Using the SuperCOSMOS Sky Surveys in a Survey for Objects of Extreme Variability

Sammy Yousef
University of Western Sydney, Nepean
Prepared as part of the subject "Astronomy Written Project" of the Astronomy Internet Masters Program, 2002.
Email: syousef@bigfoot.com

Abstract

The SuperCOSMOS Sky Surveys are described in detail in three papers: N.C. Hambly, H.T. MacGillivray et al 2001, N.C. Hambly, M.J. Irwin et al 2001 and N.C. Hambly, A.C. Davenhall et al 2001. Section 3.2 of the first of these papers describes some of the science undertaken using the survey data, with section 3.2.4 focusing on a targeted survey for objects with extreme variability. This paper discusses the detail of such a survey. A brief description of the nature of an extreme variability survey is followed by a detailed treatment of the use and limitations of the SuperCOSMOS sky surveys (SSS) data. The use of other data to supplement the SSS data and overcome these limitations is then considered.

1 Introduction

Astronomical objects that exhibit great variability over a relatively short time scale - from minutes to years - are of interest to astronomers and physicists because they are sites of extreme activity. They allow the unique, albeit physically remote opportunity to examine physical processes under extremes that cannot be found nor duplicated on Earth, but which are key to our understanding of the Universe. The discovery and identification of variable objects is therefore an important contribution to the field of astronomy.

The SuperCOSMOS Sky Surveys (SSS) include digitization of photographic data sourced from Schmidt telescope plates taken in the later half of the 20th Century. The eventual aim is to cover the whole sky. A subset of the data, called the South Galactic Cap Survey (SGC) covering 5000 square degrees, with a Galactic latitude above 60 degrees comprise the first release of SuperCOSMOS data. This includes data complete in two epochs filtered in one colour: Red or R-band data at ~700nm (SSS Website; Hambly N.C., MacGillivray H.T. et al 2001). Such data fulfils the minimum requirement for a study of variability i.e. magnitudes at two points in time for the same area of sky in the same wavelength.

In examining SGC data the SSS team has discovered several new objects (according to initial checks against the SIMBAD database). Three possible Miras or Novae, each with a magnitude difference of 5 or greater between the two epochs, as well as what appears to be an old supernova associated with an Scr spiral galaxy ESO 148-20 have been found. (Hambly N.C., MacGillivray H.T. et al 2001). Furthermore the authors state that "many more [new discoveries] are almost certainly present in the SSS data".

2 A Survey of Objects with Extreme Variability

A more thorough, variability targeted computer-automated search of the SSS data recorded at two epochs (focusing on the SGC survey) is possible on at least two levels. Firstly objects that have been identified and catalogued as part of the survey can be compared at each epoch. A computerized search of the catalogue should quickly return a listing of objects present at one epoch but absent at the other, or with the same astrometry (within a tolerance) that have changed significantly in magnitude.

Secondly, the digitized images themselves can be compared. Sections of digitized sky can be digitally aligned, scaled and adjusted for differences in exposure and in the plate etc., then mathematically subtracted from one another using computer software. (Colless, M. 2001) Done correctly this should leave only a difference in magnitudes between the images (though perhaps some image artefacts would require attention before or after subtraction) – these image deltas could then be examined manually or computer processed to validate changes in astrometry and/or magnitude.

Finally the differences between the results of each method could be compared statistically so that recommendations might be made on the most suitable approach for further variability surveys.

3 SuperCOSMOS Sky Surveys Data

3.1 Use of the Data in a Survey for Objects with Extreme Variability

The identification of interesting sources for further monitoring and study must be the primary goal of any extreme variability survey using SSS data. The SGC survey is a well defined and organized digitization of 5000 square degrees of contiguous sky. The two epochs of R-band data have been collected up to several years apart. Three types of events can be identified but not necessarily distinguished using such data:

1) Sudden one-off events such as supernovae may be detected if the event happens to begin just before one of the two plates is captured.
2) Repeated short term variability may be observed if different phases of the cycle occur around the time each plate is exposed. Several oscillations in magnitude may occur between the photographs and the data would not reveal the detail of these. The chance of photographing two objects in the same phase of their variability cycle varies depending on the light curve of the object.
3) Genuinely long-term slow changes in magnitude between epochs will be detected.

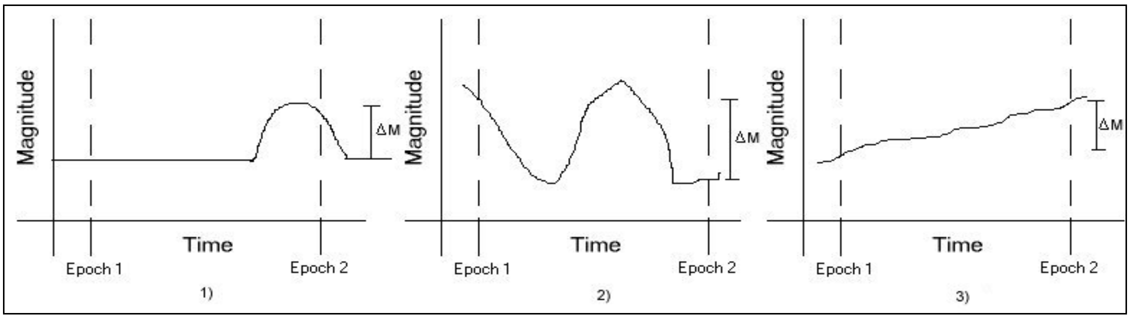

Figure 1 Types of Light Curve That Will Not Be Distinguishable From Two Epoch Data.
All the data provides is a change in magnitude (perhaps even to or from the plate limit). Note that we do not capture the maximum change in magnitude between epochs –objects may have varied more wildly between epochs when they were not being observed. The astronomer's knowledge of the nature of the object may allow a determination of its probable type e.g. Its brightness and location within a galaxy may be typical of a supernova. Light curves 1 and 3 show an increase in magnitude between epochs. Light curve 2 shows a decrease. An object which is visible at one epoch and invisible (beyond the plate limit) at another would exhibit a change in magnitude between a horizontal line that is the plate limit and its brightness at the other epoch rather than both points of its curve.

In dealing with SGC catalogue data rather than engaging in an image differencing exercise, the advantage is that the need to deal with the photometry, astrometry and object detection routines required for image analysis is eliminated. Instead the astronomer simply must be aware of the limits of the data processed to create the catalogues which are to be used. (See Hambly N.C., Irwin M.J. et al 2001 for detail). The exercise is reduced to one of data mining and data extraction, using similarly processed data for each epoch. This therefore greatly simplifies the processing required to detect variable objects. Note also that the catalogue data will be a much smaller set than the image data in terms of physical storage, making for easier and faster extraction, processing and transmission.

Catalogue data will probably show its shortcomings in terms of the number of objects identified, and may prove less suitable for comparison with other survey data where the other survey's astrometric accuracy, capture and image identification processes differ significantly from the techniques employed in SSS. Both false negative detections and false positives may become a problem. To eliminate false positives, each identification would naturally need to be cross checked and only other R-band (or complementary data for other bands covered by SSS, for significant portions of the sky) could be used without significant statistical processing.

3.2 Other Changes Between Epochs

The R-band data of the SGC survey allows for the study of other changes in the region between the two photographic plates. If an image difference analysis is carried out, it may prove more worthwhile studying overall changes between epochs rather than discarding data that does not relate to magnitude changes.

Unfortunately the one-colour nature of the data doesn't allow for in depth analysis of spectroscopic changes. However, stellar and solar system objects exhibiting large proper motions may be detected (eg. Near earth asteroids). The survey is large enough that the statistical properties of such objects, and their detection through the survey could be studied. These differences can therefore either be looked upon as additional overall data collected by the survey, or noise in attempting to isolate extreme changes in magnitude.

3.3 Limitations of the Data

If catalogue rather than image data is used, the limitations of the data as stated in the three SSS survey papers must be taken into account.

3.3.1 Selection Effects – Temporal, Positional and Waveband Selection

The main problem with surveys such as the SGC is that with only two temporal points, spaced years apart, it is not possible to follow the light curve of highly variable objects. The immediate consequence is that any variability detected between epochs will be less than the maximum variation of the objects throughout the period, as the statistic probability of capturing the maximum and minimum in variability on the two epoch samples is extremely low. This relegates the survey SGC and SSS surveys to providing:

1) Initial identification of potential targets through the comparison of object magnitudes between the two epochs of R-band data.
2) Historical Data for all catalogue objects which cannot be obtained in any way by a subsequent survey. This can provide background for newer all-sky variability surveys.

Another problem caused by using images taken years apart is that nearby objects with large proper motions may not be identified as the same object. Instead it will appear as if two variable objects have been detected – one that has faded past the plate limit and another that has brightened into view. The SSS survey astrometry goes some way towards identifying proper motions but without other data and/or observation it may not be possible to distinguish these two scenarios. Fortunately such data will likely be available and the number of objects with large proper motions is limited. This effect must however be taken into account in any automated variability detection technique.

By relying on data from the SGC survey, a positional bias is also introduced which undoubtedly will affect survey statistics. Object densities near the South Galactic Cap are less than for the Galactic plane. In fact the South Galactic Cap was chosen as the first SSS data release in part to allow for image reduction and analysis techniques to be trialled on a less dense area of sky first.

Lastly, using photographic plates, the astronomer limits the waveband at which variability is being detected. The selection is further narrowed by using only data filtered at specific wavelengths - the R-band data from both epochs of the survey, or any of the three wavebands from the SSS surveys to compare with data from a separate survey. It is important to remember that variability may occur across the entire spectrum from Radio through to Gamma and Cosmic Rays. A non-detection of variability in a specific part of the visual waveband neither rules out variability in other parts of the visual band nor other parts of the spectrum as a whole. The "extreme variability" sought would more accurately be termed "extreme visual variability".

3.3.2 Astrometry

Evidence presented in Hambly N.C., Davenhall A.C. et al 2001 suggests that the SSS data, and particularly the SGC subset is accurate to within +/- 0.2 arcsec for R-band data around magnitude 18 to +/- 0.3 arcsec for R-band data around magnitude 21 with "positional dependent systematic effects from bright to faint magnitudes at around 0.1 arcsec". Proper motions are quoted to have an accuracy of between +/- 10 $masyr^{-1}$ and +/-50 $masyr^{-1}$.

When matching objects, it is possible to combine the positional accuracy, proper motion and time between epochs for R-band data to determine if an object at one epoch matches an object at the other. Based on the above a naïve approach would be to assume that an object that is +/- (0.3 arcsec + 50 * years_between_epochs mas) distant on the second plate may match the object on the first. In practice greater care and sophistication is needed in matching objects.

3.3.3 Photometry

The quoted accuracy of the photometry of the SGC survey in Hambly N.C., Irwin M.J. et al is roughly 0.3 mag with respect to external data for magnitudes above 15, with overall accuracy possible to within 5% for well exposed plates. Since we are targeting objects which display extreme variability, the accuracy only becomes an issue near the plate limit. If we were to extend our survey to target less variable objects, the accuracy presented here might prove a problem since we cannot distinguish magnitude changes for two objects 2 * 0.3 = 0.6 mag apart based on SGC data.

Colour accuracy is also considered relatively high for the SGC survey. When compared to external sources $s_{B-R,R-I}$ is around 0.7 for B_J around mag 16.5, and increases to 0.16 at mag B_J around 20. This is particularly important when comparing data with other sources because when comparing two magnitude measurements it is important to be able to have confidence that they are measurements for the same waveband.

3.3.4 Object Detection

SSS digitization employs a sophisticated two phase image detection system that it is claimed gives 99% or greater image detection within 1.5 magnitudes of the plate limit, and around 90% overall. Both point and extended sources are identified. The technique combines aspects of instantaneous processing and post-processing. The detail of the process is beyond the scope of this paper. See Hambly N.C., Irwin M.J. et al 2001 for details and references. The effort of the SuperCOSMOS team in attaining such accuracy improves the chances of success and likely quality of a catalogue based search of the data for extremely variable objects.

3.3.5 Plate Limits, Sample Completeness and Malquist Bias

The plate limit of the SSS surveys varies between magnitudes 19 and 23 overall, with R-band plate limits between magnitudes 20 and 22. Like any survey that is magnitude limited, Malquist bias – the over-representation of brighter objects in the sample - is a problem when statistics are generated from the data. The best way to overcome this is usually to take a distance limited sample of the objects from the already magnitude limited data.

For the purposes of detecting highly variable objects however, the full range of data should be used. The primary goal of the extreme variability survey is one of discovery. If statistical data is required, a subset of the discovered objects can be chosen for analysis. This approach, while sensible, cannot however eliminate problems with sample completeness. Highly variable objects too distant to be detected at their brightest will naturally not be found.

Also a subset of the data will show an object that has faded in or out from the plate limit, further limiting the range in magnitude that is observed, and complicating any statistical analysis. As mentioned previously there is also an issue with sample completeness from a temporal point of view. Two measurements of an object’s magnitude are sufficient to declare that its magnitude varies significantly, but inadequate to form a complete picture of the object’s behaviour. For those objects remaining in view (i.e. which have not faded to the point where they can no longer be detected), additional data may be collected to obtain a more complete picture of their nature and behaviour. This will require further collection of data that is not part of the SuperCOSMOS surveys. Past events however, which are not ongoing, cannot be observed further. The only way to study these is to find other archival observations made around the time of the event, if they exist.

4 Use of Data from Other Surveys

It has been demonstrated that the SuperCOSMOS data, particularly the SGC survey can be used to detect astronomical objects that vary greatly in magnitude over a time span of many years. It has further been demonstrated that the data from this survey is limited for such a purpose and that there are many flaws in using it exclusively for the detection of variable objects. More importantly the data is of little use for follow-up research except in the sense that two historical observations of the light curve spaced years apart have been made – something which no subsequent survey could provide.

Once objects are identified, a follow-up targeted survey can then focus on the detected objects and, provided the variability continues, long term detailed observations of their light curves and spectroscopy can be made. Unlike this initial proposed “virtual” extreme variability survey which is actually based on SuperCOSMOS data, the follow-up survey would include observation and the collection of its own data, which would later form a unique archival source for subsequent research.

Data from other pre-existing survey sources may also prove useful in this respect, particularly data from surveys that monitor large numbers of objects simultaneously and resample the same parts of the sky periodically in the same wavebands. It is possible to combine data from other surveys with the SuperCOSMOS data to detect more variable objects, particularly one-off occurrences, or to provide more points on the light curve for those already detected. Problems to overcome include matching the data in terms of astrometric and photometric accuracy, overcoming incompatibilities between object detection techniques, and if image data is being used directly, scaling the data if the digitization and/or plate scales are different.

The greatest yield of data regarding the variability of objects however may come from new surveys by amateurs and professionals with an interest in variability. As technology has improved there has been a greater push for the continuous monitoring of variability for the whole sky. The recognition that only around 10% of variables brighter than apparent magnitude 12 have been discovered as yet has led to a renewed interest by amateurs and professionals with instruments as small as 4”-12” in aperture. (Blake C. 2002; TASS – The Amateur Sky Survey). The observation of variable stars has had a long history of contribution by amateurs, with detailed studies of light curves made since the 1800s. Today organizations such as AAVSO allow for co-ordinated programs to be carried out by groups of dedicated amateurs. At the same time, professionals are gaining access to increasingly sophisticated equipment capable of monitoring an increasing number of sources simultaneously. In the long term these trends may relegate all one off surveys such as SSS/SGC to a secondary role of providing historical data for objects preceding a new era of ongoing surveys. (Paczynski B. 2001; Paczynski B. undated).

Astronomers working in non-visual wavelengths such as radio are also looking at studying variability. (McGilchrist & Riley 1990). Increasingly, as orbiting telescope technology has become more pervasive opening up the entire spectrum to observation, there is a movement to study any given object in all available wavelengths to obtain the maximum possible amount of information about that object. Variability studies are no exception to this trend. A lack of detection of variability in the visual part of the spectrum does not equate to a lack of variability at other wavelengths.

5 Conclusion

The SuperCOSMOS Sky Surveys (SSS) and particularly the South Galactic Cap (SGC) subset of the data can be used to identify objects whose magnitudes have varied in the extreme. The SGC data contains R-band images for two epochs, making time-based comparison possible. Already, previously undetected objects have been found in the SGC data. These include three possible Miras or Novae, which have varied by five magnitudes or greater between the two R-band epochs, and a possible old supernova associated with galaxy ESO 148-20. An automated/computerized search should produce more sources. In searching for highly variable objects, a number of factors must be given attention. Individual objects must be identified and matched across the 2 epochs of data to a high degree of certainty given the astrometric limitations of the survey. The use of catalogue data from the SuperCOSMOS survey will provide the most immediate rewards for the least effort, however the comparison and subtraction of the R-band image data in the two epochs studied has the potential to yield more information at the cost of requiring much more careful processing.

The use of the SGC data in combination with data from other surveys opens up further possibilities for investigating highly variable objects. Years have passed between R-band pictures of the SGC survey. Other surveys can “fill in the gaps” regarding what has been happening between these snapshots. Other surveys may also yield data regarding variability at non-visual wavelengths, which may in turn produce a clearer idea of what the object in question is. However differences in the data between surveys also create new challenges. The astrometric certainty of other surveys will not precisely match that of the SGC data, placing higher demands on any automated matching algorithm. Furthermore, the wavelengths filtered in other surveys will not match those of the SGC R-band data precisely. Variations due to the use of different wavelengths rather than due to an intrinsic change in the brightness of the object will need to be estimated and corrected for. New all-sky surveys that aim at monitoring the variability of the entire sky may eventually relegate the data such as produced by the SGC survey to a historic role, for the purposes of detecting variability. That is, it will be possible to check the history of variability of an object using SSS or other old surveys pre-dating these newer efforts. The new surveys will however become a more suitable source for detecting variable objects.

6 Acknowledgements

I’d like to thank David Blank for his rapid and friendly feedback, instruction and guidance throughout the Astronomy Written Project Subject, and Alex Hons for his tireless work in support of the Astronomy Internet Masters program.

7 References

American Association of Variable Star Observers (AAVSO)
http://www.aavso.org/

Colless, M. 2001
Observational Techniques Workshop April 2001
http://www.aao.gov.au/local/www/obstech/colless.ppt

Hambly N.C., MacGillivray H.T., Read M.A., Tritton S.B., Thomson E.B., Kelly B.D., Morgan D.H., Smith R.E., Driver S.P., Williamson J., Parker Q.A., Hawkins M.R.S, Williams P.M., Lawrence A. 2001
The SuperCOSMOS Sky Survey Paper I: Introduction and Description
(Preprint) Mon. Not. R. Astron. Soc.
http://arxiv.org/abs/astro-ph/0108286

Hambly N.C., Irwin M.J., MacGillivray H.T. 2001
The SuperCOSMOS Sky Survey Paper II: Image detection, parameterization, classification and photometry
(Preprint) Mon. Not. R. Astron. Soc.
http://arxiv.org/abs/astro-ph/0108290

Hambly N.C., Davenhall A.C., Irwin M.J., MacGillivray H.T. 2001
The SuperCOSMOS Sky Survey Paper III: Astrometry
(Preprint) Mon. Not. R. Astron. Soc.

http://arxiv.org/abs/astro-ph/0108291

McGilchrist & Riley 1990
The 7C survey of radio sources at 151 MHz - a search for low-frequency variability
Mon. Not. R. Astron. Soc., 246, 123

Paczynski B. undated
The Future of Massive Variability Searches
Invited talk, presented at the 12th IAP Colloquium: Variable Stars and the Astrophysical Returns of Microlensing Surveys.
http://stupendous.rit.edu/tass/refs/searches.ps

Paczynski B. 2001
Monitoring All Sky for Variability
Publications of the Astronomical Society of the Pacific, Volume 112, Issue 776, pp. 1281-1283
Preprint at: http://arxiv.org/abs/astro-ph/0005284

Blake C. 2002
The Princeton Variability Survey
Publications of the Astronomical Society of the Pacific Preprint.
http://arxiv.org/abs/astro-ph/0201394

The SuperCOSMOS Sky Surveys (SSS Website)
http://www-wfau.roe.ac.uk/sss/

TASS – The Amateur Sky Survey
http://stupendous.rit.edu/tass/tass.shtml