Please provide feedback

APPENDIX A: INTRA-NIGHT VARIABILITY AT VHE

This section reports the single-night LCs at VHE γ-rays that show intra-night variability (INV), considered to occur when the fit with a constant value to the available intra-night flux measurements (time bins of 20 minute for FACT and 15 min for MAGIC) yield a p_{value} below 0.003 (i.e. more than 3σ significance). In case of FACT, the 20-minute binned light curves of all nights with a minimum observation time of 1 hour (196 nights) were checked for INV. From all the observations performed, INV was observed on only two nights, 2015 January 27 (MJD 57049) and 2015 March 12 (MJD 57093). In the first night, there were observations with both MAGIC (above 0.2 TeV) and FACT (E_{th} ~0.7 TeV). The INV is statistically significant only in the LC from MAGIC. In the case of FACT, the flux variations are not significant (less than 2σ) because of the larger flux uncertainties and the different temporal coverage. It seems that the flux of Mrk 421 dropped by 50% sometime between MJD 57049.20 and MJD 57049.25. In the second night, there are only FACT observations. Mrk 421 shows a decrease in the VHE flux by about a factor of 3 in the 3.5 hours that the observation spans.
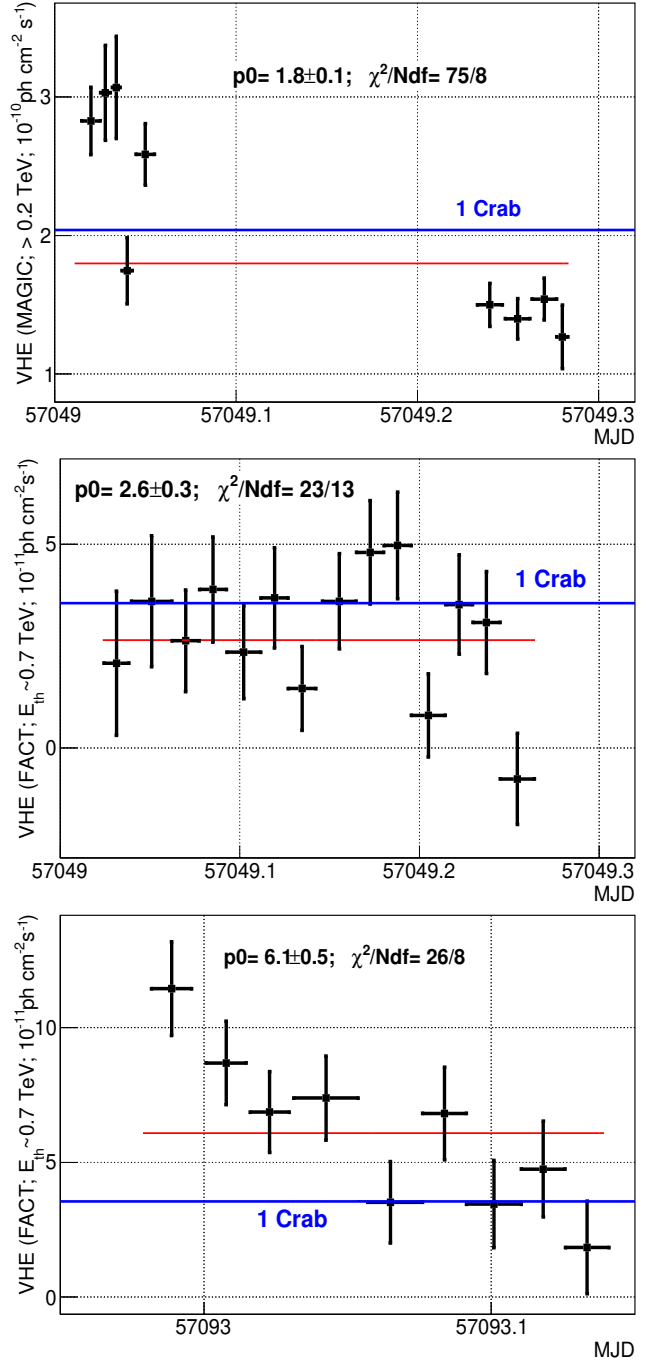
Figure A1. Single night VHE γ-ray LCs that show statistically significant intra-night variability. The first two panels show the MAGIC (above 0.2 TeV) and the FACT (E_{th} ~0.7 TeV) LCs for 2015 January 27 (MJD 57049). The lower panel shows the FACT (E_{th} ~0.7 TeV) LC for 2015 March 12 (MJD 57093). The blue horizontal lines depict the Crab Nebula flux in the respective energy band, and the red horizontal line represents a constant fit to the VHE γ-ray flux, with the resulting fit parameters and goodness of the fit reported in the panels.

APPENDIX B: MULTI-YEAR LIGHT CURVES

The studies reported in this paper are derived mostly with the extensive MWL data set collected during the campaigns in the years 2015 and 2016, when Mrk 421 showed very low flux at X-ray and VHE γ-rays. This 2-year data set is described in Section 2. However, for the correlation studies reported in Sections 5.2, 5.3, 5.4, and 5.5, and the characterization of the flux distributions reported in Section 6, the 2015–2016 data set is complemented with data from the years 2007–2014. This appendix provides a description of this additional (complementary) 2007–2014 data set.

The 2007–2016 data set, used for the above-mentioned correlation and flux-profile studies, is depicted in Fig. B1. The MAGIC VHE γ-ray and the *Swift*-XRT X-ray LCs are retrieved from various published works (Aleksić et al. 2012, 2015c; Ahnen et al. 2016; Baloković et al. 2016). The FACT fluxes from 2012 December to 2016 June were produced with the analysis described in Section 2.6. The *Fermi*-LAT fluxes in the band $0.3 - 300\,\text{GeV}$ were analyzed as described in Section 2.4. The *Swift*-BAT fluxes were retrieved from the BAT website[13], and treated as explained in Section 2.3.3. The optical data in the R-band were retrieved from Carnerero et al. (2017). The 37 GHz radio fluxes from Metsähovi were provided by the instrument team, and the 15 GHz radio fluxes from OVRO were retrieved from the website of the instrument team[14]. As done in Section 4.1, we only consider fluxes with the relative errors (flux-error/flux) smaller than 0.5 (i.e. SNR>2). In this way, we ensure the usage of reliable flux measurements, and minimize unwanted effects related to unaccounted (systematic) errors.
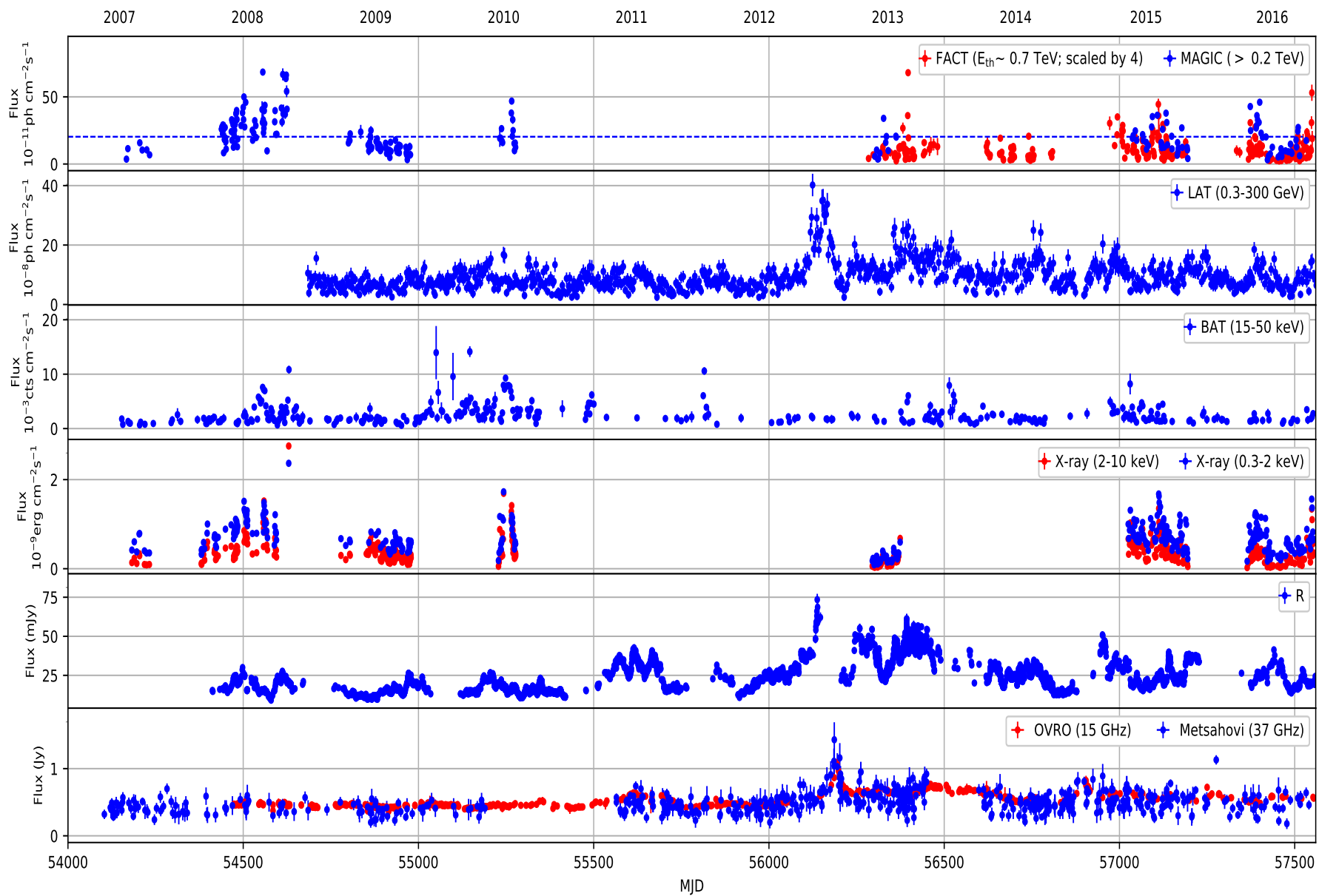
There are seven MAGIC VHE fluxes from the year 2007, from the time interval MJD 54166–54438, and five VHE fluxes from the year 2009, from the time interval MJD 54800–54835, that relate to energies above 0.4 TeV (published in Ahnen et al. 2016), and all the MAGIC VHE fluxes from the 4.5-months long MWL campaign in year 2009, from the time interval MJD 54851–54977, relate to energies above 0.3 TeV. (published in Aleksić et al. 2015c). The reason for the higher minimum energy in these two publications with respect to other publications that relate to observations performed after year 2010 (where the light curves are produced with energies above 0.2 TeV) is the operation of MAGIC in mono mode (with a single-telescope). The MAGIC observations of Mrk 421 in stereo mode, which started in the MWL campaign from year 2010, provide additional sensitivity and a lower analysis energy threshold, which allows one to reliably produce light curves with a minimum energy of 0.2 TeV. During the year 2008, Mrk 421 showed high VHE flux and, despite MAGIC operating with a single-telescope, the large VHE γ-ray fluxes and the longer exposures, permitted the reliable reconstruction of the VHE fluxes above 0.2 TeV, as reported in Aleksić et al. (2012). In order to properly compare the published VHE fluxes from the years 2007 and 2009 with those from 2008 and from 2010 onwards, we scaled VHE fluxes above 0.4 TeV and 0.3 TeV (and their related errors) by a factor of 2.83 and 1.84, respectively. These scaling factors were calculated by considering that the VHE spectral shape of Mrk 421 around the energy of 0.3 TeV can be well described with a power-law function with index 2.5, when Mrk 421 is in its typical (non-flaring) state (Abdo et al. 2011). They can then be used to convert the VHE fluxes above 0.4 TeV and 0.3 TeV to that above 0.2 TeV. The spec-

tral shape of the VHE emission of Mrk 421 does vary over time, and it is known to be related to the flux (e.g. *harder-when-brighter* behaviour). However, owing to the relatively small energy range over which one needs to extrapolate, and the relatively low VHE flux and low variability from years 2007 and 2009, including these spectral variations would vary the reported VHE fluxes by less than ±10% in most cases. These additional flux variations are typically smaller than the statistical uncertainties of the flux measurements during these low-flux periods, and hence they do not affect the reported study in any significant manner.

Figure B2 shows the 3-day binned light curves during 2012 December to 2016 June used in the correlation studies presented in Section 5.2. The VHE fluxes from FACT were derived with the analysis described in Section 2.6, but this time in 3-day time intervals. The data from *Fermi*-LAT were analyzed as described in Section 2.4.

---

[13] http://heasarc.nasa.gov/docs/swift/results/transients/
[14] http://www.astro.caltech.edu/ovroblazars/index.php?page=home

Figure B1. Multi-year data set (with flux measurements with SNR>2) used in the study reported in Sections 5.3, 5.4, 5.5, and 6. In the top panel, the horizontal blue line represents 1 Crab flux in above 0.2 TeV.
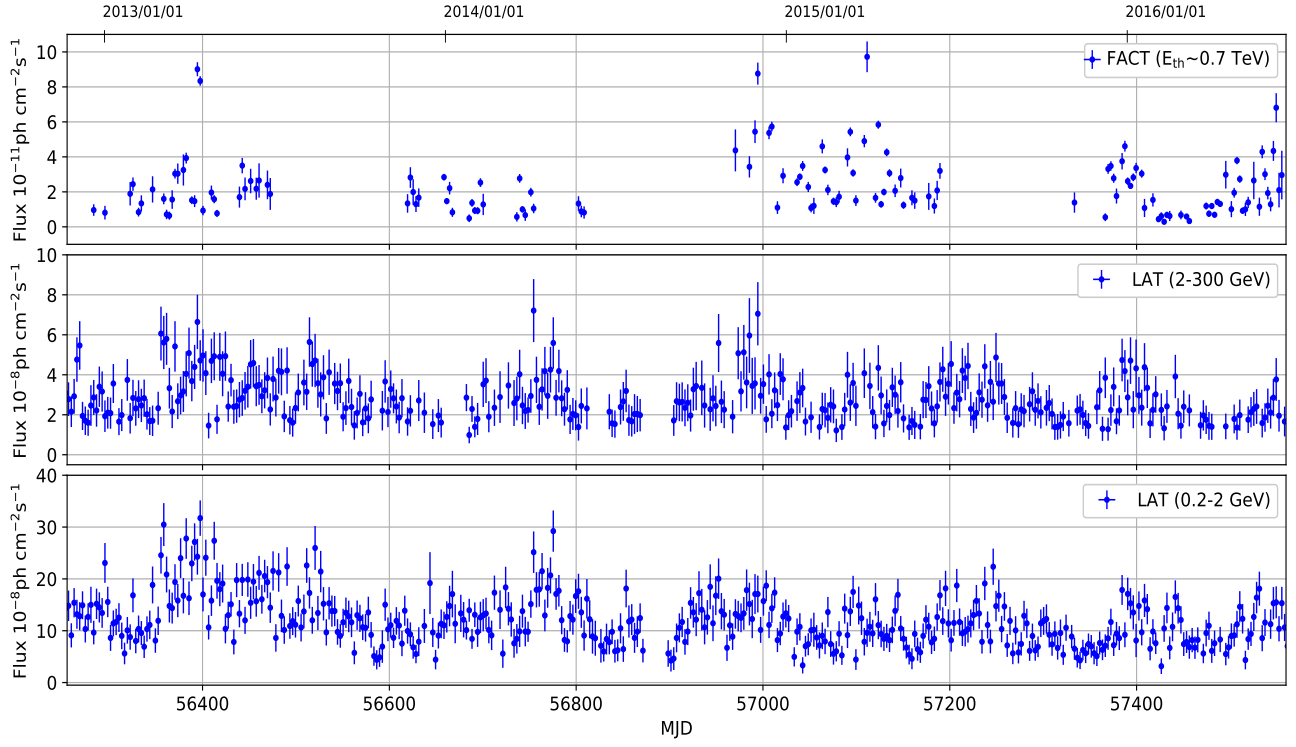
Figure B2. The 3-day binned LCs (with flux measurements with SNR>2) measured with FACT and *Fermi*-LAT in the energy bands E_{th} ~0.7 TeV (top panel), $2 - 300$ GeV (middle panel), and $0.2 - 2$ GeV (bottom panel), that were used in the study reported in Section 5.2.

APPENDIX C: MULTI-BAND FLUX-FLUX RELATIONS

This section reports the multi-band flux-flux plots related to the correlations discussed in Section 5.

Panels (a)–(d) in Fig. C1 show the integral VHE γ-ray flux from the two energy bands measured with MAGIC (reported in Fig 2, namely $0.2 - 1$ TeV and above 1 TeV), plotted against the X-ray flux in the two energy bands from *Swift*-XRT (reported in Fig 2). The panels (e)-(f) of Fig. C1 show the VHE vs. X-ray flux relations when using the VHE fluxes with $E_{th} \sim 0.7$ TeV measured with FACT. Only simultaneous observations are used in these figures. Besides the display of all the flux measurements (roughly equivalent to *unbinned* data), the panels also show the average and the standard deviation computed with data subsets of 10 observations, binned according to their flux (*binned* data). The binned data allow us to better visualize the main trend, as well as the dispersion in the single-day flux measurements. Both the unbinned and binned data are fitted with a linear function to quantify the slope in the VHE vs. X-ray flux relation. These slopes are reported in Table 2. Despite the large dispersion in the VHE vs. X-ray flux values, there is a roughly linear trend for all the bands, with the slope of the trend increasing for increasing VHE energy band, or for decreasing X-ray energy band.

The panels in Fig. C2 show the VHE γ-ray flux from FACT ($E_{th} \sim 0.7$ TeV) during the period from 2012 December to 2016 June (see Fig. B2), plotted against the HE flux from *Fermi*-LAT in two energy bands, $0.2 - 2$ GeV and $2 - 300$ GeV (see Fig. B2). As with the panels in Fig. C1, besides showing all of the 3-day flux measurements (*unbinned* data), the panels also show the average and the standard deviation computed with data subsets of 10 observations, binned according to their flux (*binned* data). Both the unbinned and binned data are fitted with a linear function to quantify the slope in the VHE vs. HE flux relation. These slopes are reported in Table 3. In contrast to what happens in the panels of Fig. C1, there is a large difference between the slopes in the linear functions fitted to the unbinned and binned data. The difference is ascribed to VHE vs. HE flux pairs which are well outside the main trend (outliers), which have a large impact on the fit to the unbinned data, but not to the binned data. The difference is also partly due to the weak (if not absent) correlation between these energy bands (see Section 5.2 for further details).

Figure C3 shows the HE vs. optical flux correlation plots for the HE γ-rays vs. optical for τ=0, the HE γ-rays vs. radio for a time shift of 45 days, and the optical vs. radio for a time shift of 45 days. The time shift of 45 days is the time for which the correlation between these two bands is the highest (see Section 5.4 and 5.5). The panels (b) and (c) of Fig. C3 show that, for a time shift of 45 days, the relation between the GeV and the radio fluxes can be approximated by a linear function. As with Fig. C1, the panels also show the average and the standard deviation computed with data subsets of 10 observations, binned according to their flux (*binned* data). In the case of LAT vs. Metsähovi, there is a large difference between the slopes from the linear functions fitted to the unbinned and binned data. This is produced by a few HE vs. optical flux pairs which are well outside the main trend; they have a substantial impact on the fit to the unbinned data, while they do not affect the binned data.
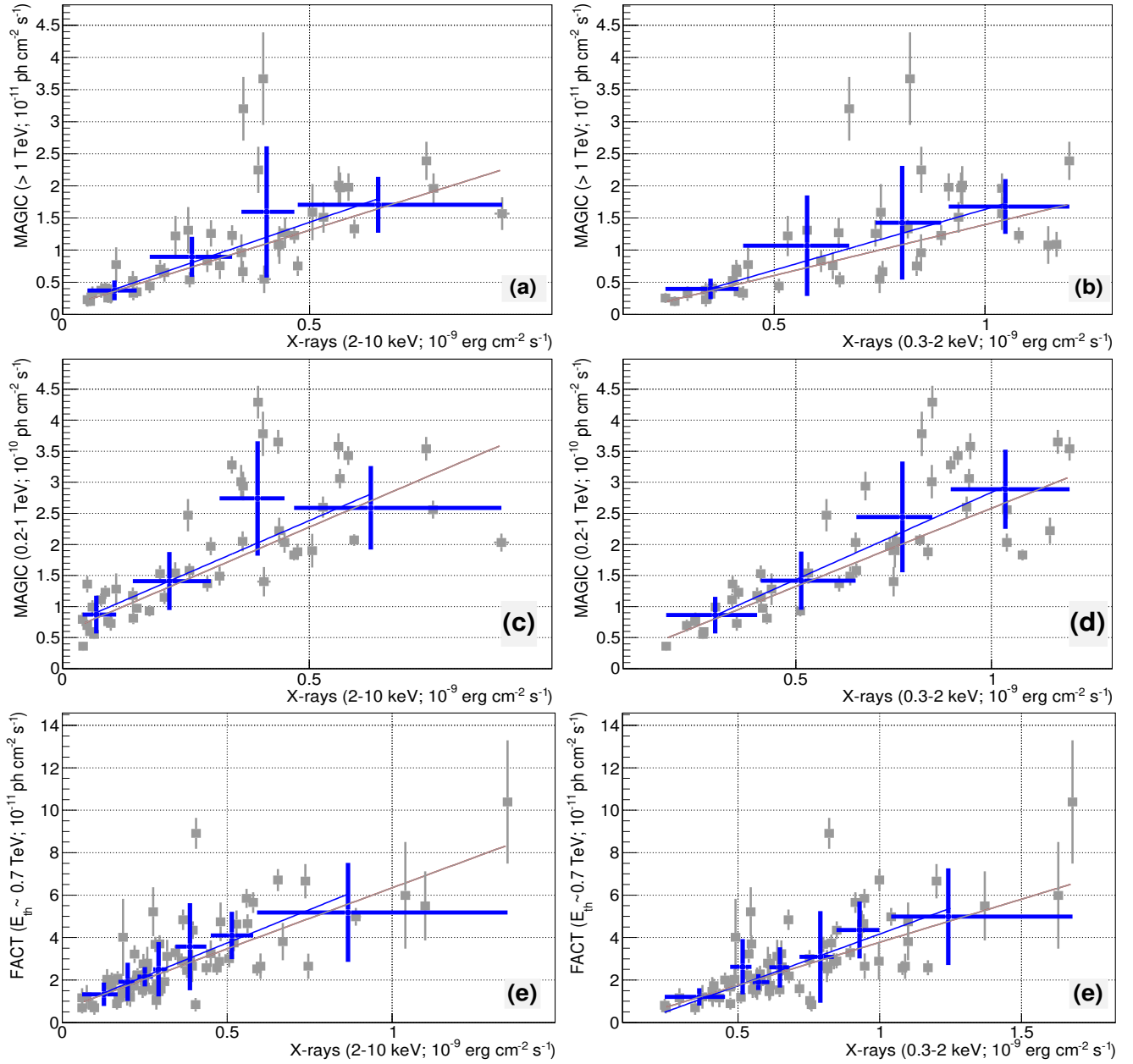
Figure C1. VHE vs. X-ray flux correlation plots during 2015–2016 campaign. The grey markers denote the individual flux measurements and related errors (unbinned data), while the blue markers show the average and the standard deviation computed with data subsets of 10 observations, binned according to their flux (binned data). The grey and blue lines depict the best linear fit to the unbinned and binned data, with the slopes reported in Table 2. Only simultaneous VHE-X-ray data (taken within 0.3 days) were used. See Section 5.1 for details.
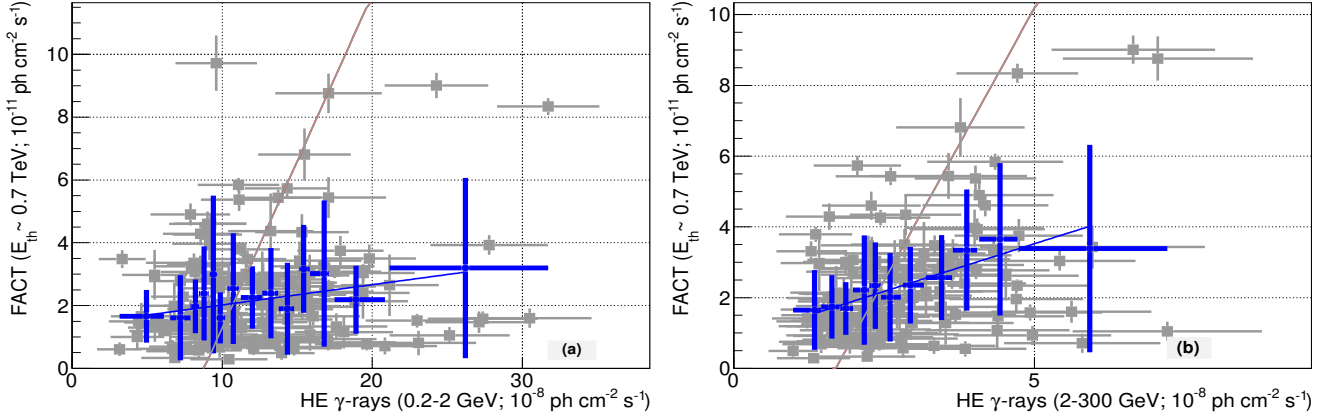
Figure C2. VHE vs. HE flux correlation plots during the period from 2012 December to 2016 June. For the description of the grey and blue markers, see the caption of Fig.C1. The grey and blue lines depict the best linear fit to the unbinned (grey) and binned (blue) data, with the slopes reported in Table 3. The flux values relate to 3-day time intervals. See Section 5.2 for details.
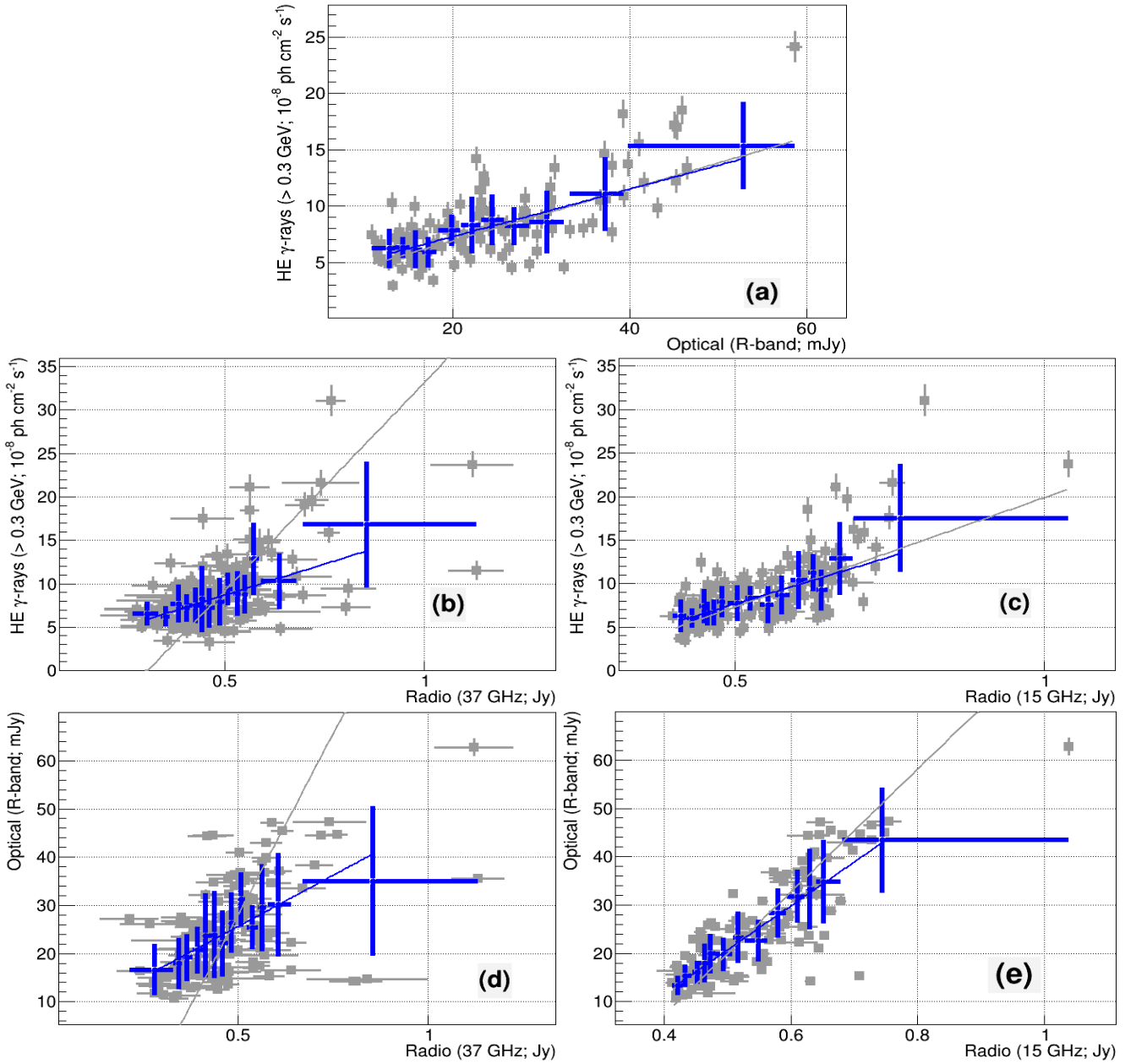
Figure C3. Flux-flux plots for several energy bands. The grey and blue markers as in Fig.C1. The grey and blue lines depict the best linear fit to the unbinned (grey) and binned (blue) data, with the slopes reported in Table 4. Panel (a) shows the flux-flux cross-correlation between the HE γ-rays (LAT; >0.3 GeV) and optical (R-band) fluxes, computed for 15-day time intervals at a zero timelag. The panels (b)-(e) report fluxes computed for 15-day time intervals, where the radio (15 GHz and 37 GHz) have been shifted 45 days earlier in order to match the time lag observed in the correlation plots from Fig. 12. See Section 5.3, 5.4, and 5.5, for details.

APPENDIX D: ESTIMATION OF THE MOST REPRESENTATIVE TIME LAG

In this section, we report an estimate of the most representative time lag and its related uncertainty for the multi-band fluxes used in the correlation studies reported in Sections 5.3, 5.4, and 5.5. We use the model-independent Monte Carlo flux randomization (FR) and random subset selection (RSS) method described in Peterson et al. (1998) and Peterson et al. (2004), which is the methodology used by Max-Moerbeck et al. (2014) to estimate the time lag of 40±9 days between the *Fermi*-LAT and OVRO fluxes. Briefly, the method employed in this study is as follows: we perform RSS of the first LC and select the simultaneous observations between the first and second LC. Then, we perform FR according to the flux uncertainties of both LCs. In this way, through this process of RSS and FR, we generate a set of 1000 Monte Carlo simulated LC pairs. Then we perform the DCF study for these 1000 simulated pairs. As in Peterson et al. (1998), a cross-correlation is considered successful if the maximum correlation coefficient is large enough such that the correlation between the LC pairs is significant above 95% confidence level. Instead of using the peak of the DCF (DCF_{max}), following the prescriptions from Peterson et al. (2004), we used the centroid of the DCF (DCF_{cen}), computed with the DCF values above $0.8 \times DCF_{max}$, which is expected to provide better results when the DCF has a broad peak. The distributions of DCF_{cen} are then obtained. The most representative value of the time lag is estimated by considering the mean of the distribution, and the uncertainties are computed using the 68% containment, that would correspond to 1σ error for a normal distribution.

Figure D1 shows the distribution of DCF_{cen} for the 1000 simulated LCs for the HE and optical R-band. The average and the 68% containment (depicted with the black and red lines in Fig. D1) is 2^{+5}_{-9}, which can be considered as good estimate of the time lag and related uncertainty between the fluxes for these two energy bands. This is perfectly consistent with no time lag, and hence simultaneous emission in these two energy bands.

The panels in Fig. D2 show the distributions of DCF_{cen} for the 1000 simulated LCs for the HE and R band vs. the two radio bands observed with Metsähovi and OVRO. Since the time lags shown in Fig D2 are statistically compatible, we decided to combine the GeV and R-band with the 37 GHz and with the 15 GHz cases, in order to estimate combined time lags for the GeV/optical and 37 GHz, and the GeV/optical and 15 GHz. The combined distributions of DCF_{cen}, derived with the 2000 simulated LCs, are shown in Fig. D3, leading to the estimation of combined time lags of 37^{+15}_{-11} days and 43^{+8}_{-5} days, respectively.
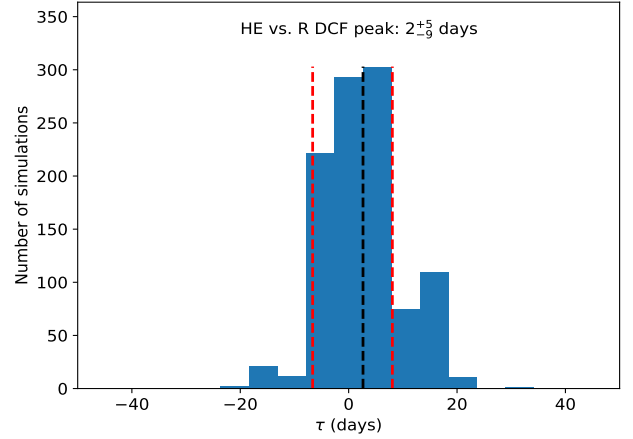
Figure D1. Distribution of DCF_{cen} derived with 1000 Monte Carlo FR/RSS simulations to estimate the time lag between the HE γ-ray and optical R-band LCs that were used to compute the DCF reported in panel *a* of Fig. 12. The average and the 68% containment are depicted with the black and red lines, respectively, and are used as the estimate of the time lag between these two bands. See text in Appendix D for further details.
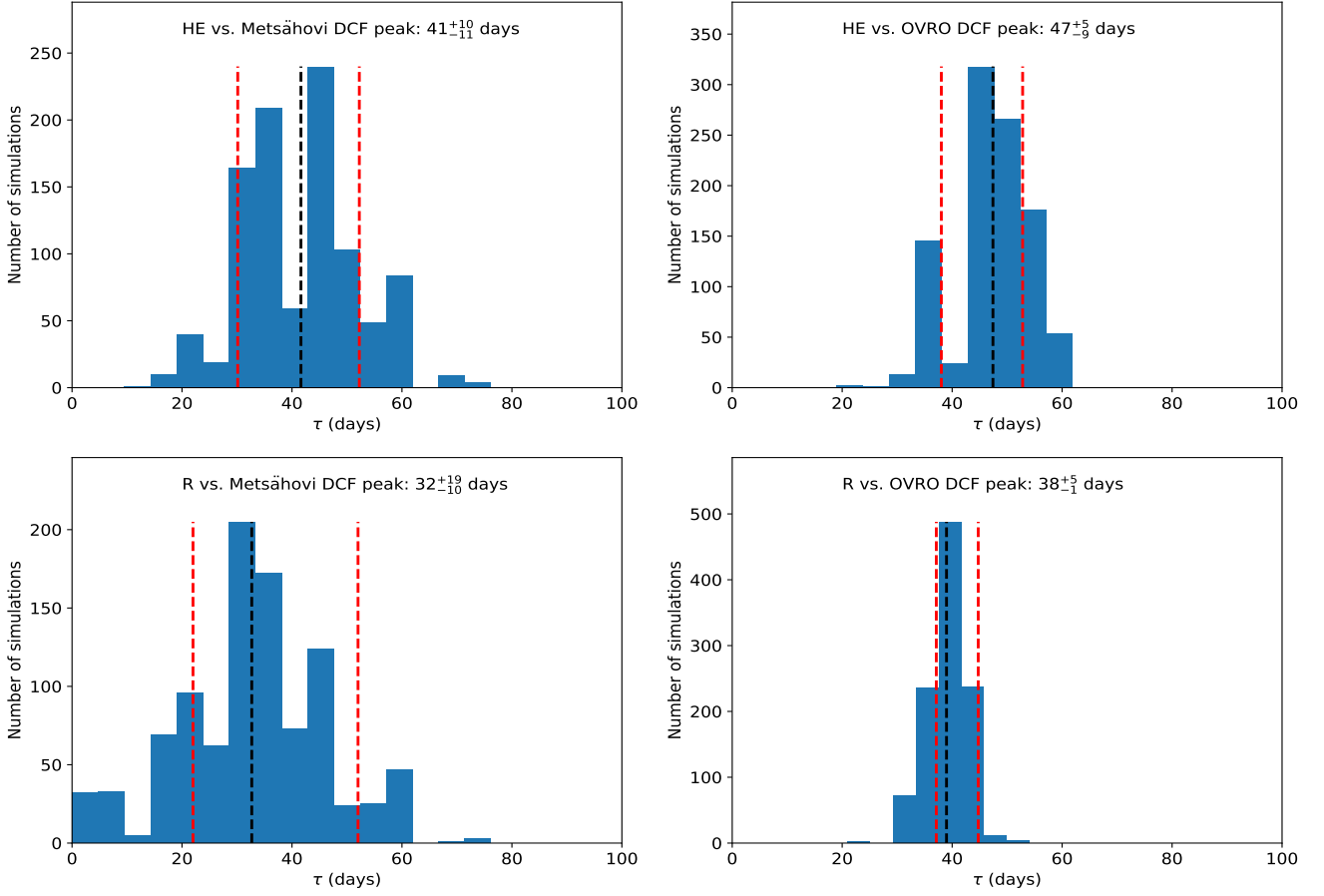
Figure D2. Distributions of DCF$_{cen}$ derived with 1000 Monte Carlo FR/RSS simulations to estimate the time lag between the various multi-band LCs that were used to compute the DCFs reported in panels *b, c, d,* and *e* of Fig. 12: a) HE vs. Metsähovi (top-left), b) HE vs. OVRO (top-right), c) R vs. Metsähovi (bottom-left) and d) R vs. OVRO (bottom-right). The average and the 68% containment are depicted with the black and red lines, respectively, and are used as the estimate of the time lag between the bands. See text in Appendix D for further details.
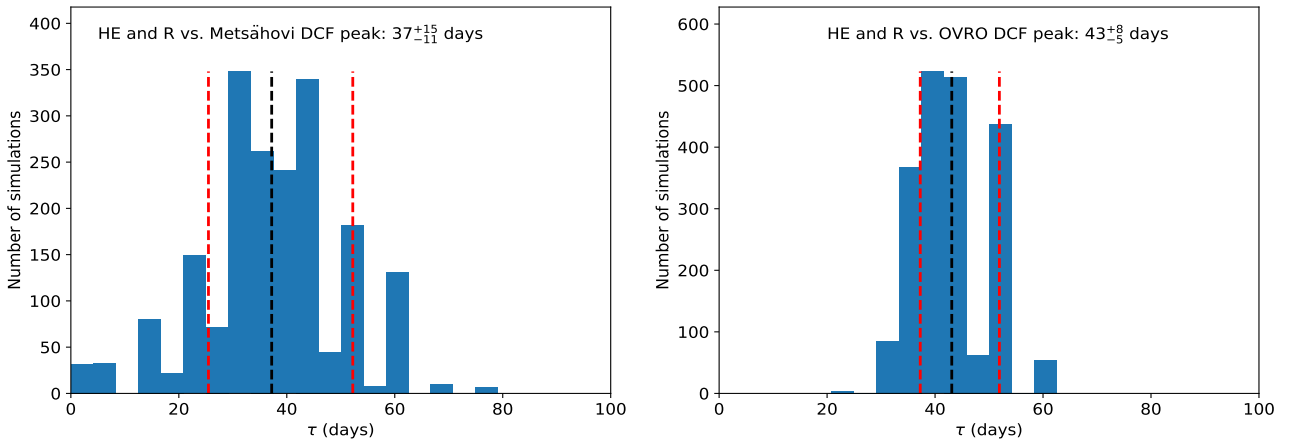
Figure D3. Distribution of DCF$_{cen}$ from the combinations of the two panels with 37 GHz Metsähovi data (left) and the two panels with 15 GHz OVRO data (right) from Fig. D2. See text in Appendix D for further details.
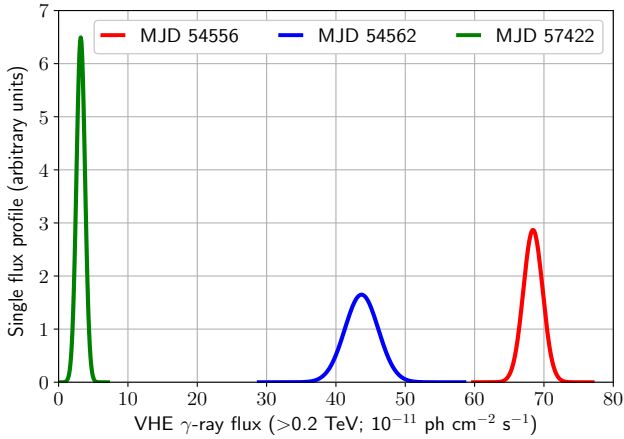
Figure E1. Examples of contributions to the VHE γ-ray flux profiles from three selected flux measurements with the MAGIC telescopes.

APPENDIX E: FLUX PROFILE

The shape of flux distribution of a source is a useful tool to study the nature of the underlying variability processes in the source. Almost all the studies done so far in this respect involve construction of Chi-square fit to the flux histograms (Tluczykont et al. 2010; Abeysekara et al. 2017; Sinha et al. 2016; Dorner et al. 2019). However, generating flux distributions from histograms has certain inaccuracies and biases related to the selection of the bin-width and the flux measurement errors, which are not considered when making a simple flux distribution. In order to address this issue, we have developed a new method, in which, we construct "flux profiles" instead of histograms.

E1 Flux profile from a light curve:

We create the flux profile by adding contributions from individual flux measurements. We assume that for individual observations, flux errors are normally distributed around the mean. At VHE γ-rays (> 0.2 TeV) during 2015–2016, the lowest number of excess events was found to be around 40, supporting this assumption. Therefore, for each individual measurement we create a Gaussian profile $G(x : \mu, \sigma)$, where μ and σ are the flux and flux error, respectively. The amplitude of the profile is normalised to $1/(\sigma\sqrt{2\pi})$, so that the area under each individual flux profile is unity. Therefore, a high uncertainty measurement will result in a smaller amplitude, but will contribute to a wider range of flux values. Finally, the overall flux profile for the whole observation period is obtained by adding contributions from individual flux profiles. A few examples of such individual flux profiles are presented in Fig. E1.

In order to create the flux profile in the VHE band for the 2007–2016 period, we have selected only flux points for which the detection significance (flux-error/flux) is less than 0.5. The highest flux in this data set, $(86.1\pm3.2)\times10^{-11}$ ph cm^{-2} s^{-1}, was observed on MJD 54555.9, while the lowest flux state of $(3.2\pm0.6)\times10^{-11}$ ph cm^{-2} s^{-1} was observed on MJD 57422. The corresponding flux profiles are presented in Fig. E1. Also, the flux profile for MJD 54562 is also shown which has a rather large flux uncertainty $(78.4 \pm8.1)\times10^{-11}$ ph cm^{-2} s^{-1}. In addition, we construct the flux profiles using the individual fluxes and flux errors scaled with the average flux of the entire observation period reported in Fig. B1 (e.g. at VHE the fluxes and errors are scaled with

2.09×10^{-10} ph cm^{-2} s^{-1} which is the long-term average flux). From the overall flux profile, we determine the most probable state to be around 60 per cent of the average flux.

E2 Validation of the flux profile method using VHE γ-ray data:

In this section, we present the validation of the flux profile method by assuming the flux distribution of the source as i) Gaussian and ii) LogNormal. We explain the procedure for this exercise for the Gaussian case and for the VHE γ-ray data set, but the same procedure also applies to LogNormal case, as well as for all the energy bands. The steps are as follows:

Step 1: We create a histogram of the fluxes in the VHE band using the long-term (2007–2016) data set, as shown in Fig. E2 (top left panel), and fit it with a Gaussian using Chi-square minimization.

Step 2: We assume that the fluxes from our source are distributed according to the fitted distribution from Step 1. We simulate 226 flux values as present in the real VHE LC.

Step 3: We then use real measurements to create a 2-D histogram of the flux vs. SNR, with 10 bins in flux and 5 bins in SNR (top right panel of Fig. E2). The SNR bins are not the same for each flux bin, rather, in each flux bin, we take the range between minimum and maximum values of the SNR and divide it in 5 bins. Finally, we take the number of points in each SNR bin and divide it with the total number of points in the whole flux bin to estimate the distribution of SNR in each flux bin.

Step 4: Using fractions of SNR in each flux bin (obtained in Step 3), we generate flux errors for each of the 226 fluxes generated in Step 2. Some high flux bins in the real data histogram are empty (see top left panel of Fig. E2). In such cases, we take the SNR to be the average SNR of the first lower flux bin.

Step 5: The 226 generated flux and flux-error pairs are now used to create a simulated flux profile.

Step 6: Steps 2–5 are repeated 1000 times in order to create 1000 generated flux profiles.

Step 7: Every generated flux profile is fitted with both the Gaussian and LogNormal functions. The fit parameters are μ_G^i and σ_G^i (μ_{LN}^i and σ_{LN}^i) for fitting with Gaussian (LogNormal), where i is the flux profile index. In addition, a parameter `redchi` is calculated (see Section 6 for details) for each flux profile and both the functions, as well as a ratio of `redchi` parameters R_{LN}^G (`redchi(G)`/`redchi(LN)`).

Step 8: Using the fit parameters for individual flux profiles, we calculate the average values of the fit parameters (μ_G and σ_G) and their standard deviations ($\Delta\mu_G$ and $\Delta\sigma_G$). A Gaussian function with μ_G and σ_G as mean and standard deviation is plotted in Fig. E2 as the reconstructed Gaussian distribution (green lines in middle left and bottom left panels). The errors on mean and standard deviation of the Gaussian are quoted as $\Delta\mu_G$ and $\Delta\sigma_G$ in the same panels. The same procedure is followed for LogNormal distribution (shown as blue lines in Fig. E2).

Step 9: We make a distribution of the R_{LN}^G which quantifies the goodness of fit.

First, we perform the described analysis by fitting the real data fluxes with a Gaussian function (Step 1). The resulting fit parameters are (μ_G= 0.67, σ_G= 0.65), and the corresponding function is shown in the top left plot of Fig. E2 with the red line. These parameters are used to generate 213 flux values (Step 2), and later to generate 1000 flux profiles (Step 6). Fitting each flux profile with Gaussian and LogNormal (Step 7) and averaging over all flux
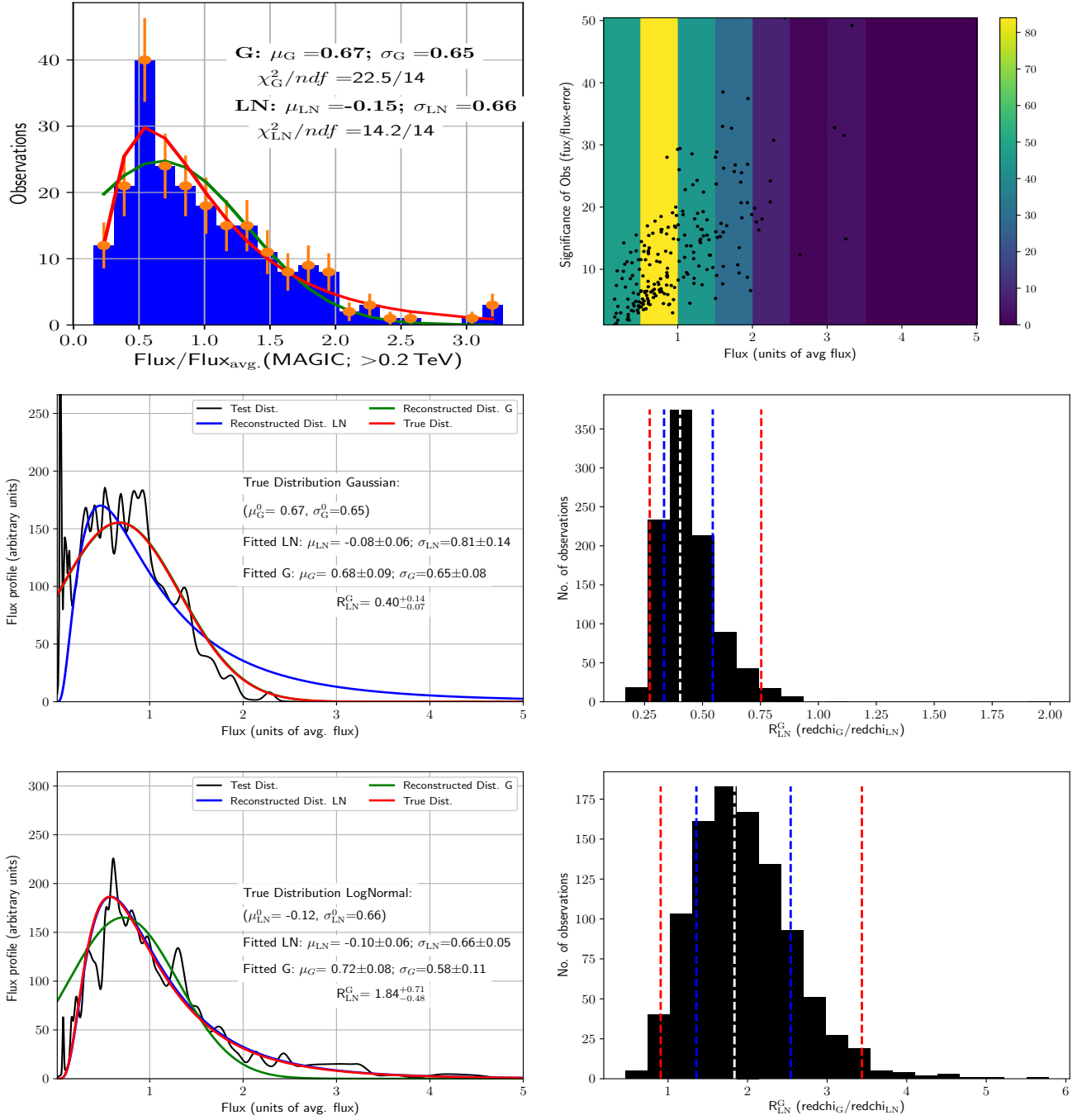
Figure E2. Validation of flux profile-method using long-term VHE observed flux. The histogram of real data (fluxes) along with fits with the Gaussian (red line) and LogNormal (green line) functions is shown in the top left plot. These functions are shown in red in the middle left (Gaussian) and bottom left (LogNormal) plots. The top right plot shows the distribution of the flux/flux-error ratio (SNR) vs. flux. The colour scale indicates the number of flux measurements in each flux bin. In the middle left and bottom left plots, the red lines represent fits to the true flux distribution with a Gaussian (middle) and LogNormal (bottom), while the green and blue lines show fits to simulated flux profiles with Gaussian and LogNormal, respectively. In each of these two plots one example of the 1000 simulated Gaussian (LogNormal) flux profiles is presented with black line. The middle right and bottom right plots the distributions of the parameter R_{LN}^G for Gaussian and LogNormal distributions, respectively. The white, blue, and red vertical dashed lines represent the weighted average of the histograms bins , the $1\,\sigma$ and $2\,\sigma$ confidence intervals.

profiles (Step 8) results in the average fit parameters (μ_G= 0.68 ± 0.09, σ_G= 0.65 ± 0.08) and (μ_{LN}= −0.08 ± 0.06, σ_{LN}= 0.81 ± 0.14), for Gaussian and LogNormal distributions, respectively. The results are shown in the middle left plot of Fig. E2. The red line indicates the fit of the real data set with the Gaussian distribution, while the
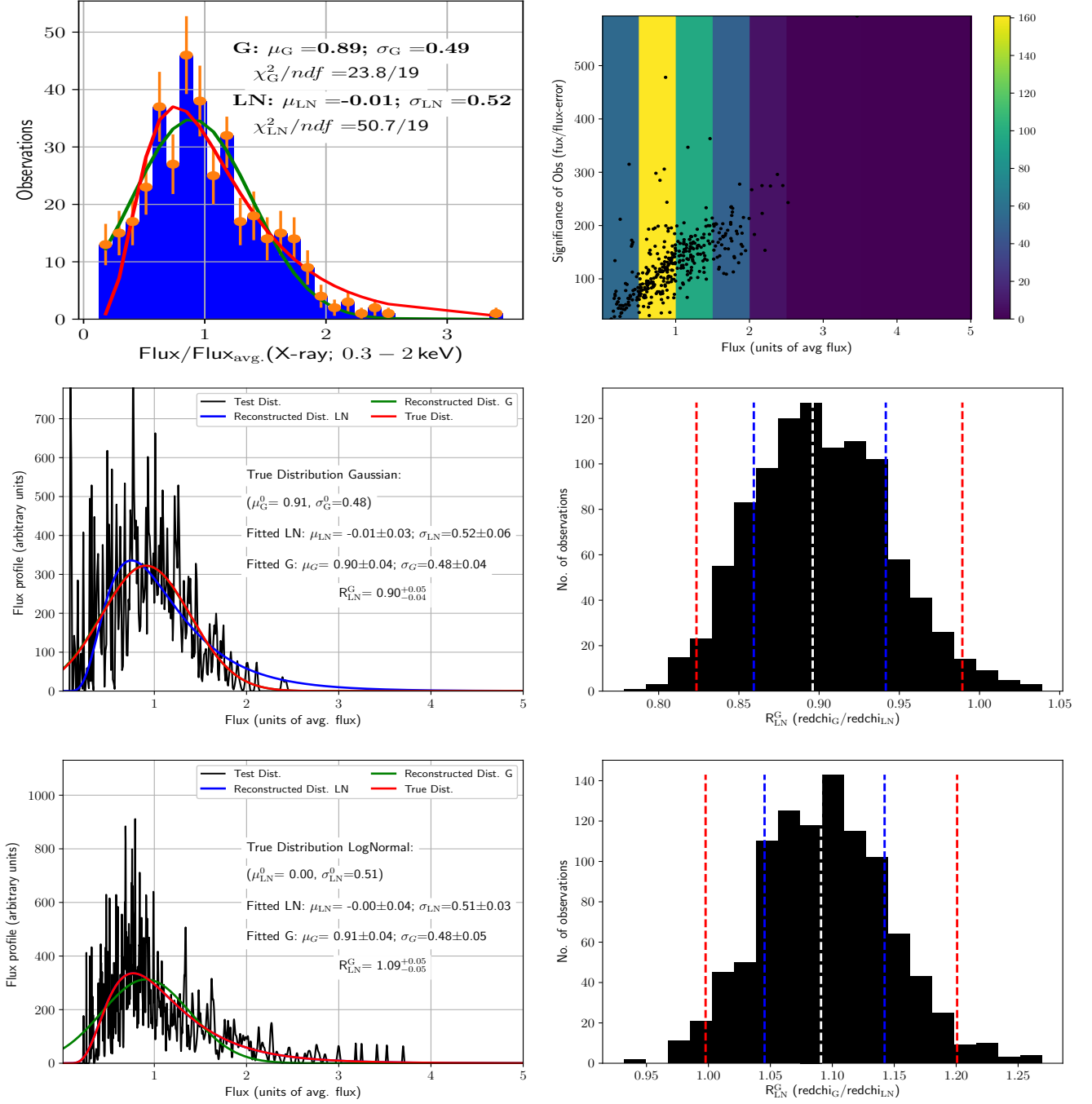
green and blue lines indicate the Gaussian and LogNormal functions, respectively. The average fit parameters for the Gaussian are consistent with the fit parameters of the initial real data distribution (the red and green lines overlapping). In addition, the ratio of the parameters `redchi` for Gaussian to the LogNormal for this case is
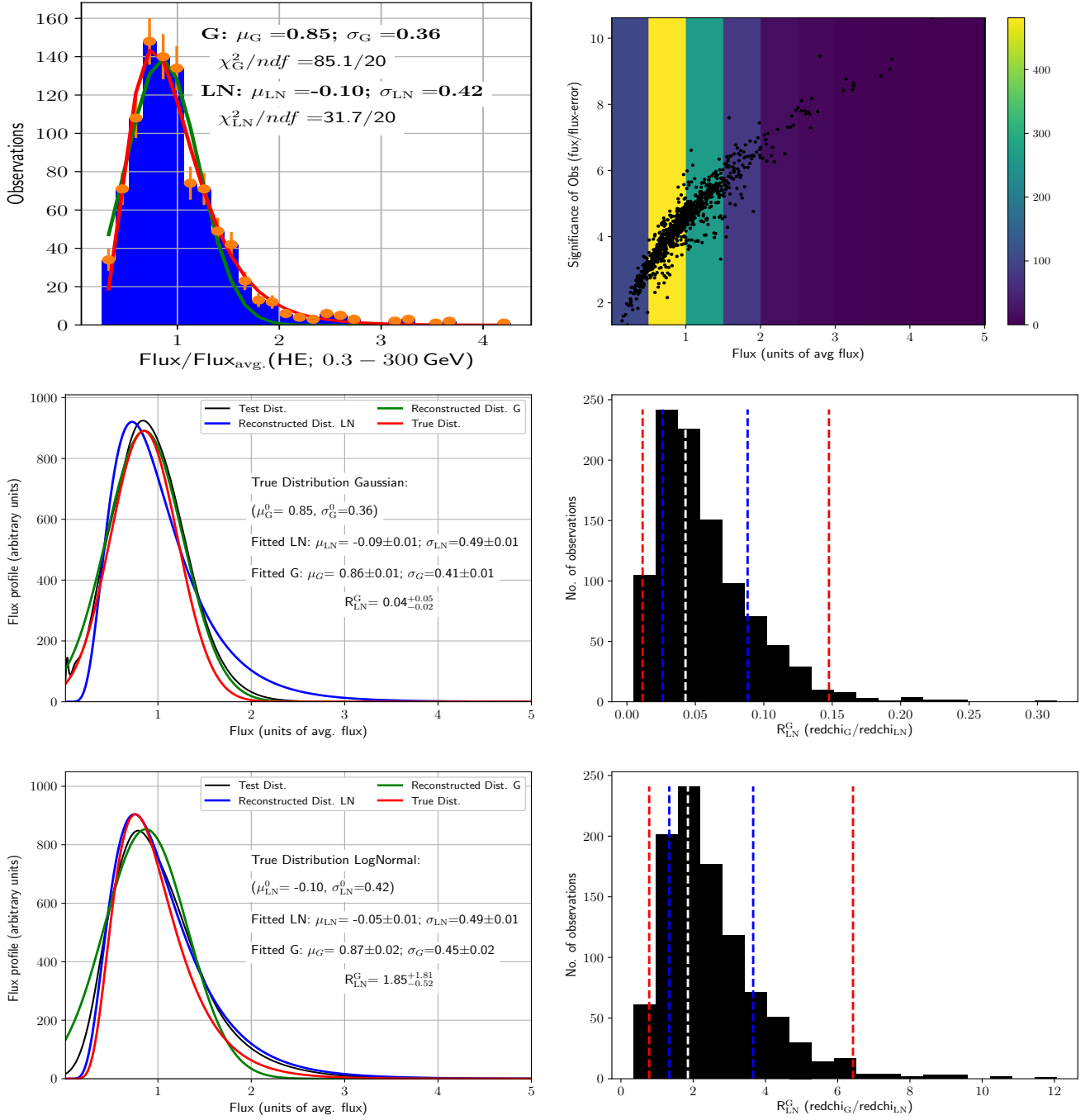
Figure E3. Validation of flux profile-method using long-term HE γ-ray (0.3-300 GeV) data

$R_{LN}^{G} = 0.40^{+0.14}_{-0.07}$, indicating that the Gaussian distribution is the preferred one, and thus proving that we correctly recovered the initial distribution. The chance probability (p), based on toy Monte Carlo, indicates the probability of wrongly reconstructing a LogNormal (Gaussian) distribution as a Gaussian (LogNormal). We calculate this by the distribution of the parameter R_{LN}^{G}. For an initial true LogNormal (Gaussian) distribution, we calculate the survival function (*sf* [15]) of R_{LN}^{G} below (above) 1 assuming the distribution to be

[15] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.skewnorm.html

a *skew-normal*[11]. This survival fraction indicates the chance probability of obtaining a Gaussian (LogNormal) flux distribution from a true LogNormal (Gaussian) distribution. The chance probability for the flux distribution in VHE γ-rays is 1.1×10^{-4}. The distribution of the R_{LN}^{G} for individual simulated flux profiles is shown in the middle right plot. Next, we repeat the analysis, this time fitting the real data fluxes with the LogNormal function (Step 1), resulting in parameters (μ_{LN}= -0.12, σ_{LN}= 0.66). The corresponding function is shown in the top left plot of Fig. E2 with the green line. The final results of the analysis are shown in the bottom plots of Fig. E2. In the left plot, the red line indicates the fit of the real data set with

Figure E4. Validation of flux profile-method using long-term X-ray $(0.3 - 2\,\text{keV})$ data

the LogNormal distribution, while the green and blue lines again indicate the Gaussian and LogNormal functions, respectively. This time, the simulated flux profiles were generated using parameters of the initial LogNormal distribution. The average fit parameters in this case are $(\mu_G = 0.72 \pm 0.08, \sigma_G = 0.58 \pm 0.11)$ and $(\mu_{LN} = -0.10 \pm 0.06, \sigma_{LN} = 0.66 \pm 0.05)$, while $R^G_{LN} = 1.84^{+0.71}_{-0.48}$ (chance probability of having a Gaussian distribution is 4.4×10^{-2}). We can see that the average fit parameters for the LogNormal are consistent with the fit parameters of the initial real data distribution (the red and blue lines overlapping), and that the LogNormal distribution is the preferred one. Therefore, we again correctly recovered

the initial distribution. The distribution of the R^G_{LN} for individual simulated flux profiles is shown in the bottom right plot.

We inspected our method on flux profiles in HE and X-ray bands. HE was chosen as an example of a band with larger relative flux uncertainties and lower variability, while the X-ray band is an example of the opposite (smaller relative flux uncertainties and higher variability). The procedure used in HE (X-ray) is exactly the same as that of the VHE band, with the exception of using 955 (374) flux points in Step 2 and 15 (10) flux bins in the 2-D histogram in Step 3. The results are shown in Fig. E3 and E4 for the HE and X-ray bands, respectively. In the HE band, the

parameters of the initial flux distributions are not recovered. This is mainly because of the relatively large flux uncertainties. However, the chance probability of having LogNormal (Gaussian) from a true Gaussian (LogNormal) is 0.0 (8.1×10^{-2}) which indicates that we do correctly reconstruct and distinguish between Gaussian and LogNormal shapes of the initial distribution. In the X-ray band, the chance probability of having LogNormal (Gaussian) from a true Gaussian (LogNormal) is 1.7×10^{-2} (3.3×10^{-2}). Therefore, in the VHE and X-ray bands, we were able to recover the initial flux distributions (including the parameters), thus validating our method for measurements with higher sensitivity.

We recognize the following types of biases that can affect our results:

a) A cut on the relative error: In this study, we only use flux measurements with a SNR > 2.0. This will bias towards slightly higher values of flux for some of the distributions (those with the largest errors), e.g., FACT, *Fermi* and BAT. It affects only the rising part of the flux distribution. For FACT, there will be some distortion in the distribution because we remove 25% of the data. In any case, it is the high fluxes what dominates the distinction between G and LN, and those remained unaffected. For the flux distribution of data from the *Fermi*-LAT, the impact is negligible (only 3% of data removed).

b) The bias for including the observations during alert (ToO) for the high flux states: The MAGIC and *Swift*-XRT observations triggered by the target of opportunity (ToO) programs during the high flux of the source may bias the flux distribution. Ideally, the unbiased observations should only be considered. The data set under consideration includes the following campaigns, 2008 (Aleksić et al. 2012) and 2010 (Aleksić et al. 2015c; Abeysekara et al. 2020) where the source showed high flux states. While the 2008 flaring episode had many ToOs involved, the 2010 March flaring activity observed consisted on observations that had been coordinated with *Swift* and *RXTE* several weeks in advanced. We have performed a study by removing all the high flux states observed during 2008 (Aleksić et al. 2012) to check for LogNormality. We have found that even with this extreme condition, a LogNormal is preferred over a Gaussian flux profile. This proves that the LogNormal distribution of the flux at VHE is a feature of the source and does not depend on the ToOs. We also note that this bias is negligible for MAGIC and *Swift*-XRT and has no effect on the observations with the FACT, *Fermi*-LAT and *Swift*-BAT.
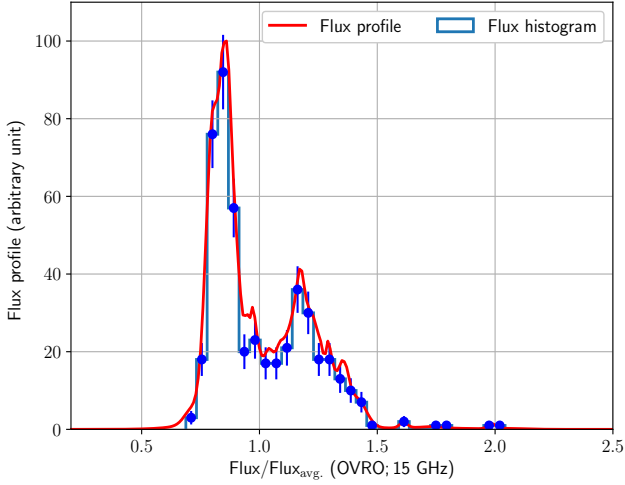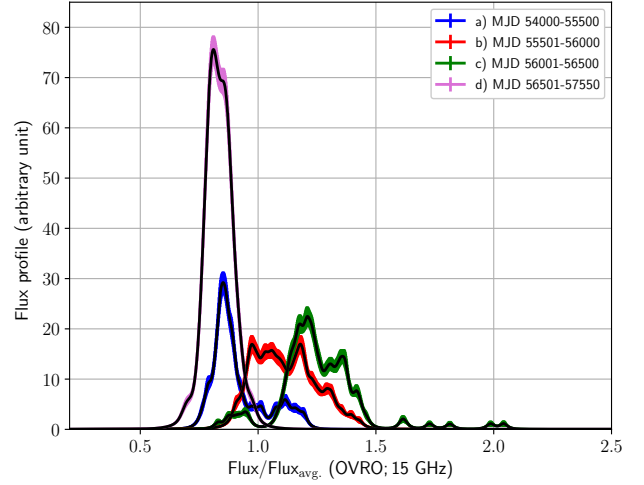
Figure F1. The flux profile and flux histogram of Mrk 421 in radio (15 GHz) band. See text for details.

APPENDIX F: YEAR-WISE VARIATION OF FLUX PROFILES

The flux distribution of the 15 GHz radio band is shown in Fig. F1, where the flux histogram and the flux profile are presented. This suggests that the flux distribution for this band is a bimodal distribution, hence, the Gaussian and LogNormal functions are not suitable. The year-wise variation of the flux profiles for the 15 GHz radio band, observed by OVRO, has been reported in Fig. F2. We have divided the multiyear radio LC into four different periods:
a) 2006 September 22 to 2010 October 31 (MJD 54000–55500),
b) 2010 November 01 to 2012 March 14 (MJD 55501–56000),
c) 2012 March 15 to 2013 July 27 (MJD 56001–56500), and
d) 2013 July 28 to 2016 June 11 (MJD 56501–57550).
For each of the periods stated above, we shuffled the uncertainties on the flux and added to the flux in order to construct a simulated flux profile. We repeated this exercise for 1000 times for a single period in order to estimate the standard deviation on the flux profiles. The bands in Fig. F2 represent the standard deviation on the flux profile (68% confidence limit) estimated from the simulations mentioned above. This study indicates that the most probable states of the source in different years are not unique. The flux profile also changes according to the flux states in different years. For example, the flux profile for the period (c) shows an isolated peak at higher flux. This is due to the huge radio flaring event in 2012. The variation in flux profiles in Fig. F2 indicates a shift form the low-flux state in period (a) to a high flux state (c) via an intermediate state (b). During period (c) the low/ typical state can also be identified.

Figure F2. Year-wise variation of flux profiles of Mrk 421 in radio (15 GHz) band. The bands for different colors indicate 1σ confidence intervals for different years. We have divided the multiyear radio LC into four different periods of observations. See Appendix F for details.

APPENDIX G: CHARACTERIZATION OF THE MWL FLUX DISTRIBUTIONS USING A (BINNED) CHI-SQUARE FIT AND A (UNBINNED) LOG-LIKELIHOOD FIT

This section reports the characterization of the flux distributions using a binned Chi-square fit and an unbinned log-likelihood fit. They are conventional ways of quantifying the shape of a distribution, and complement the results obtained with the flux profile method reported in Section 6 and Appendix E. In both exercises, we use fluxes and their errors scaled by the average flux for each of the energy bands, and present them as F and ΔF.

In order to perform the Chi-square fit, we first bin the scaled flux F. For each of the energy bands, the number of histogram bins employed permits to show the overall shape of the distribution, while keeping sufficient statistics (more than 10 entries) in most of the bins. Afterwards, we performed a regular fit with a Gaussian and LogNormal functions, starting from the minimum flux F_{min}, and obtaining the function parameters mean (μ) and standard deviation (σ) for which the Chi-square is minimum. It must be noted that the outcome of the Chi-square fit can depend on the histogram binning, and does not consider the flux uncertainties. Figure G1 shows the results of the Chi-square fit for all the bands, except for the VHE γ-ray with MAGIC, HE γ-ray with *Fermi*-LAT and X-ray ($0.3 - 2\,$keV) with *Swift*-XRT, which are presented in Fig. E2, E4, and E3, respectively.

In the log-likelihood fit, the log-likelihood function used for the Gaussian PDF, as a function of the parameters μ and σ, is given as:

$$L_g(f(F_i, \Delta F_i)|\sigma, \mu) = -\frac{1}{2} \sum \left[log(2\pi(\Delta F_i^2 + \sigma^2)) + \frac{(F_i - \mu)^2}{\Delta F_i^2 + \sigma^2} \right]$$
(G1)

In order to calculate the log-likelihood of the LogNormal distribution, we consider a grid with 3000 points (x_i), using a dynamic grid resolution, ranging from $log(F_{min})$-5 to $log(F_{max})$+5, where F_{min} and F_{max} are the minimum and maximum scaled fluxes in the corresponding energy bands. The exponential of the grid points (e^{x_i}) are then used for the defining the LogNormal PDF as a function of the parameters μ and σ, in the form given below:

$$L_{LN}^i(f(x_i|\sigma, \mu) = -\frac{1}{\sqrt{2\pi}} \frac{1}{e^{x_i}\sigma} exp\left[-\frac{(x_i - \mu)^2}{2\sigma^2} \right]$$
(G2)

Next, we calculate the Gaussian probability $G_i^j(F_j, \Delta F_j)$ for each of the flux measurements with measured flux (F_j) and flux-error (ΔF_j) using the following equation (Eq. G3) at different grid-points (x_i).

$$G_i^j(F_j, \Delta F_j) = -\frac{1}{\sqrt{2\pi}} \frac{1}{\Delta F_j} exp\left[-\frac{(e^{x_i} - F_j)^2}{2\Delta F_j^2} \right]$$
(G3)

We obtain the log-likelihood by the convolution of these two terms and integrating over the grid-range. Finally, we minimize the log-likelihood and obtain the optimal parameters for μ and σ. The results are presented in Fig. G2 where, for completeness, the flux histograms used in Fig. G1 are also shown. By construction, the log-likelihood fit considers the flux uncertainties, and does not require to bin the data, both representing big advantages over the Chi-square fit. However, we note that log-likelihood fit applied here is very simple and generic because we are using the same PDF functions for all the energy bands. One could exploit the full potential of the log-likelihood method by using dedicated PDFs for each energy band, which would allow one to address the problem in a more efficient manner. However, that would require introducing instrument response functions and physical models of emission into the PDFs, which is out of the scope of this work.

Table G1 lists the preferred flux-distributions (Gaussian or LogNormal) from the Chi-square fit, log-likelihood fit and the flux profile methods. The flux distribution for the OVRO (15 GHz) is not included in the table and in the figures because it has a bimodal shape due to the strong flare in 2012 (see Appendix F). The entries marked with "*" denote cases where the preference is not clear, either because both options are roughly equally probable, or because the methods suffer from some caveats. In the case of the Chi-square fit, this happens for FACT (E_{th} ~0.7 TeV), where the resulting Chi-square values show equally probable fits. In the case of the log-likelihood fit, this occurs for FACT (E_{th} ~0.7 TeV), *Swift*-BAT ($15 - 50\,$keV) and X-ray in the $0.3 - 2\,$keV band. In the first two cases, the applicability of the PDF (Gaussian or LogNormal) suffers from the truncation of these two distributions at low flux values (given the limited sensitivity to measure low fluxes)[16], and in the latter case, the resulting log-likelihood values are equal (within one unit) for both the functions. In the case of the flux profile method, the preference is not clear for *Swift*-BAT ($15 - 50\,$keV) because the chance probability (p; see Section 6 and Appendix F for details) for a Gaussian distribution when the true distribution is a LogNormal is only 0.16. The table shows preference for the LogNormal distribution shape in all of the energy bands, apart from the X-rays in the $0.3 - 2\,$keV, $2 - 10\,$keV and the 37 GHz radio band. The three methods prefer the Gaussian shape for the $0.3 - 2\,$keV (although the preference is not clear in the case of the log-likelihood method), while for the other two bands, the Chi-square and log-likelihood fits prefer a LogNormal shape, while the flux profile method prefers a Gaussian shape.

This paper has been typeset from a TₑX/LATₑX file prepared by the author.

---

[16] The Chi-square fit and the flux profile method are less sensitive to this effect because the fits are performed above the minimum flux F_{min}, and hence do not need to apply the entire distribution shape to the available data.
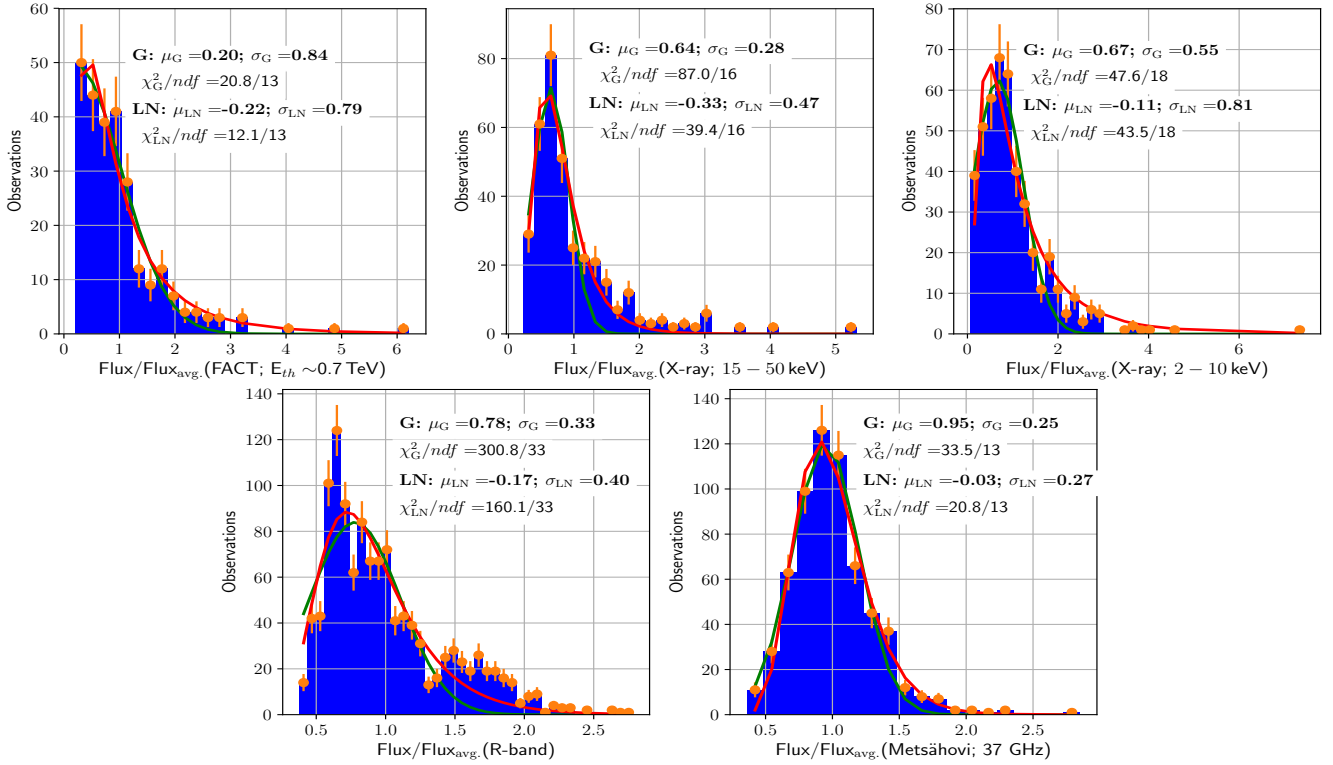
Figure G1. Characterization of the MWL flux distributions with a Chi-square fit. The X-axis shows the scaled flux and the Y-axis presents the number of observations. The green and red lines represent the best fit with the Gaussian and LogNormal functions for flux histograms (presented in blue). See text in Appendix G for details.

Energy-bands	Chi-square fit	Log-likelihood fit	Flux profile
VHE γ-rays (> 0.2 TeV)	LogNormal	LogNormal	LogNormal
VHE γ-rays (FACT; $E_{th} \sim 0.7$ TeV)	LogNormal*	LogNormal*	LogNormal*
HE γ-rays (LAT; > 0.3 GeV)	LogNormal	LogNormal	LogNormal
X-ray (BAT; $15 - 50$ keV)	LogNormal	LogNormal*	LogNormal*
X-ray ($2 - 10$ keV)	LogNormal	LogNormal	Gaussian
X-ray ($0.3 - 2$ keV)	Gaussian	Gaussian*	Gaussian
Optical (R-band)	LogNormal	LogNormal	LogNormal
Radio (Metsähovi; 37 GHz)	LogNormal	LogNormal	Gaussian

Table G1. The preferred flux-distributions based on the three methods namely Chi-square fit, log-likelihood fit and the flux profile method. Entries in the table that are marked with "*" do not have a clear preference for Gaussian or LogNormal and are discussed in Appendix G . The flux distribution for OVRO (15 GHz) has a bimodal shape, hence, it is not included in this comparison table. See Section 6 and text in Appendix G for details.
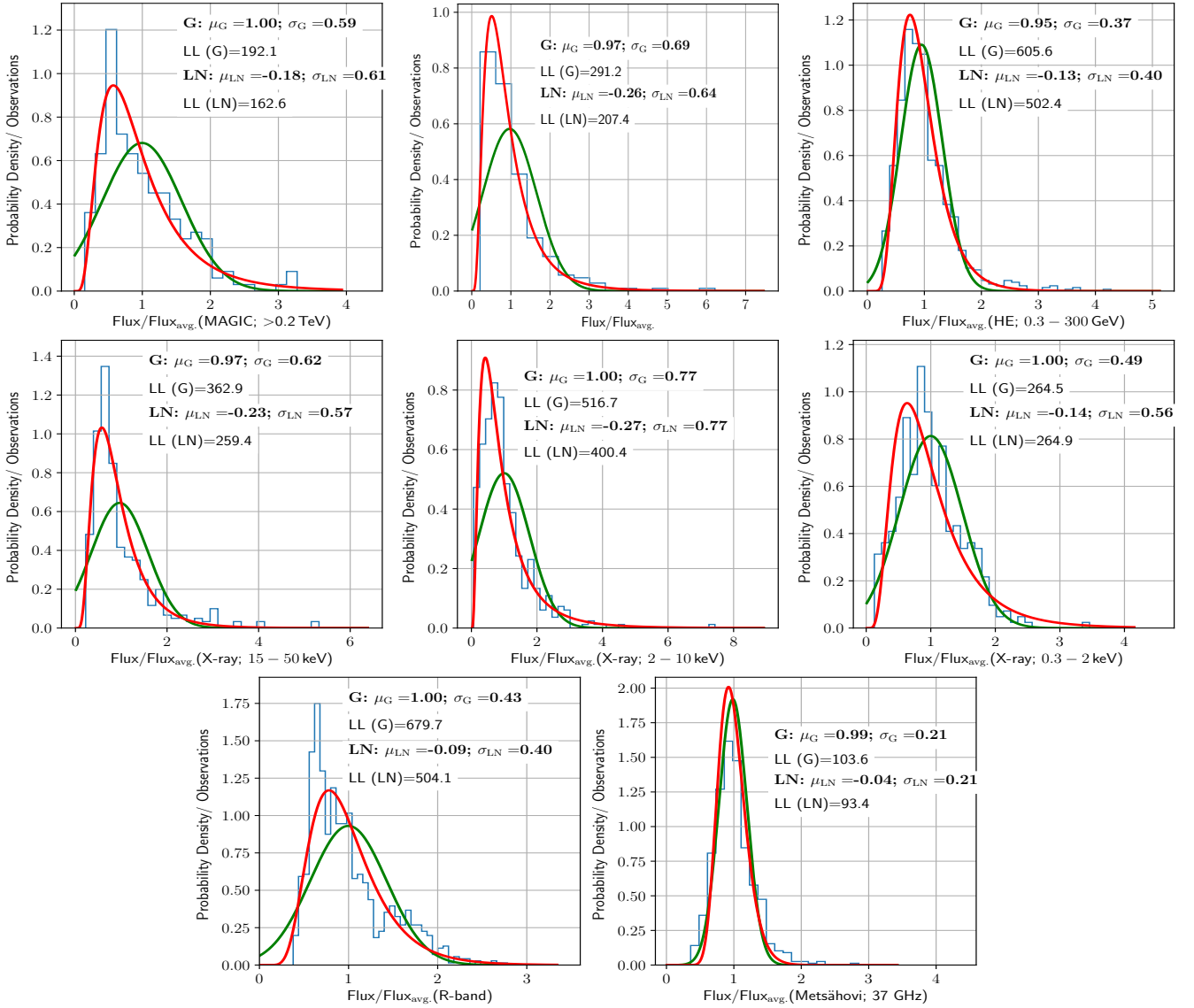
Figure G2. Characterization of the MWL flux distributions with a log-likelihood fit. The X-axis shows the scaled flux and the Y-axis the probability density. The green and red lines represent the Gaussian and LogNormal functions for which the log-likelihood is minimum. For completeness, the flux histograms used in Fig. G1 are also shown in blue. See text in Appendix G for details.