Chapter 5

Variability Analysis

With the light curve data from the monitoring program in hand, we now turn to analyzing the 15 GHz variability of each source. In this thesis, we will focus on the amplitude of that variability, treating the light curve as a population of samples drawn from a distribution defined by the processes responsible for the radio emission. For the most part, we will ignore the time coordinate in these studies. Other studies, both current and future, will examine the time dependence of these data more directly. Here, we will use a variability amplitude metric as a simple measure of source activity and aim to use this to identify a connection between gamma-ray emission and radio variability. Firmly and reliably establishing this connection is a prerequisite for performing more detailed correlation studies. The methods and the two-year CGRaBS results described in this chapter were first published in Richards et al. (2011). My colleague Dr. Vasiliki Pavlidou conceived of and developed computer codes to implement the likelihood analysis tools that we describe here. My contributions included early discussions of the standard variability methods that inspired this method, producing the light curve data for the tests, selecting the subpopulation samples for the comparisons, and the analysis and interpretation of the results of the tests.

In this chapter, we will first define the analytical tools we will use to characterize the variability amplitude of a light curve and to compare the variability of source populations. We will then present the results of a number of population studies, first demonstrating that our methods produce null results when given null inputs, then comparing the variability amplitudes of gamma-ray–loud sources with gamma-ray–quiet ones, of BL Lac objects with FSRQs, and of high-redshift with low-redshift FSRQs. We will then assess the impact of cosmological time dilation on this latter effect, and finally examine the differences between our CGRaBS and 1LAC source samples in detail.

We now have 18 months of data past the end of the set used for the Richards et al. (2011) results. These data, collected between January 2010 and June 2011, make a number of additional tests possible. First, we can verify that the properties of the various subpopulations we studied in the two-year results are stable as new data are added. Second, we now have enough data to characterize the variability of the *Fermi*-LAT–detected 1LAC sources, which were added to the OVRO sample in March 2010 (although some had been added earlier). Thus, we can now compare the 1LAC sample properties to the CGRaBS sample properties.

Finally, we have now sampled high-redshift ($z \sim 3$) sources for nearly a year of their rest frame time. In section 5.5, we will use these additional data to examine the effect of time dilation on the observed radio variability.

5.1 Analytical Tools

In this work, we are chiefly interested in characterizing and studying the *amplitude* of variability for a given blazar. This is a simple characterization of the behavior of a source that can readily be applied to the study of the collective properties of a large population. In this section, we begin by introducing the intrinsic modulation index, a variability amplitude metric that is particularly well suited to the data from our observing program. We will examine how this metric performs on our data set, then introduce a likelihood analysis framework for comparing the intrinsic modulation indices of subpopulations within our sample. We will then use these tools to explore the variability characteristics of the blazars in our program.

5.1.1 Intrinsic Modulation Index

Characterizing the variability amplitude of a source and assessing the confidence with which this can be measured are complex problems that have been addressed using a variety of measures and tests, such as the variability index (e.g., Aller et al. 1992); the fluctuation index (e.g., Aller et al. 2003); the modulation index (e.g., Kraus et al. 2003); the fractional variability amplitude (e.g., Edelson et al. 2002; Soldi et al. 2008); and χ^2 tests of a null hypothesis of nonvariability. Each of these tools provides different insights to the variability properties of sources and is sensitive to different uncertainties, biases, and systematic errors.

In this work, we will measure variability using the *intrinsic modulation index*, $\overline{m}$, which we introduced in Richards et al. (2011), where a full explanation of the likelihood analysis used to compute $\overline{m}$ is presented. Here we will give only a brief explanation of this variability measure and its properties. The intrinsic modulation index is based on the standard modulation index, defined as the standard deviation of the flux density measurements in units of the mean measured flux density, i.e.,

$$m_{\rm data} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(S_i - \frac{1}{N}\sum_{i=1}^{N} S_i\right)^2}}{\frac{1}{N}\sum_{i=1}^{N} S_i}\,. \tag{5.1}$$

The modulation index is reasonably well behaved: it is always nonnegative and is reasonably robust against outliers. However, it measures a convolution of intrinsic source variation and observational uncertainties—a large modulation index could be indicative of either a strongly variable source or a faint source with high uncertainties in individual flux density measurements. For this reason, the correct interpretation of the modulation index requires that measurement errors and the uncertainty in $m_{\rm data}$ due to the finite number of flux density measurements be properly estimated.

Our intrinsic modulation index is defined as

$$\overline{m} = \frac{\sigma_0}{S_0}\,, \tag{5.2}$$

and like the ordinary modulation index is defined by a standard deviation divided by the mean. In this case, however, σ_0 and S_0 represent *intrinsic* quantities—properties of the light curves before they are affected by observational noise, imperfect sampling, etc. Because we cannot directly measure these intrinsic quantities, we will use a likelihood analysis to estimate them from the data we collect. Observational uncertainties will affect the accuracy with which we can estimate these quantities, but we will quantify these uncertainties and propagate them into our later analysis as errors in our estimated values.

Evaluating the significance of a difference between measured values requires a good estimate of the uncertainty in those values. Thus, for the population comparisons we perform in this work, we require a rigorous estimate of the uncertainty in each intrinsic modulation index we calculate. Other methods for assessing the uncertainty in variability measures have been employed. One method that has been widely used is to evaluate each measure for a set of constant-flux-density calibrators which are known to have a flux density constant in time and which have been observed with the same instrument over the same periods of time. The value of the variability measure obtained for the calibrators is then used as a threshold value, so that any source with variability measure equal to or lower than that of the calibrators is considered consistent with being nonvariable. However, a variability measure value higher than that of the calibrators is a necessary but not sufficient condition for establishing variability. Calibrators are generally bright sources, with relative flux density measurement uncertainties lower than the majority of monitored sources; additionally, variability measures are affected by the sampling frequency, which is not necessarily the same for all monitored sources and the calibrators.

Alternatively, the *significance* of variability in a given source can be established through tests (such as a χ^2 test) evaluating the consistency of the obtained set of measurements with the hypothesis that the source was constant over the observation interval. However, such tests provide very little information on sources for which statistically significant variability cannot be established, as they cannot distinguish between *intrinsically nonvariable sources* and sources that could be *intrinsically variable* but inadequately observed for their variability to be revealed.

For our studies with the intrinsic modulation index, we will directly and rigorously propagate the measurement uncertainty for each flux density into an uncertainty on $\overline{m}$. In this way we will estimate both a measure of the intrinsic variability amplitude and our uncertainty in that estimate due to the measurement process.

5.1.1.1 Calculating the Intrinsic Modulation Index

As in Richards et al. (2011), we will assume that the "true" flux densities for each source are normally distributed with mean S_0, standard deviation σ_0, and intrinsic modulation index $\overline{m} = \sigma_0/S_0$. That is, we assume that the probability density for the "true" flux density S_t is

$$p(S_t, S_0, \sigma_0) = \frac{1}{\sigma_0\sqrt{2\pi}} \exp\left[-\frac{(S_t - S_0)^2}{2\sigma_0^2}\right]. \tag{5.3}$$

Similarly, we assume the observation process for the $j\mathrm{th}$ data point adds normally distributed error with mean S_t and standard deviation σ_j. Then, the likelihood for a single observation is given by

$$\ell_j = \int_{\mathrm{all\ S_t}} dS_t \frac{\exp\left[-\frac{(S_t-S_j)^2}{2\sigma_j^2}\right]}{\sigma_j\sqrt{2\pi}} \frac{\exp\left[-\frac{(S_t-S_0)^2}{2\sigma_0^2}\right]}{\sigma_0\sqrt{2\pi}}, \tag{5.4}$$

which after combining $j = 1, \ldots, N$ measurements and substituting $\overline{m}S_0 = \sigma_0$, gives

$$\begin{aligned}\mathcal{L}(S_0, \overline{m}) &= S_0 \left(\prod_{j=1}^{N} \frac{1}{\sqrt{2\pi(\overline{m}^2 S_0^2 + \sigma_j^2)}}\right) \times \\ &\quad \exp\left[-\frac{1}{2}\sum_{j=1}^{N} \frac{(S_j - S_0)^2}{\sigma_j^2 + \overline{m}^2 S_0^2}\right].\end{aligned} \tag{5.5}$$

By maximizing the likelihood given by equation (5.5), we find our estimates of S_0 and $\overline{m}$. In figure 5.1, we plot the most likely values and the 1σ, 2σ, and 3σ isolikelihood contours for the two-year data for CGRaBS source J1243–0218. These contours were computed to contain 68.26%, 95.45%, and 99.73% of the volume beneath the likelihood surface.

We finally obtain our estimate of $\overline{m}$ by marginalizing over all values of S_0, which yields a one-dimensional likelihood distribution, $\mathcal{L}(\overline{m})$, like that shown in figure 5.2. We compute the 1σ uncertainty by finding equal-likelihood points $\overline{m}_1$ and $\overline{m}_2$ on either side of the maximum likelihood value where

$$\frac{\int_{\overline{m}_1}^{\overline{m}_2} \mathcal{L}(\overline{m}) d\overline{m}}{\int_0^\infty \mathcal{L}(\overline{m}) d\overline{m}} = 0.6826. \tag{5.6}$$

When the maximum likelihood value $\overline{m}$ is less than 3σ from zero, we compute a 3σ upper limit $\overline{m}_3$ defined by

$$\frac{\int_0^{\overline{m}_3} \mathcal{L}(\overline{m}) d\overline{m}}{\int_0^\infty \mathcal{L}(\overline{m}) d\overline{m}} = 0.9973. \tag{5.7}$$

This method makes the assumption that the distribution of flux densities from a source are distributed normally. For many sources, this is a good description of the data. In figure 5.3, we plot the histogram of the two-year data set flux densities from J1243–0218 along with the maximum-likelihood Gaussian model we estimated from this analysis. In other sources, however, the distributions are clearly non-Gaussian, with

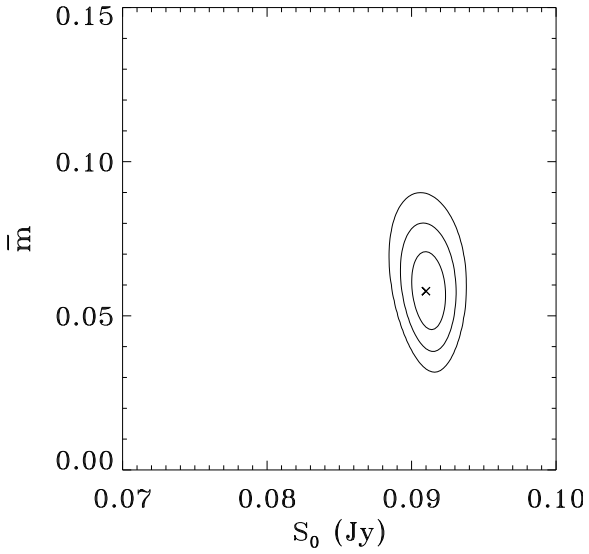

Figure 5.1. Likelihood parameter space, showing 1σ, 2σ, and 3σ contours of the joint likelihood $\mathcal{L}(S_0, \overline{m})$ for blazar J1243−0218.

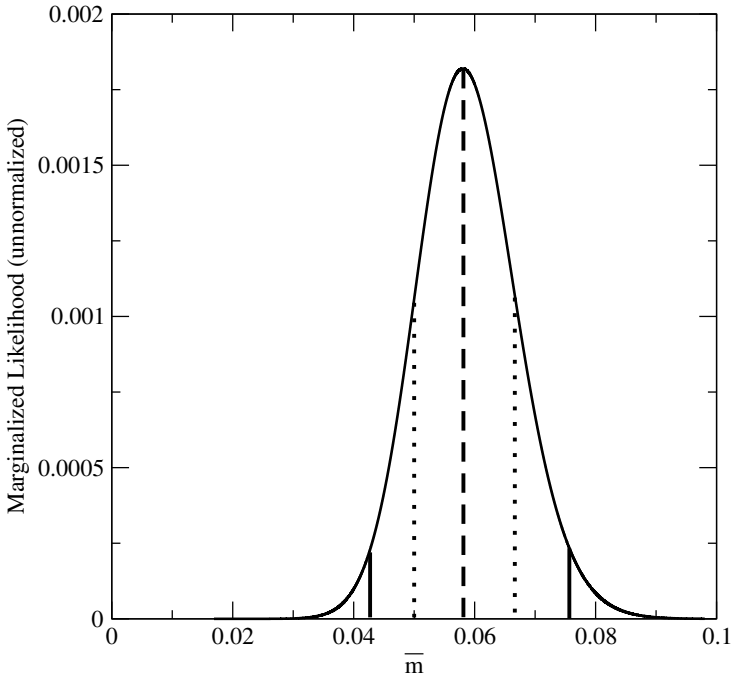

Figure 5.2. Marginalized likelihood $\mathcal{L}(m)$ for J1243−0218 (solid curve). Dashed vertical line: best-estimate $\overline{m}$; dotted vertical lines: 1σ $\overline{m}$ range; solid vertical lines: 2σ $\overline{m}$ range.

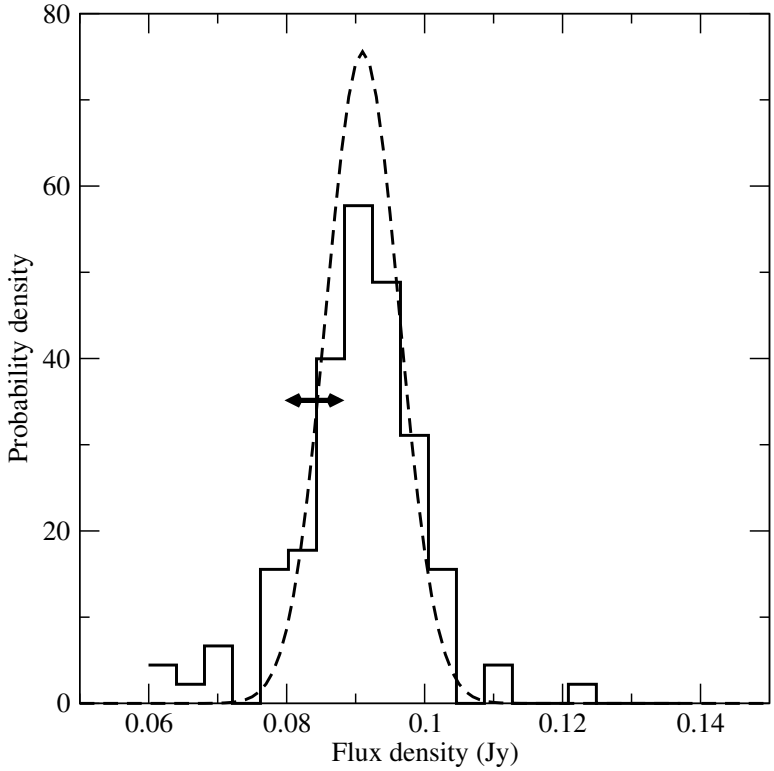

Figure 5.3. Maximum-likelihood Gaussian model for the flux density distribution (dashed line), plotted over the histogram of measured flux densities (solid line) for blazar J1243−0218. The arrow indicates the size of the typical measurement uncertainty.

possible evidence for, e.g., bimodality as shown in figure 5.4. However, we have verified that even when the true distribution is not well described as Gaussian, the modulation index and uncertainty is a reasonable description of the data. In future work, more sophisticated distributions of true flux density can be applied to this method.

5.1.2 Properties of the Intrinsic Modulation Index

We now evaluate the properties of the intrinsic modulation index using the two-year data set. Table C.1 in appendix C includes our measured values for $\overline{m}$, S_0, and their 1σ errors for the 42-month data set. Results from the two-year data are available in Richards et al. (2011). In figure 5.5 we plot the intrinsic modulation index $\overline{m}$ and associated 1σ uncertainty against the intrinsic, maximum-likelihood average flux density, S_0, for all our CGRaBS and calibrator sources. The error bar on S_0 corresponds to the 1σ uncertainty in mean flux density, calculated from the joint likelihood (equation (5.5)) marginalized over $\overline{m}$. CGRaBS sources are shown as black or magenta points or blue triangles for upper limits, while calibrators are shown as green points.

Variability could only be established at the 3σ confidence level or higher for 1139 out of 1158 CGRaBS blazars in our sample. For this study, we considered only sources for which at least 3 flux densities were mea-

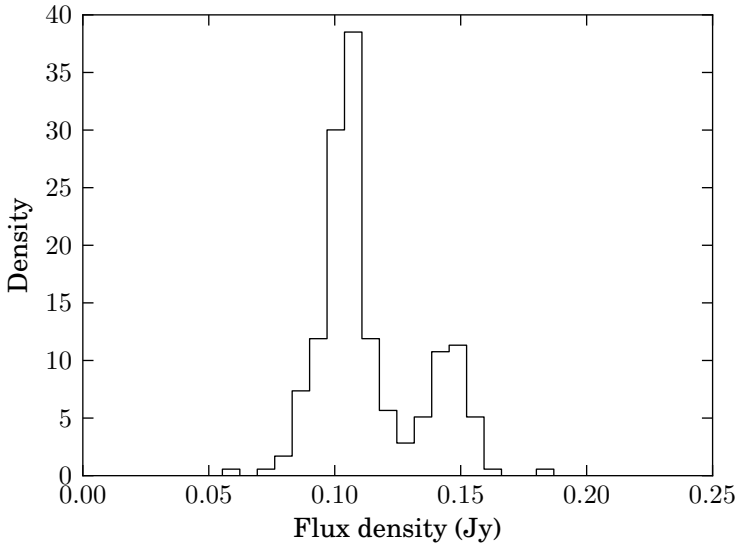

Figure 5.4. Histogram of flux density measurements for J0237+3022, demonstrating a non-Gaussian bimodal distribution.

sured, a positive mean flux density $\geq 2\sigma$ from zero was found, and at least 90% of the individual flux density measurements were $\geq 2\sigma$ from zero. These criteria excluded two sources (J1310+3233 and J1436–1846). For the other 17 sources we have calculated 3σ upper limits for $\overline{m}$. We plot these upper limits with blue triangles.

Calibration sources 3C 286, DR 21, and 3C 274 are shown in green. Although, as expected, these sources are the least variable of all sources in which variability can be established and a nonzero $\overline{m}$ can be measured, $\overline{m}$ for these sources is finite and measurable. This means that some residual long-term variability remains in our calibrators beyond what can be justified by statistical errors alone. This could conceivably result from true calibrator source variation, but more likely reflects incomplete removal of small-amplitude calibration trends. Because $\overline{m} < 1\%$ for these three sources, we quote a systematic uncertainty $\Delta\overline{m}_{\rm syst} = 0.01$ for the values of the intrinsic modulation index we produce through our analysis.

To ensure that our population studies are not affected by this residual systematic variability, in all analyses discussed in sections 5.2–5.5 only sources with $\overline{m} \geq 0.02$ will be used, so that we remain comfortably above this 1% systematic uncertainty limit. In addition, for sources with $S_0 < 60$ mJy, the number of sources for which variability can be established is of the same order as the number of sources (both CGRaBS and non-CGRaBS) for which we could only measure an upper limit, and these upper limits are very weak and nonconstraining. For this reason, we also exclude from our population studies any source with $S_0 < 60$ mJy. The part of the parameter space excluded due to these two criteria is shown in figure 5.5 as the yellow shaded area bounded by the solid black lines.

For $S_0 \geq 0.4$ Jy, no obvious correlation between flux density and modulation index is apparent, and no CGRaBS sources exist with upper limits above our cut of $\overline{m} = 2\%$. However, for sources with $S_0 < 0.4$ Jy, there is an absence of points in the lower-left corner of the allowed parameter space defined by the thick solid

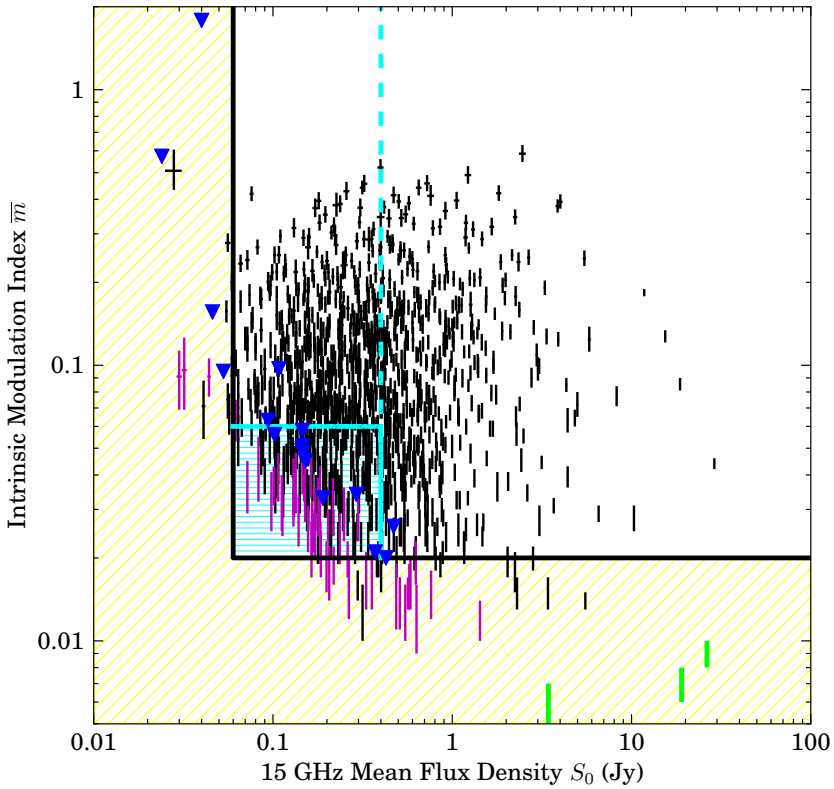

Figure 5.5. Two-year intrinsic modulation index $\overline{m}$ and associated 1σ uncertainty, plotted against intrinsic maximum-likelihood average flux density, S_0, for all sources in the program which have enough (more than 3) acceptable, nonnegative flux density measurements. Black points: CGRaBS sources found to be variable with 3σ confidence by χ^2 test; magenta points: CGRaBS sources found consistent with nonvariable by χ^2 test; green points: calibrators 3C 286, DR 21, and 3C 274; blue triangles: 3σ upper limits for CGRaBS sources for which variability could not be established at the $\geq 3\sigma$ confidence level. The error bar on S_0 corresponds to the 1σ uncertainty in mean flux density, calculated from the joint likelihood (equation (5.5)) marginalized over $\overline{m}$. Data, except for upper limits, outside the yellow and cyan shaded areas are used in the population studies of sections 5.2–5.5.

lines: for faint sources, we can only confidently establish variability if that variability is strong enough. The effect disappears for variability amplitudes greater than about 6%. In addition, there are only two CGRaBS sources with upper limits higher than 6% for sources brighter than 60 mJy (J0722+3722 and J0807+5117), $<0.5\%$ of the 452 sources measured in this region of parameter space. We conclude that we are able to measure variability at the level of 6% or higher for virtually all ($>99\%$ of) sources brighter than 60 mJy.

To ensure that our population studies are not affected by our decreased efficiency in measuring variability in sources with 60 mJy $\leq S_0 < 0.4$ Jy and $2\% \leq \overline{m} < 6\%$, we will also exclude this part of the $(S_0, \overline{m})$ parameter space from our analyses in sections 5.2–5.5. The part of the parameter space excluded due to these criteria is shown in figure 5.5 as the cyan shaded area.

For comparison, we also computed the χ^2-per degrees of freedom for each source and tested whether we could reject the hypothesis of a constant flux density at the 3σ level. Because of the long-term residual trend described in section 3.2.2.2, we added 1% of each flux density in quadrature to the reported uncertainty when computing χ^2. Of the 1139 CGRaBS sources for which we calculated the intrinsic modulation index, 51 (4.5%) are found to be nonvariable (i.e., we cannot reject the hypothesis of constant flux) with $>3\sigma$ confidence. These are plotted as magenta points in figure 5.5. All but one of these lie within the low-flux density and low-variability regions we have excluded from our population studies. The one such source not excluded, J2148+0211, is very near both the flux density and intrinsic modulation index cut lines. Of the 17 sources for which we report $\overline{m}$ upper limits, 15 are judged nonvariable by the χ^2 test. The two others are dim sources with a single outlier (J1613+4223) or very few measurements (J1954+6153), which led to large uncertainties in the estimate for $\overline{m}$ and resulted in an upper limit. Calibrator sources 3C 286, DR 21, and 3C 274 are found to be nonvariable while 3C 48 and 3C 161 (which were not used to fit the long-term calibration trend) are found to be variable by the χ^2 test, probably due to imperfect removal of the long-term calibration trend. Our estimates of $\overline{m}$ for both these calibrators are below our 2% intrinsic modulation cut level. We conclude that our analysis is generally consistent with the χ^2 test for identifying significant variability and that our data cuts for our population studies conservatively exclude the regions of parameter space where disagreements occur.

In figure 5.6 we plot the intrinsic modulation index $\overline{m}$ and associated 1σ uncertainty against the "raw" modulation index $m_{\rm data}$ of equation (5.1). The $\overline{m} = m_{\rm data}$ line is shown in blue. Green triangles are the 3σ upper limits of sources for which variability could not be established. Calibrators 3C 286, DR 21, and 3C 274 are plotted in red. Since apparent variability due to the finite accuracy with which individual flux densities can be measured has been corrected out of $\overline{m}$, the expectation is that deviations from the $\overline{m} = m_{\rm data}$ will be more pronounced for sources that are not intrinsically very variable (so that the scatter in the flux density measurements is appreciably affected, and even dominated, by measurement error). In addition, deviations are expected to be below the line, as $\overline{m}$ should be *smaller* than $m_{\rm data}$. Both these expectations are verified by figure 5.5. Note that upper limits need not satisfy this criterion, as the "true" value of the modulation index can take any value below the limit. Upper limits above the blue line are weak, indicating that the reason

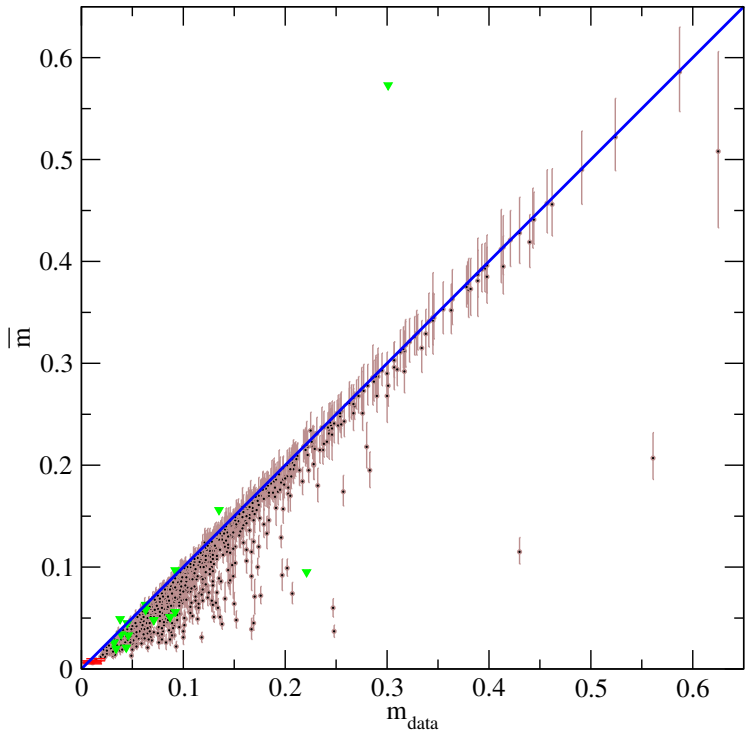

Figure 5.6. Two-year intrinsic modulation index $\overline{m}$ and associated 1σ uncertainty, plotted against the "raw" modulation index, $m_{\rm data}$, of equation (5.1) as black points with brown error bars. The $\overline{m} = m_{\rm data}$ line is shown in blue. Green triangles are the 3σ upper limits of sources for which variability could not be established. Calibrators 3C 286, 3C 274, and DR 21 are plotted in red.

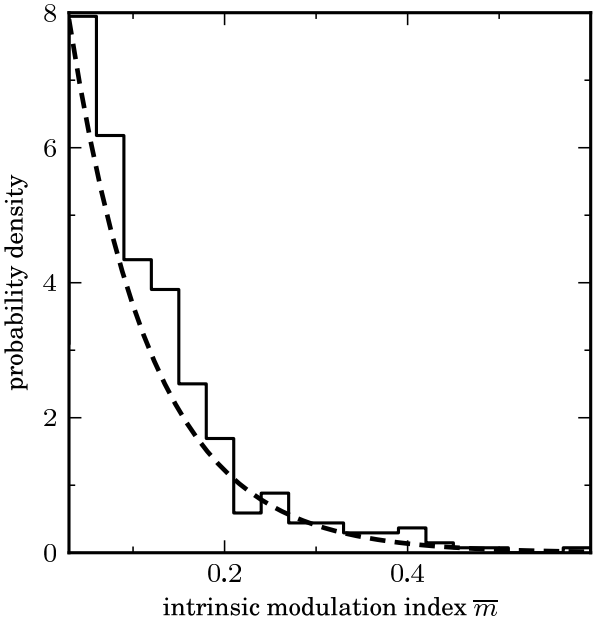

Figure 5.7. Histogram of two-year maximum-likelihood intrinsic modulation indices $\overline{m}$, for the 453 CGRaBS blazars with $S_0 > 400$ mJy, normalized as a probability density that integrates to unity. The dashed line represents an exponential distribution with $\langle \overline{m} \rangle = 0.091$.

variability could not be established is the poor sampling or quality of the data, and not necessarily a low intrinsic variation in the source flux density.

For the 453 CGRaBS objects which have $S_0 > 400$ mJy and for which variability can be established, we plot, in figure 5.7, a histogram of their intrinsic modulation indices $\overline{m}$ normalized so that the vertical axis has units of probability density. The dashed line represents an exponential distribution of mean $\langle \overline{m} \rangle = 0.091$ which, as we can see, is an excellent description of the data. Motivated by this plot, we will be using the monoparametric exponential family of distributions:

$$f(m)dm = \frac{1}{m_0} \exp\left[-\frac{m}{m_0}\right] dm, \tag{5.8}$$

with mean m_0 and variance m_0^2, to characterize various subsamples of our blazar sample.

5.1.2.1 Impact of Longer Time Series

We now compare the results of the intrinsic modulation indices computed from the 42-month data with those computed from the two-year data to look for systematic changes in apparent variability amplitude. The expectation based on the two-year data set was that additional data would tend to increase the variability amplitude on average, since many sources appeared to switch between periods of relatively steady quiescence and periods of active variability. This is because the addition of a a period of steady flux to a source with a

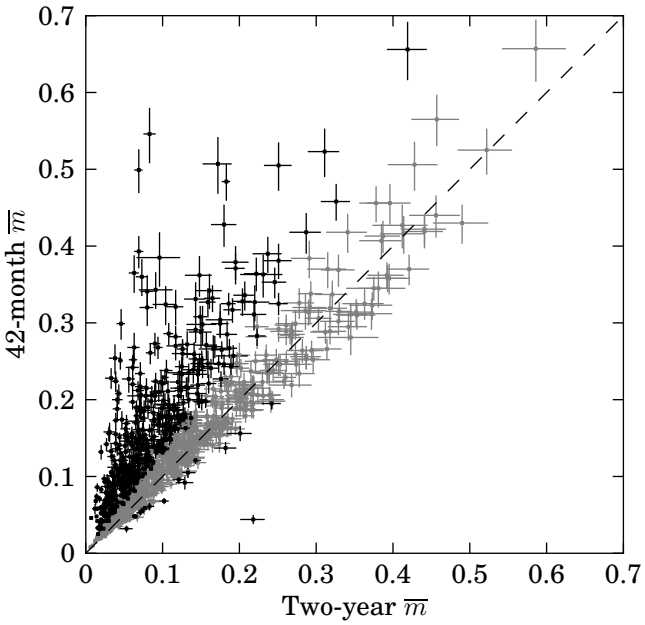

Figure 5.8. Scatter plot of 42-month versus two-year modulation indices for 1135 CGRaBS sources and five calibrators. The single outlier well below the dashed 1:1 reference line is J1154+1225. This source was affected by a single extreme high outlier in the two-year data set which was removed in the 42-month analysis. Sources for which the difference in intrinsic modulation index was less than 3σ are plotted in grey.

history of strong variability will reduce its intrinsic modulation only slightly. On the other hand, a source that has only been observed in a weakly variable state will see a large increase in its intrinsic modulation index if it begins to vary strongly.

The data confirm our expectations. In figure 5.8 we plot the 42-month $\overline{m}$ values against the two-year $\overline{m}$ values for the 1135 CGRaBS sources with measured $\overline{m}$ in both data sets, plus 3C 48, 3C 161, 3C 274, 3C 286, and DR 21. Figure 5.9 shows histograms of the two data sets. Clearly, most sources have 42-month intrinsic modulation indices that are either consistent with or greater than their two-year values. Of the 1140 sources compared, 513 changed by more than 3σ. These are plotted in black in figure 5.8. One significant exception is J1154+1225, a CGRaBS source. In the two-year light curve for this source, a single very large high outlier survived the data cuts. This outlier was eliminated in the 42-month data set, resulting in this large reduction in $\overline{m}$. The actual behavior of this source does not appear to have changed substantially.

This systematic increase in the variability index suggests that the two year interval was insufficiently long to capture the full range of behaviors in many CGRaBS sources. This is not surprising. Based on more than 25 years of monitoring at the University of Michigan Radio Observatory and the Metsähovi Radio Observatory in Finland, Hovatta et al. (2008) report typical flare durations of 2.5 years at 22 and 37 GHz. In Hovatta et al. (2007), typical flaring timescales of 4–6 years, sometimes with evidence for changes on 10 year or longer timescales, were reported. In gamma rays, Abdo et al. (2010b) found higher peak-to-mean

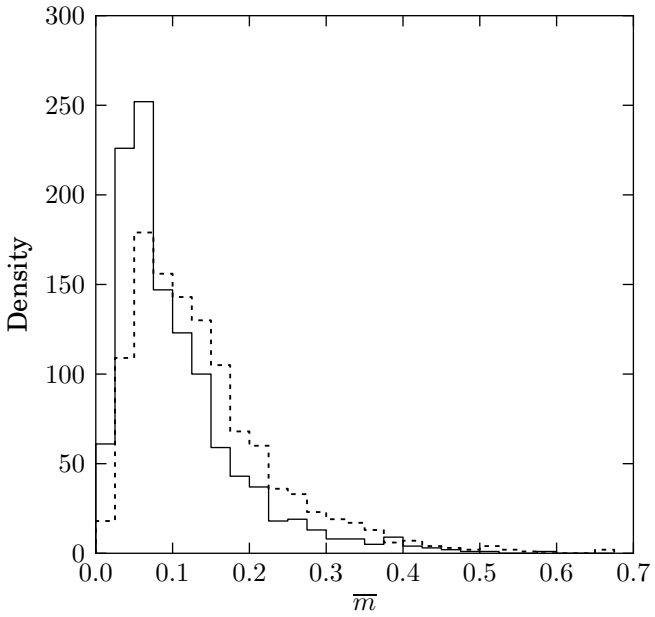

Figure 5.9. Histogram of the intrinsic modulation indices for the two-year (solid line) and 42-month (dashed line) data sets for 1135 CGRaBS sources and five calibrators.

flux ratios among *Fermi* sources compared to those reported by EGRET. It was speculated that this difference could result from the longer EGRET data collection period (4.5 years versus 11 months) permitting sources to visit more emission states. Thus, our results are consistent with these explicit timescale studies and suggest that even with the 42 month data set it is likely somewhat premature to expect our results to characterize the full behavior of the entire sample. However, unless there is, in fact, a connection between radio variability and gamma-ray emission, underestimates of typical variability amplitudes should affect different subpopulations equally. This may diminish our power to compare populations, but it should not lead us to conclude that false correlations exist.

Figure 5.10 shows a histogram of the change in intrinsic modulation index between the two-year and 42-month data sets for each source in the CGRaBS sample. The mean (median) change is 0.035 (0.021). In figure 5.11 we show the light curves of the four sources that showed the most significant change in intrinsic modulation index between the two data intervals. Not surprisingly, these sources all show steady flux density prior to 2010, then either gradually increase or decrease in brightness or exhibit a sharp change in behavior in the more recent data.

5.1.2.2 Impact of Data Outliers

The 40 m data set is not completely free of outlier points that are unlikely to represent intrinsic source variations since the filters described in chapter 3 do not detect all such points. Although most of these outliers are almost surely attributable to poor observing conditions, pointing offset measurement failures, and

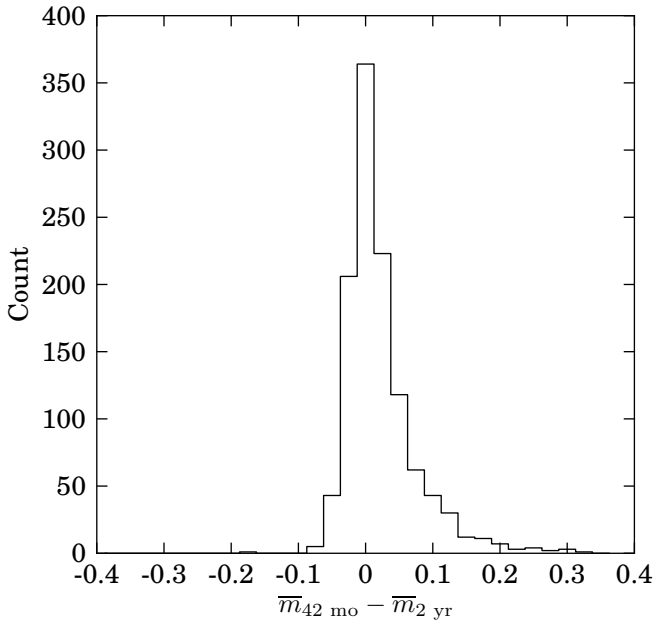

Figure 5.10. Change in intrinsic modulation indices between the two-year and 42-month data sets for 1135 CGRaBS sources and five calibrators.

other causes not related to actual behavior of the astronomical source, we do not delete a data point merely based on its apparently improbable flux density value. Unless we have an unbiased criterion by which we can eliminate the data point, we risk seriously biasing our results by rejecting source variability that does not meet our preconceptions of "reasonable." In many cases, detailed exploration of these data points has led to the discovery of such independent criteria—this is how we developed the set of data filters we employ—but at present, a fair number of these suspected unphysical outliers remain. Perhaps future work will improve the filtering and remove these, but for now we choose to study the impact of such faulty data on our analysis to ensure that our results are robust against their effect.

The most common extreme outliers we encounter are zero or near-zero (in some cases slightly negative) flux density values reported for bright sources. These probably result when the telescope obtains an incorrect pointing offset and measures blank or contaminated sky instead of the desired source. To measure the effect of such outliers, we computed the intrinsic modulation indices for each source using the 42-month data set with the addition of a single flux density value that was twice the average error above zero. Although in some cases light curves are affected by multiple outliers, in general the impact of the first such outlier is much greater than the addition of the second or third—once the probability for such an event is nonzero, the change in likelihood due to a few additional incidents is small. This was verified by comparing the modulation indices of the data set with one false outlier added to the same data with a second false outlier added. The mean change in modulation index was an increase by only 0.001, which is negligible compared to the uncertainties. We therefore only need to consider the impact of a single outlier on the intrinsic modulation index.

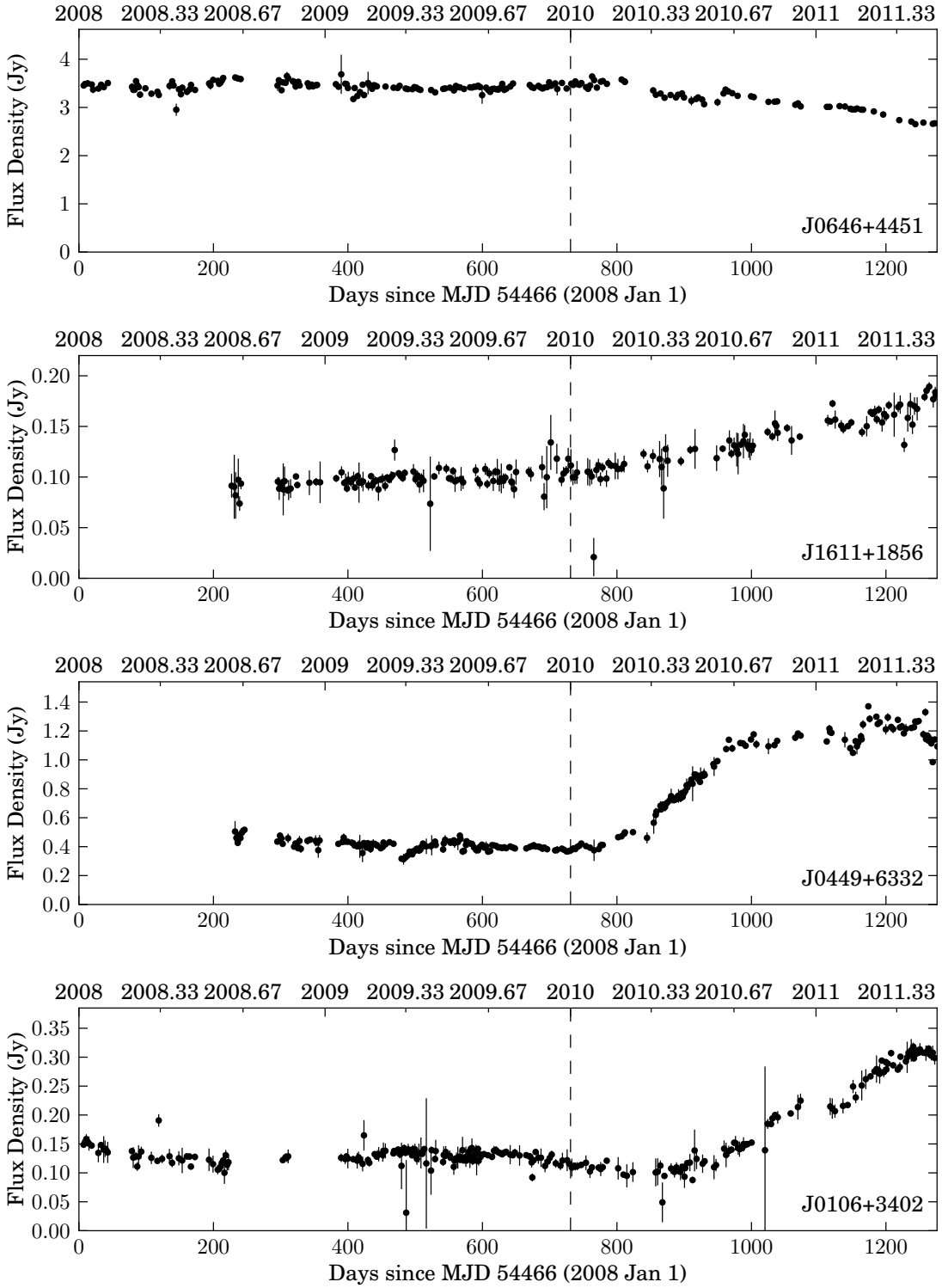

Figure 5.11. Light curves for the four sources with the most significant changes in intrinsic modulation index between the two-year and 42-month data sets. All four sources exhibit an increase in variability after 2010, the end of the period included in the two-year data set, indicated by the dashed line.

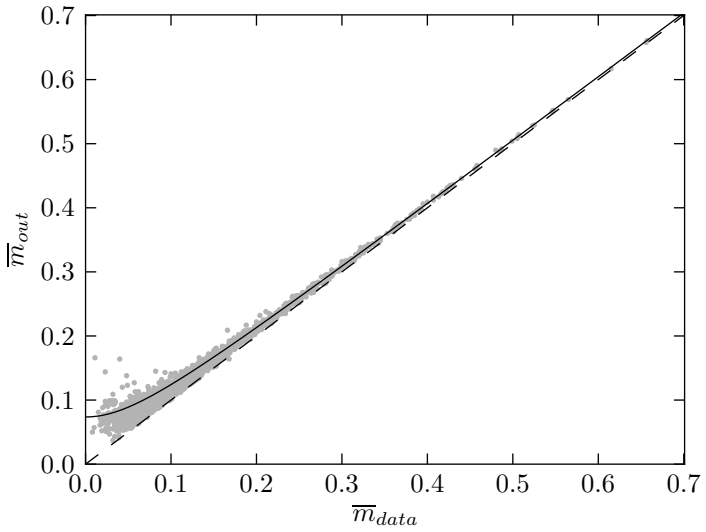

Figure 5.12. Grey points show the modulation index data computed with the addition of an extreme outlier data point plotted against the modulation index for the same source calculated from the actual data. The dashed line shows the ideal $y = x$ line. The solid line shows the effect of adding 0.066 in quadrature with the measured modulation index.

In figure 5.12, the grey points show the computed intrinsic modulation index using the data with the false outlier versus the actual intrinsic modulation index for the source. In figure 5.13, the histograms of these data sets are shown. The obvious visual difference is confirmed by a two-sample K-S test, which rejects the hypothesis of a common parent distribution at the $p \sim 10^{-20}$ level for arising by chance. The mean of the distribution for the real data is 0.143 and for the modified set is 0.160.

If the outliers induced a simple fixed increase in the modulation index for each source, to match the means this would be an increase by 0.018. Applying this increase to the real data, the two sample K-S test rejects the null hypothesis with $p = 6.9 \times 10^{-4}$. A fixed modulation increase in quadrature with the measured value would require an increase of 0.074 to match the means of the two data sets. The two-sample K-S test rejects this possibility with $p = 5.0 \times 10^{-7}$. Apparently, and not surprisingly, the impact of an extreme outlier depends on the properties of the light curve to which it is added.

First, we examine the change in intrinsic modulation index for a trend with the value measured for each source. Figure 5.14 shows the quadrature difference between the intrinsic modulation index with and without the false outlier point plotted against the actual measured value for each source. Although there is significant scatter, it appears that the binned mean of the change is well described by a constant addition of 0.066 in quadrature with the actual measured intrinsic modulation index. A K-S test using this constant quadrature increase is marginally compatible, with $p = 0.047$ to reject the hypothesis of a common distribution. The effect of adding this constant additional error is also shown in figure 5.12 by the solid line.

In figure 5.15 we show the same quadrature difference, now as a function of the maximum likelihood average source flux density. A linear fit to binned mean square of the quadrature difference as a function of

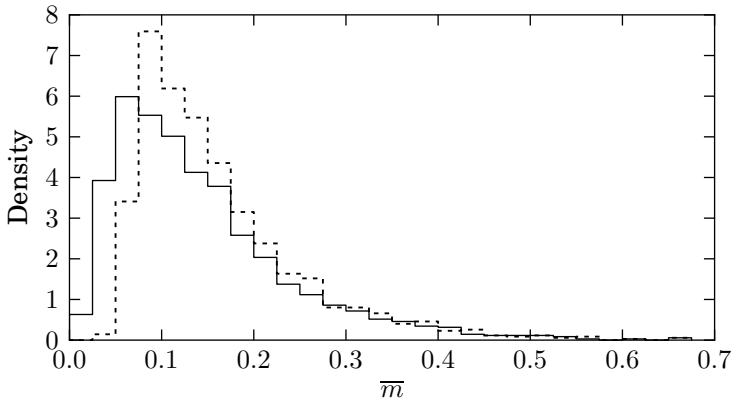

Figure 5.13. Histograms of intrinsic modulation indices for all sources using the 42-month data (solid line) and the 42-month data with an additional extreme outlier added to each light curve (dotted line). Includes 1396 sources in each, excluding those for which only upper limits were calculated. Both curves are normalized to integrate to unity.

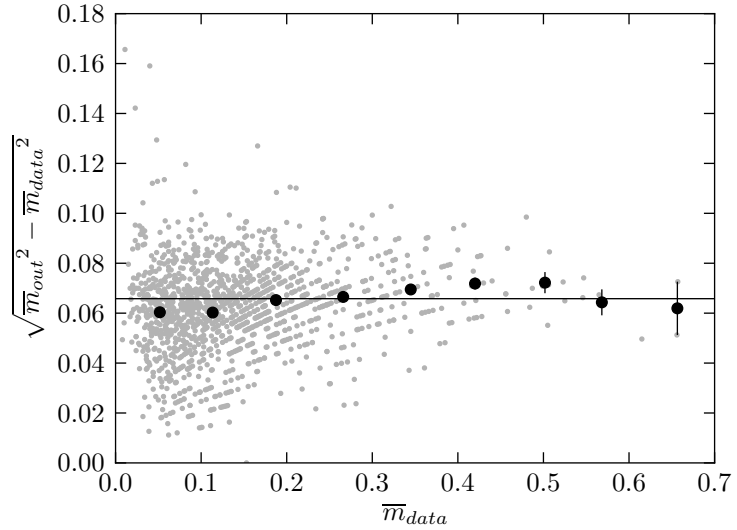

Figure 5.14. Grey points show the quadrature difference between the modulation indices calculated with the addition of an extreme outlier data point and the modulation indices computed from the actual data, plotted against the actual modulation index for the source. The black points show the binned mean difference in modulation index. The line is the constant mean quadrature difference, 0.066. The striping evident in the grey points is the result of rounding the modulation index values before computing the quadrature difference.

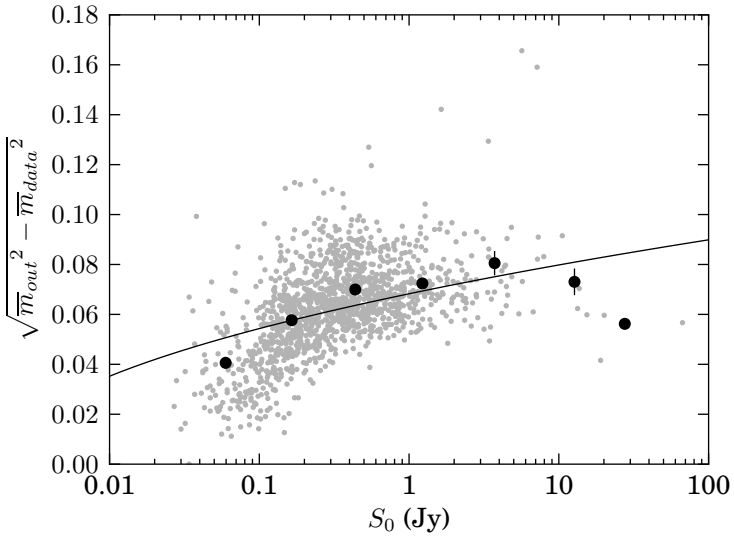

Figure 5.15. Grey points show the quadrature difference between the modulation indices calculated with the addition of an extreme outlier data point and the modulation indices computed from the actual data, plotted against the maximum likelihood average flux density for the source. The black points show the binned mean difference in modulation index using logarithmic bins in flux density. The line is a linear fit to the square of the binned y data as a function of the logarithm of the x data, excluding the last data point.

the logarithm of the flux density is shown by the solid line, which corresponds to

$$\Delta(\overline{m}^2) = (0.041)^2 \log_{10}\left(\frac{S_0}{1\text{ Jy}}\right) + (0.068)^2. \tag{5.9}$$

In this fit, we have excluded the last binned point because only a few data are in the bin. The lowest flux density bin lies significantly below the trend, suggesting that sources with mean flux densities below about 100 mJy were less affected by the addition of an outlier.

Although a trend with source flux density is apparent, the systematic quadrature change in intrinsic modulation index induced by the addition of an outlier varies only between about 0.04 and 0.08. Thus, when comparing two populations, unless the flux density distributions differ widely, this trend is unlikely to distort the results of our population studies.

It is difficult to define a robust criterion for identifying an outlier in a data set which includes sources that exhibit extreme actual variability. To estimate the number of such outliers, we choose a subset of the sources for which it is relatively easy to detect an outlier. Since the processes that result in a nonphysical outlier are connected to the instrument or observation conditions rather than astronomical effects, we can safely assume that the fraction of affected sources in such a subset is representative of the entire sample. We therefore examine bright sources with median flux density $\mathrm{median}(S) \geq S_{min}$, where low outliers are more prominent. To avoid extremely variable sources, we further restrict this study to sources with most measurements near the median. We require that at least a fraction R_{stable} of the flux densities for the source

Table 5.1. Fraction of sources determined to be affected by an extreme low outlier

Test Parameters				Test Results			
S_{min} (Jy)	$R_{central}$	R_{stable}	$R_{outlier}$	N_{total}	N_{stable}	$N_{outlier}$	Outlier Fraction (%)
1.0	0.5	0.5	0.3	163	162	8	4.9
1.0	0.5	0.5	0.2	163	162	7	4.3
0.5	0.5	0.5	0.3	408	407	34	8.4
0.5	0.8	0.5	0.3	408	407	34	8.4
0.5	0.5	0.8	0.3	408	387	30	7.8
0.5	0.5	0.8	0.2	408	387	22	5.7

Note: The estimates do not vary rapidly for small changes in test parameters. We adopt an approximate outlier fraction of 8% for our population, meaning that about 8% of sources are affected by at least one significant outlier.

be within a $\pm R_{central} \times \mathrm{median}(S)$ of the median. We then count the source as having an outlier if any flux density in its light curve is below $R_{outlier} \times \mathrm{median}(S)$. In table 5.1 we tabulate the results of this test for various values of the parameters. We conclude that no more than about 8% of sources are affected by extreme outliers.

It is extremely unlikely that the incidence of such outliers is tied to physical properties of the sources being observed, so we can reasonably assume that about 8% of the sources in any of the subsamples we select for population studies are affected by outliers. We conclude from these studies that the net effect of outliers in the data set is to add a false modulation of about 0.066 in quadrature with the intrinsic modulation index that would be measured in the absence of such outliers. This increase will affect about 8% of the sources in any physically selected sample. When comparing samples, these conclusions are valid as long as the flux density distributions of the two samples have similar dynamic ranges. If one sample is substantially different in flux density, particularly if it is clustered below about 100 mJy, then the inclusion of outliers in the data set may affect the samples differently.

More sources are likely to be affected by random outliers in the 42-month data set than the two-year data set, simply because each additional observation provides an opportunity for an outlier. Assuming the outlier incidence rate is constant, a rough estimate of the number of sources affected in the two-year data set is $8\% \times 24/42 = 5\%$. In other words, 3% fewer sources are likely to be affected in the two-year data set. For simplicity, if we assume the 0.066 average modulation index increase were added linearly, the additional 3% of affected sources would be expected to add about $0.03 \times 0.066 = 0.002$ to the mean modulation index of the population. Because the addition is in quadrature, this is an overestimate of the likely impact. The actual mean increase of 0.035 discussed in section 5.1.2.1, therefore, cannot be explained by the additional exposure to outliers and can be safely attributed to changes in the observed source behavior.

5.1.3 A Formalism for Population Studies

We now turn our attention to whether the intrinsic variability amplitude at 15 GHz, as quantified by $\overline{m}$, correlates with the physical properties of the sources in our sample. To this end, we will determine the *distribution* of intrinsic variability indices $\overline{m}$ for various subsets of our monitoring sample, and we will examine whether the various subsets are consistent with being drawn from the same distribution.

We will do so using again a likelihood analysis. We will assume that the distribution of $\overline{m}$ in any subset is an exponential distribution of the form given in equation (5.8). Since distributions of this family are uniquely described by the value of the mean, m_0, our aim is to determine m_0, or rather the probability distribution of possible m_0 values, in any specific subset.

The likelihood of a single observation of a modulation index $\overline{m}_i$ of Gaussian uncertainty σ_i drawn from an exponential distribution of mean m_0 is

$$\begin{aligned} \ell_i &= \int_{\overline{m}=0}^{\infty} d\overline{m} \frac{1}{m_0} \exp\left(-\frac{\overline{m}}{m_0}\right) \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left[-\frac{(\overline{m}-\overline{m}_i)^2}{2\sigma_i^2}\right] \\ &= \frac{1}{m_0\sigma_i\sqrt{2\pi}} \exp\left[-\frac{\overline{m}_i}{m_0}\left(1-\frac{\sigma_i^2}{2m_0\overline{m}_i}\right)\right] \times \\ &\quad \int_{\overline{m}=0}^{\infty} d\overline{m} \exp\left[-\frac{[\overline{m}-(\overline{m}_i-\sigma_i^2/m_0)]^2}{2\sigma_i^2}\right], \end{aligned} \tag{5.10}$$

where, to obtain the second expression, we have completed the square in the exponent of the integrand. The last integral can be calculated analytically, yielding

$$\begin{aligned} \ell_i &= \frac{1}{2m_0} \exp\left[-\frac{\overline{m}_i}{m_0}\left(1-\frac{\sigma_i^2}{2m_0\overline{m}_i}\right)\right] \times \\ &\quad \left\{1+\mathrm{erf}\left[\frac{\overline{m}_i}{\sigma_i\sqrt{2}}\left(1-\frac{\sigma_i^2}{m_0\overline{m}_i}\right)\right]\right\}. \end{aligned} \tag{5.11}$$

If we want (as is the case for our data set) to implement data cuts that restrict the values of $\overline{m}_i$ to be larger than some limiting value m_l, the likelihood of a single observation of a modulation index $\overline{m}_i$ will be the expression above multiplied by a Heaviside step function, and renormalized so that the likelihood $\ell_{i,\mathrm{cuts}}$ to obtain any value of $\overline{m}_i$ *above* m_l is 1:

$$\ell_{i,\mathrm{cuts}}[m_l] = \frac{H(\overline{m}_i - m_l)\ell_i}{\int_{\overline{m}_i=m_l}^{\infty} d\overline{m}_i \ell_i}. \tag{5.12}$$

This renormalization enforces that there is no probability density for observed events "leaking" in the parameter space of rejected $\overline{m}_i$ values. In this way, it "informs" the likelihood that the reason why no objects of $\overline{m}_i < m_l$ are observed is not because such objects are not found in nature, but rather because we have excluded them "by hand."

The integral in the denominator is analytically calculable,

$$\int_{\overline{m}_i=m_l}^{\infty} d\overline{m}_i \ell_i = \frac{1}{2}\left\{\exp\left(\frac{\sigma_i^2}{2m_0^2} - \frac{m_l}{m_0}\right)\times \left[1+\mathrm{erf}\left(\frac{m_l}{\sigma_i\sqrt{2}} - \frac{\sigma_i}{m_0\sqrt{2}}\right)\right] +1-\mathrm{erf}\left(\frac{m_l}{\sigma_i\sqrt{2}}\right)\right\}. \quad (5.13)$$

The likelihood of N observations of this type is

$$\mathcal{L}(m_0) = \prod_{i=1}^{N} \ell_{i,\mathrm{cuts}}[m_l]\,. \quad (5.14)$$

If we wish to study two parts of the S_0 parameter space with different cuts (as in, for example, figure 5.5, where we have a cut of $m_l = 0.02$ for $S_0 > 0.4$ Jy, and a different cut of $m_u = 0.06$ for 0.06 Jy $\leq S_0 \leq 0.4$ Jy), we can implement this in a straightforward way, by considering each segment of the S_0 parameter space as a distinct "experiment," with its own data cut. If the first "experiment" involves N_l objects surviving the m_l cut, and the second "experiment" involves N_u objects surviving the m_u cut, then the overall likelihood will simply be

$$\mathcal{L}(m_0) = \prod_{i=1}^{N_l} \ell_{i,\mathrm{cuts}}[m_l] \prod_{i=1}^{N_u} \ell_{i,\mathrm{cuts}}[m_u]\,. \quad (5.15)$$

Maximizing equation (5.15) we obtain the maximum-likelihood value of m_0, $m_{0,\mathrm{maxL}}$. Statistical uncertainties on this value can also be obtained in a straightforward way, as equation (5.15), assuming a flat prior on m_0, gives the probability density of the mean intrinsic modulation index m_0 of the subset under study.

5.2 Null Tests

Here we begin to apply the formalism introduced in section 5.1.3 to examine whether the intrinsic modulation index $\overline{m}$ correlates with the properties of the sources in our sample. We will be testing whether the distributions of $\overline{m}$-values in subsets of our monitoring sample split according to some source property are consistent with each other. Before considering physically motivated population splits, we need to be certain that our formalism is correctly implemented. To verify that our analysis does not yield spurious results, we first discuss three null test cases where the likelihood analysis *should not* find a difference in the variability properties of the different subsets considered.

Because these null tests are tests of the method rather than tests of the source populations, we do not expect any significant change between the two-year and 42-month data sets. We therefore performed two of the three tests with only the two-year data set. As a sanity check we verified that the test described in section 5.2.2 below returned a null result with both two-year and the 42-month data sets.

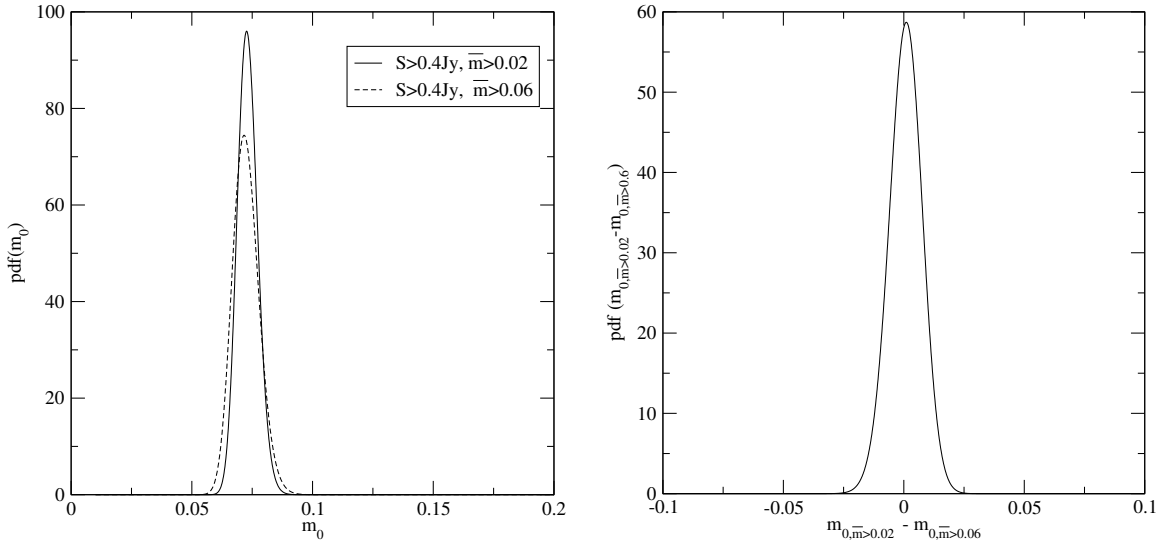

Figure 5.16. Verification that the data cuts described in section 5.1.2 are correctly implemented, using the two-year CGRaBS data. *Left:* Probability density of m_0 for the subset of bright CGRaBS blazars not found in 1LAC, for two values of the cutoff for data acceptance: $m_l = 0.02$ (solid line, maximum-likelihood value and 1σ error $m_0 = 0.073 \pm 0.004$), and $m_l = 0.06$ (dashed line, maximum-likelihood value and 1σ error $m_0 = 0.072^{+0.006}_{-0.005}$). The two distributions are consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The difference (0.001 ± 0.007) is consistent with zero within 1σ.

This and the following sections include a number of histograms of modulation index for various subpopulations of the sample. It is important to remember that we have employed the data cuts described in section 5.1.2 prior to constructing these subpopulations. This distorts the appearance of the lowest bins in these histograms where we have excluded the parameter space region in which our sampling is incomplete. The likelihood analysis is properly "informed" that this parameter space region has been excluded, so this is merely an aesthetic effect.

5.2.1 Verifying Data Cuts

The first case tests whether the data cuts discussed in section 5.1.2 are implemented correctly in section 5.1.3. To this end, we calculate $\mathcal{L}(m_0)$ for the set of gamma-ray–quiet CGRaBS blazars (blazars not found in 1LAC) in our monitoring sample with $S > 0.4$ Jy, in two different ways: first, by applying an $\overline{m}$ cut at $m_l = 0.02$; second, by applying an $\overline{m}$ cut at $m_l = 0.06$ (a much more aggressive cut than necessary for the particular bright blazar population). The increased value of m_l in the second case should not affect the result other than by reducing the number of data points and thus resulting in a less constraining likelihood for m_0. This is indeed the case, as we see in the left panel of figure 5.16, where we plot the probability density of m_0 for the two subsets. That the two distributions are consistent with each other is explicitly demonstrated in the right panel of figure 5.16, where we plot the probability density of the *difference* between the means m_0 of the two

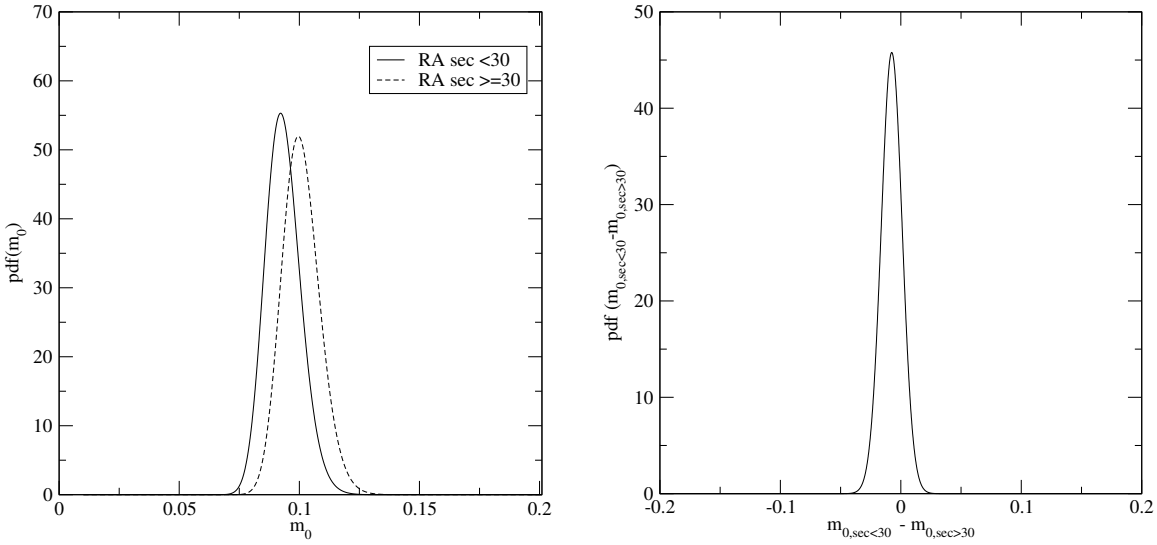

Figure 5.17. Null test verification using the two-year data. *Left:* Probability density of m_0 for bright CGRaBS blazars with seconds of right ascension < 30 s (solid line, maximum-likelihood value and 1σ error $m_0 = 0.088 \pm 0.006$) or ≥ 30 s (dashed line, maximum-likelihood value and 1σ error $m_0 = 0.096^{+0.007}_{-0.006}$). The two distributions are consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The difference (-0.008 ± 0.009) is consistent with zero within 1σ.

subsets (which is formally equal to the cross-correlation of their individual distributions). The difference is consistent with zero within 1σ.

5.2.2 Physically Insignificant Population Split

The second case tests whether a split according to a source property *without* physical meaning and with the *same* value for the cutoff modulation index m_l will yield probability densities for the m_0 that are consistent with each other. For this reason, we split the population of bright ($S > 0.4$ Jy) blazars in the sample into two subsets in the following way: we divide the right ascension of each source by 1 min. If the remainder of this operation is <30 s, we include this source in the first subsample; if the remainder is ≥ 30 s we include the source in the second subsample.

We first applied this test to the two-year CGRaBS data set. The results are shown in figure 5.17. As expected, the probability distributions of m_0 for the two subsamples, shown in the left panel of figure 5.17, are consistent with each other. This is explicitly demonstrated in the right panel, which shows the probability density of the difference between the m_0 in the two subsamples. The difference is consistent with zero within 1σ.

As a sanity check, we repeated this null test with the 42-month data set including bright ($S > 0.4$ Jy) sources from both the CGRaBS and 1LAC samples. The resulting probability distributions are shown in figure 5.18. Again, the populations are consistent with each other to well within 1σ. The most likely values for the m_0 have increased for both bins relative to the two-year results. This is because we now include the

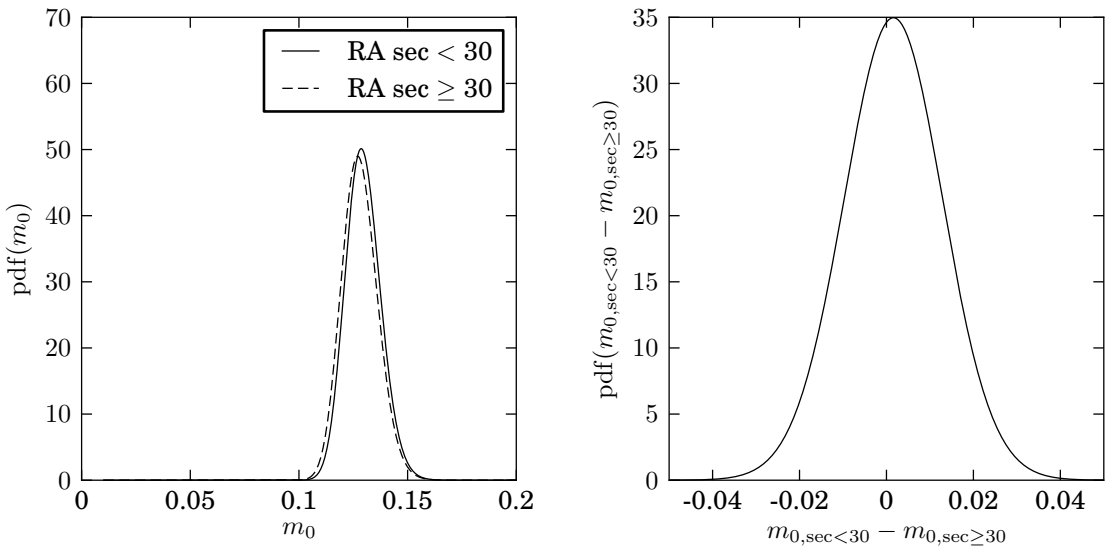

Figure 5.18. Null test verification using the 42-month data. *Left:* Probability density of m_0 for sources with seconds of right ascension <30 s (solid line, maximum likelihood value and 1σ error $0.127^{+0.009}_{-0.008}$) and ≥ 30 s (dashed line, maximum likelihood value and 1σ error 0.129 ± 0.008). The two distributions *are* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution ($0.002^{+0.011}_{-0.012}$) is much less than 1σ from zero, as expected.

more variable 1LAC sample (see section 5.3) and because we have increased the time series length for the CGRaBS sample which, in section 5.1.2.1, we showed leads to a systematic increase in measured variability amplitude. In figure 5.19 we show the histograms of the intrinsic modulation indices in the two populations for the 42-month data set.

5.2.3 Galactic Latitude Split

In the final test case, we use the two-year CGRaBS data set to examine whether a split in galactic latitude yields consistent probability densities for the two subsamples. Again, we expect consistent distributions because this division does not reflect an intrinsic physical property of the sources. For this test, we restrict the sample to bright ($S > 0.4$ Jy) FSRQs and use the cutoff modulation index $m_l = 0.02$. We split between low- and high-galactic latitude at $|b| = 39°$. This produces similarly sized subsamples (181 and 168 for low- and high-latitude, respectively). The left panel in figure 5.20 shows the probability distributions for m_0 for these two subsamples, which, as anticipated, are consistent with each other. The right panel of figure 5.20 shows the probability density for the difference between m_0 for the two subsamples, which is consistent with zero to within 1σ.

This test is not a pure null test because there is a potentially significant observational difference between the sources in the two populations. Sources at lower galactic latitudes are more likely to be affected by scintillation due to the galactic interstellar medium (e.g., Rickett et al. 2006). However, due to the galactic

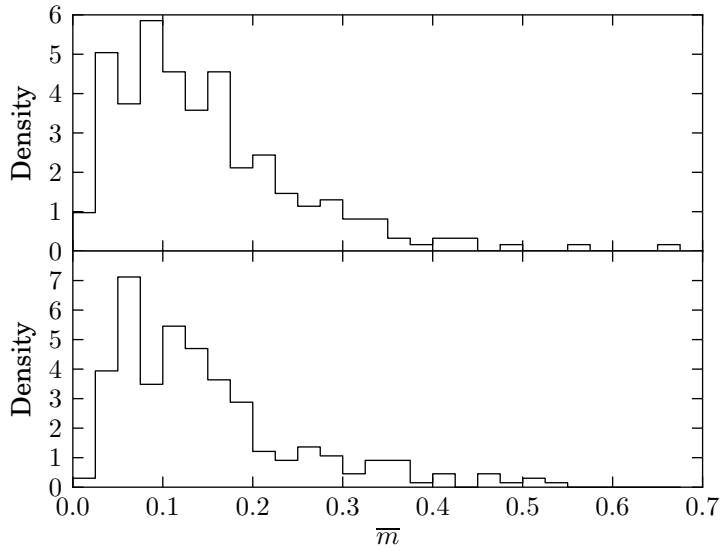

Figure 5.19. Histograms of 42-month intrinsic modulation index values for bright ($S_0 > 0.4$ Jy) sources with seconds of right ascension ≥ 30 s (top, 246 sources) and < 30 s (bottom, 264 sources). Each histogram is normalized to integrate to unity.

latitude cut $|b| \geq 10^\circ$ that defines our samples and the relatively high observation frequency, this was not expected to be a significant effect. This expectation was confirmed by our results.

5.3 Gamma-Ray Loud versus Quiet Populations

We now turn from null tests to studying subsets defined according to physical properties of the sources. The first criterion we apply is whether the source has been detected by *Fermi*-LAT at a significance level high enough to warrant inclusion in the 1LAC catalog. For sources with $S_0 < 0.4$ Jy we apply a cut $\overline{m} > m_u = 0.06$ and for sources with $S_0 \geq 0.4$ Jy a cut $\overline{m} > m_l = 0.02$. The results for the two-year data set are shown in figure 5.21. The set of sources that are included in 1LAC is depicted by a solid line, while the set of sources that are not in 1LAC is depicted by a dashed line. The two are not consistent with each other at a confidence level of 6σ (right panel of figure 5.21), with a maximum-likelihood difference of 5.7 percentage points, with gamma-ray–loud blazars exhibiting, on average, a higher variability amplitude *by almost a factor of 2* versus gamma-ray–quiet blazars.

This significant difference persists in the 42-month data set. Here again, we have considered a CGRaBS source to be gamma-ray loud if it appeared in the 1LAC sample. The likelihood distributions for the population mean intrinsic modulation indices are shown in the left panel of figure 5.22. The right panel shows the probability density for the difference in m_0 between the two populations. The values for both subpopulations are somewhat larger than before, but the most likely difference has increased and continues to be significant

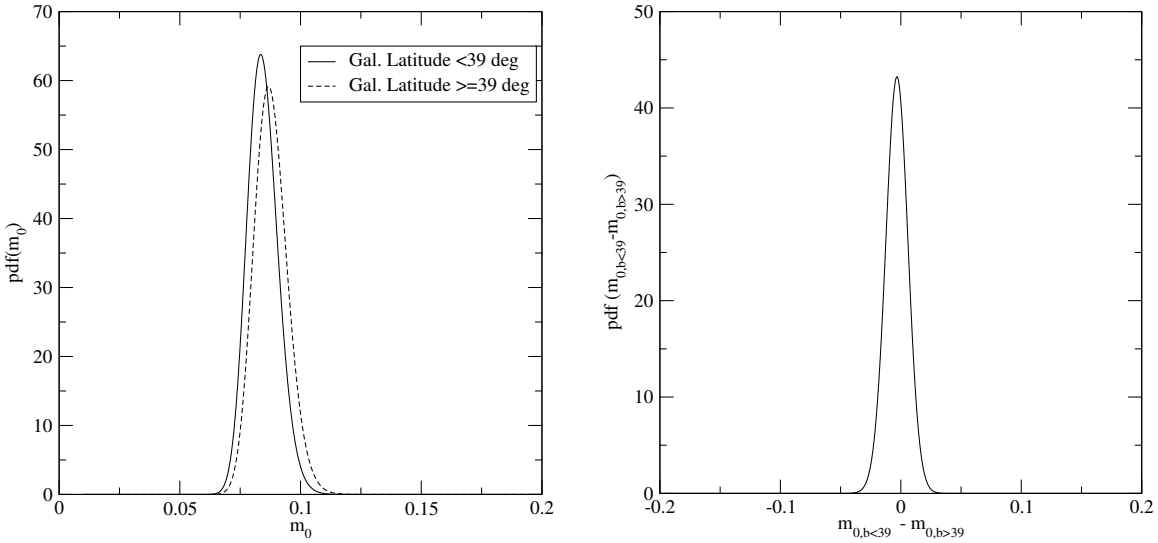

Figure 5.20. Comparison of bright ($S > 0.4$ Jy) FSRQs at high and low galactic latitude using the two-year data set. *Left:* Probability density of m_0 for the subset of bright ($S > 0.4$ Jy) CGRaBS FSRQs with $\overline{m} > 0.02$ at low galactic latitude ($|b| < 39°$, solid line, maximum-likelihood value and 1σ error $m_0 = 0.084^{+0.007}_{-0.006}$) or high galactic latitude ($|b| \geq 39°$, dashed line, maximum-likelihood value and 1σ error $m_0 = 0.087^{+0.007}_{-0.006}$). The two distributions are consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The difference (-0.003 ± 0.009) is consistent with zero within 1σ.

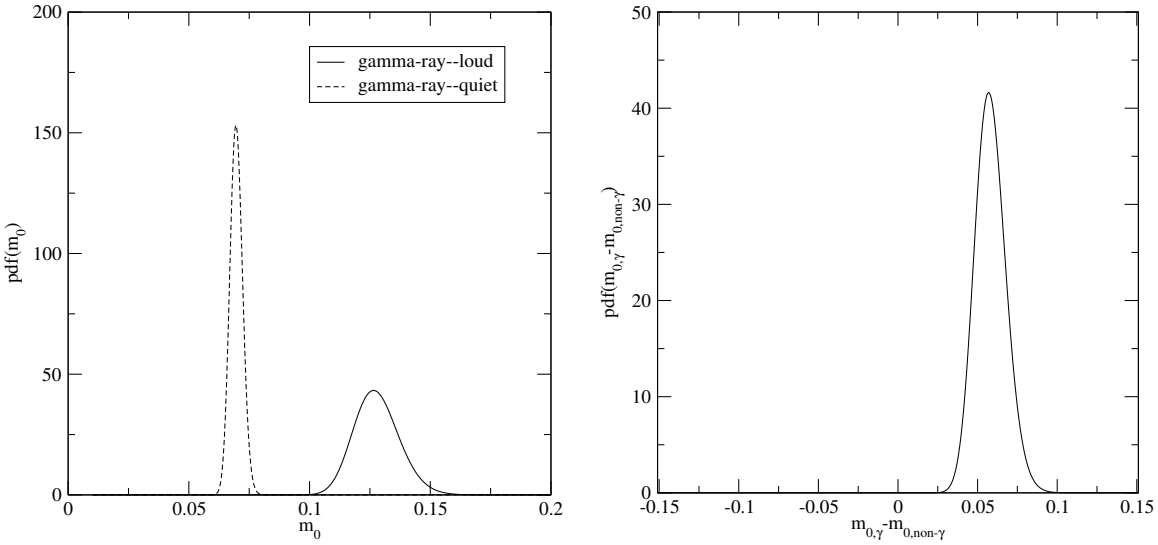

Figure 5.21. Comparison of gamma-ray–loud and gamma-ray–quiet CGRaBS populations using the two-year data set. *Left:* Probability density of m_0 for CGRaBS blazars in our monitoring sample that are (solid line, maximum-likelihood value and 1σ error $m_0 = 0.127^{+0.010}_{-0.009}$) and are not (dashed line, maximum-likelihood value and 1σ error $m_0 = 0.070 \pm 0.003$) included in 1LAC. The two distributions are *not* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution ($0.057^{+0.010}_{-0.009}$) is 6σ away from zero.

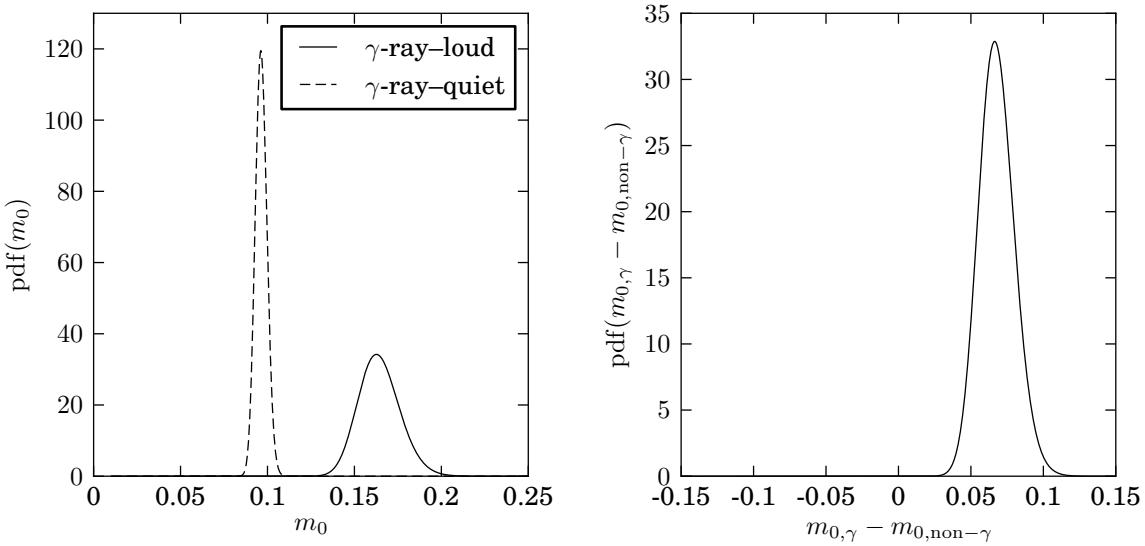

Figure 5.22. Comparison of gamma-ray–loud and gamma-ray–quiet CGRaBS populations using the 42-month data set. *Left:* Probability density of m_0 for the CGRaBS blazars in our monitoring sample that are (solid line, maximum likelihood value and 1σ error $0.163^{+0.012}_{-0.011}$) and are not (dashed line, maximum likelihood value and 1σ error 0.096 ± 0.003) included in 1LAC. The two distributions are *not* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution ($0.066^{+0.013}_{-0.012}$) is more than 6σ away from zero.

at the 6σ level. In figure 5.23, we plot the histograms of intrinsic modulation indices for the gamma-ray–loud and gamma-ray–quiet CGRaBS subpopulations using the 42-month data.

Clearly the gamma-ray–loud subset of the CGRaBS sample is much more variable, on average, than the gamma-ray–quiet subset. While this is likely an important clue about the physical conditions necessary for the production of observable gamma rays in a blazar, before further discussing this result we will examine the variability difference between other subsets of our sample. Both the CGRaBS and 1LAC samples contain a variety of source types and properties, so this may shed further light on the significant differences between gamma-ray–loud and gamma-ray–quiet sources.

5.4 BL Lac Object versus FSRQ Populations

We next examine the variability amplitude as a function of optical spectral classification. In this section we consider the BL Lac and FSRQ subsets of our samples and examine whether they differ in terms of 15 GHz variability. First, we examine the CGRaBS sample using the two-year data set. The probability densities for the mean m_0 of the two subsets are shown in the left panel of figure 5.24. The results for BL Lacs (FSRQs) are plotted as a solid (dashed) line. The two curves are not consistent with each other—the BL Lacs appear to have, on average, higher variability amplitude than the FSRQs. We verify this finding by plotting, in the right panel, the probability density of the difference between the m_0 of BL Lacs and FSRQs. The most likely difference is 3.2 percentage points, and it is more than 3σ away from zero. Note that the difference between

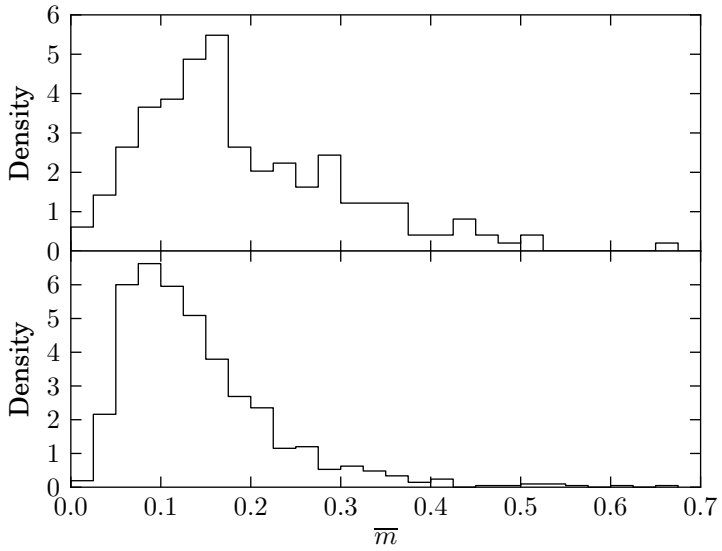

Figure 5.23. Histograms of 42-month intrinsic modulation index values for CGRaBS sources detected (top, 197 sources) and not detected (bottom, 833 sources) in gamma rays by the LAT in the 1LAC catalog. Each histogram is normalized to integrate to unity.

BL Lacs and FSRQs is less significant than that between gamma-ray–loud and gamma-ray–quiet blazars. This is both because the most likely difference in m_0 values between the BL Lac and FSRQ subsets is smaller and because the BL Lac sample is smaller than the gamma-ray–loud blazar sample: only 94 BL Lacs satisfy the data cuts we impose, versus 191 gamma-ray–loud blazars of all types. As a result, the constraints on the intrinsic distribution of modulation indices (i.e., on m_0) are stronger in the latter case.

Turning now to the 42-month data set, we first examine the CGRaBS population. The likelihood distributions for the population mean intrinsic modulation indices are shown in figure 5.25 and figure 5.26 shows the histogram of the intrinsic modulation indices for sources identified as BL Lacs or as FSRQs within the CGRaBS sample. The significant distinction in variability between CGRaBS BL Lac and FSRQ sources remains, and continues to appear at the $>3\sigma$ confidence level.

In figure 5.27, we show the 42-month results for the 1LAC sample, and in figure 5.28 we show the modulation index histograms. Using this sample, the most likely difference in variability amplitude between the two samples is now found to be less than 2σ significant. Remarkably, the difference has not only faded in significance, but the sign of the most likely difference has switched. In the 1LAC sample, the BL Lac subpopulation is found to be *less* variable than the FSRQ subpopulation, whereas BL Lacs were found to be *more* variable in the CGRaBS sample. We will discuss this further in section 5.6 below.

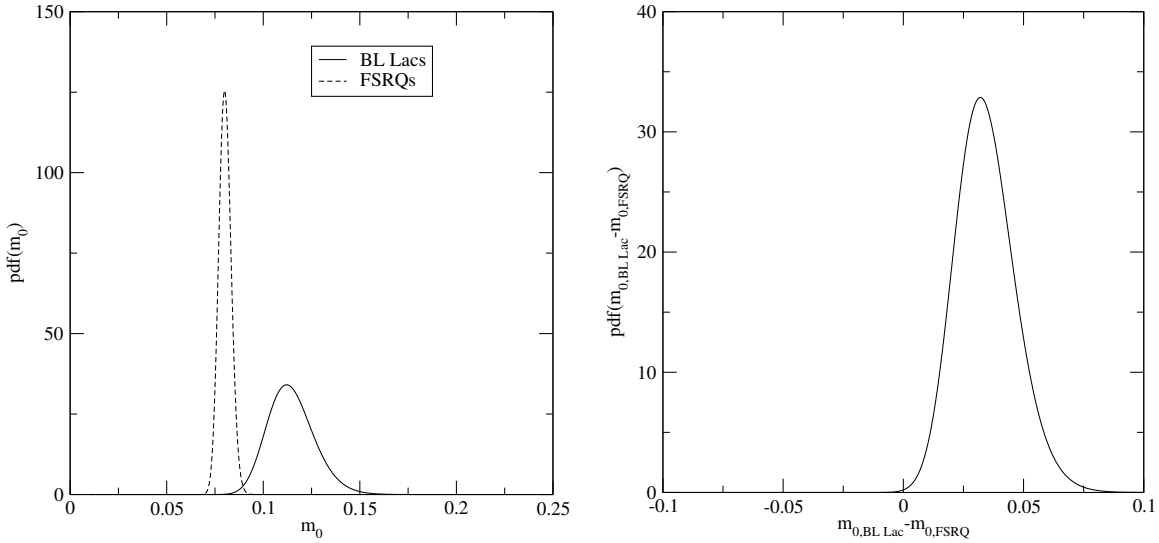

Figure 5.24. Comparison of CGRaBS BL Lac and FSRQ populations using two-year data. *Left:* Probability density of m_0 for BL Lac (solid line, maximum-likelihood value and 1σ error $m_0 = 0.112^{+0.013}_{-0.011}$) and FSRQ (dashed line, maximum-likelihood value and 1σ error $m_0 = 0.080 \pm 0.003$) CGRaBS blazars in our monitoring sample. The two distributions are *not* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution ($0.032^{+0.013}_{-0.011}$) is more than 3σ away from zero.

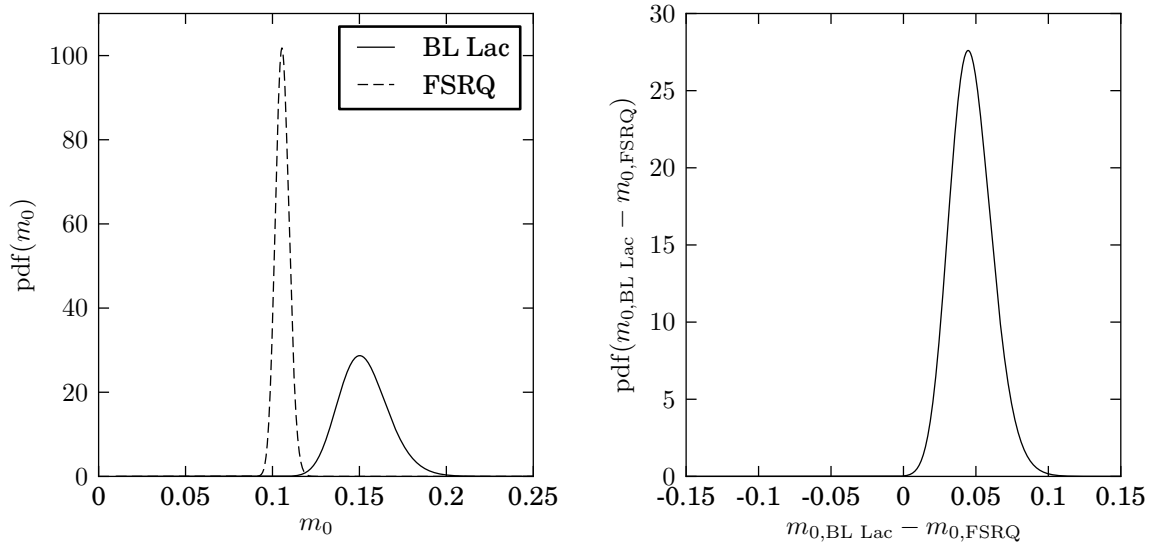

Figure 5.25. Comparison of CGRaBS BL Lac and FSRQ populations using 42-month data. *Left:* Probability density of m_0 for the CGRaBS blazars in our monitoring sample that are identified as BL Lac (solid line, maximum likelihood value and 1σ error $0.150^{+0.015}_{-0.013}$) and as FSRQ (dashed line, maximum likelihood value and 1σ error 0.105 ± 0.004). The two distributions are *not* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution ($0.045^{+0.015}_{-0.014}$) is about 3.5σ away from zero.

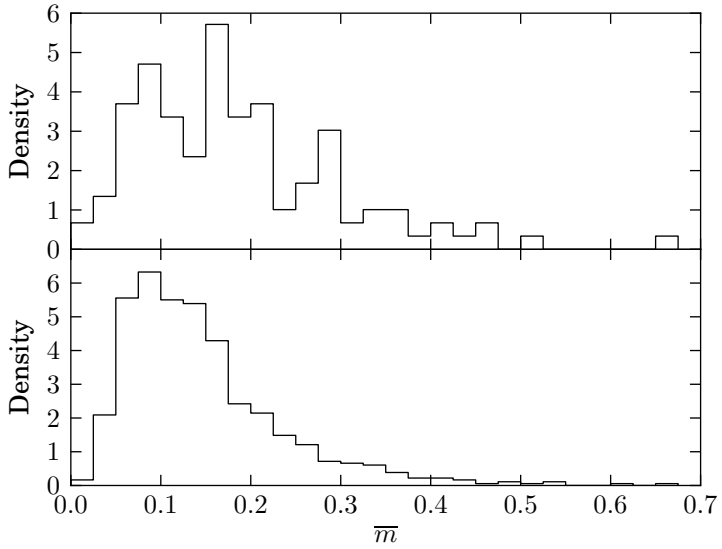

Figure 5.26. Histograms of 42-month intrinsic modulation index values for CGRaBS sources classified as BL Lac (top, 119 sources) and as FSRQ (bottom, 727 sources). Each histogram is normalized to integrate to unity.

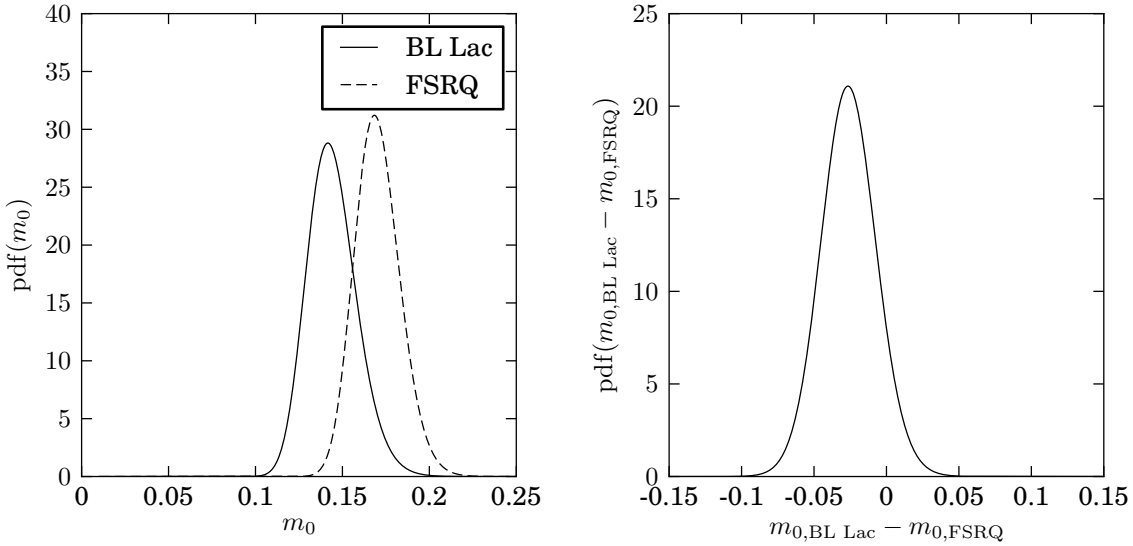

Figure 5.27. Comparison of 1LAC BL Lac and FSRQ populations using 42-month data. *Left:* Probability density of m_0 for the 1LAC blazars in our monitoring sample that are identified as BL Lac (solid line, maximum likelihood value and 1σ error $0.142^{+0.015}_{-0.013}$) and as FSRQ (dashed line, maximum likelihood value and 1σ error $0.168^{+0.014}_{-0.012}$). The two distributions *are* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution (0.027 ± 0.019) is less than 2σ away from zero.

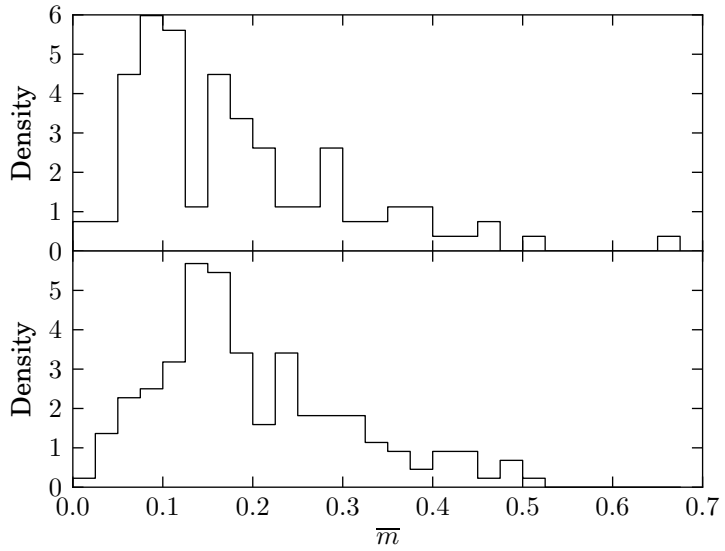

Figure 5.28. Histograms of 42-month intrinsic modulation index values for 1LAC sources classified as BL Lac (top, 107 sources) and as FSRQ (bottom, 176 sources). Each histogram is normalized to integrate to unity.

5.5 Redshift Trend

Finally, we examine the dependence of variability amplitude on redshift. In the left panel of figure 5.29 we plot the mean $\overline{m}$ (as calculated by a simple average rather than the likelihood analysis) in redshift bins of $\Delta z = 0.5$ for bright ($S \geq 400$ mJy) CGRaBS FSRQs with known redshifts in our monitoring sample, using the two-year data set. We exclude BL Lacs from this analysis so as not to bias the result, as BL Lacs with known redshifts are located at low z, and we have also already shown that they have a higher mean $\overline{m}$ compared to FSRQs within the CGRaBS sample. Although the errors are large, there is a hint of a trend toward decreasing variability amplitude with increasing redshift. We further test the significance of this result by splitting sources in our monitored sample in high- and low-redshift subsets with the dividing redshift at $z = 1$ (dashed line in figure 5.29). In the two subsets we also include faint ($S < 400$ mJy) sources, with the usual cut at $m_u = 0.06$. The probability density for the mean m_0 of each subset is shown in the left panel of figure 5.30, where the solid curve corresponds to low-redshift blazars and the dashed curve to high-redshift FSRQs. We find that low-redshift FSRQs have *higher*, on average, intrinsic modulation indices. The result is shown to be statistically significant in the right panel of figure 5.30, where we plot the probability density of the difference between m_0 in each subset. The most likely difference is found to be about 2.4 percentage points, and more than 3σ away from zero.

In figures 5.31 and 5.32, we plot the likelihood distributions and m_0 difference probability distributions for high- and low-redshift subpopulations of the CGRaBS and 1LAC samples, now computed using the 42-month data set. The histograms of the modulation indices in each sample are shown in figures 5.33 and 5.34. Although in both cases we continue to find that the low-redshift sources are characterized by greater average

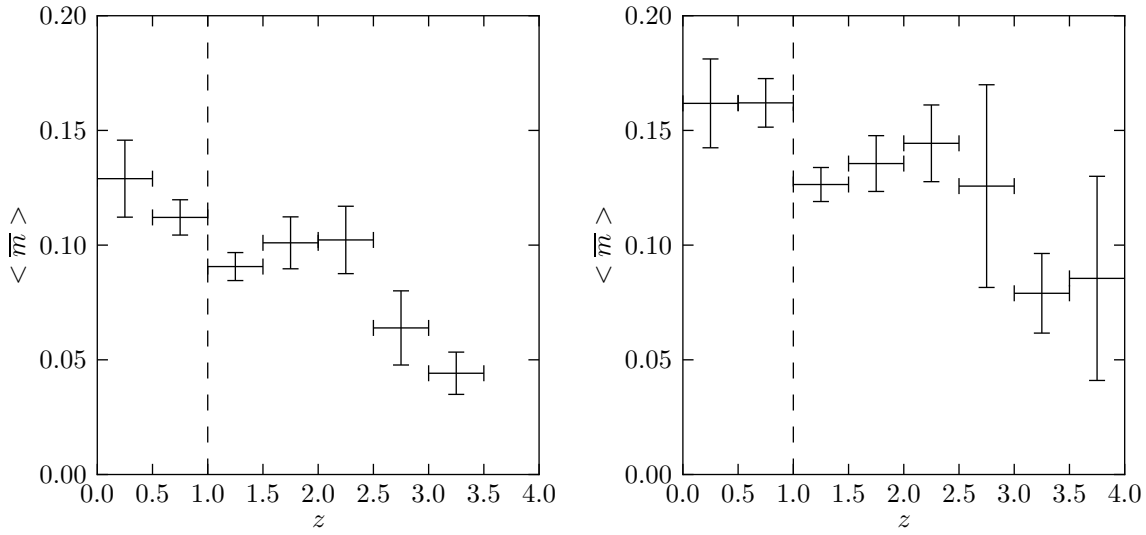

Figure 5.29. Mean $\overline{m}$ in redshift bins of 0.5 for bright ($S > 400$ mJy) FSRQs in our CGRaBS monitoring sample using the two-year (left) and 42-month (right) data. Horizontal error bars indicate the bin width, vertical error bars indicate the uncertainty in the mean computed from the scatter in the data in that bin. The dashed line indicates the $z = 1$ split between high and low redshift sources used for the population comparison.

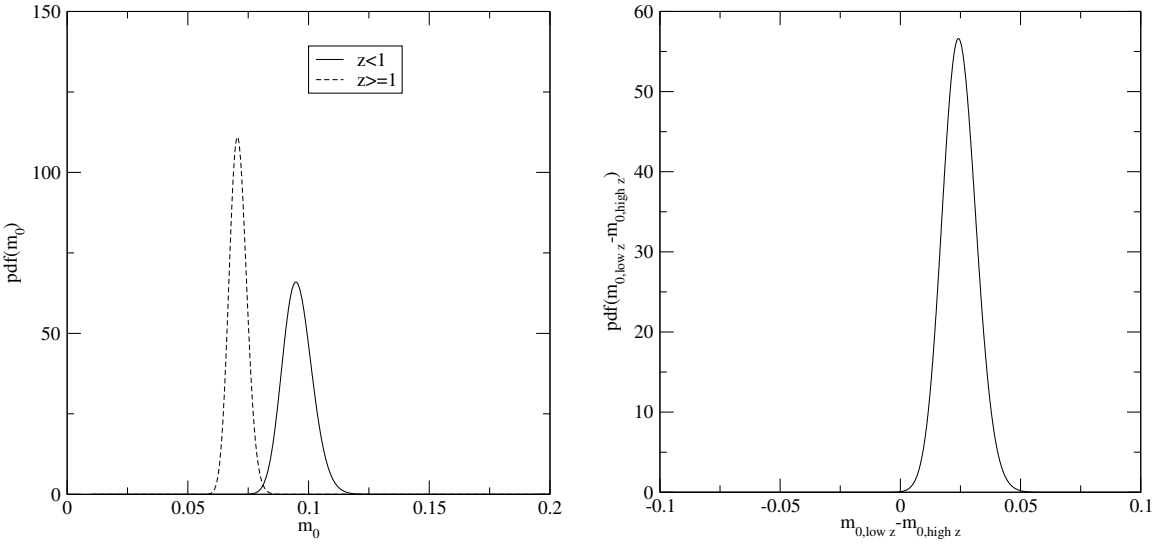

Figure 5.30. Comparison of high- and low-redshift CGRaBS FSRQs using the two-year data set. *Left:* Probability density of m_0 for FSRQs in our monitoring sample with $z < 1.0$ (solid line, maximum-likelihood value and 1σ error $m_0 = 0.095 \pm 0.006$) and $z \geq 1.0$ (dashed line, maximum-likelihood value and 1σ error $m_0 = 0.071 \pm 0.004$). The two distributions are *not* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution (0.024 ± 0.007) is more than 3σ away from zero.

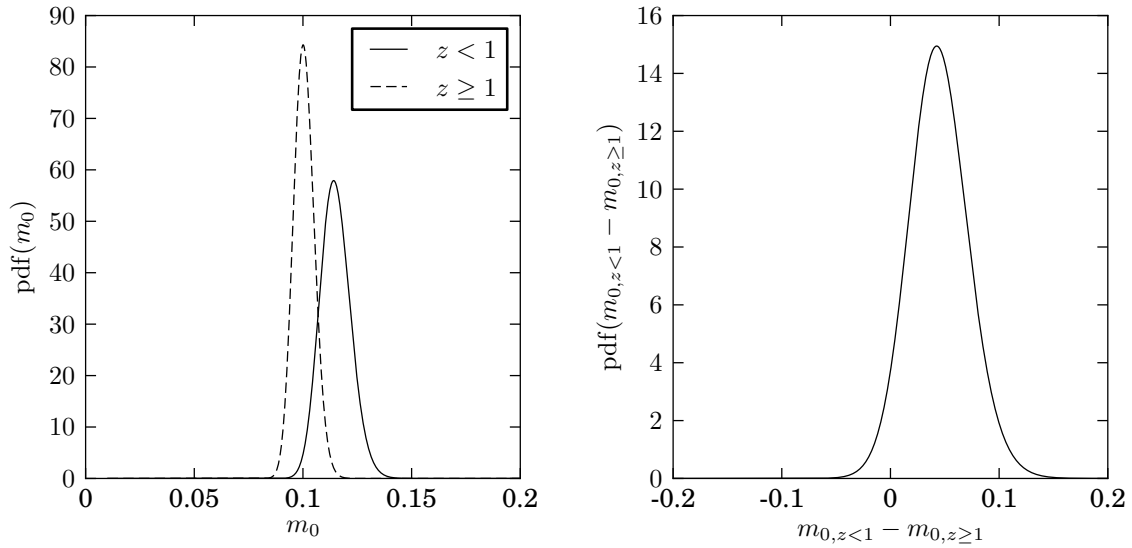

Figure 5.31. Comparison of high- and low-redshift CGRaBS FSRQ populations using 42-month data. *Left:* Probability density of m_0 for the CGRaBS FSRQs with known redshift in our monitoring sample with $z < 1$ (solid line, maximum likelihood value and 1σ error 0.114 ± 0.007) and $z \geq 1$ (dashed line, maximum likelihood value and 1σ error 0.100 ± 0.005). The two distributions *are* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution ($0.014^{+0.009}_{-0.008}$) is less than 2σ away from zero.

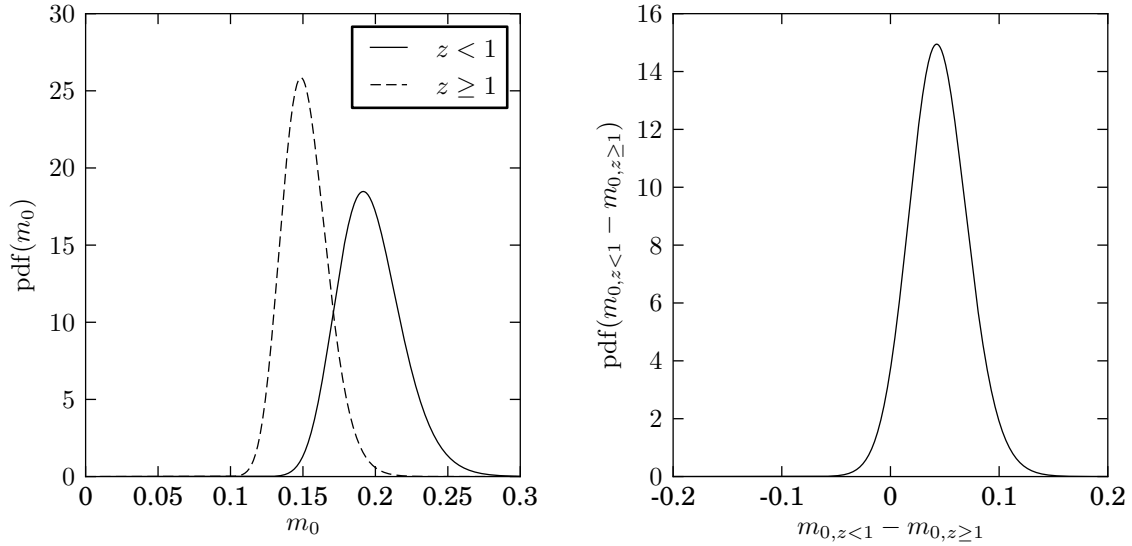

Figure 5.32. Comparison of high- and low-redshift 1LAC FSRQ populations using 42-month data. *Left:* Probability density of m_0 for the 1LAC FSRQs with known redshift in our monitoring sample with $z < 1$ (solid line, maximum likelihood value and 1σ error $0.192^{+0.023}_{-0.020}$) and $z \geq 1$ (dashed line, maximum likelihood value and 1σ error $0.149^{+0.017}_{-0.014}$). The two distributions *are* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution ($0.043^{+0.028}_{-0.026}$) is less than 2σ away from zero.

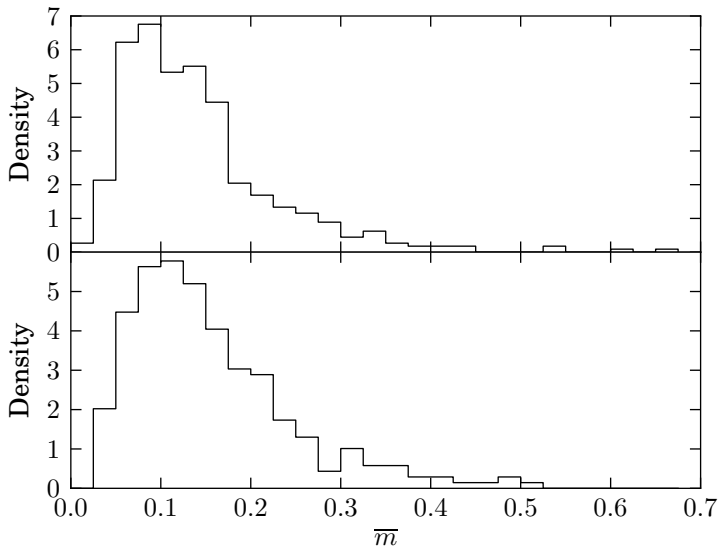

Figure 5.33. Histograms of 42-month intrinsic modulation index values for CGRaBS sources at $z \geq 1$ (top, 450 sources) and $z < 1$ (bottom, 277 sources). Each histogram is normalized to integrate to unity.

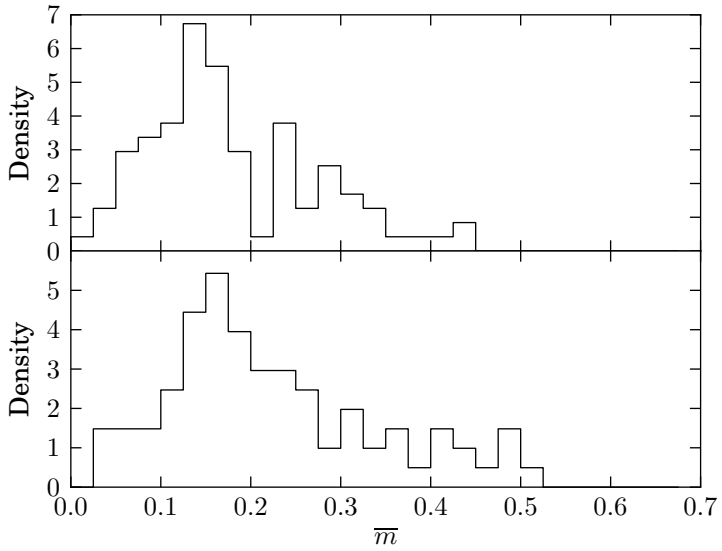

Figure 5.34. Histograms of 42-month intrinsic modulation index values for 1LAC sources at $z \geq 1$ (top, 95 sources) and $z < 1$ (bottom, 81 sources). Each histogram is normalized to integrate to unity.

intrinsic modulation index, with 42 months of data this difference is no longer even 2σ significant for either the CGRaBS or the 1LAC samples. In the right panel of figure 5.29, we plot the binned average modulation index versus redshift using the 42-month CGRaBS data set. Although the mean value in each bin still shows a hint of a trend, the scatter has also increased in the higher-z bins, diluting the appearance of a trend. That the additional data has reduced, not enhanced, the significance of this result suggests that it was merely a chance effect. Nonetheless, because evidence for cosmological evolution of blazar variability would be an important finding, we will examine this more closely.

Understanding the implications of the presence or absence of a trend of variability amplitude with redshift is complicated by competing observational and selection effects that can push the trend in either direction. To claim or rule out cosmological source evolution will require quantitative measurement or modeling of these effects. For example, at higher redshift, the 15 GHz observation frequency corresponds to a higher rest frame emission frequency. Because variability amplitude generally increases with increasing cm-wave radio frequency Stevens et al. (e.g., 1994), this would lead to an increase in variability amplitude with increasing z. The effects of Doppler beaming of the blazar jet emission also gives rise to selection effects which are somewhat more difficult to quantify (Lister & Marscher 1997). In this section, we will examine in detail the effect of one effect—cosmological time dilation—which leads to an underestimate of variability amplitude at higher z.

5.5.1 Cosmological Time Dilation

The redshifts assigned to the sources in our sample are determined spectroscopically—that is, by identifying the observer-frame wavelengths, $\lambda_{\rm obs}$, of photons that are emitted or absorbed in the rest frame of the source. By identifying patterns of emission or absorption lines that correspond to atomic or molecular line spectra, the rest-frame wavelengths, $\lambda_{\rm rest}$, of those lines can be determined. Then the redshift z of the source is obtained as

$$z = \frac{\lambda_{\rm obs}}{\lambda_{\rm rest}} - 1 = \frac{\nu_{\rm rest}}{\nu_{\rm obs}} - 1. \tag{5.16}$$

This redshift is normally dominated by the expansion of space, so this redshift is ascribed to the Doppler effect due to the cosmological motion of the source away from the observer.

Because this effect is due to the expansion of space, it is expected that the observed spectral redshift will be accompanied by a cosmological time dilation effect. This has been confirmed by, e.g., Blondin et al. (2008), who compared the evolution of Type Ia supernovae at various redshifts. Thus, a rest-frame time interval $\Delta t_{\rm rest}$ for a source at redshift z will correspond to an observed interval

$$\Delta t_{\rm obs} = (1+z)\,\Delta t_{\rm rest}. \tag{5.17}$$

Inverting equation (5.17), it is clear that, for equal observer time intervals, a source that is at a lower redshift has been observed for a longer rest frame interval. The observer is thus more likely to observe a full cycle[1] of variability behavior from a source at lower z since more rest-frame time has passed during which variability-causing events, whatever their physical nature, can occur. As a result, we are likely to sample more flares and observe more periods of variability in a source for which we have observed a long rest-frame time interval.

We need not rely on our intuition to conclude that time dilation will tend to reduce variability amplitude as redshift increases—we have data to prove it. In section 5.1.2.1, we demonstrated that in almost all cases, the intrinsic modulation index for a source increased or remained level between the two-year and the 42-month data sets. That is, by increasing the observed time period for a given source, a larger variability amplitude typically results.

5.5.2 Compensating for Time Dilation

Virtually all monitoring programs, including ours, observe their sources either over a time interval that is the same for each source—the total length of operation of the program—or that may be shorter for sources added or dropped during operation, but not selected with regard to the redshift of the source. To compensate for this effect, we simply discard some data from low-z sources in order to compare the intrinsic modulation index found in equal rest-frame time intervals.

For sources at near-zero redshift, the time dilation effect will not be significant so the rest-frame and observer-frame intervals will be nearly equal. Our sample includes redshifts as high as 5.47 (J0906+6930), however, for which the dilation factor will be $1 + z = 6.47$, which is clearly significant. Our 42 months of data correspond to only 6.5 months in the rest frame at this redshift. Because the level of short-timescale variability in a given blazar seems to vary with time, sometimes apparently switching from a stable state to a variable one, we want to keep our measurement intervals long enough to give each source a fair chance to enter a variable state. We therefore use only sources with a redshift $z \leq 3.0$, for which we have sampled a rest-frame interval of 10.5 months, or 315 d.

5.5.3 Selecting Time Intervals

A subset of the data for each source to use for the equal-Δt_{rest} comparison was selected as follows. First, from the redshift z, the time dilation factor $(1 + z)$ determines the observer-frame time period to sample, $\Delta t_{\mathrm{obs}} = (1 + z) \times 315$ d. Next, we must choose a segment of the data for the source with this length. To

[1]The term "cycle" is used loosely here: although it is clear from our light curve data that many blazars exhibit long periods of low variability followed by flares or periods of rapid variation, there is no particular reason to believe this is cyclical in a periodic sense. Searches for periodicity have not found convincing evidence, see, e.g., Hovatta et al. (2007).

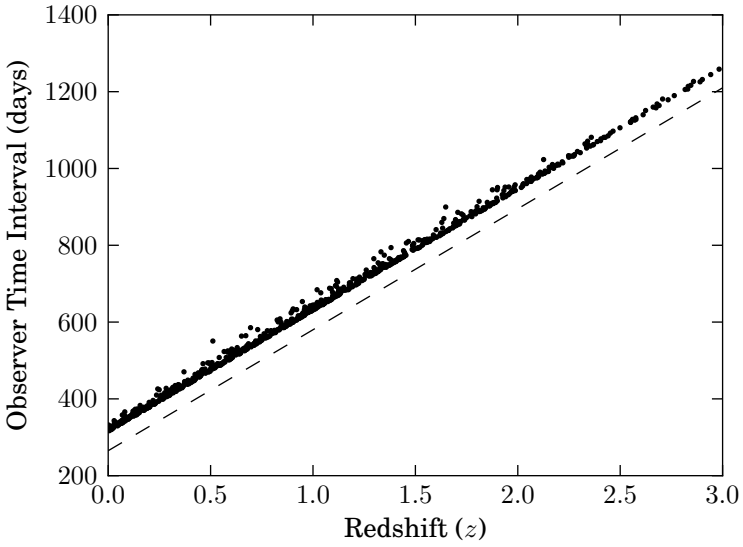

Figure 5.35. Observer time interval for each source plotted versus redshift. The dashed line shows the expected slope, $dt/dz = 315$ d, for reference.

do this, we choose a random start date within the data set for the source, at least $\Delta t_{\rm obs}$ days before the last measurement.

We choose a random start date for each source to avoid correlating the redshift of a source with the observing conditions. If, say, we simply chose an interval of the desired length that ended on the last measurement for each source, low-redshift sources would all include data observed in the winter and spring of 2011. If the conditions were slightly different during that time than the average over the 42 months of the program, this would lead to a bias. By choosing the interval at random, any such changing conditions are averaged as much as possible over the population of low-z sources.

After selecting a start date, we must ensure that interruptions in our program have not drastically reduced the number of measurements in the interval. First, we extend the end date of the selection to ensure that an outage $\Delta t_{\rm obs}$ days after the start date does not reduce the actual observed interval. Figure 5.35 shows the actual observer time interval selected for each source. We also require that the selected interval contains at least an average of one data point per 11 d (in the observer frame). If fewer data than this have been chosen, we select another random start date and repeat this process until an adequate number of data survive. Figure 5.36 shows the number of data points selected per source as a function of redshift.

From our 1LAC sample, 4 sources were dropped from this analysis because they were observed for less than 315 d in their rest frame. As shown in figure 5.37, these did not substantially alter the redshift distribution. A K-S test comparing the full data set with that of the surviving data does not reject the hypothesis that the two are drawn from the same distribution, with $1 - p = 2 \times 10^{-15}$ (i.e., $p \sim 1$).

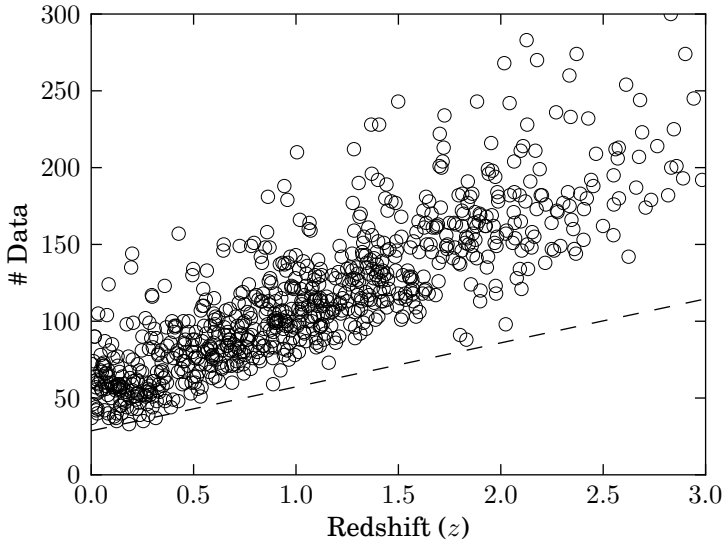

Figure 5.36. Number of data points in the equal-$\Delta t_{\rm rest}$ data sample for each source. The dashed line shows the minimum accepted number of data at each z.

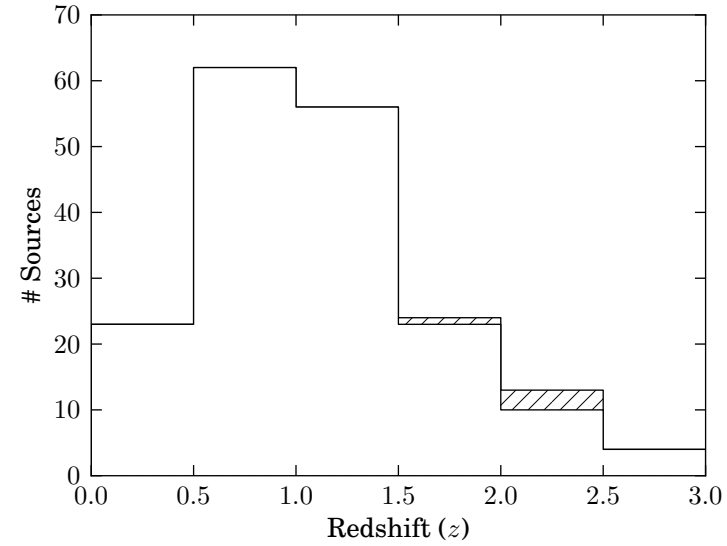

Figure 5.37. Histogram of redshifts for the 1LAC FSRQs with known $z < 3.0$. Hatched regions show the effect of dropping 4 sources that were sampled less than 315 d in their rest frames. A K-S test comparing these distributions does not reject the null hypothesis that the distributions are equal.

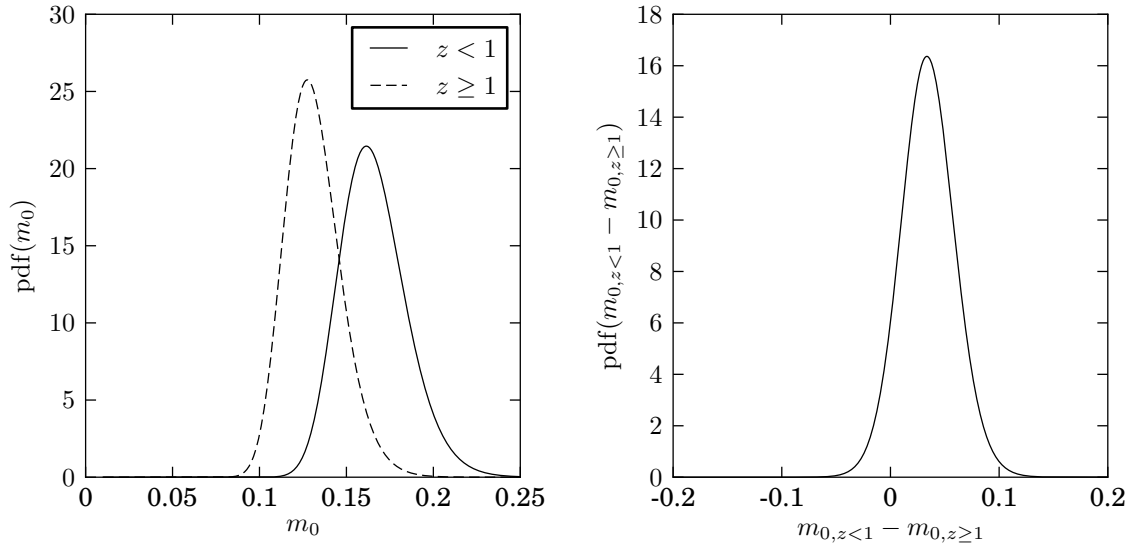

Figure 5.38. Comparison of high- and low-redshift 1LAC FSRQ populations using equal rest-frame time periods from the 42-month data. *Left:* Probability density of m_0 for the 1LAC FSRQs with known redshift in monitoring sample with $z < 1$ (solid line, maximum likelihood value and 1σ error $0.162^{+0.020}_{-0.017}$) and $z \geq 1$ (dashed line, maximum likelihood value and 1σ error $0.128^{+0.017}_{-0.014}$). The two distributions *are* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution ($0.034^{+0.024}_{-0.025}$) is less than 2σ away from zero.

5.5.4 Equal Rest-Frame Time Interval Results

Using the shortened data sets to provide equal rest-frame time intervals for all sources in the 1LAC population with $z < 3$, we remove the effect of time dilation from the observed modulation index. In figure 5.38, we plot the likelihood distribution for the m_0 parameters and for the difference in m_0 between the 1LAC FSRQs at $z < 1$ from those at $z \geq 1$. The low-redshift source population is still found to have a larger average modulation index, and this is still not a statistically significant result.

As expected, the average intrinsic modulation index in both the high- and low-redshift bins were *smaller* when the shorter equal rest-frame time intervals were used, although in both cases the difference was not a statistically significant change beyond the 1σ level. For the high-redshift population, the difference was

$$\Delta m_{z\geq1} = 0.128 - 0.149 = -0.021 \pm 0.021, \tag{5.18}$$

where we have computed the uncertainty in this difference approximately, using the average of the upper and lower error bars for the quantities being subtracted. For the low-redshift population, the difference was

$$\Delta m_{z<1} = 0.162 - 0.192 = -0.030 \pm 0.029, \tag{5.19}$$

again computing uncertainties using the average of the upper and lower error bars for each value.

We conclude that the effect of cosmological time dilation on our earlier result was not significant, but our results are consistent with there being a small effect on our data. The change was smaller for high-z sources, as expected because nearly the entire time series was used, whereas time series for low-z sources were trimmed more substantially.

5.6 CGRaBS versus 1LAC

It is clear that the population of sources predicted to be gamma-ray loud in the CGRaBS paper differed substantially from the sources actually detected by the *Fermi*-LAT instrument during the first year of operation. In section 4.2, we demonstrated BL Lac objects made up a much larger fraction of the 1LAC sample than of the CGRaBS sample. This can be attributed to instrumental differences between EGRET and the LAT. This became evident from the reported EGRET "GeV excess" (Hunter et al. 1997). This excess diffuse gamma-ray emission was not confirmed by *Fermi* (Abdo et al. 2009b) and is now believed to be a systematic artifact due to uncertain instrumental response above about 1 GeV (Stecker et al. 2008). As a result, high-energy results from the EGRET instrument are suspect, limiting its ability to detect spectrally hard sources. The LAT, which is especially efficient at detecting sources via 1 GeV photons, is thus more efficient than EGRET at detecting BL Lacs because they typically exhibit hard gamma-ray spectra (Abdo et al. 2010a, 2010c).

We found in section 5.3 that gamma-ray–loud CGRaBS sources that were part of the 1LAC sample are significantly—more than 6σ—more variable than are gamma-ray–quiet sources. Then in section 5.4 we showed that in the CGRaBS sample, BL Lac objects are significantly more variable than FSRQs, whereas in 1LAC, the situation is reversed (and no longer statistically significant). What can we conclude from these observations?

We know that the CGRaBS and 1LAC samples are not drawn from the same parent distribution. Is this true of subpopulations within those two samples? First, let us examine whether the BL Lac objects are consistent with coming from a common parent sample. In the left panel of figure 5.39 we compare the likelihood distributions for m_0 for the BL Lac objects in the CGRaBS and 1LAC samples. These samples are consistent with a common value for m_0 to within 1σ. Since the BL Lac populations appear to exhibit the same average variability amplitude, the FSRQ variability amplitudes must differ to explain the overall disagreement between the CGRaBS and 1LAC samples.

This is just what we find. In the right panel of figure 5.40, we compare the FSRQ variability amplitudes between the CGRaBS and 1LAC samples. The 1LAC FSRQ population is much more variable with $m_0 = 0.168^{+0.014}_{-0.012}$ whereas for the CGRaBS FSRQs $m_0 = 0.105 \pm 0.004$. This most likely difference of $0.063^{+0.014}_{-0.014}$ is significant at nearly the 6σ level. The CGRaBS FSRQs are, on average, much less variable at 15 GHz than the gamma-ray–loud 1LAC FSRQs.

This enhanced radio variability amplitude among gamma-ray–loud FSRQs together with the equivalence of the variability in the BL Lac population would be expected if the gamma-ray–loud FSRQs are more

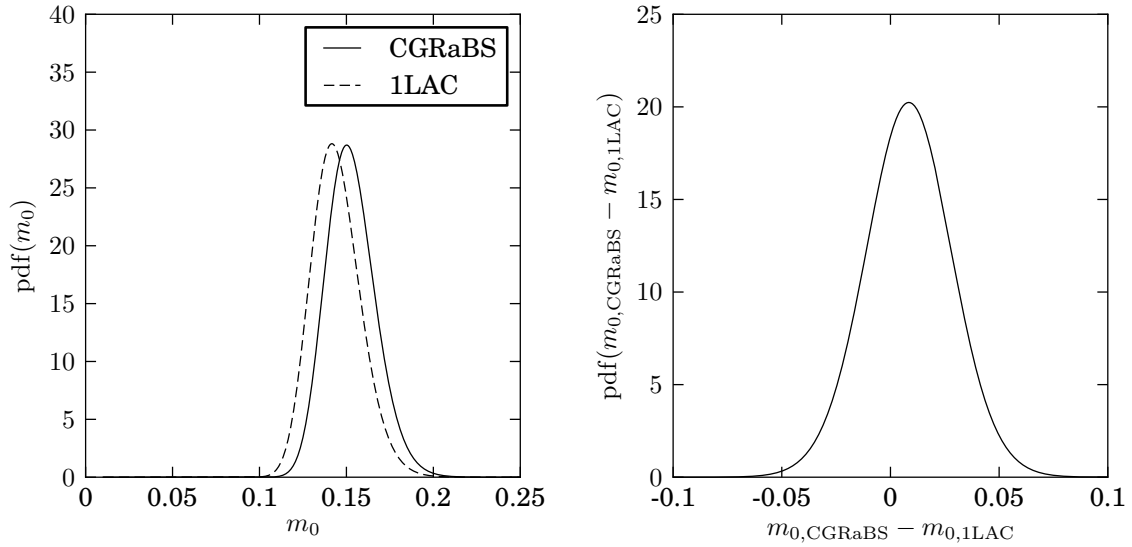

Figure 5.39. Comparison of CGRaBS and 1LAC BL Lac populations using the 42-month data. *Left:* Probability density of m_0 for the CGRaBS BL Lacs (solid line, maximum likelihood value and 1σ error $0.150^{+0.015}_{-0.013}$) and 1LAC BL Lacs (dashed line, maximum likelihood value and 1σ error $0.142^{+0.015}_{-0.013}$). The two distributions *are* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution (0.008 ± 0.020) is less than 1σ away from zero.

strongly beamed than their gamma-ray–quiet counterparts. This could be due to different inverse Compton seed photon sources between the two classes. The Doppler beaming pattern for the Compton component of EC models is narrower about the jet axis than that for the synchrotron component, while both components are affected by the same beaming pattern in SSC sources (Dermer 1995). The SEDs of FSRQ sources frequently require EC models, while BL Lac objects with high-frequency SED peaks (HBLs) are usually consistent with SSC models (e.g., Böttcher 2007; Abdo et al. 2010c). While low spectrally peaked BL Lacs (LBLs) more often resemble FSRQs in this regard, *Fermi* was much more efficient at detecting the HBLs, which may be reflected in the statistics. Thus, it seems that for FSRQ sources, radio spectral properties alone are insufficient to predict gamma-ray emission. Additional data sensitive to the beaming angle, such as variability statistics, are also needed to discriminate between the gamma-ray–loud and gamma-ray–quiet FSRQs.

5.6.1 Flux Density Comparisons

We can also compare the brightness distributions of the CGRaBS and 1LAC samples. Figure 5.41, we compare the measured average flux density (S_0) distributions for subsets of CGRaBS sources in 1LAC and not in 1LAC (left panel), and for subsets of 1LAC sources in CGRaBS and not in CGRaBS (right panel). The overlap of the samples clearly preferentially selects the brighter sources in each. That is, the brightest CGRaBS predictions are more likely to be in 1LAC and the brightest 1LAC detections are more likely to be in CGRaBS.

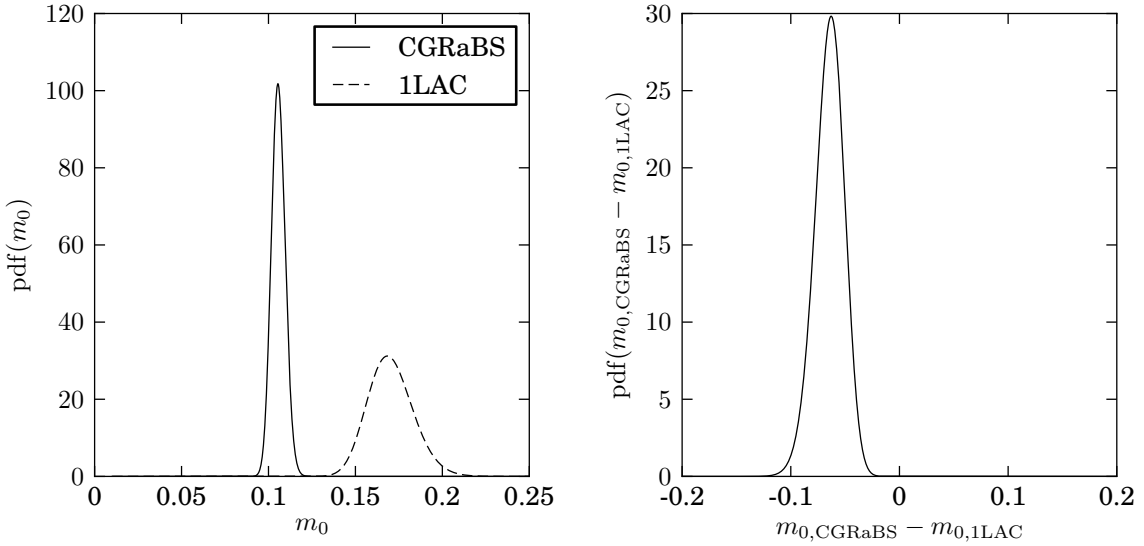

Figure 5.40. Comparison of CGRaBS and 1LAC FSRQ populations using the 42-month data. *Left:* Probability density of m_0 for the CGRaBS FSRQs (solid line, maximum likelihood value and 1σ error 0.105 ± 0.004) and 1LAC FSRQs (dashed line, maximum likelihood value and 1σ error $0.168^{+0.014}_{-0.012}$). The two distributions are *not* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution ($0.063^{+0.014}_{-0.013}$) is almost 6σ away from zero.

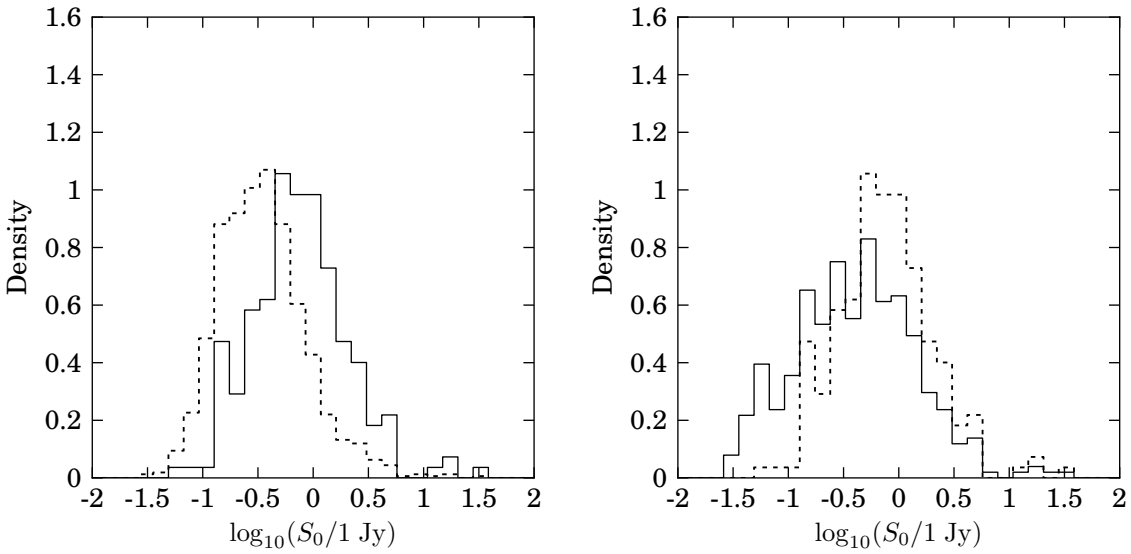

Figure 5.41. Histograms of S_0, normalized to integrate to unity. *Left:* Distribution of CGRaBS sources in 1LAC (solid) and not in 1LAC (dashed). A K-S test rejects the null hypothesis that these samples are drawn from the same distribution with $p < 5 \times 10^{-21}$. *Right:* Distribution of 1LAC sources in CGRaBS (dashed) and not in CGRaBS (solid). A K-S test rejects the null hypothesis with $p < 10^{-7}$. The dashed line in the right panel is the same as the solid line in the left panel.

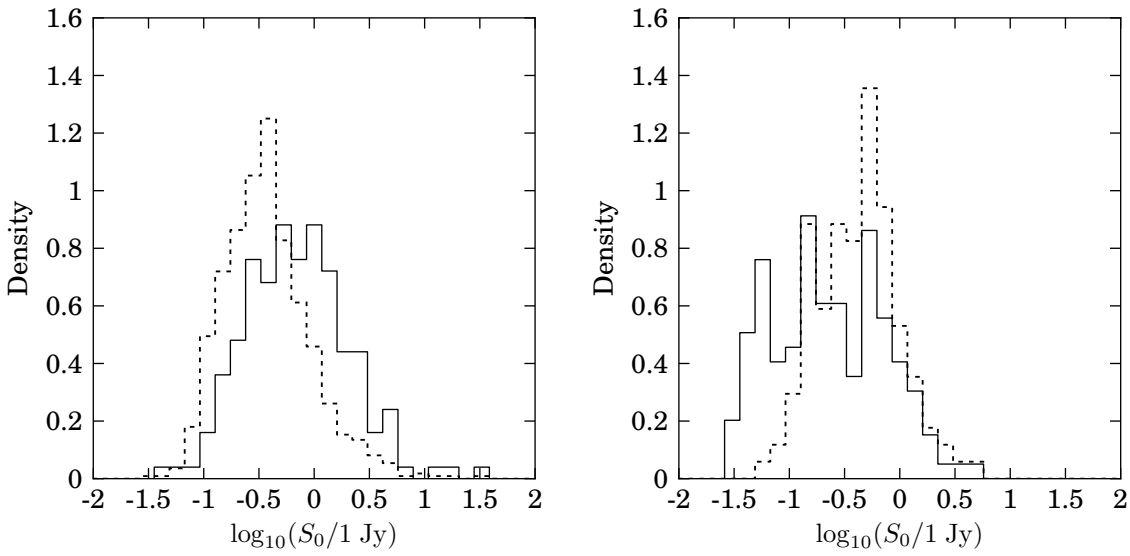

Figure 5.42. B
L Lac populations.] Histograms of S_0 for FSRQs (left) and BL Lacs (right) in the 1LAC (solid) and CGRaBS (dashed) samples. Each histogram is normalized to integrate to unity. A K-S test gives $p < 7 \times 10^{-12}$ (FSRQ) and $p < 4 \times 10^{-5}$ (BL Lac) for these data to come from the same distribution.

In figure 5.42, we again plot histograms of the distributions of the average flux density, S_0, now separating the FSRQ and BL Lac subpopulations of the two samples. The mean S_0 values for the FSRQs in the 1LAC and CGRaBS samples are (1.35 ± 0.20) Jy and (0.66 ± 0.05) Jy, respectively, and a K-S test gives $p < 7 \times 10^{-12}$ that the two data sets are drawn from the same distribution. Note that, as can be seen in figure 5.44, the 1LAC FSRQs in CGRaBS (mean $S_0 = (1.70 \pm 0.29)$ Jy) are brighter, on average, than the overall 1LAC FSRQ population. Thus, as we saw for the samples overall, brighter CGRaBS FSRQs were more likely to be detected by *Fermi* and brighter *Fermi*-detected FSRQs were more likely to have been predicted by the CGRaBS figure of merit (FoM) as gamma-ray emitters. Among BL Lacs, the mean S_0 for the 1LAC and CGRaBS samples are (0.46 ± 0.05) Jy and (0.62 ± 0.06) Jy, respectively, and a K-S test gives $p < 4 \times 10^{-5}$ that the two data sets come from the same distribution. Additionally, a number of sources were excluded from this histogram because the radio flux density measurements gave a $< 2\sigma$ detection. Most of these sources were 1LAC BL Lacs, so inclusion of these would further separate the 1LAC and CGRaBS BL Lac mean flux densities.

Among FSRQs, the result that gamma-ray–bright sources are on average brighter in the radio is consistent with the general trend found in radio-gamma connection studies, as discussed in section 1.3. If the increased variability we detect in the gamma-ray–loud FSRQs is a result of preferentially enhanced beaming in these sources, the higher Doppler factor in these sources could also give rise to this increased radio flux density. Some of this difference is probably attributable to the lower average redshift of 1LAC FSRQs, rather than intrinsic differences between sources. The reverse trend among BL Lacs—LAT-detected BL Lacs are somewhat dimmer on average than the CGRaBS BL Lacs—is most likely a result of the CGRaBS being a poor

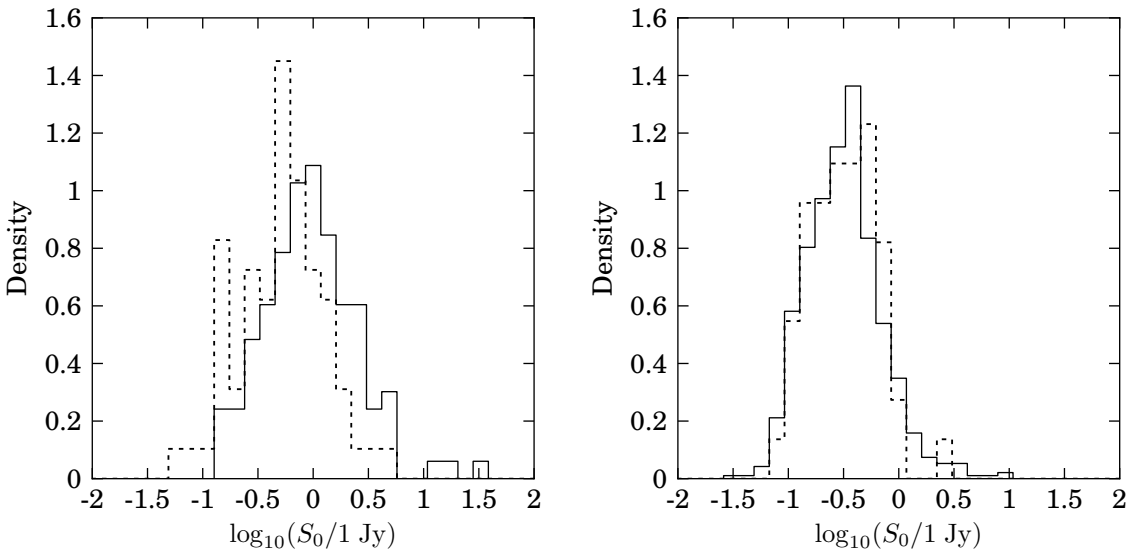

Figure 5.43. Histograms of S_0 for CGRaBS sources in 1LAC (left) and not in 1LAC (right), separately showing the distributions of FSRQ (solid) and BL Lac (dashed) samples. Each histogram is normalized to integrate to unity. A K-S test gives $p = 0.001$ (in 1LAC) and $p = 0.991$ (not in 1LAC) for these data to come from the same distribution.

prediction of the BL Lac sources likely to be detected by the LAT, reflecting the influence of the EGRET sensitivity bias on the CGRaBS selection. Thus, the LAT may simply be detecting a sample of BL Lac sources dimmer in gamma rays, and thus likely to be dimmer in radio as well.

This explanation is consistent with the average flux density distributions shown in figure 5.43, which show that among CGRaBS sources, the BL Lacs and FSRQs detected in gamma rays are quite different ($p = 0.001$ to come from the same distribution), while the gamma-ray–quiet sources in both classes are indistinguishable ($p = 0.991$ to come from the same distribution). Since the selection criteria for the CGRaBS predominantly chose FSRQs, it seems that many of the BL Lac sources in the were selected because their properties resembled those of FSRQs, not necessarily those of gamma-ray–loud BL Lacs.

Our data do not explain the population of 1LAC FSRQs that were not included in CGRaBS. It is possible this results from the use of archival radio data, which would reflect the emission state of the source at some past time rather than during the *Fermi* observation era. We do not find evidence that this is caused by an overall radio brightening: in figure 5.44 we see 1LAC detections that were well above the CGRaBS minimum radio flux density. There is no evidence that 1LAC sources are preferentially brighter in the concurrent 15 GHz data than non-1LAC sources. However, as we are comparing two different radio frequencies, it is possible that spectral index variation is masking an effect. This possibility is enhanced because the CGRaBS FoM depended strongly on the radio spectral index of the source.

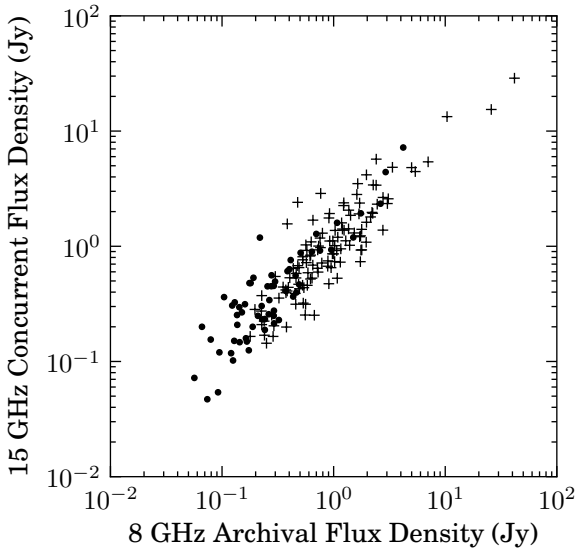

Figure 5.44. OVRO 40 m 15 GHz average flux densities (S_0) versus the archival 8 GHz flux densities tabulated in Ackermann et al. (2011) for 1LAC FSRQs. FSRQs in CGRaBS are plotted as crosses, those not in CGRaBS are plotted as points.

5.6.2 High Redshift FSRQ Populations

In addition to the differences in source classification, the redshift distributions of the 1LAC sources and CGRaBS differed. Within the BL Lac population, both 1LAC and CGRaBS were consistent with the same redshift distribution. Among FSRQs, however, the distributions were clearly not consistent. Examining the histograms in figures 4.5 and 4.6, it appears that the CGRaBS distribution is quite similar below $z \sim 1.5$ and the CGRaBS are clearly more numerous at high redshifts $z \gtrsim 1.5$.

To examine this, we compare the properties of the CGRaBS FSRQ populations in 1LAC and not in 1LAC at redshifts $z > 1.5$. Histograms of the intrinsic modulation indices are shown in figure 5.45. Although there are only 26 samples in the overlap between CGRaBS and 1LAC, the distribution extends to very high intrinsic modulation indices. The likelihood distributions for the population parameters are shown in figure 5.46. Despite the small sample size, the high-redshift CGRaBS FSRQs in 1LAC are significantly more variable than those not in 1LAC, with more than 3.5σ significance. The maximum-likelihood difference ($0.096^{+0.045}_{-0.034}$) is larger than was found between gamma-ray–loud and gamma-ray–quiet CGRaBS including all redshifts ($0.066^{+0.013}_{-0.012}$), but is consistent within 1σ of being the same.

Thus, it would seem that the CGRaBS selection criteria simply were not well tuned to picking out the systematically higher-variability sources that are bright gamma-ray emitters. By comparing *Fermi* 1LAC detections against the Australia Telescope 20 GHz survey catalog (AT20G), Mahony et al. (2010) conclude that the AGN in 1LAC are predominantly characterized by flat radio spectral indices and that there is no missing steep-spectrum population among the gamma-ray–loud population. This confirms that the most important CGRaBS selection criterion—a flat radio spectral index—is an effective method for selecting gamma-ray–

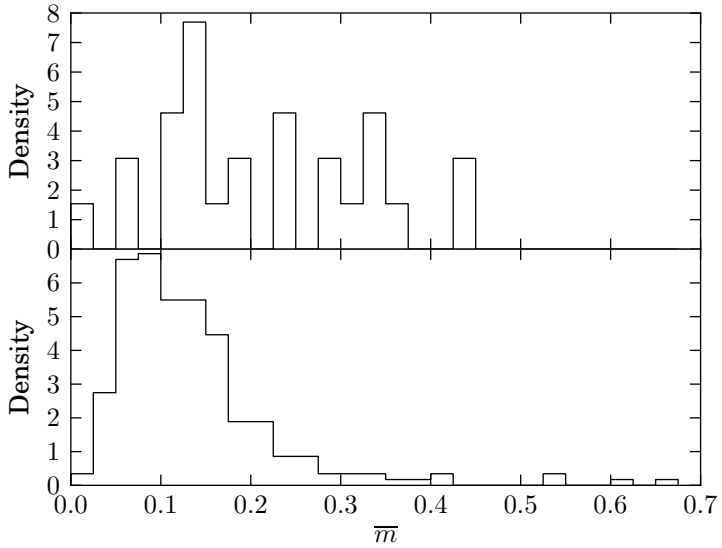

Figure 5.45. Histograms of 42-month intrinsic modulation index values for CGRaBS FSRQs at $z > 1.5$ in 1LAC (top, 26 sources) and those not in 1LAC (bottom, 233 sources). Each histogram is normalized to integrate to unity.

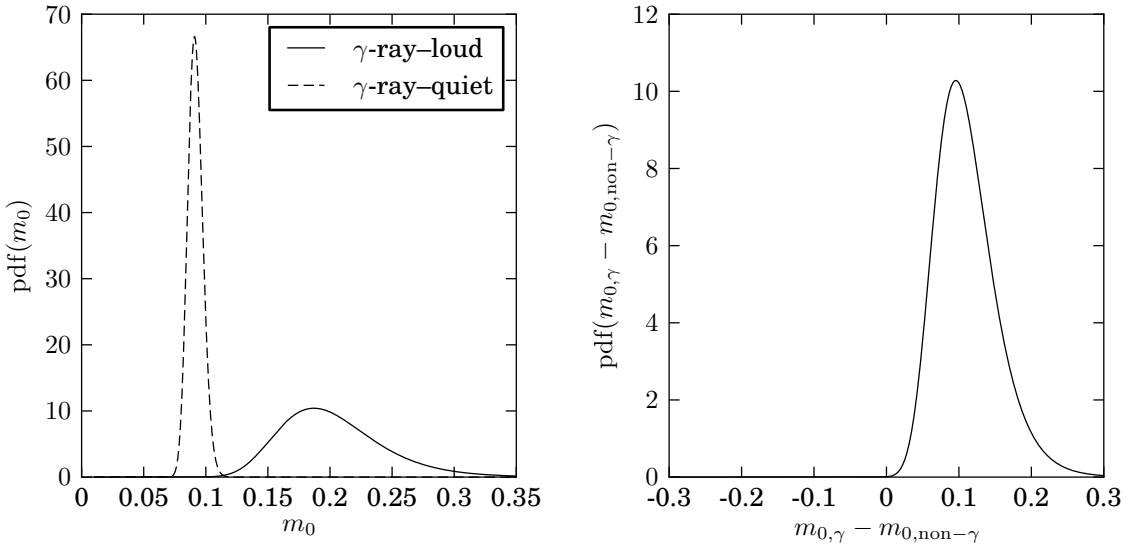

Figure 5.46. Comparison between high-redshift ($z \geq 1.5$) CGRaBS FSRQ populations in and not in the 1LAC sample, using the 42-month data. *Left:* Probability density of m_0 for the $z \geq 1.5$ CGRaBS FSRQs in the 1LAC sample (solid line, maximum likelihood value and 1σ error $0.187^{+0.044}_{-0.034}$) and those not in 1LAC (dashed line, maximum likelihood value and 1σ error 0.091 ± 0.006). The two distributions are *not* consistent with a single value. *Right:* Probability density of the *difference* between the mean modulation index m_0 for the two sets. The peak of the distribution ($0.096^{+0.045}_{-0.034}$) is more than 3.5σ away from zero.

loud AGN. That there is no evidence for a missing population is consistent with our conclusion that the CGRaBS sample is drawn from a superset of the 1LAC parent population. Although it is unclear precisely what mechanism connects between 15 GHz radio variability and gamma-ray emission, it seems evident that the presence of variability in radio is a good predictor of gamma-ray emission from a blazar.