The Optical Transient Detector (OTD): Instrument Characteristics and Cross-Sensor Validation

D. J. Boccippio, W. Koshak, and R. Blakeslee

Global Hydrology and Climate Center, NASA Marshall Space Flight Center, Alabama

K. Driscoll, D. Mach, and D. Buechler

University of Alabama in Huntsville, Huntsville, Alabama

W. Boeck

Niagara University, Niagara Falls, New York

H. J. Christian and S. J. Goodman

Global Hydrology and Climate Center, NASA Marshall Space Flight Center, Alabama

ABSTRACT

Lightning data from the U.S. National Lightning Detection Network (NLDN) are used to perform preliminary validation of the satellite-based Optical Transient Detector (OTD). Sensor precision, accuracy, detection efficiency, and biases of the deployed instrument are considered. The sensor is estimated to have, on average, about 20–40-km spatial and better than 100-ms temporal accuracy. The detection efficiency for cloud-to-ground lightning is about 46%–69%. It is most likely slightly higher for intracloud lightning. There are only marginal day/night biases in the dataset, although 55- or 110-day averaging is required to remove the sampling-based diurnal lightning cycle bias.

1. Introduction

The Optical Transient Detector (OTD) (Christian et al. 1996) is a space-borne lightning sensing device, launched in April 1995 aboard the *Microlab-1* (recently renamed *OV-1*) satellite. The instrument is an engineering prototype for the Lightning Imaging Sensor (LIS) (Christian et al. 1992, 1999a), a component of the Tropical Rainfall Measuring Mission (TRMM) (Kummerow et al. 1998) and the Mission to Planet Earth (Asrar and Greenstone 1995). The sensor detects total (intracloud and cloud-to-ground, day and night) lightning from a 735-km altitude, 70° inclination orbit. Detection is achieved by rapid (2-ms update) scanning of a 128 × 128 pixel charge-coupled device (CCD) imaging array, combined with a narrowband (0.000 845 μm) interference filter and real-time event processing. Data have been collected since mission launch and col-

lection will continue through April 2000. As such, the OTD offers unprecedented, potentially unbiased detection of global lightning activity. As with any fundamentally new sensor, quantitative validation of the instrument *as deployed* is necessary before useful science can be done with the data.

In this paper, we examine the resolution, accuracy, detection efficiency, and biases of the OTD in its first two years of deployment, and briefly discuss the false alarm rate of the fielded sensor. A priori, these characteristics can usually only be estimated, as they depend fundamentally upon the actual deployment geometry, ambient noise characteristics, and other unforseen field conditions. Validation generally requires some form of *truth dataset.* The minimum requirements for the truth sensor employed are that its precision be comparable to or higher than the sensor being validated, its accuracy be greater, its sensitivity be known (and preferably greater than the new sensor), and its biases be quantifiable. The sample data size for cross-validation must also be large enough to be statistically significant. ''Validation'' studies that do not meet these criteria may yield interesting collections of case studies, but these will be

Corresponding author address: Dr. Dennis J. Boccippio, NASA/MSFC SD-60, Global Hydrology and Climate Center, 977 Explorer Blvd., Huntsville, AL 35806.
E-mail: Dennis.Boccippio@msfc.nasa.gov

of limited value in quantitative estimation of the new sensor's field performance.

2. Precision

Sensor precision is typically the easiest characteristic to estimate a priori and the least likely to change upon deployment. Precision is usually dependent on 1) the sensor's hardware (design) specifications, 2) the mathematics of the deployed sensor/data observational geometry (in this case, line-of-sight projection distortions in the optical system), and 3) the sampling rate of the sensor itself (e.g., integration window for a charge-coupled optical device such as the OTD). Precision may be thought of as the optimal sensor performance in the absence of noise or other sources of uncertainty. It may be realized in a certain subset of the collected dataset but cannot be relied upon or used in actual error estimates. The precision of the OTD sensor is primarily governed by its design specifications and deployed orbital geometry, and is treated here in terms of its localization and radiance resolutions.

a. Localization resolution

The OTD operates by repeatedly scanning a 128 × 128 CCD pixel array, at nominal intervals of 2 ms. Background radiance maps are continuously updated, and each 2-ms scan is passed through a Real Time Event Processor (RTEP) to isolate pixel transients (candidate lightning), which exceed the background by a fixed 8-bit threshold value. The nominal time resolution (and hence precision) is thus 2 ms. The sensor hardware is similar to that described in Christian et al. (1989, 1992).

The operational spatial precision is determined by the individual pixel field-of-views (FOVs) and the Microlab orbital geometry. The OTD sensor was extensively lab-calibrated using a precision-controlled collimated beam. Among other parameters, the off-boresight angle θ_{ij} and field-of-view α_{ij} for a representative sample of array pixels (i, j) were measured (31 and 13 pixels, respectively). From the measured θ's we determine the sensor's side-to-side and diagonal FOVs:

$$\phi_{\text{across}} = \theta_{1,64} + \theta_{128,64} = 80°, \quad \text{and} \tag{1}$$

$$\phi_{\text{diag}} = \theta_{1,1} + \theta_{128,128} = 100°. \tag{2}$$

The measured pixel parameters may be used to estimate the sensor's actual ground range field-of-view and pixel resolutions. For this we consider the *Microlab-1* orbital geometry and assume a nadir-pointing sensor and spherical earth (Fig. 1). For a pixel (i, j), the side-to-side pixel resolution $r_{i,j}$ at the surface may be computed from

$$\lambda_{1i,j} = \pi - \sin^{-1}\left[\frac{R_{\text{eff}} + z_{\text{eff}}}{R_{\text{eff}}} \sin\left(\theta_{i,j} - \frac{\alpha_{i,j}}{2}\right)\right], \tag{3}$$

$$\lambda_{2i,j} = \pi - \sin^{-1}\left[\frac{R_{\text{eff}} + z_{\text{eff}}}{R_{\text{eff}}} \sin\left(\theta_{i,j} + \frac{\alpha_{i,j}}{2}\right)\right], \quad \text{and} \tag{4}$$

$$r_{i,j} = R_{\text{eff}}(\lambda_{1i,j} - \lambda_{2i,j} - \alpha_{i,j}). \tag{5}$$
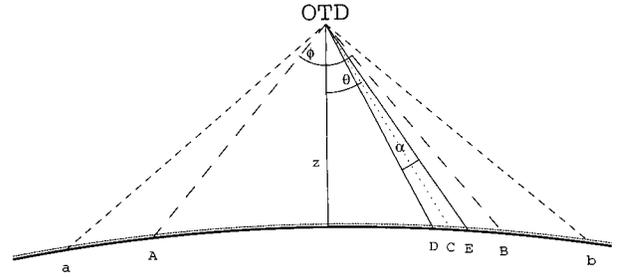
FIG. 1. *Microlab-1*/OTD deployment geometry, drawn to scale and assuming a nadir-pointing satellite. Here, ϕ is the total angular field of view; AB and ab represent the ground range fields-of-view r_{across} and r_{diag}. Also, θ is the off-boresight angle of a given pixel, α is the angular field of view of a given pixel (not to scale), and DE the ground range resolution of this pixel.

Here, R is the earth radius and z the satellite altitude. The small offset due to cloud-top altitude z_{cl} (assumed to be 12 km) is considered by taking $R_{\text{eff}} = R + z_{\text{cl}}$ and $z_{\text{eff}} = z - z_{\text{cl}}$. The total side-to-side and corner-to-corner ground range field-of-views may be computed by substituting the total angular field-of-view for α, that is, by taking $\theta = 0$ and $\alpha = \phi_{\text{across}}$ and $\alpha = \phi_{\text{diag}}$, respectively. For the Microlab altitude $z = 735$ km, we find $r_{\text{across}} = 1279.3$ km and $r_{\text{diag}} = 1915.8$ km. Note that the satellite does not orbit in a fixed-yaw mode, so the FOV width will vary between r_{across} and r_{diag} over the course of each orbit.

As noted above, $\alpha_{i,j}$ were measured in the preflight laboratory for a subset of OTD pixels. The observed $\alpha_{i,j}$ were then interpolated across the entire array using a kriging technique; the resulting values for the full CCD matrix $\boldsymbol{\alpha}_{\text{meas}}$ are shown in Fig. 2a. The instrument was designed to function in a vacuum, and hence the average value of these pixel angular resolutions is too large; that is, in the laboratory atmosphere, the sensor is out of focus. The $\boldsymbol{\alpha}_{\text{meas}}$ values do, however, reveal an asymmetry across the array, suggesting a modest canting of the lens. To examine how this asymmetry might affect the actual pixel resolutions $\mathbf{r}$, we estimate the likely actual angular resolutions $\boldsymbol{\alpha}_{\text{est}}$ by

$$\boldsymbol{\alpha}_{\text{est}} = \frac{\alpha_{\text{optimal}}}{\overline{\alpha_{\text{meas}}}} \boldsymbol{\alpha}_{\text{meas}}, \quad \text{and} \tag{6}$$

$$\alpha_{\text{optimal}} = \frac{\phi_{\text{across}}}{128} = 0.59°, \tag{7}$$

where α_{optimal} is the pixel angular resolution for perfect focusing; that is, uniform nonoverlapping pixels. The ground resolution $\mathbf{r}$ is then calculated from (5) using α_{optimal} (perfectly focused; Fig. 2b) and $\boldsymbol{\alpha}_{\text{est}}$ (possible canting; Fig. 2c), and the difference between the two shown in Fig. 2d. This difference is small and likely no more than that introduced by the (variable) effects of slight off-nadir pointing of the deployed sensor. The results are summarized in Table 1. The likely OTD pixel ground resolution (spatial precision) is thus about 8 km at nadir, 11 km on average, and decreases to anywhere from 18 to 23 km at the array corners.

(a)

(b)

(c)

(d)
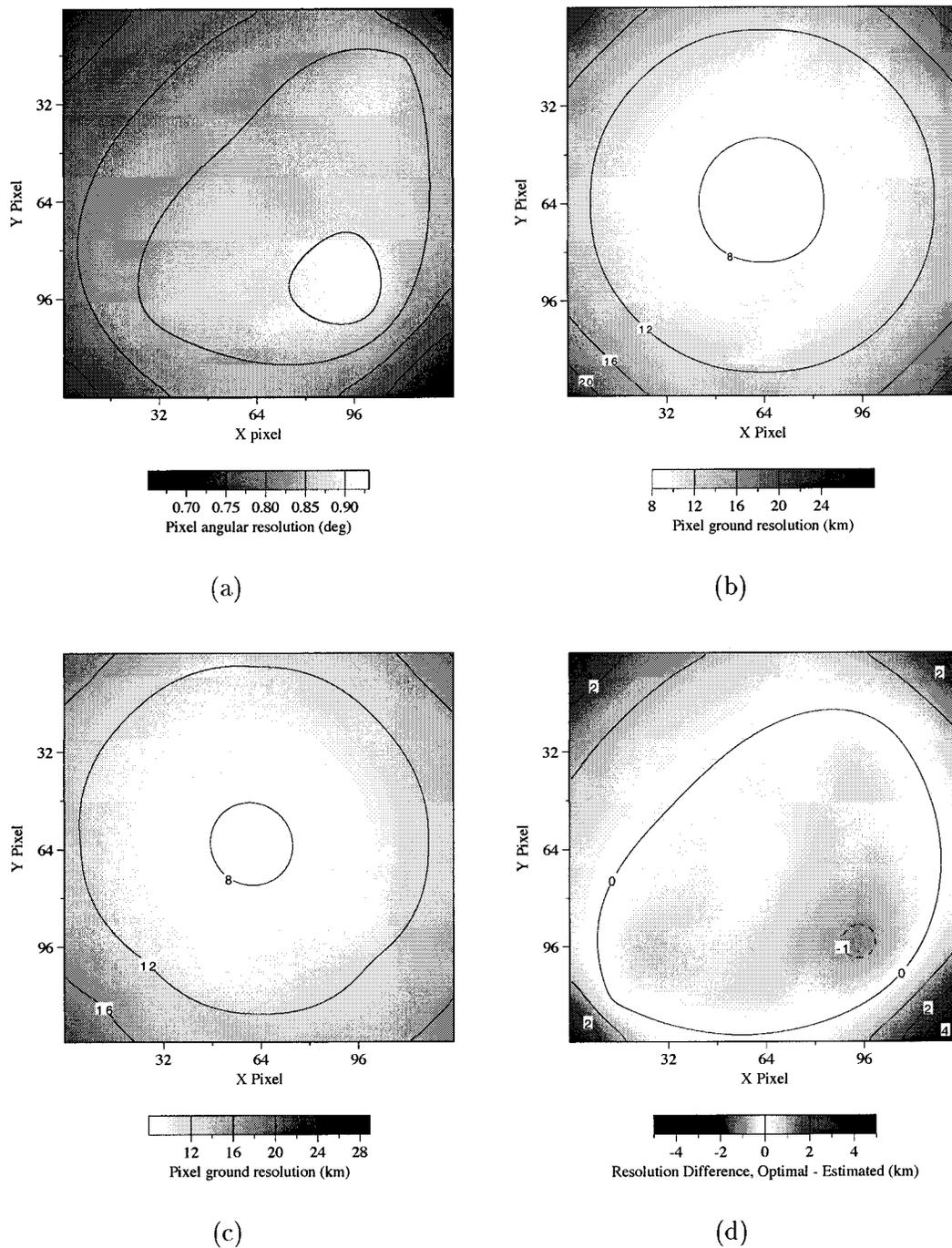
FIG. 2. OTD resolution estimates. (a) Laboratory-measured pixel angular resolutions α_{meas}. These are high-biased due to the laboratory atmosphere; asymmetries, however, likely reveal actual defects in CCD/lens construction or mounting. (b) "Optimal" pixel ground range resolutions for OTD altitude and assumed uniform, nonoverlapping pixel angular resolutions ($\alpha_{ij} = 0.59°$). (c) "Possible" pixel ground range resolutions obtained by scaling α_{ij} from (a) to have a mean equal to the value used in (b). (d) Difference in ground range resolution (b)–(c).

b. Radiance resolution

In addition to estimating the OTD's localization resolution, we may also quantify the resolution of its primary observable, lightning radiance. The sensor records transient optical pulses that stand out above a continuously updated background scene. These are quantized at 7-bit resolution. The "DC" response of the CCD array (and each of its four 64×64 pixel quadrants) has

Table 1. Summary of OTD pixel side-to-side ground range resolution estimates (km). Unfocused estimates are made directly from laboratory measurements of pixel angular resolution and are contaminated by the effects of laboratory atmosphere. Uniform estimates assume optimal, nonoverlapping pixel field of views and perfect CCD design and lens mounting ($\alpha_{i,j} = 0.59$). Possible estimates normalize the measured $\alpha_{i,j}$ to the uniform value but incorporate observed asymmetries across the array.

Pixel	Unfocused	Focused, uniform	Focused, possible
Nadir	10.9	7.5	7.7
Midquadrant	13.8–16.0	10.0–10.2	9.7–11.3
Side	18.0–18.9	13.4	12.7–13.4
Corner	25.6–28.1	23.2–23.4	18.1–20.1
Mean	15.8	11.0	11.4
Median	16.1	11.5	11.2

been laboratory calibrated under five different background intensities. As the background intensities are not routinely recorded, the distributed data (up to revision 1.1) assume three of these background levels, for "night," "twilight," and "day" conditions (determined by the solar zenith angle at each geolocated event, and recorded in the data). For each optical event, the laboratory "AC" calibration (Koshak et al. 1996, 2000) for the appropriate background and quadrant is selected, and the 7-bit raw count converted to a calibrated radiance (Fig. 3). As each calibration curve is strongly nonlinear, the inherent radiance resolution decreases at higher count levels. As seen in Fig. 4, the per-bit radiance resolution may vary from 1.5–40 μJ m^{-2} ster^{-1} over the typical range of radiance counts (7–100).

3. Accuracy

Sensor accuracy is, of course, a far more representative estimate of actual data errors. Accuracy differs from precision in that it encompasses errors introduced by both limitations in the sensor deployment (localization errors) and ambient noise (observable errors).[1] Localization accuracy errors may be introduced by uncertainty in the knowledge of sensor positioning (mounting), deployment (satellite navigation errors) or time-tagging. Observable accuracy errors may be introduced by contamination from phenomena concurrent with the observed datum that may exist in the sensor's detection band (e.g., background illumination or solar glint for optical sensors such as OTD).

An optimal empirical estimation of sensor accuracy requires *concurrent* measurements of the same phenomena (in this case, lightning flashes) by both the test sensor and a truth sensor (see above). For OTD, we shall estimate only the localization accuracy for cloud-to-ground (CG) lightning flashes, by cross-comparison of individual flashes with locations determined by the U.S. National Lightning Detection Network (NLDN) (Cummins et al. 1998). Since sensor accuracy in this case is likely independent of flash type or characteristics, we may extend the results to intracloud (IC) flashes as

---

[1] Here, "observable" is used as a noun to describe any inherent characteristic of a physical phenomenon capable of being detected by the sensor. Observables for the OTD instrument include the radiance, duration, and footprint (area) of optical pulses and flashes.
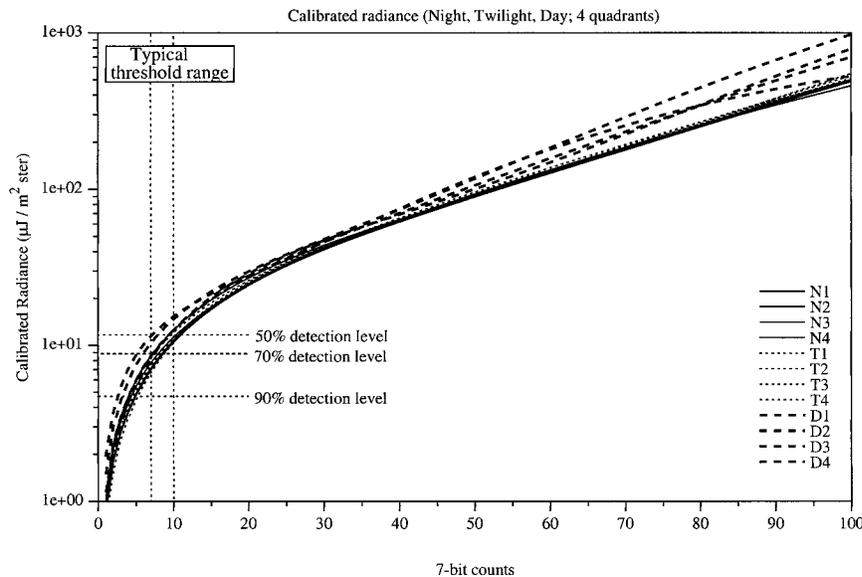
Fig. 3. Radiance calibration curves for the four CCD quadrants and three background conditions (night, twilight, day). The radiance values in the distributed OTD data files are in units of μJ m^{-2} ster^{-1} μm^{-1}; the values shown here incorporate the measurement of filter bandwidth (0.000 845 μm) reported in Koshak et al. (2000). Fiduciary marks are added showing the typical 7-bit thresholds (7–10 counts) and nighttime laboratory-estimated detection efficiency levels from Koshak et al.
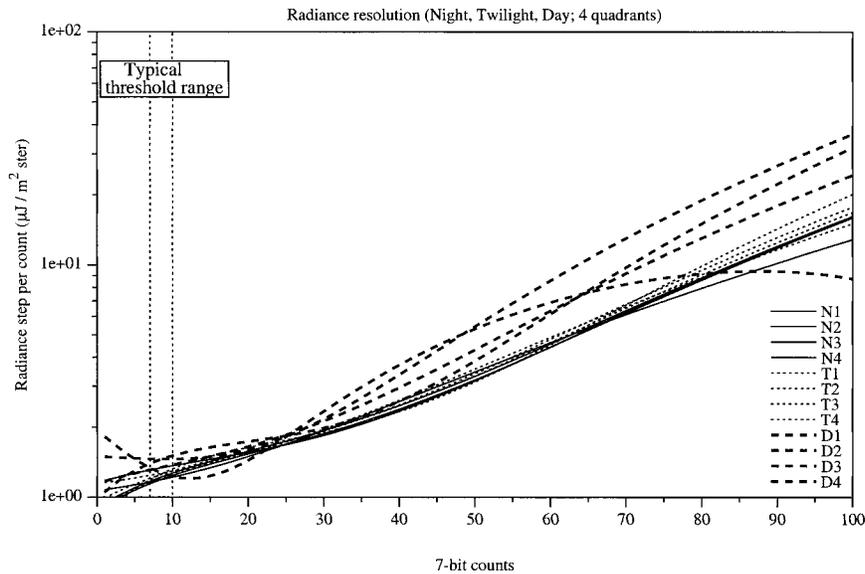
FIG. 4. Per-bit radiance resolution for the four CCD quadrants and three background conditions (night, twilight, day).

well (this is not necessarily the case with the sensor detection efficiency, as discussed below). For accuracy of the OTD observables (primarily lightning pulse radiance), we are only able to treat errors introduced by the OTD data processing and laboratory calibration procedures (an independent truth dataset for cloud-top lightning radiance is not yet available; thus, a lower bound estimate only is provided here).

a. Localization accuracy

Errors in the OTD's localization of lightning pulses arise from imperfect knowledge of the *Microlab-1* satellite's actual attitude and orbital ephemeris. Attitude and ephemeris errors are a result of both telemetry noise and poor measurement of the satellite orientation by the onboard systems. Sporadic poor navigation was suspected early in the mission when some geolocated background scenes were found to be offset from actual coastline contours; it is manifest as a spurious ''spread'' of geolocated lightning flashes around independently measured radar or IR scenes of convective cells. Occasional attitude problems were confirmed as a high-frequency noise component in the reported satellite roll, pitch, and primarily yaw. While attitude/ephemeris repair algorithms are currently being developed at the LIS SCF (Science Computing Facility), the distributed data are contaminated by these localization accuracy errors.

A limited assessment of these errors has been made by comparison with independent measurements of the Microlab ephemeris by the collocated global positioning system (GPS–MET) experiment. A more comprehensive bulk statistical estimate is also possible by cross-comparison with independently measured CG lightning

locations by the NLDN. This comparison is complicated by the fact that NLDN first return stroke times are not identical to OTD optical flash start times; up to several hundred milliseconds of intracloud optical activity may precede a CG stroke, and the OTD is not capable of uniquely isolating return stroke components within collections of contiguous optical pulses.[2] A compromise strategy of assigning nominal OTD flash times by the time of the brightest optical pulse group (assumed to be the first return stroke) was chosen and seems to work well enough for the purposes of this study. A collection of 21 069 ''coincident'' OTD/NLDN observations of the same flashes were compiled for May–September 1995, using a broad tolerance for potential coincidence of less than 2-s time offset and less than 200-km ground range offset of the independent flash localizations (minimum time offsets were used to select between multiple possible pairings). All periods when the continental United States (excluding offshore regions where the NLDN performance drops) were within the OTD field of view (and when the OTD was operating) were used to generate this dataset, hence the subset of NLDN flashes is completely representative of the entire NLDN dataset during this time window (OTD FOV subsets being independent of the underlying CG flash characteristics).

Figure 5 shows the ''errors'' (offsets) ϵ_t between NLDN first return stroke times and OTD times of brightest optical pulse groups for the 21 069 jointly observed

---

[2] An OTD flash is a collection of optical pulse groups (adjacent pixel illuminations within the same 2-ms frame), which are geographically adjacent and have no ''dead time'' between groups greater than 333 ms.

Accuracy of CG return stroke time identification assigned to OTD flashes
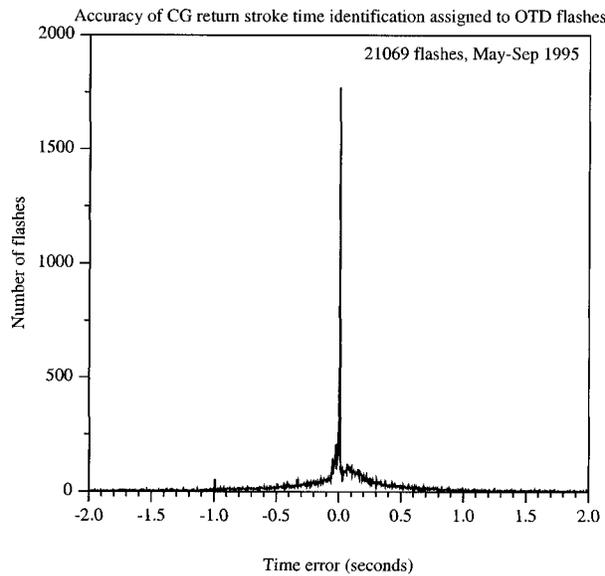
FIG. 5. Errors or offsets ϵ_t, between NLDN first return stroke times and the times of brightest optical pulse in 21 069 flashes observed jointly by the OTD and NLDN.

Empirical estimate of OTD localization accuracy

FIG. 6. Histogram of observed range localization errors of the OTD sensor, based upon a comparison of 21 069 flashes jointly observed by the NLDN. The pixel resolution uncertainty inherent in the sensor comprises part of these errors; the bulk, however, arises from poor satellite navigation. The subsets with time offsets restricted to less than 1.05 seconds (18 799 flashes) and less than 0.5 s (15 271 flashes) exhibit similar behavior; thus incorrectly paired OTD and NLDN flashes are not heavily influencing these errors.

flashes. The sharp spike near 0-s offset demonstrates that for a subset of the flashes, the brightest OTD pulse group was indeed a good indicator of the flash first return stroke [consistent with the much more limited results of Boccippio et al. (1998), who used long-range extremely low–frequency (ELF) measurements to confirm timing accuracy]. The "wings" extending out to ±600 ms represent jointly observed flashes for which the return stroke was not identifiable by this approach (i.e., flashes in which some other flash component produced the brightest optical pulse group). These wings also encompass small errors in the onboard Microlab subsecond clock, which experienced aperiodic fluctuations leading to nominal millisecond ticks having variable duration. From the "peakedness" of the ϵ_t distribution between ±100-ms offset, we infer that clock timing errors are usually less than 100 ms. On rare occasions, the subsecond clock barely overruns the second-counting clock, leading to 1-s errors in nominal time assignment. This effect can be seen by the small spike near −1 s in Fig. 5. The outermost wings on the plot correspond to cases where two different flashes were observed by the sensors, but misclassified as a joint observation of the same flash. The total number of such cases (e.g., $|\epsilon_t| > 1.05$ s) comprised less than 10% of the study sample.

The distribution of time offsets in Fig. 5 suggests that this is an acceptable technique for isolating jointly observed flashes. As such, we may use the spatial offset between the flash localizations from the two sensors as an empirical estimate of OTD localization accuracy. This is possible since the NLDN localization accuracy is reported to be quite good [0.5–2-km ground range, Cummins et al. (1998)] and of finer scale than the OTD

pixel resolution itself (8–24 km, section 2a). Figure 6 thus presents a histogram of the spatial offsets. These include both errors from the inherent OTD pixel resolution and from poor satellite navigation. We find that typical ground range errors typically run from 20 to 40 km (the median is 50 km). Twenty-five percent of the errors are greater than 100 km, and only 10% greater than 150 km. To test whether the most extreme errors were simply a result of misclassification of different flashes observed by the two sensors as "joint observations" of the same flash, the dataset was subsetted to include only those flashes with OTD/NLDN time offsets $|\epsilon_t|$ below 1.05 s (18 799 flashes) and below 500 ms (15 271 flashes) (thus increasing the confidence that the observations were indeed of the same flash). The range error histograms for these two subsets are nearly identical in shape to the original, suggesting that misclassified observations are not unduly biasing the spatial accuracy analysis.

We thus conclude that the effects of OTD pixel resolution and imperfect Microlab navigation lead to a ground range error distribution with a mode of 20–40 km, a median of 50 km, and a 90% level of 150 km. This is somewhat larger than the scale of typical lightning flashes or small storm cells. As such, the sensor can reliably be used for large-scale (2.5° or larger) climatologies (indeed, this was the original OTD mission

goal, and the sensor/satellite specifications do not demand higher accuracy). For smaller-scale case studies, the data should be carefully compared with concurrent radar or IR scenes, or the background images compared against known coastline maps, before high confidence can be attributed to specific flash geolocations.

b. Radiance calibration accuracy

The accuracy of lightning radiance estimates contained in the OTD dataset is determined by both the laboratory calibration procedures employed and the operational application of these calibrations. As described in Koshak et al. (1996, 2000), the laboratory work yielded AC calibration curves for each 64×64 pixel OTD quadrant under five different DC background illumination levels. Polynomial regression yielded highly accurate fits to the laboratory data. However, we have since observed that several of the pixels chosen for the calibration procedure (specifically those around the outermost perimeter of the array) are somewhat unresponsive and likely unrepresentative of the pixel behavior in the interior of the array. When these pixels are removed from the calibration dataset, we find that the calibration curves used in data production may be off by about 10%–20% near the lowest (near threshold) and highest (extrapolated) radiance levels.

A more significant loss of accuracy arises from the way in which these calibrations are operationally implemented. Since the curves vary under background illumination level, appropriate curves must be selected for each lightning event's actual background radiance level. However, visible background scenes are not consistently included in the Microlab-to-surface data stream; depending on the local optical event rate and data storage on the satellite, background scenes are dropped. The result is a recorded background scene rate of about one every 30 s, although this may be lower over areas of high flash rates or ambient noise rates (such as the South Atlantic Anomaly, Pinto et al. (1992).[3] Additionally, the satellite navigation may drift between background scene samples; as a result, we cannot rely on observed background radiances to calibrate the raw 7-bit count radiance data.

The approach taken in the production data was to assign nominal background levels corresponding to day, twilight, or night conditions, depending upon the solar zenith angle of the subsatellite (nadir) point. This approach works reasonably well at night but runs into trouble under highly variable daytime conditions. It also inappropriately handles background scenes including the terminator and both daytime and nighttime conditions (a better approach would have been to use the
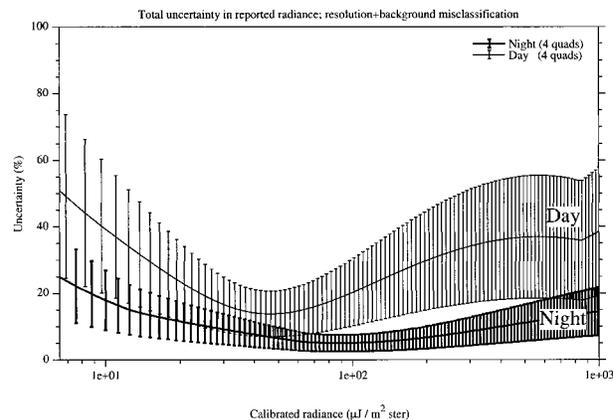
FIG. 7. Inherent uncertainty in (accuracy of) the calibrated radiances based upon the intrinsic radiance resolution and misclassification of the background scene; e.g., twilight as night (or vice versa) or day as twilight (or vice versa). Bars denote range of values over the four OTD CCD quadrants. Errors in the calibration curves themselves are not included here and may introduce another 10% uncertainty at the low and high ends of the radiance scale.

solar reflection angle of each geolocated pixel). In a sense, this can be interpreted as an uncertainty arising from the "resolution" of our contrived background levels and calibration procedure. The effect is greatest when night and twilight conditions are inappropriately assigned (1%–12% relative error in the calibrated radiances), or when twilight and daytime conditions are inappropriately assigned (6%–56% relative error).

Figure 7 shows the combined uncertainty due to both 1) the inherent 7-bit radiance resolution and 2) the effects of misclassifying background scenes using the nadir-point solar zenith angle approach. In this plot, the night, twilight, and daytime calibration curves themselves are assumed to be correct. If we include the 10%–20% error from the inclusion of unresponsive edge pixels (see above) at the lowest and highest radiances, we see that the overall uncertainty is worst both near threshold and for very high radiances. For the user undertaking a bulk-processing or statistical study (i.e., not considering event-by-event pixel quadrant, day/twilight/night conditions or actual solar zenith angle), the worst (daytime) curve must be selected as the actual uncertainty estimate. A reasonable "single-number" estimate of total OTD radiance uncertainty would thus be about 50%, with the understanding that more detailed analysis or case study examination can greatly reduce this uncertainty.[4]

---

[3] The actual average background levels at each pixel used during real time event/background subtraction are not recorded in the data stream due to bandwidth limitations.

[4] While a 50% uncertainty level at first glance appears quite poor, it is important to recall the wide effective dynamic range of the sensor (about 30 dB; Fig. 3). In this context, the 50% uncertainty is actually quite good.

4. Detection efficiency

a. Context

Detection efficiency is a sensor characteristic intimately coupled with environmental noise. It can broadly be described as the ability of the sensor to robustly identify or observe discrete physical phenomena, which are distinct from the ambient noise. It is important to remember that detection efficiency incorporates both the intrinsic sensor sensitivity and the source phenomenon characteristics. (For example, almost all electromagnetic sensors in the geosciences have range-dependent sensitivities, a natural result of signal propagation through attenuating media.) Detection efficiency is rarely a unique number; rather, it depends critically upon the observer's tolerance for noise contamination in the data, and his or her ability to independently discriminate and identify phenomenon signals from the noise. Since this skill is fundamentally probabilistic, the deterministic question ''what is my instrument's detection efficiency?'' must be recast as a question with at least one free parameter, that is, ''what is my instrument's detection efficiency, for a given confidence in signal-from-noise discrimination?'' Complications such as range dependence may add additional free parameters. It is then the scientist's task to select representative estimates of the detection efficiency for operational use.

Although the OTD observes cloud-top lightning illumination, and thus suffers from little clear-air attenuation of the optical signal, its detection efficiency may still be affected by attenuation, specifically the multiple scattering of light within the cloud (Koshak et al. 1994). Since flash altitudes generally follow a bimodal distribution [cloud-to-ground strikes originating at 5–7-km altitude, intraclouds at 8–10-km altitude; Krehbiel (1986), Boccippio et al. (1999)], we should treat the detection efficiency for each flash population separately and assume a common source altitude for each population. As noted above, we directly estimate only CG detection efficiency in this paper, although we later infer the relative behavior of the IC detection efficiency (section 8b) and quantify the sensitivity of OTD bulk detection efficiency estimates to differences between the two (section 8c).

Koshak et al. (1996, 2000) provided an estimate of the OTD bulk detection defficiency (DE) based upon laboratory measurements of the sensor transient response (sensitivity) and a distribution of lightning pulse radiances reported in Christian and Goodman (1987). After allowing for a slight increase (to 6.5 μJ m^{-2} ster^{-1}) in the operational detection threshold to reduce the sensor false alarm rate (FAR), Koshak et al. found nighttime flash detection efficiencies of 74%, 67%, 71%, and 75% for quadrants 1–4 of the OTD, respectively (other laboratory-based nighttime detection efficiency levels are illustrated in Fig. 3). These estimates account only for the intrinsic CCD sensitivity, and the FAR constraint is based upon only internal sensor (CCD readout) noise

in laboratory conditions. The estimates correspond to the lowest threshold (14) operationally used in the deployed sensor, and to the most sensitive background conditions (night scenes). They thus provide a reasonable upper bound for DE estimates derived in this empirical study.

The deployed sensor experiences a nonnegligible FAR. This is largely due to effects of the space environment. Internal electronic noise has been found to increase with decreasing temperature, peaking when the beta angle (angle between the solar vector and satellite orbit plane) minimizes solar heating and leaves the electronics anomalously cold. An additional and significant source of false alarms is the impact of high-energy radiation upon the sensor at both acute and oblique angles to the CCD array: a highly sensitive CCD camera placed in a low-earth orbit is inherently also a good radiation detector. To reduce the FAR of the deployed sensor to acceptable levels, software filtering at the data processing stage must be performed. High energy particles impacting the CCD at oblique angles are easy to remove from the data stream as they produce characteristic ''streaks,'' which are very dissimilar to lightning signatures. High energy particles impacting the CCD at acute angles present a greater problem, as the triggered events they cause fire only a few pixels and are difficult to distinguish from actual lightning signatures. As such, they must be filtered probabilistically; the approach taken was to implement an adaptive filter (dependent on the ambient triggered event rate), which passes only triggered events whose spatiotemporal clustering (in time, pixel, and geographic space) is sufficiently nonrandom in nature. This approach will, by necessity, also remove some lightning detections from the data stream if the spatiotemporal clustering of their associated optical pulses is indistinguishable from random noise. The adaptive filter was tuned to optimize the trade-off between the effective DE and FAR. As such, through cross-sensor validation we can only estimate an *operational detection efficiency,* which includes both the intrinsic sensor DE and the effects of the software filters. It is clear that this operational detection efficiency is now contingent not only upon the OTD hardware but also upon the environmental noise and the skill of our software filters.

b. Estimates

As noted above, we estimate OTD detection efficiency by cross-comparison of individual CG flashes observed by the NLDN. The NLDN is reported to have a high detection efficiency (about 80%–90%) and high localization accuracy in space (0.5–2 km) and time (1 ms) (Cummins et al. 1998). As discussed in section 3a, the OTD's spatiotemporal localization accuracy appears to be much lower. As such, our detection efficiency estimates are contingent upon the space- and time-localization ''errors'' (ϵ_r, ϵ_t) between OTD- and NLDN-

TABLE 2. Summary of trigger threshold changes for OTD. Other thresholds may have been used briefly (an orbit or two) during these periods; users should consult the metadata distributed with the data stream for orbit-by-orbit trigger levels. Final threshold of 15 is valid up until the date of publication. When correctly processed, 8-bit radiance counts not greater than the 8-bit threshold are rejected by the RTEP. Due to an error in processing from 7 Apr–20 Jul 1995, data at the actual threshold values were rejected, thus making the effective thresholds slightly higher.

8-bit threshold	7-bit in files	Effective 8-bit threshold	Date	DOY	UTC
14	8–128	15	7 Apr 1995	097	0024
20	11–128	21	9 Jun 1995	160	1834
18	10–128	19	13 Jun 1995	164	0000
17	9–128	17	20 Jul 1995	201	1115
15	8–128	15	23 Oct 1996	297	0444

observed flashes, which we are willing to tolerate in order for two observed flashes to be termed coincident. This is an important illustration of how localization and observable noise (errors) require detection efficiency estimates to be treated probabilistically.

Incorporating the OTD sensor's geographic view times (or our best estimates of them, given the platform's imperfect navigation data), we may isolate a subset $n_{possible}$ of NLDN-observed CG flashes that should have been observed by the OTD. This may be compared with the actual number of jointly observed flashes n_{seen}, and an empirical estimate of the OTD's CG detection efficiency can be derived:

$$DE_{OTD,CG} \sim \frac{n_{seen}(|\epsilon_t|_{max}, \epsilon_{rmax})}{n_{possible}}. \quad (8)$$

This formulation ignores a possible subset of real CG flashes that are undetectable by the NLDN but visible to the OTD (and hence not uniquely identifiable with these data). The existence of such a population would

FIG. 8. Cloud-to-ground lightning detection efficiency estimates for OTD 8-bit thresholds 14–19. The independent variable is the time offset $|\epsilon_t|$ between nominal (first return stroke) NLDN flash time and nominal (peak radiance pulse group) OTD flash time deemed acceptable to classify two flashes as jointly observed (or "found" by OTD). Reasonable limits on $|\epsilon_t|$ would be between 300 and 600 ms (Fig. 5). Curves for thresholds actually used in the distributed dataset are solid; curves for datasets employed only in this study are dashed.

imply that Eq. (8) is an underestimate of the true $DE_{OTD,CG}$.

During the period in which joint OTD and NLDN data were examined, the OTD sensor was set at four different threshold detection levels (14, 17, 18, and 20 8-bit counts; Table 2). Due to an error in processing, the production data distributed to the community do not contain pulses with radiance counts equal to the actual thresholds, and thus instead have effective trigger levels of 15, 17, 19, and 21 8-bit counts. The two data streams (distributed and reprocessed) thus allow us to make detection efficiency estimates for thresholds 14, 15, 17, 18, and 19 (not enough data were collected at threshold 20/21 to make statistically significant estimates). Since the trigger threshold has a direct bearing on the overall DE, only subperiods of similar threshold settings may be used for a given DE estimate. In the current study, NLDN populations $n_{possible}$ of 4571, 15 119, and 7970 flashes are used in the calculations. Recalling the difficulties in pairing OTD- and NLDN-observed flashes (section 3a), the DE estimates must be presented with free parameters, namely, the time and space errors ($|\epsilon_t|$, ϵ_r) between NLDN first return stroke time and nominal OTD flash time deemed acceptable to claim a flash "jointly" observed. From Fig. 5 we determine that $|\epsilon_t|_{max} \sim 300$–600 ms is a reasonable bound. We leave ϵ_{rmax} set at 200 km, as in section 3.

Figure 8 presents the empirical estimates of operational $DE_{OTD,CG}$. For a 300-ms acceptable $|\epsilon_t|$, the DE is estimated at 42%–58% for thresholds 14–19. For a 600-ms acceptable $|\epsilon_t|$, it is estimated at 55%–71% (a 13% gain). Recalling that the laboratory estimates of sensor-only *nighttime* DE for threshold 14 averaged to 72% (Koshak et al. 1996, 2000), it is apparent that the additional decrease in DE from the software noise filters (to 55%–71%) is, on average, small (at most 17%).

Thus, for the period 20 July 1995–23 October 1996 (threshold 17 setting), a reasonable estimate of the OTD CG detection efficiency would be 56% ± 10% (Table 3). For the period 23 October 1996–present, a reasonable estimate would be 62% ± 7%.

5. Bias

Instrumental bias is commonly associated with sensitivity. For example, lightning flash rates at far dis-

TABLE 3. Summary of CG detection efficiency estimates for OTD. Threshold 15*, 17*, and 19* data were emulated from raw data collected at thresholds 14, 14, and 18, respectively (flashes with all their optical events below 15, 17, and 19 counts were removed to form the simulated higher threshold datasets). Thresholds 15 and 17 were used during collection of the vast majority of all OTD data to date.

8-bit threshold	14	15*	17*	17	18	19*
CG detection efficiency	57%–72%	55%–70%	50%–66%	46%–60%	42%–56%	42%–55%

tances from a range-dependent radio-frequency (RF) or very low–frequency (VLF) lightning location system will clearly be biased toward higher peak current or dipole moment change flashes. Such bias will of course be more pronounced when the underlying distribution of the sensor's observable is strongly nonuniform or is comprised of distinct subdistributions. The cloud-top optical radiances of intracloud and cloud-to-ground lightning flashes, for example, belong to two distinct (but overlapping) distributions (Goodman et al. 1988). In this case, sensitivity-biased observations of the underlying distributions can lead to erroneous inferences about the behavior of the parent population, ''all lightning.''

Bias can also be introduced by inadequate or suboptimal sampling of the underlying population. Sampling-induced bias is similarly most severe when this population is strongly nonuniform in space or time; stochastic processes such as rainfall and lightning are highly variable in both space (geographic variability) and time (annual, seasonal, and diurnal cycles). Sampling bias may arise either inadvertently or of necessity [e.g., by a polar-orbiting satellite's fixed orbital ephemeris, Salby and Callaghan (1997)]. Again, this bias is only relevant when the sensor is used to characterize the observable's parent population (e.g., if constructing a lightning climatology). As this is a primary objective of the OTD mission, we shall consider sampling bias, specifically the effects of aliasing from the diurnal lightning cycle.

Finally, bias can occur if sensor performance varies under differing environmental conditions, and these differing conditions are systematically in phase with spatiotemporal variability of the parent population. This is essentially a combination of instrumental and sampling bias; in the case of OTD, this might be manifest if the sensor's detection efficiency varied under daytime and nighttime conditions, and the parent lightning population were also skewed toward one of these two conditions. Because daytime discrimination of faint lightning signals from space is an essentially new technology, this is a plausible concern. We demonstrate below that no significant day/night differences are found in the sen-

sor's estimated operational detection efficiency, and hence infer that day/night biases are not present in the observed data.

Table 4 summarizes the observed detection efficiency biases derived from the paired OTD/NLDN data. For thresholds 15 and 17 (the bulk of the OTD data), the apparent +CG detection efficiency exceeds the −CG DE by 12%–14%. Since +CGs comprise only about 10% of the total CG population, this results in an overall bias of about 1% in the total CG estimate, or about 0.25% in the total lightning estimate. Additionally, the IC contamination of 1995 NLDN data (Wacker and Orville 1999a,b) suggests that part of this ''bias'' may simply be a result of higher $DE_{OTD,IC}$, and the bias estimate thus an upper bound. The possibility of different NLDN positive and negative CG detection efficiencies has little bearing on this result; readers may use the methodology developed in sections 8a,c below to convince themselves of this. Together, the small inferred bias, the rarity of positive CG flashes, the possibility of NLDN IC contamination and the insensitivity to differential NLDN positive and negative DE suggest a negligible polarity bias in $DE_{OTD,CG}$, and an even more inconsequential effect on $DE_{OTD,bulk}$.

For these same thresholds, the nighttime DE appears greater than the daytime DE by 6%–11% (also Table 4). Using the 1995 NLDN-observed United States CG population as a proxy for actual total lightning day–night bias, we find that daytime CG flashes comprise 51%, 68%, and 64% of all CGs in May, June, and July of that year. The convolution of the apparent OTD day/night DE bias of 6%–11% with the apparent difference in actual day/night flash populations (<20%) suggests that this DE bias should not unduly affect the OTD's statistical estimates of the total global lightning population.

Diurnal bias may also be present if the data are not smoothed over sufficiently long time scales. *Microlab-1*'s revisit span for a particular equatorial earth coordinate and local hour is approximately 55 days; that is, it takes that long for the sensor to fully sample the diurnal cycle at a given location. Since the diurnal lightning cycle over land is typically very pronounced, al-

TABLE 4. Summary of detection biases for OTD.

8-bit threshold	14	15*	17*	17	18	19*
Positive/negative CG bias	12%–13%	12%–14%	12%–13%	15%–18%	14%	19%–20%
Night/day CG bias	10%	6%	11%	7%–10%	12%–15%	4%–6%
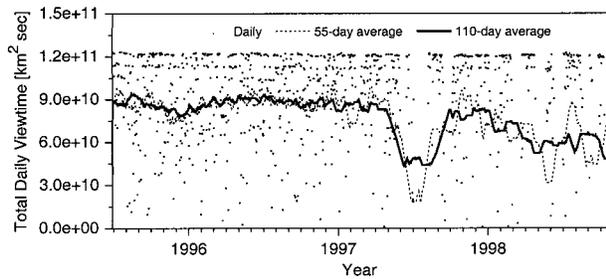
FIG. 9. Summary of total daily OTD ''viewtime'' from mission launch through autumn 1998. Viewtime can be expressed as the (*latitude, longitude, time*) sampling volume (units: km^2 s) explored by the OTD field of view (a ''tube'' in this space). Viewtime losses occur when the satellite undergoes scheduled resets, and when the satellite fails to warm adequately during low solar beta-angle periods. The 55- and 110-day moving averages are overlaid.

iasing of this cycle due to improper averaging can be quite severe. Convective phenomena with frequencies higher than 55 days that are likely subject to a diurnal modulation (such as the Madden–Julian Oscillation) are thus not observable with the OTD data. At minimum, 55-day averaging should be used in intraseasonal or longer time scale examination of the OTD data. Due to increasing data dropout after 1996 as the sensor and satellite aged (Fig. 9), 110-day averaging is recommended for the stablest regional flash rate estimates.

6. False alarm rate

As discussed in section 4, the OTD sensor detects a significant number of false alarms (though few of these are manifest in the distributed data). While a detailed analysis of the FAR is beyond the scope of this paper (and not determinable with the data used here), we briefly describe the rationale for considering the operational (distributed data) FAR to be negligible.

False alarms may arise from three sources: intrinsic CCD or sensor hardware noise, nonlightning optical transients form the earth surface, and high energy particles in the low earth orbit spacecraft environment. The sensor system was designed to keep false alarms from the first two sources below 10%, using radiometric techniques described in Christian et al. (1989), Christian (1999). Laboratory calibration by Koshak et al. (1996, 2000) confirmed the basic realization of these design goals. Further, these false alarms occur at the event (pixel) level and have little temporal persistence, thus noise from these sources lacks the spatiotemporal clustering characteristic of true lightning pixel illuminations. Such clustering is used to filter the raw data, and most of these false alarms (along with some low information content true flashes) are thus removed from the final data stream, leading to a flash FAR from these sources well below 10%.

False alarms from the local space environment comprise the bulk of spurious pixel events. Outside of the South Atlantic Anomaly (SAA) (Pinto et al. 1992) and near-polar latitudes, this noise source is geographically fairly uniform, as observed in maps of raw OTD event data. These false events are filtered as described in section 4, and again their lack of temporal persistence makes them distinguishable from true lightning signals. While a complete validation of the effective (distributed data) FAR would require a large total lightning truth dataset (not available), we can infer from geographic maps of the filtered data that FAR is a very small component of the final data. Particle noise would show up as a ''DC'' geographically uniform component of regional flash rate in maps of the filtered data; it does not (Christian et al. 1996, 1999b). Indeed, there are regions where we physically expect no lightning (the southeast Pacific cold ocean gyres, etc.) where no flashes are found in the filtered data.[5]

We thus tentatively conclude that by original design and through postprocessing, the OTD data are *overfiltered* rather than underfiltered, and that false alarms comprise an undetectably small fraction of the distributed data.[6] If a nonzero FAR occurs in the distributed data, its effects will almost certainly be too small to be of consequence to this validation study. By mission's end, a large enough coincidence dataset may be assembled with surface total lightning validation sensors (whose effective range is tiny compared to the NLDN) and this conclusion may be reexamined; however, the inherent localization ambiguity in the OTD sensor will complicate such an analysis to an even greater extent than in this study.

7. Variance

While not directly related to sensor performance and validation, estimates of variance in a new dataset are necessary prior to its use in scientific studies. This is especially true when dealing with a highly stochastic, spatially and temporally variable process such as lightning or rainfall. Accurate estimation of the sensor parameters described above (especially sensitivity and bias) are necessary before suitable dataset variances (e.g., in regional flash rate estimates) can be computed. While such an assessment is beyond the scope of this paper, well-established techniques (developed in conjunction with the TRMM program) exist for its deter-

---

[5] The same argument can be applied for sensor noise and optical artifacts from nonelectrified clouds; sensor noise would be another geographically uniform component not observed in the distributed data, and spurious OTD flashes are not observed in regions of widespread nonelectrified cloud cover (e.g., the subtropical east Pacific stratocumulus fields).

[6] This inference holds for regions outside of the SAA; the use of data quality metrics to remove radiation noise within the SAA is discussed in the documentation distributed with the OTD data.

mination (Salby 1982; Bell 1987; Bell et al. 1990; North et al. 1993).

8. Error analysis

The accuracy and detection efficiency estimates presented in sections 3 and 4 are subject to the limitations of our metholodogy, assumptions, and any errors in the NLDN truth dataset. In this section, we examine several possible sources of error, including the presence of IC flashes in the NLDN data, the effects of differential OTD CG and IC detection efficiency, the effects of errors in pairing of OTD and NLDN flashes, and the effects of errors in the OTD optical pulse grouping algorithm when counting flashes. We conclude that none of these factors leads to uncertainty greater than that already cited in section 4, and many of them lead to competing effects on the detection efficiency estimates of that section.

a. Intracloud flashes in the NLDN data

Wacker and Orville (1999a,b) have recently documented possible intracloud flash contamination of NLDN data after October 1994, appearing nominally as a large number of weak positive CGs in the distributed NLDN data. In this appendix we investigate the impli-

cations of this contamination on our OTD detection efficiency estimates and conclude the effects are negligible in comparison with the fundamental uncertainty in pairing NLDN and OTD flashes (this result would differ with the LIS instrument, which has a much higher localization accuracy).

We begin by defining the following terms: $N_{CG,true}$ and $N_{IC,true}$ are the true number of CG and IC flashes occurring within the sampled time–space windows. $DE_{OTD,CG}$, $DE_{OTD,IC}$, $DE_{NLDN,CG}$, and $DE_{NLDN,IC}$ are the appropriate true detection efficiencies for each instrument and each flash type, with the latter representing the net contamination of post-1994 NLDN data by IC flashes. The definitions of N_{seen} and N_{poss} are as given in section 4, and we define $f = N_{seen}/N_{poss}$. For purposes of this analysis we assume perfect pairing capability, that is, no time or space errors between the nominal times of OTD and NLDN flashes, and perfect skill at properly matching these flashes. This will result in worst-case estimates, as actual pairing ambiguity will tend to "dilute" the effects studied here. We also assume no covariance between flashes "missed" by the sensors; that is, the populations of flashes missed by each sensor are independent. This again will yield a worst-case estimate of contamination effects.

The f term is then immediately representable by

$$f = \frac{DE_{OTD,CG}(DE_{NLDN,CG}N_{CG,true}) + DE_{OTD,IC}(DE_{NLDN,IC}N_{IC,true})}{DE_{NLDN,CG}N_{CG,true} + DE_{NLDN,IC}N_{IC,true}}, \tag{9}$$

where we recognize that N_{seen} by OTD is composed of both CG and (NLDN misclassified) IC flashes, and the latter are quantifiable through the bulk NLDN IC detection efficiency. We further define a climatological IC:CG ratio for the domain (continental United States), $z = N_{IC,true}/N_{CG,true}$, and a ratio k of OTD IC and CG detection efficiencies; $k = DE_{OTD,IC}/DE_{OTD,CG}$. In all likelihood $k > 1$, although it is left as a free parameter in this analysis.

Using these definitions, we may now rewrite f as

$$f = DE_{OTD,CG}\frac{DE_{NLDN,CG} + kzDE_{NLDN,IC}}{DE_{NLDN,CG} + zDE_{NLDN,IC}}. \tag{10}$$

The appropriately inferred OTD CG detection efficiency, given a measured f, is thus

$$DE_{OTD,CG} = f\frac{DE_{NLDN,CG} + zDE_{NLDN,IC}}{DE_{NLDN,CG} + kzDE_{NLDN,IC}}, \tag{11}$$

where this expression reduces to our simple assumption ($DE_{OTD,CG} = f$) when the NLDN IC detection efficiency is negligible. This expression directly yields the bias in OTD CG DE estimates (under the assumption of perfect

pairing) from IC NLDN contamination. It is immediately evident that for reasonable parameter values (e.g., $DE_{NLDN,CG} = 0.85$; $DE_{NLDN,IC} = 0.02$, $k = 1.2$; $z = 3.0$), the bias is less than 2%; indeed, $DE_{OTD,IC}$ would have to be twice as great as $DE_{OTD,CG}$ (or the $DE_{NLDN,IC}$ be twice as large as estimated above) to even result in a 10% bias in the estimated $DE_{OTD,CG}$.

We can constrain the above result by estimating a worst-case NLDN IC contamination (and hence $DE_{NLDN,IC}$) using results presented in Wacker and Orville (1999a). For this we assume that their positive CG increases over 1989–93 climatological values are *all* due to IC contamination (i.e., no meteorological modulation of the flash count in 1995). The false positive CG excess in 1995 is thus $(2.1 - 0.7)$ or 1.4×10^6 flashes. We further assume the pre-1995 $DE_{NLDN,IC}$ is zero (again, the most conservative choice) and find

$$N_{IC,obs} = DE_{NLDN,IC}zN_{CG,true}, \quad \text{and} \tag{12}$$

$$DE_{NLDN,IC} = \frac{N_{IC,obs}}{z\dfrac{N_{CG,obs}}{DE_{NLDN,CG}}}. \tag{13}$$

TABLE 5. Estimated z (IC:CG) ratios over the continental United States. For these estimates, all OTD-observed flashes not paired with an NLDN-observed CG are assumed intracloud flashes.

8-bit threshold	14	15*	17*	17	18	19*
"IC":CG ratio	3.3–5.6	3.0–5.2	2.7–5.1	2.6–5.1	2.8–5.7	2.5–5.2

Under the above assumptions, the largest possible estimate of $N_{IC,obs}/N_{CG,obs}$ for 1995 is $1.4 \times 10^6/21.3 \times 10^6$, thus

$$\mathrm{DE}_{NLND,IC} \leq \frac{0.0657}{z}\mathrm{DE}_{NLDN,CG}. \qquad (14)$$

Substituting the maximum estimate from Eq. (14) into Eq. (11) and simplifying, we find

$$\mathrm{DE}_{OTD,CG} = f\frac{1.0657}{1 + 0.0657k}. \qquad (15)$$

We thus have two possible estimates for the OTD CG detection efficiency: $\mathrm{DE}_{OTD,CG,clean}$ given by Eq. (11) assuming no NLDN IC contamination and $\mathrm{DE}_{OTD,CG,contaminated}$ given by Eq. (15) (assuming the worst-case contamination estimate). The ratio of these (clean:contaminated) is simply linear in k and yields between a -2% and $+3\%$ inferred $\mathrm{DE}_{OTD,CG}$ bias for plausible values of k (between 0.7 and 1.5). If $k = 1$, that is, if OTD detects IC and CG flashes equally well, there is no net bias in the inference. A 2%–3% bias can be viewed as negligible in comparison to much larger uncertainty in flash pairing capability between the two sensors (section 3a).
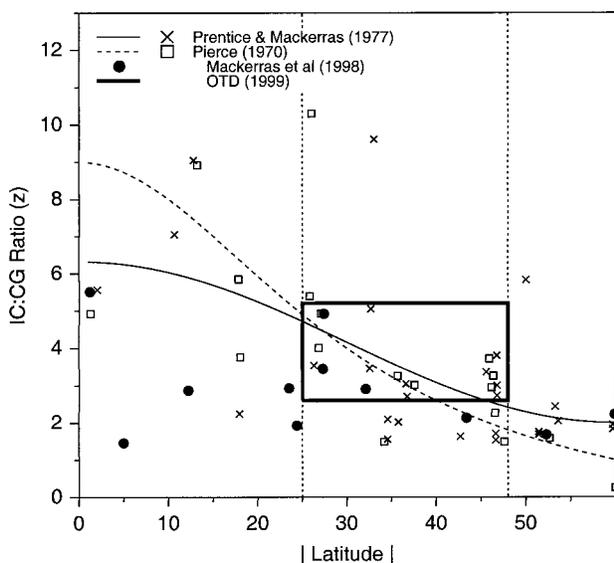
FIG. 10. Summary of current and prior estimates of the IC:CG ratio z for various latitudes. Recent results by Mackerras et al. (1998) are shown as dots, empirical fits to earlier estimates by Pierce (1970) and Prentice and Mackerras (1977) are shown as curves. Bounds using the paired OTD/NLDN data from this study are shown for 25°–48°N enclosed by the dark solid box.

b. OTD CG and IC detection efficiency

The possibility that the OTD IC detection efficiency differs significantly from its CG detection efficiency must be explored, both to help constrain error analyses in this section and to assess the impacts of operationally assuming that these efficiencies are the same to construct "absolute" flash rate estimates. Direct validation of intracloud flash DE is complicated by the rarity of ground-based IC detection systems, the limitations in their effective range, and the lack of validated algorithms for clustering individual ground-observed RF pulses into "flashes." We can, however, provide indirect evidence that the OTD CG detection efficiency more or less holds for IC flashes as well. During times when the OTD and NLDN field-of-views overlap, we can estimate the IC/CG flash ratio by tentatively classifying all jointly observed OTD flashes as CGs, and all other OTD flashes as ICs. While this estimate is imperfect due to the NLDN detection efficiency and the difficulty in isolating jointly observed flashes, it should suffice for a rough estimate.

These estimates are presented in Table 5. The bulk of OTD data (from 20 July 1995 to present) was collected under 8-bit thresholds 17 and 15. At these settings, "IC":CG ratios of 2.6–5.2 are found in the OTD data. In comparison, Mackerras et al. (1998) directly measured the IC/CG ratio at eight sites worldwide at latitudes between 23° and 60° N or S (comparable to U.S. latitudes); their measurements yielded IC/CG ratios of 2.77 ± 1.05. These estimates are shown in Fig. 10, along with earlier estimates by Pierce (1970) and Prentice and Mackerras (1977). The OTD results are consistent with these estimates (which may themselves be biased), though slightly higher. Thus, any differences in OTD IC and CG detection efficiency are likely small, and if present, favor a higher IC detection efficiency (consistent with both physical intuition and the results of Thomas et al. 1999). The use of CG DE estimates for IC flashes (and hence total lightning) thus appears, to first order, to be a reasonable (and conservative) working assumption. We next consider the impacts of this assumption.

c. OTD bulk detection efficiency

We may use the notation of section 8a to determine the maximum error likely arising from our final operational assumption that $\mathrm{DE}_{OTD,CG}$ and $\mathrm{DE}_{OTD,IC}$ are equivalent. Using the definitions of k and z above, we may write the bulk (total lightning) OTD DE as

(a) No NLDN IC contamination

(b) With NLDN IC contamination

(c) No NLDN IC contamination

(d) With NLDN IC contamination
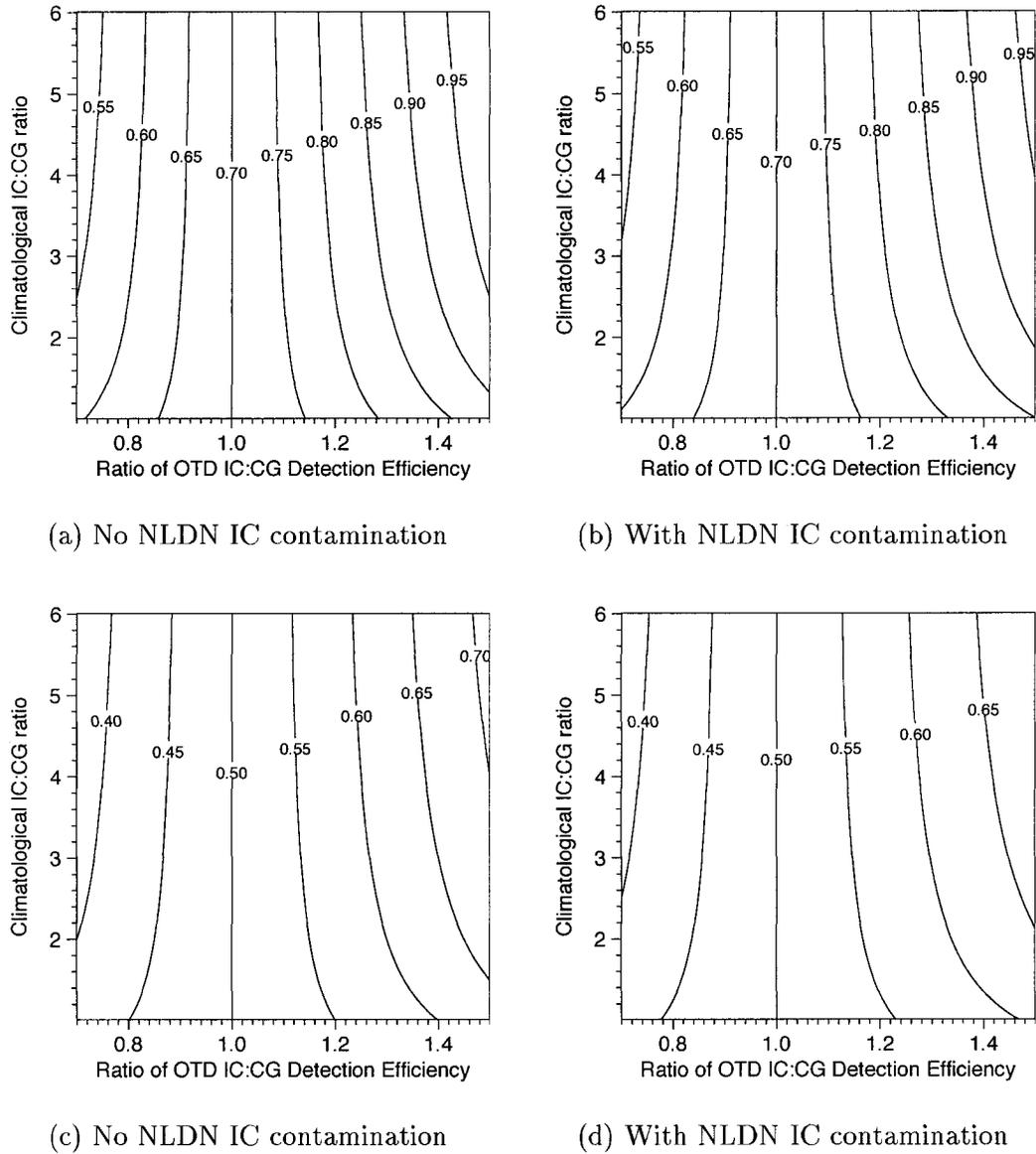
FIG. 11. Bulk OTD detection efficiency, assuming an OTD CG detection efficiency of (a), (b) 0.7 and (c), (d) 0.5 for various values of k and z. The dependence of bulk DE on z is relatively small, with a much stronger dependence on k. For the most likely values of k ($k > 1$) the simple assumption of OTD bulk DE equal to OTD CG DE underestimates the true bulk detection efficiency (i.e., OTD-inferred regional flash rates will be too high).

$$\mathrm{DE}_{\mathrm{OTD,bulk}} = \frac{\mathrm{DE}_{\mathrm{OTD,CG}} N_{\mathrm{CG,true}} + \mathrm{DE}_{\mathrm{OTD,IC}} N_{\mathrm{IC,true}}}{N_{\mathrm{CG,true}} + N_{\mathrm{IC,true}}}$$

$$= \frac{\mathrm{DE}_{\mathrm{OTD,CG}} N_{\mathrm{CG,true}} + k\mathrm{DE}_{\mathrm{OTD,CG}} z N_{\mathrm{CG,true}}}{N_{\mathrm{CG,true}} + z N_{\mathrm{CG,true}}}$$

$$= \mathrm{DE}_{\mathrm{OTD,CG}} \frac{1 + kz}{1 + z}. \qquad (16)$$

Under an assumption of $k = 1$, $\mathrm{DE}_{\mathrm{OTD,bulk}}$ is of course equal to the $\mathrm{DE}_{\mathrm{OTD,CG}}$. If we choose example inferred $\mathrm{DE}_{\mathrm{OTD,CG}}$ from Fig. 8 of 0.7 and 0.5, we may compare the "true" bulk OTD DE under various assumptions on

k and z; the results are shown in Figs. 11a,c for no NLDN IC contamination and Figs. 11b,d for worst-case IC contamination. There is very little dependence of the bulk OTD DE on plausible values of z, but (unsurprisingly) a modest dependence of $\mathrm{DE}_{\mathrm{OTD,bulk}}$ on k. For the most likely case of $k > 1$, the expected result is obtained, that is, that the (example) operational 0.7 $\mathrm{DE}_{\mathrm{OTD,bulk}}$ underestimates the true $\mathrm{DE}_{\mathrm{OTD,bulk}}$ (i.e., inferred OTD regional flash rates will be too high, here by up to 20%).

We reiterate that this analysis holds only under the assumption of perfect pairing capability. It also does not account for the small number of CG flashes that NLDN may miss but OTD may see (these are unidentifiable

with the present data). These error estimates are almost certainly worst-case bounds on the uncertainty introduced by IC contamination of the validation dataset and unknown OTD IC detection efficiency. Failure of the perfect pairing assumption probably means that these errors overlap with the uncertainty introduced by space–time pairing errors; since they are comparable or smaller than this uncertainty, we conclude that they do not measurably alter the primary conclusions of this work. If applied to a sensor with much higher localization accuracy (such as the LIS), the effects may be more significant.

d. Incorrect OTD/NLDN flash pairing

The limited spatial resolution and accuracy of the OTD sensor, and the inherent uncertainty in temporally isolating the return stroke among all OTD optical pulses in a flash, lead to the possibility that some OTD and NLDN flashes in this study are incorrectly paired; this could lead to errors in the spatial accuracy inferences and an overestimation of the operational CG detection efficiency. In this section, we attempt to estimate the severity of these potential errors.

In this study, incorrect pairing is possible when two or more (truly disjoint) flashes are occurring concurrently in time and close in space. The pairing algorithm assumes higher OTD temporal, rather than spatial, accuracy, and when multiple OTD flashes are candidates for pairing with an NLDN CG, the closest flash in time is assigned. Incorrect pairing can also occur when two concurrent flashes (one of them the NLDN CG) happen but only one (an IC or another CG) is observed by the OTD. In order to assess how often this possibility exists, we must thus estimate the frequency of true flash temporal overlaps in nature.

To first order, we can provide a bootstrap estimate using the OTD itself (without consideration of the NLDN). This is possible because the base estimate of OTD detection efficiency is so high; we can reasonably assume that we are sampling a wide portion of the true spectrum of lightning optical "intensities" (discretized as collections of optical pulses), that overlaps are infrequent, and that an instrument with higher detection efficiency would not observe significantly more overlaps. We construct two methods for identifying flash overlaps.

1) Two flashes within the FOV overlap if any part of the time intervals between their first and last optical pulses overlaps.
2) Each flash is assigned a nominal duration of 600 ms and the nominal time as that of its brightest group, and determine overlaps from these artificial time windows.

Approach 1 is the "formal" overlap definition for the OTD data, but likely will underestimate true overlaps as the OTD probably does not observe all optical pulses

in every flash. Approach 2 is much more conservative and consistent with the maximum time offset (600 ms) and flash nominal time assignments used in section 4. Under approach 1, and using the entire quality-controlled OTD dataset to date, we find that flashes temporally overlap anywhere within the OTD FOV 16.1% of the time; under approach 2, there is overlap 41.2% of the time.

Of this "background" level of overlaps anywhere within the FOV, only those within 200 km of each other would have been candidates for incorrect pairing under the criteria used in this study. Additionally, for the high end of DE estimates, only those within 600 ms of each other would have been candidates. The subset of these overlaps meeting these criteria is 5.6% for approach 1 and 15.9% for approach 2. For the lower end of DE estimates, 300 ms is the cutoff and the subsets are 4.6% and 6.6% of all flashes.

Further, of these "true" overlaps, we are only concerned with overlaps involving at least one CG. The fraction of all overlaps with at least one CG is given directly if we assume a value for z; for z of 2, 3, 4, and 5, the fractions are 5/9, 7/16, 9/25, and 11/36, respectively. Choosing a conservative z of 3, the appropriate fraction is thus 0.44, and the percentages above need to be reduced by this, yielding values of 2.0%–2.5% for approach 1, and 2.9%–7.0% for approach 2. These are estimates of the maximum occurrence of *candidates* for incorrect pairing in this study. They assume no skill of the pairing algorithm to discriminate between flashes based on their nominal times, an overly conservative assumption. Assuming that 50% of these will be correct pairings, the relative occurrence of false pairs is thus between 1.0%–1.5% for the low end of DE estimates and between 1.5%–3.5% for the high end. These represent an overestimation of the true $DE_{OTD,CG}$. However, there is also a small but unknown number of true OTD CG observations that are excluded from consideration because their location accuracy errors exceed 200 km; these would contribute to an underestimation of the true detection efficiency using this methodology. We thus conclude that erroneous pairing is not a significant source of error or bias in these results.

e. OTD flash counting

A final source of concern is the possibility that the automated OTD grouping algorithm, which assembles observed optical pulses into flashes, systematically and incorrectly merges pulses from two adjacent but separate channel structures, or fragments pulses from a single flash into multiple flash assignments, thus systematically undercounting or overcounting true flashes. We can use the flash overlap occurrence data from the previous section to address this possibility. Again using the entire OTD dataset to date, we determine the range and time separation for each overlap occurrence using approach 1 (see above), and bin the occurrences into a
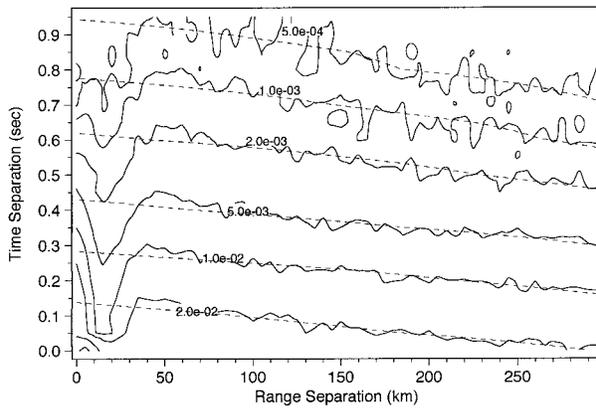
FIG. 12. Percent occurrence of flashes that overlap in time, taken from the entire OTD dataset to date. The data are binned to a 5 km by 50 ms grid, with the space and time dimensions representing the range and time separation between overlapping flashes. Solid contours represent the percent of all OTD flashes that fall into a given grid cell. Dashed contours represent the best fit to all data with range separation greater than 75 km, assumed to be truly distinct flashes correctly identified by the OTD automated pulse grouping algorithm.

TABLE 6. Errors in the OTD automated pulse grouping algorithm, derived from extrapolation of the expected number of true flash temporal overlaps. Results are presented as a percentage of error occurrence among all OTD-observed flashes. "Excess" flashes denote fragmentation of pulses from the same true flash into two or more OTD-reported flashes; "depletion" of flashes denotes merging of pulses from two or more distinct true flashes into one OTD-reported flash.

Max range separation (km)	Max time separation (s)	Flash overlap "excess" (%)	Flash overlap "depletion" (%)
25	0.5	0.255	−0.205
25	1.0	0.255	−0.241
50	0.5	0.330	−0.262
50	1.0	0.334	−0.316

density grid with cells of size 5 km by 50 ms. The results are shown in Fig. 12, in which the solid contours denote the frequency of occurrence of overlaps within the entire dataset (e.g., the grid cell with a range separation coordinate of 50 km and a time sepearation coordinate of 0.15 second falling on the 0.02 contour line represents 0.02% of the entire OTD dataset).

Since these are overlapping flashes in time and the OTD navigation is stable on timescales of seconds, absolute OTD pointing accuracy is not a consideration here. For overlapping flashes beyond, for example, 75-km range separation, it is safe to assume that most have been correctly grouped by the pulse clustering algorithm, as contiguous channel structures larger than this scale are rare. We note that the contours in this part of the parameter space are fairly stable and exhibit a slight "upward slope," perhaps related to a true tendency of lightning (or deep convection) to cluster on smaller and smaller spatial scales. Below 40-km separation, the contours of overlap occurrence experience an unusual "dip"; this decrease persists down to the OTD pixel resolution scale. Since there is no reason to expect that deep convection truly exhibits a unique tendency at these scales not to cluster, and since these scales are equivalent to the distance parameter used by the flash grouping algorithm, we may assume that this decrease represents the inability of the grouping algorithm to properly separate temporally overlapping flashes that occur near these spatial separation scales. Such flashes will be undercounted in the final data, as the algorithm (in absence of higher resolution information) is forced to identify them as single flashes. This effect is slightly offset by a sharp increase in overlapping flashes at the very smallest separations (less than 50-ms temporal and 5-km spatial separation). These represent cases where

a single pulse group of a larger flash is misidentified by the algorithm as a unique flash (hence resulting in a flash overcounting).

We can make a preliminary estimate of the severity of these algorithmic errors by assuming that the true overlap occurrence should be given by an extrapolation of the occurrence contours beyond 75-km spatial separation to smaller spatial scales (the dashed contours in Fig. 12). This is equivalent to assuming that the clustering tendency of deep convective cells increases at about the same rate as the spatial scale decreases. This clustering tendency is of course unknown and could plausibly rise sharply at cell separations below 75 km. It is, however, highly unlikely that it will *decrease,* and hence the extrapolated behavior should represent a minimum bound on the deviation of OTD-counted flash overlap occurrence from true flash overlaps. By summing all the differences between the observed and expected overlap occurrences at small spatial and temporal separations, we may thus estimate the minimum amount of flash counting bias introduced by the grouping algorithm. This difference, shown for four sample separation bounds, is presented in Table 6, again expressed as a percentage of all observed OTD flashes. It is clear from these results that while flash counting errors do occur, they are extremely infrequent, and flash overcounting at very small separations partially offsets flash undercounting at scales near the OTD pixel resolution, leading to a negligible net effect. While these estimates are minimum bounds, we note that for flash counting errors to be significant, the true deep convective clustering behavior would have to be significantly greater than estimated here (i.e., the dashed contours would have to rise very sharply at scales below 40 km). We thus tentatively conclude that our best estimates of bulk flash counting errors suggest that these represent a negligible fraction of all OTD flash observations.

9. Summary

The pixel spatial resolution of the OTD sensor, as deployed, ranges from 8- to 24-km ground range, with a mean/median of 11.3 km. Based upon intercompari-

sons with the NLDN truth dataset, errors in the *Microlab-1* ephemeris and attitude data degrade this and result in a 20–40-km overall spatial accuracy, with up to several hundred kilometers of spatial errors possible on rare occasions. The temporal resolution of the sensor is 2 ms, but errors in the subsecond onboard clock may further reduce the overall accuracy of flash times. This reduction in accuracy is not determinable from the current dataset, although for a subset of paired flashes apparent timing errors are well below 100 ms, and for some flashes below 10 ms. The flash radiance calibrations are a function of sensor quadrant and background radiance levels. Accuracy of the calibrated radiances also varies with these parameters and with radiance magnitude, may range from 5% to 75%, and is worst for very low (near threshold) and very high (near saturation) radiances. These errors must be interpreted, however, in the context of the sensor's 30-dB dynamic range.

The cloud-to-ground flash detection efficiency of the instrument, for 8-bit threshold settings 15 and 17 (comprising the bulk of the data to date) is estimated at 62% ± 7% and 56% ± 10%, respectively. These estimates are consistent with laboratory-based estimates by Koshak et al. (1996, 2000), after allowing for a slight decrease due to operational signal processing (filtering). Inferred IC:CG ratios over the United States seem higher than previous estimates for comparable latitudes, and hence we infer slightly higher IC detection efficiency [consistent with the results of Thomas et al. (1999)]. On average, the detection efficiency appears about 10%–20% higher for positive than for negative CG flashes (although this may be aliasing of IC contamination in the NLDN data), and 4%–15% higher for nighttime than for daytime flashes. These biases may be considered negligible in total lightning estimates from the instrument. The false alarm rate in the distributed data is negligible, and these data are more likely overfiltered than underfiltered. Because of the orbital precession of *Microlab-1*, 55- or 110-day averaging must be performed on composite data to remove bias (aliasing) from the diurnal cycle in the underlying lightning distribution. The OTD data, browse images, interface software, and initial results are freely available for order online at http://thunder.msfc.nasa.gov.

REFERENCES

Asrar, G., and R. Greenstone, 1995: MTPE EOS Reference Handbook. NASA Goddard Space Flight Center Tech. Rep., 277 pp.

Bell, T., 1987: A space–time stochastic model of rainfall for satellite remote-sensing studies. *J. Geophys. Res.,* **92,** 9631–9643.

——, A. Abdullah, R. Martin, and G. North, 1990: Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space–time stochastic model. *J. Geophys. Res.,* **95,** 2195–2205.

Boccippio, D., C. Wong, E. Williams, R. Boldi, H. Christian, and S. Goodman, 1998: Global validation of single-station Schumann resonance lightning location. *J. Atmos. Sol. Terr. Phys.,* **60,** 701–712.

——, S. Heckman, and S. Goodman, 1999: A diagnostic analysis of the Kennedy Space Center LDAR network. *Proc. 11th Int. Conf. on Atmospheric Electricity,* Guntersville, AL, ICAE, 254–257.

Christian, H., 1999: Optical detection of lightning from space. *Proc. 11th Int. Conf. on Atmospheric Electricity,* Guntersville, AL, ICAE, 715–718.

——, and S. Goodman, 1987: Optical observations of lightning from a high altitude airplane. *J. Atmos. Oceanic Technol.,* **4,** 701–711.

——, R. Blakeslee, and S. Goodman, 1989: The detection of lightning from geostationary orbit. *J. Geophys. Res.,* **94,** 13 329–13 337.

——, ——, and ——, 1992: Lightning Imaging Sensor (LIS) for Earth Observing System. Tech. Rep. NASA TM 4350, 36 pp.

——, K. Driscoll, S. Goodman, R. Blakeslee, D. Mach, and D. Buechler, 1996: The Optical Transient Detector (OTD). *Proc. 10th Int. Conf. on Atmospheric Electricity,* Osaka, Japan, ICAE, 368–371.

——, and Coauthors, 1999a: The Lightning Imaging Sensor. *Proc. 11th Int. Conf. on Atmospheric Electricity,* Guntersville, AL, ICAE, 746–749.

——, and Coauthors, 1999b: Global frequency and distribution of lightning as observed by the Optical Transient Detector (OTD). *Proc. 11th Int. Conf. on Atmospheric Electricity,* Guntersville, AL, ICAE, 726–729.

Cummins, K., M. Murphy, E. Bardo, W. Hiscox, R. Pyle, and A. Pifer, 1998: A combined TOA/MDF technology upgrade of the U.S. National Lightning Detection Network. *J. Geophys. Res.,* **103,** 9035–9044.

Goodman, S., H. Christian, and W. Rust, 1988: A comparison of the optical pulse characteristics of intracloud and cloud-to-ground lightning as observed above clouds. *J. Appl. Meteor.,* **27,** 1369–1381.

Koshak, W., R. Solakiewicz, D. Phanord, and R. Blakeslee, 1994: Diffusion model for lightning radiative transfer. *J. Geophys. Res.,* **99,** 14 361–14 371.

——, J. Bergstrom, M. Stewart, H. Christian, J. Hall, and R. Solakiewicz, 1996: Calibration of the Optical Transient Detector (OTD). *Proc. 10th Int. Conf. on Atmospheric Electricity,* Osaka, Japan, ICAE, 364–367.

——, ——, ——, ——, ——, and ——, 2000: Laboratory calibration of the Optical Transient Detector (OTD) and the Lightning Imaging Sensor (LIS). *J. Atmos. Oceanic Technol.,* in press.

Krehbiel, P., 1986: The electrical structure of thunderstorms. *The Earth's Electrical Environment,* National Academy Press, 90–113.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.,* **15,** 809–817.

Mackerras, D., M. Darveniza, R. Orville, E. Williams, and S. Goodman, 1998: Global lightning: Total, cloud and ground flash estimates. *J. Geophys. Res.,* **103,** 19 791–19 809.

North, G., S. Shen, and R. Upson, 1993: Sampling errors in rainfall estimates by multiple satellites. *J. Appl. Meteor.,* **32,** 399–410.

Pierce, E., 1970: Latitudinal variation of lightning parameters. *J. Appl. Meteor.* **9,** 194–195.

Pinto, O., W. Gonzalez, R. Pinto, A. Gonzalez, and O. Mendes, 1992: The South Atlantic Magnetic Anomaly: Three decades of research. *J. Atmos. Terr. Phys.,* **54,** 1129–1134.

Prentice, S., and D. Mackerras, 1977: The ratio of cloud to cloud-ground lightning flashes in thunderstorms. *J. Appl. Meteor.,* **16,** 545–549.

Salby, M., 1982: Sampling theory for asynoptic satellite observations. Part I: Space–time spectra, resolution, and aliasing. *J. Atmos. Sci.,* **39,** 2577–2600.

——, and P. Callaghan, 1997: Sampling error in climate properties derived from satellite measurements: Consequences of undersampled diurnal variability. *J. Climate,* **10,** 18–36.

Thomas, R., P. Krehbiel, W. Rison, T. Hamlin, D. Boccippio, S. Goodman, and H. Christian, 1999: Comparison of ground-based, three-dimensional lightning mapping observations with satellite-based LIS observations in Oklahoma. *Proc. 11th Int. Conf. on Atmospheric Electricity,* Guntersville, AL, ICAE, 172–175.

Wacker, R., and R. Orville, 1999a: Changes in measured lightning flash count and return stroke peak current after the 1994 U.S. National Lightning Detection Network upgrade: I. Observations. *J. Geophys. Res.,* **104,** 2151–2157.

——, and ——, 1999b: Changes in measured lightning count and return stroke peak current after the 1994 U.S. National Lightning Detection Network upgrade: II. Theory. *J. Geophys. Res.,* **104,** 2159–2162.