*Astronomical Data Analysis Software and Systems XIII*
*ASP Conference Series, Vol. 314, 2004*
*F. Ochsenbein, M. Allen, and D. Egret, eds.*

# Source Detection with Bayesian Inference on ROSAT All-Sky Survey Data Sample

Fabrizia Guglielmetti, Rainer Fischer, Volker Dose

*Centre for Interdisciplinary Plasma Science, Garching, Germany D-85748*

Wolfgang Voges, Guenter Boese

*Max-Planck-Institut für Extraterrestrische Physik, Garching, Germany D-85748*

**Abstract.** We employ Bayesian inference for the joint estimation of sources and background on ROSAT All-Sky Survey (RASS) data. The probabilistic method allows for detection improvement of faint extended celestial sources compared to the Standard Analysis Software System (SASS). Background maps were estimated in a single step together with the detection of sources without pixel censoring. Consistent uncertainties of background and sources are provided. The source probability is evaluated for single pixels as well as for pixel domains to enhance source detection of weak and extended sources.

## 1. Introduction

The RASS x-ray data were largely analyzed with the Standard Analysis Software System (SASS). The results can be found in the Bright and Faint Source catalogues (Voges et al. 1999, 2000). Nevertheless, SASS is known for lack of sensitivity for faint or extended sources. This is due to the sliding window technique which locally searches for count enhancements relative to the intensity in a surrounding area defining the background intensity. In multiple steps the window width is changed to allow for the detection of extended sources. But faint extended sources and blended faint sources in crowded fields may get lost. One reason is due to the local estimation of the background in a small region around the sliding window which may provide only poor signal-to-noise ratios (S/N). The sources are characterized by fitting the candidate sources in a further step using a Maximum-Likelihood (ML) method (Boese & Doebereiner 2001). The ML method works properly on prominent point-like sources which account for 94% of the sources published in the RASS Bright Source catalogue. The characteristics of the faint sources may not be properly estimated in case of faint extended sources (Voges et al. 1999).

The present method using Bayesian Probability Theory (BPT) estimates the background and sources in a single step neither employing pixel censoring nor using a sliding window technique. The aim is to infer simultaneously a background map for the complete field size ($6.4° \times 6.4°$ in the sky) and a probability for having source intensity in addition to the background intensity in a pixel cell or pixel domain. Bayesian inference allows reasoning on the basis of sparse data employing additional information independent of the data. The results are given by probability distributions quantifying our state of knowledge. For background estimation and source detection the additional information is the assumption that the background is smooth, e.g. spatially slowly varying compared to source dimensions. To allow for smoothness the background is modelled with a bivariate Thin-Plate spline. The coexistence of background and sources is described with a probabilistic two-component mixture model where one component describes background contribution only and the other component describes background plus signal contributions. Each pixel cell (or pixel domain) is characterized by the probability of belonging to one of the two mixture components. For the background spline estimation the photons contained in all pixel cells are considered including pixels containing additional source contributions.

This technique is applied on a data sample coming from the ROSAT PSPC in Survey Mode (0.1-2.4 keV). The ROSAT exposure map and the observatory's point spread function have been properly accounted for.

## 2. Method

Given the observed data set $D = \{d_{ij}\} \in \mathbf{N}_0$, where $d_{ij}$ expresses photon counts in a pixel cell $ij$, two complementary hypotheses arise:

$$\begin{cases} B_{ij} \rightarrow d_{ij} = b_{ij} + \epsilon_{ij} \\ \overline{B}_{ij} \rightarrow d_{ij} = b_{ij} + s_{ij} + \epsilon_{ij} \end{cases}$$

Hypothesis $B_{ij}$ specifies that $d_{ij}$ consists only of background $b_{ij}$ spoiled with noise $\epsilon_{ij}$. Hypothesis $\overline{B}_{ij}$ specifies the case where additional source intensity $s_{ij}$ contributes to the background.

Additional assumptions are that no negative values for signal and background amplitudes are allowed and that the background is smoother than the signal. This is achieved by modelling the background count rate with a bivariate Thin-Plate spline where the supporting points are chosen sparsely to ensure that sources can not be fitted. The spline fits the background component whereas count enhancements classify pixel (domains) with source contributions.

The likelihood distributions for the two hypotheses are

$$p(d_{ij} \mid B_{ij}, b_{ij}) = \frac{b_{ij}^{d_{ij}}}{d_{ij}!} e^{-b_{ij}}, \tag{1}$$

$$p(d_{ij} \mid \overline{B}_{ij}, b_{ij}, s_{ij}) = \frac{(b_{ij} + s_{ij})^{d_{ij}}}{d_{ij}!} e^{-(b_{ij}+s_{ij})}. \tag{2}$$

For background estimation we have to marginalize over the signal in eqn. 2 according to the sum rule of BPT. The prior distribution over the signal is chosen to be exponential, $p(s_{ij} \mid \lambda) = \exp\{-\frac{s_{ij}}{\lambda}\}/\lambda$, assuming we know only the average value of the signal intensity $\lambda$. The prior probability for the two hypotheses is chosen uninformative to be $p(B_{ij}) = p(\overline{B}_{ij}) = \beta = 0.5$.

Because we do not know if a certain pixel contains purely background or additional signal, the likelihood for the mixture model is

$$p(D \mid b, \lambda) = \prod_{ij} [\beta \cdot p(d_{ij} \mid B_{ij}, b_{ij}) + (1-\beta) \cdot p(d_{ij} \mid \overline{B}_{ij}, b_{ij}, \lambda)] \qquad (3)$$

where

$$p(d_{ij} \mid \overline{B}_{ij}, b_{ij}, \lambda) = \frac{e^{\frac{b_{ij}}{\lambda}}}{\lambda(1+\frac{1}{\lambda})^{d_{ij}+1}} \cdot \frac{\Gamma[(d_{ij}+1), b_{ij}(1+\frac{1}{\lambda})]}{\Gamma(d_{ij}+1)}$$

is the Poisson distribution marginalized over $s_{ij}$. The prior probability density function for $b_{ij}$ is chosen flat for positive values and 0 elsewhere. The posterior distribution is according to Bayes theorem proportional to the product of the mixture likelihood and the prior. Its maximum with respect to $b$ gives an estimate of the background map which includes the observatory's exposure map.

The probability of having source contribution in pixel cells or domains is

$$p(\overline{B}_{ij} \mid d_{ij}, b_{ij}, \lambda) = \frac{p(\overline{B}_{ij}) \cdot p(d_{ij} \mid \overline{B}_{ij}, b_{ij}, \lambda)}{p(\overline{B}_{ij}) \cdot p(d_{ij} \mid \overline{B}_{ij}, b_{ij}, \lambda) + p(B_{ij}) \cdot p(d_{ij} \mid B_{ij}, b_{ij})}. \qquad (4)$$

Details of mixture modelling in the framework of BPT can be found in von der Linden et al. (1999) and Fischer et al. (2000). In order to include the width of the instrumental PSF, a minimum of 3 by 3 pixels domain has to be considered. In addition, we allow for larger correlation lengths to enhance weak and extended source detection.

## 3. Results

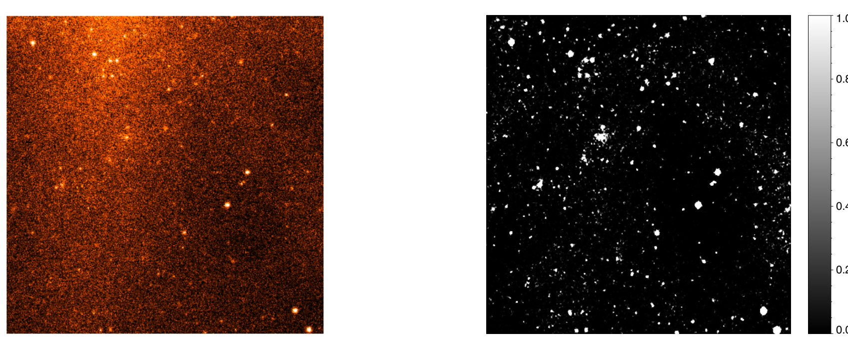


Figure 1. On the left, ROSAT field rs930625 at $\alpha = 17^h 49' 5.47''$ $\delta = +61^\circ 52^m 30.0^s$. On the right, source probability evaluated accounting for the width of the ROSAT PSPC PSF.

Fig. 1 left shows the ROSAT PSPC field RS930625 in the broad energy band ($E = 0.1 - 2.4$ keV) located in the ecliptic polar region. The observatory's exposure time varies between 1693.1-13475.5 sec. The average photon counts per

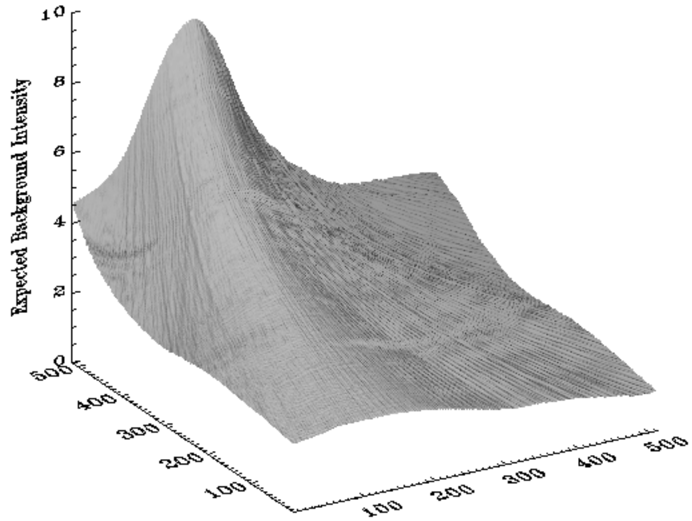


Figure 2. Background map extracted from rs930625 with a Thin-Plate spline combined with the observatory's exposure map.

pixel is 2.9. The right figure shows the probability for having source contribution in 3 by 3 pixel domains. The corresponding background map is shown in Fig. 2. The background intensity varies between $9.37 \pm 0.02$ and $1.153 \pm 0.006$ expected counts which shows the prominent variation due to the heterogeneous satellite exposure time.

## 4. Conclusions and Further Prospects

BPT allows to estimate background maps and to detect sources in a single step providing consistent uncertainties of background and sources. The source probability is evaluated for single pixels as well as for pixel domains to enhance source detection for weak and extended sources. The detection sensitivity is enhanced compared to SASS results because the full field of view is exploited for background estimation. An extensive comparison with SASS results is beyond the scope of the present paper and will be addressed in a forthcoming paper.

## References

Boese, F. G. & Doebereiner, S. 2001, A&A, 370, 649

Boese, F. G. 2000, A&AS, 141, 507

Fischer, R. & Dose, V. 2002, AIPC, 617, 143

Fischer, R. Hanson, K. M. Dose, V. & von der Linden, W. 2000, PRE, 61.1152

Voges, W. et al. 2000, IAUC, 7432

Voges, W. et al. 1999, A&A, 349, 389

von der Linden, W. Dose, V. Padayachee, J. & Prozesky, V. 1999, PRE, 59.6527