



An Empirical Background Model for the NICER X-Ray Timing Instrument

Ronald A. Remillard¹ , Michael Loewenstein^{2,3}, James F. Steiner⁴, Gregory Y. Prigozhin¹, Beverly LaMarr¹, Teruaki Enoto⁵ , Keith C. Gendreau³ , Zaven Arzoumanian³, Craig Markwardt³ , Arkadip Basak⁶ , Abigail L. Stevens^{7,8,11} , Paul S. Ray⁹ , Diego Altamirano¹⁰ , and Douglas J. K. Buisson¹⁰

¹ MIT Kavli Institute for Astrophysics & Space Research, MIT, 70 Vassar Street, Cambridge, MA 02139, USA; rr@space.mit.edu

² Department of Astronomy, University of Maryland, College Park, MD 20742, USA

³ X-Ray Astrophysics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

⁴ Smithsonian Astrophysical Observatory, 60 Garden Street, Cambridge, MA 02138, USA

⁵ Extreme Natural Phenomena RIKEN Hakubi Research Team, Cluster for Pioneering Research, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

⁶ Anton Pannekoek Institute, University of Amsterdam, Science Park 904, 1098XH Amsterdam, Netherlands

⁷ Department of Physics & Astronomy, Michigan State University, 567 Wilson Road, East Lansing, MI 48824, USA

⁸ Department of Astronomy, University of Michigan, 1085 South University Avenue, Ann Arbor, MI 48109, USA

⁹ Space Science Division, U.S. Naval Research Laboratory, Washington, DC 20375, USA

¹⁰ Department of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK

Received 2020 October 14; revised 2022 January 5; accepted 2022 January 11; published 2022 February 15

Abstract

Neutron Star Interior Composition Explorer has a comparatively low background rate, but it is highly variable, and its spectrum must be predicted using measurements unaffected by the science target. We describe an empirical, three-parameter model based on observations of seven pointing directions that are void of detectable sources. Two model parameters track different types of background events, while the third is used to predict a low-energy excess tied to observations conducted in sunlight. An examination of 3556 good time intervals (GTIs), averaging 570 s, yields a median rate (0.4–12 keV; 50 detectors) of 0.87 c s^{-1} , but in 5% (1%) of cases, the rate exceeds 10 (300) c s^{-1} . Model residuals persist at 20%–30% of the initial rate for the brightest GTIs, implying one or more missing model parameters. Filtering criteria are given to flag GTIs likely to have unsatisfactory background predictions. With such filtering, we estimate a detection limit, 1.20 c s^{-1} (3σ , single GTI) at 0.4–12 keV, equivalent to $3.6 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$ for a Crab-like spectrum. The corresponding limit for soft X-ray sources is 0.51 c s^{-1} at 0.3–2.0 keV, or $4.3 \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$ for a 100 eV blackbody. These limits would be four times lower if exploratory GTIs accumulate 10 ks of data after filtering at the level prescribed for faint sources. Such filtering selects background GTIs 85% of the time. An application of the model to a 1 s timescale makes it possible to distinguish source flares from possible surges in the background.

Unified Astronomy Thesaurus concepts: X-ray astronomy (1810); Astronomical instrumentation (799); Observational astronomy (1145)

1. Introduction

The Neutron Star Interior Composition Explorer (NICER) is a NASA mission for X-ray astronomy that has been operating on the International Space Station (ISS) since it was launched and deployed in 2017 June (Gendreau et al. 2016). The NICER X-ray Timing Instrument (XTI) consists of 56 identical and co-aligned cameras, each containing an X-ray concentrator (XRC; Okajima et al. 2016) and a customized Si drift detector positioned in the concentrator's focal plane. The other primary components of NICER are a target acquisition and tracking platform and seven electronics boxes, each of which services event processing from eight detectors (Prigozhin et al. 2016). Each detector package (detector, preamplifier, and thermoelectric cooler) is known as a Focal Plane Module (FPM), and each electronics box is referred to as a Measurement and Power Unit (MPU). The XTI sensitivity range is 0.2–12 keV, the energy resolution is typical of Si detectors (e.g., 150 eV FWHM at 6.5 keV), and detected

events are time-tagged to an absolute accuracy of 100 ns. The combined detector output from the 50 best-performing FPMs offers substantial throughput, e.g., with 10,500 counts per second (c s^{-1}) from the Crab Nebula over the range 0.4–12 keV.

The FPM and MPU designs and interworking are described in Prigozhin et al. (2016). Here, we summarize the details that are most relevant to the background model at hand. Each FPM is a single-channel device that is collimated to view a circular celestial area with radius of $3/17$. The multilayer collimator (1 mm radius) captures more than 90% of the light in the concentrator's point-spread function, and it limits the travel time to the anode for X-ray events, while the active area under the collimator extends to a radius of 2.8 mm. All NICER observations contain events from both the science target and the various types of background that are encountered while operating in space. Scientific analyses thus require a model that can predict the background spectrum so that the target spectrum can be isolated. If a model uses detector measurements to predict the background spectrum, then such events must exclude the X-rays from the science target. This strategy was used in previous X-ray astronomy missions, e.g., for the Photon Counting Array of the Rossi X-ray Timing Explorer (RXTE) Mission (Jahoda et al. 2006). To assist modeling efforts, NICER routinely schedules observations of seven sky positions

¹¹ NSF Astronomy & Astrophysics Postdoctoral Fellow.



that are void of detectable point sources. These targets (inherited from RXTE) are named “BKGD_RXTE#,” with numbers 1–6 and 8. Position #7 was eliminated as a NICER background target for the presence of a soft X-ray source (bright star).

Herein, we describe an empirical background model, named the “3C50” model, because it is based on three parameters, because the format assumes that spectral extractions will be made from standard NICER “cleaned” event lists (see below), and because the model uses the data from 50 of the 52 FPMs operating in the XTI. The remaining 4 FPMs, of the original 56, are not operating. The 3C50 model uses libraries constructed by sorting and combining the spectra extracted from observations of the background fields. Each library spectrum is the sum of the spectra within a cell defined by intervals in the adopted model parameters, as described below. We note that the selection of 50 FPMs is chosen to maximize the data contributing to the libraries, avoiding two FPMs with many occasions of noisy or unusual behavior.

The FPMs are fabricated identically, and the 3C50 model assumes that the predicted background spectrum can be rescaled for any number of selected FPMs, in a given good time interval (GTI), when there is no aberrant FPM noise. The same scalability assumption was also made for the first successful detector response calibration (i.e., the “rmf” file) distributed by the NICER team via the NASA HEASoft Calibration database. The similarity of the FPM responses is quantified, below, when observations of the Crab Nebula are considered.

Prelaunch analyses predicted that the NICER background would primarily consist of a very small contribution from the cosmic diffuse X-ray background (e.g., Wu et al. 1991), given the small field of view of the instrument (31.6 square arcmin), and a component arising from a subset of particle interactions that deposit energy indistinguishable from in-band X-rays. The 3C50 model covers these components. Additional background sources that are not considered in the 3C50 model include enhanced diffuse X-rays from hot gas in the Milky Way (dependent on galactic latitude), possible soft X-rays related to Solar activity, and possible contamination from the Earth limb or the radiation sources on the Soyuz spacecraft. In prelaunch analyses, the primary background components were expected to yield, for the majority of the ISS orbit, 0.2 c s^{-1} in soft X-rays at 0.4–2.0 keV, and an additional 0.15 c s^{-1} at 2–8 keV.

Our choice of model parameters (see below) requires additional introductory explanations about the signal-processing steps in the MPU (see Prigozhin et al. 2016). The signal line for each FPM is replicated, so that events can be found and processed independently with “fast” and “slow” measuring chains that use circuits with different time windows. The fast chain (84 ns nominal shaping time) produces time tags with higher precision, while the slow chain (465 ns shaping time) more effectively integrates the total electron yield, with lower noise, providing better measurements of the event energy. Event detections can trigger on either measuring chain, when the rate of change in the signal line exceeds a trigger threshold held in the MPU, per FPM and per chain. The trigger thresholds are chosen to admit a noise rate of $\sim 3 \text{ c s}^{-1}$ per FPM, per measuring chain, as measured below 0.25 keV when there are no sources in front of the detectors. Such noise events will be asynchronous, i.e., they will trigger only one measuring chain. On the other hand, X-ray and particle events will usually

trigger both measuring chains, yielding a single event that includes two measurements of the event energy. The caveat, here, is that the fast chain, with a higher noise level than the slow chain, has a lower trigger efficiency at energies below 1 keV. X-ray events from the source that trigger the slow chain, but not the fast chain, are likely to be in the range 0.2–0.6 keV, i.e., above the slow threshold and below the fast threshold.

When the travel path in the detector is long, from the point of incidence to the charge-collecting anode at the center of the active Si region, then the size of the charge cloud, and hence the temporal profile of the event, is elongated by charge diffusion. With its longer shaping time, the slow chain is more immune to such effects, since it has a longer time to integrate the charge. Incomplete charge collection will lower the reported event energy, and so the ratio of the slow chain energy and the fast chain energy systematically increases when the point of incidence is near the outer edges of the detector, i.e., beyond the inner ring of the collimator. This is an important detail for the background model, since prelaunch simulations had shown that particle interactions with the detector would generally produce event energies well above the 12 keV limit of the concentrator’s effective area—and thus be rejected—except for edge-clipping events near the outer edges of the active Si area. NICER’s calibrated event lists are given in the “pulse invariant” (PI) convention, where PI is the calibrated energy value from the slow chain in 10 eV units, PI_FAST is the calibrated energy in the fast chain, and the $\text{PI_ratio} = \text{PI}/\text{PI_FAST}$, for each event that triggers both of the measuring chains (and $\text{PI_ratio} = \text{INDEF}$, otherwise). The events from the predicted edge-clipping particles that can mimic X-ray events in PI value must travel a long path to the anode, resulting in the increased spread of the charge cloud, associated with the increased drift time. The measurement of such events would then show anomalously high values in the PI_ratio , since the “ballistic deficit” in the fast chain will be more substantial, compared to the slow chain.

The final topic of NICER signal processing that is pertinent to the background model is the system of NICER event flags. There are six flags, with assigned values 1 or 0, which are interpreted as “yes” or “no,” respectively. Five of these flags tie an event to a particular circuit latch in the MPU, and the circuit is designed to help distinguish events as good or bad for inclusion in scientific analyses. The flags are: first event in packet (useful only for the data pipeline software), triggered the fast chain, triggered the slow chain, forced trigger, undershoot event, and overshoot event. Forced triggers result from commands to sample the signal values in the absence of a trigger from a detector, and they serve to monitor the zero-point in the energy calibration of each FPM/MPU signal-processing combination. Since launch, forced triggers have been operating at 5 Hz for each FPM. Undershoot events have latched a circuit designed to detect the large negative pulse that is uniquely caused by a detector reset. Undershoot events signify the high-amplitude negative pulse that is seen on the signal line when an FPM discharges the capacitor that collects ambient charge running through the detector (maintained at -55°C). Overshoot events have latched a different circuit designed to safeguard against large positive pulses (roughly equivalent to 18 keV) that could cause a bit rollover in the analog-to-digital converter. Overshoots are almost always caused by high-energy particles, since there is negligible effective area from the XTI above 18 keV. The keV assignments of undershoot and overshoot

Table 1
Parent Observations for the 3C50 Background Model

Target (1)	ObsID First (2)	ObsID Last (3)	# GTIs (4)	Exposure (ks) (5)
BKGD_RXTE1	1012010101	3012010106	357	131.2
BKGD_RXTE2	1012010201	3012020201	545	309.8
BKGD_RXTE3	1012010301	3012030104	317	157.2
BKGD_RXTE4	1012010401	2012040241	451	248.4
BKGD_RXTE5	1012010501	2012050232	540	292.4
BKGD_RXTE6	1012010601	3012060201	830	548.4
BKGD_RXTE8	1012010801	3012080102	516	336.9

Note. The sum of these background observations yields 3556 GTIs and 2.024 Ms exposure time.

events are thus meaningless. The strategy for recognizing events associated with energetic particles can be summarized as follows. Most particle events would be excluded as bad events via the overshoot flag, and detector edge-clipping events with in-band energy amplitude and no overshoot flag would be identified via high values in the event’s `PI_ratio`. Observations of the background fields reveal events that escape both of these particle event recognition methods, events that would be indistinguishable from a target’s X-ray events, and these observations provide the data to develop and test background prediction models such as 3C50.

The NICER data pipeline provides cleaned event lists for each observation to enhance the users’ extraction of spectra and light curves from the target source. The cleaned events are gain-calibrated and the photon energy is restricted to slow-chain events in the energy range 0.2–15 keV, which is slightly wider than the range that could be chosen for spectral fitting. In addition, the cleaned list is stripped of events with unity value in flags for forced triggers, undershoots, or overshoots. The background spectra produced by the 3C50 model match the format of the extractions from the cleaned event lists to facilitate the subsequent background subtraction efforts to be undertaken by the user.

The scope of the remaining sections of this paper is given as follows. Section 2 discusses the observations of the background fields and the criteria for data selection. Sections 3 and 4 describe the model parameters and the construction of the spectral libraries for observations in darkness and in sunlight, respectively. Section 5 examines model residuals for the background fields, where the expectation for net flux is zero. Section 6 provides practical considerations for running the model and for filtering results as a means of quality control. Section 7 shows light curves for calibration sources—the Crab Nebula and 1E0102–72. Section 8 shows a method for distinguishing rapid source flares from background variations at timescale 1 s. Section 9 discusses the successes and problems with the 3C50 model, and possible avenues for improvements. Finally, the model is summarized in Section 10.

2. Observations and Data Selection

The 3C50 model is a phenomenological approach to predicting the in-band (0.4–12.0 keV) background spectrum using the observations of the NICER background fields. The parent data set for the model libraries includes all such observations from 2017 July 24 through 2020 March 21. The contributions from each background field are given in Table 1.

The selection filters leading to the model libraries are summarized in Table 2, and they are described in detail, below.

Data analyses utilized the NASA HEASoft package, version 6.26.1. Prior software versions had different defaults for the use of `nimake` to define GTIs (see below). The raw event lists from the NICER pipeline for the observation IDs (ObsIDs) given in Table 1 were calibrated for the 2020 gain revision specified by “GCALFILE,” “nxtflightpi20170601v005.fits.” When making the cleaned event lists, the default filters for environmental conditions were adopted, excluding times in the South Atlantic Anomaly (SAA), pointing elevations within 15° of the dark Earth limb, and those within 30° of the sunlit Earth limb. However, the default filters for bad-event count rates were effectively disabled by using extremely high values (15,000) for the maximum rates of overshoots, undershoots, and the relationship between overshoots and magnetic cutoff rigidity. These filter overrides are required to populate the cleaned event lists with good events during times when the background rates are high.

The ObsID directories from the NICER pipeline contain data for a given target, accumulated on a given day. In this work, each GTI is an interval of continuous exposure, and for NICER such intervals are usually less than 2 ks because of interruptions imposed by the rotation of the ISS with respect to celestial coordinates, once per each 93 minute Earth orbit. Many ObsIDs contain more than one GTI. Since the NICER background can change significantly at different locations in the ISS orbit, the background modeling is based on the timescale of the GTIs, rather than the ObsIDs. There is further value to measuring the amount of parameter variability that occurs within a given GTI, so as to exclude it or to redefine the time boundaries to avoid strong flares in the background. Further practical considerations for running the background model are given in Section 6.

The tool `nimake` was used to define the GTIs for every ObsID in each of the background fields. This step repeats the same filter choices used to make the revised cleaned event lists (see above). GTI selection was additionally filtered to exclude GTIs with durations less than 60 s, while disregarding any gaps of 1 or 2 s that might be imposed by a telemetry packet loss, which are corrected for, via an adjusted exposure time, by the NICER pipeline. The numbers of selected GTIs and the net exposure time accumulated per background field are given in Table 1. The total yield is 3556 GTIs, averaging 570 s per GTI and accumulating 2.024 Ms.

Finally, we choose to build the background spectrum libraries with a selection of 50 (of 56) FPMs. We label the FPMs with two digits: the first for the MPU that services it (0–6) and the second for the FPM slot (0–7) in the MPU. For example, the first FPM on the first MPU is “00,” while the last FPM on the last MPU is “67.” In this notation, the six excluded FPMs are the four that are not operating (11, 20, 22, and 60) plus two (14 and 34) that have shown episodes of unreliable spectra and high noise rates, respectively. The selection of the remaining 50 FPMs adds an element of uniformity to the background libraries. However, users can conduct target analyses with any number of selected FPMs, and then apply the 3C50 model under the assumption that the model metrics do not change, detector-by-detector, and the net background spectrum can be simply scaled by the number of selected FPMs.

Table 2
Data Selections and Filters for the Spectrum Libraries

Selection (1)	# GTIs (2)	Exposure (ks) (3)	Comment (4)
All data	3556	2024.2	from Table 1
Selected 50 FPMs operating	3435	1937.4	
Filter noise outliers	3357	1891.1	any i : $nz_i > 100$ and $nz_i > 0.15 nz$
Parameters within 3C50 limits	3264	1818.3	see Figure 3
Subset ISS night	1991	1068.6	$nz < 200$, see Figure 4
Filter out outliers	1947	1038.2	see Figure 4; GTIs for night library
Subset ISS day	1273	758.6	$nz \geq 200$
Filter for high background and stage 1 residuals	1076	627.4	$ibg_{52} < 0.4 \text{ c s}^{-1}$; $-1.0 < (C_{\text{net}} + D_{\text{net}}) < 1.0$

Note. In summary, the 3C50 spectral libraries use 3023 (of 3556) GTIs, corresponding to 1665.6 ks or 82% of the exposure time. There are 1947 GTIs contributing to the night library and 1076 GTIs for the day library.

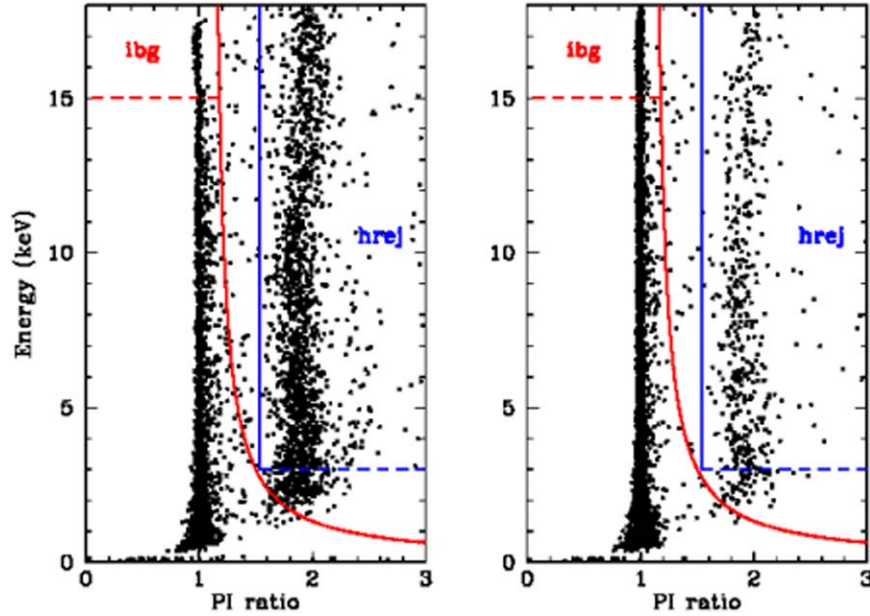


Figure 1. Sample background observations showing every good event in the plane of event energy vs. PI_ratio . The curved red line shows the NICER pipeline’s boundary between events that appear as if in focus (left of the line), vs. events originating far from the detector anode (right of the line). The integration areas for the first two background model parameters, ibg and $hrej$, are also shown. The observations were chosen to illustrate how the two background components are always present, but the fraction of events in each group can vary widely.

3. First Stage of the 3C50 Background Model

The strategy behind the 3C50 model is to sort out the background components from an empirical point of view, based on the event properties found in observations of the background fields. The first stage of the model distinguishes background components that have different spatial properties in the detector focal plane, as determined by event distributions that have different values of the PI_ratio . The second stage of the model deals with the soft X-ray excess due to noise encroachment induced by the seepage of sunlight into the instrument when observations are conducted in ISS sunlight, and this is described in Section 4.

3.1. Two Parameters for Background Components Sorted by PI_ratio

The description of the model parameter choices is framed by an examination of Figure 1, where events are plotted in the plane of (effective) photon energy (PI) versus PI_ratio , i.e., the ratio of slow-chain to fast-chain keV values for good events

that trigger both measuring chains. Each panel shows vertical spectral tracks centered on PI_ratio values near 1.0 and in the range 1.7–2.2. The curved red line shows the relationship $PI_ratio = 1.1 + 120.0/PI$, where PI is the slow-chain photon energy in units of 10 eV. This relationship is the standard cut in pipeline data processing, where events to the left of the line, plus events with $PI_ratio = INDEF$, are passed to the cleaned event lists. The value of 1.1 limits the exclusion of events to cases where the fast-chain energy differs by more than 10% from the slow-chain energy. The second term adds allowances for the broadening of the PI_ratio values due to statistical noise, especially in PI_FAST . We note that the events displayed in Figure 1 are drawn from the unfiltered but calibrated event lists (**ufa.evt*) in the pipeline’s *ObsID/xti/event_cl/* directory, which contains all of the good events (i.e., events with no undershoot, overshoot, or forced trigger flags), with no PI_ratio screening.

The two panels of Figure 1 utilize background observations that were chosen to illustrate how the same components are

always present in the plane of PI_{ratio} versus keV, while the fraction of events in each group can vary widely. The left panel combines data from Observation IDs and Mission Elapse Times (MET) as follows: 1012080149 (MET 132410965–132411942), 1012060207 (MET 152618492–152619765), and 1012020193 (MET 151521605–15152292). In these intervals, the majority of the events have high values of PI_{ratio} . A contrasting distribution is shown in the right panel, which consists of events from 1012010104 (MET 114014010–114014695), 1012080121 (MET 122750273–122751041), and 1012060219 (MET 154519499–154520301). The observations IDs correspond to background fields 8, 6, and 1 (left panel) and fields 8, 6, and 2 (right panel).

The pipeline’s PI_{ratio} filter effectively separates the two vertical distributions in background events seen in Figure 1. The right-side distribution is consistent with the expected particle hits near the edges of the Si drift detector that mimic good events with anomalously high values in PI_{ratio} . On the other hand, the left-side events are indistinguishable from “in-focus” X-rays emitted by NICER targets, and we refer to this distribution as the in-focus background component. This label is intended as a comparative reference rather than a provable statement that the XRC is involved in the process of detecting these events. Further considerations about the origin of these events are given below.

To capture and monitor the rate of events with a high PI_{ratio} , we define a “hatchet” rejection line, shown in Figure 1 as the blue vertical line at $\text{PI}_{\text{ratio}} = 1.54$. This leads to the choice of the first 3C50 model parameter, hrej , which is the count rate of hatchet-rejected events, $\text{PI}_{\text{ratio}} \geq 1.54$ in the range 3–18 keV. The high-energy cutoff represents an approximate maximum energy in the NICER calibration, while the lower limit (3 keV) avoids any overlap between the in-focus and hatchet-rejected distributions. The tailing off of rejected events below 3 keV represents the lower efficiency of the fast-chain triggers at lower energy, exacerbated by the charge diffusion pulse-broadening that further decreases the probability of detection. There is also laboratory evidence that the gain drops near the detector edges, affecting both measuring chains. However, the energy content of the particles is sometimes sufficient to endure all of these effects and trigger a pulse with a telltale high value in PI_{ratio} . In this sense, the NICER detector/electronics package has a built-in particle monitor, albeit with low efficiency.

Figure 1 shows two panels of background events, and each one represents an overlay from three GTIs, using different FPMs, selected at widely different times. The GTIs for the left panel were chosen for having roughly half of the events located in the hatchet-rejected region, while the three GTIs selected for the right panel have $\sim 10\%$ of events in the hatchet-rejected region. The intent is to illustrate common features in the energy: PI_{ratio} plane, while also showing that the fraction of events in each distribution can vary significantly. To monitor the rate of in-focus events from the background observations (i.e., the vertical track of events near $\text{PI}_{\text{ratio}} \sim 1$ in Figure 1), we define ibg as the count rate of in-focus events (i.e., left of the red curve) in the range of 15–18 keV. This restriction to high-energy events is needed to limit the ibg capture range to energies well above 12 keV, where the XRC optics have a negligible effective area, to avoid contamination of ibg by X-rays from bright sources.

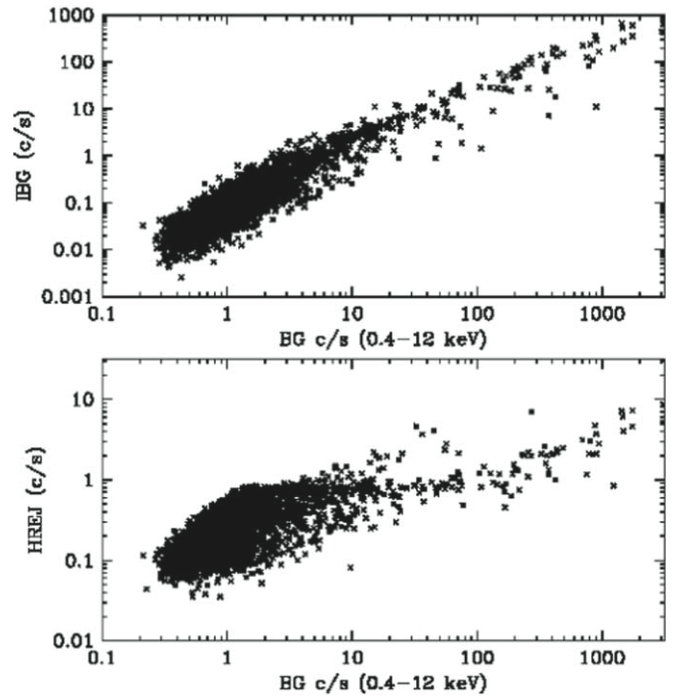


Figure 2. The in-focus background rate at 0.4–12 keV (R_{BG}) vs. the first two parameters of the 3C50 model, ibg (top panel) and hrej (bottom panel). The plotted data corresponds to all of the 3556 GTIs represented in Table 1. Rough correlations are seen in each panel, with a steeper dependence in the case of ibg .

Having defined ibg and hrej as the first two parameters for the 3C50 background model, Figure 2 shows how each parameter varies with the in-focus, in-band background count rate at 0.4–12 keV, hereafter R_{BG} . The lower limit of the energy range, i.e., choosing 0.4 keV instead of the XTI sensitivity limit at 0.2 keV, is a hedge against effects due to noise and optical light leak, which are described in the next section. The primary objective of the background model is to predict the spectrum associated with R_{BG} . Figure 2 shows that both ibg (top panel) and hrej (bottom panel) are roughly correlated with R_{BG} , with a steeper dependence in the case of ibg .

The high degree of variability in R_{BG} is apparent along the horizontal axes. The median value is $R_{\text{BG}} = 0.87 \text{ c s}^{-1}$, but the distribution, even while ignoring the 1% high and low extremes, still ranges from 0.33 to 300 c s^{-1} . Values of ibg can also vary by many orders of magnitude. Since ibg is the high-energy extension of R_{BG} , while the energy range is beyond the effective area of the NICER optics, ibg values can be used to normalize the stage 1 library selection in the 3C50 model, tuning the model to converge with the source spectrum at 15–18 keV (considered in extrapolation, since the extractions from cleaned event lists terminate at 15 keV under the default pipeline settings).

While hrej is a metric for the spatially extended events due to edge-clipping particles, the origin of the in-focus background component, tracked with ibg , is not well understood. Background components with a PI_{ratio} near unity are expected from the cosmic diffuse X-ray background, as well as possible soft X-ray emission from other sources (see Section 1). True X-ray events are expected to appear in focus, since the metal collimator above the detector surface limits the path of X-rays to radii within 1 mm ($3/17$) displacement from the anode (Prigozhin et al. 2016). However, the count rate from these

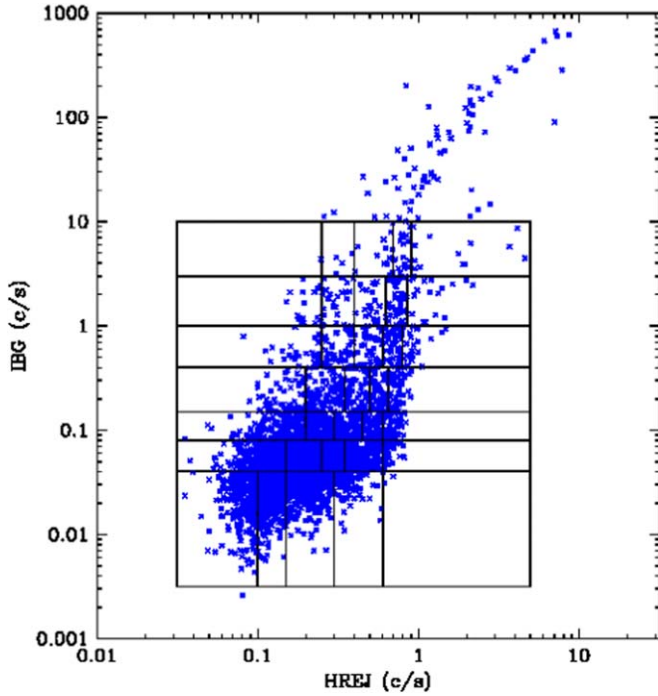


Figure 3. Plot of ibg vs. hrej . This plane is divided into five cells (horizontal) covering $\text{hrej} < 5.0$ and seven cells (vertical) with $\text{ibg} < 10.0$. The cells are referenced by $(\text{ibg}, \text{hrej})$ number with 11 on the lower left and 75 to the upper right. Cell boundaries in units of ibg , hrej are given in Table 3. These intervals define the bins in which the background spectra per GTI are combined to create the stage 1 background library for the 3C50 model. The stage 1 library contains 33 such spectra, with cells 15 and 71 unoccupied.

sources is expected to yield $R_{\text{BG}} \sim 0.5 \text{ c s}^{-1}$ over most of the sky, while the measured R_{BG} values are sometimes far brighter and highly variable. We are therefore led to view ibg as representing a second particle component that is either unable to penetrate the collimator or is guided to the detector with assistance from the XRC.

3.2. 3C50 Library for ISS Night

Figure 3 shows a plot of hrej versus ibg for all of the 3556 GTIs represented in Table 1. The hrej , ibg plane is the parent for stage 1 of the 3C50 model. A 5×7 grid of parameter values in this plane is used to bin the background spectra, per GTI, and the combined spectrum per grid cell is computed to populate the stage 1 library in the 3C50 model. The cell boundaries are chosen to follow the population pattern, rather than a regular grid. The library cells on the upper left and lower right of the grid are left vacant, and queries to those cells would select the nearest occupied neighbor, moving horizontally along hrej . Finally, the stage 1 library is restricted to $\text{ibg} < 10 \text{ c s}^{-1}$ for several reasons. This upper limit corresponds to $R_{\text{BG}} \sim 100 \text{ c s}^{-1}$, which is more than a factor of 100 times the median rate, and such conditions generally occur less than 1% of the time for targets other than the background fields and calibration sources. Furthermore, the GTIs with $\text{ibg} > 10 \text{ c s}^{-1}$ (Figure 3) show a higher tendency for variability within the GTI, creating additional strain on the goal for linearity within each library cell (see also Section 6.2, below).

The stage 1 model cells are labeled with two digits: ibg number (1–7) and hrej number (1–5). The layout is shown in Figure 3. The cell boundaries are given in Table 3, along with the average R_{BG} values, the cell accumulation time, and the

exposure-weighted normalization values in ibg_{lib} . Library selections must be standardized to a fixed number of FPMs, here chosen to be the maximum user choice of 52. However, users are free to select and specify, for any GTI, any number of FPMs (nfp_i), along with the model parameter measurements, ibg_i , hrej_i . The 3C50 model will then map a given GTI to a particular library cell, using values ($\text{ibg}_{52} = \text{ibg}_i * 52 / \text{nfp}_i$; $\text{hrej}_{52} = \text{hrej}_i * 52 / \text{nfp}_i$), and the matching library spectrum is presumed to have the correct spectral shape for that GTI. The matching library spectrum is then renormalized by a factor $\text{ibg}_i / \text{ibg}_{\text{lib}}$, and the result is the stage 1 prediction for the background spectrum in the 3C50 model.

Before proceeding to complete the stage 1 library spectra, we examine the distribution in R_{BG} that is found in each cell, as seen in the upper panel of Figure 4. Only the 1991 GTIs sorted for ISS night conditions (see Table 1) are included here. Cells are noted by bin value, $[\text{ibg}: i = 1 - 7, \text{hrej}: j = 1 - 5]$, and we use the cell number plus the fractional order of membership within a given cell to create an artificial horizontal axis that stretches out the data points simply for viewing purposes. The data from cell i, j begin at value $10 * i + 2 * (j - 1) + 1$, and the last GTI in the cell is plotted one unit later. For example, the GTIs in cell 11 are plotted in the range $11 < x < 12$, while cell 12 has the range $13 < x < 14$, and cell 75 has the range $79 < x < 80$. The larger gaps indicate the vacant cells, 15 and 71.

Even after filtering out the points above the dashed line in Figure 4, the variations in R_{BG} within each cell are much larger than the statistical uncertainties. This is shown in the lower panel of Figure 4, where the reduced chi square values (χ^2_ν) are shown for each cell, after the bright cases are removed, with the assumption that R_{BG} is constant, at the weighted average within each cell. For cell 11, where the lowest R_{BG} rates limit the statistical precision, we find $\chi_\nu = 8.7$, and χ^2_ν increases for the cells with higher rates (i.e., with better statistics). The rms variations in R_{BG} within each cell correspond to intrinsic fractional fluctuations in the 20%–40% range of the mean values. This result suggests that the ibg , hrej parameter scheme is far from a deterministic model, and additional background parameters are likely to be important. These results also motivate the strategy to normalize the library sections using the ibg_i value of a given source spectrum, to mitigate against the fractional errors of the cell parents that would be otherwise inherited. Thus, the 3C50 model assumes that the spectral shape per library spectrum is appropriate, while the normalization is fine-tuned to the target spectrum at hand.

We note that for R_{BG} or related quantities of central importance to background modeling, the exercise represented in Figure 4 can be conducted for any hypothetical set of model parameters. Intrinsic variances within cells and the progression of variances along each parameter axis can quantify how the cells organize the background measurements with minimal variance, and whether (from the measurement slopes within each cell) the background properties are divided into a grid with sufficient resolution.

The library spectra are made for each cell after filtering out the data points that are above the dashed line in the top panel of Figure 4, to avoid overweighting the results by such cases. There is a library spectrum for each cell, which is simply the sum of all the counts in the selected GTIs of that cell, divided by the total exposure time. Table 2 quantifies all of the

Table 3
Stage 1 Cell Boundaries and Normalizations

Cell (ibg, hrej) (1)	R_{BG} ($c s^{-1}$) (2)	GTI Time (ks) (3)	ibg ₅₂ Start ($c s^{-1}$) (4)	ibg ₅₂ End ($c s^{-1}$) (5)	hrej ₅₂ Start ($c s^{-1}$) (6)	hrej ₅₂ End ($c s^{-1}$) (7)	ibg _{lib} norm. ($c s^{-1}$) (8)
11	0.446	44.3	0.00	0.04	0.00	0.10	0.0218
12	0.493	149.4	0.00	0.04	0.10	0.15	0.0265
13	0.596	100.3	0.00	0.04	0.15	0.30	0.0285
14	0.775	28.5	0.00	0.04	0.30	0.60	0.0317
15	...	...	0.00	0.04	0.60	10.0	...
21	0.689	95.7	0.04	0.08	0.00	0.15	0.0503
22	0.801	109.8	0.04	0.08	0.15	0.25	0.0517
23	0.855	71.4	0.04	0.08	0.25	0.35	0.0527
24	1.008	67.5	0.04	0.08	0.35	0.60	0.0567
25	1.242	8.0	0.04	0.08	0.60	10.0	0.0583
31	1.103	41.0	0.08	0.15	0.00	0.20	0.1006
32	0.955	29.5	0.08	0.15	0.20	0.30	0.0987
33	1.322	43.9	0.08	0.15	0.30	0.45	0.1013
34	1.302	20.4	0.08	0.15	0.45	0.60	0.1012
35	1.771	16.5	0.08	0.15	0.60	10.0	0.1088
41	2.020	18.6	0.15	0.40	0.00	0.20	0.2047
42	1.724	27.2	0.15	0.40	0.20	0.35	0.2370
43	1.904	15.6	0.15	0.40	0.35	0.50	0.2128
44	2.414	20.6	0.15	0.40	0.50	0.65	0.2314
45	2.673	16.9	0.15	0.40	0.65	10.0	0.2318
51	2.683	8.6	0.40	1.00	0.00	0.25	0.5002
52	2.763	20.1	0.40	1.00	0.25	0.40	0.5659
53	3.293	10.0	0.40	1.00	0.40	0.60	0.6039
54	4.446	7.5	0.40	1.00	0.60	0.80	0.7110
55	4.288	9.3	0.40	1.00	0.80	10.0	0.5986
61	4.527	7.3	1.00	3.00	0.00	0.25	1.4975
62	6.080	13.0	1.00	3.00	0.25	0.40	1.7515
63	6.644	4.9	1.00	3.00	0.40	0.63	2.0238
64	7.604	11.4	1.00	3.00	0.63	0.85	1.6836
65	11.644	7.7	1.00	3.00	0.85	10.0	2.1318
71	...	...	3.00	10.00	0.00	0.25	...
72	9.944	2.5	3.00	10.00	0.25	0.40	3.8110
73	16.477	3.4	3.00	10.00	0.40	0.70	5.2374
74	15.311	2.9	3.00	10.00	0.70	0.90	3.6876
75	17.046	4.4	3.00	10.00	0.90	10.0	4.3715

selection and filtering steps used to construct the 33 spectra for the stage 1 library, using 1947 selected GTIs (of 1991 during ISS night) that are below the dashed line in Figure 4. The day and night assignments are made on the basis of the nz value, as explained in the next section.

The stage 1 library spectra are displayed in Figure 5, using a rebinning scheme that over-samples the XTI resolution uniformly by a factor of ~ 3 . In the range 0.2–15 keV, the number of combined PI bins is 3 (0.2–2.48 keV), 4 (2.48–6.00 keV), 5 (6.0–12.0 keV), and 6 (above 12 keV). As expected, the spectra are substantially brighter and flatter with increasing ibg cell number (vertical steps), and there is an appearance of the Si $K\alpha$ emission line (1.74 keV) in the highest two levels of ibg. With increasing hrej (i.e., from left to right), there is a more shallow increase in continuum brightness, a stronger soft component at 0.2–3 keV, and increasing emission lines indicating fluorescence at 7.47 keV (Ni $K\alpha$), 9.71 keV (Au $L\alpha$), and 11.44 keV (Au $L\beta$). These changing spectral features over the surface of the ibg, hrej plane offer some validation for the utility of choosing those model parameters. The background component tied to ibg is the main source of the R_{BG} count rate, while the hrej parameter, despite its exclusion from R_{BG} via PI_ratio filtering in the NICER pipeline, signals

systematic changes in the spectrum of R_{BG} for both the continuum shapes and the characteristics of emission lines.

4. Second Stage of the 3C50 Background Model

The second stage of the 3C50 background model is required to subtract an independent soft X-ray component tied to observations during ISS daytime. The noise in the slow chain ($\sim 3 c s^{-1}$ per FPM; see Section 1) always creates a spectral component that is centered near 0.1 keV, and it is usually invisible at 0.3 keV. However, it was recognized soon after launch that all of the NICER FPMs exhibit systematically higher levels of this low-energy noise when the XTI is illuminated by sunlight during the course of the ISS orbit (e.g., Bogdanov et al. 2019). Optical photons cannot trigger events in the MPUs, as such, but they liberate Si electrons, causing a number of secondary effects. The increased detector current, in the presence of optical light, elevates the undershoot rate (i.e., the detector reset rates) and also causes modest changes in detector gain and spectral resolution. The pipeline’s gain calibration makes corrections for such effects, while the changes in spectral resolution are also predictable, again allowing appropriate corrections for science investigations. However, the spectral broadening of the low-energy electronic

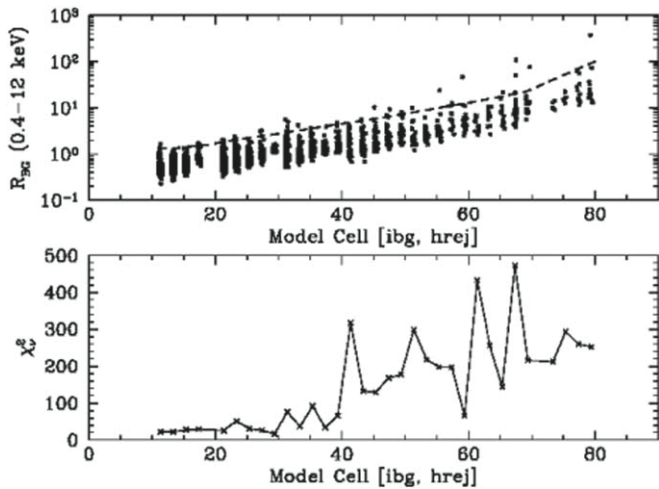


Figure 4. Top panel: values of R_{BG} , the in-focus and in-band count rate for the background GTIs, are plotted for each successive cell of the stage 1 library. The cells are counted by their *ibg*, *hrej* number, ranging from 11 to 75. The cell id numbers and their ordinal number within the cell are used to stretch out the points along the *x*-axis (see the text). The dashed line shows the high-rate filter that clips the brightest outliers to be excluded when computing the library spectra. Bottom panel: reduced chi square values for each cell, after filtering out the high points (top panel), assuming that R_{BG} is constant within each cell. The levels of intrinsic scatter in each cell are 20%–40%, motivating a step to normalize each library selection using the *ibg*, value of each target spectrum.

noise can increasingly intrude above 0.2 keV as the optical load becomes more intense. This is illustrated below in Section 4.2. Thus, the tail of the low-energy noise distribution can encroach on a portion of the in-band source spectrum during ISS daytime, making it necessary to include a quantification of this effect in the background model. There is no expected or measured correlation between excess noise and either *ibg* or *hrej*, and so we treat the daytime soft excess as an independent spectral component to be handled in a second stage of the 3C50 model. When a GTI occurs during ISS daytime, the derived spectra from model stages 1 and 2 are simply added together to form the predicted background spectrum.

While n_z and the undershoot rates per FPM do rise significantly in the presence of sunlight, geometric modeling of these effects is impeded by the complex dependencies on the Sun angles and the angles to many reflective structures on the ISS, including the large and rotatable solar panels. To first order, the FPMs respond in unison to the optical light leak, at the GTI timescale used in this paper, except for the times when one FPM goes off track to a higher noise rate, as if hypersensitized for some time interval. The prime case for this behavior is FPM 34, which was excluded from the 3C50 libraries. There are other cases with occasionally anomalous (6σ or higher) values in n_z , compared to the rates for the others, notably FPMs 10, 13, 43, and 47, and this is what prompts the screening for n_z outliers, per GTI, as accounted in Table 2. Note that a light leak reaching an FPM is vastly fainter than direct sunlight, and it is difficult to determine whether the outlier FPMs have an increased susceptibility to relatively high levels of optical light, versus being subject to geometric light leak pathways inside the XTI.

4.1. Third Model Parameter for Soft X-Ray Excess during ISS Daytime

The need to correct for low-energy noise in spectra obtained during ISS daytime is apparent in the residuals found after applying stage 1 (only) of the 3C50 model to the background spectra. Figure 6 shows the stage 1 residuals in two energy bands: 0.4–12 and 0.3–0.4 keV. These model residuals should display an average value of zero, in all bands, if the background model has completed its job. GTIs during ISS night are plotted with a black cross, and GTIs during ISS day are plotted with a red triangle. The stage 1 residuals are plotted versus R_{BG} , i.e., the original background rate at 0.4–12 keV. The horizontal axis is truncated at $R_{BG} = 3.0$, covering 89.8% of GTIs, for a better view of the details. Residuals in the range 0.4–12 keV (top panel) are fairly well contained during both day and night. But in the range 0.3–0.4 keV (bottom panel), a region in soft X-rays that is valuable to NICER science investigations, the GTIs during ISS daytime show the encroachment of noise. The vertical axis scale is chosen to be identical in both panels to highlight the significance of the problem.

To incorporate the effect of the soft X-ray excess during ISS daytime, we choose to monitor the entire low-energy noise component for the 50 FPMs during each GTI. The third parameter for the 3C50 background model, n_z , is defined as the total count rate in the slow chain in the range 0.0–0.25 keV. The GTIs chosen for the stage 1 (nighttime) spectral library show a primary distribution peak in n_z at $156 \pm 5 \text{ c s}^{-1}$, consistent with the $\sim 3 \text{ c s}^{-1} \text{ chain}^{-1}$ target for trigger threshold settings for 50 FPMs (see Section 1). During the analyses leading to the 3C50 model, it was determined that the effects of optical light become measurable at 0.3–0.4 keV only when $n_z > 200 \text{ c s}^{-1}$. This value was then used to distinguish day and night categories for the background GTIs in the 3C50 model. The quantifications of the relationships between n_z and the soft excess in various energy bands are included in Table 4.

4.2. 3C50 Library for ISS Daytime

An empirical strategy is adopted to model the soft excess during ISS daytime with an additional one-dimensional set of spectra that comprise the stage 2 library in the 3C50 model. The stage 2 library captures the mean soft-excess spectra left behind by stage 1 of the background model, using 12 steps in values of n_z , as given in Table 4. This library strategy appears to offer better performance, compared to alternative efforts to fit the noise component with a function with broad wings, e.g., a Lorentzian or modified Gaussian.

Starting with 1273 daytime GTIs, we apply two filters before combining the spectra within the designated levels in n_z (see Table 4). Both filters are intended to reduce the systematic problems that would be inherited by the stage 2 library. The first filter limits the input GTIs to moderate count rates, using $\text{ibg}_{52} < 0.4 \text{ c s}^{-1}$, which corresponds to the first four *ibg* levels (of seven) in stage 1. The second filter excludes cases in which the stage 1 residuals at 2–12 keV (i.e., away from the soft excess) are outside the range $\pm 1.0 \text{ c s}^{-1}$. This condition screens out the GTIs with stage 1 background spectra that deviate from the predicted one. These filters exclude 137 and 60 GTIs, respectively. The parent spectra for the stage 2 library then consists of 1076 GTIs, amounting to 83% of the total daytime exposure (see Table 2), divided into 12 n_{z52} levels, as defined in Table 4.

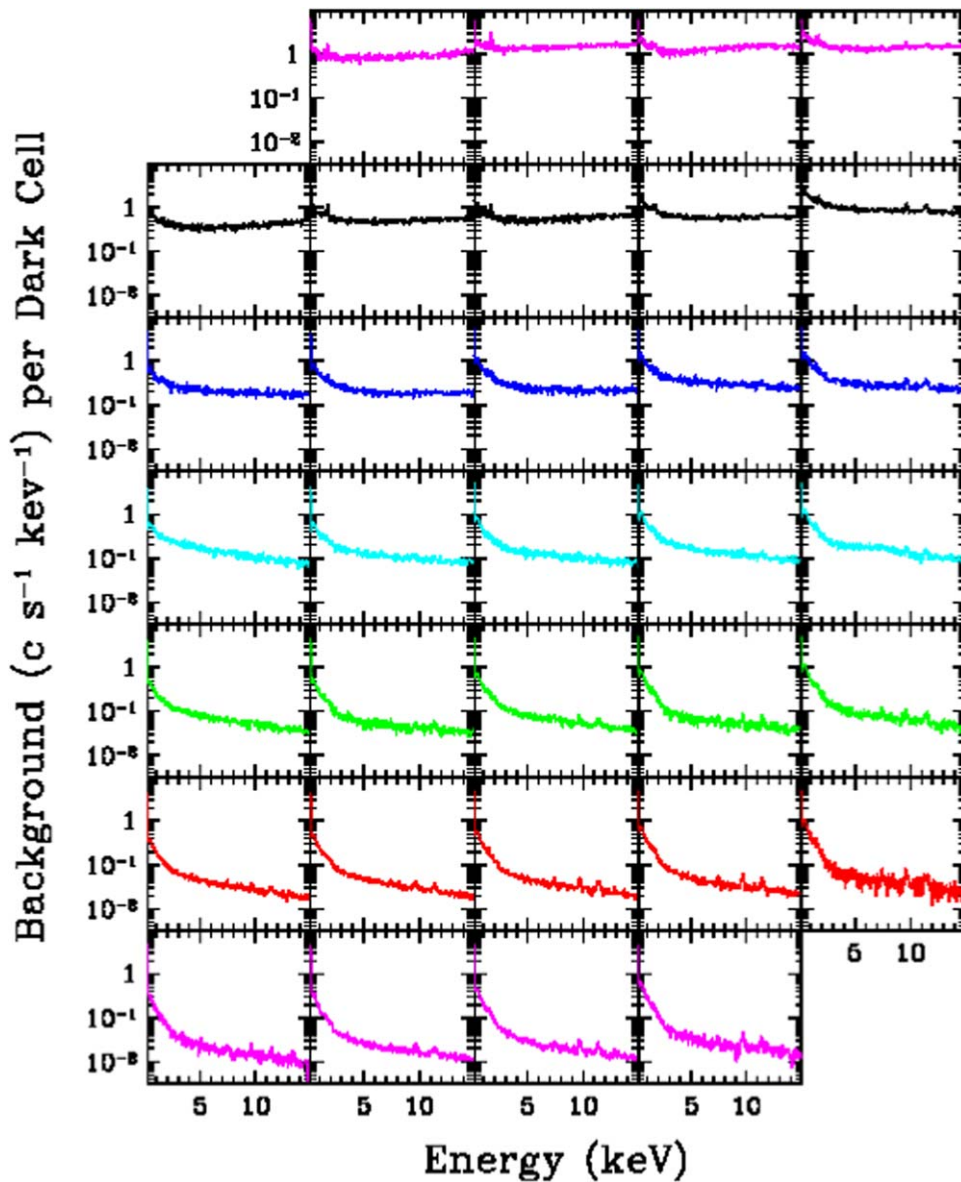


Figure 5. The stage 1 library for the 3C50 background model consists of these 33 spectra, arranged with ibg number (1–7) increasing vertically and hrej number (1–5) increasing to the right. The spectra are substantially brighter and flatter with increasing ibg cell number (vertical steps), while a more shallow brightness increase, a stronger soft component at 0.2–4 keV, and brighter emission lines are seen with increased hrej cell number (from left to right). Cell 11 (lower left) is the closest that the library comes to an isolation of the cosmic diffuse X-ray background.

The stage 2 model spectra are shown in Figure 7. The amplitude and extent in keV of the soft excess increases with the nz_{52} level, as expected. These spectra have been smoothed with a 15 PI bin “boxcar” at energies above 0.4 keV. Above 0.7 keV, the continuum levels are very faint, even at the highest nz rates (i.e., levels 10–12), where the integrated count rates correspond to an addition of less than 0.06 c s^{-1} to the count rate (i.e., 0.7–12 keV) during ISS daytime.

The normalization scheme for the stage 2 library selection follows the practices used for the stage 1 library. For any spectrum (GTI_i), a user can select and specify any number of FPMs (nfp_{m_i}), along with the noise level (nz_i) that is consistent with nfp_{m_i} . The stage 2 library intervals are again based on a standard 52 FPM scale (Table 4), so that a library selection is based on the value $nz_{52} = nz_i * 52 / \text{nfp}_{m_i}$. If nz_{52} corresponds to level j of the stage 2 library, then that spectrum is selected and

renormalized by a factor, nz_i / nz_j , and then added to the background prediction in stage 2 of the 3C50 model.

Table 4 also specifies the count rates of library spectra integrated for the softest NICER bands, S0 (0.2–0.3 keV), S1 (0.3–0.4 keV), and A (0.4–1.0 keV). The table quantifies the soft excess that would be suffered if stage 2 of the model were ignored. Stage 2 is therefore a required part of the 3C50 model for studies of faint and soft X-ray sources, including the rotation-powered pulsars that are the prime targets for NICER, since their count rates are often $< 1 \text{ c s}^{-1}$.

Table 4 can further help to evaluate residual count rates in background-subtracted spectra for science targets. We retain our conventions for labeling NICER energy bands, and we use the subscript “net” to indicate count rate queries applied to background-subtracted spectra. The manner in which noise events leak into the different energy bands (Table 4) gives some guidance as to how to use $S0_{\text{net}}$ as a quality metric to

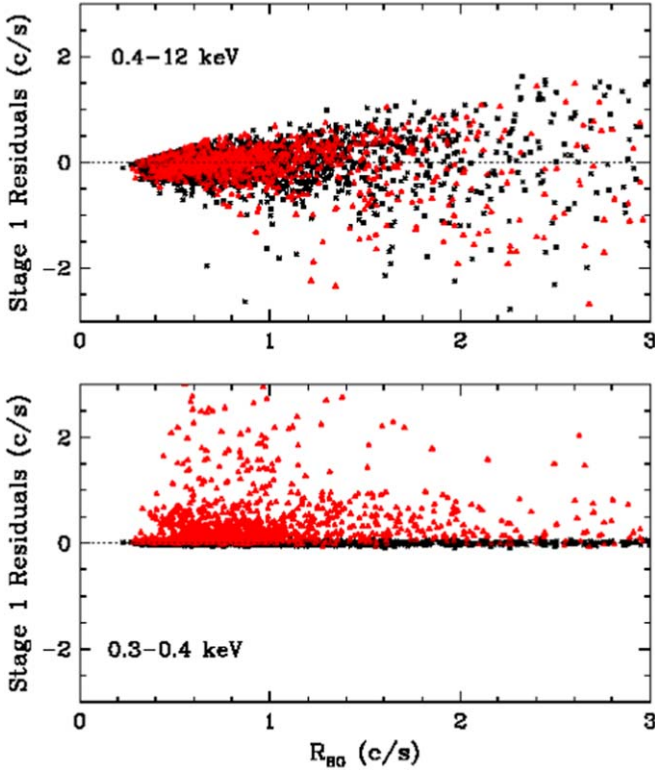


Figure 6. The 3C50 model was applied to GTIs of the background observations using only the stage 1 library. The residuals from this exercise are shown at 0.4–12 keV (top panel) and 0.3–0.4 keV (bottom panel). The plot symbols distinguish GTIs during ISS night (black cross) and ISS day (red triangle). The positive residuals during ISS daytime are associated with the seepage of sunlight into the XTI, which expands the ~ 0.1 keV noise component to the point that it can encroach into the soft X-ray region that is valuable for NICER science.

Table 4
Stage 2 Levels and Quantified Soft Excess

Level	Min. $n_{z_{52}}$ c s^{-1}	Max. $n_{z_{52}}$ c s^{-1}	Normalized $n_{z_{lib}}$ c s^{-1}	S0 c s^{-1}	S1 c s^{-1}	A band c s^{-1}
(1)	(2)	(3)	(4)	(5)	(6)	(7)
01	200	215	202.81	0.159	0.011	−0.010
02	215	250	228.52	0.264	0.022	−0.004
03	250	300	277.36	0.491	0.037	0.001
04	300	400	346.98	2.806	0.120	0.030
05	400	500	457.51	2.091	0.141	0.007
06	500	600	547.25	2.922	0.184	0.037
07	600	750	683.16	5.032	0.307	0.019
08	750	900	810.45	7.937	0.408	0.026
09	900	1100	1003.81	12.466	0.632	0.061
10	1100	1300	1198.35	31.528	1.230	0.114
11	1300	1600	1393.30	60.267	1.901	0.132
12	1600	0	1796.66	80.318	2.493	0.173

Note. The last three columns give the soft excess rates per 50 FPMs, and the energy bands are S0: 0.2–0.3 keV; S1: 0.3–0.4 keV; and A band: 0.4–1.0 keV.

estimate, per GTI, the likely level of residual contamination that may be present in $S1_{\text{net}}$, which we want to preserve for science. This is relevant because systematic errors in the background model can leave behind many more counts in $S0_{\text{net}}$, compared to the X-ray brightness of the target, since the effective area of the NICER XTI is very low in S0 and this

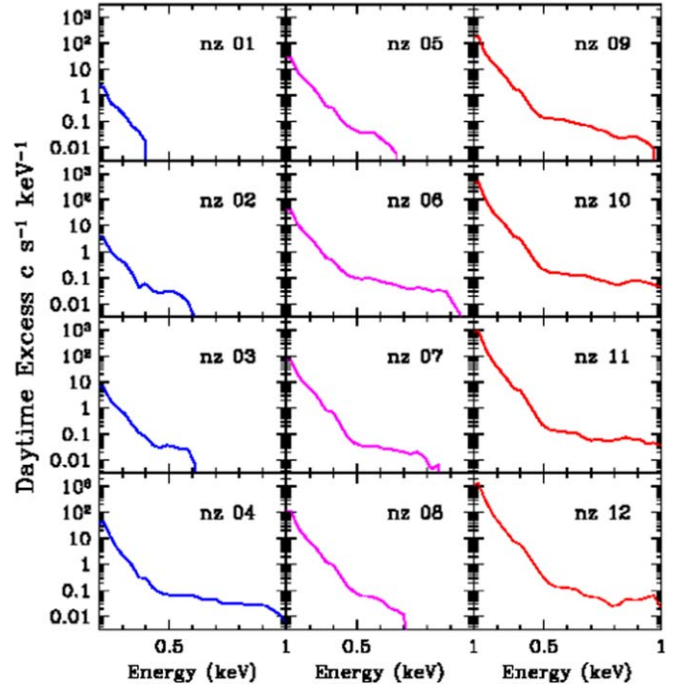


Figure 7. The stage 2 library for the 3C50 background model consists of 12 spectra that represent the soft excess caused by various levels of sunlight intrusion into the XTI. These spectra were determined from the residuals of the stage 1 model for 1076 GTIs during ISS day. The 12 spectra correspond to different levels in the $n_{z_{52}}$ parameter, which closely tracks the amount of optical light leak.

band is strongly attenuated by absorption in the interstellar medium. Depending on the brightness and softness of a given X-ray target, such filtering steps using $S0_{\text{net}}$ can, for example, inform users as to whether the soft X-ray light curve (e.g., 0.3–2.0 keV) is free from contaminated GTIs, or when spectral fitting down to 0.3 keV is likely to be safe. We offer specific recommendations and examples for filtering results via the background-subtracted spectra, in part using $S0_{\text{net}}$, in Sections 6 and 7, below.

5. Model Evaluation

To close the loop on the background observations, we apply the full 3C50 model to all of the 3477 background GTIs that have parameter values within the model limits. Residual count rates are shown versus R_{BG} in Figure 8. The top panel displays the residuals in the in-band energy range (R_{BGnet}), and the bottom panel shows the residuals at 0.3–0.4 keV ($S1_{\text{net}}$). The night/day observations are distinguished with a black cross/red triangle, respectively. What can these residual rates tell us about systematic uncertainty when applying the 3C50 background model? Considering first the in-band residuals (top panel), the 3C50 model is shown to be most effective when the background count rates are low. The need to bifurcate the evaluation into high and low count rates is tied to the pattern of points in the top panel. When $R_{\text{BG}} < 2 \text{ c s}^{-1}$, which corresponds to 82% of all background GTIs, then R_{BGnet} has rms value 0.33 c s^{-1} . In Sections 6 and 7 below, we show that the quality of these results can be improved by quality filtering the background-subtracted spectra in off-target energy bands. In Figure 8, the 3C50 model residuals above 2 c s^{-1} become more random, losing the population near zero, in marked contrast with the bottom panel. We interpret this as evidence

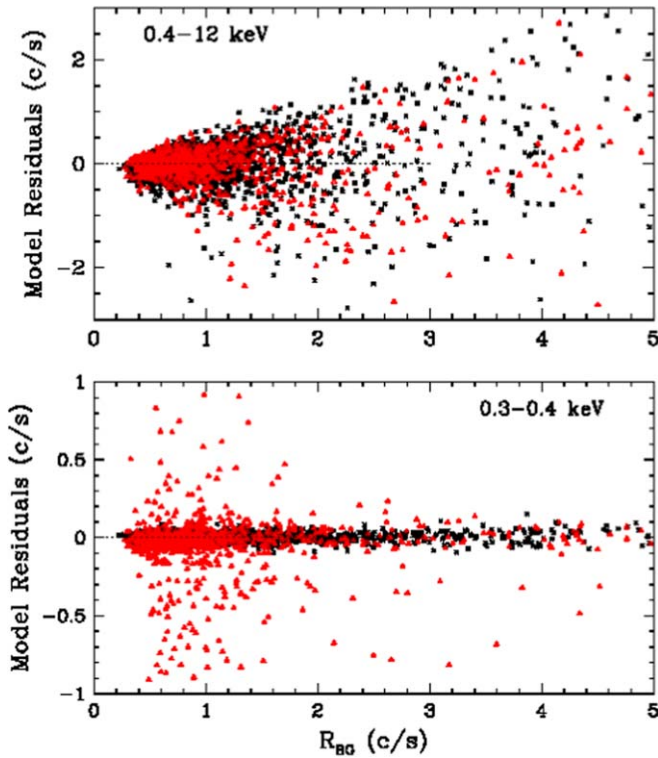


Figure 8. The final residuals from the application of the 3C50 model to the background observations, plotted vs. the original count rate at 0.4–12 keV, R_{BG} . The horizontal axis is truncated at 5 c s^{-1} for clarity; only 6% of the observations exceed this level. The model residuals are shown at 0.4–12 keV (R_{BGnet} ; top panel) and 0.3–0.4 keV ($S1_{net}$; bottom panel). The plot symbols distinguish GTIs during ISS night (black cross) and ISS day (red triangle).

that an additional model parameter, which has not been identified, has first-order significance when the background rate is high. Below, we also provide methods to identify and filter out some of the GTIs associated with high background rates, to protect the integrity of NICER science while further studies of the NICER background go forward.

The bottom panel of Figure 8 shows very small residuals in $S1_{net}$ during ISS nighttime (black crosses), but the daytime GTIs have higher residuals, particularly when the nz rates are high. The high residuals in $S1_{net}$ are matched with much higher residuals in $S0_{net}$, and this provides a method to use $S0_{net}$ to safeguard $S1_{net}$ or A_{net} , taking guidance from the relationship between the noise leaks in $S0$, $S1$, and the A-band given in Table 4. Users can screen for problematic GTIs by choosing a maximum tolerable light leak in $S1_{net}$ or A_{net} for their science, finding that level in Table 4, $S1$ or A , and then use the corresponding $S0$ value as the filter criterion for $S0_{net}$ to exclude GTIs that are likely to exceed the chosen noise limit. Systematic differences within any nz bin in the stage 2 library will leave residuals of either sign in $S0_{net}$, $S1_{net}$, and A_{net} , and filtering efforts should mirror the rejection criteria accordingly.

The origin of GTIs with high background model residuals is revealed in Figure 9, where different intervals in R_{BGnet} (see the figure caption) are plotted on a grid of Earth longitude and latitude, using the orbit location at the midpoint of each GTI. It is clear that the GTIs with the largest residuals coincide with the polar horns in the ISS orbit. The first two intervals (i.e., the residuals within $\pm 1.0 \text{ c s}^{-1}$) account for 90% of all GTIs, while the intervals with the largest residuals (1.2% of the total and $|R_{BGnet}| > 5.0 \text{ c s}^{-1}$) are largely (86%) confined to high

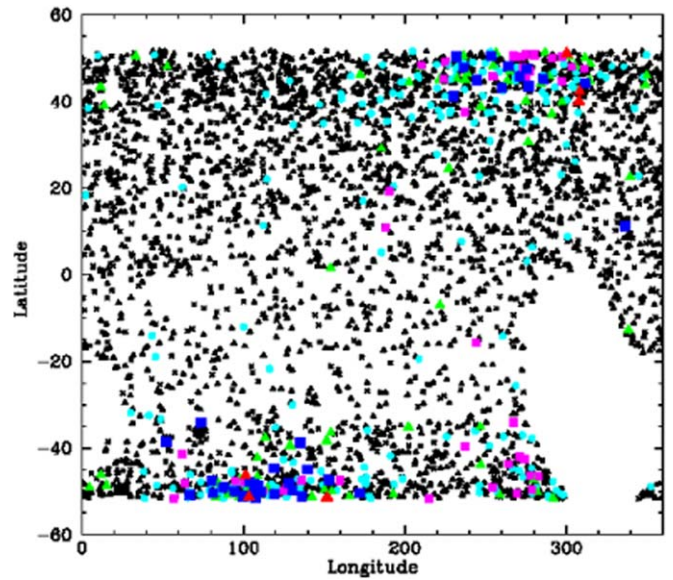


Figure 9. A map of the background observations in Earth longitude and latitude, with the magnitudes of the 3C50 model residuals, R_{BGnet} , represented with the symbol types. The c s^{-1} ranges and symbols are: ± 0.1 (black “x”s); ± 1.0 , excluding the previous level (small black triangles); ± 2.0 , excluding previous levels (cyan circles); -5.0 to -2.0 (medium magenta squares); 2.0 to 5.0 (medium green triangles); < -5.0 (large red triangles); and > 5.0 (large blue squares). The points terminate in latitude at the inclination of the ISS orbit (52 deg). The highest residual count rates from the 3C50 model reside in the polar horns in the ISS orbit. The region void of points in the lower right quadrant is the exclusion zone for the SAA, and there are no indications of problems near the chosen SAA boundaries.

latitude: $|\text{lat}| > 42.0 \text{ deg}$. On the other hand, the NICER exclusion zone for the SAA (the southern area void of points) appears to effectively exclude any background-related problems. This exclusion zone coincides with the keyword `NICER_SAA = 1` in the pipeline’s information files.

6. Practical Considerations for the 3C50 Background Model

6.1. Implementing the 3C50 Background Model

The prototype ftool task “`nibackgen3C50`” is the recommended tool for running the 3C50 background model. Although not yet formally part of the HEASoft NICER software suite, it is compliant with NASA HEASARC standards and is supported by the NICER Guest Observer Facility. There are two time intervals that are important in the use of `nibackgen3C50`. The first is the particular investigation interval for which a predicted background spectrum is desired. This is driven by the input given to the tool, which is nominally an ObsID, i.e., the reference number and top-level directory for a daily accumulation of GTIs for a single target. Users can obtain background spectra on timescales shorter than an ObsID (which generally contains multiple GTIs) by alternatively inputting the combination of a single unfiltered event file and GTI file. In either case, one background spectrum is output per each call to `nibackgen3C50`. Thus, users who plan to investigate target spectra on timescales shorter than one day would sort the unfiltered event files (i.e., the pipeline’s `$ObsID/xti/event_cl/ni*ufa.evt.gz` files) into a series of smaller files with the intended time boundaries, and then run `nibackgen3C50` sequentially on these files along with their associated GTI files.

The second important time interval is internal to `nibackgen3C50`, which generally computes the background in subintervals and then provides the exposure-weighted results in the output file. The subintervals may be the GTIs within the input file, or a shorter timescale directed by the user. Subintervals are never allowed to cross GTI boundaries, since the gaps between GTIs can be many hours, with a likelihood of different background conditions on either side of the gap. The background rate systematically varies with the ISS location in its 93 minute orbit. As noted above, NICER GTIs are seldom as long 2 ks, and a typical monitoring program has an average GTI ~ 600 s, or 11% of the orbit. As fractions of the ISS orbit, GTIs are an acceptable choice for `nibackgen3C50` subintervals. Shorter intervals are also acceptable, but below ~ 100 s, Poisson noise in `ibg`, the parameter that normalizes the nighttime library selections, can be a concern.

Another option of `nibackgen3C50` is the ability to control the FPM selections by listing the ones to ignore. Then, for each subinterval, `nibackgen3C50` reads the event lists and calculates the average count rates in `ibg`, `hrej`, and `nz` for the selected FPMs, scales these rates to the level of 52 FPMs to make library selections (Tables 3 and 4), normalizes the nighttime library selection by `ibg/inglib` (Table 3), and then, if $nz_{52} > 200$, adds the selected daytime library spectrum, normalized by nz/nz_{lib} (Table 4). Then, as noted above, the subinterval spectra are exposure-weighted to produce the modeled background spectrum. Further information about the control of time intervals and other command parameters for `nibackgen3C50` is available via the HEASARC NICER tools website,¹¹ from where it may be downloaded and locally installed.

Additional implementation notes are as follows. The 3C50 model enforces a minimum `ibg52` value of 0.016 c s^{-1} , to avoid cases where short intervals and low `ibg` rates would lead to occasions where `ibgi` is zero, and the background prediction would also be zero, since `ibgi/ibglib` is the normalization factor for the stage 1 contribution to the predicted background. This lower limit on `ibg52` was estimated from the distributions in `RBG` and `ibg52` for the 3556 GTIs examined in this study. Finally, the current version of the `nibackgen3C50` tool described above allows for the application of the `hbgnet` quality check, with the user selecting the maximum allowed absolute value. The `S0net` check is not implemented in the current release, but will be in subsequent releases.

The NICER data archive contains event files that were calibrated with a series of gain solutions, which can be identified by the keyword “GCALFILE.” Users are recommended to bring their data sets to a uniform calibration level with the “`nicerl2`” tool, using either the calibration used in this paper, `nixtflightpi20170601v005.fits` (2020), or the more recent one, `nixtflightpi20170601v006.fits` (2021), for which there are no differences below 12 keV. The `nibackgen3C50` tool attempts to match the GCALFILE in the input event lists with an appropriate set of library spectra, and 3C50 model libraries have been prepared for two previous gain calibrations, governed by the GCALFILES `nixtflightpi20170601v002.fits` (2018) and `nixtflightpi20170601v004.fits` (2019). By default, if a match cannot be made, then the user is warned, but a background spectrum is produced using the most recent model library.

Tables 3 and 4 and the details in this section provide sufficient information (e.g., cell boundaries, library normalizations, minimum `ibg52` values, and GCALFILE issues) for users who wish to download the model library files, which are a simple directory of standard “pha” files, and script their own implementation of the 3C50 background model. The calculation steps would be the same as those given above in the functional outline of `nibackgen3C50`, and the scripts would need to include commands to sort out the events corresponding to the definitions of `ibg`, `hrej`, and `nz`, e.g., using “fselect” on the unfiltered and calibrated event lists, given choices of time intervals and lists of FPMs (“DET_ID”) to exclude. The conversion of the model parameter event lists to event rates would complete the assembly of the required 3C50 parameters— N_{FPM} , `ibg`, and `nz`—for each time interval needing a background prediction.

6.2. Variations in `ibg`, `hrej`, and `nz` within a GTI

The selection of GTIs for populating the model libraries did not have an explicit screening step for variations in the values of the model parameters within a GTI (Table 2). We address this issue here. To evaluate the 3C50 parameter values for each GTI interval, we extracted both the spectra and light curves (1 s bins) for each parameter. The light curves were routinely used to calculate the mean (μ), standard deviation (σ), and variance in excess of Poisson statistics ($\sigma_{\text{int}}^2 = \sigma^2 - \mu$) for each parameter. Trends with brightness were investigated, and we explored several filtering criteria and their ramifications.

For `ibg` and `hrej`, the low values of μ make it inappropriate to use the fraction of intrinsic deviations (i.e., $\text{sqrt}(\sigma_{\text{int}}^2/\mu)$) as a screening tool. Instead, an ad hoc relationship was favored, with a rejection criterion: $\sigma_{\text{int}} > 1.0 + 0.2 \times \mu$. At the lowest count rates (see Figure 2), the intrinsic deviations must be above 1 c s^{-1} to prompt rejection, while at the highest rates (i.e., 10 c s^{-1}), the intrinsic deviation must be 20% of the mean rate. Considering all of the GTIs with average parameter values within the 3C50 model limits, `hrej` variations fail this test in only 0.8% of the intervals, while the `ibg` rejection rate is 5.4%, and this group includes all of the `hrej` failures. Furthermore, all but 50 of the `ibg` failures were already rejected for library use by the other criteria listed in Table 2. These remaining 50 cases are all among the 183 GTIs in the brightest three levels of the night library, i.e., in cells 51–75. Rather than exclude these cases, we concluded that variability in `ibg` is another characteristic of the high background conditions that do not fall in line with the 3C50 model.

Variations in `nz` are an entirely different matter. The count rates are high, and systematic differences in the responses of individual FPMs to optical light leak may occur during ISS daytime. The effects of the light leak on the X-ray spectrum are not the same, at a given FPM-integrated count rate in `nz`, if the distribution is skewed toward one FPM rather than being more evenly distributed. The most striking example is FPM #34, which was eliminated from this study because of its frequent extreme response to ISS daytime. In practice, it was found that the disparity in FPM noise rates was a more important issue than the changes in `nz` within a given GTI. Different rejection criteria were investigated, using 3C50 model residuals in the *S1* and *A* bands as the metrics for quality assessment. We adopted the criterion that a GTI would be excluded from consideration (during ISS daytime) if FPM_{*i*} with the highest noise rate yields a GTI average rate $nz_i > 100$, while FPM_{*i*} also contributes more

¹¹ https://heasarc.gsfc.nasa.gov/docs/nicer/tools/nicer_bkg_est_tools.html

than 15% of the total noise counts. This step is included in the data selection outline given in Table 2. GTIs that fail this test were rejected from further consideration, in order to maintain the strategy of using the same 50 FPMs to build the model libraries. For general investigations with NICER, users could alternatively choose to excluding the offending FPM_i and recompute the extractions from the event lists for that GTI with a reduced set of selected FPMs.

6.3. Filtering Steps after Background Subtraction to Improve Data Quality

To deal with systematic errors in the background subtraction process, the strategy was introduced (in Section 5) of filtering out results on the basis of residual count rates in spectral bands that are not needed for science. The energy range for any spectra extracted from cleaned event lists is 0.2–15.0 keV, and in the context of this background investigation, this can be seen as $S0 + S1 + A + B + C + D + \text{gap} + \text{hbg}$, corresponding to energy bands 0.2–0.3, 0.3–0.4, 0.4–1.0, 1–2, 2–4, 4–12, a gap, and 13–15 keV. The background-subtracted rates are expected to be near zero in $S0_{\text{net}}$ and hbg_{net} , while the other bands contain the target spectrum to be used for science analyses. Tabulating the count rates in $S0_{\text{net}}$ and hbg_{net} provides a basis for quality filtering the background modeling process. One can view hbg_{net} as a quality diagnostic for the stage 1 background component, while $S0_{\text{net}}$ is a diagnostic for the stage 2 component.

Three levels of filtering are advised for NICER investigations of targets with different levels of X-ray brightness. They are detailed below in the sense of data selections for quality purposes, to be applied to GTIs prior to science analyses of light curves or spectra. The next section offers two examples of this process. The filter levels given below are illustrative, and users should explore the tradeoffs in coverage versus data quality to decide the optimal filter criteria that are consistent with investigation goals and requirements.

1. The level 1 filter selects GTIs with $((-30.0 < S0_{\text{net}} < 30.0) \text{ c s}^{-1} \text{ and } (-0.5 < \text{hbg}_{\text{net}} < 0.5)) \text{ c s}^{-1}$. This filter should be applied to even the brightest X-ray sources.
2. The level 2 filter selects GTIs with $((-10.0 < S0_{\text{net}} < 10.0) \text{ c s}^{-1} \text{ and } (-0.1 < \text{hbg}_{\text{net}} < 0.1)) \text{ c s}^{-1}$. The level 2 filter is appropriate for moderately bright sources, e.g., $20.0 < R_{\text{net}} < 300 \text{ c s}^{-1}$. For moderately bright sources with very soft spectra (e.g., with detections limited to energy below 2 keV), filter level 2S can additionally impose $-0.5 < D_{\text{net}} < 0.5 \text{ c s}^{-1}$, where the D-band (4–12 keV) is given up as an additional background band.
3. The level 3 filter selects GTIs with $((-2.0 < S0_{\text{net}} < 2.0) \text{ c s}^{-1} \text{ and } (-0.05 < \text{hbg}_{\text{net}} < 0.05)) \text{ c s}^{-1}$. This filter is appropriate for faint sources, e.g., $R_{\text{net}} < 20.0 \text{ c s}^{-1}$. For a faint source with a very soft spectrum, filter level 3S can again impose $-0.5 < D_{\text{net}} < 0.5 \text{ c s}^{-1}$.
4. It has been shown that the majority of GTIs with the highest background rates and largest model residuals occur in the polar regions of the ISS orbit (Figures 8 and 9). However, there is no effective way of screening results by orbit position without incurring significant data losses. To illustrate this, we define a group of “bad” model residuals as 153 GTIs (of 3477) with

$R_{\text{BGnet}} < -2.0$ or $R_{\text{BGnet}} > 2.0 \text{ c s}^{-1}$. A broad definition of the polar region, with latitude $\text{lat} < -42 \text{ deg}$ or $\text{lat} > 42 \text{ deg}$, captures 80% of the bad GTIs. However, only 9.4% of the polar GTIs are bad, and GTI exclusion on this basis would be costly. An ad hoc definition of the polar horns can be made from Figure 9, using the same polar region with additional constraints for longitude intervals: $200 < \text{lon} < 320$ (north) and $60 < \text{lon} < 180$ (south). This region captures 63% of the bad GTIs, but only 19% of the GTIs are bad.

6.4. Model and Filtering Considerations for the Brightest Source

Considerable attention has been paid to the brightest and the softest X-ray sources observed with NICER, to investigate the effects of source counts on the background parameters and also the background-subtracted count rates pertinent to quality filtering. We first consider the the high-energy range of the spectrum, specifically ibg , a 3C50 model parameter (15–18 keV, in focus), and hbg_{net} (13–15 keV, in focus), which is used as a data quality filter. We note that ibg is a raw measurement, while hbg_{net} is a background-subtracted quantity (subscript “net”). Both of these energy bands are outside the effective area of the concentrator optic, but there is still a finite probability that a high-energy photon may pass straight through to the detector, without interacting with the concentrator foils. Thus, it is relevant to investigate whether any extremely bright X-ray sources may elevate the count rates of either parameter. Of particular interest are Scorpius X-1, the brightest X-ray source in the sky ($116,000 \text{ c s}^{-1}$ when normalized to 50 FPMs); the black hole transients MAXI J1820+070 (65,000 normalized c s^{-1} at maximum) and MAXI J1348–630 (47,000 normalized c s^{-1}); and the neutron star transients with high-mass companion stars (HMXBs) and relatively hard X-ray spectra, Swift J0243.6+6124 (28,000 c s^{-1} at maximum) and A0535+26 (6000 c s^{-1}). The first three cases are the only targets (2017–2020) for which there was a commanded reduction of the number of active FPMs, so that the telemetry rate would remain below the maximum event rate ($\sim 30,000 \text{ c s}^{-1}$) for the cables connecting the output of the MPUs to the telemetry stream. For all of these sources, we find that contamination of ibg is not an issue. Values of ibg are found to be uncorrelated with changes in source intensity, when comparing these quantities on the GTI timescales. Variations in ibg are dominated by seemingly random changes in the background conditions, with no evidence of the outburst profile of the X-ray sources.

However, the impact of these bright or hard sources on hbg_{net} is somewhat different. After background subtraction, a residual count rate in hbg_{net} is seen in Scorpius X-1 (up to 1 c s^{-1}), and similar residuals are seen for the pair of bright HMXBs at times of maximum intensity. All of the other bright transients show $\text{hbg}_{\text{net}} < 0.2 \text{ c s}^{-1}$ when the sources are near maximum intensity. Since the prescription for the level 1 quality filter is to reject background-subtracted GTIs with $\text{hbg}_{\text{net}} < -0.5$ or $\text{hbg}_{\text{net}} > 0.5$, observations would be falsely rejected, at level 1 filtering, for Scorpius X-1, Swift J0243, and A0535-26. The solution to this problem is either to refrain from filtering these three sources, near times of maximum intensity, or to predict the NICER background with the “Space Weather” model (see Section 9.3), which has no parameters related to

measured count rates. Users of the 3C50 background model are advised to compare the light curves for exceptionally bright or hard X-ray sources with the light curves of the derived $hb_{g_{net}}$ values in order to determine the customized filtering values that are appropriate. Furthermore, the level of filtering should be approached as a function of source count rate, particularly for transients that NICER observes with more than five magnitudes of dynamic range between intensity maxima and the final measurements as the source returns to quiescence. The increased susceptibility to X-rays from bright and hard sources for hb_g (13–15 keV), relative to ibg (15–18 keV), can be understood as a combination of the decreasing absorption cross section at 13–18 keV in silicon, combined with the decreasing photon spectrum in that same range, for most X-ray subclasses.

An analogous search was made for the target contributions to n_z , the raw count at 0.0–0.25 keV, and the filtering parameter, SO_{net} , the count rate at 0.2–0.3 keV in the background-subtracted spectrum. The investigation included the same bright sources noted above, plus very soft X-ray transients, e.g., MAXI J0637-430 (6000 $c s^{-1}$ at maximum) and the coronal flares in HR1099 (reaching 675 $c s^{-1}$). It was found that the n_z values are generally not affected by exceptionally bright or soft sources. However, the level 1 filtering condition, $SO_{net} < 30 c s^{-1}$, is exceeded in the brightest GTIs for four sources: MAXI J1820, Scorpius X-1, MAXI J0637, and HR1099. Again, users can suspend data filtering near times of maximum intensity for these sources, or alternatively they can use the “Space Weather” background model (Section 9.3). The comparison of light curves in SO_{net} versus the broadband source intensity (0.4–12 keV) is a prudent step in the effort to customize the filtering and optimize the data quality for exceptional sources.

7. Background-subtracted Light Curves

As noted in the introduction, the 3C50 model libraries match the energy range and PI bins of nominal extractions from the cleaned event lists (0.2–15 keV), which is the expected starting point for obtaining light curves and spectra for the GTIs chosen by the user. In this section, we show background-subtracted light curves and the effects of quality filtering for two commonly used calibration sources, the Crab Nebula and 1E0102-72, which is a supernova remnant (SNR) in the Small Magellanic Cloud (SMC). Here, the light curves are derived from the background-subtracted spectra by integrating the counts within the desired energy range, for each GTI. For the Crab, which has a broadband spectrum, we choose the energy range to be 0.4–12 keV, as also done for the raw background count rate (R_{BG}). For 1E0102-72, which has a very soft X-ray spectrum, we show the light curve and filtering improvements for the energy range 0.3–2.0 keV. We note that 0.3 keV is a practical lower limit for analyses, based on the XTI effective area and the quality of current background models, while using 0.4 keV as a lower limit affords extra protection against noise intrusion during ISS daytime (see Table 4), with only a minor reduction in the count rate for broadband sources.

7.1. Observations of the Crab Nebula

The Crab Nebula is commonly used as a bright reference source in X-ray astronomy. However, the Crab’s intensity is not truly constant; long-term variations up to 7% were detected in multi-satellite observations (Wilson-Hodge et al. 2011). NICER observations (with the target name “PSR_B0531+21”) over the interval 2017 August 5 to 2020 April 27 were reprocessed, applying the same calibrations used for the background fields (Section 2). The query for GTIs, again using `nimake` while excluding the undershoot/overshoot rate filters, netted 418 GTIs with duration >50 s.

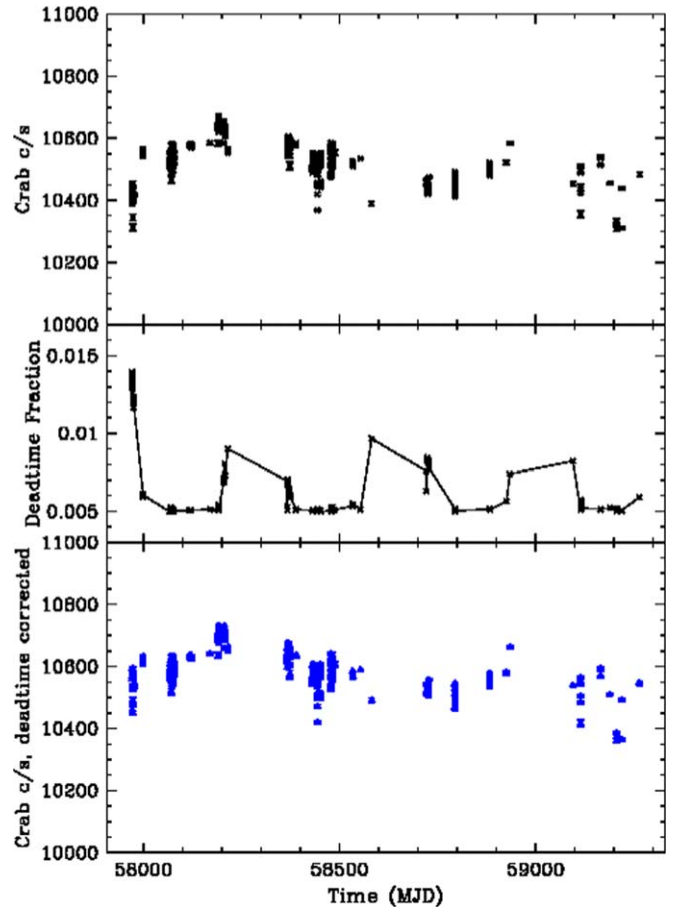


Figure 10. Top panel: NICER light curve of the Crab Nebula at 0.4–12 keV, after background subtraction with the 3C50 model and the application of level 1 filtering, which excluded only two GTIs. The selected data (431 GTIs) have a mean and rms deviation of $10,526 \pm 65 c s^{-1}$. Most of the the 0.6% variations can be seen as a gentle 900 day wave in the light curve. Middle panel: the deadtime fraction for Crab observations (see the text), showing the low level of NICER deadtime (0.5%–1.5% for the Crab) and the annual cycle tied to increases in undershoot and noise events when Crab observations are conducted in sunlight. Bottom panel: the deadtime-corrected light curve for the Crab Nebula, suggesting that the subtle 900 day waves in intensity may be real.

over the interval 2017 August 5 to 2020 April 27 were reprocessed, applying the same calibrations used for the background fields (Section 2). The query for GTIs, again using `nimake` while excluding the undershoot/overshoot rate filters, netted 418 GTIs with duration >50 s.

The application of the 3C50 model yielded background predictions for 433 GTIs, while four GTIs had ibg_{52} values that exceeded the 3C50 model limit ($10 c s^{-1}$). Count rates were integrated from the background-subtracted spectra in the range of 0.4–12 keV, and level 1 filtering was applied (see Section 6). This filtering step excluded only two GTIs as quality risks, and the count rates for the remaining 431 GTIs are shown in the top panel of Figure 10, scaled to 50 FPMs. The average GTI exposure is 811 s. These data display mean and rms values of $10,526 \pm 65 c s^{-1}$. In contrast, the two filter-eliminated GTIs yielded count rates of 10,655 and 10,277 $c s^{-1}$, respectively. The Crab data were also examined in terms of hard color, which is the ratio of count rates at 4–12 and 2–4 keV (or D/C in terms of energy band labels). The hard color measurements (not shown) have a mean and rms of 0.251 ± 0.001 . These intensity and color results indicate the photometric precision

that can be achieved with NICER spectra, using modest quality filtering, over the time interval 2017 August to 2021 February.

The Crab light curve shows that the 0.6% variations have a systematic temporal profile that can be seen as a gentle ~ 900 day wave in intensity. In the middle and bottom panels of Figure 10, we show that the intensity variations are not related to instrument deadtime. The NICER deadtime depends on the total event rate (all energies and all event flags) and the event processing time, which is approximately $20 \mu\text{s}$. This processing window freezes only the event-triggered FPM, but not the entire MPU and all of its 8 FPMs, since other FPMs remain free to trigger and hold the signal line data until the MPU is free to measure the triggering pulse. The X-ray light curve for the Crab averages $10,500 \text{ c s}^{-1}$, but the total event rate (R_T) varies between $11,000$ and $33,000 \text{ c s}^{-1}$, with the higher end caused by increased noise and undershoots (i.e., detector resets) during the most unfavorable observations conducted in sunlight. The deadtime fraction can be accurately computed for NICER, since the MPUs record the deadtime for every event as the difference between the event trigger time and the FPM re-enable time. This includes any hold time that an FPM may experience before a given event signal is analyzed. For a GTI with exposure t_i and a set of selected FPMs numbering N_{FPM} , one can sum the deadtime for all events, t_{dead} , from the selected FPMs during the GTI interval with a query to the “DEAD-TIME” column in the event lists. We note that this query must be applied to all of the relevant MPU files in the pipeline’s ObsID/xti/event_u/ directly, where no event types are excluded, and the query must also exclude any operating FPMs that are not selected for data analyses. The deadtime fraction is then simply, $f_{\text{dead}} = t_{\text{dead}}/t_{\text{GTI}}/N_{\text{FPM}}$. The deadtime fraction for the Crab observations is shown in the middle panel of Figure 10. Despite the bright intensity of the Crab Nebula, the deadtime fraction is confined to the range 0.5%–1.5%, with a distinct 1 yr oscillation tied to the Sun angle. The deadtime correction factor for each GTI is $1.0/(1.0 - f_{\text{dead}})$, and the deadtime-corrected light curve for the Crab is shown in the bottom panel of Figure 10. The differences between the top and bottom panels are very subtle, and we conclude that deadtime is not the cause of the 900 day wave seen for the Crab’s light curve. Observations of the Crab Nebula with other space missions will help to determine whether these variations are intrinsic to the Crab or arise from systematic factors that have escaped the current investigation.

Observations of the Crab Nebula were also analyzed to quantify the similarity of the individual FPM responses, helping to assess the 3C50 model assumption that the background predictions can be scaled for any number of FPMs selected for a given GTI. Individual FPM spectra were made from 202 ks of Crab exposures conducted during ISS nighttime. The 52 operating FPMs yield a distribution with mean (rms) $209.9 (9.0) \text{ c s}^{-1}$ at 4–12 keV. This corresponds to FPM response deviations of 4.3%, which is substantially smaller than the residuals, as percentage of the measured background, in the 3C50 model predictions. We further note that the Crab deviations are subject to small variations in FPM boresight positions, which would not be expected to matter for background observations.

7.2. Light Curve of 1E 0102.27219

The SNR in the SMC, 1E0102.2–7219 (hereafter “E0102”), is a faint calibration source that serves as a flux and spectral

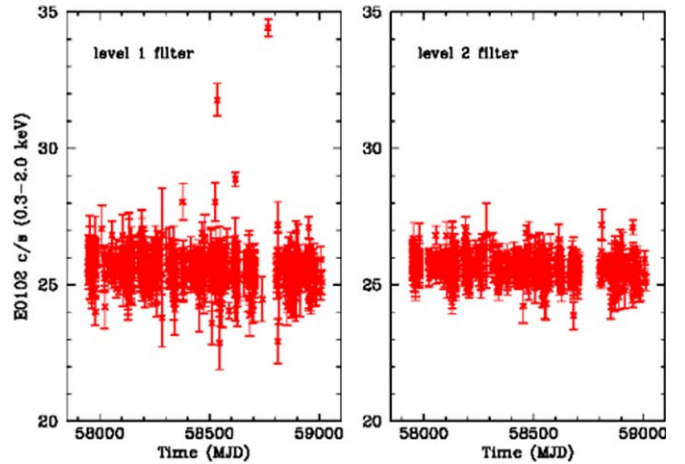


Figure 11. NICER observations of the SNR in the SMC, E0102, yielded background-subtracted spectra with the 3C50 model for 941 GTIs through 2020 June 12, accumulating 408 ks exposure time. The soft light curve (0.3–2.0 keV) extracted from all of the GTIs is shown in the left panel, while the same data with level 2 filtering (804 GTIs, 352 ks) is shown in the right panel (see Section 6). These measurements have a mean and rms of $25.60 \pm 0.62 \text{ c s}^{-1}$ and $25.63 \pm 0.42 \text{ c s}^{-1}$, respectively, while the rms statistical uncertainty is 0.32 c s^{-1} .

line reference for many X-ray instruments (Plucinsky et al. 2017). The NICER observations from 2017 July 17 to 2020 June 12 netted 965 GTIs with an average exposure of 438 s. The 3C50 model yielded 941 background-subtracted spectra, while filtering steps (see Section 6) left 916 GTIs at level 1 and 804 GTIs at level 2. We note that the fraction removed by the level 2 filter (15%) is larger than normal, because E0102 is observed in a wider range of conditions, as a calibration source, compared to many NICER science targets.

Since the source is soft, we examine the background-subtracted light curve in the range 0.3–2.0 keV. Figure 11 shows the results for the level 1 filtering (the left panel) and the level 2 filtering (the right panel). The measurements have a mean and rms of $25.61 \pm 1.15 \text{ c s}^{-1}$ and $25.62 \pm 0.41 \text{ c s}^{-1}$, respectively. The average statistical uncertainty at 0.3–2.0 keV is 0.3 c s^{-1} . Many of the lowest points and the highest points in the light curve of 1E0102 are eliminated with level 2 filtering. This shows that the filtering steps help to predict when the 3C50 model and the actual background do not agree (see Figure 8). Quality filtering with $S_{0,\text{net}}$ and $h_{\text{bg},\text{net}}$ sacrifices some amount of temporal coverage to improve the accuracy of the light curves and X-ray spectra.

8. Background Modeling at 1 s Timescale

The background parameters, ibg and hrej , normally have count rates below 1 Hz, and one must integrate for a few hundred seconds to produce an average value with reasonable statistical precision. However, the occasional surges in the background rates show corresponding variations in ibg and hrej , and the relationship between these quantities can help to diagnose whether the rapid changes in the NICER light curves may originate from either the X-ray target or from the background.

In Figure 2, it was shown that ibg and hrej are both roughly correlated with R_{BG} , with somewhat different average slopes. This motivates a strategy of estimating R_{BG} as a linear combination of ibg and hrej with different coefficients. Using a least-squares fit confined to the range $0.5 < R_{\text{BG}} < 300 \text{ c s}^{-1}$,

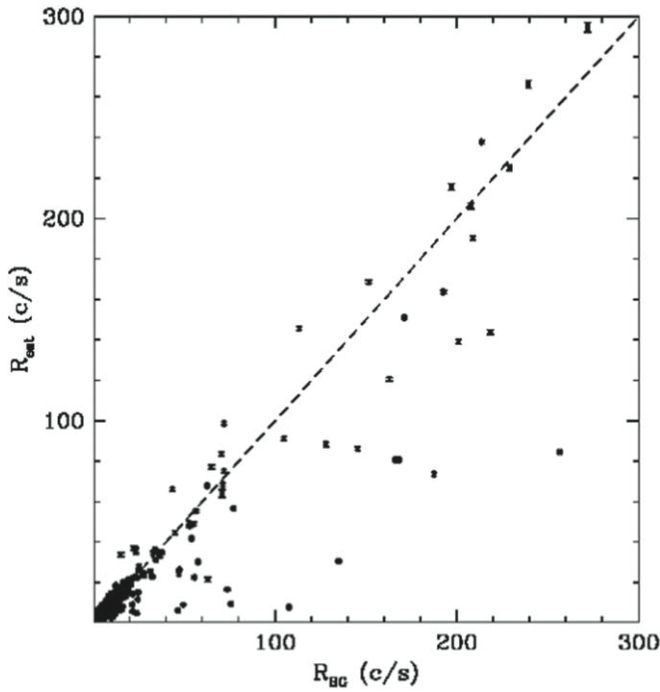


Figure 12. Estimates of the in-band background rate using the relationship $R_{\text{est}} = 2.91 * \text{ibg} + 4.67 * \text{hrej}$. Each point corresponds to one GTI, and the equation is determined with a least-squares fit confined to the range $0.5 < R_{\text{BG}} < 300 \text{ c s}^{-1}$. The error bars consider only the statistical uncertainties in R_{BG} , ibg , and hrej . The dashed line is a reference for the match between this relationship and the measured rate (R_{BG}). The estimator is valuable for showing background temporal variability, e.g., at 1 s timescale, that can help to determine whether fast flaring behavior originates from the target or from activity in the background.

the best-fit relationship is $R_{\text{est}} = 2.91 * \text{ibg} + 4.67 * \text{hrej}$, with the results shown in Figure 12. There is significant scatter in the ability of the background estimator to predict R_{BG} at high count rates, pointing to the same problem seen with the 3C50 model residuals (Figure 8). Nevertheless, the background estimator might show rapid increases and a temporal structure that resembles the NICER light curve in short time bins, and this would convincingly implicate the background as the origin of the fast flares.

To illustrate the use of the background estimator, we consider the case of the X-ray transient source Swift J1858.6–0814 (Krimm et al. 2018), hereafter “SwiftJ1858.” NICER observations from 2018 November 1 through 2019 November 17 show dozens of GTIs containing fast variability, where the source intensity ranges from nondetectable levels to multiple sharp maxima in the range $100\text{--}1600 \text{ c s}^{-1}$ (Ludlam et al. 2018). It was later shown that much of this “flaring” is actually driven by variable absorption along the line of sight to a nearly eclipsing binary system (Buisson et al. 2021). After a data gap imposed by a low Sun angle during the interval MJD 58805–58903, NICER found SwiftJ1858 to be in a more conventional state of quasi-steady emission, with eclipses and absorption dips. Then type I X-ray bursts were detected, identifying the source as an accreting neutron star (Buisson et al. 2020).

Despite the propensity of SwiftJ1858 to vary rapidly, during the first part of its outburst, it is necessary to distinguish the few cases in which rapid flaring originated in the background, rather than the X-ray source. Figure 13 shows two GTI light curves, in 1 s bins, with contrasting findings regarding the

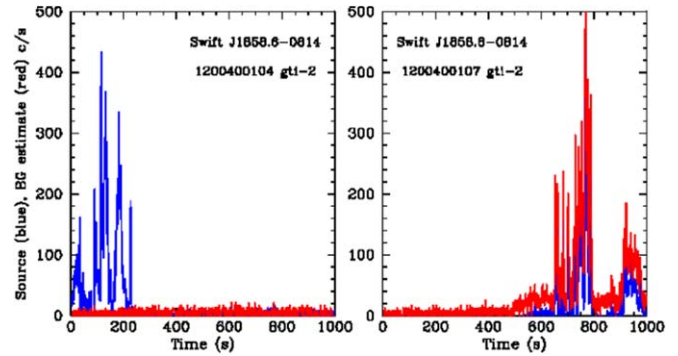


Figure 13. The application of the 1 s background estimator to two GTIs from observations of SwiftJ1858 when the source, an X-ray binary system containing an accreting neutron star, was in a flaring state. The first GTI shows that the rapid flaring is from the X-ray source (blue), and this behavior has been observed many times with NICER and other instruments. However, the second GTI shows a light curve with unusual rapid flaring in the background (red), demonstrating that this particular sequence of flares is from the background, rather than the X-ray source.

origin of the fast flares. In both cases, the light curve shown in blue is the background-subtracted count rate using the 3C50 model with parameters averaged over the respective GTIs. The values for the background estimator are shown in red. In the first case (the second GTI on MJD 58426), the background intensity remains low and quiet, implying that the flares originate in SwiftJ1858.

In the second case (the second GTI on MJD 58429), the background estimator shows that the high-amplitude variations coincide with significant activity in ibg and hrej . A precise match is not expected between the blue and red curves, since an effort has been made to background-subtract the blue curve on the GTI timescale (note the negative values at times before 800 s), and since the predictability of the background is compromised by systematic error in the model at high background rates. Nevertheless, it is clear that the second set of flares originated in the background and not in SwiftJ1858. In this example, the background flares did occur in the southern polar horn, during a time interval with a range in orbit latitude and longitude of $-51.7 \text{ deg} < \text{lat} < -47.5 \text{ deg}$ and $79.0 \text{ deg} < \text{lon} < 107.2 \text{ deg}$.

9. Discussion

9.1. Sensitivity Limits with the 3C50 Background Model

Quantitative analysis of the residuals in a background model, when applied to observations of blank sky regions, provides information on both the instrument sensitivity limits and systematic problems regarding model performance. Given the large range in R_{BG} (Figure 2), and the limited success of the 3C50 model when $R_{\text{BG}} > 2.0$ (Section 5), these topics must be approached with qualifications. After applying the 3C50 model to 3447 GTIs with model parameters within limits, the residuals at 0.4–12 keV are within $\pm 0.5 \text{ c s}^{-1}$ in 80% of the GTIs. Applying the level 3 filtering criteria (see Section 6), which exclude 15% of these GTIs, the rms value of R_{BGnet} is 0.40 c s^{-1} . This, in turn, implies a detection limit (3σ in a single GTI) of 1.20 c s^{-1} at 0.4–12 keV, which is equivalent to $3.6 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$ at 0.4–12 keV, assuming the spectral shape of the Crab Nebula. In the soft X-ray band, the corresponding detection limit (3σ , single GTI) is 0.51 c s^{-1} at 0.3–2.0 keV. These limits would improve by a factor of four if

the exploratory GTIs for a given target accumulate 10 ks of level 3–filtered exposure time, since the filtering removes a substantial portion of the systematic error associated with the background model. The percentage of GTIs that will pass the level 3 filtering is estimated to be $\sim 90\%$, since the scheduling of science targets would be more favorable than the background targets, which widely sample the background conditions, with GTIs surviving the level 3 filtering 85% of the time.

9.2. Model Limitations at High Background Rates

Our assessment of the 3C50 model noted evidence of a missing model parameter that is particularly important at high values of the raw background rate at 0.4–12 keV (R_{BG}). The correlation between R_{BG} and ibg implies that most of these events are associated with the in-focus component implied by the values in the PI_ratio . This is reminiscent of the early Chandra discovery that protons could scatter off the mirrors and come into focus on the detectors when the satellite passed through the radiation belt (O’Dell et al. 2000). The hypothesis that a similar condition is affecting NICER would imply that the background model should consider the angle between the camera’s pointing direction and the local magnetic field lines during the course of the NICER orbit. Such attention was suggested by Fukazawa et al. (2009) in the background model for the Suzaku HXD Instrument.

Further motivation to track the camera’s viewing direction, relative to its position in the Earth orbit, is provided by a study of an archive of particle rate measurements built with a series of NOAA polar-orbiting satellites (Fidani et al. 2010). Since 1998, these satellites have been equipped with the Space Environment Monitor system, which contains two sets of instruments that monitor the energetic charged particle environment above the Earth. One of these systems, the Medium Energy Proton and Electron Detector, has two sets of detectors mounted perpendicular to each other. The different viewing angles produce differential particle fluxes for electrons in the ranges of both 30–100 and 100–300 keV (see Figure 1 of Fidani et al. 2010). The pitch angle, i.e., the angle between the charged particle flow and the local magnetic field, varies systematically with the position in the low Earth orbit, suggesting that the particle flux should depend on the longitude, latitude, the camera angle with respect to the local magnetic field, and the local pitch angle. This context will be explored further in the effort to link the direction of particle flow to the in-focus component in the NICER background.

The limitations of the 3C50 model were first apparent in Figure 4, where the raw background rates showed significant variations when binned in model cells for the nighttime library. This, in part, motivated the strategy of renormalizing the selected library spectrum for a given GTI, by the factor ibg_i/ibg_{lib} . An assessment of the impact of this step can be made by comparing the residuals with and without the renormalization step, while making use of the quality filtering criteria described in Section 6.3. Closing the loop on the background pointings with (without) ibg renormalization leaves residuals that pass the level 1 filtering in 93% (88%) of all GTIs, while passing the level 2 filtering 85% (85%) of the time. The conclusion is that ibg renormalization provides only a modest improvement in the quality of the 3C50 model when the background rate is high. This is likely due to both the need for an additional parameter in the background model, and also

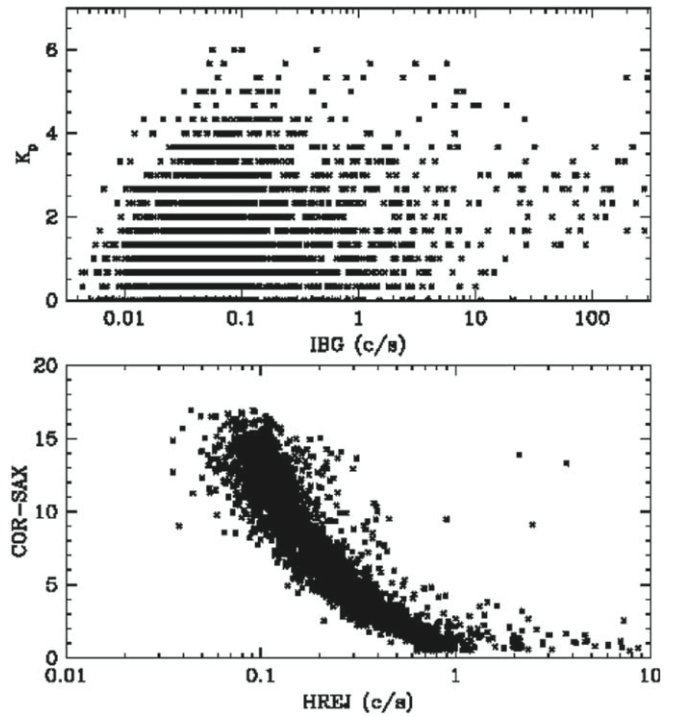


Figure 14. The relationship between the 3C50 model parameters and those of the NICER Environmental Model. The plot of $hrej$ vs. the cutoff rigidity (COR_SAX) is the best correlation seen between any two hypothetical background parameters in this study. On the other hand, ibg and K_p pair with their model partners in different ways.

the limited statistical accuracy of the ibg values because of the low count rate in this parameter.

9.3. Comparisons with the Parameters of the Space Weather Background Model

The “Space Weather” background model is an alternative to 3C50 that predicts NICER background spectra on the basis of the local spacecraft environment (K. C. Gendreau et al. 2022, in preparation). It is implemented in the FTOOL, “*nicer_bg_estimator*,” which is also available via the HEASARC NICER tools website (see the prior footnote). The model parameters are the local cutoff rigidity (“COR”; Smart & Shea 2005), which is a measure of the shielding provided by the Earth’s magnetic field, and the K_p index, which is a global measure of the disturbances in the magnetic field. The NICER pipeline furnishes the values of “ COR_SAX ,” every second, in the filter files, and the values of the cutoff rigidity are computed with a particular model developed for the BeppoSAX Mission (Amati et al. 2002). COR_SAX shielding has units in the range 0–17, while the K_p index is given for each 3 hr interval, with a range 0–6 quantized in steps of 0.333. The values for the K_p index are obtained from the GFZ site in Potsdam.¹²

The relationship between the model pairs for the 3C50 and Space Weather models is examined in Figure 14. It is apparent that $hrej$ and COR_SAX are measuring the same background component, i.e., the amount of magnetic shielding at the ISS orbital position is inversely correlated with the rate of spatially extended events due to particles, as measured with $hrej$. The COR_SAX parameter would be a desirable substitute for $hrej$, in future versions of the 3C50 model, because it is readily

¹² <https://www.gfz-potsdam.de/en/kp-index/>

accessible in NICER filter files, and is free from the statistical accuracy limits that confront hrej, due to its low count rate.

On the other hand, there is no correlation between ibg and K_p , implying that the two models handle the other background component(s) in different ways. Users are currently invited to use both models and investigate their performance in ways customized to their science goals. Finally, we note that the filtering steps described in this paper (Section 6.3) can be applied to background-subtracted spectra derived from any any NICER background model.

10. Summary

NICER has a comparatively low background rate, typically 10^{-4} times the broadband count rate of the Crab Nebula, but it is highly variable in both amplitude and spectral shape. The silicon drift detectors are high-throughput but single-channel devices, and so the background spectrum must be predicted using measurements that are not affected by the targeted X-ray source. The 3C50 model predicts the background spectrum using an empirical approach and three model parameters. Data analyses are based on recurrent observations of seven pointing directions that are void of detectable sources. Spectra from a wide range of observing conditions are sorted by the values of the model parameters to build a two-stage library of spectra that are the core of the background model. The first stage applies to all observations, while the second stage removes a soft X-ray excess due to detector noise that is encountered, in a predictable way, when observations are conducted in sunlight. It is noted that most particle hits are automatically excluded from either the target spectrum or the background model because the energy of the event trips the overshoot flag, removing such events from spectral consideration.

Two model parameters, ibg and $hrej$, track background components associated with particle-induced events. They are distinguished by the values of the PI_ratio , which is the ratio of event energies in the fast measuring chain, relative to the slow chain in the instrument electronics. The values of the PI_ratio can discriminate detector ionization locations near the center of the silicon drift detector (i.e., the events appearing “in focus”) from those near the outer edges of the detector (hence associated with spatially extended events). We define ibg as the rate of in-focus events at 15–18 keV (beyond the effective area of the optics), while $hrej$ is the rate of particle events at 3–18 keV that originate near the outer edges of active silicon, underneath the metal collimator. A grid of values in these two parameters is used to bin and average the GTI-based collection of background spectra to form the stage 1 library of the model. The measured values in these parameters for any given target observation are then used to select a matching library spectrum. That spectrum is renormalized by ibg/ibg_{lib} to form the stage 1 prediction of the NICER background. The third parameter, nz (the count rate at 0–0.25 keV), is used to predict a low-energy excess that is tied to observations conducted in sunlight, when $nz > 200 \text{ c s}^{-1}$. For each of the twelve intervals in nz , the residual spectra from stage 1 of the model are averaged to form the stage 2 library. Then, for target observations, the measured value nz_i is used to select a spectrum from the stage 2 library. That spectrum is renormalized by nz_i/nz_{lib} , and the result is added to the stage 1 background spectrum to complete the background prediction.

The small contribution from the cosmic diffuse X-ray background is carried into the background model by the

manner in which the stage 1 library is constructed. This component is always present in the NICER field of view, and its inclusion in the 3C50 model is guaranteed by the imposition of a minimum value for ibg , 0.016 c s^{-1} . There are no provisions in the model for diffuse Galactic emission, components local to the Earth or the solar system, or contaminating sources in the field of view. Such contributions, when anticipated, must be considered externally.

An examination of 3556 GTIs, with an average duration of 570 s, shows that the in-band count rate of good events at 0.4–12 keV, scaled to 50 selected detectors, has a median value 0.87 c s^{-1} . However, the distribution is quite broad, ranging from 0.33 to 300 c s^{-1} , after excluding the 1% outliers on each end. After applying the 3C50 model to 3447 GTIs with model parameters within limits, the residuals at 0.4–12 keV are within $\pm 0.5 \text{ c s}^{-1}$ in 80% of the GTIs. However, residuals persist at 20%–30% of the initial rate for the brightest cases, which tend to occur in the polar horns of the ISS orbit (mixed with many quiet GTIs at the same polar locations). The inaccuracy of the model when the background rate is high suggests one or more missing model parameters. Evidence from satellites that measure particle energies in the environment of low Earth orbits suggests that the direction of the NICER cameras with respect to the local magnetic field lines may be an important consideration for improving the 3C50 model.

Quality filtering criteria are developed to warn users when the predicted background spectrum is not likely to be satisfactory. The background-subtracted (“net”) count rate at 0.2–0.3 keV ($S0_{net}$) and at 13–15 keV (hbg_{net}) should have values near zero for most targets. Screening criteria were advised for sources at various levels of intensity in Section 6.3. When such filtering criteria are applied at “level 3” (appropriate for faint X-ray sources), selecting GTIs only if $-2.0 < S0_{net} < 2.0 \text{ c s}^{-1}$ and $-0.05 < hbg_{net} < 0.05 \text{ c s}^{-1}$, then the systematic uncertainty in the model, which is an estimate of the detection limit, is 1.20 c s^{-1} at 0.4–12 keV (3σ , for a single GTI), and 0.51 c s^{-1} at 0.3–2.0 keV. For a Crab-like spectrum, the detection limit at 0.4–12 keV is equivalent to $3.6 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$. The limiting count rate in soft X-rays is equivalent to $4.3 \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$, assuming a 100 eV blackbody spectrum, with an interstellar medium column density of $5 \times 10^{20} \text{ cm}^{-2}$. These limits would improve by a factor of four if the exploratory GTIs accumulate 10 ks of level 3-filtered exposure time. The GTIs that pass such filtering criteria amount to 85% of the total, while higher success rates would be expected for general targets scheduled more favorably than the background observations.

Under normal conditions, the empirical model’s background predictions are limited to timescales of minutes or longer, because of Poisson noise in ibg and $hrej$, which often have count rates $< 1 \text{ c s}^{-1}$. However, the crude background estimator, $R_{est} = 2.91 * ibg + 4.67 * hrej$, can be applied on timescales of seconds or less, to help assess whether observations of fast variability originate in either the background or the X-ray source. Background flares that are associated with significant variations in the raw in-band count rate produce momentarily high values of ibg and $hrej$, and the temporal structure in R_{est} will be highly correlated with the in-band variations under scrutiny.

ORCID iDs

Ronald A. Remillard  <https://orcid.org/0000-0003-4815-0481>
 Teruaki Enoto  <https://orcid.org/0000-0003-1244-3100>
 Keith C. Gendreau  <https://orcid.org/0000-0001-7115-2819>
 Craig Markwardt  <https://orcid.org/0000-0001-9803-3879>
 Arkadip Basak  <https://orcid.org/0000-0002-6575-2153>
 Abigail L. Stevens  <https://orcid.org/0000-0002-5041-3079>
 Paul S. Ray  <https://orcid.org/0000-0002-5297-5278>
 Diego Altamirano  <https://orcid.org/0000-0002-3422-0074>

References

- Amati, L., Frontera, F., Tavani, M., et al. 2002, *A&A*, **390**, 81
 Arenberg, J., Matthews, G., Atkinson, C., et al. 2014, *Proc. SPIE*, **9144**, 91440Q
 Bogdanov, S., Guillot, S., Ray, P. S., et al. 2019, *ApJ*, **887**, 25
 Buisson, D. J. K., Altamirano, D., Armas Padilla, M., et al. 2021, *MNRAS*, **503**, 5600
 Buisson, D. J. K., Altamirano, D., Bult, P., et al. 2020, *MNRAS*, **499**, 793
 Fidani, C., Battiston, R., & Burger, W. J. 2010, *RemS*, **2**, 2170
 Fukazawa, Y., Mizuno, T., Watanabe, S., et al. 2009, *PASJ*, **61**, 17
 Gendreau, K. C., Arzoumanian, Z., Adkins, P. W., et al. 2016, *Proc. SPIE*, **9905**, 99051H
 Hare, J., Tomsick, J. A., Buisson, D. J. K., et al. 2020, *ApJ*, **890**, 57
 Jahoda, K., Markwardt, C. B., Radeva, Y., et al. 2006, *ApJS*, **163**, 401
 Krimm, H. A., Barthelmy, S. D., Baumgartner, W., et al. 2018, *ATel*, **12151**
 Ludlam, R. M., Miller, J. M., Arzoumanian, Z., et al. 2018, *ATel*, **12158**
 O'Dell, S. L., Bautz, M. W., Blackwell, J. C., et al. 2000, *Proc. SPIE*, **4140**, 99
 Okajima, T., Soong, Y., Balsamo, E. R., et al. 2016, *Proc. SPIE*, **9905**, 99054X
 Plucinsky, P. P., Beardmore, A. P., Foster, A., et al. 2017, *A&A*, **597**, A35
 Prigozhin, G., Gendreau, K. C., Doty, J. P., et al. 2016, *Proc. SPIE*, **9905**, 99051I
 Smart, D. F., & Shea, M. A. 2005, *AdSpR*, **36**, 2012
 Wilson-Hodge, C. A., Cherry, M. L., Case, G. L., et al. 2011, *ApJL*, **727**, L40
 Wu, X., Hamilton, T., Helfand, D. J., & Wang, Q. 1991, *ApJ*, **379**, 564