  is the satellite zenith angle for the  $j$ th pixel. This weighting function progressively downweights off-nadir pixels in favor of near-nadir pixels. Tanaka et al. (1991) also used a zenith angle weighting based on only  $\cos(\theta)$ . The function we have used does not fall off as severely as  $\cos(\theta)$  for small  $\theta$  but falls off more quickly for extreme limb values of  $\theta$ . This is shown in Fig. 1 compared with  $\cos(\theta)$ . The form of the function in expression (3) is chosen to reflect the functionality often used for limb darkening correction factors (see Minnis et al. 1991 and later discussion on limb-darkening correction). The cutoff value of  $\cos(\theta) = 0.1$  is used to exclude the extreme limb pixels, which are the most affected by the extreme viewing angle and that might be viewing a mixture of on-planet and off-planet regions. Larger values can be specified on a per-geostationary-satellite basis to exclude bad pixels associated with navigation problems as discussed earlier (usually for GOES satellites in the ISCCP B3 data).

The spatial weighting function  $k_{i,j}^{(s)}$  is given by the linear spherical kernel function (Hodges 1996),

$$k_{i,j}^{(s)} = \begin{cases} \frac{C}{\pi(C-1)^2} (C\mathbf{X}_i \cdot \mathbf{X}_j - 1), & C\mathbf{X}_i \cdot \mathbf{X}_j - 1 \geq 0 \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

where  $C$  is the kernel smoothing parameter, which controls the bandwidth of the kernel; and  $\mathbf{X}_i$  and  $\mathbf{X}_j$  are the unit vectors representing the grid point and pixel position, respectively, on the sphere. Note that this function is also a probability density function since it integrates to unity on the sphere. The three smoothing levels are represented by a choice of three values for the parameter  $C$ , these are given in Table 3 with their



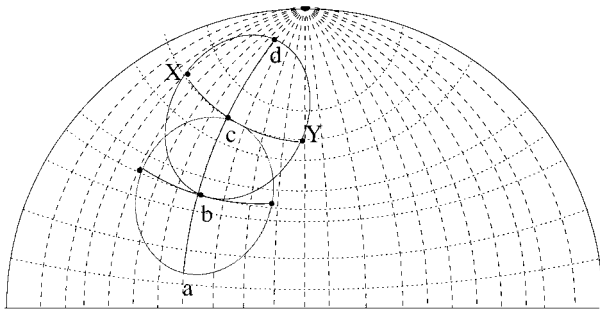


FIG. 2. Schematic of the method for determining the grid points local to a pixel position for fixed-width kernels.

equivalent arc band radius (ABR), obtained from the constraint  $\mathbf{CX}_i \cdot \mathbf{X}_j - 1 \geq 0$ .

Note these are global parameters; ideally we would like to use a fully adaptive scheme in which case the parameter  $C$  becomes a local parameter dependent on the data density (Hodges 1996), but this is very expensive.

The spatial data density is given by

$$\hat{f}_i = \frac{1}{N} \sum_{j=1}^N k_{i,j}^{(s)}, \quad (5)$$

and is used to determine which estimate to use of the three smoothing levels.

As mentioned earlier, rather than test each grid point for a contribution from the current pixel, only those grid points local to the pixel position are tested for a contribution. This is possible because constant width kernels are being used. The method for determining the grid points local to a pixel position is illustrated in Fig. 2.

For each latitude circle in the grid a longitude point is chosen, for example, point (b) in Fig. 2. The number of longitude nodes that can possibly fall within a fixed kernel centered on this point is computed by translating the kernel to point (c) in Fig. 2, and counting the longitude nodes along the latitude line XY in Fig. 2. This is done by computing the value of  $\mathbf{CX}_1 \cdot \mathbf{X}_2 - 1$ : positive values indicate that the point  $\mathbf{X}_2$  is within the kernel support for a kernel centered at  $\mathbf{X}_1$ . This ensures that all required grid points are captured for any pixel position. This process is performed once for all the lines of latitude in the grid for a hemisphere (the process is symmetric about the equator). Clearly, as we move to higher latitudes the number of longitude nodes that fall within the kernel support increases, so that near the poles all the longitude nodes are likely to fall within the kernel support. The number of latitude nodes within the kernel extent is constant and only needs computing once. This process is repeated for each kernel width used, that is, three in this case. The way these subgrid extents are used is to find the nearest grid point to the current pixel position; the subgrid extents then provide the local grid points that need testing for possible con-

tributions to the WBT estimate and density estimate at each of these grid points.

An example of the result for just the primary regression, that is, using the smallest regression kernel, is shown in Fig. 3, together with the primary data density, the result of applying the three levels of spatial regression (all scaled to 8-bit images) and its temperature calibrated form.

Notice in Fig. 3 that the result for the primary regression is a much more complete image than was obtained by Tanaka et al. (1991) for their first spatial stage. The final image, obtained using all three regression levels, is shown in Fig. 3c. This still has some regions with no data; these are subsequently filled using the temporal interpolation part of the algorithm described in section 6. Figure 3b shows the spatial probability density function (i.e., ignoring the temporal weighting of the polar orbiter data) scaled to an 8-bit image. This shows the complex way the satellite views overlap with each other and the patterns formed by the data sampling using spherical kernel estimators. Figure 3d shows the same data but calibrated to temperature.

## 5. Limb-darkening correction

The reasons for including a limb-darkening correction have already been discussed. Here the form of this correction is described. The approach of Tanaka et al. (1991) to the problem of limb effects, is to weight the pixels with their cosine of satellite zenith angle to reduce the contribution of the limb pixels relative to pixels closer to nadir. However, this only works when satellite views overlap: it has no effect when there is no overlap. The effect of limb darkening, when combining data from different satellite platforms, is often apparent when there is a polar orbiter swath adjacent to a geostationary limb. The geostationary limb shows much lower temperatures (brighter image values) than the adjacent polar orbiter, that is, there is a discontinuity in temperature. To correct for these effects for data from geostationary satellites there have been several previous studies, resulting in empirical limb correction functions for application to the radiances which correct the radiances to nadir view. However, these can never really fully take account of the geometric effects. The correction of the radiances to nadir is provided by the following expression:

$$R(0) = \frac{R(\theta)}{\lambda(\theta)}, \quad (6)$$

where  $R(\theta)$  is the radiance at a zenith angle  $\theta$ ,  $R(0)$  is the radiance at nadir, and  $\lambda(\theta)$  is the limb correction function as a function of the satellite zenith angle. Minnis et al. (1991) derived an empirical function based on limited radiative transfer calculations (Minnis and Harrison 1984) in aerosol-free atmospheres for the spectral window 10.0–12.0  $\mu\text{m}$ . The Minnis function is given by

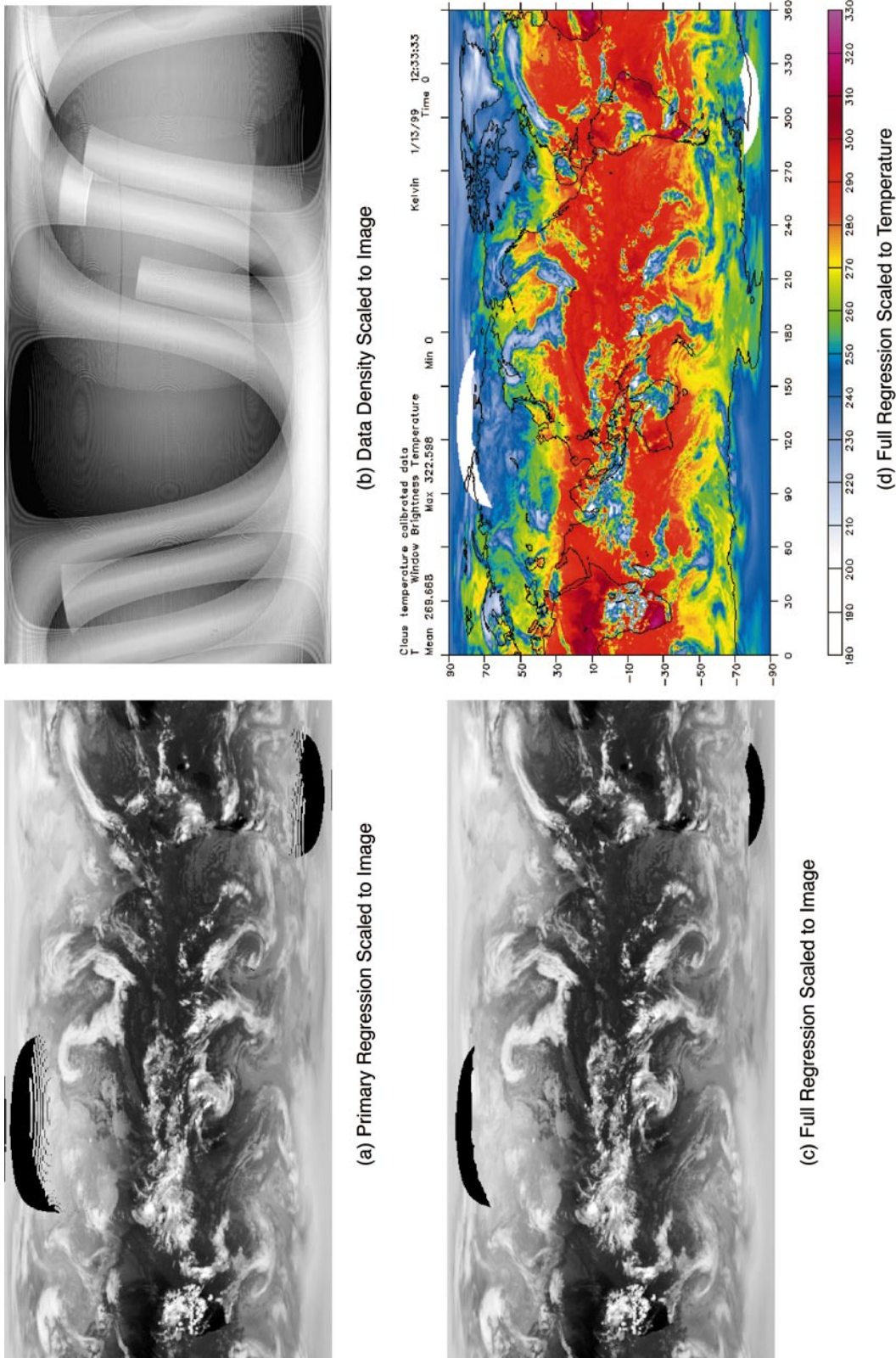


FIG. 3. Examples at 1500 UTC 12 Jan 1984 of (a) applying the nonparametric regression at the primary level, scaled to 8-bit image format (light gray levels are cold, while dark gray levels are warm); (b) the probability data density, scaled to 8-bit image format (light gray levels are low densities, while dark gray levels are high densities); (c) the full regression, scaled to 8-bit image format; and (d) the full regression converted to temperature.

$$\lambda(\theta) = \begin{cases} 1, & \text{if } \theta < 11^\circ \\ b + a \ln[\cos(\theta)] & \text{otherwise,} \end{cases} \quad (7)$$

where  $a = 0.03247$  and  $b = 1.000602$ .

However, we have found that the response of this function does not correct sufficiently for the limb-darkening effect. This can be attributed to the lack of any clouds in the radiative models from which the expression is derived, or the presence of aerosols that can result in more absorption. An alternative empirical means to determine the nature of the correction would be to compare the geostationary limb pixels with polar orbiter pixels suitably localized in time and space to that of the geostationary satellite. This will be discussed in more detail shortly.

A similar approach to determining the limb-darkening relationship was explored by Smith et al. (1994) using data from the Earth Radiation Budget Experiment (ERBE) scanning radiometer operating in along-track mode. This is a broad band instrument (5–50  $\mu\text{m}$ ) and allows the comparison to be made between limb views and nadir views of the same position on the earth with only a short time difference. The data were partitioned according to scene type, day/night and season and the limb-darkening relationship derived for each of these scenarios. However, Smith et al. (1994) were interested in radiative fluxes, so their limb correction functions are normalized differently than in expression (6). However, for most scenes, irrespective of time of day or season the relationship between limb and nadir data is remarkably similar, except for scenes with snow cover. This stable relationship between nadir and limb data is in part due to the size of the ERBE footprint, which is quite large and likely to contain some cloud and also because the broad spectral band used which will tend to smooth the differences due to cloud cover. Additionally, the averaging over a large number of observations will tend to smooth out any variation; the relationship may show more variation for a single pass. Smith et al. (1994) fit their results to a simple physical model, which resulted in a relationship dependent on special functions in the form of the exponential integral of order 4.

To explore the limb-darkening problem further, the gridded values of geostationary satellites were compared with those of the polar orbiters. The look angle of the polar orbiter imagery was restricted to  $\theta < 36^\circ$  to approximate the nadir view of scenes imaged on the limbs of the geostationary satellites. The polar orbiter scans were further restricted to times within half an hour from the geostationary image to reduce effects of temporal variability in clouds or land. To compare the derived limb-darkening functions with previous results, the gridded temperatures for both the polar orbiter and geostationary satellites were converted to a standardized radiance based on a common window profile. We chose a typical window profile for the GOES-E satellite based on the ISCCP calibration tables. We find that the results are not sensitive to the detailed form of this standardized

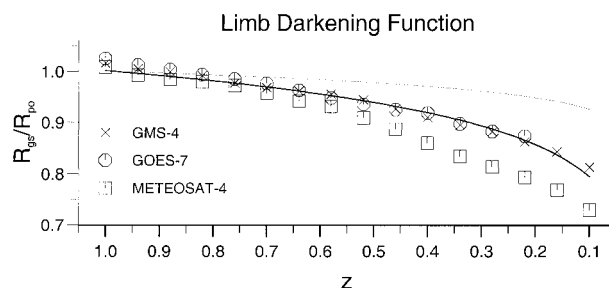


FIG. 4. Limb-darkening function estimated by the ratio of geostationary to polar orbiter standardized radiances. The lower curve represents the limb correction applied to the CLAUS data. The upper curve shows the limb-darkening function predicted by the radiative transfer calculations by Minnis et al. (1991).

window. For every grid location which has both geostationary and polar orbiter coverage (subject to the constraints listed above), the ratio  $R_{gs}/R_{po}$  of the geostationary to polar orbiter radiances are computed. The average value and standard deviation of these ratios for a given range  $dz$  of the cosine nadir view angle are then computed.

Figure 4 shows the limb-darkening profiles for three geostationary satellites over the period 1–16 January 1992. The limb functions for the *GOES-7* and *GMS-4* satellites are in good agreement with each other and are consistent with the broadband measurements of Smith et al. (1994). *METEOSAT-4* appears to exhibit a greater amount of limb darkening for reasons, which are not clear at present. One possibility that may account for these differences in limb behavior is that for these geostationary instruments their may be a danger of including the effects of  $\text{CO}_2$  absorption/emission at the long-wave end of the transmission window. However, the overlap of the spectral response functions of each of these satellites barely overlap the  $\text{CO}_2$  region with the greatest overlap occurring for *GMS-4* (according to the ISCCP documentation). What may be more relevant is the shape of the spectral response functions: this would probably require a radiative transfer study along the lines of the forward modeling described in the introduction to resolve this issue. Inspection of images of the spatial variation of  $R_{gs}/R_{po}$  reveals that much of the scatter associated with each data point in Fig. 4 arises from cloudy regions containing high spatial variability. Thus, the scatter may be caused by slight navigation errors, differing footprint sizes due to viewing geometry angle, projection effects of vertical cloud distribution, and/or temporal variability. The upper curve shows the radiative transfer predictions of Minnis et al. (1991) ( $a = 0.0325$ ) discussed above. The empirical estimates of the limb function clearly have a faster fall-off with nadir angle than the Minnis function. The reason for this discrepancy, we suggest, is due to the neglect of aerosols and clouds in the theoretical calculation. The lower curve shows the limb correction [eq. (7) with  $a = 0.09$ ] adopted for the CLAUS dataset. This admittedly crude



approach may lead to some artifacts remaining associated with limb darkening as a result of the limited study performed here. But there is a significant reduction in the number and magnitude of artifacts that can be identified with no correction. Ideally, the limb correction should be determined on a per-satellite basis and for different scene types, etc. This is a substantial task that we hope to tackle in the future.

The limb correction is applied in the following way. For each individual geostationary satellite image a quartic polynomial fit is produced for the radiance as a function of window brightness temperature using the calibration tables available in the ISCCP BT version 1 files. This is used to obtain the radiance, which is corrected for the limb effects by applying expression (6). This is then converted back to temperature by applying Newton's method to the quartic expression. The computational overhead is small, since the expression for the polynomial and limb correction function can be evaluated rapidly and the Newton scheme has a quadratic order of convergence. Also, we only need to store the quartic coefficients for each geostationary image, whereas using an alternative lookup table approach would require a significantly greater amount of storage (following the reformat stage the satellite calibration tables which might be used as lookup tables are no longer required).

## 6. Temporal interpolation

Gaps in the spatially sampled images because of incomplete satellite coverage cause problems for applications which require complete data, such as storm tracking, scale-dependent analysis, and some GCM validation activities. For this reason, we use a temporal interpolation method to "fill in" the data holes to produce a spatially and temporally homogeneous dataset. The interpolation scheme is based on a cross-correlation method to estimate a field of cloud motion vectors over each data void using adjacent images in the series. Interpolation is then performed along the local velocity vector over the neighbouring image frames. As will be shown below, the motion compensation reduces blurring of moving cloud systems, especially in midlatitude regions, as compared with the simple linear interpolation method used by Tanaka et al. (1991), which neglect motion effects.

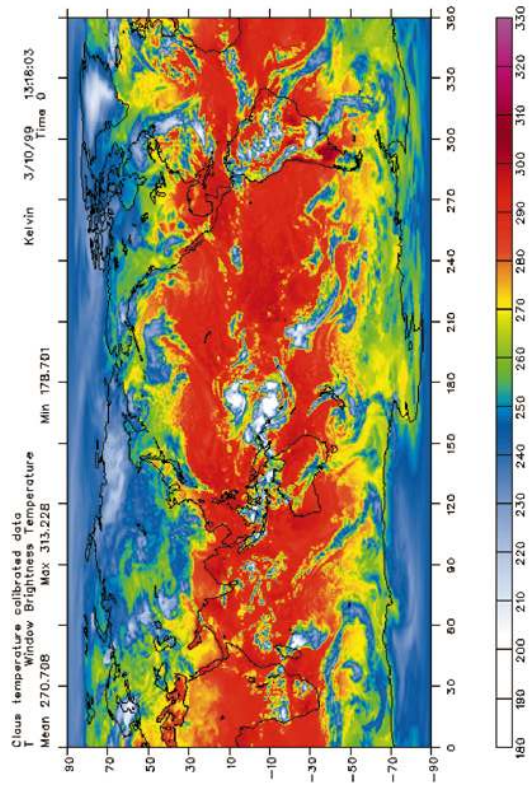
Central to this method is the estimation of cloud motion vectors. Much work has been done in this area as a means of deriving mesoscale and synoptic wind fields for use in forecasting and model validation. This work largely uses cross-correlation methods to search for highly correlated regions between pairs of successive satellite images. Straightforward application of cross-correlation methods, however, often produces irregular velocity fields containing non-physical motion vectors. By thresholding the velocity field according to the correlation coefficient many (but not all) of the erroneous velocity



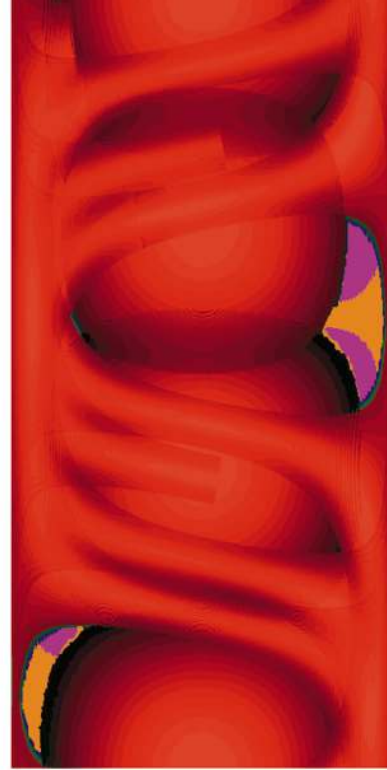
FIG. 5. Comparison of interpolated images of a Southern Hemisphere midlatitude storm using motion compensation (left) and straight temporal averaging (right). The true image is shown in the center panel. The features indicated by the arrows are discussed in the text.

vectors may be rejected, but this can lead to sparsely sampled velocity fields (e.g., Wahl and Simpson 1990). Attempts to produce spatially smooth fields often rely on iterative methods involving, for example, vector median filters (Simpson and Gobat 1994), energy minimization (Larsen et al. 1998), or fluid motion models (Hasler et al. 1998). However, these techniques tend to be computationally expensive, prohibiting their use for large datasets such as that here. In addition, not all these methods clearly demonstrate that the smoothed velocity field actually represents a better estimate of the true field. We have found that for the present dataset, confusion between multiple correlation maxima is a large contributor to the presence of erroneous velocity estimates. This is most likely due to a breakdown in the implicit assumption of the cross-correlation method, which holds that at small scales the image structure evolves entirely through rigid translation and not rotation, dilation, or differential brightness variations. The degree to which these assumptions are valid for cloud imagery depends on the spatial scale and temporal sampling of the image sequence. As, for example, the temporal sampling rate decreases, larger differences may occur in the cloud structure between subsequent images, violating these assumptions and resulting in a poor correlation match for the true displacement vector. If, in addition, a given cloud feature happens to resemble a neighbouring cloud system, then an erroneous correlation peak may appear, even though the corresponding displacement vector does not correspond to true physical motion.

In the present approach, we use a hierarchical method of selecting the final correlation peak. Initially, relatively large blocks are chosen which are designed to follow the motion of synoptic systems such as midlatitude storms or tropical cyclones. We find that for blocks with linear sizes of  $20^\circ$ , the correlation surfaces for the ISCCP B3 data often contain a single peak corresponding to the average motion at that spatial scale. Thus, confusion between multiple peaks rarely occurs on these scales. However, since the velocity field is underresolved at this scale, we divide each large block into four smaller blocks and recompute the correlation surface. The velocity estimate in the large block is used to select the local maximum in the correlation surface produced by each smaller block. (In principle, a quad-tree algorithm



(b) Final Image Scaled to Temperature



(d) Second Quality Byte Image, Color Coded



(a) Final Image



(c) First Quality Byte Image, Color Coded

FIG. 6. An example of the final result of applying the algorithm at 2100 UTC 6 Jan 1992: (a) grayscale image, (b) temperature calibrated, (c) first quality field, and (d) second quality field.



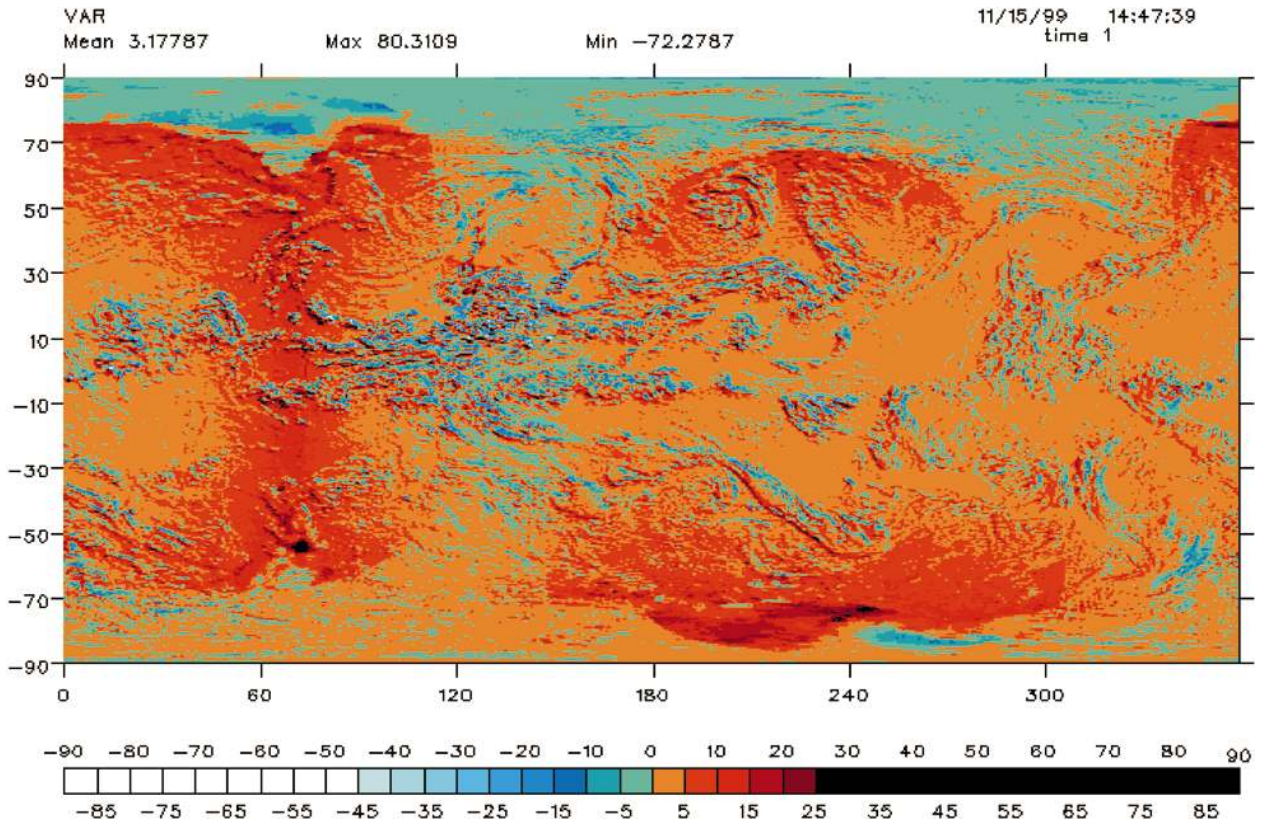


FIG 7. Difference temperature plot for the CLAUS algorithm and the Tanaka et al. (1991) algorithm for 0600 UTC 4 Jan 1984: CLAUS–Tanaka.

could be used, so that the correlation surface need not be recomputed. The total computation time would therefore not be sensitive to the number of hierarchical levels used.) This method implicitly assumes that smaller-scale motions may be regarded as perturbations on a larger-scale flow. For midlatitude systems, squall lines, and tropical cyclones this appears to be a good assumption, with the method producing velocity fields that are substantially more coherent, and that result in smaller interpolation errors than those made from straightforward cross-correlation methods. However, for complex dynamics at small spatial scales, common to tropical convection, this assumption breaks down and the method provides little (if any) improvement over simple linear interpolation as used by Tanka et al. (1991). Finally, we note that this method has little computational overhead, unlike other more sophisticated iterative techniques, which makes it ideal for large datasets such as CLAUS.

Implementation of this method proceeds as follows. Given a data void in the temperature image  $T_i^t$  at time  $t$ , a grid of contiguous blocks centered on the void is overlaid on the preceding image  $T_i^{t-1}$ . Each block is then translated over a portion of the image  $T_i^{t+1}$  to find the displacement vector  $\mathbf{v}_f$  that produces the largest correlation coefficient. If any of the blocks contain incomplete data, that block is repositioned on the next frame, for

example, from image  $T_i^{t-1}$  to  $T_i^{t-2}$ . As a quality control, the matching is then performed in reverse, where the grid of blocks is overlaid on image  $T_i^{t+1}$  and matched against the preceding frame  $T_i^{t-1}$  to produce the vector  $\mathbf{v}_b$ . If the displacement vectors  $\mathbf{v}_f$  and  $\mathbf{v}_b$  agree to within a specified value  $\delta_v$  and have correlation coefficients larger than a specified threshold  $R_0$ , the average of the vectors is used for the local velocity estimate. The vectors are rejected if  $|\mathbf{v}_f - \mathbf{v}_b| > \delta_v$  or if both correlation coefficients fall below  $R_0$ . If only one vector satisfies  $r > R_0$ , then the corresponding vector is selected. Rejected velocity estimates are replaced by local averages of the velocity field. In the event that no neighboring velocity values are defined, the velocity estimate is set to zero. In this case, the interpolation reduces to the simple non-motion-compensated method. The resulting set of motion vectors are then “fed in” to the hierarchical scheme described above. At each stage in the hierarchy, the quality control just described is implemented.

Once the velocity field has been estimated the temperature values  $T_i^t$  in the data void are interpolated according to

$$T_i^t = \frac{\sum_{\Delta t} W(\Delta t) T_{i+\mathbf{v}_i \Delta t}^{t+\Delta t}}{\sum_{\Delta t} W(\Delta t)}, \quad (8)$$

TABLE 4. First quality byte structure. The spare bit fields are used when additional satellites are operating, e.g., an additional METEOSAT.

Bit	Satellite series
1	GMS
2	METEOSAT
3	GOES-E
4	GOES-W
5	NOAA-PM
6	NOAA-AM
7	Spare
8	Spare

where  $\mathbf{v}_i$  is the local cloud motion field, the summation runs over the two neighboring images on either side of the target image, and the temporal weighting function  $W(\Delta t)$  is temporally symmetric, with  $W(\Delta t = \pm 3 \text{ h}) = 5W(\Delta t = \pm 6 \text{ h})$ . Only those images that have spatially sampled data values contribute to the sum. The values of the weighting function were chosen so that the image contributions at times  $\Delta t = \pm 6 \text{ h}$  are negligible when the  $\Delta t = \pm 3\text{-h}$  frames are present. When  $\mathbf{v}_i$  is locally zero, the interpolated value is given by simple temporal averaging.

The left panel in Fig. 5 shows the results of the motion-compensated interpolation technique applied to a midlatitude storm in the Southern Hemisphere. The true image is shown in the center panel, while temporal interpolation without motion-compensation was used to produce the image in the right panel. In this example, motion compensation reduces the amount of blurring due to cloud motion (see lower arrow), reducing the corresponding interpolation error by 30% from 6.5 to 4.6 K. Notice that the cloud band that appears doubled in the right panel (see upper arrow) is much better represented by the motion-compensated interpolation. An example of a complete image obtained from the spatial and temporal parts of the algorithm are shown in Fig. 6, along with its temperature scaled version and images of the two quality fields.

A comparison of Tanaka et al.'s (1991) results and the result of applying the full CLAUS algorithm is shown in Fig. 7. This figure shows that the algorithm reported here consistently has lower cloud-top temperatures. This may be due to using the more recent ISCCP calibration tables, but is more likely to be due to the difference in approach of the spatial sampling. As has already been mentioned, the distance weighting scheme fills more of the grid values in the primary pass than that of Tanaka et al. (1991). The Tanaka algorithm fills the remaining small holes by local averaging which smooths out the temperatures, resulting in warmer cloud tops. Additionally, the box averaging used by Tanaka is a discrete process, while our distance weighting method is continuous, so there may be an element of misalignment in this comparison. Clearly evident is the effect of the limb-darkening correction for the geostationary satellites, which show warmer temperatures than

TABLE 5. Second quality byte packing order.

<i>M</i>	<i>I</i>	<i>Z</i>
8	7 6 5	4 3 2 1

Tanaka et al. (1991) around the geostationary limbs; the Indian Ocean region shows this most clearly; this is where there is only a very small overlap between METEOSAT and GMS and at this particular time there is no polar orbiter coverage for this region.

## 7. Quality indicators

As well as producing the gridded WBT data, a large amount of quality data is also produced, which is stored as two bytes per grid point. Although this amounts to twice as much data as the gridded data, this is justified because it provides users with as much information as possible to allow assessment of the quality of a grid point value based on a knowledge of how it was derived. The first byte contains the contributing satellite information, one bit per satellite and packed according to Table 4.

The second byte (Table 5) contains a missing data value flag in the high end bit, with a value of zero for present and one for missing. The first 4 bits (low end) contain the average satellite zenith angle (*Z*) computed as

$$\hat{z}_i = \frac{\sum_{j=1}^N z_j k_{i,j}^{(s)}}{\sum_{j=1}^N k_{i,j}^{(s)}}, \quad (9)$$

where  $z_j = \cos(\theta_j)$  is the cosine of satellite zenith angle for data point  $j$  and  $\hat{z}_i$  is the mean cosine of satellite zenith angle computed from all contributing satellites to the  $i$ th grid point. The values of  $\hat{z}_i$  are scaled into the integer range 0–15. The remaining three bits (*I*) are used to indicate the means by which the gridpoint data has been derived, for example, primary regression, temporal interpolation etc. The values take the range from 0 to 7, that is, they fit into 3 bits. The various values are listed in Table 6. The temporal values are defined to reflect the number of frames at times  $\pm 3 \text{ h}$  from target which contribute to the interpolation, since they gen-

TABLE 6. Values of the interpolation flags in the second quality byte.

Value	Regression/interpolation level
0	Primary spatial
1	Secondary spatial
2	Tertiary spatial
3	Primary temporal
4	Secondary temporal
5	Tertiary temporal
6	Reserved
7	Reserved

erally provide a more reliable estimate than the  $\pm 6$ -h images. The primary, secondary, and tertiary temporal interpolation values indicate that both, one, or neither of the  $t_0 \pm 3$ -h frames were used in the interpolation. When tertiary temporal interpolation is used, the estimated image value is solely dependent on the image frames 6 h from the target image, and thus represents the least accurate interpolation level. The reliability of the local image estimate therefore increases with decreasing value of the interpolation fields taken as a whole.

The two quality bytes for each grid point are stored as images in two separate files. Color-coded examples of these are shown in Fig. 6, and clearly reveal the complex way that the satellite coverage of the individual satellites overlap, as well as the distribution of methods by which the data have been combined.

## 8. Applications

The types of data that are produced by the algorithm described here can be used for a variety of climatological studies and GCM validation. Processing the whole of the ISCCP B3 archive will produce a dataset from July 1983 to the present, that is,  $\sim 16$  yr, at the 3-hourly time steps. This will allow a range of climatological studies to be performed on cloud systems from meso-scale to supercloud cluster scales, for example African squall lines, modulation of convection by easterly waves, diurnal variations, the Madden-Julian oscillation and the effects of El Niño to name but a few. Salby et al. (1991) used their limited period of data to generate time means, filtered variance and power spectra. The CLAUS data will allow these simple statistics to be generated over a much longer period, to explore the natural variability of tropical cloud systems in particular. The motion and evolution of cloud systems can also be explored (Hodges 1998), although new techniques are currently being investigated to deal with differing cloud structures.

For GCM validation the data can be used to explore whether models produce the correct diurnal cycle, an important diagnostic of the radiation scheme and convective parameterization; also whether the models are capable of simulating such processes as the Madden-Julian oscillation, for example.

As an example, the usefulness of the CLAUS data is demonstrated by a limited study of the diurnal variation in the Tropics using a years worth of data for 1992. The results of this study are shown in Fig. 8 for the diurnal amplitude and Fig. 9 for the diurnal phase (in terms of the local time) for the four seasons obtained by performing a time spectral analysis at each grid point.

The diurnal amplitude of the WBT in Fig. 8, shows that the maximum WBT at this frequency occurs over the land, as would be expected as a result of the low heat capacity of the land surface, which allows it to respond to solar heating and infrared cooling, faster than

the ocean. Regions over the land with lower amplitudes of the WBT indicate the main regions of convective activity. For example, over the African region of the Sahel and Congo for JJA there is a band of lower amplitudes from the Ethiopian highlands to the Guinea coast, which indicates the main region of convective activity for this period.

More general observations of the diurnal phase indicate that for clear-sky regions over the land the maximum WBT occurs in the early afternoon after the peak in the solar heating, whereas for clear-sky regions over the ocean the maximum WBT occurs in the late afternoon due to the larger heat capacity of water, resulting in a slower temperature response to the solar heating. In convectively active regions over the land the maximum in WBT occurs in the morning when the convective activity is decaying. Over the convectively active oceanic regions the maximum in WBT shows less organization with a lot of variability, although around the land margins the phase is consistently in the late evening, possibly associated with the formation of stratiform clouds.

More extended studies will be performed along the lines described above once the full CLAUS dataset is completed.

## 9. Conclusions

A new algorithm has been presented for generating gridded WBT from multiple satellite images. Although the skeleton of the algorithm follows that of Tanaka et al. (1991), the individual elements of the algorithm use alternative techniques that are more comprehensive in their approach, for example the use of spherical kernel estimators for the spatial regression, the introduction of a temporal weighting for the asynoptic polar orbiter data, a correction for limb darkening, and the use of motion compensated temporal interpolation to fill data voids. This results in a reduction in many of the artifacts apparent in the data of Tanaka et al. (1991) (observed in the years' worth of data processed by Tanaka et al. 1991). The CLAUS project will process the whole of the ISCCP B3 archive, amounting to about 16 years' worth of data. This will allow meaningful climatological studies to be performed as described in the previous section, as well as the validation of several aspects of GCM integrations.

Further improvements to the algorithm are ongoing, including studies of the limb-darkening problem: at the moment empirical, but also in the future using radiative transfer models. Further work will also be performed to study the effects of data bias in the source data and due to the sampling of the raw data to B3 resolution. The efficacy of applying the algorithm to full-resolution data will also be explored.

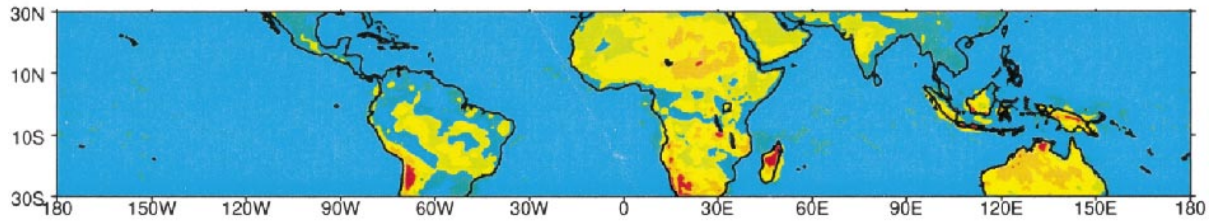
*Acknowledgments.* The data used for this study were obtained from the NASA Langley Research Center



### Amplitude of the diurnal harmonic of BT (CLAUS)

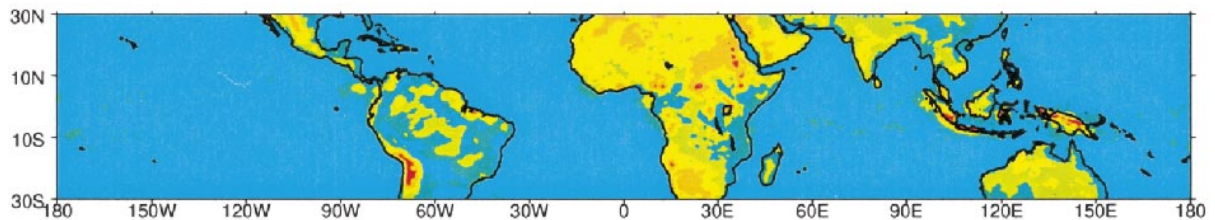
DJF 91/92

Mean 2.93777    Max 23.201    Min 0.00463521



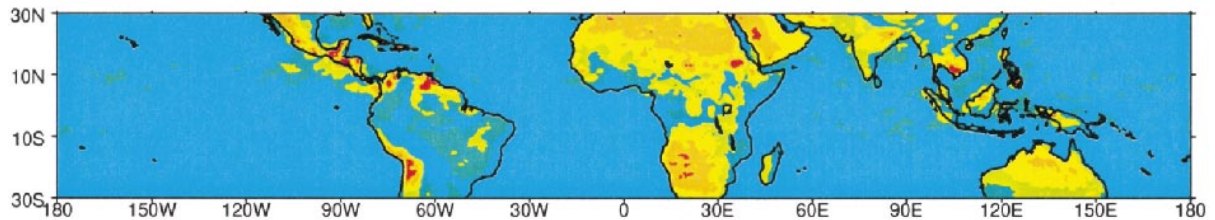
MAM 92

Mean 2.94089    Max 21.0109    Min 0.0029255



JJA 92

Mean 3.04056    Max 23.4387    Min 0.00468916



SON 92

Mean 3.14882    Max 22.8631    Min 0.00439906

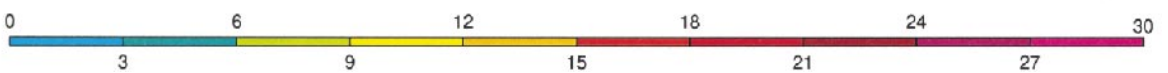
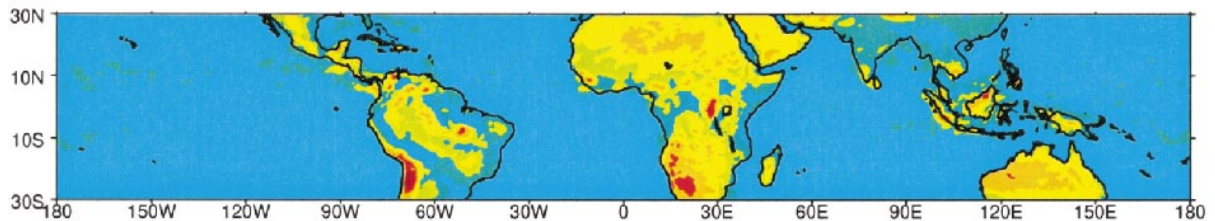


FIG. 8. Diurnal amplitude for the four seasons: DJF, MAM, JJA, and SON.

### Phase of the diurnal harmonic of BT (CLAUS)

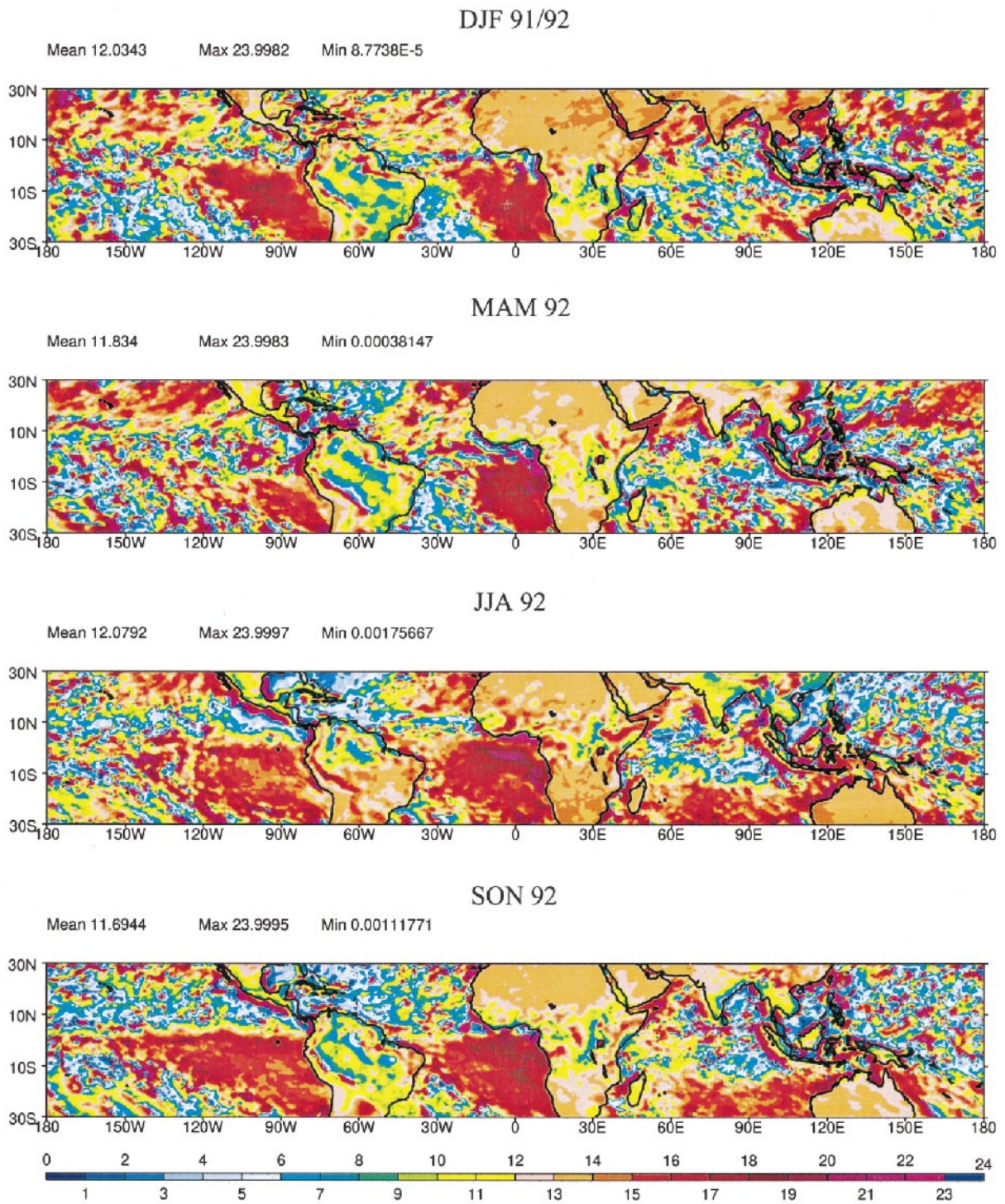


FIG. 9. Diurnal phase in local time for the four seasons: DJF, MAM, JJA, and SON.



EOSDIS Distributed Active Archive Center. The CLAUS project is part funded by the European Commission under Contract ENV4-CT96-0356 (<http://www.ceo.org/clus.html>). The algorithm implementation and archive processing is being undertaken by the The National Remote Sensing Centre (NRSC). The authors would also like to thank the reviewers for their helpful comments and bringing to our attention related work and also Professor Murray Salby for the supply of the data derived from the Tanaka algorithm.

## REFERENCES

- Brest, C. L., W. B. Rossow, and M. Roiter, 1997: Update of radiance calibrations for ISCCP. *J. Atmos. Oceanic Technol.*, **14**, 1091–1109.
- Hasler, A. F., K. Palaniappan, C. Kambhummetu, P. Black, E. Uhlhorn, and D. Chesters, 1998: High-resolution wind fields within the inner core and eye of a mature tropical cyclone from GOES 1-min images. *Bull. Amer. Meteor. Soc.*, **79**, 2483–2496.
- Hodges, K. I., 1996: Spherical nonparametric estimators applied to the UGAMP model integration for AMIP. *Mon. Wea. Rev.*, **124**, 2914–2932.
- , 1998: Feature-point detection using distance transforms: Application to tracking tropical convective complexes. *Mon. Wea. Rev.*, **126**, 785–795.
- Larsen, R., K. Conradsen, and B. K. Ersboll, 1998: Estimation of dense image flow fields in fluids. *IEEE Trans. Geosci. Remote Sens.*, **36**, 256–264.
- Minnis, P., and E. F. Harrison, 1984: Diurnal variability of regional cloud and clear-sky radiative parameters derived from GOES data. Part 1: Analysis method. *J. Climate Appl. Meteor.*, **23**, 993–1011.
- , D. F. Young, and E. F. Harrison, 1991: Examination of the relationship between outgoing infrared window and total long-wave fluxes using satellite data. *J. Climate*, **4**, 1114–1133.
- Rossow, W. B., A. Walker, and M. Roiter, 1997: International Satellite Cloud Climatology Project (ISCCP)—Description of reduced resolution radiance data. WMO Tech. Document 58, 163 pp. [Available from ISCCP Global Processing Center, NASA Goddard Institute for Space Studies, New York, NY 10025.]
- Salby, M. L., H. H. Hendon, K. Woodberry, and K. Tanaka, 1991: Analysis of global cloud imagery from multiple satellites. *Bull. Amer. Meteor. Soc.*, **72**, 467–480.
- Saunders, R., M. Matricardi, and P. Brunel, 1999: An improved fast radiative model for assimilation of satellite radiance observations. *Quart. J. Roy. Meteor. Soc.*, **125**, 1407–1425.
- Simpson, J. J., and J. I. Gobat, 1994: Robust velocity estimates, stream functions, and simulated Lagrangian drifters from sequential spacecraft data. *IEEE Trans. Geosci. Remote Sens.*, **32**, 479–493.
- Smith, G. L., N. Manalo-Smith, and L. M. Avis, 1994: Limb-darkening models from along-track operation of the ERBE scanning radiometer. *J. Appl. Meteor.*, **33**, 74–84.
- Tanaka, K., K. Woodberry, H. Hendon, and M. Salby, 1991: Assimilation of global cloud imagery from multiple satellites. *J. Atmos. Oceanic Technol.*, **8**, 613–626.
- Wahl, D. D., and J. J. Simpson, 1990: Physical processes affecting the objective determination of near-surface velocity from satellite data. *J. Geophys. Res.*, **95**, 13 511–13 528.