THE ASTROPHYSICAL JOURNAL, 388:595–602, 1992 April 1

TIME VARIABILITY STUDIES WITH PHOTON-COUNTING IMAGING DETECTORS. I. A MAXIMUM LIKELIHOOD TECHNIQUE

S. SCIORTINO AND G. MICELA
Osservatorio Astronomico di Palermo, Palazzo dei Normanni 1, 90134 Palermo, Italy
Received 1990 December 27; accepted 1991 October 7

ABSTRACT

We present a new method for characterizing the variability properties of a preselected class of celestial objects which are repeatedly observed with the same photon-counting imaging detector. Instrument-specific characteristics, such as detector instrument point response function and mirror system off-axis response, are separated and taken into account in the form of input parameters. All collected source data, both in the form of detections and upper bounds at putative source positions, are retained and fully used in determining the maximum likelihood (ML) estimator of the intrinsic source count rate. For each distinct observation we can determine, at the chosen significance level based on Poisson statistics, whether the observed flux (or upper bound) undergoes a statistically significant variation with respect to the derived ML flux estimator, and we can evaluate variation amplitude, or an upper bound for those cases where significant variations have not been detected. The derived variability amplitudes (or upper bounds) allow us to build, for the chosen class of celestial objects, the ML variability amplitude distribution function, which allows us to characterize the variability features of the class under study. We illustrate the capabilities of the proposed technique, applying it to the samples of the Pleiades and the Hyades stellar X-ray sources surveyed with the Imaging Proportional Counter flown on the *Einstein Observatory*. We conclude that, on the basis of presently limited sensitivity data, the variability behavior of the X-ray coronal emission of Pleaides and Hyades stars cannot be distinguished.

Subject headings: methods: numerical — methods: statistical — stars: coronae — X-rays: stars

1. INTRODUCTION

The study of temporal variations is one of the most powerful ways to characterize and study the physical conditions of any celestial source. Since the beginning of X-ray astronomy the study of variability of X-ray emission has been quite an active area of research. Studies of time scales, variability amplitude, and other related characteristics furnish useful information on dimensions and physical conditions of the regions where X-ray emission originates and on the mechanisms responsible for the observed emission. Comparative studies of the variability properties of homogeneous classes of X-ray sources are well suited to diagnostic and to differentiating the mechanisms at the origin of their X-ray emission. To accomplish these results in an effective way requires access to a mass of data homogeneously reduced and an *objective* way to characterize and to compare the variability properties of different samples of X-ray sources. While selected samples of the more intense sources have been searched for both short- and long-term X-ray variability (Snow, Cash, & Grady 1981; Montmerle et al. 1983; Zamorani et al. 1984; Ambruster, Sciortino, & Golub 1987; Peres et al. 1989; Collura, Pasquini, & Schmitt 1988; Collura et al. 1989; Collura, Reale, & Peres 1990; Pallavicini, Tagliaferri, & Stella 1990; White et al. 1986, 1987, 1990), little has been said on the long-term X-ray variability of the new class of lower flux emitters (largely stars and AGNs) made accessible with the successful *Einstein* and *EXOSAT Observatories* (see Maccacaro, Garilli, & Mereghetti 1987; Mereghetti & Garilli 1987).

The advent of imaging X-ray observatories and of photon-counting imaging detectors, with their impressive increase in sensitivity, has drastically improved the capability to explore the log (N)–log (S) relation in the low flux region, increasing the number of X-ray sources by a factor ~ 100 (for a recent report of many of *Einstein Observatory* results cf. Elvis 1990). The long uninterrupted observations of the *EXOSAT* satellite have provided a way to monitor X-ray variability from hundreds of seconds to hours in hundreds of Galactic and extragalactic sources (see, e.g., the review papers of White & Peacock 1988; Giommi, Tagliaferri, & Angelini 1988; Pallavicini 1988; Osborne 1988; Parmar & White 1988; Stella 1988); the substantially larger sensitivity and larger field of view of the *Einstein* IPC (Imaging Proportional Counter; Giacconi et al. 1979b; Gorenstein, Harnden, & Fabricant 1981) have allowed the determination of fluxes for $\sim 12{,}000$ X-ray sources and significant upper bounds for a large fraction of the objects falling in the $\sim 10\%$ of the sky surveyed with this detector (Harris et al. 1992); ~ 3000 of the detections and more than 30,000 significant upper bounds belong to the class of the stellar coronal sources (Sciortino et al. 1988; Harnden et al. 1990). Approximately 15% of the IPC images cover overlapping regions of the sky, yielding a similar fraction of X-ray sources observed more than once, and resulting in a total of ~ 1500–2000 celestial objects seen several times and detected at least once. Hence the IPC observations contain a large amount of information on long-term X-ray variability, but this information requires proper methodology for being extracted and studied. Our aim in the present paper is to present a new approach for this kind of study.

Motivated by our specific interest in undertaking an objective comparative analysis of the X-ray variability properties of several samples of X-ray sources, we have developed an original methodology, especially suited to search for long-term variability in low-statistic X-ray sources, observed with photon-counting detectors having imaging capabilities. This

methodology makes use of both source counts and background counts and takes into account exposure time and source position in the detector field of view.

While this method has been originally motivated by a device to analyze variability of sources observed more than once with the *Einstein Observatory* Imaging Proportional Counter, we want to stress, and our presentation attempts to show, that the method is quite general and obviously can be applied also in fields different from X-ray astronomy. The determination of variability significance level and variability amplitudes is specifically tailored for imaging photon-counting detectors and properly accounts for the uncertainty associated with low count number statistics. The study of sample properties using both determination of, and upper bounds to, variability amplitude is of more general application: any variability analysis technique that allows one the determination of upper bounds for the variability amplitude can benefit by its application.

Our paper is organized as follows: a description of the detailed technique to determine variability amplitude from raw data is given in § 2; in § 3, to illustrate the characteristics of the method and its capability to characterize sample properties, we present its applications to IPC X-ray data of stars in two open clusters. A brief summary of our results and of the main features of the proposed technique is given in § 4. We intend to present in forthcoming papers of this series a comparative analysis of the long-term X-ray variability of solar-type stars of different spectral type as a diagnostic of the occurrence of magnetic activity phenomena, and to characterize the occurrence of X-ray flux variability in early-type stars at different evolutionary stages.

2. THE METHOD

The imaging capability of an hypothetical photon-counting detector makes possible the selection of the photons falling in any chosen subsection of the detector field of view, i.e., at the detected (or presumptive) source position. The number of counts collected in each subsection of the detector is a realization drawn from a Poisson distribution whose mean value depends both on the contribution of diffuse background and possibly pointlike (X-ray) sources at the given position. We want to derive a maximum likelihood (ML) estimator that takes into account all instrumental and observational effects, to evaluate the "true" source rate in the presence of repeated observations of the same region and under the hypothesis that the contribution from (X-ray) pointlike sources is dominated by the source under investigation (i.e., we assume that the contributions from other sources mimic a diffuse background). With these data in hands one can test, at the chosen significance level, the "null hypothesis" that the source flux is constant, comparing for each observation the expected number of counts (i.e., that derived from ML estimate of rate) with the observed number of counts. This method allows us to search for long-term variability using the counts collected at source position in all the available images where the source occurs to fall in; in particular *all available observations*, both positive detection and upper bounds are retained in the analysis.

For the sake of clarity we present the application of this method to IPC data resulting from the final standard data processing (see, e.g., Harnden et al. 1984; Harris et al. 1992). In this data processing X-ray source count rates are determined on the basis of two detection methods the "Map" and the "Local" Detection algorithm. Both scan an image with a moving cell window, and identify significant fluctuations above the background as detections. The two methods differ in the way the background and the source rate are evaluated. In the Map algorithm the background is determined from a reference map calibrated using the long exposures of the Deep Survey (Giacconi et al. 1979*b*; Primini et al. 1991) so that the number of background counts is known with high accuracy, and its error can be neglected. In this case the number of counts, C, collected in the detection cell is

$$C = \phi_{\rm PRF} S + B \,, \tag{1}$$

where $\phi_{\rm PRF}$ accounts for the fraction of source counts, S, collected in the detection cell due to the width of the point-response function, and B are the background counts falling in the detection cell. In the Local algorithm the background is evaluated locally and is known with less accuracy than in the Map algorithm. The source counts and the background counts have to be determined from the image using two independent measurements. This is accomplished measuring at the same time C and the number of counts, F, collected in a larger cell concentric to the detection cell (see Fig. 1). In this geometry we have the further relation (see, e.g., Harnden et al. 1984):

$$F = \beta S + \eta B \tag{2}$$

where $\beta(\sim 1)$ is the fraction of source counts falling in the larger cell and η is the ratio of large cell area and detection cell area. The relations between the two measured numbers of counts C and F and the derived quantities S and B are

$$S = \frac{1}{\phi_{\rm PRF} - \beta/\eta}\left(C - \frac{F}{\eta}\right), \tag{3a}$$

$$B = \frac{1}{\eta - \beta/\phi_{\rm PRF}}\left(F - \frac{\beta C}{\phi^{\rm PRF}}\right). \tag{3b}$$

The intrinsic source rate can be deduced from the value of S given by equations (2) or (3a), taking into account the reduction in collecting efficiency introduced by the mirror system. The two main effects are (1) the scattering at large angles about the desired images of a fraction of gathered X-rays due to small-scale mirror imperfections and (2) the reduction of reflection efficiency for X-rays striking mirror surface along directions not parallel to the mirror optical axis, leading to vignetting of off-axis sources. Introducing proper correction factors $\phi_{\rm MS}$ and $\phi_{\rm vign}$ for the mirror scattering and mirror vignetting effects, respectively, we can write the intrinsic source rate as

$$R_S = \frac{r_S}{\phi_{\rm MS}\,\phi_{\rm vign}} = \frac{S/T}{\phi_{\rm MS}\,\phi_{\rm vign}}; \tag{4}$$

where T is the net exposure time.

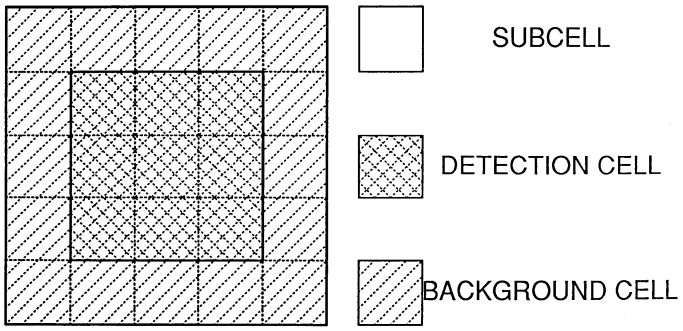

FIG. 1.—Sliding window detection cell layout of the IPC data processing system. Note the internal source cell and the surrounding background frame.

In real applications, one wants to infer conclusions on temporal variations of source rates, after having tested the occurrence of temporal variations of counts collected in the detection cell. For this purpose one has to distinguish between two cases, namely: (*a*) the error associated with the background is negligible—as for the so-called Rev-1 Map detection method—in such a case the background, in each observation, can be described with a very narrow distribution;[1] (*b*) the number of background counts is known (from eq. [3a]) with finite precision—as for the so-called Rev-1 Local detection method—because both the F_i and C_i counts follow a Poisson distribution. In this case the uncertainty in background determination could be of the same order of that of C_i and F_i and must be taken properly into account to assess source variability. To properly estimate the source rate we have to determine the background rate taking into account, at least in a first-order approximation, the various components that make the measured background and their variations across the field of view.[2] In the case of the *Einstein* IPC the background can be modeled as the sum of a non–X-ray (not imaged) component—largely due to particle events—B_P, and an X-ray (imaged) component, $B_{\text{X-ray}}$. The nonimaged component does not suffer mirror losses, while the X-ray component is affected by the vignetting effect, hence the "observed" number of background counts, B, at the given position in the field of view, is

$$B = B_P + B_{\text{X-ray}} = (R_{B_P} + \phi_{\text{vign}} R_{B_{\text{X-ray}}})T \,, \tag{5}$$

where R_{B_P} is the nonimaged background rate and $R_{B_{\text{X-ray}}}$ is the rate of the X-ray component. The overall intrinsic background rate R_B can be evaluated from the "observed" number of background counts, B, if one knows the fraction of imaged intrinsic background, f_B. The expected number of "intrinsic" background counts, B_{true}, is

$$B_{\text{true}} = \frac{B}{(1 - f_B)\phi_{\text{vign}} + f_B} \tag{6}$$

or in terms of rate

$$R_B = \frac{B/T}{(1 - f_B)\phi_{\text{vign}} + f_B} \,. \tag{7}$$

The best estimate of the fraction of nonimaged IPC background, $f_B \approx 0.25 \pm 0.07$, has been recently computed by Micela et al. (1991) in the course of an extensive analysis of IPC soft X-ray background. This mean value has been adopted in the following.

We point out now that the technique we propose—as discussed in detail in §§ 2.2 and 2.3—does not require the restrictive assumption that the errors in the background counts are negligible (Maccacaro et al. 1987; Mereghetti & Garilli 1987). We assume only that (1) the counts collected at *any* position in the detector are Poisson distributed, and (2) the instrumental properties are constant with time, or, that all temporal changes of instrument properties have been properly corrected, as done in the final processing of IPC data. This more general framework allows us to include in the analysis upper bounds to the flux that are (and had to be) computed "Locally" on the basis of the image background evaluated according to equation (3) (for more details, see Harnden et al. 1984). Notice that because of the presence of entrance-support structure partially shadowing part of the detector field of view, we have to exclude from variability analysis all those detections and upper bounds whose count collection cells fall, also partially, in the "shadowed" region where count collection efficiency is reduced in an unknown way. Hence in the following only completely "unobscured" positions will be considered.

2.1. *Background Intensity Known with Negligible Uncertainty*

Let us consider an (X-ray) source observed N times. The ith observation is characterized by the exposure time (T_i), by the number of counts collected in the detection cell at source position (C_i), and by the number of background counts collected at source position (B_i) derived from a predetermined background map known with negligible uncertainty. The instrumental and detection algorithm properties are described by correction factors that account for the loss of collection efficiency due to mirror scattering ($\phi_{\text{MS},i}$) and mirror vignetting ($\phi_{\text{vign},i}$), and for the finite extent of detection cell size ($\phi_{\text{PRF},i}$). If the source were constant, the "true" source rate R_S remains unchanged in each observation. The expected number of counts in the ith observation, $\tilde{C}_i$, as function of R_S, is

$$\tilde{C}_i = \phi_i T_i R_S + B_i \,, \tag{8}$$

where the overall correction factor ϕ_i is defined as $\phi_i = (\phi_{\text{PRF},i}\, \phi_{\text{MS},i}\, \phi_{\text{vign},i})$ and accounts for the reduction of time spent in "effectively" observing the source due to reduction of sensitivity for all instrumental effects.

According to the Poisson statistics, the probability $P_i(C_i)$ to collect C_i counts in the ith observation is

$$P_i(C_i) = \frac{(\phi_i T_i R_S + B_i)^{C_i}}{C_i!}\, e^{-(\phi_i T_i R_S + B_i)} \,. \tag{9}$$

The probability to observe $\{C_1, C_2, \ldots, C_N\}$ counts in N distinct observations of the same objects is

$$P_{1,2,\ldots,N} = \prod_{i=1}^{N} \frac{(\phi_i T_i R_S + B_i)^{C_i}}{C_i!}\, e^{-(\phi_i T_i R_S + B_i)} \,. \tag{10}$$

The value $\tilde{R}_S$ that maximizes equation (10) is the ML estimate of R_S. Taking the logarithm of equation (10) and equating to zero the derivative of the resulting expression we get

$$\sum_{i=1}^{N} \frac{\phi_i T_i [C_i - (\phi_i T_i R_S + B_i)]}{\phi_i T_i R_S + B_i} = 0 \,, \tag{11}$$

The solution of this equation gives the searched ML rate estimator, $\tilde{R}_S$. Only in the *special* case in which one can assume both $\phi_{\text{PRF},i}$ and $\phi_{\text{MS},i}$ independent of observation, and negligible contribution of particle background (this implies that the background rate $R_B = \phi_{\text{vign}} R_{B_{\text{X-ray}}}$) the solution, $\tilde{R}_{0,S}$, to this problem can be expressed in an explicit form, namely:

$$\tilde{R}_{0,S} = \frac{\sum_{i=1}^{N} C_i - \sum_{i=1}^{N} B_i}{\phi_{\text{PRF}}\, \phi_{\text{MS}} \sum_{i=1}^{N} T_i \phi_{\text{vign},i}} \,. \tag{12}$$

While this is *not* the general solution, however, it furnishes a valuable starting point for searching the exact solution of equation (11) that had to be performed iteratively, adopting for example standard numerical methods, such as the Van

[1] This case is quite similar to that already considered by Maccacaro et al. (1987); however, as outlined in § 2.4, our approach in assessing source variability and variation amplitude is slightly different from that adopted by those authors.

[2] For a more detailed discussion of IPC background components and their relative importance, the reader is referred to Harnden et al. (1984), Harris et al. (1992), Micela et al. (1991).

Wijgaarden–Dekker-Brent methods (for a description see Press et al. 1976).

To go further, we assume that the ML rate derived solving equation (11) *is* the "true" source intrinsic rate and substitute $\tilde{R}_S$ for R_S in equation (8). We then test, at a given chosen significance level, for the ith observation the "null hypothesis" that the observed counts C_i are compatible with being drawn from a Poisson-like distribution with mean value $\tilde{C}_i$. This test is easily accomplished because the probability to measure C_i counts is given by equation (9) when R_S takes the value $\tilde{R}_S$. If the computed probability does not meet the chosen acceptance criterion one can conclude that, in the ith observation, C_i is incompatible with the predicted value $\tilde{C}_i$, i.e., the collected counts have undergone a fluctuation (positive or negative) that is incompatible, at the chosen significance level, with the statistics of the counting process; therefore we are in presence of a "real" variation of observed rate. In this case the measured variation of C_i is entirely due to change in source rate (by definition, in this case, the background rate does not change). From equation (8), noting that for each observation $C_i = \phi_i R_{S,i} + B_i$, where $R_{S,i}$ is the actual source rate in the ith observation, it follows that

$$C_i - \tilde{C}_i = (R_{S,i} - \tilde{R}_S)\phi_i T_i = \Delta R_i \phi_i T_i \,. \tag{13}$$

Where the value of C_i turns out to be compatible, at the chosen confidence level, with the value of $\tilde{C}_i$, one can determine the upper limit to the amplitude of undetected variations. In fact, for each given confidence level, one can determine from equation (9) two values C_i^+ and C_i^-, namely the number of counts required to detect, at the chosen significance level, brightening or dimming of the source, respectively. Substituting C_i^+ and C_i^- for C_i in equation (13) one can derive the searched upper limits to brightening and dimming variability amplitude, respectively. We want to stress the importance of this step because it is the capability to determine upper limits—hence the attained sensitivity to amplitude variation—that it is crucial for the further steps of the proposed methodology (see §§ 2.4 and 3). We note in passing that, when one is in presence of a "variability detection," it is obvious whether it is a source brightening or dimming that one is dealing with.

2.2. *Background Intensity Known with Finite Accuracy*

The second case (that, for instance, applies to the *Einstein* Rev-1 Local determination of cell and background counts) is more elaborate because it implies two distinct measurements (see eqs. [1] and [2]). In this case the inference on the (X-ray) variability of the source requires us also to take into account possible fluctuations of the background component. In fact, the collected counts at source position are the sum of "source" counts S and "background" counts B, and only if one can separate the effect of background fluctuations (or real temporal variation, if any) on intrinsic source rate variability, one can determine the characteristics of variability of the source. The expected numbers of counts in the detection cell $\tilde{C}_i$ and the expected number of counts $\tilde{F}_i$ in a larger cell containing the detection cell (see Fig. 1) can be written in term of the source rate and background rate (R_B) as

$$\tilde{C}_i = T_i(\phi_i R_S + \kappa_i R_B) \,, \tag{14a}$$

$$\tilde{F}_i = T_i(\beta_i' R_S + \eta_i' R_B) \,, \tag{14b}$$

where $\phi_i = \phi_{\mathrm{PRF},i}\,\phi_{\mathrm{MS},i}\,\phi_{\mathrm{vign},i}$, $\kappa_i = [(1 - f_B)\phi_{\mathrm{vign},i} + f_B]$, $\beta_i' = \beta\phi_{\mathrm{MS},i}\,\phi_{\mathrm{vign},i}$, and $\eta_i' = \eta_i \kappa_i'$. The two measured quantities F and C are independent, hence, in a given set of observations, the probability of getting $\{F_1, F_2, \ldots, F_N\}$ and $\{C_1, C_2, \ldots, C_N\}$ counts can be written as the product of Poisson distributions:

$$P_{1,2,\ldots,N}(C, F) = \prod_{i=1}^{N} \frac{[T_i(\phi_i R_S + \kappa_i R_B)]^{C_i}}{C_i!}\, e^{-T_i(\phi_{\mathrm{PRF}} R_S + \kappa_i R_B)} \times \prod_{i=1}^{N} \frac{[T_i(\beta_i' R_S + \eta_i' R_B)]^{F_i}}{F_i!}\, e^{-T_i(\beta_i' R_S + \eta_i' R_B)} \,. \tag{15}$$

An explicit expression for computing the ML rate $\tilde{R}_S$ and $\tilde{R}_B$ maximizing equation (15) cannot *generally* be derived. In the *restricted class* of cases in which ϕ_{PRF} and ϕ_{MS} (hence β and η) are independent on ith observation and the contribution of particle background is negligible the search solution, $\tilde{R}_{0,S}$, $\tilde{R}_{0,B}$, can be expressed in explicit forms, as follows:

$$\tilde{R}_{0,S} = \frac{1}{(\phi_{\mathrm{PRF}} - \beta/\eta)} \frac{\sum_{i=1}^{N} C_i - 1/\eta \sum_{i=1}^{N} F_i}{\phi_{\mathrm{MS}} \sum_{i=1}^{N} T_i \phi_{\mathrm{vign},i}} \,, \tag{16a}$$

$$\tilde{R}_{0,B} = \frac{\sum_{i=1}^{N} F_i - (\beta/\phi_{\mathrm{PRF}}) \sum_{i=1}^{N} C_i}{(\eta - \beta/\phi_{\mathrm{PRF}}) \sum_{i=1}^{N} T_i \phi_{\mathrm{vign},i}} \,. \tag{16b}$$

These last equations are formally identical to equations (3a)–(3b), where C and F go into

$$\sum_{i=1}^{N} C_i \quad \text{and} \quad \sum_{i=1}^{N} F_i \,,$$

respectively, and where the total observing time is substituted with

$$\sum_{i=1}^{N} T_i \phi_{\mathrm{vign},i} \,,$$

hence we conclude that, in these restricted class of cases, the ML estimate of source rate is equivalent to sum up all available observations taking into account exposure times and positions in the field of view. The above solution does not hold in general cases; however, $\tilde{R}_{0,S}$ and $\tilde{R}_{0,B}$ are valuable starting points for the iterative numerical method (minimization of a two-variable function) we adopted for finding, in general case, the searched exact solutions, $\tilde{R}_B$ and $\tilde{R}_S$.

The expected number of source counts falling in the detection cell is

$$\tilde{S}_i = \tilde{R}_S T_i \phi_{\mathrm{MS},i}\, \phi_{\mathrm{vign},i} \,, \tag{17}$$

and it should be compared with the actual number of source count S_i—given by equation (3a)—to test, at the given chosen significance level, the "null hypothesis:" that $\tilde{S}_i$ is compatible with S_i.

The random variable S, whose expectation in the ith observation, $\tilde{S}_i$, is given by equation (17), has a rather complicated probability distribution. Apart from a constant factor, $(\phi_{\mathrm{PRF}} - \beta/\eta)^{-1}$—dependent on instrument point response function and constants specific of the detection algorithm—S is distributed as the difference of C, a Poisson distributed variable, and F/η, a Poisson distributed variable scaled by the area ratio of detection and frame cell. Neglecting the constant factor, one has to deal with the distribution function $G(H, \tilde{H})$, of a random variable

$$H = C - F/\eta \,, \tag{18}$$

with mean value $\tilde{H} = \tilde{C} - \tilde{F}/\eta$. The mean values $\tilde{C}$ and $\tilde{F}$ are computed substituting in equations (14a) and (14b) the values $\tilde{R}_S$ and $\tilde{R}_B$ for R_S and R_B, respectively. Because C and F can assume only integer values, the distribution function $G(H, \tilde{H})$

is defined only on a discrete number of H values. In the following the set of these values is indicated as $[H]$.

The knowledge of the explicit functional form of $G(H, \tilde{H})$ is not required to compute the probability—hence the significance level—at which a given value H can be drawn from the underlying distribution $G(H, \tilde{H})$. This requires the evaluation of the lower and upper percentage points, H_{α_-} and H_{α_+}, respectively, verifying the condition

$$\sum_{[H\alpha_+]} G(H, \tilde{H}) = \alpha_+ , \tag{19a}$$

$$\sum_{[H\alpha_-]} G(H, \tilde{H}) = \alpha_- , \tag{19b}$$

where $[H_{\alpha_+}]$ is the set of H values defined by the condition $H \geq C - F/\eta$ and $[H_{\alpha_-}]$ is the set of H values defined by the condition $H \leq C - F/\eta$. It is noteworthy that the domain to which each sum extends it depends also on the shape of the distribution $G(H, \tilde{H})$. The need for two distinct equations derives from the asymmetry of the probability distribution $G(H, \tilde{H})$ and from the interest in searching for fluctuations both in excess (brightening) and in defect (dimming) with respect to the source "true" rate. H_{α_+} and H_{α_-} are computed noting that F and C can assume only integer values, are independent, and have expectation $\tilde{F}$ and $\tilde{C}$. Hence the probability, $P(H)$, to get the value H, when one expects $\tilde{H}$, can be computed summing up the probability $P(C, F)$ of getting C and F counts in a given observation and extending the sum to the entire domain of C, F pairs compatible with the attained H value, namely:

$$P(H) = \sum_{i,j} \left(\frac{e^{-\tilde{C}} \tilde{C}^i}{i!} \frac{e^{-\tilde{F}} \tilde{F}^j}{j!} \right) . \tag{20}$$

When $H > \tilde{H}$, the sum has to be extended to all the pairs of values of C, F fulfilling the condition $H \geq (C - F/\eta)$. In this case the above expression gives the probability to collect from a constant source a number of counts greater or equal to H when one expects to measure $\tilde{H}$ counts. Hence $1 - P(H > \tilde{H})$ gives the significance level to reject the hypothesis of source being constant with respect to a putative brightening variation. For $H < \tilde{H}$ the sum has to be extended to all the pairs of values C, F fulfilling the condition $H \leq (C - F/\eta)$. In this case $1 - P(H < \tilde{H})$ gives the significance level to reject the hypothesis of source being constant with respect to a putative dimming variation. While, in principle, the number of pairs on which to extend the sum is infinite, it is obvious that only a finite fraction of these pairs need to be considered, namely those within few "standard deviations" from the expected values $\tilde{C}$ and $\tilde{F}$, because the contributions of all other pairs to the sum will be negligible. We have verified the correctness of this approach comparing the $H_{\alpha_\pm}$ values derived from application of equation (20) with those derived through Monte Carlo simulation performed on a grid of 35 values of $\tilde{F}$, ranging from 1 to 4000, and 34 values of $\tilde{C}$, ranging from 1 to 3000. This comparison has been performed for $\eta = 25/9$, i.e., the value adopted in the IPC standard data processing system. For each pair $(\tilde{C}, \tilde{F})$ we have generated 10^5 pairs (C, F) of Poisson distributed variables and computed the corresponding 10^5 values of H. From each of these sets of realizations of the $G(H, \tilde{H})$ distribution function we have determined the H_{α_+} values for several α_+ values ranging between 0.8413 (1 σ point of one-tail Gaussian integral distribution) and 0.99997 (4 σ point of one-tail Gaussian integral distribution), and the corresponding values of H_{α_-}.

The Monte Carlo approach cannot be adopted as the general solution to the problem of computing the $H_{\alpha_\pm}$ values, because it requires significant computer resources (the above simulations required $\sim$20 hours of CPU time on a $\sim$10 MIPS cpu). Moreover, it has three clear disadvantages: (*a*) the number of $(\tilde{C}, \tilde{F})$ pairs on which to extend the grid is necessarily finite, and any pair not exactly present in the grid requires an interpolation; (*b*) very high significance levels, i.e., above 3.5 σ, cannot easily be reached since this would require a very large number of simulations; (*c*) more important, the results are dependent from the adopted value for η and a new set of simulations has to be performed to determine the H_α's for any new distinct value of η. This is particularly disadvantageous because it is conceivable that, for a given detector, the ratio between detection and frame cells could be varied to account for the change of detector point response function width across the field of view.[3] In such a case several values of η should (or could) be used for the same detector and detection technique making more difficult the managing of data.

The significance level computed according to equation (20) allows us to determine whether the source counts S_i undergo a significant fluctuation with respect to the expected number of counts $\tilde{S}_i$. The amplitude of detected variability is computed from equation (3a) as

$$\Delta R_i = R_{S,i} - \tilde{R}_{S,i} = \frac{S_i - \tilde{S}_i}{T_i \phi_{\mathrm{MS},i} \phi_{\mathrm{vign},i}}$$

$$= \frac{1}{(\phi_{\mathrm{PRF}} - \beta/\eta)} (C_i - \tilde{C}_i) - \frac{1}{\eta} (F_i - \tilde{F}_i) . \tag{21}$$

The determination of upper limits to the variability amplitude in this case is more involved to derive. The upper limits are computed substituting in equation (21), S_i with S_i^+ and S_i^-, i.e., the number of counts required to obtain a statistical significant source brightening or dimming, respectively. The evaluation of S_i^+ and S_i^- requires to search for the zeros of equations (19a) and (19b), respectively. Because the distribution $G(H, \tilde{H})$ is known only in an implicit form, this search needs to be performed numerically.

2.3. *Mixed Cases*

Finally we need to consider the most general case, i.e., when for a given X-ray source we have both Local and Map measurements. This usually occurs considering together upper bounds ("Locally" evaluated) and Map detections. There are also cases in which the Map detection method cannot be applied[4] (for more details on the IPC data see Harnden et al. 1984 and Harris et al. 1992) and in such a case one needs to consider also Local detections.

Let us consider the case with N measurements, N_{L} obtained with the Local method and N_{M} with the Map method. In this case the Local measurements can be used as described in § 2.2 to determine the ML estimate of background rate. Adopting this estimate, $\tilde{R}_{B_{\mathrm{L}}}$, as the "true" value of R_B one can determine the ML estimate of source rate $\tilde{R}_S$ maximizing the conditional

[3] Indeed this kind of varying size moving window source detection algorithm has already been implemented in the *ROSAT* PSPC data processing system.

[4] This generally occurs because the measured background is much higher than the expected one, indicating the presence of a background component not modeled by the background map.

probability

$$P(C_1, \ldots, C_N | B_1, \ldots, B_{N_M}, \tilde{R}_{B_L}) = \prod_{i=1}^{N_M} \frac{(\phi_i T_i R_S + B_i)^{C_i}}{C_i!} e^{-(\phi_i T_i R_S + B_i)} \times \prod_{i=1}^{N_L} \frac{[T_i(\phi_i R_S + \kappa_i \tilde{R}_{B_L})]^{C_i}}{C_i!} e^{-T_i(\phi_i R_S + \kappa_i \tilde{R}_{BL})} \times \prod_{i=1}^{N_L} \frac{[T_i(\beta_i' R_S + \eta_i' \tilde{R}_{B_L})]^{F_i}}{F_i!} e^{-T_i(\beta_i' R_S + \eta_i' \tilde{R}_{BL})} . \quad (22)$$

to get $\{C_1, \ldots, C_N\}$ counts when the number of Map background counts $\{B_1, \ldots, B_{N_M}\}$ are given and the Local background rate is $\tilde{R}_{B_L}$. Again this problem in the general case needs to be solved numerically. However, in the *restricted* class of problems in which both ϕ_{MS} and ϕ_{PRF} are independent from position in the field of view and the contribution of nonimaged background is negligible, we can obtain a solution, $\tilde{R}_{0,S}$, $\tilde{R}_{0,B_L}$, in explicit form:

$$\tilde{R}_{0,S} = \frac{\sum_{i=1}^{N_M} C_i - \sum_{i=1}^{N_M} B_{M,i}}{\phi_{PRF}\, \phi_{MS} \sum_{i=1}^{N} T_i \phi_{vign,i}} + \frac{\sum_{i=1}^{N_L} C_i - (1/\eta) \sum_{i=1}^{N_L} F_i}{(\phi_{PRF} - \beta/\eta)\phi_{MS} \sum_{i=1}^{N} T_i \phi_{vign,i}} . \quad (23a)$$

Together with the expression for $\tilde{R}_{0,B_L}$:

$$\tilde{R}_{0,B_L} = \frac{\sum_{i=1}^{N_L} F_i - (\beta/\phi_{PRF}) \sum_{i=1}^{N_L} C_i}{(\eta - \beta/\phi_{PRF}) \sum_{i=1}^{N_L} T_i \phi_{vign,i}} . \quad (23b)$$

For the purpose of testing the occurrence of source rate variations, one has to distinguish the cases of Map deduced rates from those of Local deduced rates. In the first case one has to follow the recipe given in § 2.1. In the second case one has to adopt the approach discussed in § 2.2. Obviously in these last cases the attainable sensitivity is lower because of the uncertainty introduced by the local background determination.

2.4. *Final Considerations*

It is important to note some key aspects of the proposed technique:

1. Our procedure allows us to use all available observations of a given celestial object, both in the form of (X-ray) detections and upper bounds; indeed the only quantities required are cell counts B, C, F, exposure times, and correction factors.

2. For each given observation it is possible to test for the occurrence of variability and to define the variability amplitude as

$$\frac{\Delta R_i}{\tilde{R}} = \frac{\tilde{R} - R_i}{\tilde{R}} , \quad (24)$$

where $\tilde{R}$ is the ML source rate and R_i is the count rate measured in the ith observation.

3. Because sensitivity to detection of variability is dependent on counting statistic and, in principle, could be different in different observations it is of importance to evaluate the sensitivity thresholds (i.e., upper bounds) to the variability amplitude. This is computed as

$$\frac{\Delta R_{thr,i}}{\tilde{R}} = \frac{\tilde{R} - R_{thr,i}}{\tilde{R}} \quad (25)$$

where $R_{thr,i}$ is the threshold rate for the ith observation at a given chosen significance level.

In this way each observation can be characterized by a variability amplitude, if the source meets the criterion for being variable, or by an upper limit to variability amplitude if the source does not meet the criterion for "variability detection."

The adoption of techniques to build ML integral distributions in presence of upper bounds (Avni et al. 1980; Schmitt 1985; Feigelson & Nelson 1985) allows us to use all available information to construct the ML amplitude variability integral distribution for a given class of sources.

We propose to use this kind of distribution, together with the distribution of time separations between pairs of observations of same objects, to statistically characterize variability properties of well-defined classes of sources and to compare in an objective way the "variability behavior" of different classes of sources. The proposed approach is analogous to that adopted in studying the X-ray emission of several classes of stars; in this case, *after having taken into account the limiting sensitivity of observations*, one compares X-ray luminosity functions of different classes of sources, and infers conclusions on typical X-ray emission level, occurrence of a tail of population at higher luminosity, etc. (compare the reviews of Rosner, Golub, & Vaiana 1985; Vaiana & Sciortino 1987; Vaiana 1990 and references therein cited). In the context of variability studies we want to infer conclusions on frequency of variations of a given amplitude, on the sensitivity of available observations, and on the range of time scale searched for variability occurrence.

To conclude this section we note in passing that the approach of properly "merging" all available data, allows us, in principle, to increase the signal-to-noise ratio for the detection of a given celestial object, and in some cases could allow us to discover faint emitters, otherwise undetected.

2.5. *Comparisons with Other Similar Techniques*

A similar technique has been developed by Maccacaro et al. (1987). Their method uses Poisson distribution of detection counts as well as upper limits; however, it does not allow to evaluate sensitivity to variability detection. This limits the capability to obtain a statistical description of variability occurrence in a sample of X-ray sources such as the distribution variability functions that we can estimate with the new proposed approach. We have checked that individual results of Maccacaro et al. (1987) are usually consistent with those derived with our technique, namely, sources that result to be variable with their approach are also variable with our technique, and sources are seen "constant" with both techniques. The approach presented here and that adopted by Maccacaro et al. (1987) differ in the way of determining the "true" source count rate. In the Maccacaro et al. (1987) method the assumed "true" rate for a source seen twice is fixed by the condition that the probability to see deviations greater than or equal to those observed assumes the same value in the two observations; if this probability level is lower than the chosen acceptance threshold, the source is considered variable, otherwise it is considered constant. The method presented in this paper assumes a ML estimate for the "true" source rate, i.e., the deduced rate is weighted according to the statistical quality of the observations. Count rates for a "constant" source evaluated with the two methods differ usually by a few percent; the difference is larger in those cases in which the statistical quality

(i.e., the statistical weight) of individual observations is very different. In such a case our approach gives a "true" rate that is nearer to the observation(s) with higher count statistics.

Notice that this comparison is necessarily restricted only to the data gathered with the Map method, in fact the Maccacaro et al. (1987) approach can treat only this kind of data, and, differently from our technique, cannot be applied to measurements where the background counts are known with finite uncertainty.

3. AN APPLICATION

In this section we intend to elucidate the capability of our method, presenting its application to a sample of late-type stars belonging to the Hyades and to the Pleiades open clusters.

We use data taken with the Imaging Proportional Counter (IPC) on board of *Einstein Observatory*, reduced with the standard IPC Rev-1 processing (Harnden et al. 1984, 1990); the final processing results are organized in a easily retrievable way in the Stellar *Einstein* X-ray data base resident on SUN™ computers at the Center for Astrophysics (Cambridge, MA), and on DEC™ computers at Osservatorio Astronomico di Palermo (Sciortino et al. 1988; Harnden et al. 1990).

The Hyades and the Pleiades stars are from the most recent complete *Einstein Observatory* surveys of these clusters (Micela et al. 1988, 1990). The Hyades and Pleiades samples consist of the 14 multiply observed stars with $L_x > 3 \times 10^{29}$ ergs s^{-1}, and of the 14 multiply observed stars with $L_x > 3 \times 10^{28}$ ergs s^{-1}, respectively. The subdivision according to spectral type is summarized in Table 1. We note that four Hyades stars and eight Pleiades stars have been observed 3 times or more. Exposure times for the IPC images containing the Hyades and the Pleiades stars considered here range between 10^3–1.3×10^4 and 10^3–1.5×10^4 s, respectively. Most of the observations of the Pleiades stars cover the time separation of few days. In contrast, the Hyades observations are distributed on a wider range of time separations.

TABLE 1

COMPOSITION OF HYADES AND PLEIADES SAMPLES

	Hyades		Pleiades	
Spectral Type[a]	Stars[b]	Observations	Stars	Observations
dF	1	3	1	3
dG	7 (2)	17 (3)	5	16
dK	3	6	7	17
dM	2	4	1 (1)	2 (1)
gK	1 (1)	2(1)	...	...

[a] Spectral-type subdivision for main-sequence stars is based on color index: dF corresponds to the 0.3–0.5 $(B-V)_0$ range, dG to the 0.5–0.8 range, dK to the 0.8–1.45 range, and dM to $(B-V)_0$ greater than 1.45.

[b] In parentheses we indicate the number of stars and number of observations with detected variability.

Applying our technique to these two samples of data, we have been able to detect significant—at significance level greater than the one-tail Gaussian 3 σ level—brightening variability in $\sim$20% of the Hyades stars (see Fig. 2*a*, solid line) and in $\sim$10% of the Pleiades stars (see Fig. 2*a*, dotted line). Due to the paucity of the involved stars these two distributions cannot be distinguished. It is noteworthy that in the case of the Pleiades the sensitivity to variations (Fig. 2*a*, short-dashed line) is significantly—at a significance level of 99.5% according to Kolmogorov-Smirnov test—lower than for the Hyades (Fig. 2*a*, long-dashed line); indeed, variability amplitudes similar to those detected in the Hyades could have occurred in the

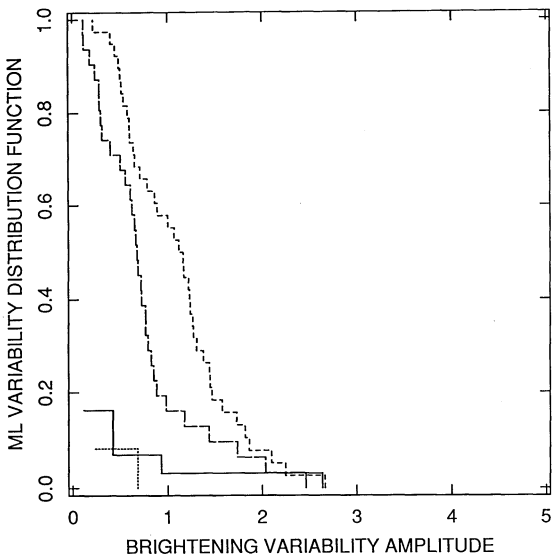

FIG. 2*a*

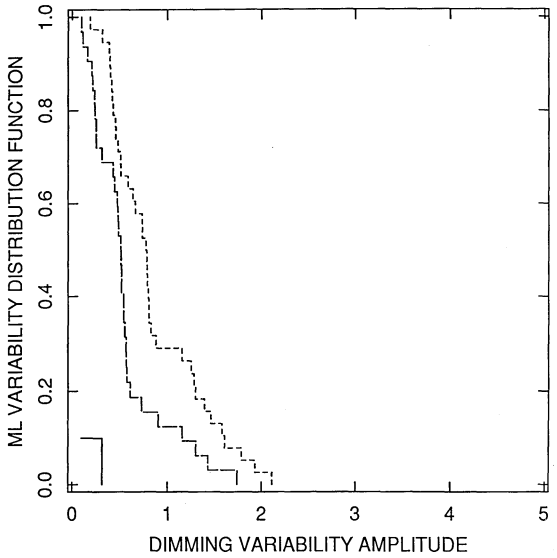

FIG. 2*b*

FIG. 2.—(*a*) ML brightening variability distribution function for the Hyades late-type stars (*solid line*) and the Pleiades late-type stars (*dotted line*) compared with the ML brightening variability threshold distribution for the Pleiades late-type stars (*short-dashed line*) and for the similar threshold distribution function for the Hyades (*long-dashed line*). The threshold distribution functions are marginally statistically distinguishable (significance level 99.5%). Only one dM Pleiades star has been recognized as variable according to the adopted 3 σ significance level; however, the sensitivity to variability attainable with present data does not exclude that the Pleiades stars could undergo long-term variations of amplitude similar (and even greater) to that detected in the Hyades stars. (*b*) Like (*a*), but for dimming variability. In this case the threshold distribution functions for the Hyades and Pleiades stars are statistically distinguishable (significance level 99.8%). In this case only one Hyades star shows detectable dimming variability; however, available data do not rule out the presence of similar (and even greater) amplitude variation in the Pleiades stars.

Pleiades and pass undetected due to the lower count statistics of these more distant stars. A similar consideration holds for the case of dimming variability (Fig. 2*b*). In this case the distribution of sensitivities to variations for the Pleiades can be better distinguished from that of the Hyades (significance level of 99.8%), and only a single Hyades star has undergone a significant variation in a single observation. Hence, on the basis of this analysis of all presently available data, we conclude that the Pleiades and the Hyades stars, at least with the limited data and the limited sensitivity they allow us to reach, show an indistinguishable variability behavior in the X-ray band.

The above example should give a clear perception of the need for evaluating upper limits to variability amplitudes and of the capability of our methodology. We intend to present elsewhere a detailed account of its application to other stellar samples.

4. SUMMARY AND CONCLUSION

We have presented a new approach for studying long-term variability of low count-rate sources observed more than once with a photon-counting imaging detector. This new method allows us to detect the occurrence of statistical significant temporal variations, to derive variability amplitude upper bounds in the cases of undetected variability, and to characterize variability properties of a class of sources in terms of the ML variability distribution functions. It is noteworthy that our approach allows us to compare in an *objective* way the occurrence of temporal variations in different classes of sources by making use of all available information.

We have illustrated the capability of the method through its application to X-ray data, gathered with the *Einstein Observatory* Imaging Proportional Counter, of stellar sources in the Hyades and Pleiades open clusters.

Our analysis of the presently available, limited sensitivity, Hyades and Pleiades data has allowed us to conclude that the variability behavior in the soft X-ray of these two stellar samples cannot be distinguished.

We conclude by noting that the new data gathered with the *ROSAT* PSPC and *ROSAT* Wide Field Camera will be an elective area of application for the proposed methodology.

We wish to acknowledge extensive discussion with A. Collura and G. S. Vaiana, and many suggestions of S. Serio that have improved the presentation. We thank the referee, T. Subba Rao, whose suggestions have helped us to clarify our work. This work has been partially supported by Ministero della Università e della Ricerca Scientifica e Tecnologica, and by Agenzia Spaziale Italiana.

REFERENCES

Ambruster, C., Sciortino, S., & Golub, L. 1987, ApJS, 65, 273
Avni, Y., Soltan, A., Tananbaum, H., & Zamorani, G. 1980, ApJ, 238, 800
Collura, A., Pasquini, L., & Schmitt, J. H. M. M. 1988, A&A, 205, 197
Collura, A., Sciortino, S., Serio, S., Vaiana, G. S., Harnden, F. R., Jr., & Rosner, R. 1989, ApJ, 338, 296
Collura, A., Reale, F., & Peres, G. 1990, ApJ, 356, 119
Elvis, M., ed. 1990, Imaging X-Ray Astronomy (Cambridge Univ. Press)
Feigelson, E. D., & Nelson, P. 1985, ApJ, 293, 192
Giacconi, R., et al. 1979a, ApJ, 230, 540
———. 1979b, ApJ, 234, L1
Giommi, P., Tagliaferri, G., & Angelini, L. 1988, Mem. Soc. Astron. Ital., 59, 33
Gorenstein, P., Harnden, Jr., F. R., & Fabricant, D. G. 1981, I.E.E.E. Trans. Nucl. Sci., NS-28, 869
Harnden, F. R., Jr., Fabricant, D. G., Harris, D. E., & Schwarz, J. 1984, SAO Spec. Rept. 393
Harnden, F. R., Jr., Sciortino, S., Micela, G., Maggio, A., Vaiana, G. S., Schmitt, J. H. M. M., & Rosner, R. 1990, in Imaging X-Ray Astronomy, ed. M. Elvis (Cambridge Univ. Press), 313
Harris, D. E., & Irwin, D., ed. 1984, *Einstein Observatory* Revised User's Manual
Maccacaro, T., Garilli, B., & Mereghetti, S. 1987, AJ, 93, 1484
Mereghetti, S., & Garilli, B. 1987, AJ, 94, 1616
Micela, G., Harnden, F. R., Jr., Rosner, R., Sciortino, S., & Vaiana, G. S. 1991, ApJ, 386, 495
Micela, G., Sciortino, S., Vaiana, G. S., Schmitt, J. H. M. M., Harnden, F. R., Jr., & Rosner, R. 1990, ApJ, 348, 557
Micela, G., Sciortino, S., Vaiana, G. S., Schmitt, J. H. M. M., Stern, R. A., Harnden, F. R., Jr., & Rosner, R. 1988, ApJ, 325, 798
Montmerle, T., Koch-Miramond, L., Falgarone, E., & Grindlay, J. E. 1983, ApJ, 269, 182
Osborne, J. P. 1988, Mem. Soc. Astron. Ital., 59, 117
Pallavicini, R. 1988, Mem. Soc. Astron. Ital., 59, 71
Pallavicini, R., Tagliaferri, G., & Stella, L. 1990, A&A, 228, 403
Parmar, A. N., & White, N. 1988, Mem. Soc. Astron. Ital., 59, 147
Peres, G., Reale, F., Collura, A., & Fabbiano, G. 1989, ApJ, 336, 140
Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1986, Numerical Recipes (Cambridge Univ. Press)
Primini, F. A., Murray, S. S., Huchra, J., Schild, R., Burg, R., & Giacconi, R. 1991, ApJ, 374, 440
Rosner, R., Golub, L., & Vaiana, G. S. 1985, ARA&A, 23, 413
Schmitt, J. H. M. M. 1985, ApJ, 293, 176
Sciortino, S., Harnden, Jr., F. R., Maggio, A., Micela, G., Vaiana, G. S., Rosner, R., & Schmitt, J. H. M. M. 1988, in ESO Conf. Workshop Proc. 28, Astronomy from Large Database, ed. F. Murtagh & A. Heck (Gerching bei Munchen: ESO)
Snow, T. P., Jr., Cash, W., & Grady, C. A. 1981, ApJ, 244, L19
Stella, L. 1988, Mem. Soc. Astron. Ital., 59, 185
Vaiana, G. S. 1990, in Imaging X-ray Astronomy, ed. M. Elvis (Cambridge Univ. Press), 61
Vaiana, G. S., & Sciortino, S. 1987, in IAU Symp. 122, Circumstellar Matter, ed. I. Appenzeller & C. Jordan (Dordrecht: Reidel), 333
White, N. E., Culhane, J. L., Parmar, A. N., Kellet, B. J., Kahn, S., van den Oord, & Kuijpers, J. 1986, ApJ, 301, 262
White, N. E., Culhane, J. L., Parmar, A. N., & Sweeney, M. A. 1987, MNRAS, 227, 545
White, N., & Peacock, A. 1988, Mem. Soc. Astron. Ital., 59, 7
White, N. E., Shafer, R. A., Horne, K., Parmar, A. N., & Culhane, J. L. 1990, ApJ, 368, 776
Zamorani, G., Giommi, P., Maccacaro, T., & Tananbaum, H. 1984, ApJ, 278, 28