

Latitudinal Variation of Bolide Flux as Detected by GOES GLM

Anthony Ozerov¹ and Jeffrey C. Smith^{2,3}

¹Columbia University, New York, NY

²SETI Institute, Mountain View, CA

³NASA Ames Research Center, Moffett Field, CA

Abstract

The large combined field of view of the Geostationary Lightning Mapper (GLM) instruments on the GOES satellites makes them useful for studying bolides (bright meteors). The thousands of GLM bolide detections allow us to empirically investigate the latitudinal variation of bolide flux. We discretize the data using a randomized method and develop a Bayesian Poisson regression model which simultaneously estimates GLM bolide detection biases and the latitudinal flux variation. The estimated bias due to the angle of incident light upon the instrument corresponds roughly with the previously measured sensitivity of the GLM instruments. We compare our latitudinal flux variation estimates—including for particular meteor showers—to existing theoretical models and discuss how much a bias due to bolide impact velocities may affect estimates derived from GLM data.

Contents

1	Introduction	2
2	Methods	3
2.1	Data	3
2.2	Poisson Regression	4
2.3	Nonparametrics	5
2.4	Discretization	6
2.5	Bayesian representation	7
2.6	Theoretical computations	8
3	Results	8
4	Discussion	10
4.1	Computed distribution and bias	10
4.2	Comparison to previous empirical calculations	11
4.3	Comparison to previous theoretical calculations	11
4.4	Impossibility of completely debiasing the distribution	12
4.5	Other approaches to probing biases	13
4.5.1	Diurnal variation	13
4.5.2	Sensitivity to meteor showers	14
4.6	Future work	15
5	Conclusion	15
6	Acknowledgments	15
A	Supplementary Figures	17

1 Introduction

It is well-established that asteroids do not necessarily impact the Earth uniformly (Robertson et al., 2021). Based on the distribution of orbits of Near-Earth Objects (NEOs), particularly the ecliptic latitude at the point where they cross the Earth’s orbit and their velocity relative to the Earth, the distribution of impacts across latitudes can be estimated. This has been done, but different results have been obtained based on the NEO databases and methods used (Robertson et al., 2021; Le Feuvre and Wieczorek, 2008). In terms of empirically verifying the estimates, the two main methods used in the literature are cratering records and impact detections. Cratering records are applicable to bodies like the Moon (Robertson et al., 2021; Le Feuvre and Wieczorek, 2008), but are difficult to employ on the Earth. For impact detections, there are several possible sources of data.

US Government sensors Previous work has used impact detections from United States Government (USG) sensors (Robertson et al., 2021; Evatt et al., 2020). This set of data is appealing as it has global coverage. However, due to the classified nature of the sensors, the detection biases of this data set are unknown. Due to the other priorities of these sensors, they may not necessarily have a uniform detection efficiency across different latitudes. For instance, the sensors might generally have a lower detection efficiency for objects hitting the poles (which would explain all of the discrepancy described in Robertson et al., 2021). This makes it impossible to use this data to probe the distribution of bolide impacts across latitudes.

Ground-based systems Automated ground-based photo, video, and radar meteor observation stations detect thousands of meteor impacts daily. Yet they mostly detect small objects from known meteor showers, and their limited coverage in area makes it more difficult to study the distribution of rarer, larger objects of greater concern for planetary defense. Moreover, some systems are unable to detect objects during the day, and there are other factors like weather and the specifics of a given system that may nonuniformly affect detection efficiency across different latitudes.

Geostationary Lightning Mapper Data from the Geostationary Lightning Mapper (GLM) instruments aboard the Geostationary Operational Environmental Satellites (GOES), which are known to detect bolides, have also been used (Smith et al., 2021; Robertson et al., 2021). This data set is appealing as it has near-hemispheric coverage and much more is known about the sensors than with the US Government data. In previous studies, only human-vetted data, which has a bias towards the stereo region, has been used. Other biases, like the angle of incident light upon the sensor, have not previously been accounted for. This paper seeks to account for these biases and produce a more accurate estimate of the variation in bolide impacts across latitudes.

2 Methods

2.1 Data

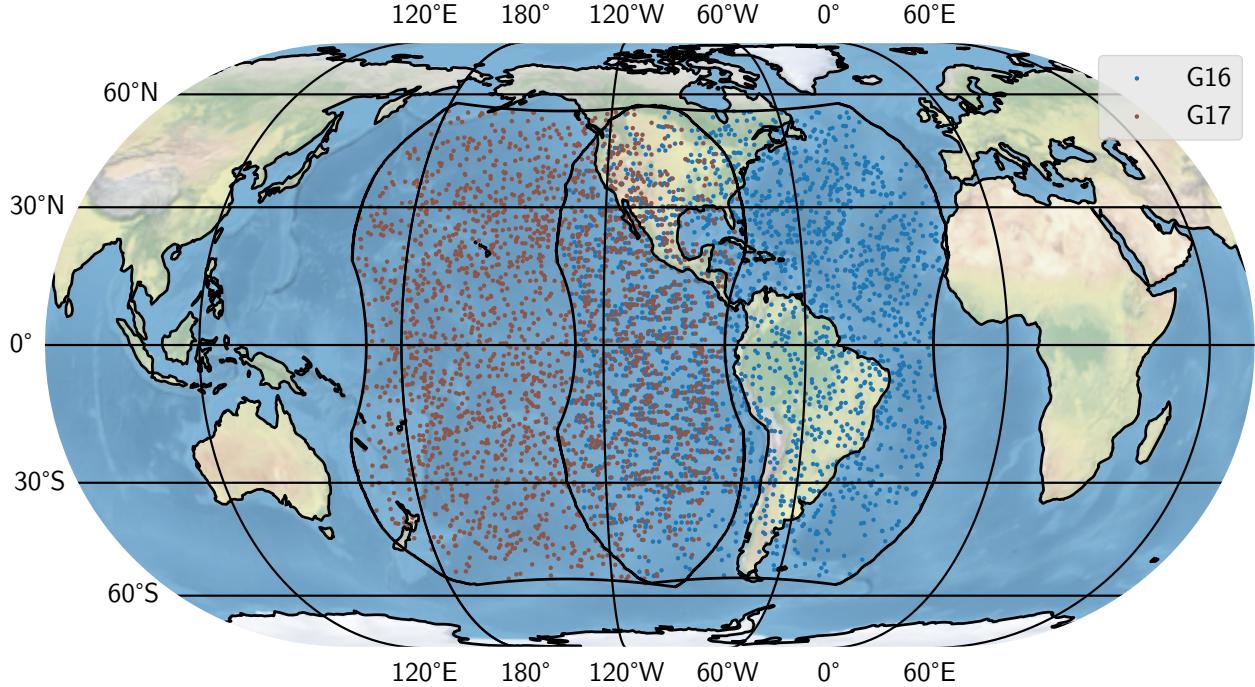


Figure 1: All detections in the database used, colored by which satellite detected them. There are double-detections in the stereo region, where the same object was detected by both satellites. The boundaries are the two platforms’ FOV boundaries. The FOV boundaries closely match the limits of the bolide detections. Eckert IV projection. Figure generated using `bolides` (Ozerov and Smith, n.d.).

We use bolide detections from the GLM bolide detection pipeline (Smith et al., 2021) run for days between 2019-07-01 and 2022-06-03. We only keep data with a confidence threshold of ≥ 0.7 , as at that threshold the false positive rate is evenly distributed over the field of view—with lower thresholds, glint and lightning activity introduce a bias. For the purposes of this investigation we assume that all of these detections are true positives. The pipeline also runs separately for each satellite, i.e. each satellite makes its detections independently. This means that, in the data used here, some detections in the stereo region (where both satellites may detect an event) appear twice—once for each satellite—which is important as we treat each satellite as an independent detector. Using the unvetted pipeline data reduces the human selection bias, which overemphasizes the stereo region (as a stereo detection is a very confident one for a human vettor). Due to the shape of the stereo region—which is not a uniform proportion of the total GLM FOV across latitudes—previous attempts to use GLM data to study the latitudinal distribution of impacts likely exhibit artifacts of this human bias. We do note, though, that even though we study unvetted data, the underlying classifier is still trained on human-vetted data and may exhibit some bias as a result of this.

2.2 Poisson Regression

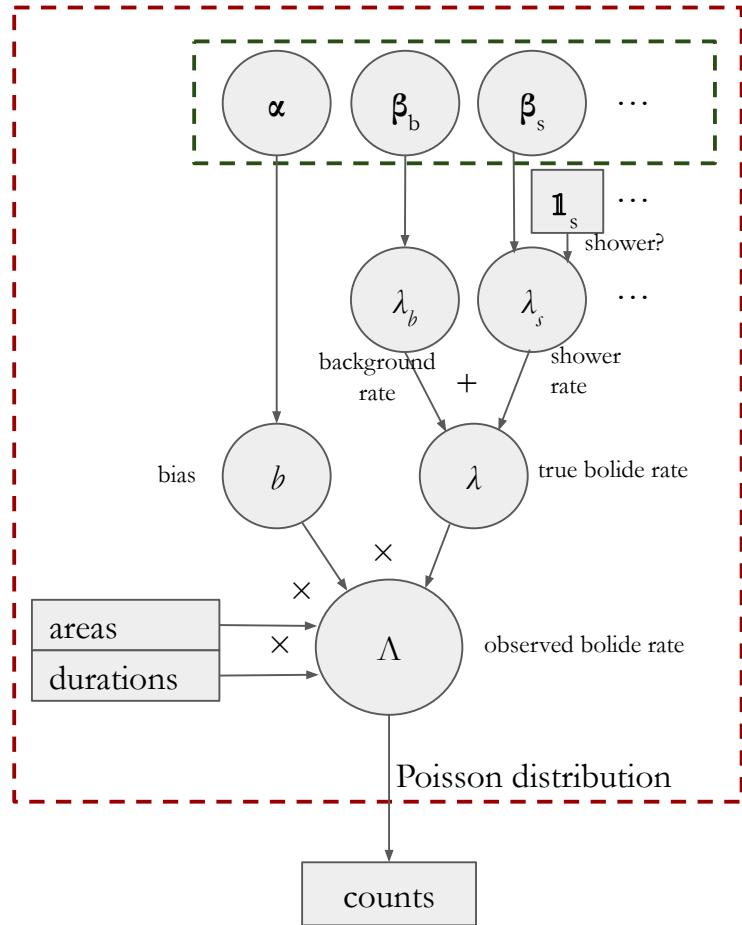


Figure 2: Diagram of the hierarchical Bayesian Poisson regression model. Outlined in red is the data-generating process. Outlined in green are the vectors of coefficients to be estimated. Areas, durations, and counts come from the discretization discussed below.

Since we are investigating bolide flux as a function of latitude, with distance from nadir accounted for, we can model the observed bolides as coming from a Poisson distribution with an inhomogeneous rate. The Poisson distribution comes with the assumption that impacts are not autocorrelated. The data do not show any apparent clustering in both space and time (as they would if, say, an object broke up just before reaching the Earth), so we accept this assumption. The model can then be expressed as:

$$y(A, T) \sim \text{Poisson} \left(\int_T \int_A \Lambda(d, l, t) dd dl dt \right),$$

where $y(A, T)$ is the number of bolides in region A (any set of latitudes and longitudes) over a set of times T , and $\Lambda(d, l, t)$ is the inhomogeneous rate function (in terms of d , the distance from the nadir, l , the latitude, and t , time) of the Poisson distribution that generates $y(A, T)$. Since we want to simultaneously estimate the bias due to the angle of incident light upon the sensor and the variation over different latitudes for the base rate and the targeted showers, $\Lambda(d, l, t)$ is a product of two terms:

$$\Lambda(d, l, t) = b(d) \sum_{s \in S(t)} \lambda_s(l),$$

where $b(d)$ is a multiplier due to the angle of incident light upon the sensor and $S(t)$ is a set of bolide sources active during time t , which always includes the background rate and may include a meteor shower. $\lambda_s(l)$ is a function representing the rate attributed to a particular bolide source, as a function of latitude. $b(d)$ is defined as:

$$b(d) = \exp(\alpha_0 + \alpha_1 d + \alpha_2 d^2 + \alpha_3 d^3),$$

where the α 's are estimated coefficients and d is the distance, in kilometers, from the nadir of the corresponding satellite.

$\lambda_s(l)$ is defined as:

$$\lambda_s(l) = \exp(\beta_{s,0} + \beta_{s,1}l + \beta_{s,2}l^2 + \beta_{s,3}|l^3|),$$

where the β_s 's are estimated coefficients for source s and l is latitude.

Expressing the model in this way allows for the simultaneous estimation of the detection bias as a function of distance from nadir and the additive rates due to different meteor showers and the background rate. Multiplying the detection bias by the sum of the rates gives the total observed rate of GLM.

2.3 Nonparametrics

To confirm that the efficient and more interpretable parametric method is able to recover the distributions, we also perform a fit where the bias term and functions of latitude are instead exponentiated Gaussian processes:

$$\begin{aligned} b(d) &= \exp(f_b(d)); \quad f_b(x) \sim GP(m(x), k(x, x')) \\ \lambda_s(l) &= \exp(f_s(l)); \quad f_s(x) \sim GP(m(x), k(x, x')) \end{aligned}$$

We use mean function of zero and the Exponentiated Quadratic kernel:

$$k(x, x') = \exp\left(-\frac{(x - x')^2}{2\ell^2}\right),$$

where $\ell \sim Exp(1)$. We choose the Exponentiated Quadratic kernel as it makes for smooth results. There is no reason to expect the true latitudinal distribution to be jagged.

2.4 Discretization

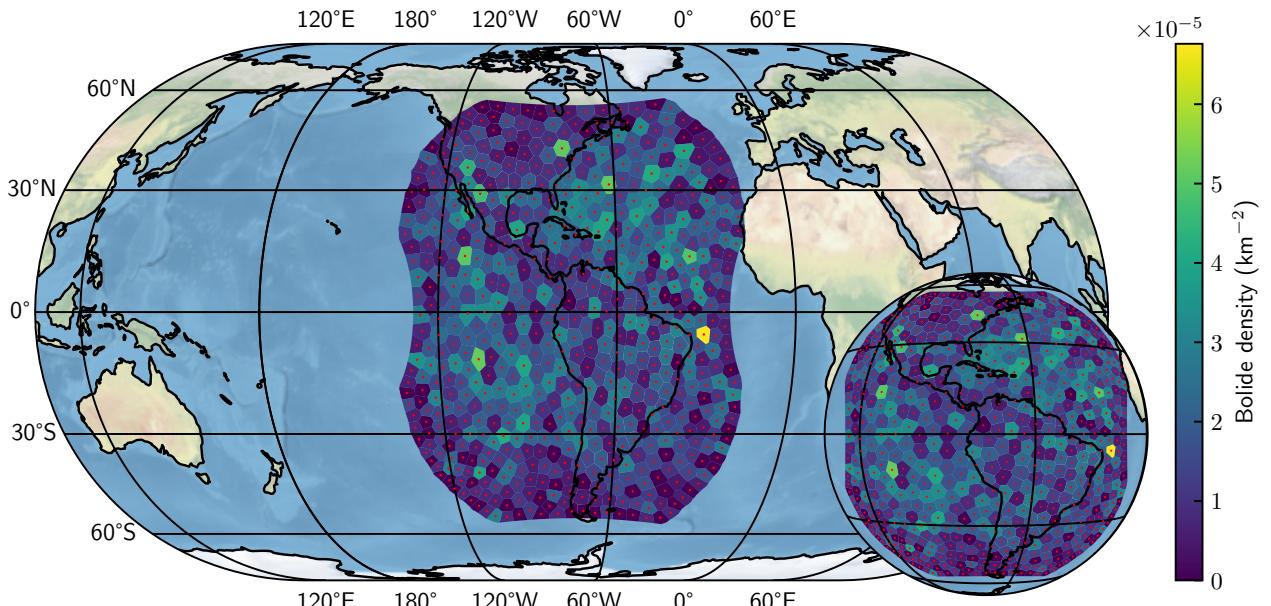


Figure 3: Sample discretization of GOES-16’s bolide detections. The field of view is split into small polygons, and the density of bolides in each polygon is shown. The red points are the centroids of the polygons, which are used to compute the latitude, longitude, and distance from nadir. The main map is in the Eckert IV projection centered on GOES-16’s nadir. The inset globe is from the perspective of GOES-16 (the GLM field of view, though it takes on an odd shape in most projections, is really a square with the corners masked).

As we use a spatial Poisson regression model, which has a computationally intractable likelihood function, we must discretize the point data into a set of counts. It is possible to simply grid the data, but we instead use a randomized method which can be repeated and shown to be robust. This allows us to sample from the model multiple times, with different random polygons, and be confident that an arbitrary gridding is not affecting the results. For each satellite (GOES-16 and GOES-17), we:

1. Get a uniformly random set P of points within the satellite’s collecting area in the Azimuthal Equal Area coordinate reference system centered at the satellite’s nadir. *Rationale:* An equal area projection means that there is no systematic effect on polygon size of latitude, longitude, or distance from nadir. The rotational symmetry of the Azimuthal Equal Area projection in particular ensures that there are no differences between the general shapes of the polygons made at the extreme edge of the FOV in longitude and at the extreme edge in latitude.
2. Repeat the following:
 - (a) Compute the Voronoi diagram V from P using the Qhull library (Virtanen et al., 2020; Barber C.B. and H.T., 1996).
 - (b) Let P be the centroids (geometric centers-of-mass) of the polygons in V , as computed in the Azimuthal Equal Area projection.

Rationale: This iterative algorithm approaches a centroidal Voronoi diagram, which creates polygons that are uniform in size. (Du, Faber, and Gunzburger, 1999) In principle, it is not actually necessary to use polygons of uniform size, but doing so guarantees a fine resolution across the field of view with fewer polygons.

3. Clip the polygons of the Voronoi diagram to the sensor’s FOV. *Rationale:* This ensures that the shape of the FOV is taken into account, and e.g. the lack of observations in the

extreme North and South just outside the FOV isn't taken to imply a decrease in flux at high latitudes.

4. Count the number of detections within each polygon.
5. Compute the area of each polygon.
6. Compute the “observation duration” of each polygon. This is 1 for most, but for those which are only observed for half of the year by GOES-17 this is $1 - a/2$, where a is the proportion of the polygon which is only in GOES-17’s inverted FOV or non-inverted FOV (i.e. the proportion of the polygon within the symmetric difference of the inverted and non-inverted FOV). *Rationale:* This ensures that fewer detections in the regions where GOES-17 only observes for half of the year, which also happen to be at extreme latitudes, are not taken to imply a decrease in flux at high latitudes.
7. Compute the latitude, longitude, and distance from nadir of each polygon according to its centroid.

This discretization works because, within a small polygon, the distance from the satellite’s nadir and the latitude will vary little, and Λ will be roughly constant. Thus we get the relation that:

$$\int_A \Lambda(d, l, t) dd dl \approx a_A \Lambda(d_A, l_A, t)$$

Where a_A is the area of polygon A , d_A is the distance of polygon A ’s centroid from the nadir, and l_A is the latitude of the centroid.

For additionally discretizing time, since we assume that bolide sources are either active or not active given a time, we see that

$$\int_T a_A \cdot \Lambda(d_A, l_A, t) dt = \mu(T) \Lambda(d_A, l_A, t_0),$$

as long as $S(t) = S(t_0) \forall t \in T$. We thus group the data into bolides within each target meteor shower, and bolides not associated with any of the target meteor showers. For every shower, and for the base rate, we do the process above, and then add S new indicator variables to the polygons indicating which showers they belong to. We are thus left with a table of polygons with the following variables:

- Count
- Latitude ($^{\circ}$)
- Distance from nadir (km)
- Area (km^2)
- Duration (arbitrary units)
- Is [shower 1]? (0/1)
- Is [shower 2]? (0/1)
- ...
- Is [shower n]? (0/1)

2.5 Bayesian representation

The above model could be fit by simply selecting the α ’s and β ’s to maximize the likelihood of the observed data. However, this would not allow for proper uncertainty quantification. Instead, we specify priors on the the α ’s and β ’s as Normal distributions centered on zero with a standard deviation large enough to be non-informative in the parameterization of the variables that is used when fitting. We then sample the posterior distributions of the model using the No-U-Turn Sampler (Hoffman, Gelman, et al., 2014) in PyMC (Salvatier, Wiecki, and Fonnesbeck, 2016).

2.6 Theoretical computations

To contextualize and understand the empirical results, we also expand upon the theoretical calculations of latitudinal impact distributions performed in previous work by providing calculations for given radiant velocities and specific meteor showers. We do both using the method of (Robertson et al., 2021). For given radiant velocities, rather than performing Monte Carlo calculations using the joint velocity and ecliptic latitude distribution, we use the conditional distribution of ecliptic latitude given a velocity. For specific meteor showers, we perform the Monte Carlo calculations starting with the given meteor shower’s radiant and radiant velocity. Since the assumption of North-South impact symmetry no longer applies for specific meteor showers as it does for the total NEO population, we do not fold the distribution.

3 Results

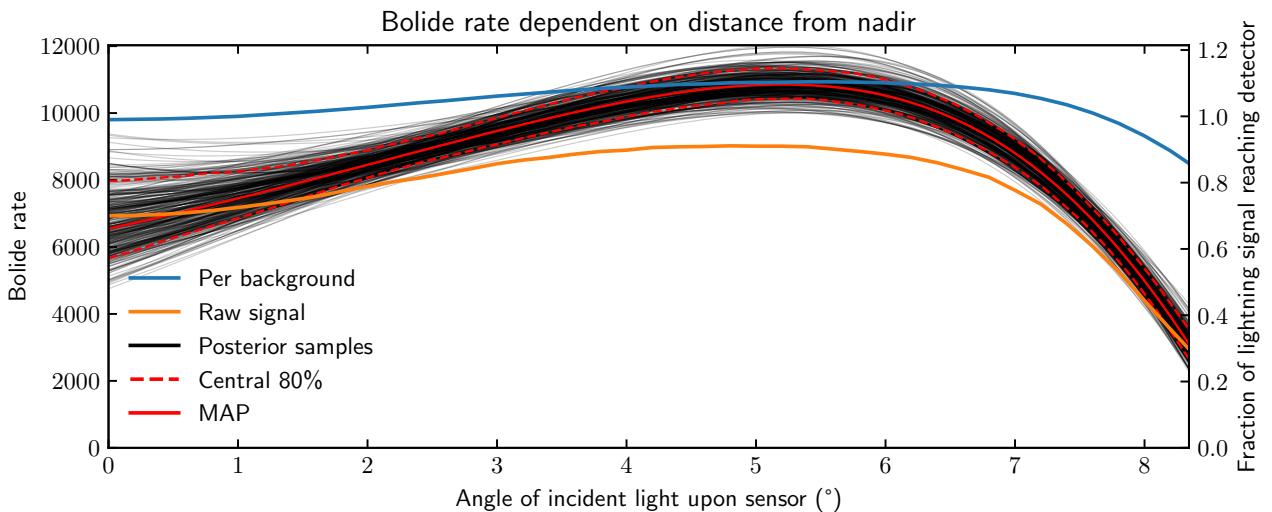


Figure 4: Bolide rate as a function of the angle of incident light upon the sensor, along with curves representing unpublished sensor performance data. Rate is per unit area per unit time. The regression was done as a function of distance from nadir, but this figure reparameterizes $b(d)$ in terms of angle. The samples from the posterior roughly follow the curves representing the fraction of lightning signal reaching the detector, peaking at about 5°.

We run the model described above with just the Leonids and the background rate as the shower sources. For the Leonids, we set them as active for 10 days from their peak solar longitude as computed from ground-based meteor cameras (P. Jenniskens et al., 2016). Running the model with the Perseids—the second most prominent meteor shower in GLM data—included too was deemed unsuccessful as some coefficients for the Perseids achieved distributions as wide as their priors, intendend to be uninformative.

After running the No-U-Turn sampler to obtain the posterior ditributions, we can separately examine the posterior distributions of the polynomials for $b(d)$ and $\lambda_s(l)$. Figure 4 displays the variation of the bolide rate as a function of the angle of incident light upon the sensor. The alignment of the posterior bolide detection rate curves with the known efficiency of the sensor suggests that the bandshift due to incidence angle does indeed affect bolide detection efficiency.

Figure 5 displays the variation of the non-Leonid bolide rate as a function of latitude. The steep drop-off at high latitudes is consistent with GLM being strongly biased towards detecting faster objects. We believe the bias to the North is a true bias, as the flux of meteor showers is higher in the Northern sky (American Meteor Society, n.d.).

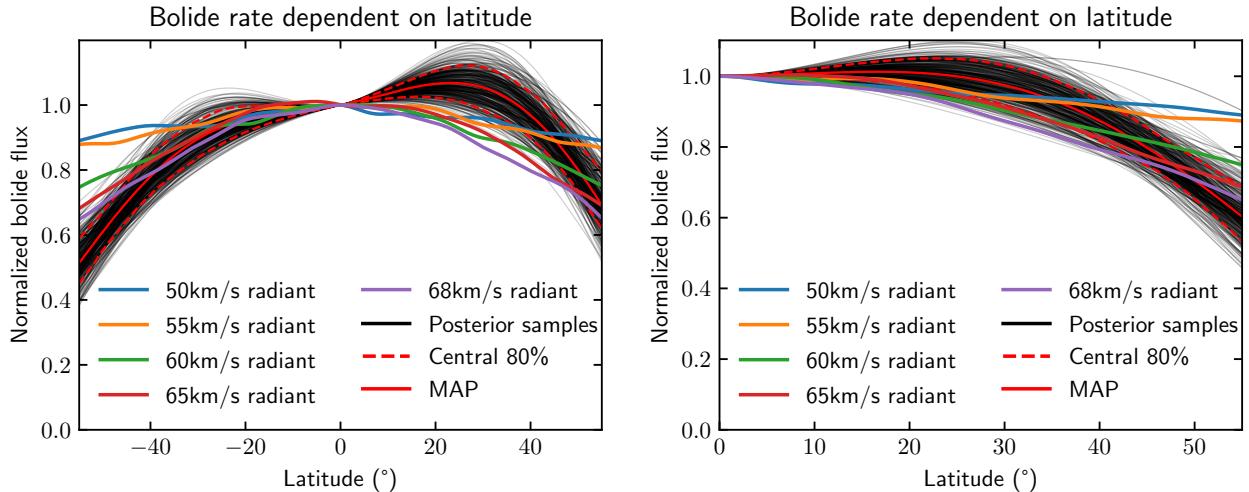


Figure 5: Non-Leonid bolide rate as a function of latitude. The figure on the right folds latitude over the equator, so -50° becomes 50° . Rate is per unit area per unit time. Theoretical distributions obtained using the method of (Robertson et al., 2021) with fixed radiant velocities. In the plot on the left, we see that GLM bolide detections seem biased towards the North, even after removing the Leonids. In the plot on the right, while the observed curve does not precisely match any of the theoretical curves, the drop-off at high latitudes is consistent with GLM being biased towards very fast objects.

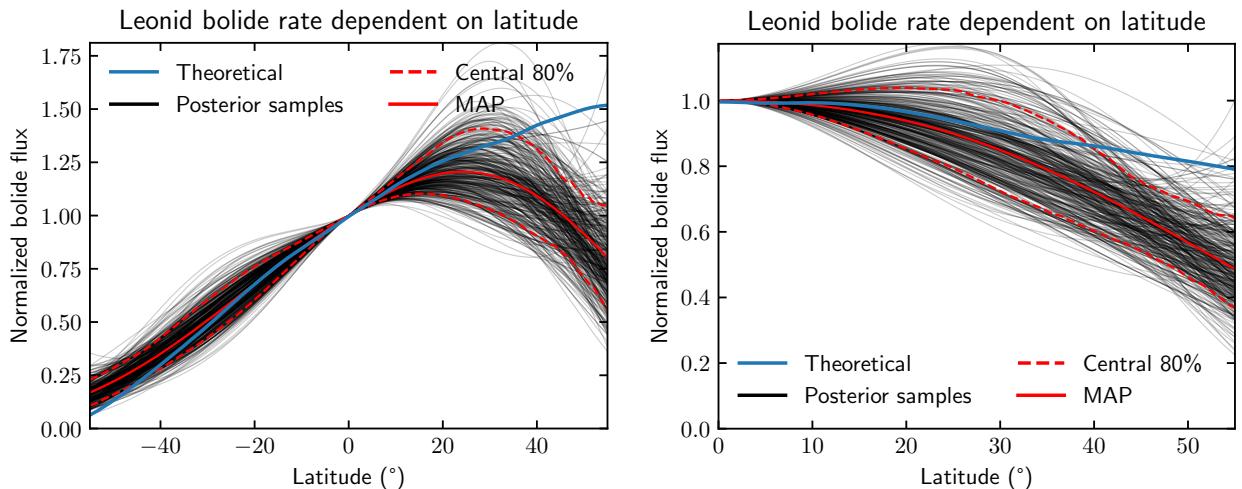


Figure 6: Leonid bolide rate as a function of latitude. The figure on the right folds latitude over the equator, so -50° becomes 50° . Rate is per unit area per unit time. Theoretical distribution obtained using the method of (Robertson et al., 2021) with radiant velocity, radiant declination, and solar longitude fixed to median values for the Leonids (P. Jenniskens et al., 2016). We see that the theoretical distribution is consistent with that observed by GLM.

There are several ways to verify the robustness of the model. Figure 6 displays the variation of the Leonid bolide rate as a function of latitude, and the observed distribution is consistent with the theoretical distribution. Figure 12 shows the headline result for several different confidence thresholds, showing that it is similar for thresholds between 0.5 and 0.9. Finally, Figure 13 shows the results when bolides detected by GOES-16 and GOES-17 are considered separately, again showing similar results and demonstrating how the method pools the results from both satellites to obtain a tighter fit that is less dependent on any possible effects of the geography in a satellite’s FOV.

4 Discussion

4.1 Computed distribution and bias

The key result, that GLM does indeed detect a decrease in bolide flux at high latitudes, is consistent with GLM being biased towards faster objects. The fact that GLM is disproportionately sensitive to the Leonids, which have a radiant velocity of 70km/s suggests that GLM does indeed have this bias. This is consistent with the observed bolide detections by solar hour depicted in Figure 9, which shows an overabundance of detections at points in the direction of the Earth’s motion even while there does not appear to be no large effect of sunlight on bolide detection efficiency. It would also make physical sense, as faster objects deposit more energy into the atmosphere. However, because most GLM detections, being outside of the stereo region, do not have associated velocity data, and most detections in the stereo region are single-pixel observations with unreliable velocity data, adjusting for velocity is not currently possible. A full analysis of GLM velocity calculations, their uncertainty, and velocity-derived detection biases remains to be done.

Figure 6, showing Leonid bolide rates across latitudes, shows good agreement between the distribution detected by GLM and the theoretical distribution (which is well-understood). This suggests that the debiasing according to nadir distance does indeed bring the GLM-detected bolide rates close to the true rates. Further development of the methods will allow studying other, less-prominent meteor showers that appear in GLM data.

Figure 4, showing bolide rate as a function of the angle of incidence of the light entering the sensor, shows a strong dependence. This dependence aligns quite well with the measured performance of the sensor, particularly the “Raw signal” curve showing the efficiency of the sensor for the light emitted by lightning. However, GLM detects a combination of continuum emission and oxygen line emission during bolide impacts (Peter Jenniskens et al., 2018; Smith et al., 2021), and angle of incidence affects both the total efficiency of light transmission and shifts the band of the narrowband filter nominally at 777.4nm. Thus the detection efficiency of bolides as a function of angle of incidence is a combination of different optical effects affecting both continuum and line emission and depends on bolide velocity, mass, and composition. In turn, the observed bolide rate as a function of angle of incidence depends on the population of bolides. Fully disentangling and understanding these effects is the subject of future work.

4.2 Comparison to previous empirical calculations

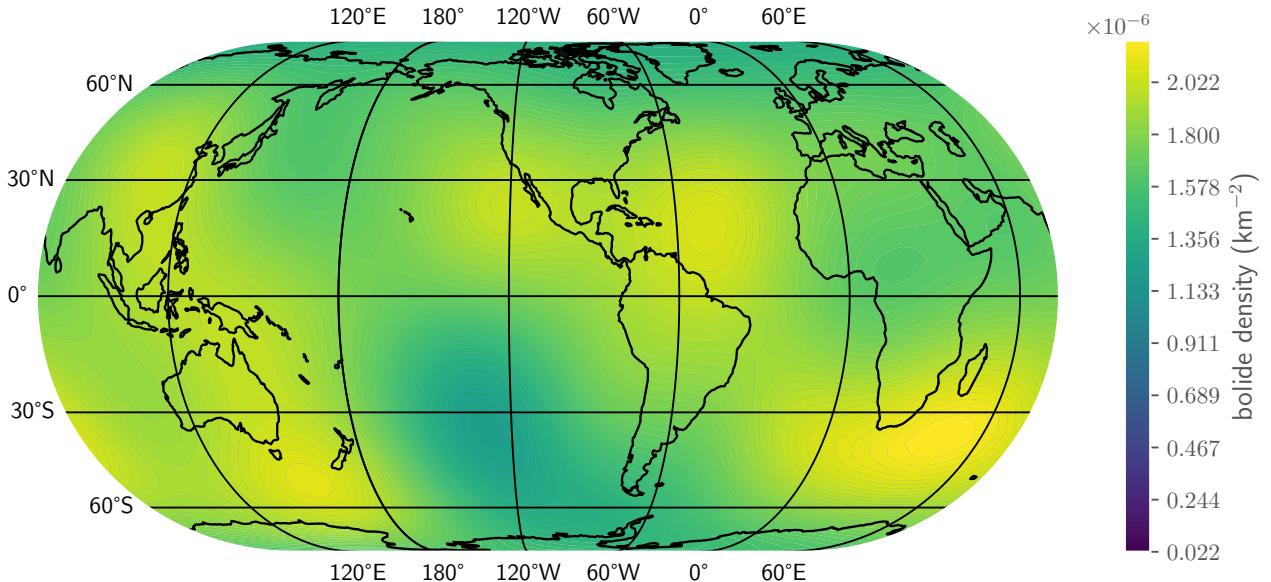


Figure 7: Density of USG bolide detections produced by a kernel density estimator with the Haversine distance metric and a bandwidth of 20° . There is no reason to assume that the USG sensors have the same sensitivity across the globe. Eckert IV projection. Figure generated using `bolides` (Ozerov and Smith, n.d.).

Robertson et al., 2021 and Evatt et al., 2020 include empirical calculations for USG data. They both demonstrate a flux that appears to increase then decline as one moves away from the equator. In the USG data, the large error bars may indicate that the initial rise might be due to sampling error. However, the decline at high latitudes appears to be significant in both studies. A decline at high latitudes is consistent with our findings in GLM data, though as Robertson et al., 2021 note, the USG data does not seem to show a decline until latitudes that are outside the GLM field of view. But in principle USG data cannot be used for such studies as the detection efficiency of the sensors cannot be assumed to be uniform across the surface of the Earth (see Figure 7). While USG data is highly useful for studying individual events, insufficient information is available about the sensor capabilities and pipeline to apply something like the model in this paper to it in an attempt to debias it for population studies.

Robertson et al., 2021 also includes empirical calculations for GLM data. They again show an increase followed by a decline as one moves away from the equator, though a much sooner decline than for USG data. Our Figure 4 shows that detection rates peak at some distance from the nadir, which happens to correspond with about twenty degrees of latitude. Hence our Figure 5 does not show as much of an increase in flux as one moves away from the equator. Similarly, our results show that detection rates plummet at high angles of incidence of light into the sensor, and hence our results do not show as much of a decrease at high latitudes, though there is still a decline. In both cases, estimating and taking into account the bias due to the angle of incidence makes the variation observed across latitudes less extreme. We also use pipeline data rather than human-vetted data, and use the computed FOV rather than the nominal FOV, which no doubt also makes our results differ from previous work.

4.3 Comparison to previous theoretical calculations

Several previous studies have made theoretical calculations of the latitudinal distribution of impacts. These studies fall into two categories:

1. Halliday, 1964 and Evatt et al., 2020 proceed from the assumption that the vast majority of impactors come from the ecliptic, and with some calculations obtain the resulting distributions for given velocities. The former simply reports these distributions, while the latter uses the velocity distribution in USG data to derive a final distribution.
2. Le Feuvre and Wieczorek, 2008 and Robertson et al., 2021 proceed from data on known NEOs and their impact computed probabilities to produce distributions of ecliptic latitude and velocity, and run Monte-Carlo simulations to obtain final distributions (the difference lies in the NEO database and the impact probabilities).

The studies in (1) predict a large decrease in flux at high latitudes, while those in (2) predict either a roughly uniform distribution or an increase in flux at high latitudes. This difference is understandable, as including objects with high inclinations in the calculations will increase the relative flux at the poles. Taking our results at face value would suggest that the distributions in (1) are more accurate, and that the studies in (2) over-emphasize objects at high inclinations (perhaps assigning them a too-high impact probability). However, due to the possible velocity bias in GLM detections, which would in fact create such an observed dropoff at high latitudes, the results cannot be taken to support the studies in (1) over those in (2).

4.4 Impossibility of completely debiasing the distribution

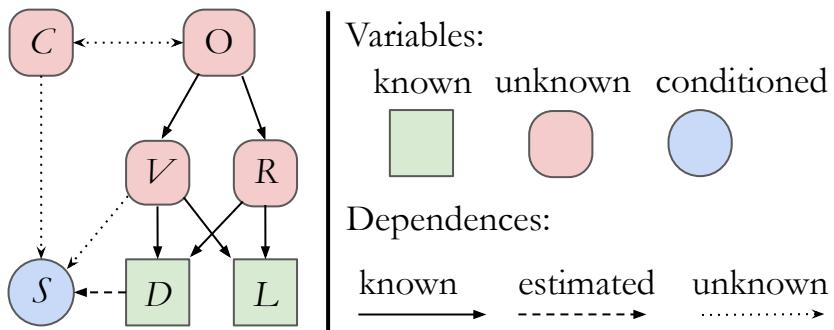


Figure 8: Diagram representing the observed and unobserved variables and the relationships between them. C : bolide's physical characteristics (mass, composition, etc.). O : bolide's orbit prior to the encounter with the Earth. V : radiant velocity. R : radiant. D : impact distance from nadir. L : impact latitude. S : a selection variable representing whether or not the bolide is detected. Both D and L are affected by both V and R , which in turn are determined by O , the orbit. S is conditioned upon in the data because all bolides in the data have necessarily been detected. It is unknown how C , V , and D affect the probability of being detected, but the effect of D is estimated here.

Only using GLM data, it is not possible to say what proportion of bolides with a certain mass and velocity are detected, which makes it not possible to adjust for these factors and get closer to the true distribution. Figure 8 is a causal representation of how the data are generated. With such a diagram in hand, it is easy to see that it is impossible to recover $P(L)$ (the distribution over latitudes, not conditioned on anything) from the data $P(L, D|S = 1)$ without making strong assumptions that V and C don't affect S . In fact, to estimate $P(L)$ independent of any GLM detection biases, one would have to define more carefully what exact objects are meant (1-meter objects, 0.5-meter, or smaller? Or perhaps some threshold of energy deposition?). If one were to assume that C does not affect S , then knowing V would be enough, as:

$$P(L) = \sum_{v,d,s} P(L|S = s, V, D)P(S = s|V, D)P(V, D) = \sum_{v,d,s} P(L|S = 1, V, D)P(S = s|V, D)P(V, D)$$

The key step is that L is independent of S when conditioned on V and D (assuming that C has no effect on S). The distribution $P(S = s|V, D)$ could be estimated similarly to how, in effect,

$P(S = s|D)$ is estimated in this paper. There is an additional relationship between V and D which would need to be disentangled, but it is reasonable to assume that V does not have a significant effect on D . Then all that is needed is the assumption that D is uniform, which is reasonable because it is determined by a uniform point on the impact disk, and knowledge of $P(V)$, which can be obtained from other sources.

The GLM data does contain estimates of V in the stereo region, and can in principle be used to obtain velocity vectors and compute the radiant and the orbit, but this is only effective in multi-pixel detections, which are relatively rare. More work is needed to study GLM's estimates of V .

4.5 Other approaches to probing biases

4.5.1 Diurnal variation

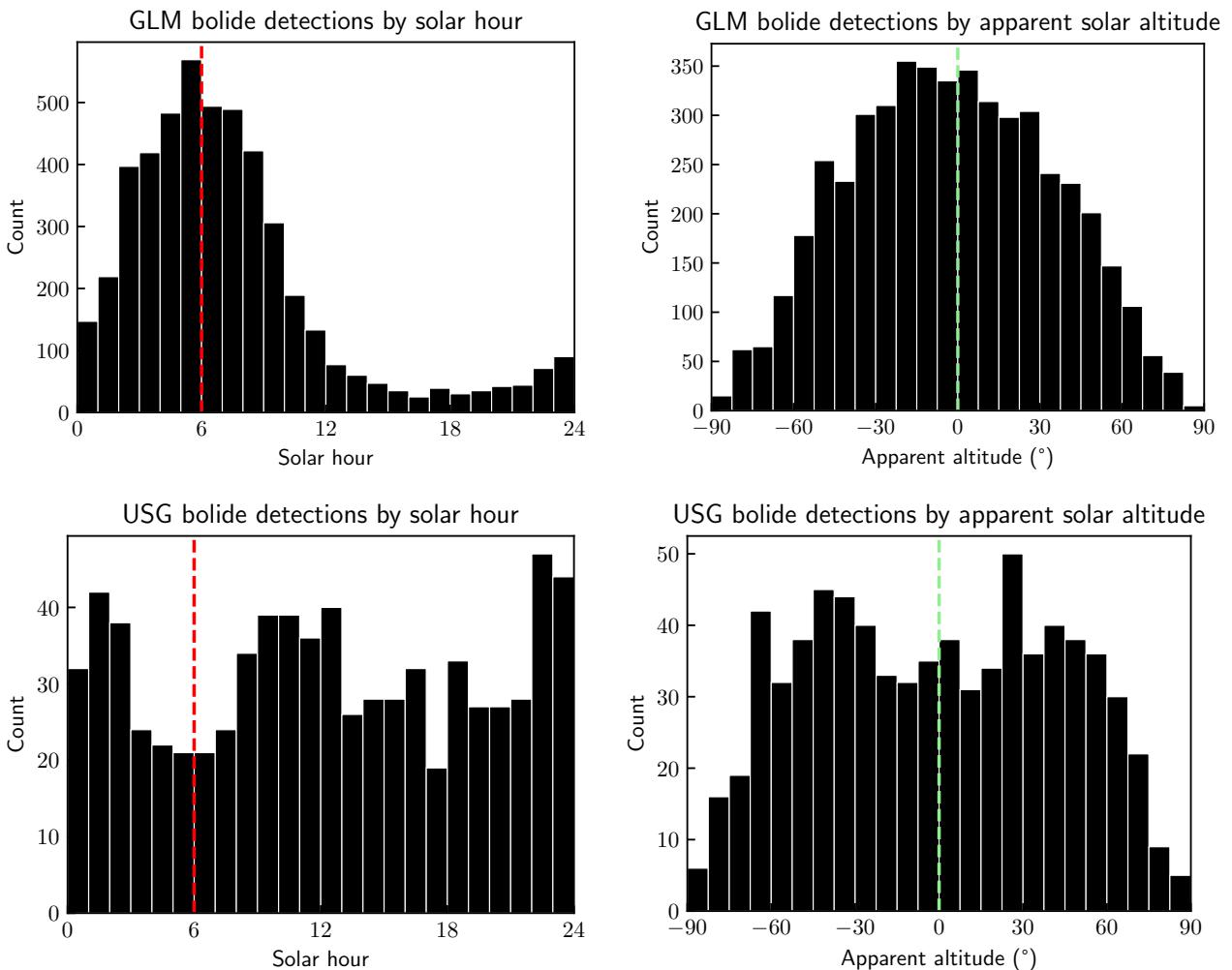


Figure 9: The histograms on the left are bolide detections by the local solar hour at the time of the impact (12 being noon). The red line corresponds to a solar hour of 6, which is when, for an observer at the equator, the direction of the Earth's motion is directly above. Those on the right are bolide detections by the apparent solar altitude, i.e. the altitude, in degrees, of the Sun above the horizon (no atmospheric refraction). The light green line corresponds to the Sun being directly on the horizon (again, assuming no refraction). The top two figures are for GLM data, the bottom two are for USG data as of 2022-12-20.

As another piece of evidence supporting the idea that GLM is much more sensitive towards faster objects, consider Figure 9. This figure shows that, in local solar time at a bolide's location,

detections peak at six in the morning in the GLM data (in fact, this is the case in all seasons, as seen in Figure 11). This is also the local solar time when a standing observer at the equator is travelling head-first in the direction of the Earth’s motion. USG data does not show such a peak. Given that, in both, there does not appear to be a major effect of solar illumination on detection rates, this again suggests a detection bias towards faster objects in GLM data, as, generally, the fastest objects (in terms of relative velocity to the Earth) come from retrograde orbits and hence tend to impact at around 06:00. If there were no such bias, we would expect a bimodal distribution, or perhaps a weaker peak, as there would be more detections at midnight and noon in local solar time (which correspond to objects crossing the Earth’s orbit from further and closer to the Sun, respectively) (Robertson et al., 2021). This is, in fact, what is seen in the USG data.

GLM’s distribution over solar hour can also be compared to the diurnal variation of meteor rates as detected by ground-based meteor radars. The ratio of GLM detection counts from the peak at 05:30-06:30 to the trough at 17:30 to 18:30 is $516/32 = 16.125$. This is very much on the extreme end of the diurnal variations reported by several meteor radar systems in various locations and times of year (Lovell, 1954; Okamoto and Maegawa, 2008; Singer et al., 2005; Szasz et al., 2004), which again suggests that GLM has a strong bias towards high-velocity objects.

4.5.2 Sensitivity to meteor showers

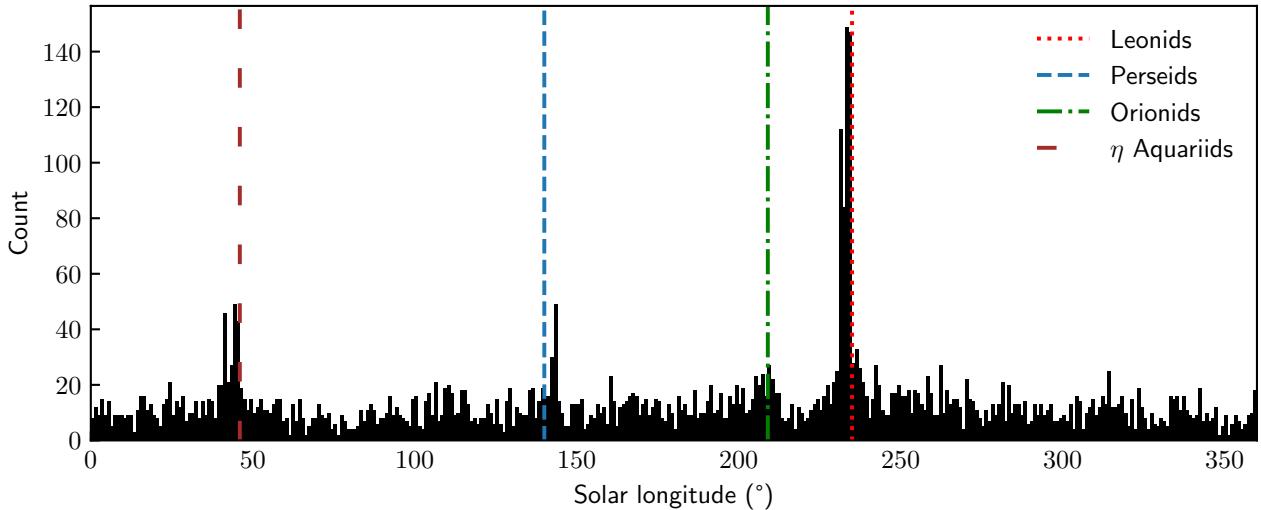


Figure 10: GLM bolide detections by the solar longitude at the time of the impact (0 being the vernal equinox). Dashed and dotted lines represent the solar longitudes of the peaks of different meteor showers (P. Jenniskens et al., 2016).

As seen in Figure 10, there is an increase in detections around the peak of the Leonids. Though it is not possible to attribute GLM detections to specific meteor showers, this temporal correspondence suggests that GLM is particularly sensitive to the Leonids, which also has one of the fastest radiant velocities at 70km/s (P. Jenniskens et al., 2016). GLM appears to be more sensitive to the Leonids than to the Perseids (59km/s), Orionids (66km/s), and η Aquariids (66 km/s). The Leonids are, in fact, not at all more prominent than these other showers in ground-based data (P. Jenniskens et al., 2016; Vida et al., 2021), again suggesting that GLM has a bias due to velocity. Given the small differences in radiant velocity, the bias must be quite large, unless other factors, like differences in meteor size or composition in the different showers, play a role. Other prominent showers, like the Geminids (34km/s), do not appear to be associated with any spike in flux in the GLM data.

4.6 Future work

Several ground-based meteor networks have cameras in the GLM field of view, and their ground-based observations allow precise estimates of orbits, velocities, and some estimation of physical characteristics. Given that ground-based data contains much fainter objects than GLM data does, it may be possible to assume that ground-based data will detect all impacts above a certain mass and velocity. Given this, it is possible to compute a distribution $P(S|V, C)$ by seeing what proportion of ground-based detections at certain characteristics and velocities are detected by GLM. With some assumptions on how $P(S|V, C, D)$ acts as a function of V , M , and D (one reasonable one being that the effects of (V, M) and D are multiplicative in decreasing the probability of detection), it may then be possible to completely debias GLM's observed distribution. However, there are challenges associated with this, including that ground-based data is affected by weather and, in particular, sunlight, which may skew the data in unexpected ways. Alternatively, more work could be done on fusing ground-based data with GLM data to improve models estimating velocity and physical characteristics from GLM data alone.

5 Conclusion

We have partially debiased the latitudinal variation of bolide flux detected by GOES GLM by estimating and accounting for optical effects using a Poisson regression. Using this parametric method, we also offer much tighter distributions than completely nonparametric methods like binning and using Poisson errors which do not take full advantage of the data and lead to larger errors than necessary. It is reasonable to assume that the variation with latitudes is smooth and does not jump wildly from one bin to the next.

Our method is able to recover the well-understood latitudinal distribution of a meteor shower like the Leonids. Yet we advise against using the results of this work in confirming or rejecting theoretical models of overall impact distributions, as GLM's bolide detection biases are still not well-understood enough to obtain a completely debiased estimate. We are unable to recover an overall debiased distribution as velocities are not measured accurately enough, and bolide velocity is related to both its radiant and, presumably, its probability of detection by GLM.

6 Acknowledgments

Darrel Robertson provided code to calculate theoretical distributions across latitudes and much helpful discussion. Katrina Virts provided the data on the GLM field of view, without which a proper analysis would not have been possible. We would also like to thank developers of open-source software, including: Python, the scientific Python stack, Astropy, and PyMC. Work carried out for this project is supported by NASA's Planetary Defense Coordination Office (PDCO). Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center. AO and JS are supported through NASA Cooperative Agreement 80NSSC19M008.

References

- American Meteor Society (n.d.). *Major Meteor Showers*. URL: <https://www.amsmeteors.org/meteor-showers/major-meteor-showers/>.
- Barber C.B., Dobkin D.P. and Huhdanpaa H.T. (Dec. 1996). “The Quickhull algorithm for convex hulls”. In: *ACM Trans. on Mathematical Software* 22.4, pp. 469–483.

- Du, Qiang, Vance Faber, and Max Gunzburger (1999). "Centroidal Voronoi Tessellations: Applications and Algorithms". In: *SIAM Review* 41.4, pp. 637–676. DOI: 10.1137/S0036144599352836. eprint: <https://doi.org/10.1137/S0036144599352836>. URL: <https://doi.org/10.1137/S0036144599352836>.
- Evatt, G.W. et al. (Apr. 2020). "The spatial flux of Earth's meteorite falls found via Antarctic data". In: *Geology* 48.7, pp. 683–687. ISSN: 0091-7613. DOI: 10.1130/G46733.1. eprint: <https://pubs.geoscienceworld.org/gsa/geology/article-pdf/48/7/683/5073455/683.pdf>. URL: <https://doi.org/10.1130/G46733.1>.
- Halliday, Ian (1964). "The Variation in the Frequency of Meteorite Impart with Geographic Latitude". In: *Meteoritics* 2.3, pp. 271–278. DOI: <https://doi.org/10.1111/j.1945-5100.1964.tb01433.x>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1945-5100.1964.tb01433.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1945-5100.1964.tb01433.x>.
- Hoffman, Matthew D, Andrew Gelman, et al. (2014). "The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo." In: *J. Mach. Learn. Res.* 15.1, pp. 1593–1623.
- Jenniskens, P. et al. (2016). "The established meteor showers as observed by CAMS". In: *Icarus* 266, pp. 331–354. ISSN: 0019-1035. DOI: <https://doi.org/10.1016/j.icarus.2015.09.013>. URL: <https://www.sciencedirect.com/science/article/pii/S0019103515004182>.
- Jenniskens, Peter et al. (2018). "Detection of meteoroid impacts by the Geostationary Lightning Mapper on the GOES-16 satellite". In: *Meteoritics & Planetary Science* 53.12, pp. 2445–2469. DOI: <https://doi.org/10.1111/maps.13137>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/maps.13137>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/maps.13137>.
- Le Feuvre, Mathieu and Mark A. Wieczorek (2008). "Nonuniform cratering of the terrestrial planets". In: *Icarus* 197.1, pp. 291–306. ISSN: 0019-1035. DOI: <https://doi.org/10.1016/j.icarus.2008.04.011>. URL: <https://www.sciencedirect.com/science/article/pii/S0019103508001802>.
- Lovell, B. (1954). *Meteor Astronomy*. International series of monographs on physics. Clarendon Press. URL: <https://books.google.ca/books?id=Yje4AAAAIAAJ>.
- Okamoto, Sadao and Kimio Maegawa (2008). "Annual and Diurnal Variation of Meteor Rates by the Forward-Scatter Radio Observation". In: *Earth, Moon, and Planets* 103, pp. 65–68.
- Ozerov, Anthony and Jeffrey C. Smith (n.d.). *Bolides*. URL: <https://bolides.readthedocs.io>.
- Robertson, Darrel et al. (May 2021). "Latitude Variation of Flux and Impact Angle of Asteroid Collisions with Earth and the Moon". In: *The Planetary Science Journal* 2.3, p. 88. DOI: 10.3847/psj/abefda.
- Salvatier, John, Thomas V. Wiecki, and Christopher Fonnesbeck (Apr. 2016). "Probabilistic programming in Python using PyMC3". In: *PeerJ Computer Science* 2, e55. DOI: 10.7717/peerj-cs.55. URL: <https://doi.org/10.7717/peerj-cs.55>.
- Singer, W. et al. (Aug. 2005). "Diurnal and annual variations of meteor rates at latitudes between 69°N and 35°S". In: *17th ESA Symposium on European Rocket and Balloon Programmes and Related Research*. Ed. by Barbara Warmbein. Vol. 590. ESA Special Publication, pp. 151–156.
- Smith, Jeffrey C. et al. (Nov. 2021). "An automated bolide detection pipeline for GOES GLM". In: *Icarus* 368, p. 114576. DOI: 10.1016/j.icarus.2021.114576.
- Szasz, Csilla et al. (Dec. 2004). "Latitudinal Variations of Diurnal Meteor Rates". In: *Earth, Moon, and Planets* 95.1, pp. 101–107. ISSN: 1573-0794. DOI: 10.1007/s11038-005-9007-0. URL: <https://doi.org/10.1007/s11038-005-9007-0>.
- Vida, Denis et al. (2021). "The Global Meteor Network—Methodology and first results". In: *Monthly Notices of the Royal Astronomical Society* 506.4, pp. 5046–5074.

A Supplementary Figures

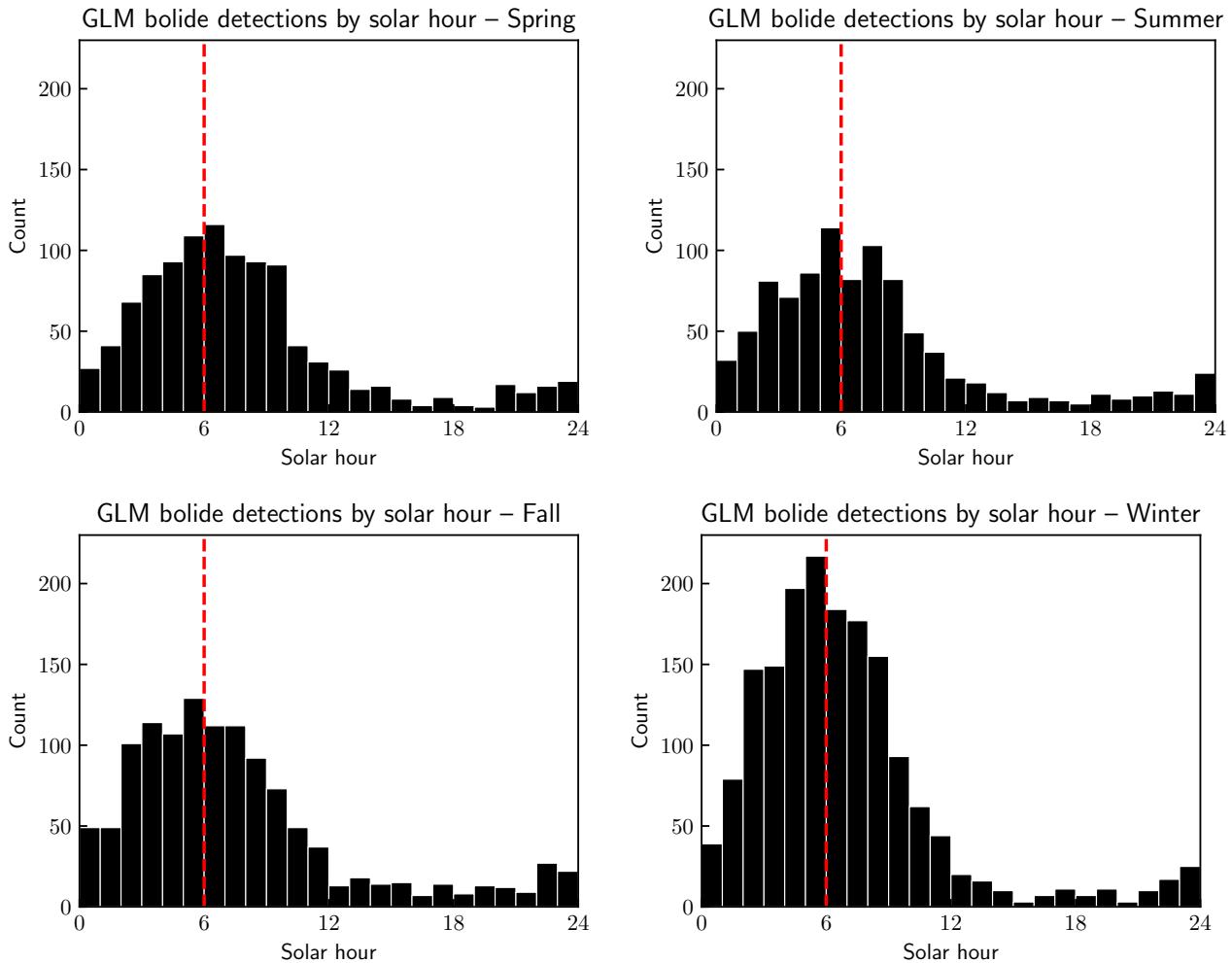


Figure 11: GLM bolide detections by the local solar hour at the time of the impact (12 being noon), split apart by season. The red line corresponds to a solar hour of 6, which is when, for an observer at the equator, the direction of the Earth’s motion is directly above. Detections are assigned to “seasons” according to the solar longitude at the time of the detection. Spring corresponds to a solar longitude in the interval $[-45, 45]$, Summer to $[45, 135]$, Fall to $[135, 225]$ and Winter to $[225, 315]$. The extreme diurnal variation is present in all seasons.

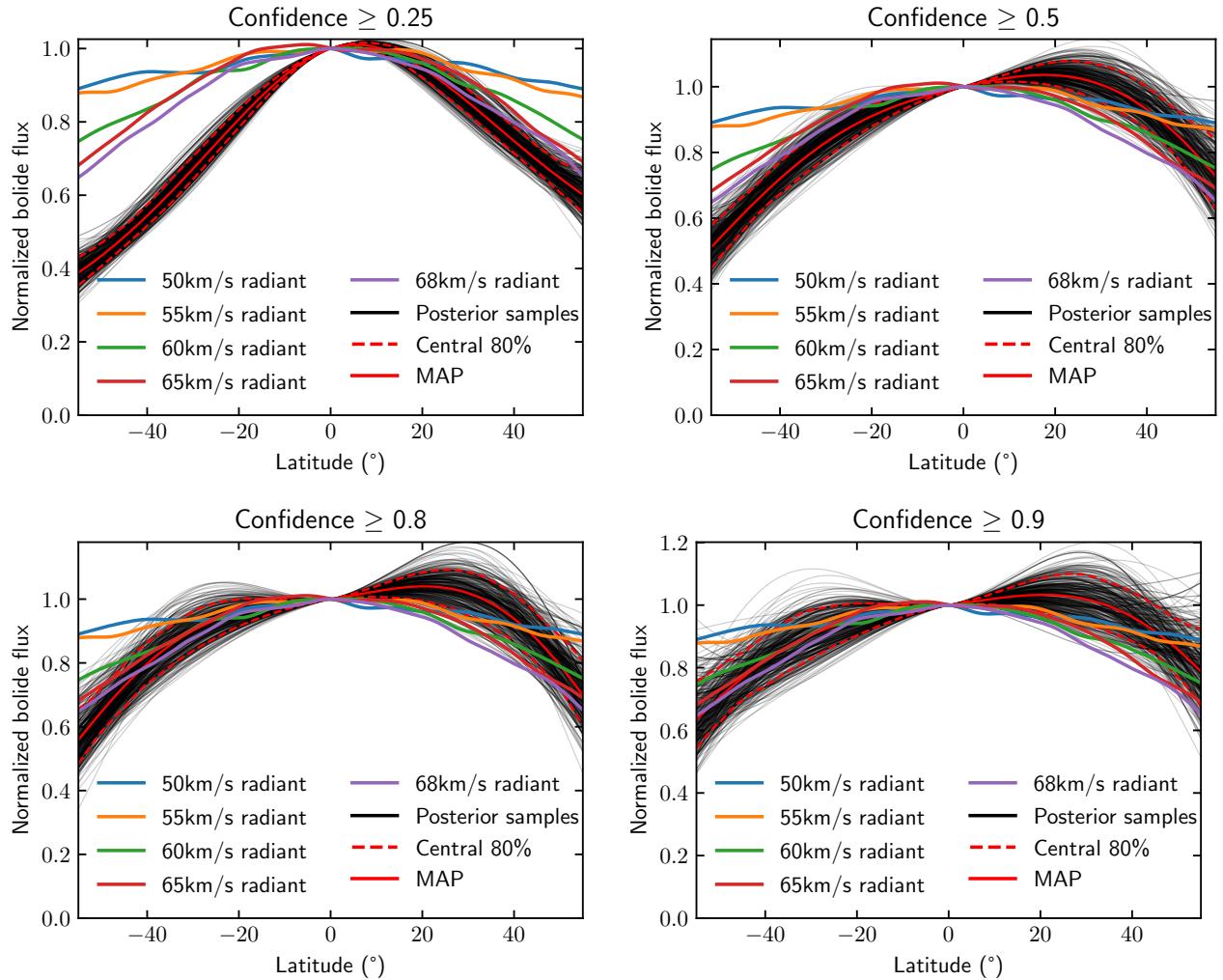


Figure 12: Non-Leonid bolide rate as a function of latitude, computed by fitting the model at several confidence thresholds. Rate is per unit area per unit time. Theoretical distributions obtained using the method of Robertson et al., 2021 with fixed radiant velocities. Higher thresholds entail a wider posterior, as there are fewer detections. Lower thresholds emphasize the equator more, which may be due to false positives from solar glint.

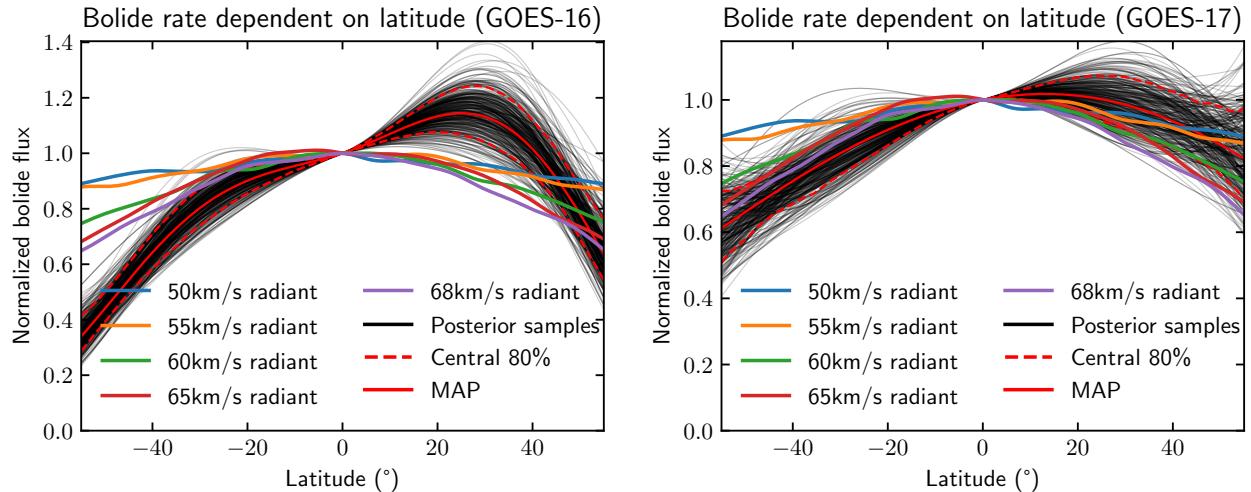


Figure 13: Bolide rate as a function of latitude, fitted separately for GOES-16 and GOES-17 (rather than combining the detections from both). Rate is per unit area per unit time. Theoretical distributions obtained using the method of Robertson et al., 2021 with fixed radiant velocities. The distributions are broadly similar, especially given the uncertainties.