

## NORTHERN SKY VARIABILITY SURVEY: PUBLIC DATA RELEASE<sup>1</sup>

P. R. WOŹNIAK,<sup>2</sup> W. T. VESTRAND,<sup>2</sup> C. W. AKERLOF,<sup>3</sup> R. BALSANO,<sup>2</sup> J. BLOCH,<sup>2</sup> D. CASPERSON,<sup>2</sup>  
 S. FLETCHER,<sup>2</sup> G. GISLER,<sup>2</sup> R. KEHOE,<sup>3,4</sup> K. KINEMUCHI,<sup>4</sup> B. C. LEE,<sup>5</sup> S. MARSHALL,<sup>6</sup> K. E. MCGOWAN,<sup>2</sup>  
 T. A. MCKAY,<sup>3</sup> E. S. RYKOFF,<sup>3</sup> D. A. SMITH,<sup>3</sup> J. SZYMANSKI,<sup>2</sup> AND J. WREN<sup>2</sup>

Received 2003 December 3; accepted 2003 January 2

### ABSTRACT

The Northern Sky Variability Survey (NSVS) is a temporal record of the sky over the optical magnitude range from 8 to 15.5. It was conducted in the course of the first-generation Robotic Optical Transient Search Experiment (ROTSE-I) using a robotic system of four comounted unfiltered telephoto lenses equipped with CCD cameras. The survey was conducted from Los Alamos, New Mexico, and primarily covers the entire northern sky. Some data in southern fields between declinations  $0^\circ$  and  $-38^\circ$  are also available, although with fewer epochs and noticeably lesser quality. The NSVS contains light curves for approximately 14 million objects. With a 1 yr baseline and typically 100–500 measurements per object, the NSVS is the most extensive record of stellar variability across the bright sky available today. In a median field, bright unsaturated stars attain a point-to-point photometric scatter of  $\sim 0.02$  mag and position errors within  $2''$ . At Galactic latitudes  $|b| < 20^\circ$ , the data quality is limited by severe blending due to the  $\sim 14''$  pixel size. We present basic characteristics of the data set and describe data collection, analysis, and distribution. All NSVS photometric measurements are available for on-line public access from the Sky Database for Objects in Time-Domain ((SkyDOT)) at Los Alamos National Laboratory. Copies of the full survey photometry may also be requested on tape.

*Key words:* astronomical data bases: miscellaneous — catalogs — stars: general — stars: variables: other — surveys

### 1. INTRODUCTION

An amazing fact about modern astronomy is that the global time variability of the optical sky is largely unexplored for objects fainter than those observable with the naked eye (Paczynski 1997). As a result, the existing samples of known variables are quite incomplete. The commonly accepted standard catalog of variable stars is the General Catalogue of Variable Stars (GCVS; Kholopov 1998). However, the GCVS was compiled from a multitude of heterogeneous observational sources—in many cases based on analysis of photographic plates—and does not present any light curves. Spatial coverage in the GCVS is strikingly patchy at magnitude 10 and fainter. With the data currently available for most stars, we cannot answer the simple question of whether or not the star is variable. This is unfortunate because temporal flux changes not only carry useful physical information about the star, but they are of concern for experiments where variability could degrade the quality of the comparison star grid (e.g., the *Space Interferometry Mission*, Frink et al. 2001).

The reasonably sensitive, inexpensive CCDs that have become available in the last few years, coupled with the improved affordable data processing capabilities, have opened new windows for discovery in astrophysics. The large, uni-

form samples being collected with CCD imagers are enabling the temporal study of objects in new detail. Microlensing surveys, for example, have shown that massive photometric monitoring programs can return numerous scientific results, often unrelated to the original goal (Paczynski 2000a; Ferlet, Maillard, & Raban 1997). The added value of large number statistics and good sky coverage is also evident in catalogs from digitized POSS plates (Djorgovski et al. 2001) and multicolor surveys like the Sloan Digital Sky Survey (SDSS; Stoughton et al. 2002; Abazajian et al. 2003) and the Two Micron All Sky Survey (2MASS; Skrutskie et al. 1997) when applied to rare objects such as high- $z$  quasars and galaxies or brown dwarfs.

A new generation of small robotic sky patrol instruments is making all-sky temporal monitoring of point sources possible (Paczynski 2000b; Chen, Lemme, & Paczynski 2001). New results from these variability studies are contributing to our understanding of stellar evolution and Galactic structure. Resulting improvements in the local distance scale and discoveries of supernovae enable more accurate estimates of cosmological parameters and stellar ages. A substantial contribution to the understanding of gamma-ray bursts (GRBs) has already been made by robotic follow-up telescopes of modest size such as ROTSE (Akerlof et al. 2000b; Kehoe et al. 2001) and LOTIS (Park et al. 2002). Development of autonomous systems searching for optical flashes in real time will enable monitoring of a variety of fast and rare phenomena including the onset of optical emission from GRBs. The RAPTOR system (Vestrand et al. 2002) is a stereoscopic sky-monitoring system that is designed to find such flashes and provide instant notification and response while events are occurring. In the context of extrasolar planets, a shallow but very large area time-domain survey with high cadence and subpercent photometry currently offers the best prospect of discovering bright systems with transiting

<sup>1</sup> Based on observations obtained with the ROTSE-I robotic telescope, which was operated at Los Alamos National Laboratory.

<sup>2</sup> Los Alamos National Laboratory, MS D436, Los Alamos, NM 87545; wozniak@lanl.gov.

<sup>3</sup> Department of Physics, University of Michigan, 500 East University Avenue, Ann Arbor, MI 48109-1120.

<sup>4</sup> Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824-2320.

<sup>5</sup> Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720-8160.

<sup>6</sup> Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550.

Jovian planets that are suitable for detailed studies using high-resolution spectroscopy (Horne 2002; Charbonneau 2003).

To date, the largest volumes of data on time variability are those collected by microlensing searches, containing typically  $10^7$  light curves, each with a few hundred individual photometric measurements spread over several years (Paczynski 2001; Brunner et al. 2002). Microlensing surveys can be considered medium angle surveys limited to specific areas of interest, namely the Galactic bulge and several galaxies of the Local Group. In the category of very wide field surveys, the All Sky Automated Survey (Pojmański 1997) has made a substantial fraction of the data available in the public domain and returned  $\sim 7000$  variable stars brighter than 15 mag discovered primarily in the  $0^h$ – $6^h$  quadrant and additional scattered fields of the southern hemisphere (Pojmański 2000, 2002). The search for transits by extrasolar Jupiters resulted in recent proliferation of large-area photometric monitoring projects (Horne 2003), but little data has been published to date.

This paper marks the first release of data from the Northern Sky Variability Survey (NSVS), a CCD-based synoptic survey covering the entire sky north of the declination  $\delta = -38^\circ$ . This data release provides light curves for  $\sim 14$  million objects down to  $V \sim 15.5$  mag, with hundreds of repeated observations spanning 1 full year. NSVS photometry is a great improvement over photographic work despite disadvantages of using a single unfiltered photometric band. Most importantly, all observations were collected with the same instrument resulting in uniform data quality limited largely by crowding near the Galactic plane. We are working toward making the NSVS the best possible tool for studies of stellar variability and Galactic structure using bright stars. Akerlof et al. (2000a) published a preliminary variability analysis covering nine out of 161 survey tiles and 3 months of observing time. While we repeat parts of their discussion of the observing system and data reduction, there are significant differences in data processing between this work and Akerlof et al. (2000a).

The outline of the paper is as follows. Section 2 describes the telescope, cameras, and the process of data collection. Section 3 presents the details of data reductions and photometry followed by the discussion of survey quality and coverage. In § 4, we describe public access to the data and some technicalities of the data products. Section 5 concludes with the summary and future prospects for NSVS.

## 2. INSTRUMENTS AND DATA ACQUISITION

Numerical data for this section is summarized in Table 1.

### 2.1. ROTSE-I Robotic Telescope

All data in the present NSVS data set were collected by the first-generation Robotic Optical Transient Search Experiment (ROTSE-I). The primary goal of that experiment was prompt response to GRB triggers from satellites in order to measure the early light curves of GRB optical counterparts (Akerlof et al. 1999, 2000b; Kehoe et al. 2001). The normal operation of the ROTSE-I instrument was completely automatic requiring only periodic maintenance. The telescope consisted of four Canon 200 mm lenses with  $f/1.8$  focal ratio, each covering  $8.2^\circ \times 8.2^\circ$  for a total field of view about  $16^\circ \times 16^\circ$ . All four optical elements were carried by a single rapidly slewing mount and designated with the symbols A through D. The telescope and mount were located on the roof of a military

surplus enclosure, which housed the instrument control computers, and protected by a clamshell during the day or in bad weather. The instrument was located at the Los Alamos National Laboratory ( $W106^\circ 15' 13''$ ,  $N35^\circ 52' 9''$ ), just outside the town of Los Alamos, New Mexico. Light pollution at this site, although detectable, had only a minor impact on final survey photometry. Given the very dry local climate and 85% fraction of useful nights it is well suited for survey astronomy. Significant cloud cover is confined to the monsoon season lasting from July to mid-September.

### 2.2. Imaging Cameras

Each of the four Canon lenses was equipped with a thermoelectrically cooled AP-10 camera, which employs a Thomson TH7899M CCD. The  $2K \times 2K$  chip format covers an  $8.2^\circ \times 8.2^\circ$  field of view with  $14\frac{2}{3}$  pixels. The spatial resolution of the system was limited by instrumental seeing. The Canon lenses delivered a typical point-spread function (PSF) with full width at half-maximum (FWHM) of  $\sim 20''$  and therefore marginally undersampled images of point sources. The ROTSE-I telescope was operated without any filters so the spectral response is primarily limited by the sensitivity of the CCD, resulting in a very broad optical band from 450 to 1000 nm that covers the photometric bands from mid  $B$  to mid  $I$ . The quantum efficiency of the front-illuminated thick CCD chips makes the effective band most comparable to the Johnson  $R$  band. To optimize the readout speed for GRB response measurements, the images were read in 14-bit mode. There is no loss of information, however, because of the relatively narrow dynamic range of the AP-10 cameras. The typical gain setting was  $\sim 8 e^- \text{ADU}^{-1}$ . Images are sky background limited primarily because of large pixel size. The limiting  $V$  magnitude of the faintest stars recorded in 80 s exposures was typically 14.5–15.5. Saturation occurred at 10–10.5 mag in normal exposures (80 s) and at about 8 mag in bright time exposures (20 s). Vignetting in the lenses is very significant and amounts to about 40% loss of sensitivity near the corners of the CCDs. This effect is very stable and easily corrected by flat-fielding procedure. The shutters in the AP-10 cameras did not perform according to their specification. Especially in cold conditions, shutters did not operate smoothly and caused “anomalous vignetting” near the frame edges in some of the images. Photometric corrections explained in § 3.2 remove this and other effects. Dark frames and small-scale flat-field features were stable over a few days to a week. A few bad columns in the CCDs did not affect the overall quality of the data set. Cameras A–D did not perform equally. The slightly lower photometric quality of camera D can be seen in survey statistics presented in § 4.1. The loss of observing time due to temporary failure of camera C is also visible.

### 2.3. Observing Protocol

Despite the fact that the primary goal of the ROTSE-I project was rapid response to GRB triggers and not sky patrols, almost all observing time was actually spent in the latter mode. An accessible GRB position would be posted by the GCN network approximately once every 10 days. Upon receipt of the coordinates, the ROTSE-I system would abort the current patrol activity and immediately start observing the field around the position for approximately 1 hr of imaging.

At the beginning of each night, about 12 dark frames were collected for calibration purposes. No special flat-field exposures were made (§ 3.1.1). The large combined field of view

TABLE 1  
EQUIPMENT AND OPERATIONS IN THE NSVS

Parameter	Value
Telescope and Site	
Geographic position.....	Los Alamos, New Mexico: W106°15'13", N35°52'9"
Elevation.....	2300 m
Mount speed.....	100° s <sup>-1</sup>
Telescopes.....	Four 200 mm, f/1.8 Canon lenses
Seeing.....	FWHM ~ 20", instrumental
Vignetting.....	Up to 40% near frame corners
Lens offsets.....	$\Delta\alpha \cos \delta$ , $\Delta\delta$ with respect to mount position
A:.....	-4°, +4°
B:.....	-4°, -4°
C:.....	+4°, -4°
D:.....	+4°, +4°
Imaging Cameras: Four Apogee AP-10's	
CCDs.....	2K × 2K Thomson TH7899M chip
Gain.....	8 e <sup>-</sup> ADU <sup>-1</sup>
Read noise.....	13–25 e <sup>-</sup> pixel <sup>-1</sup>
Dynamic range.....	>74 dB
Read mode.....	14 bit, 1.3 MHz
Frame size.....	2035 × 2069 pixel
Image scale.....	3.50 mm deg <sup>-1</sup>
Pixel size and scale.....	14 $\mu$ ; 14.4 pixel <sup>-1</sup>
Filter.....	Unfiltered optical response ~450–1000 nm, effective wavelength of <i>R</i> band
Field of view.....	8.2 × 8.2 per camera
Data Collection	
Survey area.....	33,326 deg <sup>2</sup> ( $\delta > -38^\circ$ ), best coverage for $\delta > 0^\circ$
Time baseline.....	1 yr, from 1999 April 1 to 2000 March 30
Number of fields.....	644 = 161 × 4 (cameras A, B, C, D)
Number of frames.....	184,006
Number of nights.....	275 out of 365
Time sampling.....	Pairs of frames 1.5 minutes apart, up to two pairs per night
Calibration.....	500–1000 Tycho stars per frame

delivered by the ROTSE-I system requires only 206 tiles to cover the entire sky, with 161 tiles observable from Los Alamos. The list of fields with elevation above 20° was prepared by start-up scripts. During normal execution of a patrol, for each of those fields, two 80 s exposures would be taken, separated only by the 1.5 minute duty cycle. Reduced exposure time of 20 s was used in bright moonlight conditions (~30% of data). No frames were taken when the moon was closer than 12° from the field center. Time keeping accurate to 20 ms was implemented using Network Time Protocol. On a good night, it was possible to cover the entire local sky (~10,000 deg<sup>2</sup>) twice. Paired observations are useful for detecting variability and aperiodic transients. They also provide a handle on spurious detections due to man-made space objects, cosmic rays, hot pixels, and other effects.

The position angle of the cameras was fixed at P.A. = 0° in all fields, except for the near-polar region where the control software allowed P.A. = 180°. In the latter case, fields normally assigned to cameras A and B would be imaged by cameras C and D, respectively. Such observations are flagged appropriately and excluded from parts of the analysis (§ 3.2.3).

### 3. DATA PROCESSING

Analysis of the data presented in this release was conducted off-line on archival ROTSE-I images. Here we briefly discuss

the data reduction pipeline, schematically shown in Figure 1, which was employed to analyze that data.

#### 3.1. Image Reductions

##### 3.1.1. Basic Frame Corrections

Between 1997 August and 2001 December, ROTSE-I collected 7 TB of image data, however the performance of the system was not optimal in the first few months of the project and near the end of its lifetime. To build the NSVS, we selected observations covering 1 full year between 1999 April and 2000 March, when the system delivered the best overall data quality. This limited the raw data set to 225,000 images (~2 TB). The system would automatically prepare a median dark frame for each exposure time using all dark images collected on a particular observing night. Dark subtraction removes a small fraction of pixels ( $\ll 1\%$ ) with high dark current rates. Flat-field frames were obtained from a median of all individual patrol images made during a given night. This is possible with a large number of independent fields (~80) and statistics limited by sky noise due to large pixel size. Stellar profiles are completely removed by the procedure. We found that shutter problems (§ 2.2) affected some of the flat-field images. Therefore, we visually evaluated all flat-field frames and their ratios with frames made on a few other nights to

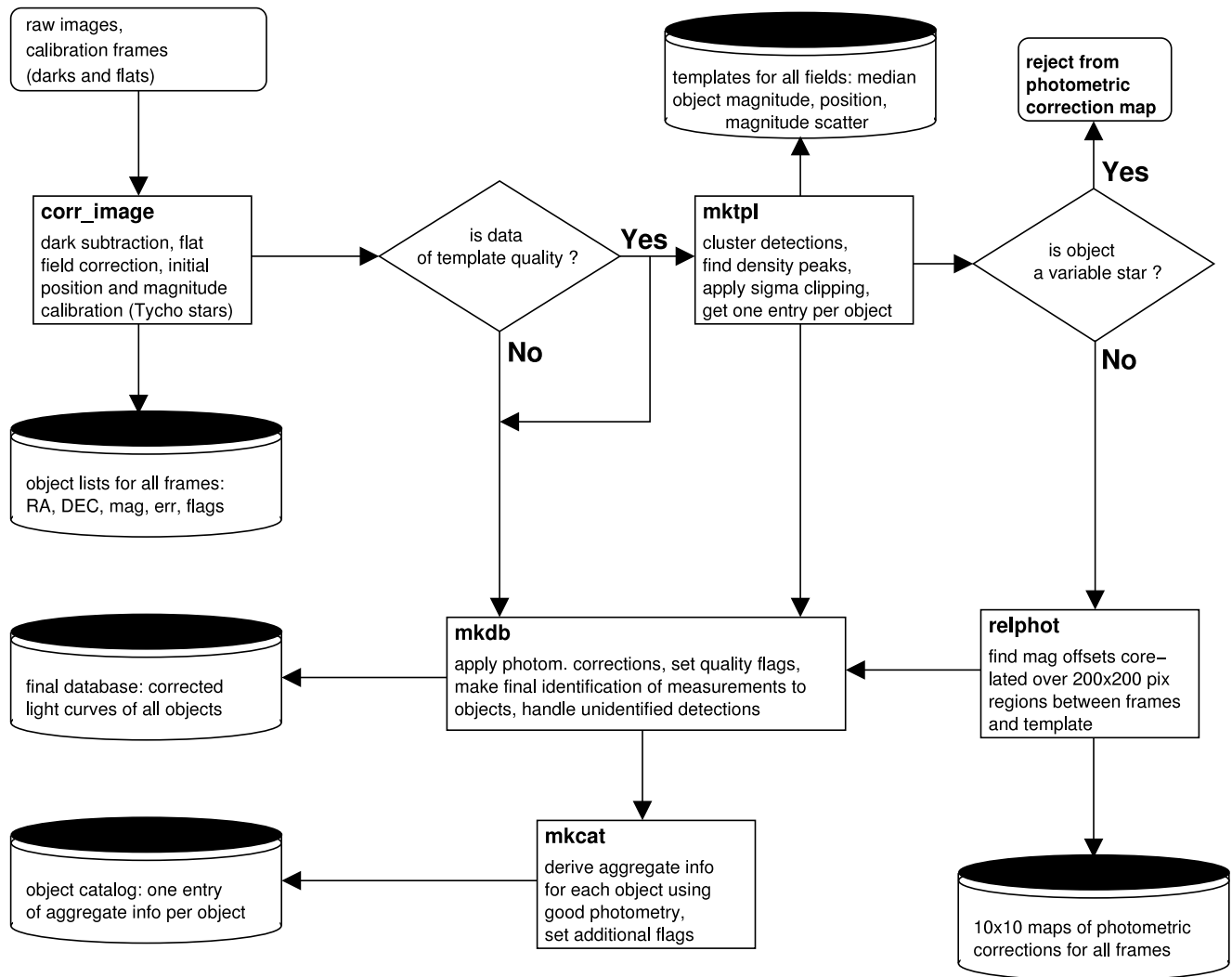


FIG. 1.—NSVS data-processing pipeline

select calibration sets with consistent large-scale properties. As a result, it was possible to correct the vast majority of the frames using good calibration frames from the same night or the night before.

### 3.1.2. Source Extraction

The corrected images are passed through SExtractor software (Bertin & Arnouts 1996), which reduces them to object lists. The choice of this source extraction package was motivated by the undersampling of stellar profiles, the significant gradients of the PSF shape in very wide field images, and overall reduction speed. SExtractor was optimized for reduction of images in galaxy surveys, but it is known to perform well in moderately crowded stellar fields. In order to optimize sensitivity, our images are filtered before object detection with a Gaussian kernel employing a FWHM of 2.5 pixels and requiring a minimum of 5 connected pixels in an object. These basic detections are further thresholded by the software and an attempt is made to break up blended objects. We use SExtractor aperture magnitudes calculated with the 5 pixel (72") aperture diameter. Since the ROTSE-I images are almost completely dominated by stars, we store only a small fraction of information available for each object: position, magnitude, magnitude error, and processing flags. The observed errors at

the bright end of the magnitude range are larger than predicted by simple photon noise. Such discrepancies are common for CCD measurements and are typically generated by residual systematic effects of flat-field errors, thin clouds, PSF variations, and sampling. In order to account for these systematic errors, we had to add, in quadrature, a 0.01 mag contribution to the formal error bars.

### 3.1.3. Blending

SExtractor does not perform PSF photometry, and in general it is unable to deblend light distributions without a saddle point. This sets the distance limit of about 3.0 pixels for separation of stellar blends. We found that the default parameters of the deblender were very conservative, resulting in very large patches of the sky at low Galactic latitudes being assigned to the same objects. After some experimentation, we were able to partially control this process, however, there are still cases when tight groups of several objects with merging wings are considered to be a single object. Typically, such aggregates extend up to 10 pixels across with a bright object in the middle, but occasionally in dense parts of the Milky Way, they can be up to 30 pixels wide. This makes the completeness of the NSVS at  $|b| < 20^\circ$  depend strongly on stellar number density at small spatial scales of around  $\sim 3'$ .

The exact assignment of pixels to objects may differ from frame to frame, and therefore some measurements lost due to severe blending may still be present among “orphaned” measurements, unidentified with any of the light curves (§ 3.2.3). The term “orphan” used here in the context of a photometric detection should not be confused with orphan GRBs and other optical transients.

### 3.1.4. Astrometric and Photometric Matching

Initial calibration of object lists consists of transformation of instrumental positions to celestial coordinates and conversion of raw instrumental magnitudes to a photometric system that can be understood by users. In wide-field imaging, this can be done on an image by image basis. Each ROTSE-I field covers 64 deg<sup>2</sup> and contains on average about 1500 stars from the Tycho catalog (Hog et al. 1998). The Tycho catalog, a product of the *Hipparcos*/Tycho mission, provides very accurate astrometry and two-color (*B* and *V*) photometry. Most Tycho stars are fainter than the 10 mag saturation limit of the ROTSE-I observations. The astrometric matching is performed in the detector plane after deprojecting the corresponding part of the catalog using a canonical gnomonic projection (Calabretta & Greisen 2002). Using an approximate mount position for a given image, the first set of roughly 30 Tycho stars can be identified with the triangle algorithm. This first-order transformation is used to further match a few hundred bright, but unsaturated, Tycho stars. A third-order polynomial warp adequately describes the transformation between the observed and catalog positions of stars in the detector plane. Finally, transformed (*x*, *y*) positions are converted back to ( $\alpha_{J2000.0}$ ,  $\delta_{J2000.0}$ ) using the inverse of the initial gnomonic projection.

Photometric calibration is somewhat complicated by the very wide unfiltered spectral response of the ROTSE-I imaging system that spans a large part of the Johnson-Cousins system from mid-*B* to mid-*I* (§ 2.2). The best empirical prediction of a ROTSE magnitude  $m_{V,ROTSE}$  for Tycho stars is

$$m_{V,ROTSE} = m_V - [(m_B - m_V)/1.875].$$

The median shift between instrumental magnitudes and the above color-corrected magnitudes of Tycho stars is then applied to all stellar magnitudes from a given image. This procedure puts ROTSE-I measurements onto a *V*-equivalent scale, in the sense that the mean Tycho star has  $m_{V,ROTSE} = m_V$ . These intermediate object lists are passed through light-curve building software and are subject to further refinements (§ 3.2). Matching to Tycho stars occasionally fails in very crowded areas of the Galactic plane and/or because of substantial cloud cover. Initial calibration was successful for 184,006 frames.

### 3.2. Object Identification and Photometric Corrections

A set of object lists for all exposures of a given field has to be collated in order to identify measurements that belong to the same objects and hence construct light curves. In the process, a collective look at the temporal behavior of stars across the field can provide useful information on systematics of the photometry. Very wide field images show some complications that are usually unimportant in data from narrow field instruments. Pronounced gradients of background, color-

dependent atmospheric extinction and gray extinction from thin clouds are common. Color-dependent effects cannot be fully corrected using only a single photometric band. Some ROTSE-I images have additional complications near frame edges caused by shutter problems (§ 2.2). After the processing steps outlined in § 3.1, most photometric residuals are still correlated over spatial scales of a few hundred pixels. These gradients are handled by local photometric corrections. The procedure allows measurements of an object originating from each frame to be expressed on a relative scale with respect to stars in the neighborhood, as given by a master list of stars called a template. At the same time, we can collect diagnostics providing useful measures of the data quality in the final database. They are used to set the measurement quality flags (Table 2), making it easier to select measurements satisfying requirements of a particular application.

#### 3.2.1. Template Construction

The process begins with preparation of a template object list for each of the 644 fields. Frames with the MOUNTFLIP flag (see Table 2) or fewer than 1000 detected objects are rejected for this purpose. We also apply a cut on standard deviation of the position and magnitude offsets around the fit to Tycho stars ( $\text{pos\_sigma} < 0.3$  pixels and  $\text{zp\_sigma} < 0.4$  mag, see § 4.2). Measurements of all field objects from all admitted images are clustered in order to identify persistent objects. To accomplish this, we count the number of individual detections on a simple grid with pixels equal in size to detector pixels. Local maxima of this histogram containing at least 15% of all possible detections are declared as objects. From this counting exercise, we derive crude centroids that are further refined using a 1 pixel identification radius and subsequent sigma clipping. The main properties of objects stored in a template list for each field are median ( $\alpha_{J2000.0}$ ,  $\delta_{J2000.0}$ ) positions, median object magnitudes, and standard deviations of individual magnitudes around the median. These are essentially aggregate parameters of the preliminary light curve for each object. By using median object magnitudes to construct the template, we avoid biasing all field photometry toward a single frame that may be affected by systematic effects of the kind described above. As a measure of scatter that is robust against occasional strong outliers, we adopt half size of a centered range that includes 68% of all magnitude points.

#### 3.2.2. Photometric Correction Maps

In the next step, template lists are used to derive two-dimensional maps of photometric corrections. The most significant contribution to systematics in frame-to-frame photometry is caused by clouds, atmospheric extinction gradients, and occasional shutter problems. The photometric errors are therefore correlated over the scales of a few hundred pixels. By examining a median shift between magnitudes of stars detected in a program frame and template magnitudes for the same stars in any 200 × 200 pixel area, one obtains a correction to the photometry of stars in the defined macropixel. Stars for which the scatter is more than 4 times the median error bar are rejected as potentially variable. The frames are then divided into macropixels, each covering 200 × 200 detector pixels, to obtain 10 × 10 maps. The map records relative photometry corrections, the scatter of all magnitude differences in a macropixel and the number of objects available for those calculations. The global scatter of the map is also recorded. All these characteristics can be used

TABLE 2  
PROCESSING FLAGS

Hexadecimal Bit	Decimal Bit	Flag	Description
Frame Flags			
0x0001.....	1	MOUNTFLIP	Mount flip near north pole, fields AB observed by cameras CD
0x0002.....	2	ODDMNTPOS	Nonstandard mount position, only fields 001 ABCD in year 2000
Object Flags			
0x0001.....	1	TPLSTATS	Light-curve statistics in “Object table” without photometric corrections
0x0002.....	2	BIGSHIFT	Final median object centroid more than 1 $\sigma$ from preliminary position
Measurement Flags: SExtractor			
0x0001.....	1	NEIGHBORS	$\geq 10\%$ of object area affected by a neighboring object or bad pixels
0x0002.....	2	BLENDED	Object is a result of deblending procedure
0x0004.....	4	SATURATED	Object has at least 1 saturated pixel
0x0008.....	8	ATEDGE	Object is truncated by image boundary
0x0010.....	16	APINCOMPL	Aperture data incomplete or corrupted
0x0020.....	32	ISINCOMPL	Isophotal data incomplete or corrupted
0x0040.....	64	DBMEMOVR	Memory overflow occurred during deblending
0x0080.....	128	EXMEMOVR	Memory overflow occurred during extraction
Measurement Flags: Photometric Correction			
0x0100.....	256	NOCORR	Relative photometry correction could not be calculated
0x0200.....	512	PATCH	Map of relative photometry corrections was patched to derive correction
0x0400.....	1024	LONPTS	Low number of points in a macropixel ( $<10$ )
0x0800.....	2048	HISCAT	High scatter of magnitude differences in macropixel ( $>0.2$ mag)
0x1000.....	4096	HICORR	High value of correction ( $>0.1$ mag)
0x2000.....	8192	HISIGCORR	High scatter of corrections across the map ( $>0.1$ mag)
0x4000.....	16384	RADECFLIP	Mount flip near north pole occurred (fields 001– 032 ABCD only)

to assess the quality of the data. The pipeline will assign a “null” value to a macropixel in cases when there are fewer than 10 stars to work with, available stars cover less than half of the macropixel area, or the value of the correction or its error are unreasonably high ( $>1.0$  mag).

### 3.2.3. Compilation of Light-Curve Database and Object Catalog

Applying relative photometry corrections and building intermediate database files is the final stage in the data reduction process. Individual corrections are obtained by bilinearly interpolating the maps. In the case of a missing macropixel surrounded by valid ones, we allow a patch based on linear interpolation to be applied. Such measurements are flagged accordingly but were observed to remain reliable. We also flag measurements with large corrections ( $>0.1$  mag), large scatter of magnitude differences used to derive the correction ( $>0.2$  mag), and large scatter of all macropixels in the map ( $>0.1$  mag). If the patch cannot be computed, there is insufficient information to derive the correction. This condition sets the NOCORR flag signaling a very unreliable measurement. Table 2 describes all the data-processing flags. The correction could not be calculated for fewer than 0.2% of the database measurements. For most frames, 95% of all correction values are distributed within  $\pm 0.05$  mag. In case of a frame affected by a shutter glitch, the distribution can be strongly non-Gaussian with the 95% interval typically extending to  $\pm 0.1$  mag. In both cases, there is a small tail of high-correction values extending to about  $\pm 1.0$  mag.

Measurements that can be associated with the position of a template object to within 4  $\sigma$  are tagged with the object identification and frame identification and stored in the main

light-curve database file. The error estimate is available from the template list. There are no additional restrictions on detections admitted to the main light-curve database. The final number of database light curves is almost  $2 \times 10^7$  (the same as the number of template objects). There are  $\sim 3.35 \times 10^9$  individual observations with identified parent objects. Remaining observations are in the “orphan” category. We decided to keep only high S/N orphan measurements, those with magnitudes 14.5 or brighter and magnitude errors below 0.1. The database contains over  $2 \times 10^8$  orphan measurements. All measurements are treated individually, in particular joined detection in paired exposures (§ 2.3) is not required. Such criteria, based solely on frame epochs, can always be applied to the data as a post-processing step.

At this point, we formed a revised catalog of objects with average properties based on the fully processed light-curve files. For each template object, we examine a corrected set of measurements and recalculate the median magnitude, magnitude scatter (significantly improved in most cases), median error bar, and other useful characteristics. Only the subset of “good” measurements defined in Table 3 is used in these calculations. These criteria primarily reject error codes and several flags associated with known problems. Currently, it is a recommended choice for working with NSVS data that offers a sensible compromise between the amount of data and data quality. Other equally good or better selection cuts may be possible. A small fraction of all objects below the equator and at the far field edges have fewer than 15 “good” measurements and, as such, their data are of limited use. Corresponding catalog entries have light-curve statistics copied from the template list and are flagged as having TPLSTATS.

TABLE 3  
DEFINITION OF A GOOD PHOTOMETRIC POINT

Condition	Serves Primarily to Remove
$5.0 < \text{mag} < 16.0$ .....	SExtractor error codes
$0.0 < \text{mag error} < 0.4$ .....	SExtractor error codes
!SATURATED .....	Saturated measurements; important for stars in and out of saturation
!NOCORR .....	Generally unreliable macropixels
!LONPTS .....	Measurements from far frame edges or shallow images
!HISCAT .....	Macropixels with unreliable photometry
!HICORR .....	Macropixels substantially affected by cloud cover or shutter problem
!HISIGCORR .....	Frames substantially affected by cloud cover or shutter problem
!RADECFLIP .....	Possible camera-dependent systematics; harmless in many applications

### 3.2.4. Multiple Detections in Overlap Regions

Pointing imperfections in repeated exposures of a given field result in object detections outside the area around the field center covered by a single CCD format (Table 1). This effectively extends each template by a margin of  $\pm 150$  pixels on all sides. Together with similarly sized intended overlap between the fields, this causes multiple detections of the same physical objects in more than one field. Overlap regions tend to grow closer to the celestial north pole for simple geometric reasons. About 30% of the database light curves belong to object references of second order or higher. In § 4.2, we use those independent light curves to assess systematic errors in positions and magnitudes. On the other hand, removing multiple references to the same objects requires perfect control over systematic effects. Even if the influence of shutter problems and atmospheric gradients were known with certainty, some irreducible effects would remain. The problem is not trivial in the case of unfiltered photometry, especially for fast variable sources, but it becomes much easier to handle for a few sources of particular interest after the data has been extracted from the database. This is why a global merge of the physical database has not been performed. We provide a database table with pairs of object identifications that refer to the same physical object (§ 4.2). Implementation details of this kind can be hidden by the user interface to the database by defining a mapping between the low-level database structures and the top-level view where object data appear as merged

(§ 4.2). That approach provides full flexibility to refine the definition of what is the same, and what is not, as more is known about the data set.

## 4. PUBLIC DATA DISTRIBUTION

### 4.1. Survey Coverage and Quality

Our NSVS covers the entire sky visible from Los Alamos, New Mexico. This includes the entire region north of the declination  $\delta \simeq -38^\circ$ ; more than 30,000 deg<sup>2</sup>, 75% of the celestial sphere. The completeness, quality of the data, and the number of available measurements are noticeably lower at low declinations and low Galactic latitude. The overall quality of the survey measurements is summarized in Table 4. In Figure 2, we show positions of roughly  $5 \times 10^5$  NSVS stars brighter than 11 mag out of about 14 million stars brighter than 15.5 mag present in the survey. The main feature in this plot is the overdensity of stars near the plane of the Milky Way. In that region, one can notice relatively less populated areas due to dust lanes. An empty spot near the Galactic Center (field 156 D) is an artifact. The region is basically lost from the survey due to the very low number of available frames. Each dot in Figure 2 represents a time series of typically a few hundred points spread over a total time baseline of 1 yr. Temporal coverage is subject to yearly visibility patterns. The trade-off between this enormous monitoring coverage on one side, and on the other, relatively low-resolution compounded by complexities of very wide field photometry, shaped the

TABLE 4  
NSVS DATA QUALITY

Parameter	Value
Pointing offsets .....	$\sigma \sim 0''.3$ (75 pixels) per coordinate
PSF .....	FWHM $\sim 1.5$ pixels, undersampled, spatially variable, temporally stable
Saturation .....	10–10.5 mag, up to 8 mag in bright time
Limiting magnitude .....	$\sim 15.5$ mag, 14.5 mag in bright time
Astrometric errors:	
Random <sup>a</sup> .....	Magnitude-dependent, $1 \sigma = 0''.7\text{--}4''.3$ , $1 \sigma = 1''.4\text{--}5''.8$ at Galactic $ b  < 20^\circ$
Systematic <sup>b</sup> .....	Median deviation $\sim 1''.2$ in general field, $\sim 2''.5$ at Galactic $ b  < 20^\circ$
Photometric errors:	
Random <sup>a</sup> .....	Limiting scatter $1 \sigma = 0.02$ mag in median field, $1 \sigma = 0.02\text{--}0.05$ mag at Galactic $ b  < 20^\circ$
Systematic <sup>b</sup> .....	Median deviation $\sim 0.04$ mag near frame edges, up to 0.2 mag in extreme cases, improving near field center
Blending .....	Stars closer than 3 pixels (84'') generally merged, severe at Galactic $ b  < 20^\circ$ , $\sim 5\%$ loss of survey area
Time sampling .....	Twice per night to once every $\sim 4$ nights
Number of epochs .....	$\sim 200$ for average light curve, follows yearly visibility across the sky
Number of objects .....	$\sim 14$ million

<sup>a</sup> Frame to frame within the same field.

<sup>b</sup> Systematic difference for the same object in overlap region between fields.

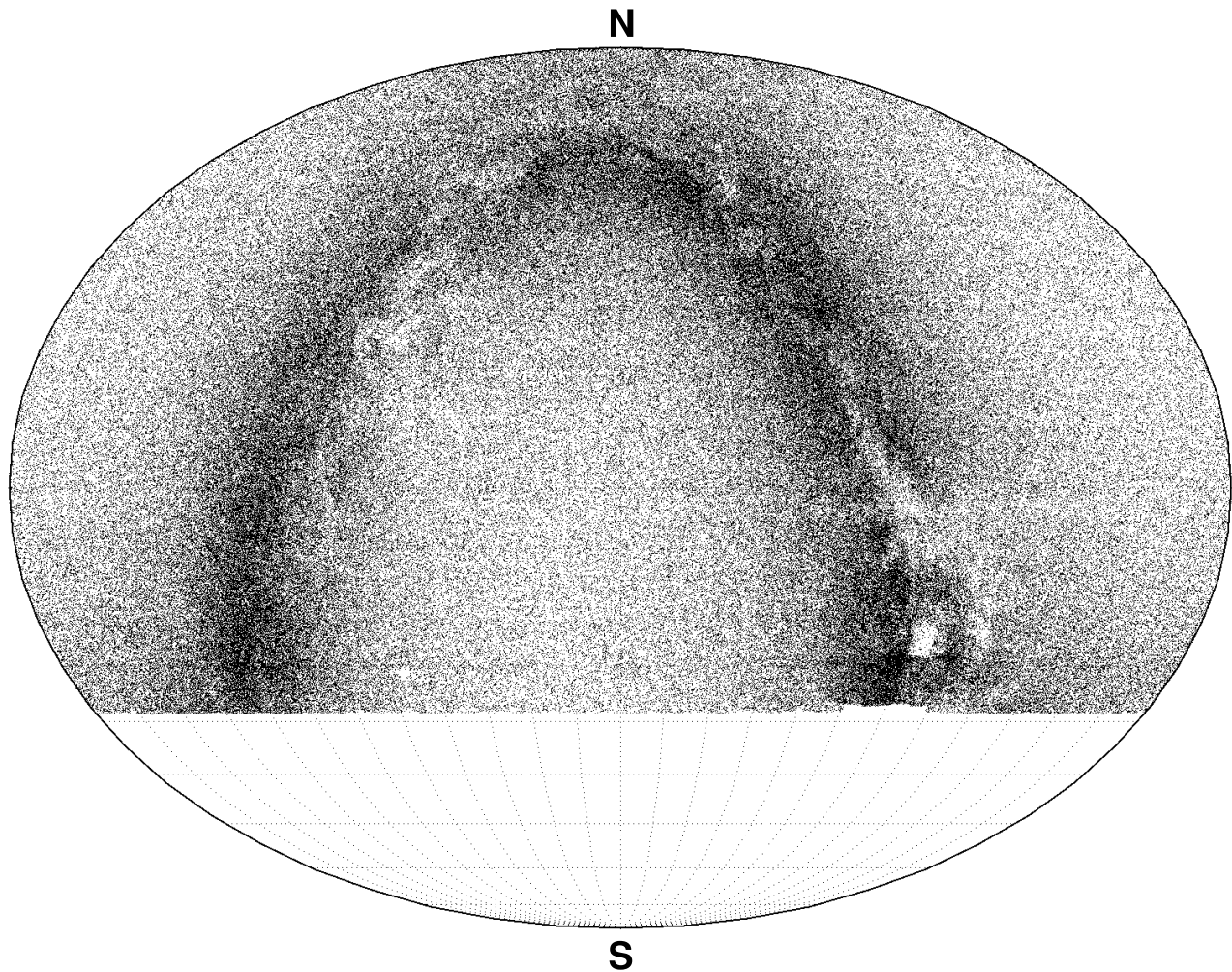


FIG. 2.—Positions of NSVS objects brighter than 11 mag (about 500,000) in equal area Mollweide's projection. Each dot represents an NSVS object with a temporal history typically composed of a few hundred measurements covering the 1 yr baseline.

final data quality. The faintest objects recorded in the survey have  $V \sim 15.5$  mag, however, the incompleteness starts increasing sharply at 15th magnitude. Saturation may occur in stars as faint as 10.5 mag on some nights, but it mostly affects stars brighter than 10 mag. Because of exposure times shortened by a factor of 4 during bright time, some stars as bright as 8.0 mag still have a number of unsaturated measurements sufficient for analysis.

Figure 3 summarizes important statistics of the NSVS. Numerous survey parameters strongly correlate with the Galactic latitude and declination. Several artifacts due to the low volume and quality of available data near the Galactic plane at low declinations are visible. The bright spot in the plots of number of available frames is due to a special data run in field 072 (Kehoe et al. 2002). The darker areas near the Galactic plane are a result of a lower success rate in matching frames to the Tycho catalog. Strong differences near the celestial pole between Fig. 3c and 3d are caused by exclusion of flag MOUNTFLIP in the definition of a “good” measurement. Lower than average performance of camera D is evident from a periodic pattern in the map of photometric scatter (Fig. 3b). The field pattern in the number of database frames reveals that camera C was not collecting data for about 3 months, when it had to be serviced after an electronics failure. Although the number of useful measurements below the equator drops

dramatically to fewer than 100, this is still sufficient to detect variables and the NSVS remains useful over a large section of the southern hemisphere.

#### 4.1.1. Astrometric Errors

Statistical scatter of object positions in individual frames can be better than 0.1 pixel ( $1''.4$ ) within a single field (lower part of Fig. 4). Median positions from a large number of measurements should be much more accurate than that, but that ignores systematic errors of the coordinate system derived separately in each field. A better measure for overall positional accuracy is the difference of median positions (in two-dimensions this time) for the same bright unsaturated stars observed in overlapping parts of adjacent fields. Such differences should be dominated by a systematic contribution and are shown in Figure 5. The distribution of these offsets turns out to be comparable to that from Figure 4 and fits well within a single image pixel. Figure 5 also shows how typical position uncertainties are affected by the high density of stars in the vicinity of the Galactic plane. The distribution peaks at a higher value for the error and develops a much longer tail due to strong blending.

#### 4.1.2. Photometric Errors

Figure 4 (top) presents magnitude scatter as a function of median object magnitude in a random field away from the



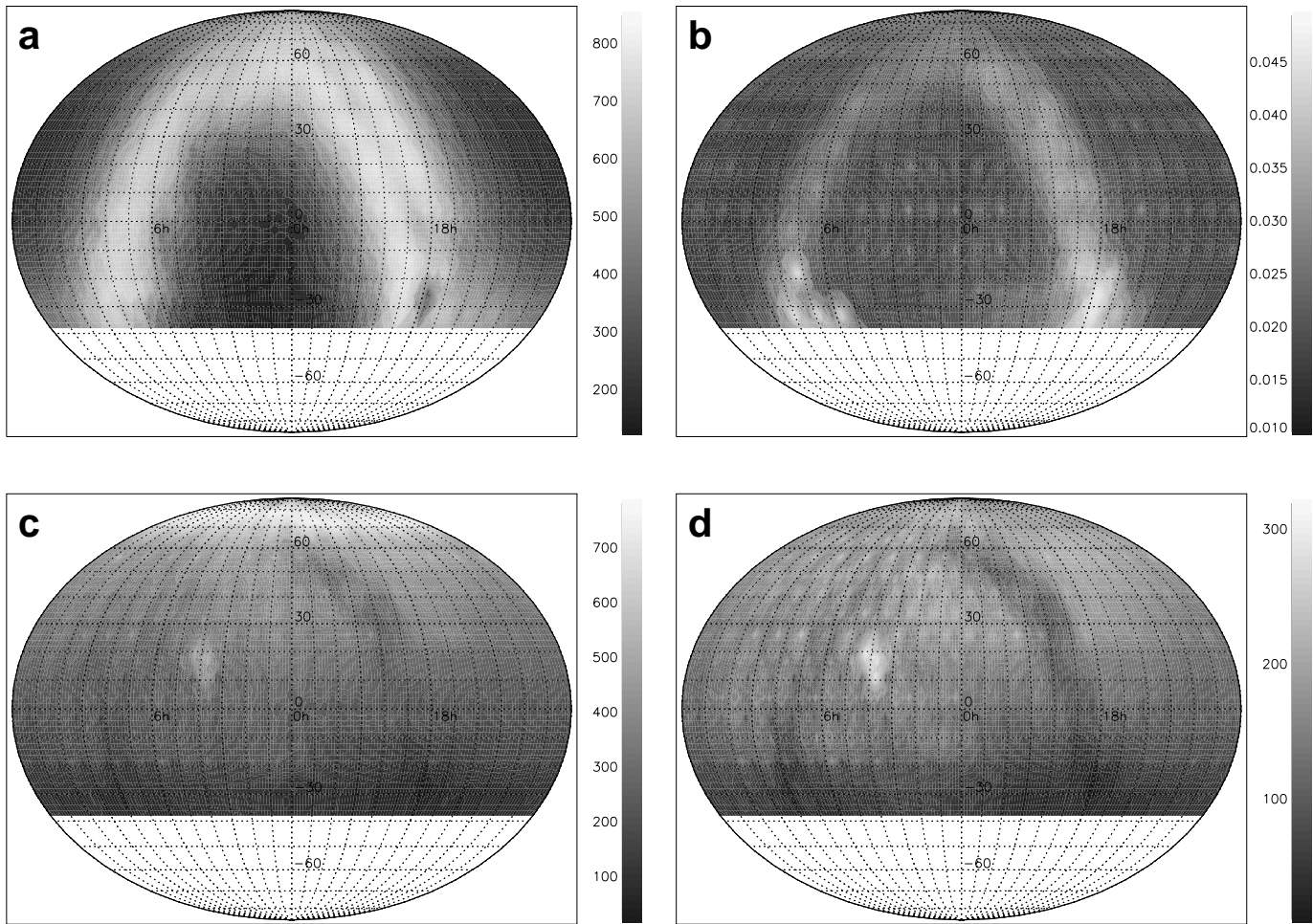


FIG. 3.—NSVS at a glance. Four panels show all-sky gray-scale maps smoothed over  $8^\circ$  spatial scale: (a) number density of NSVS objects, (b) photometric scatter for bright unsaturated stars calculated using “good” measurements, (c) median number of points per light curve, and (d) median number of “good” photometric points per light curve.

Galactic equator. Photometric scatter is estimated using “good” photometric points (§ 3.2.3) corresponding to about 75% of the best data. We consider magnitude scatter for a median star between 11 and 12 mag in each field to be the “limiting scatter,” that is, the best attainable for a significant fraction of bright stars in a given field. Observations in this magnitude range are limited by various systematics of photometric conditions and data reductions rather than the statistics of the background. There is a strong correlation between limiting scatter and Galactic latitude that is evident in Figure 6. The median limiting scatter over the total survey area is 0.02 mag. However the Galactic plane region, where photometric accuracy suffers significant degradation, has a higher number density of objects. Averaged over an  $8^\circ \times 8^\circ$  field, spatial density of the NSVS objects can vary between 150 and 1000  $\text{deg}^{-2}$ .

Similar to astrometry, systematics of photometry can be investigated by examining the offsets between magnitudes measured for stars in overlapping parts of adjacent fields. Histograms of these differences for two random overlap regions at low and high Galactic latitudes are dominated by constant stars and are shown in Figure 5. Half of the stars in the figure differ by 0.04 mag or less in their median magnitude obtained from light curves constructed independently

in neighboring fields. In individual cases, however, such differences can reach 0.1–0.2 mag. They result primarily from residual shutter problems that propagated to the construction of field templates, but also from irreducible atmospheric color effects in single broadband photometry. The fact that the histogram width does not change in a significant way in the proximity of the Galactic equator agrees with our assessment that the differences arise due to systematics associated with a particular set of field images and not because of high number density of stars or blending. It must be stressed that the overlap regions between the fields provide the upper bound on the systematic errors. This is where many instrumental effects (§ 2) manifest themselves strongest. The internal consistency over the remaining area is certainly much better, although not easily studied without a suitable external comparison catalog.

#### 4.2. Public Database

All photometric time-series data in the NSVS is available for public access. The primary means to search and extract the data is Sky Database for Objects in Time-Domain (SkyDOT);<sup>7</sup>

<sup>7</sup> See <http://skydot.lanl.gov>.

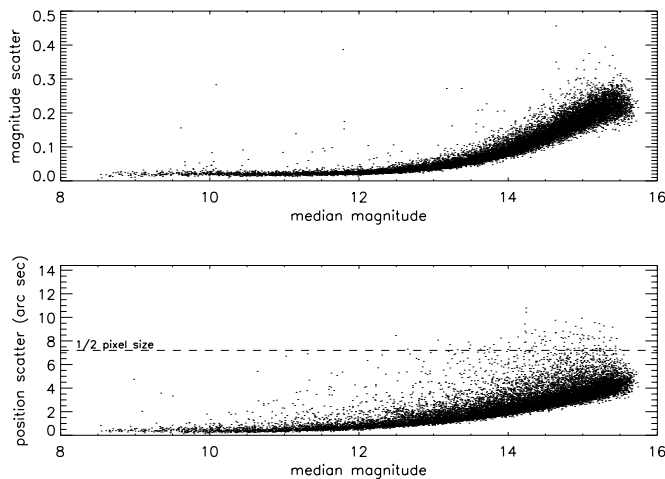


FIG. 4.—Random errors as given by frame-to-frame scatter in a typical field. Photometric errors (*top*) and position errors (*bottom*) are shown as a function of median object magnitude.

Woźniak et al. 2002). SkyDOT is intended to become a virtual observatory for variable objects. The site will provide a uniform interface to several large temporal data sets with a number of time-series analysis tools for on-the-spot application to currently viewed data. SkyDOT is implemented using

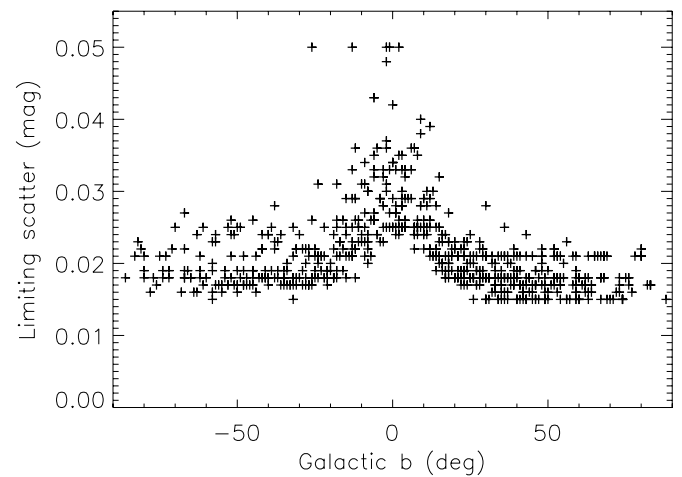


FIG. 6.—Limiting photometric scatter as a function of the Galactic latitude. In each field, we show light-curve scatter for bright unsaturated stars calculated using “good” measurements.

PostgreSQL,<sup>8</sup> an Object-Relational Database Management System with support for practically the entire SQL92 standard.

<sup>8</sup> See <http://www.postgresql.org>.

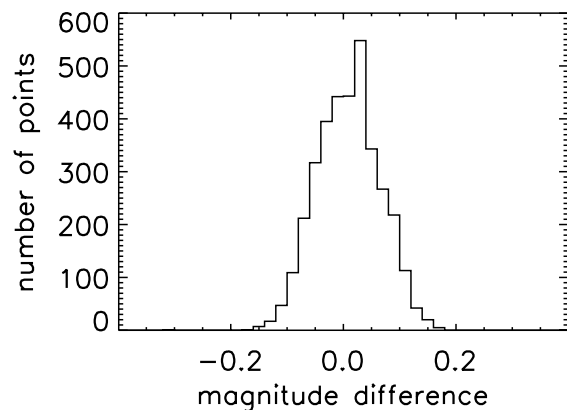
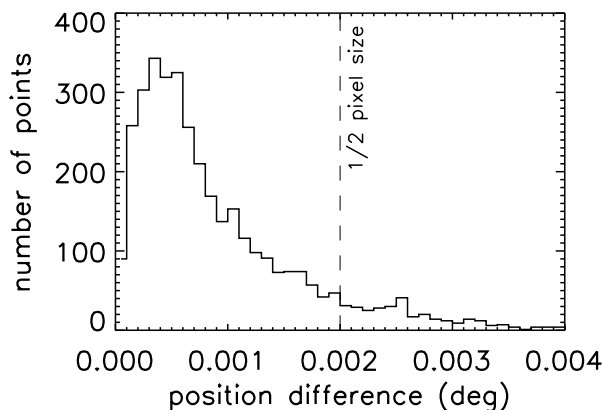
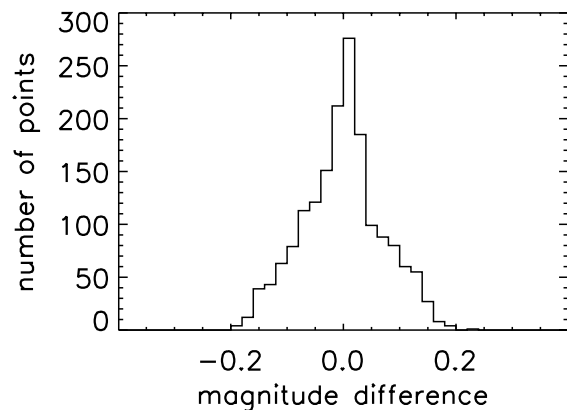
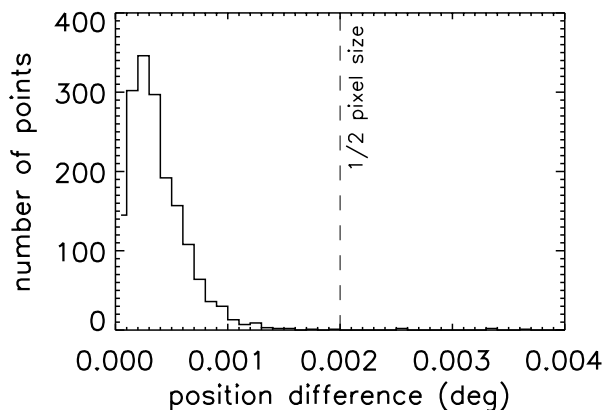


FIG. 5.—Systematic errors as given by differences between multiple detections of the same objects in overlap regions between the adjacent fields in the Galactic plane (*bottom*) and near the Galactic pole (*top*). Shown are differences in median object positions (*left*) and median object magnitudes (*right*) for bright unsaturated stars. Position differences fit well within a small fraction of a pixel. Magnitude offsets result from residual shutter problems and properties of very broad band photometry in the presence of intrinsic spread of object colors. Estimates based on overlap regions, where numerous instrumental effects are strongest, provide an upper bound on systematic errors.

TABLE 5  
DATABASE TABLES

Table	Number of Rows	Description
Field .....	644	ROTSE-I patrol tiles, each camera counted separately
Frame .....	184,006	Image header and postprocessing frame quality information
Object .....	19,995,106	One record of aggregate information for each light curve in the database counting separately same object detections from different fields
Synonym .....	14,582,566	Pairs of light curve IDs referring to the same physical object
Observation .....	3,353,171,900	All measurements for all light curves
Orphan .....	208,106,474	Measurements unidentified with any of the light curves but brighter than 14.5 mag and with errors less than 0.1 mag

The present database includes six entities listed in Table 5. The columns are explained in Tables 6 and 7. Five of those tables represent major entities in temporal work: Field, Frame, Object, Observation, and Orphan. The remaining one, Synonym, helps to identify multiple references to the same physical objects (§ 3.2.4). It implements “is the same” relationship between entries of the Object table. Users familiar with SQL

can submit queries to the engine with some limitations on the size of output imposed. All users can access the database through a graphical interface offered by the Web site. The most popular SQL queries can be accessed through browser buttons and search forms. Currently, only NSVS photometry is available in the public domain. There is neither a browsing capability nor a direct download option for ~2 TB of the

TABLE 6  
EXPLANATION OF TABLE COLUMNS: TABLES FIELD AND FRAME

Column Name	Data Type	Unit	Description
Field Table			
id .....	int32		Field ID (primary key)
name .....	char[4]		Field name
rac .....	float32	deg	$\alpha_{J2000.0}$ of field center
decc .....	float32	deg	$\delta_{J2000.0}$ of field center
glc .....	float32	deg	Galactic $l$ of field center
gbc .....	float32	deg	Galactic $b$ of field center
nobs .....	int32		Number of frames
nobj .....	int32		Number of catalog objects
sig_ph .....	float32	mag	Limiting photometric scatter
Frame Table			
id .....	int32		Frame ID (primary key)
field_id .....	int32		Field ID (foreign key)
fname .....	char[20]		Image file name
camera .....	char[2]		Camera ID (ABCD)
mjd .....	float64	day	JD – 2,400,000
obstime .....	float64	s	Time of observation in UT seconds
date_obs .....	char[22]		UT date of observation
exptime .....	float32	s	Exposure time
bkg .....	float32	counts	Sky background
bkg_sigma .....	float32	counts	Standard deviation of sky background
pos_sigma .....	float32	pixels	Standard deviation around the fit to positions of Tycho stars
zp_offset .....	float32	mag	Median magnitude offset with respect to Tycho stars
zp_sigma .....	float32	mag	Standard deviation of magnitude offsets
m_lim .....	float32	mag	Limiting magnitude
sat_mag .....	float32	mag	Saturation magnitude
nobj_det .....	int32		Number of objects detected by SExtractor
nobj_ext .....	int32		Number of objects actually measured
nmatch .....	int32		Number of Tycho stars used in magnitude matching
dmoon .....	float32	deg	Angular distance between frame center and the Moon
elev .....	float32	deg	Elevation of frame center with respect to horizon
azimuth .....	float32	deg	Azimuth of frame center
mount_ra .....	float32	deg	$\alpha_{J2000.0}$ position of telescope mount
mount_dec .....	float32	deg	$\delta_{J2000.0}$ position of telescope mount
offst_ra .....	float32	deg	$\Delta\alpha \cos \delta$ offset between frame center and mount position
offst_dec .....	float32	deg	$\Delta\delta$ offset between frame center and mount position
map_rms .....	float32	mag	Standard deviation of the photometric correction map
map_npix .....	int16		Number of valid pixels in photometric correction map
flags .....	int16		Frame flags

TABLE 7  
EXPLANATION OF TABLE COLUMNS: TABLES OBJECT, SYNONYM, OBSERVATION, AND ORPHAN

Column Name	Data Type	Unit	Description
Object Table			
id .....	int32		Object ID (primary key)
rao .....	float64	deg	Median $\alpha_{J2000.0}$ of object centroid
sig_rao .....	float64	deg	Standard deviation of individual $\alpha_{J2000.0}$ positions
deco .....	float64	deg	Median $\delta_{J2000.0}$ of object centroid
sig_deco .....	float64	deg	Standard deviation of individual $\delta_{J2000.0}$ positions
htm_id .....	int64		HTM ID for quick spatial queries on the sphere
mag .....	float32	mag	Median object magnitude from “good” points
rms_mag .....	float32	mag	Standard deviation of “good” points around median
med_err .....	float32	mag	Median error bar of “good” points
n_obs .....	int16		Number of “good” points
n_noflip .....	int16		Number of all points without RADECFLIP flag
n_points .....	int16		Number of all object detections
flags .....	int16		Object flags
Synonym Table			
id1 .....	int32		First object ID (composite primary key)
id2 .....	int32		Second object ID (composite primary key)
separation .....	float32	deg	Spherical distance (<1 pixel)
Observation Table			
obj_id .....	int32		Object ID (composite primary key and foreign key)
frame_id .....	int32		Frame ID (composite primary key and foreign key)
dra .....	int16	1/32000 deg	$\alpha_{J2000.0}$ offset from median centroid
ddec .....	int16	1/32000 deg	$\delta_{J2000.0}$ offset from median centroid
mag .....	int16	mmag	Corrected object magnitude
err .....	int16	mmag	Magnitude error
flags .....	int16		Measurement flags
Orphan Table			
frame_id .....	int32		Frame ID (composite primary key & foreign key)
rao .....	float64	deg	$\alpha_{J2000.0}$ of object (composite primary key)
deco .....	float64	deg	$\delta_{J2000.0}$ of object (composite primary key)
mag .....	int16	mmag	Corrected object magnitude
err .....	int16	mmag	Magnitude error
flags .....	int16		Measurement flags
htm_id .....	int64		HTM ID for quick spatial queries on the sphere

survey image data. However, image access is planned for future versions of SkyDOT.

We take advantage of limited precision and range of some database quantities (like magnitudes) and store them as short integers rather than floating point numbers. This simple rescaling trick generates close to 50% storage savings due to a small record size in the main table that dominates the size of the database. Only users of low- and intermediate-level data products need to be concerned with these technical details, since they are transparent to top-level users who access data through the presentation layer.

The data is searchable by celestial coordinates. The Hierarchical Triangle Mesh (HTM; Kunszt, Szalay, & Thakar 2001) is used for very rapid indexing of positions on the sphere. The NSVS database uses HTM partitioning of depth 14 with corresponding  $\sim 1'$  cell size. Searching for position matches within a small tolerance radius can be performed at a rate of  $\sim 20,000$  matches per minute. Access time is nearly uniform across the sphere. In particular, extraction of objects in regions including the north celestial pole is handled seamlessly. HTM runs as a shared library extension to the database server with SQL wrappers.

#### 4.3. Other Options for Data Access

All data discussed in the present release can also be downloaded from the SkyDOT Web site,<sup>9</sup> but transfer of the entire database over the network may be impractical. Therefore, we will also, upon request, distribute the data on user provided DDS 4 tapes (DAT size). The most efficient way of storing the database as files in terms of the size and speed of rebuilding the database proved to be gzip-compressed, semicolon-separated ASCII files. To save space, we skip columns with row identifications running from one to the number of entries. Those can always be deduced. The total size of the data set in this form is around 35 GB. It is strongly dominated by the Observation table with all object light curves, followed by the Orphan table with observations that could not be identified with any of the objects. Each of the tables, Field, Frame, and Object, fit within a single file. Tables Observation and Orphan are broken up into 644 files, one for each field.

<sup>9</sup> See <http://skydot.lanl.gov>.

#### 4.4. Acknowledging NSVS Data

We request that the researchers using NSVS in their published work include the following statement to acknowledge the source of data: This publication makes use of the data from the Northern Sky Variability Survey created jointly by the Los Alamos National Laboratory and University of Michigan. The NSVS was funded by the Department of Energy, the National Aeronautics and Space Administration, and the National Science Foundation.

#### 5. DISCUSSION AND FUTURE WORK

We have presented the Northern Sky Variability Survey (NSVS), to this date the most extensive temporal record of the sky on large spatial scales. All of the survey data is available to the astronomical community and can be searched efficiently using the public SkyDOT database. The database contains a total of 3.35 billion measurements for approximately 14 million objects in the 8–15.5 mag range. Time sampling over 1 full year is between twice per night and once every four nights, on average. The ROTSE-I instrument has achieved a complete spatial coverage of the northern hemisphere and a large fraction of the southern sky using remarkably low cost hardware. These two factors pose limits to the level of detail at which variability of the sky was recorded: low spatial resolution, spatial sensitivity variations, a nonstandard filter, and complicated systematics near the Galactic plane.

Despite its limitations, the NSVS is a truly rich source of information on stellar variability. Among stars in the Galaxy, the fraction of variables with amplitudes detectable by the NSVS is about  $\sim 1\%$  (Eyer 1999; Eyer & Cuipers 2000). Based on that and on preliminary results in Akerlof et al. (2000a), one can expect that tens of thousands of new variable stars with good uniform quality light curves are present in the data set. Current database schema needs to be expanded along the lines described in Woźniak et al. (2002) to accommodate various types of variables and provide classification capability. The NSVS combined with astrometric catalogs providing distances and motions, as well as multicolor surveys (2MASS, or even SDSS in a narrow magnitude range) will enable a comprehensive look at the Galaxy as traced by variable stars. The NSVS objects are bright and therefore the preferred targets for detailed spectroscopic and astrometric work. The spatial resolution of the survey is not far from that of high-

energy sky catalogs like the *ROSAT* All Sky Survey (Voges et al. 1999) or the *XMM* Catalog of Serendipitous Sources (Watson 2003), and therefore it is well suited for cross-correlations. Perhaps the most exciting questions to be attacked using the NSVS are regarding rare, hard to find objects. The astronomical literature provides numerous unexplained reports of variability events on normal stars (e.g., Schaefer, King, & Deliyannis 2000 and references therein).

Photometric monitoring data for active galactic nuclei (AGNs) providing diagnostics of accretion flows is valuable, but limited. Only a few bright AGNs are within the magnitude limit of the NSVS, so the real contribution to AGN physics will require deeper flux limits and better resolution in future projects. A major but low cost improvement in data usability would be the use of a set of standard filters before starting deeper surveys with more frequent time sampling.

Small robotic telescopes with automated data-processing pipelines are the best candidates for closing the gap in the current level of temporal monitoring of the sky. The computing power to perform on-line photometry is available. Experiments like RAPTOR (Vestrand et al. 2002) are starting to tackle the problem of real-time detection and immediate follow-up of short timescale phenomena. One can envision a monitoring system capable of partial interpretation of various events occurring on a variety of timescales and notifying subscribers about interesting changes of objects in their scientific problem domain. The main challenge is making the immense data stream comprehensible by putting enough smarts into the software. The sky itself is the ultimate astronomical database that should be mined continuously and in real time.

This work was supported by the Laboratory Directed Research and Development funds at LANL under DOE contract W-7405-ENG-36 to the RAPTOR project and NASA grant NAG5-5281 to the ROTSE-I collaboration. P. W. was supported by the Oppenheimer Fellowship at LANL. K. K. was supported by the NSF grant AST 02-05813 to Michigan State University, and S. M. was supported under the auspices of the DOE, NNSA by UC, and LLNL under contract W7405-ENG-48.

#### REFERENCES

- Abazajian, K., et al. 2003, *AJ*, 126, 2081  
 Akerlof, C., et al. 2000a, *AJ*, 119, 1901  
 ———. 1999, *Nature*, 398, 400  
 ———. 2000b, *ApJ*, 532, 25L  
 Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393  
 Brunner, R. J., Djorgovski, S. G., Prince, T. A., & Szalay, A. S. 2002, in *Handbook of Massive Datasets*, ed. J. Abello, P. M. Pardalos, & M. G. C. Resende (Dordrecht: Kluwer), 931  
 Calabretta, M., & Greisen, E. W. 2002, *A&A*, 395, 1077  
 Charbonneau, D. 2003, preprint (astro-ph/0302216)  
 Chen, W. P., Lemme, C., & Paczyński, B., ed. 2001, *Small Telescope Astronomy on Global Scale* (San Francisco: ASP)  
 Djorgovski, S. G., et al. 2001, in *Mining the Sky*, ed. A. J. Banday, S. Zaroubi, & M. Bartelmann (Heidelberg: Springer), 305  
 Eyer, L. 1999, *Baltic Astron.*, 8, 321  
 Eyer, L., & Cuipers, J. 2000, in *ASP Conf. Ser. 203, The Impact of Large-Scale Surveys on Pulsating Star Research*, ed. L. Szabados & D. Kurtz (San Francisco: ASP), 71  
 Ferlet, R., Maillard, J. P., & Raban, B., ed. 1997, in *IAP Colloq. 12, Variable Stars and the Astrophysical Returns of the Microlensing Searches* (Gif-sur-Yvette: Ed. Frontières)  
 Frink, S., Quirrenbach, A., Fischer, D., Röser, S., & Schilbach, E. 2001, *PASP*, 113, 173  
 Hog, E., Kuzmin, A., Bastian, U., Fabricius, C., Kuimov, K., Lindegren, L., Markarov, V. V., & Roeser, S. 1998, *A&A*, 335, L65  
 Horne, K. 2002, in *Proc. First Eddington Workshop on Stellar Structure and Habitable Planet Finding*, ed. B. Battrick, F. Favata, I. W. Roxburgh, & D. Galadi (ESA SP-485) (Noordwijk: ESA), 137  
 ———. 2003, in *ASP Conf. Ser. 294, Scientific Frontiers in Research on Extrasolar Planets*, ed. D. Deming & S. Seager (San Francisco: ASP), 361  
 Kehoe, R., et al. 2001, *ApJ*, 554, L159  
 ———. 2002, *ApJ*, 577, 845  
 Kholopov, P. N. 1998, *General Catalog of Variable Stars* (4th ed.; Moscow: Nauka)  
 Kunszt, P., Szalay, A. S., & Thakar, A. R. 2001, in *Mining the Sky*, ed. A. J. Banday, S. Zaroubi, & M. Bartelmann (Heidelberg: Springer), 631  
 Paczyński, B. 1997, in *IAP Colloq. 12, Variable Stars and the Astrophysical Returns of Microlensing Surveys*, ed. R. Ferlet, J. P. Maillard, & B. Raban (Gif-sur-Yvette: Ed. Frontières), 357  
 ———. 2000a, in *ASP Conf. Ser. 203, The Impact of Large-Scale Surveys on Pulsating Star Research*, ed. L. Szabados & D. Kurtz (San Francisco: ASP), 71

- Paczynski, B. 2000b, *PASP*, 112, 1281  
———. 2001, in *Mining the Sky*, ed. A. J. Banday, S. Zaroubi, & M. Bartelmann (Heidelberg: Springer), 481  
Park, H. S., et al. 2002, *ApJ*, 571, L131  
Pojmański, G. 1997, *Acta Astron.*, 47, 467  
———. 2000, *Acta Astron.*, 50, 177  
———. 2002, *Acta Astron.*, 52, 397  
Schaefer, B. E., King, J. R., & Deliyannis, C. P. 2000, *ApJ*, 529, 1026  
Skrutskie, M. F., et al. 1997, in *Impact of Large Scale Near-IR Sky Surveys*, ed. F. Garzon et al. (Dordrecht: Kluwer), 25  
Stoughton, C., et al. 2002, *AJ*, 123, 485  
Vestrand, W. T., et al. 2002, *Proc. SPIE*, 4845, 126  
Voges, W., et al. 1999, *A&A*, 349, 389  
Watson, M. G. 2003, in *ASP Conf. Ser. 295, Astronomical Data Analysis Software and Systems XII*, ed. H. E. Payne, R. I. Jedrzejewski, & R. N. Hook (San Francisco: ASP), 107  
Woźniak, P. R., Borozdin, K. N., Galassi, M. C., Priedhorsky, W. C., Starr, D., Vestrand, W. T., White, R., & Wren, J. A. 2002, *Proc. SPIE*, 4846, 147